INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
A Xerox Education Company
PROBLEMS OF TRUTH AND REFERENCE IN FICTION

DISSertation

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Mary Jeannette Sirridge, B.A., M.A.

The Ohio State University
1972

Approved by

Advisor
Department of Philosophy
PLEASE NOTE:

Some pages may have
indistinct print.
Filmed as received.

University Microfilms, A Xerox Education Company
ACKNOWLEDGMENTS

I wish to thank my dissertation advisor, Professor Alan Hausman, for his continued help and encouragement, and the members of my dissertation committee, Professors Lee Brown and William Lycan, for their careful reading of the manuscript and many helpful suggestions. Finally, I wish to thank other members of the philosophy department of The Ohio State University for their encouragement and help with various sections of the thesis, particularly Professors Richard Garner, Robert Turnbull, and George Schum.
VITA

May 7, 1946  .  .  .  .  Born - Cleveland, Ohio

1967  .  .  .  .  B.A., St. Mary's College, Notre Dame, Indiana

1967-1968  .  .  .  .  Fulbright Fellowship, University of Heidelberg, Heidelberg, Germany

1968-1969  .  .  .  .  NDEA Fellowship, University of Rochester, Rochester, New York

1969-1970  .  .  .  .  Teaching Assistant, The Ohio State University, Columbus, Ohio

1970-1971  .  .  .  .  Research Associate, Center for Medieval Studies, The Ohio State University, Columbus, Ohio

1971-1972  .  .  .  .  NDEA Fellowship, The Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Philosophy
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>I. HISTORICAL POSITIONS</td>
<td>10</td>
</tr>
<tr>
<td>1.1 The Proposition Theory</td>
<td></td>
</tr>
<tr>
<td>1.2 The Expressive Theory</td>
<td></td>
</tr>
<tr>
<td>II. THE LANGUAGE OF FICTION</td>
<td>70</td>
</tr>
<tr>
<td>2.1 Speech Act Analysis of the Language of Fiction</td>
<td></td>
</tr>
<tr>
<td>2.2 Marks of the Language of Fiction</td>
<td></td>
</tr>
<tr>
<td>2.3 The Historical Novel</td>
<td></td>
</tr>
<tr>
<td>2.4 The Analysis of Direct and Indirect Quotation and Related Phenomena</td>
<td></td>
</tr>
<tr>
<td>III. THE SEMANTIC BASE OF CRITICISM</td>
<td>138</td>
</tr>
<tr>
<td>3.1 The Nature of the Language of Criticism</td>
<td></td>
</tr>
<tr>
<td>3.2 Individuation of the Fictional World</td>
<td></td>
</tr>
<tr>
<td>3.3 Formal Semantics</td>
<td></td>
</tr>
<tr>
<td>IV. EXTENSIONS OF THE BASIC SEMANTICS</td>
<td>192</td>
</tr>
<tr>
<td>4.1 Schematic Analysis of Extensions</td>
<td></td>
</tr>
<tr>
<td>4.2 Specific Application of Theory</td>
<td></td>
</tr>
<tr>
<td>V. FURTHER EXTENSIONS</td>
<td>223</td>
</tr>
<tr>
<td>5.1 Schematic Analysis of the Logic of Extensions of Basic Criticism</td>
<td></td>
</tr>
<tr>
<td>5.2 Specific Application of Theory</td>
<td></td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>275</td>
</tr>
</tbody>
</table>
Introduction

I propose to take up the question of the cognitive value of what I shall call "works of epic literature" and the question of the justification of interpretive critical claims which have to do with the cognitive value of a work of epic literature. Obviously the two questions are closely related.

I am interested in works of literature consisting of an ostensible narrative about persons and objects and the events by which they are related. In addition, it is characteristic of the works I shall be considering that the purported persons and objects talked about do not and never have existed and the events never really occurred, or at least that whether the persons and objects exist and whether the events occurred is irrelevant to the success of the literary work. Hence these works are often called "fictional." Such works I shall call "works of epic literature," taking the name from the oldest members of the group; sometimes I will call them "works of epic fiction" or merely "fictional works." The classification cuts across many distinctions customarily made with regard to works of literature. Some of them are prose works, while others, like The Nibelungenlied, are in verse. Some establish complex worlds, e.g., Remembrances of Things Past, while others concentrate on a single episode. Broad as the scope of the classification is, however, it does not include all works of literature. It does not, as here presented, include the drama or lyric poetry. And I shall not be dealing with
every aspect of the works included under "epic literature." I am
interested primarily in their alleged cognitive value.

It is widely claimed that what I have called "works of epic
literature" have cognitive value, i.e., function to expand and clarify
our knowledge and beliefs about the actual world. But they cannot
function as evidence about the world, since they are only stories,
and, in view of their fictional character, not even true stories at
that. "I read it in a novel," is never per se good evidence for a claim
about the way the real world is. The question is how such works exer­
cise their distinctive cognitive function.

The justification of critical claims about the cognitive value of
a work of epic literature is another side of the same coin. I hope
that certain traditional answers to these questions are unacceptable.
It turns out that an acceptable answer can be given only by giving a
full scale account of the nature of works of epic literature and of the
justification of critical claims. This is what I hope to accomplish.

In chapter I, I will examine certain historically important
answers to the sorts of questions I have raised. The first of these
is what has been called "the proposition theory." On such a theory,
works of fiction are construed on the model of fact-stating discourse.
It is alleged that they constitute or contain as their meaning true or
false propositions about the actual world. On this view, an inter­
pretive critical claim about the cognitive significance of a work to
the effect that (a part of) the meaning of the work is p, where p is
some proposition about the actual world, can be justified, just in
case p can be shown to be (part of) the meaning of the work in question.
The proposition theory is, I think, counterintuitive as a theory about the nature of works of fiction and about the fictional use of language. The main reason for holding such a theory is the need to account for the cognitive function of works of fiction and the justification of critical claims. In section 1.2, I shall attempt to show that there is no clear formulation of the theory which actually gives an answer to these questions. Indeed, the examination of the proposition theory points out a difficulty in the justification of any interpretive critical claim whatsoever, a difficulty which it cannot itself resolve.

In section 1.2, I shall consider the so-called emotivist approach to the problem as advanced by Stevenson and Richards. On this view, works of literature are not construed on a model of fact-stating discourse, but rather along the lines of emotive discourse. Although they cannot serve as evidence for beliefs about the actual world, it is alleged that they cause or persuade readers to entertain such beliefs. Interpretive claims in general are also construed as functioning as emotive language. Although there is no "fact of the matter" in interpretive criticism, it is claimed, interpretive claims may persuade us to have certain beliefs about the work of fiction or may persuade us to see a work in such a way that it affects our beliefs about the actual world. Rather than posing a solution to the problem, then, the emotivist position offers a way to dissolve it. I shall argue that it is correct to construe works of epic literature as non-fact-stating discourse, but that the emotivist never offers a detailed alternative account of how such works function in influencing our beliefs. And I shall argue that the emotivist position offers an unsatisfactory
account of critical discourse in general, and a fortiori of interpretive critical claims.

Chapter I is thus mainly polemical. But it does have some constructive force in that it shows that to get at questions relating to the cognitive value of a work, we need a detailed account of the nature of a fictional work. And in the course of chapter I, it becomes obvious that interpretive critical claims about the cognitive value of a work of fiction are very complex critical claims whose treatment demands a systematic account of the justification of the simpler critical claims on which they appear to be based, at least in part.

Obviously, the question of the nature of a fictional work comes first. Only after it is dealt with can we begin to look at the question of how to justify critical claims about such works. In chapter II, I start from the basic claim that works of epic literature are plausibly taken as products of complex speech acts, i.e., as products of actions of a given kind performed by means of language in accord with rules or conventions for their performance and for their understanding. Unlike most common uses of language, the fictional use of language is not directed at the actual world. Instead fictional language is used to present fictional worlds and their characters. This suggestion is not new, but it has commonly been held to lead to the view that the language of fiction is not meaningful. In section 2.1, I attempt to show that this conclusion does not follow. I argue that the language of fiction is not meaningless, although it is not directed at the actual world, because in certain crucial ways it is parasitic on language which is
actual-world-directed. This latter kind of language I call "the language of normal narration." One use of language is properly said to be parasitic on another if it is the case that we would not understand the first if we did not understand the second, although the two uses are distinct. In section 2.2, I list certain characteristics of fictional language by which it can be recognized and attempt to show that in each case, the characteristic is traceable to the features of fictional language which make it distinct from the language of normal narration. Ultimately, such features are a consequence of the fact that the language of normal narration is directed toward the actual world, and fictional language is not. Historical novels seem prima facie to constitute a counter-example to the general claim that fictional language is quite different in function from normal narration. Such works do seem to be actual-world-directed. In section 2.3, I attempt to show that historical novels are only apparently recalcitrant to the treatment I propose. They are related to the actual world in a complex way. But this is not because they directly convey or purport to convey information about the actual world. Instead they are genuine fictional works upon which the actual world imposes certain more or less strong structural constraints. In the final section of chapter II, I attempt to exploit the extensive parallel between the language of normal narration and the language of fiction to suggest a solution to the critical problem of works which, intuitively speaking, seem to consist of a "narrative frame" and an "interior narrative." The problem has to do with what we take the "facts" of the work to be, the events of the interior narrative or the "intentional facts" of the narrator's performance. I suggest that we need not take the
latter course if we are willing to construe the narrative frame as a characteristic indication that the work is to be treated as fiction, analogous to the opening 'Once upon a time'.

In chapter III, I begin to deal with the problem of the justification of critical claims by setting up a procedure or assigning truth values to a small segment of critical claims about a work, i.e., those which match the sentences of the work or which can be derived from claims which do.

Since the truth values will be assigned relative to a given work, the first task is to sharpen our commonsense conception of what constitutes a single work. I take up the problem of the individuation of fictional works in section 3.1. It has been a standard philosophical objection to assigning truth values to critical claims that in assigning the value 'true' to sentences containing ostensible references to fictional entities, we involve ourselves in undesirable ontological commitment to non-existents. In section 3.2, I attempt to meet this objection by setting up a formally correct recursive truth definition for basic critical claims which does not automatically involve such an undesirable commitment, although it might be interpreted so as to do so. In section 3.3, I offer an interpretation of the nature of fictional entities and the truth values of critical claims which I think avoids the undesirable ontological commitment completely. Basically, this involves treating a fictional world as the result of an author's speech act; fictional characters are construed literally as parts of these fictional worlds, as what I shall call "expression types." In section 3.4, I attempt to show that such a position does not make nonsense of ordinary critical
discourse, although it does involve a reinterpretation of critical claims. But this need for reinterpretation, as I attempt to show, is a quite common feature of critical discourse, and is desirable on independent grounds. Quite simply, we often (although not always) talk critically as if fictional entities or the figures of a painting were real persons standing in actual world relations, but we certainly do not thereby commit ourselves to meaning literally that they are.

Chapter III is thus primarily directed at philosophers. With a few exceptions, it merely makes explicit the way in which literary critics proceed. But in addition, it shows just how fragmentary a critical account would be which restricted itself to what is actually in the text and what is derivable from it. In chapter IV, I point out that with certain qualifications, such a critical account is precisely what would result if we took seriously the requirement that the critic should restrict himself to factors internal to the work. It is, I suspect, the only clear way to specify this requirement, a requirement which is at least theoretically, basic to the formalist theory of criticism. In section 4.1, I show that such criticism is obviously inadequate. I then propose a way of going beyond these "internal considerations" by an appeal to factors which are external to the work. In section 4.2, I show that critics whose position is explicitly formalist often make assumptions in practice which can only be justified on some such "externalist" approach as I propose.

Once the explicitly formalist position has been given up, the way is open to accounting for the justification of a large number of more complex critical claims. In section 5.1, I set out a schematic account
of how such a project can be carried out. This in effect furnishes a logic of criticism. Among those claims for which I propose an account are what I call "cognitive significance claims," claims about the relation of a literary work to beliefs about the actual world. Thus, in chapter V, we return to the questions of chapter I. I offer a solution which is distinct from those offered by the proposition theory and the emotivist position, and which is, I think, distinctly preferable to them. On my view, works of epic literature do not serve as evidence for beliefs about the actual world. In this respect my view resembles the emotivist position. But on my view, there are strong constraints on what beliefs may properly be attributed to a fictional work. In addition, my view leaves open the possibility that the relation between works of epic literature and the actual world may be construed as quasi-evidential in the sense that thought experiments have quasi-evidential force with respect to beliefs about the actual world.

A logic of criticism such as I propose would, I think, be suspect, if it proved to be the case that no critics used arguments which could be shown to have the structure I suggest. In addition, the enterprise would perhaps be superfluous if it could be shown that no gain in clarity resulted from such a systematization of critical arguments. In section 5.2, I attempt to show that there are critical arguments which are, intuitively speaking, good arguments, which may plausibly be construed as having the structure I suggest, and that making their logical structure explicit is a great help in deciding whether they are valid, what their suppressed premises are, etc.
In summary, then, I hope to have given a logic of descriptive and interpretive critical discourse and an account of the relevant features of works of epic literature. With respect to the whole field of literary criticism, this is quite a narrow result. I will not have dealt with questions of evaluation of works of epic literature or with questions related to their purely aesthetic effects, although the results I arrive at will, if correct, have some bearing on these other issues. In any event, if I am right, the questions I consider are an important preliminary to the consideration of these other questions.
Chapter I

There are two perennial problems regarding those works which I have called "works of epic literature." First, works of this kind are often alleged to have cognitive value; in some way, it is claimed, they add to or clarify our knowledge of the actual world. The perennial question is how the work of epic literature, which is merely a story—and not a true story at that—can have such a cognitive value or function. The second question concerns the legitimacy and justification of what I shall call "critical significance claims," those which claim that a given work has as a part of its content a "meaning" or "moral" which is a general truth (or falsehood) about the actual world.

The justification of critical significance claims is an instance of a more general problem in aesthetics, i.e., the legitimacy and justification of what I shall call "interpretive claims" about the content of a work. A concrete example makes it clear, I think, what kind of critical claims are involved and why they present a problem. Consider the following event description.

The grass-plot before the jail, in Prison Lane, on a certain summer morning not less than two centuries ago, was occupied by a pretty large number of the inhabitants of Boston.

Clearly a critic may say that the expression 'Arthur Dimmesdale' or the proposition that on a certain morning, most of the inhabitants of Boston were gathered before the jail is part of the content of the novel; he can defend his claim by pointing to the expressions in the
work. But this is certainly not the entirety of what is meant when
the content of the work of art is under discussion. The critic may
also say, for example, that it is part of the content of The Scarlet
Letter that unacknowledged guilt leads to perdition. He may mean this
merely as a claim about the characters of the work; in this case he
would merely be making an interpretive claim. Or he may mean that the
novel contains a general truth about the actual world; in this case
he is making an interpretive claim which is also a critical significance
claim. In either case, a problem arises. Such claims cannot be
justified by showing that 'Unacknowledged guilt leads to perdition'
occurs on the pages of the novel. And even if the critic could so
justify his claim (even if the sentence occurred, say, as a chapter
heading), that is not what the critic intends. What he wants to claim
is that the work as a whole gives rise to or contains this truth. The
procedure for justifying such claims about the content of a work is un-
clear, because the way in which the alleged contained truth (or false-
hood) is related to the work is far from clear. Interpretive claims
about the content of a work are not justified merely by pointing to
expressions which occur in the work. 'Contains' and 'gives rise to'
are vague terms. They themselves call for an explanation before we can
proceed to the defense of critical claims in which they occur. In this
chapter, I will consider two historically important ways of solving
these perennial problems.

One suggestion, that which I shall call "the proposition theory,"
is that we can solve both perennial problems at once by claiming that
works of epic literature contain general truths or falsehoods about the
actual world as statements or at least as quasi-statements. On this
view, a critical significance claim to the effect that a given work has a given "meaning" or "moral" will be considered justified just in case the proposition is contained in the work. And works of epic literature, on this view, have cognitive value because by means of them we are "told" something about the world. Whether we believe what we are told or not, of course, depends upon the reliability and breadth of experience of the author and the antecedent plausibility of the things he tells us. Works of literature have cognitive value on this view because after reading them we have it on the authority of the author that something is the case. In section 1.1, I will argue that no version of the proposition theory enables us to deal effectively with the questions relating to the cognitive value of a work of epic literature. This is because we are unable on this theory to explain the justification of critical significance claims or how they may be taken as evidence about the way the world is, or both. As a consequence, we cannot explain the alleged cognitive value of works of literature by claiming that they "literally tell us something."

The temptation at this point is to claim that the alleged cognitive value of works of epic literature is only apparent and that there is no way of securing objectivity with regard to interpretive claims, and a fortiori with regard to critical significance claims. In section 1.2, I will deal with this "emotive" analysis of the alleged cognitive value of works of epic literature as advanced by Stevenson and Richards. I will argue that in some respects, this kind of theory is preferable to the proposition theory in that it denies that works of epic literature are disguised statements. But the "emotivist" position goes too
far in denying that we can justify interpretive claims at all. It is this wrong move which leads ultimately to the denial that works of epic literature have cognitive value.
1.1

In this section, I will examine the suggestion that at least some works of epic literature have cognitive value because they somehow contain as propositions about the actual world the general truths we claim to learn from them, and which we can choose to take as evidence about the way the world is (depending on our view of the author's competence), construing them as the author's statements.

Let us understand a proposition to be whatever is true or false. Then this suggested solution to the problems of the cognitive value of works of literature and the justification of critical significance claims can be seen as an appeal to the proposition theory for literature. In its general form, the proposition theory is a theory about the nature of a work of art, i.e., the theory that a work of art contains truths or falsehoods, or perhaps itself is true or false. It has been pointed out on numerous occasions that the proposition theory is extremely implausible as a general theory about the nature of art. But in the present case, we are dealing with works of epic literature, which seem prima facie to contain propositions. They are, at least, linguistic. The suggestion, then, is that we account for the apparent cognitive value of works of epic literature and their relation to general truths (or falsehoods) about the actual world by claiming that these general truths (or falsehoods) are part of the propositional or true/false content of the work. This view has some intuitive plausibility. I shall attempt to show in this section that this plausibility is only apparent, that the proposed solution does not work. I will advance three ways of explicating the notion of the propositional content of a work and then argue that none of these is any help in explaining the cognitive value of works of
epic literature. I think that these three alternatives are exhaustive, that almost any version of the proposition theory can be shown to be a variant of one of them.

The first specification of the propositional content of a work of epic literature is as follows: The propositional content of a work of epic literature consists of whatever is present explicitly in the work and is true or false. The notion of "being explicitly present in a work" can be clarified as follows. In a novel, we find such apparent event-descriptions as the following, which occurs in The Scarlet Letter:

The stranger had bent his eyes on Hester Prynne. It was carelessly at first . . . very soon, however, his look became keen and penetrative. A writhing horror twisted itself across his features. Such sentences I will call explicit reports, by which I shall mean a sentence of a novel which, outside a work of literature, would normally be used to state a fact. In addition, sometimes we find in a work of epic literature expressions such as the following:

Such is frequently the fate, and such the stern development, of the feminine character and personality, when the woman has encountered and lived through an experience of peculiar severity. If she be all tenderness, she will die. If she survive, the tenderness will either be crushed out of her, or—and the outward semblance is the same—crushed so deeply into her heart that it can never show itself more. Such sentences I will call explicit reflections, by which I mean those sentences in which the narrator ostensibly states an opinion about the events of the narrative or generalizes on them. For the present we leave it open whether this is a comment on Hester and her contemporaries or a comment on women of the actual world or both. What is being claimed is that the explicit reports and the explicit
reflections of the work comprise the sentences of the work which can be regarded as straightforwardly true or false, i.e., its propositional content. We can now give a precise formulation of the first version of the proposition theory:

(FL) The propositional content of a work of epic literature consists of all and only the explicit reports and reflections of the work.\(^1\)

(FL) has the advantage of allowing us to state clearly the sense in which works of literature are propositional, and, given an appropriate semantics, to locate the true propositional content of a work. There is a clear sense in which the author may, on this account, sometimes be said to tell the truth, i.e., when he happens to include in his work truths about the actual world. But of course, no one interested in the proposition theory as an account of critical significance claims or the cognitive value of a work of epic literature would take (FL) seriously. It is, however, interesting to see why the suggestion is unhelpful for these purposes. (FL) fails for an obvious reason.

It turns out that on the theory just propounded, the number of propositions contained in a work is very small indeed. It is sometimes true that reflections are made explicit in a novel, as in the above example. But very often this is not the case. Nowhere in The Scarlet Letter, for example, do we find the reflection made explicit that unacknowledged guilt leads to perdition. Yet we are inclined to say that this or something like it is true of the characters and, as applied to the actual world, is the "moral of the story." Such a reflection I will call an implicit reflection.\(^{11}\) It is easy to see that according to (FL), such implicit reflections are not part of the propositional
content of the work, whether they are construed as being about the characters of the work or about the actual world, or both. And this group includes many of the truths we are most interested in when we are concerned with the cognitive value of the work. In addition, presented with the above explicit report, we are inclined to say that Roger Chillingworth recoiled in horror because he saw Hester, and that this is somehow present in the passage in question. But this is in fact not the case. We conclude—we are not told—that the causal relation between the events obtains. The causal reportive proposition is not explicit; one of the symbols is missing, i.e., 'because'. We have what I shall call an "implicit report." Critical claims about such implicit reports are also interpretive. And if (PL) gives a correct account of the propositional content of the work, then this report is not part of that propositional content. Thus (PL) circumscribes the content of the work much too narrowly to be of much use in any interpretive claim, and a fortiori in justifying a critical significance claim. Obviously, therefore, it does not help in explaining the cognitive value of works of epic literature.

So much for the usefulness of the first attempt to specify the propositional content of a work. But there are other moves left for the proponent of the proposition theory. One such move is the following. The problems posed for the proposition theory interpreted as in (PL), he might say, result from taking the explicit reports and reflections of the work as the sole constituents of the propositional content of the work. This, the proponent of the proposition theory may claim, is all a mistake, for the explicit elements of a work are not even the most important thing which we have in mind when we claim that works of
epic literature have propositional content. We mean rather that
certain general or universal truths about the actual world follow
from the work. Implicit reports are explained in a similar way.
Insofar as the explicit reports and reflections of a work do have truth
values, they are part of the propositional content of the work. But
the general truths which follow from them are more important. We thus
need to extend (FL) accordingly:

(P2) The propositional content of a work of epic literature is
constituted by all and only (i) those true or false reports
and reflections which are explicit in the work; and (ii)
those reports and reflections which are implied by the
explicit reports and reflections.

The rationale behind this extension of (FL) is clear enough. What we
are trying to explicate is how works of epic literature "tell the
truth (or falsehoods)," in order to explain their cognitive value and
the justification of critical significance claims. Certainly A can be
said to have told B by implication something which follows obviously
from what A said explicitly. If works of epic literature tell us im-
portant truths by implication in this sense, then their cognitive function
can be explained. And a critical significance claim would be justified
just in case the proposition it singled out was indeed explicit in or
implied by the work.

One thing which should be perfectly clear is the distinction
between A's telling B the truth by implication and A's telling B by
implication something which is true. It is not sufficient for A's
telling B the truth by implication that something which follows from
A's explicit linguistic performance happens to be true. If A is
lying or mistaken in his explicit performance, then he cannot be telling
the truth by implication, even if something which follows from what he says is true. It is only in the case where A tells B the truth by implication that A's linguistic performance can be construed as evidence of the reportive kind for B's knowledge claim. The importance of this distinction will become evident in the succeeding discussion. It is only if the author can be shown to be telling the truth by implication that (P2) offers an explanation of how works of literature tell the truth and hence an explanation of the cognitive value of literature. It is also clear that if (P2) is to be an adequate explanation of the cognitive value of literature, it must allow us to explain why some critical significance claims and not others are justified.

The main problem with assessing the adequacy of (P2) for our purposes is that, whereas the meaning of "explicitly present" is tolerably clear, the meaning of "implied" is very murky indeed. On (P1) it was easy to show what the propositional content of a work was. But in order to evaluate the real usefulness of (P2), we need to know precisely what is meant by "implied" in (P2ii). Once we know this, we can decide whether (P2) is useful in explaining the cognitive value of works of epic literature and the justification of critical significance claims.

What we need is a precise account of the implication relation which is claimed in (P2) to hold between the work of literature and the proposition it allegedly implies. I will examine a series of possibilities. It may be that I have not got them all, but some can be handled more conveniently in my subsequent discussion of (P3).

Let us start with a fairly standard sense of implication in which
both the antecedent of the implication relation and the consequent are propositional. Now the most pressing question becomes: What constitutes the antecedent of the implication relation? The most obvious answer to the question posed in this way is that the set of explicit reports and reflections of the novel constitute the antecedent. But at the outset we note the following problem. The reports and reflections explicit in a work of epic literature can be considered propositional; but if they are so considered, they are also often clearly untrue. The story which the novelist tells may, and often does, consist of many explicit reports which are straightforwardly false on a historical model. This becomes a major problem in trying to find a sense of implication in which the novelist can be said to tell us the truth by implication.

I think we can discard at the outset any strictly formal sense of logical implication. Similarly, the claim that the implied propositions are implied by the work precisely because the antecedent propositions are largely false is unacceptable. This amounts to saying that we should treat 'implies' as \( \Rightarrow \).\(^{13}\) Merely note that when a false proposition occurs in the antecedent, we can claim that, where \( P \) stands for the conjunction of sentences in the antecedent, \( P \implies Q \) is true where \( Q \) is any proposition whatsoever. The implication relation we are looking for to explicate the cognitive value of works of epic literature is certainly not one which holds between a work and every proposition whatsoever.

For part of the reason for adopting (P2) is to show us how the author tells us some general propositions and not others. Therefore, if (P2) is to be helpful, 'implied' in (P2 ii) must have some other sense.

Let us consider the suggestion that the alleged implied propositions
are such that the statement of the implication relation is a truth table tautology. To this suggestion, there are two objections. First, in fact most of what we take to be clear cases of implicit reflections are not such that the connection between them and the explicit propositions of the work is a logical tautology. Secondly, consider the following instance of such an implication: 'The stranger had bent his eyes on Hester Prynne ⊃ the stranger had bent his eyes on Hester Prynne v Pearl was playing in the woods'. This does not seem to be the sort of thing which (P2) was formulated to account for.

To say that the propositions covered by (P2 ii) are those which are related to the explicit sentences of the work by truth table tautologies thus gives us a criterion which seems both too wide and too narrow. It allows us to justify too many claims which we would not want to advance as interpretive claims about the work, and a fortiori not as critical significance claims. And it does not allow us to justify many antecedently plausible interpretive claims about the work or about what the work "says" about the actual world. The defender of (P2 ii) might try to meet the claim that the criterion is too wide by adding a sort of "criterion of importance." All of the tautological consequences of the work are legitimate cases of (P2 ii), it might be claimed; but only some of them are significant and important. But this proposal also meets with an objection. Consider the following:

The stranger bent his eyes on Hester Prynne

The stranger bent his eyes on Hester Prynne v Nixon is president of the United States.

This inference would seem to be valid. But in fact, the conclusion of
the above inference does not seem to follow from the novel in the sense that it is one of the things we consider the author to be telling us at all. This suggests that it is not merely the case that some of the propositions included in the propositional content of a work by (P2 ii) interpreted in this way are unimportant. Some of them, it seems, are not the sort of thing we want to say the novelist tells us by implication at all.

But there are other possibilities for interpreting (P2 ii). Let us consider the suggestion that the connection between the work and the implied propositions is one of entailment. Clearly, on this account, we can explain how some interpretive claims and a fortiori some critical significance claims are justified and some are not. But here we come up against a further problem. In the case of a work of epic fiction, a good many of the explicit reports and reflections are literally false evaluated on a historical model; countless claims made in novels about their settings are false, for example. In addition, many reports are, depending on one's view, either false or non-truth-valued because they have non-referring subject terms. In such a case we cannot plausibly say, it seems, that the novelist has told us the truth by implication just because something entailed by his (largely untrue) performance happens to be true. If A says, "The present King of France is a bald Frenchman," which clearly entails 'There is at least one bald Frenchman,' which happens to be true, A certainly cannot be said to have told B the truth by implication. And B cannot be said to have learned the truth that there is at least one bald Frenchman from what A said, because A's remark is not good evidence for the entailed proposition.
And to restrict ourselves to cases where the explicit reports and reflections which constitute the antecedent are true is out of the question. We would thereby avoid the problem, but we would also exclude from consideration the sentences in which we are most interested. Thus if we interpret "implied" in (P2 ii) as "entailed," we seem in general unable to account for the cognitive function of works of literature.

I am not aware that anyone has suggested the following explanation of how "implied truths" are in literature, but there is no harm in exploring it. Given that a large number of the purported antecedent propositions are false (though some of them may be true), yet the conclusion is held to follow anyway, it seems that counterfactual analysis might be fruitful. A first attempt:

If Hester and Arthur had committed adultery and then paid for it in proportion to their acknowledged guilt, then it would have been the case (or would be the case) that unacknowledged guilt brings the greatest suffering.

is not quite satisfactory. We want to be able to assert, not merely a connection between the unfulfilled indicative associated with the antecedent and the consequent, but the indicative claim associated with the consequent detached from the antecedent. This seems to be captured by a second formulation, in which the implication has counterfactuals as antecedent, and an implied proposition of a lawlike character as consequent. Thus we formulate:

If Hester had sinned and confessed, she would have suffered comparatively little.

If Arthur had sinned and not confessed, he would have suffered horribly.

Unacknowledged guilt causes more suffering than acknowledged guilt.
If this account were correct, then we should expect that with more counterfactuals of the same kind, the conclusion would seem more strongly supported. As luck will have it, there is at least one more instance, Hester's husband. We can assert that if Hester's husband had not acknowledged his guilt at all, he would have suffered the most horribly of all. And it is pretty clear that the third case does round out and support the implied reflection.

This account is rather seductive. But there is one serious problem. The present theory is supposed to show how propositions which are not explicit in the work of fiction are implied by reports and reflections which are explicit in the work. And the antecedent counterfactuals are not directly asserted in the work, any more than the law-like conclusion. What is explicitly present is that Arthur and Hester committed their sin, that Hester had no choice but to own up and suffered less than Arthur, that Arthur concealed his part and suffered long and horribly. Hawthorne does not explicitly connect the events. At least he does not explicitly connect the propositions as counterfactuals.

It seems clear, however, that if we could legitimize the counterfactuals, they would support the general truth. We might, taking this line, claim that the counterfactuals themselves are not explicit—because they themselves are implied by the explicit reports and reflections of the work which serve as evidence for the counterfactuals. This accords with the general understanding of counterfactuals as formulated on the basis of evidence. But clearly the same difficulty will arise here as arose with the very first version of the theory that general truths are implied by the explicit elements of the work. The basic
evidence consists primarily of false propositions. The indicative 
associated with the antecedent of a counterfactual may be false without 
disturbing the truth of the counterfactual, but if the evidence for the 
counterfactual is false, we cannot very well claim that it is acceptable.

It is fairly easy to see that any attempt to treat the relation 
between the explicit reports and reflections and the alleged implied 
propositions as an inductive relation will have the same problems. The 
implied reports and reflections themselves are supported by predominantly 
false premises. And knowledge claims are out of order since the evidence 
central to the conclusion consists primarily of falsehoods.

I do not see any other sense of propositional implication which 
will escape these difficulties. If works of epic literature themselves 
are treated as pieces of fact-stating discourse, there is no way, I 
think, to escape the conclusion that we are not often in general told 
general truths in works of literature, in cases where those truths are 
not explicitly present. Very simply, if B is straightforwardly lying, 
there are, of course, truths we can learn from his lie. We can learn, 
for example, that B is not a trustworthy character, etc. But we are 
not told these by B. Nor are we told the truth by B if some obvious 
consequence of his lie happens to be true. The assumption that works of 
literature are fact-stating discourse, coupled with the observation that 
the reports and reflections are predominantly false is fatal to the 
view that works of epic literature tell the truth in any significant 
sense.

These problems seem to a large extent to be due to the assumption 
that works of epic literature are to be treated simply as predominantly
false pieces of fact-stating discourse. From this assimilation of works of epic literature to normal narration, two consequences follow: (i) the reports and reflections of the work are assessed for their truth on a historical model, i.e., as if they were history of the actual world; and (ii) provided it is from the work as discourse that we gain the alleged knowledge which corresponds to what we have characterized as implicit reflections, then what we learn must be told us by implication. It is easy to see that (i) and (ii), together with the obvious fact that works of epic literature contain a good many untrue sentences, suffice to lead to the undesirable conclusion that we learn almost nothing from epic literature. For this reason, if there is some alternative to treating works of literature as fact-stating discourse which does not lead to these conclusions, it would be preferable. Finding such an alternative means rethinking our treatment of the explicit reports and reflections of the work of literature.

One alternative is to hold as a special thesis about the language of epic literature that it is simply a mistake to attach any importance to the truth values of fictional sentences, whose function is based instead on their meaning, that in concentrating on the literal truth of the reports, we confuse poetry with fact-stating discourse. In fiction, one may say, we learn conventionally to ignore the truth values of explicit reports and reflections and go on to other levels of meaning. This theory is specifically a theory about the language of literature, and it is in open conflict with the assumption that fictional language is merely predominantly false fact-stating discourse. As a theory, it opens up a new specification of the propositional content of a work of
epic literature, in which the implicit reports and reflections are accounted for, not on the basis of the truth values of the explicit reports and reflections, but on the basis of their meanings.

This view is essentially the view held by Weitz and Beardsley. Beardsley seems to hold flatly that the explicit reports of a work are never used to make truth claims and that the same is true of reflections, implicit or explicit—for they do not "evince belief," are not advanced with the claim that they are true.\(^{14}\) Weitz holds in his later work only that our theory must be able to explain the presence of general truths even on the supposition that the explicit reports and reflections are emotive in nature, by which he means that they are not used to make assertions or to communicate information.\(^{15}\) On these views, the explicit reports and reflections are not (or at least not all) to be treated as truth claims, although they may be true or false. In fact, the theory goes, they seem to have a totally different function; and it is from the explicit reports and reflections in their non-assertive function that the implicit propositions follow. Naturally, the implicit propositions do not follow by propositional implication, but it is nonetheless true that they are contained in the work by virtue of the meanings of the explicit reports and reflections. Heuristically speaking, the kind of implicit containment is this. Note, the proponent of the theory points out, that in ordinary descriptive language words function as symbols of something else; in this sense we call such language "descriptive" or "symbolic."\(^{16}\) The language of fiction is not, on the level of explicit reports and reflections symbolic in this sense. But the entity—and event—descriptions of works of fiction function
with regard to the implicit reports and reflections in a manner similar to descriptive language with regard to the world. We can thus at long last state a third version of the proposition theory for epic literature:

(P3) The propositional content of a work of literature is constituted by all and only (i) those true or false reports and reflections which are explicit in the work; and (ii) those reports and reflections which are entailed by the explicit reports and reflections; and (iii) what is implicit in the work by virtue of its (symbolic) meanings and is true or false.17

Clause (P3 iii) is nebulous. Weitz and Beardsley offer two distinct ways of explicating (P3 iii): in terms of an "aesthetic theory of types," as Weitz proposes; or in terms of a paraphrase procedure, as Beardsley suggests. I think that in discussing these two attempts we can discuss the possibilities and limitations of this kind of approach exhaustively. I will attempt to show that neither kind of attempt is acceptable and that neither shows what it purports to show, how implicit reports and reflections are contained in the work.18

What Weitz says is that in some works of literature we find marks on paper which serve as signs of concepts or ideas. These concepts or ideas in syntactical juxtaposition constitute the sentences of the work. These are the "surface" or "first order" meanings. Some works also contain "depth meanings" or "second order meanings," which may be propositional in nature and which function as truth claims, though they do not appear in print. Weitz calls these depth meanings "second order" because "they are logical functions of the first order meanings."18

Examining the depth meanings of Native Son, Weitz claims that the character Bigger represents modern man in his own tortured self-
destructiveness; as a consequence of this, the novel signifies the tragedy of modern man, that he can attain autonomy only through destruction and eventual self-destruction.\textsuperscript{19}

Weitz's references to philosophical theories of types are somewhat confusing. But perhaps by seeing what Weitz might hope to gain by his account, we can see what notion of type theory he has in mind.

One possibility is that Weitz has in mind a distinction between language used to talk about the world (object language) and language which is used to talk about language as used to talk about the world (metalanguage). Such a distinction has sometimes been called a "type distinction." Thus we might say that "Rome is burning" is an object language or a first order sentence and that "'Rome is burning' is a purported fact-stating sentence" is metalinguistic. In view of the situation in works of literature which seems to disturb Weitz, something would be gained if the implicit reports and reflections of works of literature were second order sentences in this sense. Notably, it is certainly the case that the truth of metalinguistic claims (with the exception of those containing semantic predicates such as 'true' or 'false') are independent of the truth values of sentences of the object language. We can make true claims in the metalanguage about object language sentences which are false or which have no truth value because they are not the sort of syntactic unit to which we properly assign truth values. This would be welcome to Weitz in the context of the problem under discussion, for he concedes the possibility that the sentences of literature may be "emotive," by which he means, presumably, at least that they are non-truth-valued. There would be a sense in
which the author could "tell the truth," while writing sentences which are not true.

But there is an objection to this view as so far stated. First, the view that the implicit reflections of a work of literature are "about" the first order sentences is counterintuitive. Secondly, it would be unwelcome to Weitz, for his examples of "second order truths" are clearly general propositions about the world. Weitz considers this objection serious enough to answer. And this is good reason to suppose that at least some of the time he has the object-language/metalanguage distinction in mind.

But the answer Weitz gives to this objection is thoroughly confusing. He says:

But I cannot see why these truth claims cannot be both about the printed sentences in the novel and about certain non-linguistic social conditions in exactly the same way that "Napoleon had all the characteristics of a great general" is both about certain first order sentences and about a non-linguistic person named "Napoleon." 20

Weitz gets this example from Russell. "Having all the characteristics of a great general" he supposes to be of higher type than 'having cunning', from which he supposes it to follow that the alleged sentence of higher type is about the sentence 'Napoleon had cunning' and about the world, i.e., about Napoleon. But, contrary to what some philosophers have thought, the alleged sentence of higher type is not about the sentence 'Napoleon had cunning', though it may in some informal sense imply it. If it is of higher type, it is so precisely because it is about Napoleon and a real property, a real property which seems in fact to consist of a group of properties. It is a little odd to call this "being of higher type." And the sentence in question is assuredly not
about Napoleon and a group of sentences. Thus, once the matter is straightened out, if Weitz has in mind the object language/metalanguage model, on which the truth value of the sentences of higher type is independent of that of the sentence of lower type, he must hold that the alleged higher type sentences are merely about the sentences of the work, a view he explicitly rejects.

Well, perhaps he is serious about his example, at least. What is the relation between 'Napoleon had cunning' and 'Napoleon had all the characteristics of a great general'? We might say that the latter implies the former in the sense that the natural kind "great-general-making-attribute" subsumes "cunning." There are systems for working out such implications, so-called "theories of types for natural languages." But then the truth value of the alleged sentence of higher type (presumably the implicit reflection) is not independent of the value of the alleged sentence of lower type, the explicit reports. However loose, the implication relationship, if 'Napoleon had cunning' is false or neutral, then 'Napoleon had all the characteristics of a great general' cannot be true. That is, reverting to the original problem, the novelist cannot on this interpretation of "type theory" be held to "tell the truth" by means of his untrue sentences.

Neither theory to which Weitz could be appealing will allow him to say both that the truth values of the implicit reflections are independent of that of the explicit reports and that the implicit reflections are about the actual world, not about the sentences of the work. Weitz is trying to have the best of both worlds, but the theories are incompatible as accounts of the same phenomenon. We could go still further into what
Weitz might mean, but it seems best to regard his claims at best as an illuminating metaphor couched in misleading formal terms.

What Weitz means on an intuitive level is fairly clear. He means, obviously, that the character Bigger "stands for" the whole "type" of which Bigger is a representative (whatever "representative" means).

Weitz emphasizes the fact that we more or less see through Bigger to the totality of tortured modern man. Bigger may in this sense to be archetypal. Why Weitz ever thought his formal presentation was a clarification of this kind of informal theory is puzzling. Possibly he was misled by his misreading of Russell's view. The main problem with the archetype theory posed as such is that it provides no explication of (P3 iii). It is never clear how the character stands for the "type" he supposedly represents. Does Bigger represent modern man simply by being a modern man? And even if it were clear how a character represents a type, so that we could say that the type was contained in the work via the character, we would still not have explained how propositions are contained implicitly in the work.

Beardsley's theory may be seen ultimately as an extremely sophisticated version of the archetype theory proposed to account for the sense in which we can say that implicit reflections are in the work intrinsically as its meaning. Beardsley argues against the proposition theory as a general theory of art. His argument is that the propositions which we hold are contained by the art work are in the case of many art works, just not there. Such works are not linguistic. They have no identifiable syntax and semantics on the basis of which we can locate a subject (indexical) element and a predicate (characterizing) element
for the alleged propositions. We are therefore not justified in making claims about the meaning of such a work, where the alleged meaning is envisioned as some general proposition about the actual world. But in the case of literature, he concedes, we are dealing with a linguistic object to start with. Works of literature may thus be propositional. The basic problem with works of literature is that the language used explicitly does not account per se or by implication for what we have called implicit reports and reflections. Beardsley's theory is an attempt to show how these propositions are in the work. I will argue that he does not show this; he fails to show what elements intrinsic to the work are responsible for containing the implicit propositions. What Beardsley means by "intrinsic to the work" is far from clear. Basically, he means that for a critical significance claim 'W means p' to be true of a work W, it is required that p be contained in the work in some way. The requirement is the ultimate product of the "new critical" approach, on which we must confine our attention to the aesthetic object under consideration without importing information from outside the work. Thus it is the intrinsicist thesis which compels the theorist who recognizes some cognitive significance claims as legitimate to hold some version of the proposition theory. But once the proposition theory is espoused, the intrinsicist position is a natural (perhaps even a theoretically indispensable) adjunct to his position. For if works of epic literature are in fact a way of conveying truths (or falsehoods) about the actual world, then the reader, once he understands the "code" (the special syntax, semantics, and context relevant features of the "language"), should be able to understand just what is being said. No doubt readers
of allegories are able to do just this. It is important to note the connection between the proposition theory and the intrinsicist thesis, for it is the latter which proves fatal to Beardsley's position; and it is not an idiosyncratic weakness.

On Beardsley's account: (i) it is the activity of the narrator which conveys by suggestion the implicit reports and reflections which we infer; (ii) we arrive at the implicit reports of a work by noting that the juxtaposition of elements of the narrative suggest further reports; (iii) we arrive at the implicit reflections of a work on the basis of the narrator's performance, the generalized significance of what he says, which "suggests" further reflections.

Let us first consider the problem of implicit reports, for it is here, I think, that Beardsley's account is most obviously inadequate. Consider the two sentences of The Scarlet Letter quoted earlier: "His look became keen and penetrating. A look of writhing horror twisted itself across his features." Earlier we said that the element which would provide a causal link between the two reports is simply missing. Beardsley says that event descriptions "suggest" by their juxtaposition the causal connection between them. What is it about the juxtaposition of these two event descriptions which "suggests" the causal connection? Certainly not the fact that one follows the other. The sentences 'The day was chill and somber. Overhead was a gray expanse of cloud, slightly stirred, however, by the breeze. are juxtaposed in this way. And here there is no temptation to supply the causal connective.

The exposition Beardsley gives earlier in the book of how a proposition or attitude can be "suggested" by others is not particularly helpful.
There is, of course, a sense in which "He has nice handwriting," advanced as a recommendation for a candidate for an academic position, suggests that the candidate is not professionally competent. It does so by ironic understatement. But surely this sense of "suggest," dependent on the "special effects" properties of language, cannot be what Beardsley means in the commonplace cases of causal connections and their absence.

Perhaps we might try to explicate Beardsley's "juxtaposition" as follows. It is in some sense the meanings of the two reports in the first instance taken together which combine to suggest the causal connection, though not in the second case. The answer to this, I think, is that it is not the meanings of the two sentences which lead to the causal connection. There may be "meaning implications" in the cases of truths usually described as "analytic" or "synthetic a priori." But this is certainly not such a case. I should be inclined to say that it is our knowledge of the way the actual world is, of perception and human psychology in the first instance, and of the weather in the second, and our claim that the fictional world is constructed in relevant respects in a way similar to the actual world, which suggests the inference in the first case and excludes it in the second. But this commonsense alternative is not open to Beardsley, given his strict requirement that the proposition in question be identifiably internal to the work. On Beardsley's own view, the commonsense little proposition we are looking for is just not there.

Beardsley's suggestion that it is the narrator's performance which suggests the connection does not improve on this in any important
way. His suggestion is that:

What is suggested or implicitly predicated is what the speaker purports to believe about the situation he describes, but without saying so in so many words.27

There is a clear sense in which the narrator's performance is just the sentences we are considering. I have already argued that the proposition is not to be found there. But even if we allow that the narrator element adds something to the facts as he relates them, we merely move the problem up a level. For if something is added, it has to do with the psychology of the narrator. Now it is our knowledge of the world as it is and our knowledge of the relation of the narrator's performance to usual human psychology which backs the inference. I should be perfectly prepared to accept this, but it certainly does not show that the predication in question is intrinsic to the work. And on Beardsley's own view, the interpretive claim is therefore indefensible, although perhaps not inexplicable genetically. Beardsley has not shown successfully that there are elements intrinsic to the work which have the meaning which is allegedly present in the implicit report.

Beardsley's account fares little better in explaining implicit reflections. The claim, it seems, is that somehow the narrator's manner of relating the events suggests their generalized significance, that he is judging the events and reading some generalized meaning into them.

As an illustration of narrator emphasis, Beardsley writes:

A long novel in which poor people are constantly exploited by their employers, landlords, stores, politicians, and policemen can hardly help being a predication about social relations, even if the narrator makes no explicit generalizations.23

But the only sense in which the narrator "suggests" the conclusion is that he recounts the relevant incidents. It is not anything about the
narrator's performance, it seems, which leads to the conclusions, but rather our knowledge of the way the world is, the laws of induction, etc. Of course, there are cases where the narrator "suggests" an implicit reflection by his irony, etc. But it is clear from the above reference that Beardsley does not want to restrict his theory to cases of the "intrusive" or "editorializing" narrator.

We may yet try to save the theory in the following way. Certain individuals and events which turn up prominently in the narrative symbolize the theme of the work, providing the indexical part of the implicit reflection (we can push aside the nasty question of whether we have to wait until the element has occurred repeatedly to realize that it has special symbolic significance). The rest of the work furnishes symbolically the characterizing element. In The Scarlet Letter, it might be claimed, Hester is pretty clearly the symbol of expiated guilt. Her expiated guilt is at least part of the cause of her ultimate state. Thus it is at least part of the propositional content of the work that expiated guilt brings peace and dignity. This we might represent schematically as follows:

```
Expiated guilt  leads to  peace and moral integrity
symbolically means  symbolically means  symbolically means
Hester            causal connection         Hester's lot as described by the sentences of the work
```

The alleged implicit reflection is thus contained "symbolically" in the work. I think this sort of thing is pretty clearly what Beardsley has in mind.

But there are some problems which this proposed way of explicating
the sense in which a work of literature has a given meaning. The claim that a poem of a certain kind has a given meaning because its elements are symbolic has some plausibility. Consider Beardsley's example:

I

Among twenty snowy mountains,
The only moving thing
Was the eye of the blackbird.

IV

A man and a woman
Are one
A man and a woman and a blackbird
Are one.

VII

I know noble accents
And lucid inescapable rhythms;
But I know, too,
That the blackbird is involved
In what I know.

It is fairly plausible that even to make sense of the poem we must interpret the prominent figure of the blackbird. And whatever the blackbird stands for, it is being called to our attention with each occurrence of the blackbird symbol. But how plausible is this as a claim about the events and characters of a work of epic literature? Prima facie we can make good sense of The Scarlet Letter without understanding what Hester "stands for." There seems to be good reason to maintain that, whereas the blackbird has to be interpreted primarily as a symbol, nothing of the sort is true of Hester Prynne. However nebulous our conception of a symbol, there are in many novels elements which function straightforwardly as symbols in this primary sense. But the main characters are usually not among these elements. The view that a plot
must be treated as an extended metaphor or symbolic structure is a view thoroughly typical of the critic who holds, as Beardsley does, that the cognitive value of a work must be explained on the basis of propositions contained in it symbolically or metaphorically. And this view has led to some exemplary instances of critical madness. Beardsley is too sane a critic to say that a novel is an allegory; but talk of symbolic paraphrase is a coverup for reading it as one.

But perhaps there is a commonsense way of construing this talk of symbols which does not have such extreme consequences. One plausible suggestion is that the supposed individuals of novels are merely picturesque and interesting devices for presenting us with congeries of properties, those properties which are ascribed to the pseudo-individuals. We arrive at the meaning of the novel by noting the properties of the pseudo-individuals and then noting how the pseudo-individuals interact. The "individuals" become a convenient and picturesque code for shuffling properties, and the "paraphrase procedure" becomes a matter of looking at the properties and relations and their interrelations. With all due apologies to those who love symbol-talk for its vagueness, this account is coherent and initially plausible.

But there is a serious problem with the account thus presented. The procedure described is most suited to crude allegories. In a crude allegory, nearly everything attainable by paraphrase is, intuitively speaking, part of the content of the work. This is presumably not the case with epic literature. Hester expiates her guilt, but she also has black hair. The problem, in short, is how much of the literal fact of the work is carried over into the symbolic paraphrase. The heroic path
is simply to say "all of it." But surely this provides us with too wide a criterion for what the "implicit reflections" are. We need, in addition, it seems, a criterion of importance. The work, we will say, has all the symbolic meanings we can dredge out of it, but only some of them are important. I am at a loss to think of such a criterion which is not flatly question-begging or which does not make an appeal to the way things are in the actual world.

The heroic path also provides us with a criterion which is too narrow. It is never actually said in The Scarlet Letter that Hester expiates her guilt—we are told that she sits on the scaffold and bears the ill-will of the townspeople. We will need some way to go from Hester's having these properties to her having expiated her guilt. Of course, an appeal to the way the world really is is probably the answer to this question. In taking this view we construe fictional works as an imitation of narrative and fictional worlds as an "imitation of life." If we take this view, most of Beardsley's comments and explanations fall into line. They are even helpful and insightful. Proponents of "imitation type" theories are intensely disliked by Beardsley; they go outside a work of art to explain its meaning and function as a work of art. But the requirement that the "meaning" of a work must be traced entirely to its "intrinsic elements" seems to be too strong. We will return to the actual methodological flaws of this requirement in chapters IV and V.

One thing seems clear. We are not really getting anywhere with (P3 i ii). And (P3 i ii) was more or less the last alternative for showing that the work of epic literature has cognitive content because
the author somehow literally tells us by means of his performance the propositions we claim to apprehend in a work. I conclude that the proposition theory as explicated here is not an answer to the question of how to explicate the alleged cognitive value of works of epic literature.
The appeal to the propositional character of works of epic literature as explicated in (F1)--(F3) has proved unproductive. We are left with certain pre-analytically obvious features of works of epic literature to explain—or explain away, as the case may be. First, works of epic literature seem to expand and clarify our knowledge of the actual world, their "apparent cognitive value." It is often claimed that certain general propositions about the actual world are associated with certain works; such claims we have called "critical significance claims." In addition, it has emerged in the discussion that there are certain complex propositions about the world of the work itself which are claimed to be justifiably derived from the work, but which are not in any obvious sense "contained" in the work. Such propositions, and the propositions about the actual world associated with the work all fall under the classification of "implicit reports and reflections." To claim that such a proposition is associated with or derived from a work is to make what has often been called "an interpretive claim." In the discussion which follows, I will often use this latter term. This should cause no confusion, so long as it is clear that an interpretive claim may be about either a proposition about the world of the work or a proposition about the actual world. The second problem, then, is that we still have no explanation of how implicit reports or reflections are related to works of epic literature, and thus no procedure for justifying interpretive claims.

The expedient of treating a work of epic literature on the model of fact-stating conversation, in which we are told certain things which we then construe as reportive evidence for knowledge claims has proved
unsuccessful. We are not told explicitly or by implication, as suggested by (P1) and (P2), all the general truths we associate with such works. And what we are told explicitly or by implication is not told us in such a way that we can account for the apparent cognitive value of the work of epic literature. The attempt to treat works of epic literature as a special kind of conversation in which we treat the performance as a kind of code with special meanings (P3), by which we are presented with congeries of elements which convey general truths is also unacceptable for explaining how the wide range of implicit reports and reflections in which we are interested are related to works of epic literature.

The obvious alternative to an appeal to the propositional character of the language of epic literature is to try to give a different kind of account of this language which does account for our pre-analytic intuitions. A common suggestion is that we deny altogether that works of epic literature are to be treated on the model of fact-stating conversation in which we are said to have been told something which we then construe as reportive evidence for a knowledge claim. We are asked to note the many cases in which ostensible referring expressions in such a work of literature simply do not refer to the world and the irrelevance of such reference when it does happen to occur. This is good evidence, it is claimed, that such uses of expressions do not ever per se have the "symbolic" or "referring" function. We can draw further support for this view, it is claimed, by noting that 'I read it in a novel' is never per se good evidence for a knowledge claim about the world, but that, on the other hand, the fact that we have
just been reading a given novel is often a perfectly good explanation for a current emotion we feel or view we hold. It must be that the language of epic literature, though it does not tell us about the world, brings us to have certain feelings and attitudes about the actual world and the world of the work of art.

This is a fairly common phenomenon in language; we certainly are affected by non-fact-stating language, sometimes even caused to entertain beliefs. The correct model for the use of language in works of epic literature is thus, it is held, those utterances which function primarily to express emotions and attitudes of the speaker and to produce attitudes and emotions in the audience. Such language I will call expressive language. The theory that epic literature consists of such language I shall call the Expression Theory. We can thus state as the first tenet of the expression theory:

(E₁) The language of epic literature as such is expressive language and is used to express the attitudes and emotions of the author and to evoke attitudes and emotions in the audience.

The meaning of 'emotion' I shall take as fairly clear. An emotion is an occurrent feeling, such as fear or anger. What expression theorists have meant by 'attitude' is far from clear. On the one hand, they seem to want to distinguish the notions of "belief" and "attitude." The notion of an attitude as a general notion is that of a disposition to act in a given way. A belief, at least as Richards defines it, is said to be an inclination to act as if whatever its content is, is true. On this kind of account, it is difficult to see why beliefs should be opposed to attitudes. Accordingly I shall treat beliefs as a special kind of attitude. I shall allow that there are attitudes
which are not beliefs because they do not have propositional objects. I shall also allow that there are what Richards calls "emotive beliefs," which resemble genuine beliefs, but which are not "scientific" or "genuine" beliefs because their objects are not genuine propositions, although they may appear to be so at first glance. It has been held that readers do not acquire genuine beliefs from reading literature. Instead they acquire attitudes which are not beliefs. There are two reasons why theorists such as Richards might hold this. They may hold that the attitudes generated by reading works of fiction simply do not have propositional content; hence a fortiori we cannot act as if their propositional content were true. Or they may mean that such attitudes may have propositional content, but that we customarily refrain from acting as if that content were literally true. Either way, it seems, they are wrong.

It is certainly true that a reader may form a disposition to act in a given way upon reading a work; he may acquire the disposition to fly in the face of convention. In such a case, there is perhaps no proposition involved. Or a reader may acquire the attitude that Roger Chillingworth is evil, which Richards would probably call an emotive, i.e., non-propositional attitude because of the evaluative term 'evil'. But it seems equally likely that a reader may acquire on reading the novel the attitude that Roger Chillingworth is a very confused man, or that culture-related strictures are responsible for a good deal of unhappiness. In the first case, Richards might still cling to the view that this is an emotive belief, given that Roger Chillingworth does not exist, hence that the object of the belief is not a genuine proposition. But the second example seems to be a clear, commonsense case of a belief.
In a more extreme case, say that the reader by some quirk acquires the belief that the sea is salty. It seems clear that there is no basis on which even the historical expression theorists could deny in these cases that genuine beliefs are acquired. Once the issues are sorted out, and we have conceded that Richards may be right in holding that often what we acquire from reading works of literature are not beliefs at all, we must conclude that he has nonetheless certainly not shown that we never acquire beliefs from reading works of epic literature.

What characteristic might all attitudes acquired from reading works of fiction have in common, which would lead Richards to call them "emotive" or "pseudo-beliefs"? The distinguishing feature of the attitudes about the actual world acquired from reading works of epic literature is that, if they are so acquired, they are unsubstantiated by evidence, even bad evidence; they are acquired a-logically, from emotions. Henceforth, I shall use "mere attitude" to mean a belief or conglery of beliefs which is characteristically neither presented as a knowledge claim by the author (although perhaps it could be, had he chosen to deliver a speech), nor per se substantiated by anything which could be called "evidence", even bad evidence, for the reader in whom it is generated by reading the work. And I shall allow that there be "moral beliefs" and "beliefs about fictional worlds," as Richards would deny. By restricting 'attitude' in this way I have perhaps excluded one of the most common historical meanings of the expression, that of nonpropositional readiness to act. But we seem to have a general feature of the interesting attitudes, and we may suppose that this is
what Richards was getting at. It does not follow from the fact that such beliefs are not justified, Richards seems to think, that they never have genuine content, or that they are not unequivocally held as beliefs.40

On this sort of theory, then, the language of epic literature is construed as a species of expressive language, language whose primary function, on the one hand, is to convey the emotions and attitudes of the speaker, and on the other hand, to evoke emotions and attitudes in the audience. The language of epic literature is not, like the language of some lyric poems, what we should call *prima facie* "expressive." In another context, much of the language of epic literature could be used for a simple reporting of what was the case. It is the context of use, not the emotion-laden character of the language or the predominance of poetic ornamentation, which leads us to subsume such expressions under "expressive language." The basic scheme seems to be the following. By some generally understood convention, when presented with such language, we suspend concern with the literal truth of what we are "told." We are caused in some way by our apprehension of such a use of language to entertain certain (unsubstantiated) beliefs about the actual world and about the features of the world of the work, beliefs which do not, strictly speaking, correspond to what we are told. Similarly, we are caused to have certain feelings or emotions about the actual world and about the characters and events of the world of the work. In *The Scarlet Letter*, for example, we acquire the belief that Roger Chillingworth is a profoundly evil and sad man, and the general attitude that hate is a protean emotion. Or we may, for example, come to feel
indignation at the moral blindness of the people of the town and to feel indignation at the moral blindness of custom-bound moralizers in general. On the Expression Theory, the work simply does express these emotions and beliefs on the part of the author and excite such attitudes and emotions in the audience. The work is held to cause us to have the emotions or feelings about the actual world or about the world of the work. In addition, the causal force of the language explains on the one hand, our reactions to the work; such reactions form the basis of our interpretations restricted to the world of the work. On the other hand, the theory explains away what we were inclined to call the "cognitive value" of the work in terms of the causal efficacy of the work in shaping attitudes directed at the world. Works of epic literature on such an account cannot serve as evidence, even of a reportive kind, for beliefs about the actual world (excluding such trivial beliefs as that there is at least one novel with a character named Hester). It is, strictly speaking, false to say that we have been told a genuine proposition about the world by the author, and certainly wrong to say that we learned it from the work. But works of epic literature appear to serve as evidence because they cause us to acquire attitudes, which we mistake for genuine or justified beliefs. We make this mistake because often these attitudes and feelings about the actual world turn out on subsequent examination in the light of genuine information about the actual world to be justifiable or appropriate. Most of us would concede that the above example of The Scarlet Letter is such a case. This completes the illusion that we have learned something about the world from the work.42

Most expression theorists do not restrict themselves to claims
about the nature of language used in literature. They are also interested in giving an account of the critical language used to talk about a work. Stevenson, for example, is interested almost exclusively in this second question. According to both Richards and Stevenson, there is scientific or non-expressive discourse about a work of epic literature. We can discourse non-expressively about the connection between the author's state of mind and the work, when the relevant information is available. We are also discoursing scientifically about a work when we explain causally the reactions people do in fact have to a work of literature or delineate a complete set of possible responses to the work. According to the expression theorists, in many cases, we objectify our reactions to a work by attributing to the work an aesthetic property. What they seem to mean by "objectification" is better seen in a simple example. In many cases, instead of saying that an object presents a red appearance to an observer, we simply say that the object is red, omitting the reference to the observer and to standard conditions of observation. Such objectification does no harm, so long as we realize the nature of such claims, i.e., that they are actually descriptions of a reaction to a work with a built-in reference to observer and standard observation conditions. The meaning of a work seems to be taken by most expression theorists to be a complex aesthetic property of the work. To say that a work means $p$, where $p$ is a proposition, is just to say that the work has a $p$-meaning-property. One peculiarity of such complex aesthetic properties in general, according to Stevenson, is that there is no set of commonly accepted standard conditions of observation. When the critic objectifies a reaction and says that a work means $p$, he cannot
be saying that the work has the \( p \)-meaning-property when observed under standard conditions; instead, he is saying that the work has the \( p \) meaning-property when observed under the proper conditions, where 'proper' conceals recommendation of the critic to see the work in a given way in order to have the "correct" reaction to it. Thus, on the view of most expression theorists, when we speak of the interpretation of a work of epic literature, or an appropriate reaction to such a work, we as critics are using expressive language. Thus most expression theorists hold:

\((E_2)\) Interpretive discourse about a work of epic literature is itself expressive language.

This does not mean that there are no constraints at all on the interpretation of such a work. Most expression theorists hold at least that critics come to agree about the appropriateness of certain reactions to the world of the work or ways of seeing the work. In fact, they hold the stronger view that there are more or less correct ways of interpreting the work. But critical interpretations are not factual judgments about the work. Instead they are instructions for having appropriate responses to it. Granted that beliefs about the actual world may be associated with a work, they will say that a proposition \( p \) may be justified as a claim associated with a work if entertaining \( p \) is an appropriate response to the expression of feeling and attitude in the work.

The relation between \((E_1)\) and \((E_2)\) is somewhat murky. One might think at first glance that \((E_2)\) is entailed by \((E_1)\). If the language of epic literature functions to evoke attitudes and emotions, then, it might be claimed, when we make an interpretive claim about the work, we
merely re-express on our part the emotions or attitudes evoked by the work. Or at least, we record them. Of course such claims are emotive. And if we have said that a given work has a p-meaning-property, then to say that it is a correct interpretation of that work that it has the property is redundant.

But this is to miss the point that for the expression theorist there is a difference between reacting to a work or even announcing that one has so reacted, and claiming that a given response is a proper reaction. In isolation, the critic's claim that the work has a given p-meaning-property is ambiguous. It may be elliptical for the claim that the work seen under the proper conditions has the p-meaning-property. In this case, the claim that the interpretation is correct is redundant, for it is simply to say that to see the work under the proper conditions, under which it has the p-meaning-property is to see it under the proper conditions. But the critic's claim that a work has a p-meaning-property may merely be an objectification of the critic's reaction to the work seen in some given way, which he is not sure is the proper way. It might, for example, be part of the following sequence: "Sometimes (seen as an allegory) the work means p; sometimes (seen as a parable) it does not mean p; and I'm not sure which it ought to be seen as." In this case the claim that it is a correct interpretation of the work that it means p is not redundant, for it says that to see the work as an allegory, under which it has the p-meaning-property, is to see the work under the proper conditions. From the list of possible reactions under different observation conditions, it singles out the proper way to see the work, i.e. as an allegory. On the expression theory, all interpretive
claims, if they are not merely records of responses with a set of conditions noted, are covertly expressive, because they contain a suppressed recommendation to see the work in a particular way. Proper conditions cannot be reduced to 'standard conditions'. On this view, the critic hands out such directions because he, as well as the author, is engaged in creating and shaping the aesthetic responses of the audience by making him "see" the work in a special way. He takes up where the author leaves off.

But even if this view of the critic's enterprise were correct, the connection between (E₁) and (E₂) is a de facto connection, not a logical one. It rests on a shaky distinction between aesthetic and non-aesthetic properties and a specific claim about the function of criticism. If (E₂) is correct, then all interpretation claims are to be taken as quasi-imperatives. Their appearance of descriptive objectivity is merely their imperative force coming through. A good many theorists who hold some version of (E₁) apparently do not hold (E₂). And those who hold (E₂) argue for it on independent grounds. This is a fairly important point. For some version of (E₁) is pretty clearly true. The language of epic literature is not descriptive of the actual world. And it is pretty clearly true that our belief that, for example, Roger Chillingworth was a confused man is in some sense suspended, restricted to the world of the work. If this is what Richards means by "emotive belief," then we can concede his point, although we dislike his terminology. We can also concede that works of epic literature do not function as evidence for beliefs about the actual world, given the results of section 1.1; again, if this is what is meant by "attitude," then we can accept the point,
however much we dislike the nomenclature. In addition, we often do have, phenomenologically speaking, strong emotional reactions to works of epic literature which in some sense carry over to the actual world. Even though this is not always the case, it is interesting enough to have it pointed out to us by Richards. But all this is evidence for a very attenuated form of (E₁).

There really is no such plausibility argument for (E₂). There is an additional consideration. (E₁) as spelled out by the Expression Theory cannot solve the problems of what constitutes an appropriate reaction to work of literature, i.e., on (E₁) we cannot justify interpretive claims. We should expect (E₂), which is a thesis about critical language, specifically about interpretive discourse, to do this. But the addition of (E₂) in fact reduces the possibility of such a solution to two alternatives, both of which are clearly unacceptable. I will first consider the sense in which (E₁) is inadequate and then attempt to show that (E₂) closes off the possibility of making up the deficit.

There is no point in attacking (E₁) on the basis that there is no hard and fast distinction between expressive language and descriptive language or on the grounds that it assimilates the language of epic literature to the language of lyric poetry. The proponent of (E₁) needs merely to make the antecedently plausible claim that the language of epic literature is not used to describe the actual world and proceed to give separate accounts for epic literature and lyric poetry. The main criticism of such a theory, I think, is that it is too underdeveloped to explain adequately what there is to explain. But this sort of objection must be stated carefully. For example, we certainly cannot
complain that the theory fails to explain why some critical significance claims and not others, are justified. This would be parallel to complaining that a theory which shows that there are no witches and explains the apparent witchy features of some persons in terms of simple natural phenomena does not explain how some people come to be witches and some do not. The main problem with \((E_1)\) is that it offers an explanation of how works of epic literature produce attitudes and feelings related to the actual world which is theoretically unsatisfying, since it merely gives a set of causal descriptions and gives no general characterization of the function of such language. As a consequence, it provides no adequate criterion for deciding which implicit reports and reflections are appropriately associated with a given work.

The apparent cognitive value of a work of epic literature is purportedly explained by the Expression Theory in terms of attitudes engendered by the experience of the work. But how these attitudes and feelings, particularly those which apply to the actual world, are engendered is not very clear. The proponent of the theory may of course respond as follows: We can give an explanation of the function of epic literature in generating attitudes and emotions by giving a complete description of the causal chain in terms of the stimuli and responses in each case, then generalize on a number of such accounts; we can make the account as detailed as we like and as our present day methods permit. There is no doubt that such a theory furnishes a kind of explanation of the function of epic literature. But it is somehow unsatisfying. To see the respect in which this is so, let us contrast it with the sort of explanation we might give of the function of reports which we take to
express another's beliefs about the world in generating beliefs on our part about the way the world is.

Let us say that by our own experience we have amassed a bank of world-information which we can, if we choose, make explicit in a set of assertions about the world. The encyclopedia thus compiled is the source of premises which we use in making and inferring knowledge claims about the world. The reports of others, we might say, replace our own personal observations for us in gathering information. We, in effect, augment our own encyclopedias with the content of the encyclopedias of others then proceed to draw our conclusions about the world from the composite encyclopedia. On this provisional account, we can explain why some reports and not others, or why reports and not exclamations, generate beliefs. And we can characterize the future members of the group of reports which will generate beliefs; we can do so recursively, taking our own encyclopedias as the base, and designating a set of reports which add to our store of information about the world. Of course, we can also add onto the theory a complete causal account of how the generation of beliefs occurs when a given kind of report is apprehended.

\( \psi \), however, does not explain in terms of a general characterization of the function of the language why it should function in generating beliefs and attitudes about the actual world. Clearly it does not add to our information directly as a case history presumed to be true does. If the language of epic literature were clearly emotive in the narrow sense, then we might claim to have at least as good a theory about the function of such language as we have for exclamations.
and imperatives, etc. But it has been conceded that the language of epic literature is not clearly emotive in this sense. Lacking such an explanation of the function of the language of epic literature, we cannot characterize the future members of the group in any general way. On this kind of theory about such language, about the best we can do in the way of prediction is to invoke a "nearly-like-cause, nearly-like-effect" rule, based on the instances we have in fact observed. Given the enormous variety of new literary works, this, I think explains our dissatisfaction with the kind of explanation provided by (E₁). It may not be wrong in principle, but it is surely inadequate.

This leads to a related, and more serious difficulty with the theory, this time with (E₂). The causal theory proposed does not designate some responses and not others, as appropriate relative to the world of the work and to the actual world. We should hardly expect it to, since it is a causal theory, although not a very good one. To see the force of the problem, we need merely pose the case of the reader who hears the words of the work of epic literature in a language which he does not understand, then through some quirk comes to entertain a belief about the actual world as a result of hearing it. If this case is disallowed as farfetched, there are also less extreme cases, those in which the listener comes to entertain what we might call "random" or "arbitrary" beliefs as a result of experiencing the work, believing that the sea is salty, for example. The theory sketched above can, of course, "explain" the generating of such attitudes. And it can tell us that they are infrequent. But it is fairly clear that even in these extreme cases, it cannot explain our discrimination between appropriate and inappropriate
attitudes and emotions relative to a work. Some additional criterion of appropriateness must be proposed by the Expression Theory. If we had only \((E_1)\) to deal with, it would be left open how we decided what interpretive claims were justified. We could secure a good deal of objectivity by treating the language of epic literature as presenting a world, then leaning on our functional description of how second-person reports contribute to our knowledge of a world, set up a way of constructing a world description of the fictional world. "Belief" in such propositions, we could, allowing Richards his technical term, call "emotive beliefs." By this we should mean merely that they are not directed at the actual world, i.e., that it is inappropriate to jump onto the stage to save the leading lady from the villain. We could then go on to outline a complex procedure for the transfer of such claims to the actual world, perhaps on the basis of certain describable similarities between the world of the work and the actual world. Again, deferring to Richards, we could allow that acceptance of such a proposition on the basis of reading a fictional work is an "attitude," i.e., a belief not per se justified by evidence. We would, proceeding in this way, have a procedure for deciding what interpretive claims about the world of the work and what "attitudes" toward the actual world are appropriate to a given work, and which of them are not. \((E_1)\), which is merely an under-developed theory about the kind of language we have to deal with, i.e., says that it is not appropriately assigned truth values, leaves open such alternatives. They will of course be limited by the precise way in which we fill out \((E_1)\). But \((E_2)\) closes off the possibility of securing objectivity in this way. If \((E_2)\) is correct, then interpretive
discourse about a work of epic literature is not descriptive of anything, but is itself expressive in every case. An instance of \((E_2)\) is a quasi-imperative.

Accordingly, given \((E_2)\), there seem to be only two ways to differentiate between appropriate and inappropriate interpretations. The Expression Theorist may say that we must consider that a claim \(p\) associated with a given work is appropriately so associated insofar as the interpretive claim is persuasive. An appropriate response to a work is just a response which we can convince others is appropriate, or one which we can support with considerations which provide persuasive support for the interpretive claim.\(^54\) This may in fact describe the way in which differing critics seem to resolve their differences of opinion over the correct interpretation of a work. But as a general theory of interpretation, it has obvious defects. First, presumably, we need some specification of the group which counts as the group of people to be convinced. One proposed solution is that we should consider the persuasive power of the interpretation for an ideal body of critical observers over an ideal time span. But if we merely mean the body of critical observers over a long time, we assume an undue optimism about the progress of criticism and tend to forget that the ability of critics to judge and interpret a given artwork may very well decrease as the work ceases to be a product of their own age.\(^55\) This suggests that a diagram of the dialectic of critical differences, which is what Stevenson has given, may not actually be the logic of interpretation. If the group to be persuaded is determinately specified, the question arises of how some particular body of critics, say the body of English critics, come to
serve as the arbiters. But if it is alleged that the suggestion that the question be referred to an ideal body of observers is not a proposal to consult any particular body of observers under specified conditions, then, it seems, the suggestion is in danger of begging the question. The body of ideal observers may turn out to be just the people with the right responses. In addition, if \( E_2 \) is taken seriously, it rules out the possibility that certain internal features of the work can serve as reasons or evidence for our interpretive judgments that at least some implicit reports and reflections are objectively associated with the work and some are not. This is most serious in the case of interpretive claims which are about the world of the work. It is difficult to view the implicit report that the stranger winced because he saw Hester or the implicit reflection that unexpiated guilt leads to perdition merely as suggested ways of seeing the work which cannot be justified, but which are very persuasive. Nor it is very useful at this point, I think, to say that a critical significance claim about an implicit report is "persuasive" just in case we (or someone) likes the work much better when seen in this way, i.e., when we incorporate a given implicit report. What counts as a legitimate implicit report on this view will be very idiosyncratic. Someone might very well claim that he chose to see no causal connection between the events because to assume such a connection would create an air of suspense about the scene, which he for one would find unendurable—or because he happens to hold on independent metaphysical grounds that there are no causal connections. This is the kind of idiosyncrasy which is not very acceptable commonsensically. There may well be times when the reader
can choose between competing interpretations on the grounds that one or the other of them makes the work more pleasing to him. But this does not seem to be the only such criterion, and it is particularly unintuitive in the above example. This kind of idiosyncrasy is certainly not something which Richards would sanction in practice; he is unsparing of his criticisms of "mnemonic irrelevancies." And it is also something which Stevenson would like to avoid, although he seems to be aware that he cannot avoid it. And even though the suggestion seems to fail most obviously for implicit reports and reflections about the world of the work, we have no reason to be optimistic about the effect of works in causing appropriate emotions and attitudes with regard to the actual world either.

The Expression Theorist may try to secure some measure of objectivity in the preferring of some interpretive claims to others by making the following move. He may say that the entertaining of an attitude or emotion on the part of the reader is justified just in case the work is the expression of a relevantly similar attitude or emotion on the part of the author. If we take this suggestion literally, those cases where the relevant information is permanently lost are de facto undecidable. This is prima facie implausible. In addition, we can conceive of a case of which an aberrant author writes a work which we would call radically inexpressive of his known attitudes and emotions—are his emotions and attitudes still to be the criterion for a correct response to the work?

At this point, the Expression Theorist may make a radical change in his original proposal. The work once created, he may say, stands on
its own feet. This is not surprising. In a parallel case, we are able to delineate to some extent a range of appropriate responses to "Alas!" and "Damn!", even when we lack information about the context of utterance. We ought in the case of epic literature to be able to some extent to do the same sort of thing, to delineate a range of attitudes and emotions which are appropriate to the work, even when we lack information about its creator and the context of its creation. The question is how, given \((E_2)\), we are to do this.

The obvious solution, it seems, is to give up \((E_2)\) as a general thesis about interpretive discourse. We can retain its shred of truth by allowing that interpretive discourse often has as one of its aims making us see aspects and interpretations of a work which we did not see before. We can then concentrate on \((E_1)\). The strong points of \((E_1)\) are:

(i) it admits that we have to do in epic literature with a kind of linguistic performance which is radically different from ordinary fact-stating discourse; and (ii) it claims that works of literature are not direct evidence, even of the reportive kind, for knowledge claims about the actual world. What we need is an account of the language used in epic literature and a way of delineating in the case of a given work the range of responses which are appropriate, both for justified descriptions or interpretations which are relevant to the world of the work and for those which are assumed to extend to the actual world. The causal account proposed by the Expression Theorist can then be worked out and tacked on as a psychological explanation of how this kind of language in fact fulfills its function.
I shall use 'proposition' throughout to mean whatever can properly be said to be true or false. I shall assume that there are certain conditions for something's being propositional in character, roughly speaking that it have elements in syntactic arrangement such that it has meaning, and such that we can give a general semantic procedure for assessing the value of objects of this kind and can at least give a procedure for determining the truth value for the object in question. In most cases, I think propositions are merely sentences (or statements for those who like this distinction). It will become obvious, however, that we must leave open the possibility that there are arrangements of elements which are linguistic in a broader sense which might be allowed to qualify as propositions. What I have said about my use of 'proposition' also applies to my use of 'truth (or falsehood)'.

Morris Weitz makes explicit this link between the possibility of learning of truths from literature, which I have called its "cognitive value," and the claim that works of literature tell us something, i.e., are propositional. "The only thing I am saying is that some literary works of art do try to tell the truth, i.e., convey knowledge." "Does Art Tell the Truth?" Philosophy and Phenomenological Research, III (1943), 342. Weitz is, of course, looking at only one side of the coin. If works of literature can "tell the truth" and thereby "convey knowledge," they can also "lie" and "convey false beliefs." Arguments based on works of literature would thus be subject to all the problems involved in the argument from authority.

It may seem that this definition of propositional content is circular, that it in fact specifies the propositional content of a work by means of the propositions of the work. But this is at least not viciously circular, since, so long as it is the explicit elements of the work which are in question, we can specify the propositions of a work as its declarative sentences. By adopting an appropriate semantics, we can then specify which of the sentences are true and which are not. Of course, some decision must be reached as to the semantic status of sentences which contain ostensible references to fictional entities. In the present section, and the next, I shall treat them as false. The modifications of the presentation for a non-truth valued account are easily supplied, and it makes no difference to the argument which view is espoused.

I am aware of the weight that this way of specifying explicit reports will not carry, due to the difficulty of specifying in advance what can and what cannot be used to state a fact. Nonetheless, there are some rather clear instances, and I will be dealing primarily with these.

Notes


3 Morris Weitz makes explicit this link between the possibility of learning of truths from literature, which I have called its "cognitive value," and the claim that works of literature tell us something, i.e., are propositional. "The only thing I am saying is that some literary works of art do try to tell the truth, i.e., convey knowledge." "Does Art Tell the Truth?" Philosophy and Phenomenological Research, III (1943), 342. Weitz is, of course, looking at only one side of the coin. If works of literature can "tell the truth" and thereby "convey knowledge," they can also "lie" and "convey false beliefs." Arguments based on works of literature would thus be subject to all the problems involved in the argument from authority.

5 I am aware of the weight that this way of specifying explicit reports will not carry, due to the difficulty of specifying in advance what can and what cannot be used to state a fact. Nonetheless, there are some rather clear instances, and I will be dealing primarily with these.

6 Hawthorne, The Scarlet Letter, p. 82.


Here I follow Beardsley's terminology, *Aesthetics*, p. 409. Clearly, there are borderline cases, but not much hangs upon the distinction being a sharp one. It is merely a convenient way to divide the sentences we are interested in. It should be noted in passing that the division into explicit reports and reflections by no means exhaustively divides the expressions of the novel. Exclamations such as "Alas!" and injunctions such as "Consider our hero's distress!" fall in neither group.

Obviously, if one holds that sentences containing ostensible references to fictional entities are non-truth valued, one will be dealing in (P1) with a subclass of the whole class of reports and reflections, those which do not contain such ostensible references.

This, again, is Beardsley's term, *Aesthetics*, pp. 409-10. I use this terminology because it seems apt, although Beardsley himself is none too clear with regard to what the implicit reflections are to be considered to be about, the characters and events of the work or the actual world.

Most informal attempts to explicate the sense of "implied" in (P2) suffer from a central defect: they provide at best a rule of thumb for arriving at the alleged implied truths (or falsehoods), while leaving the nature of the implication murky. One such attempt is, for example, the view that the alleged implied truths are implied by crucial speeches of the characters. How they are implied is never clear.

Morris Weitz, "Does Art Tell the Truth?", pp. 346-7, once suggested that the failure to see that works of literature could assert general truths despite the character of the explicit reports could be traced to a failure to appreciate the paradoxes of material implication. The consequences of such a view were brought to his attention in R. Hookstra's reply to his article, *Philosophy and Phenomenological Research*, V (1945), 377.


The importance of the notion of a symbol is clear in M. Weitz, "Does Art Tell the Truth?", pp. 342-4.
17 I have included clauses (i) and (ii) because (P1) and (P2) (with "implied" read as "entailed" or "inductively implied") were shown to be too narrow, not too broad. Neither Weitz nor Beardsley ever denies that he holds (P3 i) or (P3 ii) as a part of the notion of propositional content. Both, so far as I can see, admit that the reports and reflections may be true or false, although both hold that they are not advanced as truth claims and that their truth value is incidental to the important truths contained in a work.

18 It should be noted at the outset that between these two views there is a difference of opinion as to whether the symbolically implicit content is asserted in the work or not. Weitz holds that the propositional content we have described as the implicit reflections of a work is advanced as a truth claim. Beardsley, however, argues that even such content is not asserted, that the work of literature contains only "non-asserted predications," which may be true or false, but which are not advanced as truth claims.

19 Morris Weitz, Philosophy of the Arts, ch. VII.


22 Beardsley, Aesthetics, p. 415.

23 Ibid.

24 The awareness of the problem of implicit reports in Beardsley's work is probably traceable to the fact that he is more concerned with vindicating what we have called "critical significance claims" and other interpretive claims than with the question of truth in literature.


26 Ibid., ch. III.


28 Beardsley, Aesthetics, pp. 405-7, 414-5, 432-35.

29 Ibid., p. 415. A similar view is advanced by R. Gale in "The Fictive Use of Language", Philosophy, XLVI (1971), pp. 324-339. Gale is clearly after a sense in which the author may be said by his explicit performance to "imply" general truths about the actual world. Insofar as Gale means that the author tells the truth by implication, his proposal falls prey to the objections connected with (P2). But insofar as he relies, as Beardsley does, on an unexplicated notion of "suggestion,"
loosely specified in terms of the emotive and connotative features of ordinary language, his proposal has the same problems as Beardsley's: The notion of "suggest" simply does not explain the relation of complex truths to works of epic literature.

30Beardsley, Aesthetics, p. 402.

31Robert E. Heilman, This Great Stage: Image and Structure in 'King Lear' (Louisiana: 1948).

32This sort of ploy harks back to Gilbert Ryle in "Imaginary Entities," PAS Suppl. XII (1933), although Ryle's reasons for adopting it are somewhat different from Beardsley's.

33An example of such a move in a slightly broader context is that made by K. B. Price in "Is There Artistic Truth?", Journal of Philosophy, XLVI (1949), pp. 285-291. He argues first that it is not in general true that works of art are themselves propositionally true. The basis of his argument is that many works of art lack a suitable combination of a semantic element ("definitional rules") and a syntactic element ("order rules") requisite for meaningfulness as propositions. He then argues against Hospers, if I understand the point of his argument, that Hospers' notion of "truth to" cannot explain the purported insight into reality afforded by the art object, because it is not, as Hospers seems to think, the object as such, nor yet its resemblance to reality, but rather our judgment that the work resembles reality, which is responsible for the cognitive value of art. He then concludes that the alternative to the view that a work of art has cognitive value because it is true in some sense is that "The contemplation of certain works of art evokes certain emotions... The important connection between works of art and other things brought about in this way is not meaning in the sense in which truth presupposes the latter. It is rather a causal relation between the observation of works of art and the assertion of predicative propositions."

34Such language has usually been called "emotive language."

I. A. Richards, Practical Criticism (New York: 1928), p. 353, also "Emotive Meaning Again," in Speculative Instruments (New York: 1954). Beardsley, in Aesthetics, calls such theories "emotive theories." Even Richards admits that the term was not well chosen. The reason for this, I think, is that 'emotive' is rather narrower in connotation than the users of the term intended. The use of 'emotive' leads critics of the view to the conclusion that what is being proposed is the assimilation of the language of poetry to the class of expressions such as 'Alas!' and 'Damn!', expressions which may be said to express occurring states of feeling. Beardsley does not make this error, but he does fail to notice the similarity of his theory of "implicit" or "secondary meanings" (p. 123) with the theories actually advanced by emotive and quasi-emotive theorists. My use of 'expressive language' is not unprecedented; and it avoids most of these problems. It is the term adopted by Wheelwright, The Burning Fountain (Bloomington: 1964), pp. 50-1, to present his quasi-emotive theory.
Obviously what I am calling the Expression Theory is specifically a theory about the language of epic literature. No doubt the name could be and has often been attached to a general theory of art, but I am here considering a narrower theory. This theory differs from, say, the expressionist theory of Tolstoy in two ways: (i) It is a theory of language in literature of a particular kind, not a general theory of art; and (ii) It does not hold a work is a good work if it adequately expresses a laudable emotion of the author; it is not a theory of evaluation at all. I am now aware that in using the technical term "Expression Theory" for the theory held by Richards and Stevenson, I have preempted a term which has a long history, even in literary criticism. I mean the term to apply only to Richards and Stevenson and to those who follow them. Much of what follows in the text does not apply to, say, Collingwood, who would deny the second half of \((E_1)\), p.44 and who would not accept \((E_2)\).

A. Richards, Principles of Literary Criticism (New York: 1928), p. 112. "These imaginal and incipient activities or tendencies to action I call attitudes."

Ibid., 177. Richards characterizes a scientific (or genuine) belief as "readiness to act as though the reference symbolized by a proposition is true." Unless there is some unexplicated difference between a tendency to act and a readiness to act, then a belief is a special kind of attitude. Richards then distinguishes between an "emotive belief," a pseudo-belief held for the sake of aesthetic experience, and a "strong emotional attitude" left as the residue from emotive beliefs. What he has in mind is probably the distinction between our "emotively" entertaining the pseudo-belief that Hamlet was a Dane, and the "emotional attitude" left over from seeing the play (which perhaps feels like a belief about the actual world) that those embroiled most closely with one another destroy each other and their common world of interaction.

Elsewhere Richards lumps all reactions to emotive language together under 'attitude.' Ogden and Richards, Meaning of Meaning (New York: 1959). Here it is clear that "belief" and "attitude" are straightforwardly opposed, presumably on the basis of their origin. I shall simply ignore this earlier terminology. Either the later section is badly muddled, or there has been a shift.

Richards comes very close to saying just this.

The restriction that such attitudes need not be per se substantiated by evidence is included for purposes of generality. A reader could have "reactivated" a belief he already held on adequate grounds; the work would nonetheless not have generated a justified belief.

I would be inclined as I have indicated to hold against Richards that there are evaluative beliefs and that there are genuine beliefs about the world of the work, but this depends mainly on the semantics
one adopts, whether evaluative claims and claims about the world are treated as referenceless. Richards thinks they are without reference. I think they are referential.

41 It need not, for example, describe an emotional state, as the following passage does:

There's a wailing of wind in the chimney nook,
And I vow that my love lies dead.

which occurs as an example in Richards' Practical Criticism. Nor need it have the sonorous arrangement of sounds sometimes held responsible for the effectiveness of a poem, such as:

My love is slain, I saw him goe
O'er the white alps alone.

Quoted as an example in Beardsley's Aesthetics.

42 It will later emerge (cf. chapter V, section 5.1) that I agree to some extent with (E1). My main complaint against (E1) as advanced by Richards, is that though it correctly assesses the logical link between reading a work of literature and entertaining beliefs about the actual world, i.e., claims that there is none, it somewhat misleadingly attributes the entertaining of such beliefs to emotions on the part of the reader, and thus fails to give an adequate causal account of the process.

43 Stevenson, "Interpretation and Evaluation," Philosophical Analysis, ed. Black, 334. And Richards clearly thinks that this is what he is doing throughout Practical Criticism, which he describes repeatedly as "scientific."


45 Stevenson makes this point clearly in "Interpretation and Evaluation," pp. 330-340. He expresses it by saying that the critic's judgment is based upon a judgment which is non-scientific. On Stevenson's analysis, the interpretive judgment is to be interpreted as a quasi-imperative. He makes the same claim in "The Analysis of a Work of Art," Philosophical Review, LXVIII (1958), 83.

46 Richards clearly holds throughout Practical Criticism that there are deviant and irrelevant interpretations of a poem, as well as correct ones. Stevenson is more cautious. He holds merely that there is "reasonable" as well as "unreasonable" support for an interpretation, just as there is "reasonable" support for an imperative. This he attributes to the pragmatics of the context. He claims that there is never, strictly speaking, evidence for interpretive claims, whether they are claims about the work itself or claims about its purported meaning with regard to the actual world.
I have argued earlier that we can dismiss Stevenson's and Richards' quibbles about whether "beliefs in the strict sense" can follow from works of literature. They have at the very least failed to note the possibility that works of literature might cause readers to have random beliefs, say that the sea is salty. In any event, theorists such as Price (fn. 1) do talk about propositional beliefs resulting from works of literature.

It ought to be noted that the expression "justified" is a strange one for such theorists (although Richards may in practice mean it in quite a conventional sense). Stevenson would no doubt accede to this vague formulation. There is no question here of reducing "justification" to matters of fact because 'appropriate' in the definiens also has emotive force.


It is interesting that (E₁) has traditionally been proposed as a thesis about lyric poetry. In fact, if (E₁) is taken to entail the view that lyric poetry is non-propositional it is pretty clearly false in many cases. Obviously language can be both expressive and fact-stating. And in cases such as Pope's work, it seems clear that we merely have essays in verse.

This is not all that Richards means by emotive belief, but we can argue that it is what he should mean, that he has seized upon the phenomenon more widely known as "willing suspension of disbelief" and tagged it "emotive belief." In such a case we in a suspended sense "believe" that Dante is going to get to heaven eventually, even though if someone were to ask us about our theological beliefs, we might claim that there was no such place. And we in some sense "believe" in the prison scene of The Scarlet Letter that Roger Chillingworth is going to cause a lot of trouble, although we know that he does not really exist, and a fortiori that he is not going to cause trouble for anyone in the actual world. If Richards wants to call this "emotive belief," then he can do this. It is his term, after all.

I at least accept the view that they do not function in any standard evidential way, although I will argue in chapter V that they do have a specifiable relation to our knowledge of the actual world.

Richards argues that a good deal of criticism just is metaphorical (or expressive). This may very well be true; he produces some horrendous examples. But it is certainly not obviously universally true. And this is the only argument that he ever gives for the view, so far as I can see.

Similarly, it may be true that some criticism does, as Stevenson alleges, shape and guide our aesthetic perceptions. But this fails to establish (E₂), since (i) it is not clearly true for all of criticism; and (ii) there is no incompatibility involved in saying that criticism both guides our perceptions and objectively describes the work. This double function of language is a daily occurrence.
Stevenson's strongest argument is that every interpretive claim involves a decision about how to see the work. Even if this were true, it would not establish \((E_2)\). We may say, for example, that whether we describe Wittgenstein's notorious figure as a duck or as a rabbit depends on a decision to see it in a given way. Nonetheless, the resulting judgments 'It's a duck' or 'It's a rabbit' are clearly not quasi-imperatives.

We might add as the first recursive clause, for example, (1)
\[ p \text{ is a proposition such that } \neg p \text{ is asserted by a reliable witness } W \text{ is a member of my encyclopedia.} \]

I think this is, in fact, Stevenson's view of the matter, "Interpretation and Evaluation," pp. 344-350.

Here it seems that Stevenson himself is a bit overly optimistic in his view that people just will rule out certain interpretations, since no one "in the light of knowledge is content to decide in their favor," and his prediction that people's decisions "will not be capricious." *Ibid.*, p. 343.

Richards at some point seems to take this view. But it is more likely that his view is that the correspondence of the emotion evoked with the emotion expressed is an ideal toward which poetry and criticism should strive, not a criterion useful in sorting out good and bad interpretations. Stevenson rejects this solution altogether because he finds it in conflict with \((E_2)\). It is not clear that it actually is incompatible with \((E_2)\). Presumably we could hold that an interpretation is an expressive utterance while still, for whatever reason, requiring that it attribute to the work the emotion which matches that of the author.

It is possible that Richards sees no difficulty whatsoever with regard to the relation of his theory and his practice. He may in fact hold a much weaker version of \((E_2)\), on the basis of which there is interpretive criticism which is not non-scientific in the strong sense, although it is expressive in its aim, i.e., seeks to guide our aesthetic perceptions.

Stevenson raises precisely these objections to this suggestion.
Chapter II

In this chapter, I want to look more closely at the language of fiction. There are three principal sources of information about fictional language. First there are works of fiction themselves, the phenomena under discussion. Secondly, there is literary criticism. Here we often find a wealth of material about individual works and their genesis. But the methodological assumptions involved are often taken for granted and are frequently half hidden. The form of the arguments is often similarly unclear, and when clear, undefended. That part of literary criticism which describes and interprets works and bodies of works thus seems to have the status of an informal theory. Third, there are the data from the philosophy of language, from linguistics, and from special critical investigations of the language of fiction. The last group of investigations listed furnish a good deal of information about the phenomenon under discussion, but such work is often very narrow in scope; and in general it lacks a clear understanding of the kind of analysis needed to explain the speech act of writing fiction. All too often, for example, literary theory has veered into armchair psychological analyses of creativity and receptivity, instead of exploring the language use to be considered qua linguistic activity. And all too often, the crucial technical terms are left unexplicated, and the structure and function of language in non-fictional contexts is taken as a relatively unproblematic guide, whereas it is the case that crucial areas of non-
fictional language are not yet clarified. The philosophy of language approach tends to remedy these defects, but in general philosophy of language and linguistics have lacked the broad empirical base necessary for dealing with the language of fiction as a particular kind of speech act and for demarcating the relation between the language of normal narration and the language of fiction. This suggests that all three sources of information might be combined to complement each other. In this chapter, I propose to try this combined approach.

My main contention will be that writing fiction is a kind of speech act which is strongly dependent upon the linguistic activity of describing and recounting events of the actual world, although it is distinct from this latter practice. The novelist does indeed set up a world, people it with characters and connect them by events, comment on this world, etc. In this sense, he is like the person who might be described as "the complete and concerned historian." Throughout this chapter, I will be comparing and contrasting the activity of such a historian and his language, which I will call "the language of normal narration," with the activity of the author of a work of epic literature and his use of language, "the fictive use of language." Our historian may in the course of his narrative ask questions, exhort, etc. And in this chapter I will take 'normal narration' to include all of these activities.

Fiction may be analyzed as a kind of speech act. What do we mean by saying that something is a speech act? We can take as relatively unproblematic the claim that linguistic performances are one of the things human beings typically do. They are of a kind of action. In a weak or broad sense, perhaps any action done using language may be called a speech
act. I will use "speech act" in a narrower sense. A speech act in this narrower sense is an instance of a kind of linguistic performance which is rule-governed in the sense that there are guidelines, i.e., rules or conventions, laid down for performing an act of the kind in question using language, and for apprehending a given linguistic performance as being of that kind. These conventions specify, at least broadly, what counts as performing a speech act of the kind in question. Their employment enables the audience to decide (in the absence of evidence to the contrary) that an act of the specified kind is being performed. The existence of such conventions and the performer's internalization of them and intention to employ or abide by them is a condition of his being able to perform the kind of act they specify. And his being able to intend that his audience will construe him as employing them or abiding by them is also a condition of his being able to perform the act. This in turn presupposes that he can presume that his audience has internalized the rules or conventions of fiction to the extent of being able to construe him as employing them or abiding by them. And, once the audience has so internalized the conventions, there is a clear sense in which the audience can treat a speech act as if it were of a given kind, as if the author intended that he be recognized as employing certain conventions, despite his explicit intentions to the contrary or in the absence of information about his intentions. Usually the occasion for doing this is the fact that the conventions seem to be being employed, no matter what the author of the speech act says. This approach by no means commits us to a naive view on which language is taken to be some sort of game with definitive and meaning-conferring rules. We need only
note that certain ways of performing a linguistic action lead to its being and being understood as a performance of a given kind of action. It may well be that linguistics will eventually present us with a more satisfying account of these rules for performance of actions of a given kind. Talk of conventions and rules may be taken as a stopgap for that more detailed account; and what I shall call conventional features of fictional language may be taken as stopgap features of such language of which any complete theory must eventually take account.

A large number of kinds of linguistic activities seem to depend upon the ability of the speaker to specify or demarcate things or events and to characterize or classify them as being of a kind or in a state. I will simply call this element of such a speech act a "referring-characterizing component." To say that referring-characterizing is basic to certain other linguistic activities is simply to say that to be able to perform certain kinds of linguistic activity, we must first be able to pick out individuals (refer) and know how generally to say things about them (classify or characterize them, or predicate). Alternatively, we can say that many linguistic activities have a referring-characterizing component. The most obvious examples of this are acts of asserting or stating that a state of affairs obtains. It is also rather intuitive to view commands as, say, instructions to bring about a state of affairs. Analyzed in this way, commands may be seen to contain a referring-characterizing component. Similarly, if questioning is viewed as an inquiry about whether a state of affairs obtains, then, again, a question contains a referring-characterizing component. The situation here seems to be that the ability to perform certain simple linguistic acts which are commonly
called "illocutionary acts" is dependent upon our knowing how to refer
to something and characterize or classify it. And certain more complex
illocutionary acts, such as promising or warning, seem to be similarly
dependent upon referring-characterizing.

In addition, there are certain speech acts (commonly called "per-
locutionary acts") aimed at producing an effect in the hearer in addition
to his understanding of what is said. I may, for example, attempt to
frighten or amuse my hearer by what I say. I may frighten someone by
saying, "The pass will be watched tomorrow," or by exclaiming, "Don't
go by the pass tomorrow!" In the first case, the dependence of frightening
the hearer is on the act of assertion. In the second case, frightening
him is similarly dependent upon commanding or exhorting. In both cases,
if the above hypotheses about the basicness of referring-characterizing
is correct, the illocutionary act itself contains a specification of a
state of affairs, and is thus dependent upon referring-characterizing.
And in general, perlocutionary acts are dependent for their success upon
the hearer's understanding of the illocutionary act which is used to
produce the effect.

All these linguistic activities have something in common. They are
related by being in some sense about or directed toward the world. This
is the crucial point, whether or not the above slightly controversial
hypothesis about their structure proves ultimately satisfactory. All of
them may occur in the course of what I have called "normal narration."
Their common actual world directedness is particularly clear if they are
seen as containing a referring-characterizing element, for the reference
is clearly directed at some individual of the actual world. And, it can
be argued, our ability to classify or characterize is dependent upon our having learned to describe things by applying characterizing expressions to them and withhold them from others when certain conditions do or do not obtain in the actual world.\textsuperscript{10}

In the practice of writing fiction, we find kinds of expressions parallel to the full range of kinds of speech acts which may occur in normal narration.\textsuperscript{11} The expressions which occur in the language of fiction differ from those in normal narration, however, in that they are not directed at the actual world. This has given rise repeatedly to the question of how such language can be meaningful at all.\textsuperscript{12} In section 2.1, I will attempt to give an account of the language of fiction as a kind of language used in performing a particular kind of speech act. I will argue that we can account satisfactorily for the meaningfulness of the language of fiction in the following way. We can say that with regard to certain important parts of the language of fiction, the linguistic activity in question is parasitic upon the function of syntactically similar expressions of the language of normal narration, although in each case, the use to which the kind of expression is put in fiction is fundamentally different from the use of syntactically similar expressions in the language of normal narration. In section 2.2, I will list certain features of the commonest characteristics of the language of fiction which are such that their presence leads us to construe the performance as fiction in problem cases, where we have a lack of information about the context of the speech act or where we have reason to suppose that the speaker has misdescribed his own action, as sometimes happens.\textsuperscript{13} In section 2.3, I will take up the question of the treatment of historical
novels, which seem prima facie to constitute an objection to the treatment of fictional names I advance in section 2.1. In section 2.4, I will pose a central problem in the analysis of literature, that of dialogue, indirect quotation and related phenomena having to do with narrative point of view. In this case, I will attempt to show that the account we use for such expressions in normal narration suggests a treatment for expressions which have parallel syntax, but different function from these expressions in normal narration.
There has always been a question about the source of meaning of the language in fictional contexts. In this section, I will proceed as follows. I will first give a general characterization of the fictive use of language as a kind of speech act to show how problems of meaning arise with regard to such language. I will then proceed to a detailed analysis of the central features of the language of fiction to answer this question.

Works of epic literature are certainly linguistic performances in an obvious sense. They are actions which the author performs using language. He intends to write a work of fiction and to be understood as doing this; to achieve this result, he follows certain guidelines. If he is successful, he acquires an audience which does in fact understand the work as a work of fiction. All of this is just to say that writing a work of fiction is an instance of a kind of speech act in my narrow sense. That is to say, in the paradigm cases, writing fiction is a rule-governed kind of linguistic activity which is performed typically with a specific kind of intention and in accordance with certain conventions or rules. In the paradigm cases, the author intends that his work be understood as a work of fiction at least partially by having his audience note that he is abiding by the characteristic conventions of fiction writing.

There are more and less successful performances and apprehensions of works of epic literature. It is certainly not necessary that writer and readers be able to state clearly and explicitly what they are doing. But we as analysts of such activities, can certainly strive to give a characterization of the activities of writing and reading works of epic
fiction which explains with maximum clarity and precision the main features of the actions which most agents perform, perhaps unconsciously.

Heuristically speaking, we can characterize the kind of speech act in question by saying that the author of a work of fiction is constructing and presenting a fictional world; and we can presume that his audience understands him as doing this. This is a fairly common way of talking about fiction. He is not and does not understand himself to be describing the actual world in the sense in which the biographer or the historian does. He may, of course, expect that his work will clarify and expand his readers' knowledge of the actual world. Many works of fiction are in this sense "novels of purpose," in which the author regards himself as "telling the truth." But, as we say in chapter I, the sense in which the author of a work of epic literature "tells the truth" is by no means simple to explain. It is the point at which we intend ultimately to arrive in chapter V. Of course, it could be argued that certain myth-constructors of ancient times did regard themselves as historians in the strict sense or as offering causal hypotheses about the generation of the world. But to generalize on the basis of such cases to the case of present-day writers of fiction is to get things backwards. The proper procedure, it seems, is to analyse the activity of present-day writers of fiction, who at least self-consciously claim to be writing fiction, then to explain how we can construe other acts as being of this kind, even if they were not intended as such or if the speakers' intentions were ambivalent.

Works of fiction certainly are meaningful as fiction, i.e., as non-world-directed. One of the perennial questions about the language of
fiction is how it can be meaningful at all, given that it is not world-directed. And, given that such language is meaningful, there is a further question concerning what special conventions govern this use of language. These are the central questions to be answered in this section. In doing so, I will employ the notion of one use of language being parasitic upon another. To say that one use of language is parasitic upon a second, is merely to say that we would not understand the second if we did not understand the first. The fact that fictional language is to a large extent parasitic upon normal narration in this sense has often been noted. One obvious connection between the two kinds of language is the following. The predicates used in a fictional work retain to a large extent their intensions relative to the time the author produced the work. Commonsensically, I cannot ask a child who has not learned the meaning of color words to imagine a red balloon (leaving open what precisely counts as learning the meaning of color words). Similarly, if I were presented with a work of fiction in which all the English color words had been changed to the corresponding expressions of some language I did not know, I simply would not understand (completely) what the transformed sentences said. New predicates used in a fictional work must be defined in terms of other predicates whose intensions do remain constant (or as corresponding to or opposed to predicates of this kind). This is a clear sense in which our ability to write and understand language in fiction is dependent upon the ability to write and understand language in non-fictional contexts.

In this section I want to show that further features of the language of epic literature are parasitic upon normal narration. The function of
expressions in fiction, though significantly different from the function
of similar expressions in normal narration is parasitic upon this
latter function. I will attempt to do this by showing that the view
is correct for various important linguistic functions in normal narra­
tion. I will argue (i) that understanding an author as presenting a
character is parasitic upon our understanding acts of naming and re­
ferring in normal narration; (ii) that understanding an author as es­
tablishing a fact of a fictional world is parasitic upon our under­
standing true or false reports about the actual world; and (iii) that
understanding commands, imperatives, etc., within a work of fiction is
parasitic upon our understanding them in non-fictional contexts. In each
of (i) - (iii) the fictional use of language is significantly different
from the normal narration use to which it is parallel. In each case I
shall attempt to point out what the difference is and why it is signifi­
cant. (i) - (iii) constitute the main rule-like or conventional features
of the language of epic literature. To read a work as fiction is to
read it as if its syntactic units have the functions described in this
section. The individual claims I am defending are not uncontroversial.
But they are, I think, individually plausible. And insofar as each claim
is plausible, as a group they present fairly impressive evidence for the
general thesis.

It is precisely this parallel between fictional language and normal
narration which answers the traditional question: How is fictional
language meaningful? It turns out that although we can maintain that
expressions and constructions acquire their meaning only in relation to
the actual world, there are meaningful uses of such language so acquired,
notably the use in fiction, which are not meaningful by virtue of being directed at the actual world.

(i) In normal narration, proper names and pronouns are used to pick out and identify individuals of the actual world. Let us say that in such a case a speaker $S$ refers to or makes a reference to an individual $i$ by using a name $n$, which is conventionally attached to $i$. A hearer $H$ who knows the name $n$ and knows that $n$ is conventionally attached to individual $i$ understands $S$'s reference to $i$. When $H$ comes to know what individual $S$ is talking about on the basis of what $S$ says and the name system they have in common, we will say that $S$ successfully speaker-identifies $i$ to $H$, that $H$ hearer identifies $i$, and that successful speaker-hearer identification has taken place.\(^{15}\)

In fiction, syntactically similar expressions are used to present a character. We rely on the syntactic similarity of expressions in fictional and non-fictional contexts to decide which expressions in fiction are character-presenting.\(^{16}\) In normal narration, when $H$ hears what he takes to be a name, he assumes that it is being used to refer to an individual.\(^{17}\) Name-like expressions in fiction constitute an "occurrence" of the character each time they occur in the narrative. A reader of a novel presented with a name-like expression for the first time adds a character to the furniture of the fictional world, accepts a character. When the author presents again a character which the reader accepts, I will say that speaker-hearer accord has taken place. The process in fiction is analogous to that in normal narration, but in fiction the reader is not himself a part of the world in which the characters move. This is the crucial dissimilarity. In normal narration we gradually chart
a continuous part of the individual as related in space and time to ourselves. In fiction, on subsequent occurrences of the namelike expression, the reader notes the kind of character which has been added to the fictional world and its relationships to the other individuals of the world of the work, although the normal narration relation of particulars to himself is lacking.

Along similar lines, in normal narration, a speaker S may use a definite description d to identify an individual by means of a property which it alone has. S speaker identifies or refers to i by means of d if i is the individual S has in mind and i is in fact the only individual which has the property in question. Hearer-identification occurs when H identifies the individual i which S has in mind and picks out i by the use of d, on the basis of S's use of d and the background knowledge common to S and H. It is obvious that H can come to know what particular S has in mind on the basis of d, even if d, strictly speaking does not pick out a unique individual, and even if H knows this. In such cases, I will simply say that S successfully picks out i for H, reserving 'identifies' and 'speaker-hearer identification' for those cases where d does in fact apply to i uniquely.

In fiction, the author presents or re-presents characters as having certain unique properties by the use of expressions syntactically similar to definite descriptions. The author cannot fail to refer to or identify an individual by means of d, for, strictly speaking, d in such a context does not refer to something or identify it at all. When an expression is used in fiction to present, it has no referent. Can an author, however, fail to present a character by using d as a speaker in normal
narration can fail to refer to something? The answer, I think, is "yes." He may fail to present a character using 'the grandmother of Lucy', for example, if, according to the story, Lucy has two grandmothers. Of course, even if he fails to present a character in the strict sense, the context may save him, as in normal narration.

(ii) In normal narration, a good many declarative sentences are used to report or describe facts of the actual world. A sentence 'Fa' in an utterance context will be true of the actual world just in case the individual referred to or identified by 'a' does have the property specified by 'F'. In the context of fiction, declarative sentences are used, not to describe, but to establish the facts of the fictional world. Sentences used to establish fictional facts are not properly assigned truth values. First, the facts of the fictional world are what the author says they are (with the exception noted below). It is hard to see what force 'true' would have here, since they are established by fiat and cannot be wrong. Secondly, since the namelike question expressions in such sentences do not refer to anything, it is hard to see what the property in question would have to be a property of in order for the sentence to be true. Declarative sentences of fiction are simply not the kind of linguistic performance to which we assign the value 'true' or 'false'. It is as inappropriate to do so as to jump onstage at a play to rescue the heroine. The author's declarative sentences are not, strictly speaking, true or false reports, since they are not reports at all. An author may be unsuccessful in establishing a universe in a case where he establishes two facts which are incompatible in the context of the novel. At this point, since we have denied truth values to the
sentences of the novel, we will say that two such sentences are incompatible if the inclusion of them (and perhaps related sentences of the work) in a set of sentences would render that set inconsistent under an assignment of truth values on which all the novel sentences were counted as true.

Again, it is clear that the practice of establishing fictional universes is parasitic on the practice of describing the actual world. The parallel syntax of the two uses of language tells us which sentences of fiction are candidates for being fact-establishing. The constant intensions of predicates tell us what facts are being established. And the notion of consistency for sentences used in establishing a fictional universe is dependent upon the notion of consistency for sets of sentences which do have truth values. We would not know quite what to do if someone were to present us with an alleged description of the actual world which contained sentences which were incompatible. Similarly, we would not know quite what to do with a set of sentences which allegedly presented a fictional world, but which established contradictory facts.

(iii) In normal narration, we have a difference between making a statement, asking a question, and issuing an imperative. These differences we can ascribe to the differences in the illocutionary force of certain kinds of expressions. The speaker, whoever it is that does the ordering, the asking, the stating, is what I shall call the "illocutionary subject" of the expressions in question as they are used. In normal narration, given that questions, exclamations, etc., may occur in the course of telling a story, they demand an illocutionary subject. Clearly, the specification of the illocutionary subject of the history
or biography taken as a whole distributes over the individual expressions which occur in the story. The whole performance is, say, a biography, and the individual declarative sentences of the biography are statements or assertions of the biographer. In addition, the author may intend for the biography as a whole to horrify his audience, move them to pity, etc., although this may not be true of all the individual assertions. It may be, for example, that certain sentences of a biography of Caligula do not horrify the audience; but the biography as a whole will certainly do so. The biography, we may say, has perlocutionary force as a whole, and the biographer is the perlocutionary subject of the whole.

In fiction, we have no trouble, I think, in identifying the perlocutionary subject of the work as a whole. It is the author. Nor do we have any trouble identifying the illocutionary subject of the work, qua work of fiction. Again, it is the author. But when we ask about the illocutionary subject of the individual expressions within the work, we encounter a problem. Granted that the sentences of a work of fiction as sentences of such a work, are not reports, nonetheless we do need to mark the differences between expressions occurring in such a work which are used simply for fact-establishing and expressions which are, intuitively speaking, questions or exclamations. We were accustomed to do this in normal narration by saying that the different functions of such expressions are explained by the difference in their illocutionary force. And we can advance a parallel account of the difference in force of syntactically similar expressions which occur in the language of fiction. But who is the illocutionary subject of such expressions when
they occur in a work of fiction? In many cases, I think that common
sense tells us that their illocutionary subject is not the author;
we need only consider those works told in the first person by a narrator
who is not identical with the author. And even in works in which the
author is not so intrusive, e.g., Tolkien, it is not at all clear that
the author is the illocutionary subject of the simple illocutionary
acts. The author does not have to ask questions, and yet the narrator
may say such things as 'Who can say what passed through his mind?' It
is plausible to suppose, I think, that it is almost always the narrator
as distinct from the author who is the illocutionary subject of the force
of these expressions. In normal narration, the illocutionary force of
telling a story distributes over simple illocutionary acts which occur
in the story. In fiction, the illocutionary force of writing fiction,
which resides in the author, does not, it seems, distribute over the
expressions involved in establishing the universe. Insofar as the in-
dividual expressions of the work are treated as having different force,
we need an illocutionary subject or placeholder to mark that difference.
The account given in the previous section tells us how to read a work, given that it is fiction. But that account is naive in the sense that it assumes that the writing and reading of fiction is a "no secrets" affair, that we always know when a work is intended to be read as fiction and when this intention is fulfilled, when we should read it and understand it as fiction. This is certainly not always the case; in many instances, our knowledge of the context and the kind of speech act involved falls far short of the proposed ideal. In this section, I will propose some methods by which we can decide that a work is fiction, even if our knowledge of it is suboptimal, then, proceed to read it as fiction, i.e., as if it had the features described in the previous section. These are also features of language which would justify us in reading a work as fiction, even in the face of knowledge that it was not intended as such.

The work of fiction is then a linguistic performance. As such, it is performed at a given time or over a time span by a performer who has a specific intention. The performer intends to perform the complex illocutionary act of writing a work of fiction. There is no doubt that the author's intention in this simple sense, i.e., what he wants to do by his action, is important in how we judge any action, particularly a speech act. We should not be surprised that often our best indicator that a linguistic performance is a work of fiction is the fact that the author says it is, e.g., that he refers to his work as "my latest novel." If the author does not explicitly intend in this sense to write a work of fiction, then we would certainly balk at saying that he has not succeeded in doing so; and it would most likely be misleading to claim
that he has performed the action of writing a work of fiction, since one of the commonest conditions of saying that someone has performed an action of a given kind is that he intended to do so, i.e., would assent to this or some equivalent description of his action. In a case where the author has denied or would deny that he has written a work of fiction, we may nonetheless construe his performance as fiction. We might do this for various reasons, not all of which need be "internal" or author independent. Massive failure of reference, the fact that the author was a member of a group who typically wrote fiction, the fact that the work resembles a group of other works which are unproblematically fiction, could all count against the author's disclaims, for example. And of course, in the terminology of the first section, the author may intend that his illocutionary act of writing fiction in addition have perlocutionary force. He may intend that his fictional work amuse his audience. Or he may intend that the audience acquire certain true beliefs about the actual world as a result of reading his fictional work. How the work can be used to achieve this sort of perlocutionary effect is of course the question posed in chapter I. We will return to it in chapter V.

Convenient as an appeal to the author's stated intention in this simple sense is, it is not the whole story. There are anonymous works and works where we know only the author's name. Nonetheless in many such cases, we presume, despite the absence of information about the author's biography that we have a work of fiction before us. We are rather confident of our ability to do this; the question is how or why we do this. Secondly, it is not in general true of speech acts that
intention in the simple sense of wanting to perform a kind of act is sufficient to be said to be performing it. Ordinarily one cannot, for example, ask someone what time it is by shouting 'Stop!' The point here is not that one cannot in the simple sense intend to do whatever one likes with any piece of language whatsoever. The point is rather that language is a public affair; certain things count as performing an act of a given kind. There are recognized ways of performing such acts and of indicating that one is doing so. There is scant hope of being understood to be performing an act of a given kind for which there are previously established rules or conventions of some sort without obeying these rules or formally announcing a departure from them. Hence, one cannot intend in the stronger sense to perform the act in question unless one is abiding by the rules (or unless one is operating on a prearranged departure from them). We are looking in this section for certain conventional features of fictional speech acts which mark them as being of this kind and order our responses to them. Of course, ordinarily we presume that when such features are present the author intended to incorporate them in his performance, that he knew what he was doing. But once the features are recognized, we will be able to see how a work can be "read as fiction," even when we have reason to believe it was not meant as such.

The writing of fiction as we have it today is well established as a practice. By following established conventions or reproducing well-known features of previous performances of the kind, the present-day performer is able to indicate to the audience what he is doing without telling them explicitly. Though an innovator who does not formally announce his
departure from standard practice may stand to gain from his omission, he also runs the risk of being misunderstood. Throughout the history of fiction, innovators have tended to protect themselves from misunderstanding by announcing their departure from standard rules or procedure for performing the act of writing fiction. A good deal of informal theory about literature originates in this way.

The present-day writer of fiction has a choice of many such sets of conventions. Some of these are highly stylized and very familiar. Instances suggest themselves readily, e.g., the preliminary invocation to the muse of poetry and conventional openings and closings such as 'once upon a time' and 'they all lived happily ever after'. But much of present day fiction attempts to do without blatantly stylized devices. Given that it is plausible that the relation between normal narration and writing fiction is as I have indicated, we should expect that certain conventional features of the language of fiction will serve the purpose of presenting a fictional universe effectively and of differentiating this enterprise from normal narration. The writer of fiction has language as his medium. He does not have the stage presentation of the dramatist or the screen representation of the film producer at his disposal. We will expect that certain of the common features of fictional language will be obvious means of making the most of the medium at the author's disposal. The language in use is not directed at the world as is the language of normal narration. And so we should expect that certain of the common features of fiction, will be those on the basis of which the reader, even in the absence of the needed information about the context of utterance, will be able to tell that he has fiction before him. The
author of a work of fiction does not in these times often have the explicit context of a campfire around which people traditionally sit telling fantastic tales at his disposal. And so he must rely on certain features of his language to provide him with the understanding and co-operative audience which full context might otherwise secure. I want to turn now to some of the more common features of such uses of language by virtue of which we know that a given performance is fiction. They are especially helpful in problem cases where we lack information about the context of the speech act or where we have reason to believe that the author has misdescribed his own action. These may not be features which we commonly note, although for the experienced reader of fiction they might well become more or less automatic.

(1) One common internal clue that the work which is before us is fiction is the recurrent use of poetic devices. Even a predominance of simile and metaphor, or a large amount of highly descriptive language, is evidence that we are dealing with a fictional world. It is not conclusive evidence, for such devices have uses in normal narration. No doubt stylistic usage in fiction and in normal narration fluctuates. Newspaper reports of the 1800’s are replete with poetic usage, and some naturalistic novels lack poetic devices almost altogether. What does not fluctuate is the following special feature of such language which makes it an ideal tool for the author of a work of fiction (and by the same token, a suboptimal tool for the historian or biographer, if accuracy is his object). Poetic language presents a situation in maximum detail and vividness. It does not, of course, in any sense, create in us the illusion that we are actually experiencing a sensibly presented
world scene. Even the most imaginative reader, if he is normal, presumably knows very well the difference between the experience of looking out a window and seeing a scene or experiencing a situation and reading a written account of it. But the author does want to create a vivid aesthetic surface in his work. And poetic language often functions to vivify the presented universe through intensifying the experience of the reader and suggesting images to his imagination. Due to the highly suggestive character of poetic language, its use results in an account, which if intended as a report about the actual world, would be highly precise. If such accounts were really descriptions of the actual world, we would in many cases know the facts of a situation in enormous detail, down to the colors of the queen's robe, the intangible quality of the atmosphere in the hall, etc. In addition, such language is precise in a way best shown by an example. To be told that something was red is not to be told very much. To be told that it was scarlet is to be told a good deal more: the range of application of 'scarlet' is much narrower than that of 'red'. To be told that it was scarlet with a glow of muted incandescence like the play of lamps on old velvet is to be told a great deal—more than we know of most historical situations. Just how such language supplies the additional information, I am not sure. Common sense suggests that it does so partly by appeal to memory and association. But such descriptions do furnish enormous detail, however they do it. When this much evocative detail is furnished, we begin to suspect that we are dealing with fiction, not with a factual report. Other indications of fiction which fall in the same group for similar reasons are descriptions of states of mind or intentions presented as fact, not as con-
jecture from external facts about overt behavior, etc. Because such events are in general not directly accessible as matters of fact, we suspect that a universe is being constructed, not reported on, when they occur in the account as statements of fact. They may be seen as degenerate cases of such straightforwardly poetic devices as soliloquy and interior monologue.

An even stronger indication that a work is fiction is recurrent use of symbolic imagery. When a given element which is clearly symbolic turns up so often that its presence in a description of the actual world would be a matter of extraordinary coincidence, the chances are very good that we are dealing, not with a description of the actual world, but with a manipulated universe. The scarlet letter on Hester's dress is acceptable—but by the time the scarlet letter turns up on Arthur Dimmesdale's breast and Pearl has been likened to a dancing image of the scarlet letter, and the symbol occurs flaming in the sky, we pretty well know that we are dealing with fiction, not a report on the facts of the actual world.

It is fairly obvious that such devices are useful in fiction because the speech act involved is one of presenting a universe. On one level, the use of poetic language is functional in that it allows the universe to be presented completely and vividly. And when we find a high preponderance of such language, we begin to suspect that it is a work of fiction which we have before us. We begin to read the work as if the vividness of the presented scenes and their interconnections are the object of the author's activity and not strict accuracy with respect to the actual world. And when we read a work as if accuracy with respect
to the actual world did not matter, concentrating on its vividness
and drawing conclusions from the interconnections of the events and
objects without regard to whether they ever really happened or existed,
we are reading a work as fiction. Thus, a preponderance of poetic
language devices does serve as a marker for the fictive use of lan-
guage. We now turn to slightly more important indicative features of
such language.

(ii) One of the most important features of fiction is significant
failure of reference. The simplest form of failure of reference involves
widespread failure of ostensible names and descriptions to refer to indi-
dividuals of the actual world. If the failure of reference is not com-
plete, one consequence of failure of reference in some cases will be
that numerous reports involving terms which do seem to refer to indivi-
duals of the actual world will be untrue. For example, 'David Copper-
field' as it occurs in the novel does not refer to an individual of the
actual world. Let us allow for the moment that 'London' as it occurs
in the novel does refer to the actual city. If 'David Copperfield was
in London' occurs on the pages of the story it will in any event be un-
true, if construed as a report about the actual world which relates David
Copperfield with London (taken as a particular). One of the ways of
seeing whether a work is a work of fiction is finding out whether the
majority of apparent names in it name individuals of the actual world.

There are more sophisticated versions of failure of reference. Some,
for example The Lord of the Rings, have to do with "sources of informa-
tion which turn out to be nonexistent." Some have to do with seemingly
orderly and complete spatio-temporal frameworks of reference which are
not continuous with our own. How this affects reference may be seen as follows. Let us construe times and places as particulars of a special kind such that they constitute a framework which is complete and continuous, i.e., every time-place particular is spatio-temporally related uniquely to every other. Let us say that we can locate a particular by placing it in this framework with respect to our present position, and that our reidentification of particulars in general is dependent upon our being able to say that a particular p located at location $l_2$ at time $t_2$ is identical with a particular which was located at location $l_1$ at time $t_1$. It has been argued convincingly that our ability to identify particulars in general is dependent upon our being able to reidentify some particulars and that this latter ability is dependent upon our ability to reidentify some places and vice versa.\footnote{26}

If the spatio-temporal framework of a fictional universe is not continuous with our own, the problem is not merely that some of the names are not attached conventionally to particulars of the actual world, but that we would not know how to use the time-space specifications of the individuals of the work to locate the particulars which might be attached to the names. In such a case, we may say that reference fails in principle.

A less striking instance of the same phenomenon is the work in which the action is set in a "microcosm," a ship, a hospital, etc., where the world of the work might be assumed to be a segment of our spatio-temporal framework and is orderly and complete throughout, but is not connected with our spatio-temporal framework at any point in a specifiable way. There is no general way to verify a sentence whose spatio-temporal conditions are not locatable in the actual world, if the sentence is one
to whose truth time and space would be relevant. (We could not, for example, assess the truth of the sentence 'Ben Franklin died the next day', found on a sheet of paper). An even less extreme case are novels about the future. They are seemingly related to us in space and time, but they are set in a time when the reference of terms is still completely indeterminate.

It is not surprising that failure of reference is one of the most important internal marks of fictional works, if we recall that one of the basic differences between normal narration and fiction is that normal narration purports to describe the actual world, whereas fictional works present and constitute fictional universes. (This suggestion suggests that not even 'London' in David Copperfield should be construed as referring to the actual world, but this is a point I will take up in section 2.3). Because the namelike expressions of a work of fiction present characters rather than referring to individuals of the actual world, the fact that they fail to refer to actual individuals is, of course, irrelevant to the value of the work. By contrast, if it should turn out that Aristotle did not exist, then most histories of philosophy would have to be judged faulty. One of the signs that we have decided to read a work as fiction is that we decide that failure of reference is not relevant to our judgment of it as a linguistic performance.

(iii) Another common device for marking a performance as fiction is the introduction of a narrator who is not identical with the author of the work. In many cases, the work is told in the first person by a narrator about whom we learn enough in the course of the work to know that he has traits which the author did not have. The author of Moby
Dick was not named 'Ishmael', and he did not, so far as we know, experience the events of Moby Dick. The narrator of Remembrances of Things Past is neither of Jewish ancestry nor a homosexual; Marcel Proust was both. Stream of consciousness and center of consciousness writing are often amenable to the same sort of treatment, depending on how much we know about the center, or the locus of the stream. Third person narration presents something more of a problem. Since the narrator often "disappears" behind the work, it is often nearly impossible to show that he is not identical with the author. This is not always so. Sometimes enough of a narrator element remains to demand that we provide an illocutionary subject for the "questions" and "exclamations" of the work. If I am right in claiming that the illocutionary subject of the "assertions" and "questions" of the work is almost always the narrator as distinct from the author, then in these cases too we have interposed narration, or as it might be called, a "narrative screen." As a last resort, in our treatment of direct and indirect quotation it will emerge that often there are syntax based ways of locating the utterance context of the "assertions" of the narrative in time in such a way as to rule out the author as their illocutionary subject. If none of these is the case, then we are dependent upon other marks of fiction to tell us that the performance is fiction.

Again, this mark of a work of fiction should not be surprising, given that one of the differences between many works of fiction and normal narration is that the illocutionary force of the act of writing fiction does not distribute over the "simple" illocutionary acts within the work. The "interposed narrator" becomes a sort of "place holder" for
the illocutionary force of these acts. Where we do not have interposed narration, the "author" who appears is not in propria persona, but the author in a fiction-telling role. If at some point in an acknowledged work of fiction, the author appears in propria persona, then we have exposition, not fiction.

Of the features of fictional works, I think that these last two are undoubtedly the most help when we are dealing with a work where our knowledge of its context, even of the author's intention, is deficient. Conventional openings are presently out of style. If we find failure of reference throughout the work or an obvious instance of interposed narration, we are usually justified in assuming the work was intended as fiction, and certainly justified in reading it as fiction. In this connection, I want now to turn to two further problems. In 2.3, I will take up the question of historical novels. They seem prima facie to give rise to objections to the view expressed in section 2.1 that the namelike expressions in fictional works do not refer to anything, certainly not to objects of the actual world. I will attempt to show that they do not present a genuine difficulty, but that we must rely on certain features of the use of such expressions in normal narration to explain their function in fiction. In 2.4, I want to show that by exploiting the parasitic relation of language in a fictional context to the language of normal narration, we can solve a general problem for fictional language, the problem raised by the relationship between dialogue and indirect discourse and narrative techniques which resemble them. One by-product of the account I shall propose is that the phenomenon of the interposed narrator can be shown to be more widespread than is at first evident.
Historical novels such as *Desire* pose a special problem for an account of fictional language. They are not clearly identifiable as fiction by any of the features of fiction listed in section 2.1. And they fail spectacularly to exhibit one of the most striking features of the fictional use of language: Failure of reference and historical inaccuracy are not, it seems, irrelevant to how we judge such a work. Nonetheless we have a strong intuition that such works are to be treated as fiction.

But the problem presented by historical novels is not merely an interesting conundrum. One reason for paying special attention to them is that certain of their features, e.g., ostensible reference to individuals of the actual world, are fairly common throughout fiction. (We need only consider the number of works which are "set in real cities.") I will proceed in my examination of historical novels and related phenomena as follows. (i) I will sketch the problem for clear cases of historical novels and then set out a range of cases which show how many works present this problem or similar problems. (ii) I will then suggest a fairly simple solution for cases which contain only a limited number of ostensible references to the actual world, then propose a more radical solution for works which are historical novels in my stricter sense.

(i) We do want to classify historical novels as fiction. Otherwise we would not speak of them as "historical novels" or "historical epics." But on a commonsense level they seem to be about this actual universe in a way that works like *The Lord of the Rings* is not. At least some of the "names" in the novel seem to refer to actual indivi-
duals, persons, places, and events. That being so, there is some basis for regarding sentences in which such terms occur as reports about the actual universe, not as constituting a fictional universe. And so, not surprisingly, if we consider the common "marks" of fiction considered in section 2.2, we find that such works are not clearly identifiable as fiction on the basis of such features.

Historical novels certainly do not employ conventionalized openings and closings. It is hard to imagine someone beginning a historical novel with 'once upon a time'. The author is in fact far more likely to appeal to the conventions of history writing.

It is certainly true that we tend to find in historical novels a richness of detail, sometimes involving poetic devices, that may lead us to suspect that we may not be dealing with straightforward history. On the one hand, the descriptions are often so detailed and so evocative that we suspect the author of intending to present, not report. And such works typically involve access to mental states of the characters, accounts of events for which there were no witnesses, etc., presented as fact, not as conjecture from facts. On such grounds, we often come to suspect that strict historical accuracy is not the object being pursued. But on the other hand, we do not find the striking use of recurrent symbols and images which often characterizes non-historical fiction. And the absence of such devices is not merely coincidental. At the most, the author of a historical novel in the very strict sense can wax lyrical about the appropriateness of some event or object. He cannot implement multiple occurrences of the object or invent a chain of similar events to reinforce its symbolic force. 27 An author can, for example,
comment upon the cold beauty and utter uselessness of Marie Antoinette's diamonds. But he cannot strengthen the sense of symbolic appropriateness by claiming, for example, that she wore them to her beheading, or that certain other empty-headed ladies had necklaces of the same kind, if we know that this was not in fact the case. If he does, we simply accuse him of cheating, or at least of embellishing the tale at his own convenience.

There is typically some failure of reference in a historical novel. The names of minor characters often fail to refer to individuals of the actual world. But in general, in a novel such as Desireé, even this failure of reference is kept to a minimum. And failure of reference is not irrelevant to the success of the work as a speech act of writing this kind of fiction. We are inclined to regard a historical novel as "faulty" if it "distorts the facts" or if it introduces major non-historical figures which interact with the historical figures, or if it gets the geography of the city in which it is set wrong. We are inclined to praise the writer of a historical novel who does enough research to get the main features of the period and the historical events right. The sources of information, if they are given, we assume to be genuine; and the spatio-temporal framework is ostensibly our own. 28

Some works of fiction classed as "historical" do indeed introduce a narrator not identical with the author or take the viewpoint of one of the characters. In such cases there is little difficulty in deciding that the performance is a work of fiction. Ben and Me is pretty clearly fiction. (Of course, being able to decide that the performance is fiction is one thing; understanding the complex relation between such a
work and straightforward history is quite another, and continues to be a matter of interest). In such cases, we probably classify the work as fiction simply because the interposed narrator is a well-known conventional device in fiction, or in some cases because the references to the narrator figure do not refer to an existent person—or because the author is not identical with the person whose viewpoint is presented. But a good many historical novels employ a third-person "omniscient" author, who more or less "disappears behind the work," as the author of a history book does. Thus often we cannot rely on the interposition of a narrator not identical with the author to explain how we recognize the work as fiction or why we classify it as such.

We seem to be left in a strange predicament. Historical novels have some of the common characteristics of straightforward history. But if they are treated as such, they are bound not to be good ones, since there is commonly minor failure of reference, and a great overspecificity, relative to the available facts. If, on the other hand, they are fiction, then it is hard to explain why there are constraints related to reference of terms and accuracy with regard to historical fact.

A look at the wider context of fiction in general shows that this kind of problem is not peculiar to the case of what we have called "historical novels in the strict sense." We need only consider the following range of cases.

Case A: We are told that the events of a novel which is a clear case of a work of fiction are set in London and that the fictional character passing into the prison is reminded of the queens who have entered there before him. Surely we are supposed to know that the prison
is the Tower of London and that the ill-fated queens were Anne Boleyn and Katherine Howard, among others. But we know this, not from the novel, but from history. A similar case is that of a novel in which, say, Napoleon occurs as a minor figure and the events of Napoleon's Russian campaign form the background for the events of the novel. We know in advance what the outcome of the war will be; the Russians will win. The rest of the book, however, exhibits such massive failure of reference and other features of fiction that it must be treated as fiction. In such cases, has the author in question referred to the city of London or to Napoleon and his Russian campaign? If he has, then we can explain easily how we fill in the facts from history. But it would be strange for the author to be both describing the actual world and presenting a fictional world. And if he is describing the actual world throughout, then failure of reference and historical falsity, however minor, are relevant to the success of the work. Or has the author merely presented us with a fictional character and with fictional settings for fictional works, contrary to appearances? If so, then we are at a loss to explain our appeal to historical fact. Either way there are problems.

**Case B:** We have the case of a genuine historical novel in my strict sense, such as *Desirée.* The book coincides with history with regard to major events; certainly it does not clash with known history. But certain minor characters are fictional, and the book tells us a good deal more about the characters and incidents than anyone, historically speaking, knows about them.

**Case C:** We are presented with a work about someone, say the Cid. In all probability, some such person existed, although much of what is
accepted about him is legendary. There are certain constraints on such a work. Given the bit of history we are more or less sure of, the author cannot very well say that the Cid was in the service of the King of Italy. Given the pattern of the legends, he cannot say that the Cid lived to old age and was buried near Madrid. (Everyone "knows" that he died young and was carried off into the clouds still riding). A similar case is that of recurrent myths and semi-myths. The Oedipus story and some of the characteristics of Oedipus constitute an immutable framework for the author's work although within this framework, the author has a good deal of freedom of movement. The same is true, for example, of the basic pattern of the Oresteia. This was already true in antiquity, when there was perhaps some lingering doubt about the historicity of the plot. And it is true of present day versions of the old story. Once a character is named 'Electra' or the story has taken a certain number of familiar turns, the outcome is certain. But there is a good deal of room for reinterpretation and elaboration within the confines of the traditional story.

Widely divergent as these examples are, they have similar features: certain constraints are imposed on the author's universe by a framework external to his work. I shall first propose a solution to the difficulty associated with examples under case (A). I will then propose a solution to the problem for works which fall under case (C) and argue that, though we have a choice about which of the two solutions to apply to examples which fall under case (B), it is generally more easy and natural and sometimes necessary to subsume such examples under the model provided by case (C).
In case (A), the problem is that in a fictional work we seem to have a reference to a historical person, event, or location, though this is not a major element of the work. I shall call such expressions "historical names." This should do no harm so long as it is remembered that the question of whether they are really names in the sense of referring expressions is still open. We very naturally add to the presented "facts" of the fictional world, facts based upon the properties and relations of the actual or historical individual to which 'London' or 'Napoleon' would refer in normal narration.

In order to explain my proposed solution to the problem, I raise two seemingly unrelated points. First, most predicates used in fictional language, we said, retain their normal narration intensions at the time of utterance (the time of the writing of the work). Secondly, although in normal narration we use names to pick out or refer to individuals, the successful use of names to identify individuals to a hearer (speaker-hearer identification) is dependent upon speaker and hearer having background information about the individual. When the reference is to historical individuals of which neither the speaker nor the hearer has had direct experience, the use of the name is even more strongly tied to background information linking the individual to other individuals.

In the context of normal narration, we should probably not want to say that names have intensions. At the most we want to say that for S and for H, a name has a connotation, i.e., a name is backed by background knowledge or beliefs (not necessarily overlapping for S and H) about the individual. (S and H might know disjoint subsets of a complete set of beliefs about the individual). In addition to the individual
connotations attached to the name by speaker and hearer in a particular situation, such names have at a given time what we might call a "usual connotation" consisting of the individuating descriptions commonly held at that time to be true of the individual in question. The usual connotation of a name is dependent upon what people at a given time generally believe true of the individual to whom the name is attached. Thus the usual connotation of a name may change over an interim. (Say that we suddenly find the long-lost public works of Aristotle. The 1972-usual connotation of 'Aristotle' will include, as the 1970-usual connotation did not, 'The author of _____', where the blank is filled in with the names of the newly-discovered works).

As a solution to the problems posed in case (A), I suggest that, just as predicates used in a work of fiction retain their intensions relative to the time the work was written, so the "historical names" used in works of fiction retain their usual connotations (at least relative to the time the work was written). That is to say, for whatever individuating descriptions would commonly have been held to be true of the historical individual in a global description of the world at that time (given that we know the time), we may characterize the character in the novel in a parallel way. Thus we need not claim that in our example 'London' refers to the actual city in order to be justified in filling into the "facts" of the fictional world whatever we know about the actual city. This leaves open the question of what 'London' as used in the novel names, if it names anything. (If my remarks about the language of fiction in the first section are correct, then 'London' used in the novel does not name anything, but instead presents an entity of the fictional world).
A problem arises immediately. The sentences of the fictional work will in some cases relate the "historical character" to the fictional characters of the work. Such relations, it might be objected, do not belong to the usual connotation of the name. London did contain the Tower of London, and Thomas More was beheaded in London; so far the name 'London' does retain its usual connotation in the novel. But the sentences relating Thomas More's manservant to London are not true—because the name of Thomas More's manservant does not refer to an actual person of the actual universe. Hence, it might be objected, 'London' does not retain its usual connotation in the novel.

There are two ways to handle this objection. One avenue of approach is to pick out certain features of London as important or essential and maintain that 'London' maintains its usual connotation in the novel if the city in the novel is presented as having at least those characteristics. I think there are several flaws in this kind of solution.

First, it involves making a decision about which traits of London are essential and which are not. There are certain problems with a philosophical position of this sort; and although the issue is far from decided, it would not be wise to rely uncritically upon the assumption that some form of essentialism is correct. In addition, in practice, even if the position is sound, it is difficult to decide and get agreement about whether a given characteristic is or is not essential to the city's being London. The second problem with this proposed solution is more serious. Even if we grant that London has certain properties which are clearly essential, e.g., containing the Tower of London, some of the characteristics we insist upon in a given case are not by any stretch of
the imagination essential. It is probably not essential to the city's being London that Thomas More was beheaded there. Nonetheless we would resent, I think, an author's setting his work in a city he calls 'London', then denying that Thomas More was beheaded there. At least we would resent it if the case before us were that of a realistic novel in a "real" setting. There may be works which have deliberately confused settings; but their effectiveness, it seems, depends on our expectations in the normal case. It seems that we simply do not accept having our knowledge about London breached in some ways, whether or not we regard the fact in question as somehow essential.

It is for these reasons that I reject the proposed solution and suggest another. What is important, it seems, is that the connections between names which do have a usual connotation in normal narration remain intact. When one or more of the names related is purely fictional, there is usually no problem in accepting the fictional fact, provided the fictional relation is not incompatible with a relation we know obtains between historical individuals. The condition, then, is that "historical names" in fiction retain at least their usual connotation. This in turn imposes two conditions: (i) the relations between two or more "historical names" must be in accord with the usual connotation of both names; and (ii) the relations between "historical names" and purely fictional names must not conflict with the usual connotation of the "historical name."

This condition may seem too strong at first sight. I think if we consider carefully our reactions to the fictional use of names which have usual connotations, it turns out to be correct. At least it represents an ideal. And the ideal requirements are probably fulfilled
more often than one would think. In practice, the requirement is not so stringent as it sounds. First, the usual connotation of a name is usually quite small (the fact that this is not the case with recent events is probably the reason why there are seldom "historical novels" in the strict sense about the recent past). Secondly, historical novels are notoriously inexact about the precise dates and times of events; and without this information, it is extremely difficult to come up with an incompatibility between the fictional fact and the usual connotation of the name. But the explanation in terms of (i) and (ii) helps to explain why a certain amount of historical research is needed even to set a novel in an "actual setting." As to whether we would make a negative judgment about the success of a work because the author's use of "historical names" used as setting or background is incorrect, the answer is probably "no." In the first place, for most of us the connotation we personally attach to a given name is likely to be very small. Secondly, although we are bothered as soon as there is an obvious clash with what we take to be the usual connotation of a name, we make allowances for minor incompatibilities. We require, it seems, that 'London' used in a novel as a background or setting satisfy much the same conditions as we would impose upon its use in an encyclopedia article on Thomas More. In the latter case, we note inaccuracies about London, but so long as the article is correct in the information about its subject, we are more or less satisfied. In fiction, as long as the work itself is effective and plausible, we overlook minor deviations from conditions (i) and (ii).

A further problem for this account arises from limitations on the author's knowledge of the usual connotation of a "historical name" which
he uses. It is certainly true that we do not expect an author to incorporate into the connotation of a "historical name" facts which occur or become known after his own time. The conditions on intensions of predicates and usual connotation of "historical names" are relative to the time at which the work was written. This is not particularly surprising, for in reading historical documents, an adjustment must often be made for the state of knowledge at the time the document was written. But authors tend to fall short, even of the usual connotation of a name, regardless of whether the facts in question were in their own time part of the usual connotation of the name. Yet such a work may be a successful work. Again, this is no real shock. There is an analogous phenomenon in normal narration; as noted in section 2.1, the hearer may make adjustments for the speaker's deficient knowledge or false beliefs, so that he succeeds in discovering what particular the speaker has in mind, even if the words uttered plus context do not pick out a unique object. We can treat the author in such a case as a speaker with deficient knowledge or false beliefs; and it is not surprising that we make similar adjustments for an author's deficient knowledge or false beliefs. (But, just as in normal narration, we do not say that the speaker has correctly used an expression to identify a particular in such a case, so in the parallel case, the author has used a "historical name" incorrectly).

As previously pointed out, if the historical name is merely part of the setting, failure to observe ideal requirements for the use of historical names may not count significantly against the work, although we are inclined to view the artistic accomplishment of the work in general as less if the author does not meet the requirements.
The difficulty posed by examples which fall under case (A) are, I think, fairly plausibly handled in this way. If an author uses a name which has a usual connotation, he must take some trouble to avoid adding facts to his work which conflict with parts of that usual connotation. But, as already noted, we need not claim that such names refer to actual individuals or that sentences containing them must be true of the actual world, to explain our reluctance to accept what we originally termed "historical inaccuracies" and our tendency to fill out the facts of such a work with our knowledge of the actual world. We can instead treat such names as if they were predicates and had a sense, where the "sense" is constituted by the usual descriptive backing for the name.

We could try to extend this treatment to the case of historical novels such as Desirée, and say that in such a case we have a large number of "historical names" with usual connotations and that many of the names are interconnected by virtue of their usual connotations. The task of the author on this suggested treatment of instances of case (B) is to arrange the facts of his work in accord with the constraints imposed by the "grid" of historical names he uses and their usual connotations.

But this solution will not quite do, although it does explain a part of the constraints imposed on the author. It is hard, for example, to see how we can fit the fact that it was raining on the day Mary Stuart set out for France into the "usual connotation" mold. Surely someone in Mary's entourage saw some of the raindrops. But the author may not have the information necessary to anchor the event in question to some particular name he uses; the information may simply not be available.
And even if the author has the information, he may choose not to use that name. Nonetheless, it seems clear that he is constrained to take the event in question into account (at least by not including the report that it was not raining on that day), even if he is not so constrained by the usual connotations of the names he does use. Similarly, no doubt there were numerous Frenchmen milling about the palace courtyard on the day of the St. Bartholomew's Day massacre and numerous stones thrown. But an author writing a novel about this period may in fact use the names of the stones. Still he must, it seems, take account of the event; at least he may not include the report that the courtyard was still all day long or that no one in the palace heard a sound from outside.

This constraint does not seem to come from the usual connotation of the names the author uses. No doubt if we knew all the names involved, and if the author had to use them all, we could always explain this sort of constraint on the author in terms of the treatment of case (A). But since these conditions do not obtain, we need some other explanation for the additional restrictions.

One suggested solution to the problem is that the events in question ought to be treated as particulars, that they should be assigned names or designations with a usual connotation, quantified over, etc. This is a suggestion that has been opposed by philosophers for various reasons; and it has yet to be shown that it is in general a necessary or a desirable procedure. But more importantly, it does not solve the problem at hand. For the author in question need not use the designating phrase 'The St. Bartholomew's Day massacre'. He need not even use 'the palace courtyard'. And if he does not use it, then the constraints upon
him cannot very well spring from its usual connotation. The question of where the additional constraint comes from remains. For a solution to this question related to instances of case (B), I suggest that we turn to the treatment of instances of case (C), and see if an appropriate solution suggests itself.

(ii) The situation in examples which fall under case (C) is actually not that we are presented with names that are names of individuals of the actual world about whom we have considerable information. And although the earlier playwrights may have thought of themselves as writing history, by the time of Aristotle (given that he reported the matter correctly in the Poetics), no one knew if that was the case, and no one cared. The point of writing a work about Oedipus or a work dealing with the Oresteia was rather that there were certain well-known plot patterns such that once events began to take a familiar turn, the general outlines of what would subsequently happen were set. Usually, as Aristotle saw, there were certain stable congeries of personal characteristics, usually bearing a familiar name, from which the action developed, although these were none too specific. But, as Aristotle pointed out, the names used, although convenient, were no longer of crucial importance. It makes little difference whether the main character is called 'Antigone' or 'Lostcause'. The basic plot is set, once the 'Antigone-figure' defies authority to bury her brothers. She will get caught, and will be condemned to death, taking the fortunes of her family with her. There is thus a sense in which the treatment of legendary and mythic plots are dependent on the modal of cases which fall under case (A). But name-independent plot constraints represent a significant
step forward, for they are no longer name or history dependent. The genius and challenge of the playwright is precisely his ability to interpret and provide an explanatory structure for a series of foregone conclusions. The exact nature of the main character's motivation, however, is left up to the author. She may be portrayed as motivated by a fairly primitive compulsion to fulfill a religious duty, by a misguided attachment to a misguided parent, by the desire to introduce rationality into a savage and authoritarian culture, by the desire to do something which will establish her radical individuality within a narrow range of possibilities, etc. Similarly, we know that the Agamemnon-character (whoever it is that comes home, having done in a daughter for business purposes) will be murdered by the Clytemnestra-character, touching off a family disaster. It is a mark of the genius of Aeschylus that he arranges the events leading up to this foregone conclusion, portrayed as a final purgation by fire, by a subtle interplay of light and fire imagery, and an intricate arrangement of portentous double meanings.33

We have a similar situation in fictional works which are based on legendary characters. In our time, we may be unsteady about the factual bases of the story. Moreover the question is of no great importance. The author, for example, is bound by the usual connotation of 'St. Joan of Arc', whether the connotation is due to history or due to legend. In addition, he has to work within a framework of events, some deriving from history, some from legend, even if those events do not fall within the usual connotation of the names he uses. The source of the constraints is a matter of little importance, even though we still
L

115
regard the name as having a h is to r ic a l usual connotation.

The c ru c ia l

job of the w rite r in th is case i s to f i l l in , say, the ch aracter of S t.
Joan, so th a t the course o f events (we a l l know what they have to be)
becomes p la u sib le and in te r e s tin g .

We are in c lin e d to view S c h ille r in

h is S t. Joan as ch eatin g , when he allows S t. Joan to d ie on the b a tt le ­
f ie ld , in stea d of a t th e s ta k e .
usual connotation of 'S t . J o a n ',

In th is case, he has run afoul of the
S im ilarly , i f he scheduled a period of

peace throughout France during a period when we know h is to r ic a lly th a t
th e re were b a ttle s going on, he would have run afoul of h is to r ic a l p lo t
c o n s tra in ts .

Honegger's v e rsio n has S t. Joan dying a t th e stak e (as we

know she d id ) looking a t a cross (as legend has i t she did) and singing
(which we assume she d id n o t, but accept as a v iv id versio n of an in te r ­
n al monologue).

She sings th e f in a l stan za of a fo lk song about giving

up dinner to buy a candle to burn before th e a l t a r , and changes th e l a s t
l i n e from 't o buy a candle b rig h t* , to 'And I myself s h a ll be the candle
b r i g h t '', thus connecting h er humble o rig in s to her stran g e d estin y and,
by augmenting the sim ple meter of th e o rig in a l w ith th e th re e heavily
s tre s se d s y lla b le s o f *1 m yself' emphasizing her aloneness, her agony,
and her consummate p r id e .

Honegger has done a f a i r l y good job, ju s t as

S c h ille r has done a bad jo b .
In h is to r ic a l novels o f case (B), I su g g est, we are presented with
an extreme case of th i s kind of phenomenon.

The author i s constrained

by th e usual connotations o f th e h is to r ic a l names he u se s.
ad d itio n , th e p lo t i s s e t .

But in

We can g e t i t out o f any h is to ry book.

Of

course the reason th e p lo t i s s e t by h is to ry , and n o t, say, by legend,
i s th a t the author i s using a la rg e number of fu ll-fle d g e d h is to r ic a l


names. But the important fact is that the plot is set, not that it is set specifically by history. The broad outlines of, say, Napoleon's and Josephine's characters are set, and the outlines of the plot are determined, insofar as the story must develop out of the characters and must accord with historical fact. The challenge to the writer of a historical novel is to work within a very constricted framework of a prearranged plot by introducing minor characters, ascribing thoughts and motives to the characters, etc.

This approach in terms of a plot framework explains the constraints upon instances of case (B) which are not solely a result of the usual connotations of the names used in the work. It explains our concern for historical accuracy, even with regard to events which are not part of the usual connotation of any name the author actually uses. Within these constraints, the author adds to the usual connotations and the plot as imposed by history, the material which makes the flow of the plot plausible and interesting. His ability to do this within a very narrow framework is his peculiar artistic accomplishment.

This analysis of the wide range of phenomena related to historical novels is consistent with the basic account of the language of fiction I have sketched, although such cases seem prima facie to constitute a counterexample to this kind of account. In cases (A) and (B), the treatment of the use and understanding of expressions in fiction is shown to be parasitic upon the rules for use of and our way of understanding these expressions in normal narration, although in each case, the fictional use is significantly different from the normal narration use. The basic difference turns out to be that what are truths about individuals in
normal narration are rules for the use of namelike expressions in fiction. In legendary and mythic plots we have constraints imposed upon the use of namelike expressions and construction of plot which are similar in their function to the constraints on "historical names," although the source of the constraints is myth and legend, rather than truths about the actual world. In none of (A)-(C) need we claim contrary to the view expressed in section 2.1, that the "historical names" which occur in historical fiction refer to individuals of the actual world.
The language of fiction is parasitic on normal narration. Some of the senses in which this is true we considered in sections 2.1 and 2.2. In this section, I want to draw attention to a further element of this parasitism, which we noted only in passing in the previous section: The syntactic and semantic complexity of declarative sentences in fictional language runs parallel to that of sentences of normal narration. We find in fiction such sentences as 'The stars were fading and a grey light was slowly growing', which are clearly conjunctions of simpler sentences. And just as we can have in normal narration such sentences as:

(1) Mary Tudor said, "You will find Calais in my heart."

and:

(2) Galileo said that the earth moves.

we may have in fiction such sentences as:

(3) "It is gone forever, and now all is dark and empty." 35

and:

(4) Mr. Bilbo Baggins of Bag End announced that he would shortly be celebrating his eleventy-first birthday. 36

Thus the syntactic parallel is more complete than is immediately obvious.

An account of the function of a given kind of expression of the language of fiction will not be precisely the same as the account of the function of expressions of normal narration which are, syntactically speaking, of the same kind. We saw three examples of this parasitism in section 2.1 and yet another example in section 2.3. In this section I want to take up the question of such sentences as (3) and (4) and related contexts.
Sentences of direct and indirect quotation in fiction, represented by (3) and (4) constitute one of the most important technical problems in the analysis of any work. In the first part of this section I will first give a set of cases which show the extent and complexity of the phenomenon of direct and indirect quotation and then explain why such cases pose a problem for criticism. The problem, in short, is that some works which, syntactically speaking, might be construed as one long intentional context seem to resist such treatment. I will advance a tentative solution to the problem, on which the refractory cases turn out to be only apparent intentional contexts. This leaves us with a residual problem—what function does the apparent intentional context serve. I will then argue that, given that an analysis of real intentional contexts should proceed along certain lines, there is an important function which such apparent intentional contexts could have.

In this section I will give only general rules of thumb for the decision about whether a given work is or contains a real or only an apparent intentional context. A more precise procedure will be given in chapter V.

In order to delineate the extent and complexity of the phenomenon of direct and indirect quotation in fiction, we need only consider the following range of cases.

Case A: In a novel which has a third person omniscient author, we are presented with a sequence of sentences consisting of dialogue, or with a sentence or sentences which appear to fall under an intentional operator such as 'he said that' or 'he thought that', etc. A more extreme example is that in which at some point in the narrative, a
character tells a story in which he himself figured. We may also consider cases in which the "testament" of a character is found and read, and cases where we find ourselves momentarily seeing events through the mind of a character or party to a character's train of thought.

Case B: Some novels are partially or wholly epistolatory; they consist of or contain diaries or letters of the characters. The Sorrows of Young Werther combines third-person narration and Werther's diary. Jane Austen's Lady Susan consists almost wholly of the correspondence of the lady and her friends.

Case C: A set of characters in a situation decide to tell each other tales to while away the plague, the pilgrimage, etc., tales in which they themselves do not figure as characters. A similar case is the following: an old man comes in, sits down, and begins to tell a tale; 342 pages later, he rises and leaves. A third kind of case is Storm's Immensee, in which an old man sits in a darkening room and "remembers" the rest of the novella; the events of the darkening room are only tenuously linked to the events of the story by the old man, who is, we presume, the younger man of the story.

Case D: The entire work consists of a first person narrative. An example of this is Melville's narrative beginning "Call me Ishmael." A similar problem is presented by stream of consciousness novels. Such a work, one might argue, has to be taken as one long quotation, or as prefixed by 'The narrator thought'. One might argue here that we have merely an extensive case of the sort of example which falls under Case A, or that the work is one long chain of thoughts of a person.
With a little ingenuity, we can think of various combinations of these (dialogue in a first-person narrative, etc.) which make the variety of examples seem to defy systematic treatment. Nevertheless, I would contend there is an intuitive difference between examples which fall under cases (A) and (B) and those which fall under cases (C) and (D), a basic difference which persists throughout the bewildering mass of cases. In case (A) the omniscient narrator presents linguistic performances of various kinds or intentional facts among the non-linguistic "facts" of a fictional world. In case (B), the world as we are given it seems to consist almost wholly of linguistic facts. It seems plausible that before we assess the truth of the reports, we should prefix any claim about the content of Lady Susan's letters with 'Lady Susan said' and the parts of Werther's diary with 'On that day Werther wrote'.

But the examples in cases (C) and (D) resist such treatment. In many instances of case (C), it seems that there is a "frame" which is wholly divorced from the story or stories inside it. Do the "facts" of the work consist of the few facts of the "frame" plus a series of linguistic facts, each prefaced with, for example, 'The prioress said...'? This is a little implausible, since we sometimes talk about the events within the tale as if they were non-linguistic facts of a fictional world, i.e., the world of the Prioress' Tale. And in the Storm case, we do not always qualify our critical claims with 'The old man thought'. In examples of case (D), we are disinclined to say that Moby Dick, for example, is a fictional world which consists of one linguistic performance or one string of mental events; we are much more inclined to
say that the first person work is to be treated like a third person work, that the events of Ishmael’s tale just are the events of the fictional world, not merely parts of the one intentional event of the world, i.e., Ishmael’s linguistic performance. We do not preface every description of Ahab’s action with 'Ishmael thought that' or 'Ishmael said'.

We have to decide what the facts of the fictional universe are. Thus the question of how much of a work is in an intentional context is important. If the cases under (C) and (D) have to be assimilated to cases (A) and (B), our intuitions about such works (and nearly everybody’s practice) are just wrong. The problem then is to decide how our intuitions about (C) and (D) can plausibly be accounted for, how we can treat such cases differently from the cases under (A) and (B). And if cases under (C) and (D) are treated differently from those under (A) and (B), we will still have to account for why the author should have used an apparent intentional context.

Roughly speaking, if the story inside the frame is independent of the frame, the frame may be treated as a stylized way of introducing the fictional world. The frame functions like 'Once upon a time... and they all lived happily after', or like the prologue of a play in which an actor comes out and delivers an opening monologue ending with 'Let the play begin', or 'And this is what happened ...'. We need not prefix our claims about the events of The Wife’s Tale with 'The Wife said', anymore than we would prefix our descriptions of the play with 'A character said'. The intentional context here is only apparent. We need not specify that the claims are about The Wife's Tale. The
rest of the cases under (C) have an interesting common feature. Once
the story inside the frame begins, the supposed intentional context
erodes into an omniscient narrator situation. Storm’s old man
"remembers" events he could not remember because he was not present
when they occurred. When this sort of thing happens, i.e., when the
narrator conveys information inappropriate to the supposed intentional
context, I think we can easily say that the intentional context is only
apparent.

Cases (B) and (D) are distinguished in a similar way. Although
Ishmael could be construed, formally speaking, as a limited first person
narrator, this is implausible. Ishmael knows too much. He reports on
scenes from which he was absent, reports on other characters’ states of
mind as if he had privileged access to them, etc. He is most plausibly
construed as a third-person omniscient narrator. A good many cases of
first-person narration follow this pattern. The narrator knows what
he could not qua character know or knows with certainty what he qua
character could only surmise. In such cases, the intentional context
is merely apparent.

The guidelines laid down here are only rules of thumb. In chapter
V, we shall consider the problem in greater detail; and the more precise
account will enable us to deal with borderline cases between (A) and (C)
and between (B) and (D). In broad outlines the solution is clear, I
think. But we are left with a residual problem. Why does the author
use apparent intentional contexts in cases under (C) and (D)? Is this
a slip on his part, or perhaps a deliberate attempt to confuse and mis-
lead the reader? For an answer to this question, let us turn to an
analysis of the use of intentional contexts in (1)-(4) in normal narration and see if it suggests a reason for the author's use of such apparent intentional contexts.

There is as yet no definitive analysis of such contexts for normal narration. Nonetheless, I think enough of the guidelines are clear to enable us to answer our question about apparent intentional contexts. It is perfectly clear, I think, that the "inscriptionalist analysis," on which "'You will find Calais in my heart'" and 'the earth moves' name or refer to utterances of Mary Tudor and Galileo, respectively, is unacceptable. The quoted expressions and the expressions following 'that' must be taken as semantically structured parts of the containing sentences for the following reasons. In order to make sense of:

(5) Mary Tudor said, "You will find Calais in my heart"; but they didn't find it there.

'it' must be taken to cross-refer to 'Calais', and both expressions to refer to a city. On the inscriptionalist analysis, the cross-reference simply cannot be explained. In addition, if expressions in quotation contexts have no semantic structure, then:

(6) Quine says that quotation "has a certain anamalous feature." which looks correct and clear, would be ill-formed.

On a view recently advanced by Davidson, the quotation marks in (1) plus the expression inside the quotation marks contain a reference in some sense to an utterance of Mary Tudor which had a given shape. The expression inside the quotation marks furnishes an example of the shape of Mary Tudor's utterance (in this case, the written shape representing the spoken shape). On this analysis, using a paratactic colon to display the example expression, (1) becomes:
(N) Mary Tudor spoke a sentence of this shape: you will find Calais in my heart.

What follows the colon is the utterance of the speaker of (1), whoever says all of (1). It is merely an utterance and has no assertive force. (1) is true just in case on the basis of what Mary Tudor said, she and the speaker of (1) are co-utterers. A and B are co-utterers just in case they produce linguistic tokens of the same type. The displayed utterance is used, tokened, by the speaker, although not to perform an assertive act. It is not the name of an expression. There is thus no reason why pronouns from the surrounding context cannot cross-refer into it.

I think we can simply adopt some version of Davidson's account for sentences which fall under cases (A) and (B). In sentence (3), we are told that Bilbo spoke a sentence of a given shape. I am inclined to think that short movements to stream of consciousness or center of consciousness passages should also be treated in this way, since they seem to add to the facts of the novel a mental event of the form 'The character thought...........'. Diaries and letters, when introduced into the narrative of a work of fiction (Case (B)), may be taken as extended cases of quotation, the speaker's assertion that such and such a character produced a linguistic performance of a given shape, followed by an example of the shape.

We can represent Davidson's account of sentences such as (2) as:

(2') Galileo said that: the earth moves.

which is an abbreviation of:

(2'') The earth moves: (∃x) (Galileo's utterance x and my last utterance make us samesayers)
There are some problems with Davidson's analysis. The notion of same-saying is somewhat unclear. And it has recently been pointed out that (2''') has at least one entailment which (2) lacks: (2'') entails the existence of the speaker of (2), although (2) does not. Nonetheless, the analysis is on the right tract, I think. And it has an interesting consequence. Let us suppose that what Galileo said in 1640 was, "I will be sick of it all." The true report which we as present day speakers make is no doubt:

(6) Galileo said that he would be sick of it all.

The use of 'would be' in our utterance is a result of our being in the future relative to Galileo's utterance. We are trying to produce something which "has the same import" as what Galileo said. And whatever the exact analysis of "same import," what we say must include the possibility of the earth moving between 1640 and 1972, as (6) does, and

(7) Galileo said that he will be sick of it all
does not. (7) is certainly odd, and it assuredly does not say what it should say, even if it is nondeviant. Whenever we find such a construction, we can assume that the speaker of the report on someone else's utterance (or thought, wish, etc.) is located in the future relative to the verbal or mental event reported on. And this has an interesting consequence for the analysis of indirect quotes in fiction. For now it is by no means clear who the samesayer is in (4). Prima facie we might be inclined to say that it is Tolkien who is the samesayer. He is, after all, the author of the work of fiction in which (4) occurs. But there is in fact no reason to assume that the world of The Lord of the Rings is spatio-temporally continuous with this universe, the one in
which Tolkien is an individual. It is in fact pretty clear that the events of *The Lord of the Rings*, including the utterances, are not for Tolkien past events. If Tolkien is taken to be the speaker of (4), then we cannot explain the syntax of (4) by running the process which produced (6) backwards. And I do not know how else we could explain it without *ad hoc* solutions. But there is someone for whom the events are past events, i.e., the narrator, as distinct from the author. Tolkien makes this clear in the narrator's preface and the epilogue. All of this points to the conclusion that the samesayer, the illocutionary subject of (4), is the narrator, as distinct from the author, and that Tolkien, at least in practice, recognizes this. (Similar considerations lead us to deny that the narrator of novels set in the future and written in the past tense is identical with the author, for whom the events are in the future). These cases are interesting because they involve a third-person omniscient narrator. It is in such cases that it is most tempting to claim that the narrator just is the author. If I am correct, the syntax of fiction itself almost always bears out the view that this is not the case. There is in most cases a narrator distinct from the author. The narrator is what we have called the "illocutionary subject" of the work; and the narrator, as the illocutionary subject of (3) and (4) is the co-utterer of Frodo in (3) and the samesayer with Bilbo Baggins in (4).

This result of the treatment of cases under (A) and (B) suggests a plausible explanation for the author's use of apparent intentional contexts in examples which fall under (C) and (D). In such cases, the apparent quotation context or thought-context might be construed as a
making explicit of the fact that we are to construe what follows as a whole as fiction by dissociating the author from the narrator. The situation is analogous to that in which the "stage manager" or one of the actors comes out in front of the curtain and delivers a prologue. Of course, in some sense, the play begins with the prologue. But the action begins afterward. In a narrative frame or a case of first person narration which falls under (C) or (D), we are presented with the illocutionary subject of place-holder of the work. And this subject is presented or described in such a way that he is clearly not to be identified with the author. In the cases under (D), this is accomplished by introducing a first person narrator clearly not identical with the author. In the cases which fall under (C), Storm's old man who "remembers" and the "frame" characters who introduce the tale inside the frame, serve the same purpose. In the one kind of case, it is the quotation context which is merely apparent. In the other kind of case, it is the thought- or memory-context which is merely apparent.

Thus it seems that despite the bewildering variety of examples, there are only two basic kinds of cases to be considered when we talk about direct and indirect quotation contexts in fiction. (i) There are cases where we are presented with a linguistic or mental performance among the events of the work. (ii) There are cases where an ostensible intentional context is a conventional means of introducing the illocutionary subject of the work. Of course, we cannot tell on purely formal grounds which individual cases are instances which fall under (ii). Sometimes a first person narrative is an extended instance of case (B), and some stream—or center of consciousness works are extended
instances of case (A). And sometimes we have iterations of cases of
type (i) and type (ii). A method for handling borderline cases and
iterative contexts will be discussed in chapter V.

This, I think, can serve as a model of how to handle problems
which arise in the language of fiction. We use works of literature
themselves and informal theories and intuitions about them to de­
lineate the scope of the problem. We then turn to standard philosophy
of language solutions for syntactically similar expressions in normal
narration for a treatment of expressions which are parallel to them in
function and for a rationale for using the construction in question when
the parallel breaks down.
Notes

1 I shall restrict the present discussion to prose and verse narrative, excluding for the time being theatrical performances and written dramas. The reason for this is obvious; the latter are not clearly linguistic or wholly so. In addition, I shall speak of a single author writing a work, ignoring for a moment the problems with multiple authorship and the obvious fact that the author's performance results in his manuscript, whereas we all read copies, sometimes illustrated or abridged copies.

2 This comparison of the author of fiction with the historian or the biographer is intuitively appealing. Not surprisingly, it has been a popular comparison throughout discussions of the nature of fictive language. In modern times, we find it as early as Ryle's article "Imaginary Objects," Proceedings of the Aristotelian Society, Suppl. XII (1933), pp. 19-43. In that article Ryle speculates on the conditions under which we would have called Dickens' novels "biographies."

3 It has been objected that this claim that the employment of conventions allows us to tell what kind of act is being performed is by no means uncontroversial. Davidson attacks this view in his article "On Saying That," reprinted in Words and Objections, ed. Hintikka and Davidson (New York: 1969), pp. 153-173, and he has made the more general point repeatedly in unpublished lectures. Certainly it is true that without knowledge of the context of a linguistic performance, we cannot tell, for example, if someone has used a declarative sentence to make an assertion or not. But this sort of general objection does not raise appreciable problems for the kind of account I propose. There are two reasons for this. (i) The objection decreases in plausibility as the speech act in question becomes more complex. (ii) There are many cases which are clear instances of the kind of speech in question, and we can work with these.

My account owes a good deal to the method set forth by Searle in Speech Acts (Cambridge: 1969). But my account differs from Searle's in several ways. I am not, as Searle takes himself to be doing, looking for necessary and sufficient conditions for the fictive speech act's being performed. I take myself to be advancing an empirical theory about a kind of language of which we have clear-cut instances. Finding necessary and sufficient conditions for a speech act's being a successful work of fiction seems unfeasible because (i) the act in question is extremely complex and varied; and (ii) a work of fiction is a work of art, and designating necessary and sufficient conditions for a successful work of art is notoriously pointless.

4 Searle, Speech Acts, expresses a somewhat similar point when he says that "what we can mean is a function of what we say," p. 43. Where Searle would use 'mean', I generally use 'can intend in the stronger sense.'
This is what Searle, Speech Acts, calls the "reflexivity of intention." "The way the reflexive intention works then, as a pre-
liminary formulation, is: the speaker S intends to produce an illo-
cutionary effect \( \text{IE} \) in the hearer H by means of getting H to recognize
S's intention to produce \( \text{IE} \)." p. 57. The speaker intends to get the
hearer to "know (recognize, be aware of) these things ... in virtue of
his knowledge of the rules for the sentence uttered." p. 48.

Here we might ask what counts as "internalization" of the con-
ventions or rules on the part of the performer and audience. Need
they, for example, be conscious of a complete set of rules and be able
to make them explicit? Clearly the speaker need not, anymore than a
man who has heard a number of verses of a kind and developed a certain
"feel" for them has to be explicitly conscious of or be able to express
his guidelines in order to be able to produce verses of the kind in
question. The same kind of considerations apply on the part of the
audience. Searle, Speech Acts, makes a similar point, p. 35.

These are what Searle, Speech Acts, calls "propositional acts."
It is not clear whether Searle considers both referring and predicating
(my "characterizing") propositional acts, or whether he is talking about
the combined act, as I am. Even on this latter interpretation, it is
somewhat misleading to designate the conglomerates per se as acts, since,
as Searle points out (pp. 23, 25, 29), they are not performed and do not
occur alone for their own sake, but rather always occur as components
of a complete speech act.

In certain simple situations, it may seem that we perform a com-
plete speech act merely by referring, for example, the person who
shouts "Danger!" But the speaker in this case must be construed as
saying elliptically that something is dangerous, not merely as referring
to the situation.

Not all linguistic acts have a referring-characterizing component.
'Hello!', for example, does not.

In an unpublished paper, Richard Garner explicates this as follows:
If in describing his experience (making a statement or assertion about
it) "S utters a sentence \( (E \text{ is } \phi) \), then S classifies \( E \) as a \( \phi \) and S
discriminates \( E \) from other experiences from which he would withhold
\( \phi \)." I am dealing with the general question of statement making, not
with the restricted context of statements about experience.

If lying is construed as an illocutionary act, it constitutes an
exception to this pattern. The perlocutionary effect of misleading the
hearer is dependent in this case on the hearer's not apprehending the
illocutionary force of the speaker's (lying) speech act, or, alternatively,
upon his apprehension of the speaker's misdescription as an assertion.

Garner, in the above mentioned paper, makes this point about
description of experiences as follows: "If S classifies \( E \) as a \( \phi \) and
S discriminates \( E \) from other experiences from which he would withhold
\( \phi \), then S relates \( E \) to past experiences to which he applied (or would
have applied $\emptyset$ and distinguishes $E$ from past experiences to which he
would not have applied $\emptyset$ . . . $S$ recalls these past experiences."

11It is this feature which prompted Moore, among others, to call
fiction "a pretend use of language." This is a rather unhappy way of
speaking, for fictional uses of language are legitimate uses, although
such language is not directed toward the actual world.

12The answers have been many and varied. The question led Ryle,
"Imaginary Objects," pp. 35-6, to construe 'Mr. Pickwick' as a "pseudo
designator," and to construe the whole work as one big complex
predicate which is not ordinarily true of anything. The meaning of
such a predicate he takes to be unproblematic, since we know what it
would be for the predicate to be true of something (perhaps we know
what it is for its individual parts to be true of something). This
leads Ryle to the view that if there should happen to be one or more
things of which the complex predicate were true, the alleged "novel"
would become a biography of those objects, pp. 38-39. (Ryle does balk
at saying that Dickens would thereby become a biographer, p. 39).

Brathwaite, "Imaginary Objects," Proceedings of the Aristotelian
Society, Suppl. XII (1933), notes an ambiguity in Ryle's treatment of
the notion of a "pseudo-designator." Basically, he disagrees with Ryle
that if the expression in question is a designator, i.e., equivalent to
the predicate 'the thing named Mr. Pickwick,' it would be true of more
than one object. Such a "predicate" can be satisfied by only one
individual. If Ryle holds that the complex predicate could be true of
more than one individual, then he is at that point treating 'Mr. Pickwick'
as a variable, not as a designator. Thus, he concludes, Ryle's treat-
ment of 'Mr. Pickwick' is ambiguous. And so, for that matter is
Dickens' use of 'Mr. Pickwick'. Hence the language of fiction, though
we understand it, i.e., we understand every element of the compound
predicate, is meaningless because it is ambiguous, i.e., has no deter-
nimate meaning. On Brathwaite's analysis, it is still possible for
Pickwick Papers to turn out to be a biography, but it will be a bio-
ography of only one individual, namely someone who was called 'Mr.
Pickwick' by Dickens and to whom all the events of Pickwick Papers
happened.

G. E. Moore, "Imaginary Objects," Proceedings of the Aristotelian
Society, Suppl. XII (1933), makes a radical break with this sort of
account. We need not, he says, reduce fictional names to predicates to
show that the language of fiction is meaningful. Dickens uses "Mr.
Pickwick" as the name of some one individual which he pretends to have
in mind. Of course, there is no such individual. Pickwick Papers
will never be a biography of, i.e., be a set of assertions about, a
real individual, since, even if a real individual were to satisfy all
the predicates, he would still not be the individual which Dickens
had in mind.

13As R. Gale, "The Fictive Use of Language," Philosophy, LXVI
(1971), 329, point out, "if we were to make authors of professed
fictional works immune from libel suits, it would encourage flagrant
abuses, since a person could defame another with impunity, provided he
preceded his work with the words 'A Fictional Story'."
Thus we are told in Tolkien's *Fellowship of the Ring* that to be an orc is to be an evil being, capable of only the lowest thoughts and impulses, where we understand these latter predicates, though we did not previously understand 'is an orc'.

This terminology is borrowed from Strawson's *Individuals* Methuen, London, p. 16, although Strawson spends very little time on the analysis of proper names.

Syntactic similarity is based on occurrence in similar contexts, in English by capitalization, in Russian by patronymics, etc.

It has been pointed out that syntactical considerations alone will not tell us that an expression is being used as a proper name (cf. fn. 14), i.e., that the expression is being used to refer in a namelike way. Ordinarily fairly complex specifications of the context in which the expression is used are needed to determine this. These involve, among other things, the speaker's intention, his background knowledge, the conventional attachment of the expression to the individual in question, etc. I agree. Throughout this section I will be using cases where we suppose that the contextual requirements are fulfilled, that we can assume that the expression is being used in a referring context.

This, I think, is Strawson's considered view in *Individuals* (London: 1959), although in that book he gives two different sets of conditions for speaker-hearer identification. In one set (p. 20) he does not require that there be an object such that it is in fact correctly and uniquely specified by the speaker's description (but rather only that the hearer correctly identify the particular which the speaker has in mind on the strength of what the speaker says). Later he adds this requirement (p. 240).

The qualification 'in the context of the novel' is added because the specific sentences in question may not be incompatible, logically speaking. It is not logically contradictory to present a character as being in two places at the same time. But if the universe is one in which the normal laws of time and space hold, i.e., one in which a sentence expressing such a law is explicitly present or may be derived, then the two "fact" sentences plus the law are incompatible.

At present we have no choice but to define consistency with respect to what would be the case on a hypothetical assignment of truth values. In chapters III and IV, I will introduce a more convenient procedure related to assigning truth values to sentences about the novel.

Gale, "The Fictive Use of Language," maintains that expressions used in a fictional work are fictive just by virtue of the author's desisting from performing the illocutionary acts typically performed by such a use of a sentence. This leaves him no way, so far as I can see, of accounting for the difference between "fictional imperatives" and "fictional assertions."
Gale, "The Fictive Use of Language," 332, maintains that the author is the illocutionary subject of the work and, for example, that, if an adult story teller begins his story with "I am five years old," he has certainly said (although perhaps not asserted) something false. This move is motivated by Gale's concern that (i) sentences used in fiction mean what they mean in non-fictional contexts; and (ii) the meaning of a sentence is tied up with its truth conditions. If sentences of a novel do not have the truth conditions they would have if they were written about the actual world, then we do not know what they mean. It is, however, doubtful that he will be able to get around the argument I advance in section 2.4 that if sentences in fiction mean what they would mean in non-fictional contexts, the narrator is in most cases not identical with the author.

This may very well be an instance of a more general truth about actions. We would hardly say that someone was knitting a sweater if he were putting cardboard squares together, even if he protested loudly that he was knitting a sweater. And we would hardly say that someone was playing chess if he were abiding by none of the rules, even if he said that he was playing chess and had the proper equipment. The second case is more pertinent to the case at hand as it is not the pragmatic inappropriateness of the person's behavior, but his failure to play by the rules, which disqualifies him.

This is similar to the sense in which a present day speaker can indicate that he is asking a question by inverting word order or raising his voice and need not prefix his performance with 'I ask'.

This should be distinguished from the related phenomenon of seeing the world of the work, if only momentarily, from a character's viewpoint. An example, I think, will make this clear:

He took his change and, as he turned to go back to the waiting car, two men stepped from around the corner. Don Corleone knew immediately what was to happen.

Mario Puzo, The Godfather (New York: 1969), p. 81. Here we are told something as fact which in the actual world we would have to infer from the Don's actions. The clue that it is fiction is that the presentation of his state of knowledge is not prefixed by 'In all probability' or 'It was obvious that'.

Very coolly, very deliberately, Michael fired the next shot through the top of his white-haired skull. The air seemed to be full of pink mist.

Mario Puzo, The Godfather (New York: 1969), p. 81. Here we are definitely seeing the events through the eyes of a character, if only momentarily. The obvious prefix for the second sentence would be 'To Michael', rather than 'In all probability'.
This may be because we commonly expect novels to have some perlocutionary force, although we do not expect this from history. It would be a mistake to suppose that all symbols are recurrent, or that they are ordinarily so gratuitous with regard to the plot that they would be remarkable coincidences in fact. Sometimes they are neither; they are simply "too appropriate to be anything but a symbol," e.g., Jake's impotence as a result of the war in *The Sun Also Rises*.

This is essentially the account which occurs in Strawson, *Individuals*, pp. 17-26, 31-37.

Obviously some sort of continuum runs from, say, *War and Peace*, which has a historical setting, but fictional characters, to, say, *Desiree*, in which nearly every character is a historical figure, and which agrees with history on the main events. But for the purposes of this section I shall use "historical novel" to refer to works from this latter end of the spectrum, which have as their main characters historical figures and whose plots consist primarily of historical events.

Of course, as remarked before, a symbol need not occur more than once to be effective.

This describes how most of us do in fact respond to historical novels; we might possibly be misguided in our reactions. But I am working under the assumption that a good many people know, at least in practice, how to perform and respond to the kind of performance in question. The object of this section is to get behind these data, taken at face value, and see if there is a plausible theory to explain the facts. Linguistic practice is not always good evidence, but it is good evidence for theories about linguistic practice.

This is always true of names used in natural language, I believe. Even in the case where S and H are in the presence of the individual denoted by the name, and S points at it as he utters the name, there is background information, for example, where the object is with respect to other things.

If it turned out to be the case that, say, Strawson did not write the books usually attributed to him, I should merely say that I have been given a lot of false information about Strawson, for I have seen him and know what particular I attach the name to. If, on the other hand, it were to turn out that all the works traditionally ascribed to Aristotle were written by Plato as a joke, I would cease to use the name at all.

A name does not fail to denote an individual because he is not the unique possessor of a property, because it is not true of him, but rather because it is not conventionally assigned to him as the expression which denotes him.

Obviously I am not using "connotation" in Frege's sense, but in the sense of Strawson's "background information."
33 This is essentially Wheelwright's interpretation of the imagery of the *Oresteia*, and I think it is the correct one.


37 This case is easily identified as presenting the question of how point of view in fiction is to be construed. (Some less extensive cases fall under Case (A)). W. Booth, *Rhetoric of Fiction* (Chicago: 1961), argues fairly convincingly that a novel is always told from some point of view. This may be true, but the philosophically interesting cases are those in which the narrative might be construed as a long quotation or as a description of a mental event. I agree with Booth in his general claim that what point of view is taken, and even consistency of point of view, is a matter of fairly arbitrary choice, not per se a basis for judging the competence of the author. No such technique is "uniquely appropriate to" or "best for" fiction.

Handling the whole question of point of view in the context of quotation is a little strange. It is useful, I think, for I hope to show that, all aesthetic subtleties aside, there are only two logically relevant categories of such devices. Since we are out to determine the facts of the fictional world, we have to decide whether a mental or linguistic event has been established, or whether the appearance of this is merely a vivid way of adding facts not of this kind.

38 This example is from notes on a paper read by Donald Davidson.

39 In what follows, I will adopt what is essentially Donald Davidson's treatment of such expressions as presented in "On Saying That," re-printed in *Words and Objections*, ed. Davidson and Hintikka (New York: 1969), and elsewhere.


41 We might think of this utterance along the lines of that of a teacher, who goes to the board and says, "Take this example," then writes a sentence token on the board. We might, it has been suggested, say that it has "exhibitive illocutionary force."

42 Of course, if the sentence occurs in a third-person omniscient author narrative, as (3) does, there is no question of the truth of (3) (even if one assigns truth values to the declarative sentences of a work of fiction—as I have argued one should not). The facts are as Tolkien says they are.

43 William Iccan, unpublished manuscript, "Davidson on Saying That."
This ties in remarkably well with my suggested solution to the question of who the illocutionary subject of the questions and exclamations of the work is. One standard move has been to deny altogether that such expressions have illocutionary force. This is unsatisfactory, since we still have to explain the difference between ostensible questions and ostensible statements.
Chapter III

The second task set in chapter I was that of setting up a procedure for assigning truth values to sentences of criticism. In this chapter I will begin to handle this task of outlining such a procedure for a certain small segment of the language of criticism. In chapter IV, I will argue that although our first task in the literary criticism of a work is arriving at this minimal description of the world of the work we have to deal with, it is by no means our only task. But it is an indispensable preliminary to the other kinds of criticism which I will discuss schematically in chapters IV and V. Among these further kinds of criticism are those bearing on the question of the first chapter, the question of the alleged connection between our experience of fictional works and the expansion and clarification of our knowledge of the actual world. In this chapter, I will set out the first stages of the description of a work of fiction.

It has fairly often been argued by philosophers that assigning truth values to sentences of criticism which contain references to fictional entities involves us in making undesirable metaphysical assumptions. I want to show that at least with regard to sentences of criticism, there need be no such problem. No one has ever thought that there was a special problem with assessing the truth values of sentences such as:

(1) Jane Austen wrote Emma.
which contain a reference to a whole work. Works of fiction, whatever their metaphysical status, are part of the furniture of the world.

(On the view expressed in chapter II, (1) simply says that Jane Austen performed a given speech act). The difficulties, it is alleged, begin with the assessment of the truth values of such sentences of criticism as:

(2) Emma Woodhouse was clever.

where (2) is taken to be a critic's claim about a fictional character. Here, it is alleged, we are going beyond the level of the work itself to make a reference to its "content." Of course, if a given sphere exists, then so do its parts. But it is not obviously true that when we speak of Emma Woodhouse, we are speaking of a part of the work in the sense in which a hemisphere is a part of a given sphere. (I will argue when I discuss the language of criticism that this is just what we are doing, but the view is not prima facie obvious).

In section 3.1, I will elaborate on the view of fictional language proposed in chapter II and will take up the questions of the individualization of fictional works and their worldlike constitution. All of this is a preliminary to the enterprise undertaken in section 3.2, the establishment of a procedure for ascribing truth values to a group of basic critical claims. In section 3.2, I will proceed to set up a formally correct recursive truth definition which gives us the results that we know intuitively are correct for a large number of critical claims, those which "match" the sentences of the work and those which are derived from such claims. Because of the way truth values are assigned in this section, particularly to singular and quantified sentences, it
will become clear that we can carry out such an enterprise without encountering any of the difficulties long predicted by philosophers. In section 3.3, I will argue that the formal device of treating fictional entities as expressions which is used for setting up the procedure of section 3.2 is in fact a plausible metaphysical account of such entities, and is preferable to the alternatives. In section 3.4, I will attempt to make plausible the view that a large number of critical claims are in fact claims about the author's speech act. The account of the meaning of such critical claims is by no means complete, but I think that it is correct in its basic outlines.

Chapter III is thus aimed primarily at philosophers who claim that we cannot assign truth values to such sentences of criticism without courting disaster and who worry about the ontological status of fictional entities. The basic claim, that at least some sentences of criticism are true or false of a work on the basis of what is in the book, is one which any critic takes for granted. Of course, the procedure I propose does not provide an answer to the broader philosophical problem of how to assign truth values to statements with subject terms which refer to nothing in the actual world where those terms are not used in a work of fiction to present fictional entities or used by critics to talk about works of fiction. (The broader problem arises again in this context if someone claims to have referred to a "character" which occurs in no extant work of fiction).

Although chapter III is aimed primarily at philosophers, it does have some consequences for criticism. Section 3.2 does give rise to a criterion for judging which critical arguments from a basic description
of a work are deductively valid and for designating some critical arguments as invalid on the basis of their form. But the main importance of chapter III for criticism is negative, as will emerge in chapters IV and V. It provides a good way of delineating those critical claims which are true of a work on the basis of its "intrinsic" features. Chapter III thus furnishes an extremely effective method for assigning truth values to sentences of criticism which are true on the basis of the explicit elements of the work or which are derived deductively from critical sentences which are true on this basis. In chapters IV and V, I will raise the question of how many critical claims can be accounted for in this way. It will be obvious, I think, that much of criticism which is commonly conceded to be legitimate cannot be so accounted for, that we need far more complicated logical techniques and many external assumptions to account for portions of criticism which we would not want to rule illegitimate. At that point, I will attempt to show schematically how such an extension of the procedure sketched in chapter III can be carried out.
If the view presented in chapter II was correct, then there is no problem in saying that the ostensible referring expressions in a novel are not used to refer to anything, but rather to present the characters of a work, and that the ostensible reports of a novel are not used to describe anything, but are used rather to establish the facts of a fictional world. We need not treat the work as the author's description of something which is somehow beyond the work. In addition, I argued that when an English sentence is used in this way, the assignment of truth values is inappropriate. It is inappropriate, not merely in the sense that the assignment of truth values is a matter of indifference, but because implementing such a procedure is a mistake, just as it is a mistake to assign a truth value to 'Simmons, you will go to Bristol' when the sentence is used as an imperative. Briefly recapitulating the arguments for the non-assignment of truth values, the "reports" of a work of fiction have a peculiar feature. If they are treated as descriptions of some fictional universe, they are all true by fiat. This alone should lead us to suspect that they are not reports at all. The author can put anything he wants into his world, provided that he is consistent. It might be objected here that the very requirement of consistency means that the author is constrained by something beyond his creation and that in this case at least, some of the author's sentences may be called false. But to argue in this way is to wrongly identify the basis of the consistency requirement, I think. Consistency is a condition of a set of reports constituting a fictional universe, not because the author is describing something about which he sometimes makes a mistake, but rather because we understand the establishment of
fictional universes on the basis of our understanding of the features of state descriptions of the actual world. Hence we are understandably convinced that something has gone wrong when we are presented with a set of sentences which purport to establish a fictional world, but are inconsistent. But in such a case, some of the ostensible reports are discarded, not because they fail to accord with something beyond the novel, but because they render the set of ostensible reports inconsistent. It is the coherence of the set, not the correspondence of the reports with some further non-linguistic entity, which dictates our procedure. Rather than saying that the author has included one or more false reports in a list of predominantly true ones, it is more natural simply to say that one of the sentences simply is not a fact of the fictional world. The work stands alone. It is the product of a speech act. It consists of expressions, some of them grouped into sentences which we understand because we understand the words and the parallel function of expressions in the language of normal narration. On this view, there simply is no semantics for the fictive use of language, in the sense of a systematic assignment of truth values to the sentences as used by an author to present the facts of a work of fiction. Of course, since the language of fiction is parasitic on the language of normal narration, it is true that insofar as semantic issues have a bearing on what we often take to be syntactic formation rules, e.g., the use of modal and tense operators, they have a bearing on how expressions are used in a work of fiction.

There is no point in claiming that since sentences of fiction are syntactically indistinguishable from those used to state facts about the
actual world, we cannot deny them truth values on the basis of how they are used. For it seems completely reasonable that a declarative sentence used to ask a question, e.g., 'You're going then', or to issue an imperative, e.g., 'Simmons, you will go to Bristol', should be denied a truth value—because of the way they are used. Given that this is correct, there is no reason in principle why we may not deny truth values to sentences used in fiction, provided they have a distinctive use for which this is not reasonable. Insofar as the points raised in chapter II are correct, then, they comprise an argument for denying truth values to sentences used in fictional contexts.²

Sentences such as:

(1) Emma Woodhouse was clever.

however, may also be used by a literary critic to describe a work of fiction. The main thesis of this chapter is that since descriptive sentences of literary criticism are about works of fiction, i.e., about complex linguistic phenomena, they may be assigned truth value on the basis of which expressions do and which do not occur in the literary work, and that such an assignment forms the semantic basis of literary criticism. In this section, I want to deal with the preliminary steps for establishing a semantics for sentences such as (1) and critical claims which are logically derivable from them. We proceed by taking fictional worlds and their entities, at least for the purposes of semantics, as linguistic phenomena and the expressions which are parts of these phenomena, respectively. It follows from this that at least some sentences of criticism may be assigned truth values on the basis of whether they match the novel or do not. Assigning truth values in this
way avoids traditional questions about the metaphysical status of fiction entities and thereby avoids the traditional objections to assigning truth values to sentences of criticism. The present sections are neutral on the questions surrounding the metaphysical status of fictional entities and the meaning of sentences of literary criticism. I will turn to these issues in sections 3.3 and 3.4. The commonsense basis for setting up the semantic base in this way is that any literary critic proceeds in practice in the way I suggest, no matter what his metaphysical views. The world of the work comes into existence with the author's action. After his action, and only after his action, is there discourse about his work. And, at least on the strictly descriptive level, the criterion for the truth of critical claims is what the author says, what facts he establishes. But before we turn to the actual truth definition for basic sentences of criticism, there is one matter which demands attention. Since it is the individual work of literature which will serve as the basis on which truth values are assigned to claims about it, we need to be clear about what counts as an individual work of fiction.

For the purposes of semantics of criticism, I shall rely on the view defended in chapter II that a work of fiction just is (or is the result of) the author's speech act. We cannot accurately report on someone's speech act if we make changes in what he says. (The one exception to this rule is the case in which a work is inconsistent). In that case, we take the work to be a proper subset of the set of sentences written by the author). From this position it follows that there are three criteria for the individuation of works of fiction.
The speech act criterion. The requirement that we stick to the sentences actually used by a single author furnishes us with one of the necessary conditions for a work's being considered a single fictional universe. This requirement is loose enough to allow us to consider the whole *Sherlock Holmes* series a single work (just as we can say that A, who builds a boat at home over a period of several years has performed the single action of building a boat). We can make the requirement loose enough to include the case of genuine joint authorship, i.e., the case in which A and B agree to write a work together. In this case we have a single work which is the result of A and B's speech act. But if there is no such prior agreement, then A's work and B's work are to be treated as two separate works. And a "sequel" written by another author is not part of the same fictional universe (just as B's gratuitous work on the boat is not part of A's action of building the boat). Completions of *The Mystery of Edwin Drood* will remain forever theories about a truncated universe, not additions to the work. On this view, if Dickens and Thackery had just happened to write down identical lists of sentences and had each titled the manuscript 'David Copperfield the Obscure', we would still have two different works, since the works would be the products of the actions of two non-co-operating agents. And, although Pamela appears in *Joseph Andrews*, the *Joseph Andrews* universe is not continuous with the *Pamela* universe (although it might have been, had Richardson written *Joseph Andrews*). This is also the general reason why a character of one fictional universe is not a character of another (except in the case of "nested" universes, where the characters of the narrative-within-the-narrative are fictional characters of the containing
narrative, although not actual characters.

The consistency requirement. We have already seen that consistency of the facts presented is also a necessary condition for a work's being a fictional universe. The critical alternatives when this requirement is not met will be discussed in chapter V. They range from denying that we are dealing with a work of fiction at all to shrinking the work until we get a subset of sentences which is consistent. But we have not been very precise about the exact nature of the consistency requirement. At this point, since we are dealing with the non-truth valued sentences of a fictional work, we still have to define consistency in terms of the logical relationships which would obtain between sentences of a work on the supposition that we were to assign them the value "true." A work is consistent if such a set would be logically consistent. In particular, such a set must be consistent with regard to the predicate calculus. In addition, novels must be consistent with regard to the calculus of identity, although the usual metaphysical interpretation of the calculus of identity, dependent as it is upon the usual criteria of identification and extension, may not work for a novel. (In Farnham's Freehold, for example, one of the characters, coming backwards in time encounters "himself." Here there is no obvious breach with respect to the calculus of identity, though we do have to change our usual apparatus for individuation, since ordinarily we consider different time stages of an individual to be "part" of a single individual). In addition, the modal sentences of a novel must, I think, be consistent with regard to the modal system which expresses the logic of logical necessity. The tense sentences of a novel must be consistent in the
sense that they constitute or determine a model for a tense logic which is consistent and complete. And the epistemic sentences must be consistent with whatever epistemic logic is finally shown to be correct. And if there are propositions such that they must be conceded to be theorems of any acceptable epistemic logic, they must be consistent with them. Furthermore, the employment of semantic predicates in non-intentional contexts must be consistent with the generally acceptable rules for the employment of such predicates. Ordinarily, these last requirements are not particularly important. In most novels, we have very few modal sentences, and we can assume the normal conception of time. But in science fiction novels, for instance, we are sometimes presented with logically possible alternative universes or with non-standard time models. In such cases, the modal and tense logical consistency requirements become important.

Another interesting question is what we should say about so-called "conceptual truths," such as that all bachelors are unmarried, or that every father has either a son or a daughter. Must we require that fictional worlds preserve these truths, that they be "conceptually possible" with regard to the actual world? The answer in general is "no." There is, of course, the requirement that the predicates used in constituting a fictional universe retain their normal narration intensions or be defined in terms of normal narration intensions. But to require that all "conceptual truths" be preserved in a fictional universe is no doubt too strong. I think we could claim that the predicate 'is a man' had retained its meaning or intension, even given that in a work men had received the gift of living forever, although 'All men are
mortal might be (or at least has been claimed to be a "conceptual truth.") Many fictional universes become interesting precisely because they are constructed around the breakdown of one or more truths which we take to be conceptually true in the actual world, or perhaps what often comes to the same thing, because commonplace inductive generalizations about the actual world do not hold in the fictional world. The situation can be summed up as follows. If there are definitional equivalences and what we sometimes describe as antonymy relations, these must remain intact in fictional universes. And if there are combinations of predicates which are logically contradictory, then authors must refrain from combining predicates in these ways. This requirement arises out of the fact that we understand predicates in the language of fiction in terms of their application in normal narration, not from the requirement that fictional universes must be consistent. But certain weaker relations of predicates sometimes described as "conceptual truths" can be strained or broken in fiction. The extent to which this occurs may well be a good way of characterizing the style of a given work of epic literature, ranging from "normal" or "realistic" to "distorted" universes. Certain groups of works may also be causally possible relative to the actual world, i.e., governed by the same causal laws which hold in the actual world. In such works, "normal" or "realistic" works, the standard "conceptual truths" hold. Sometimes the delineation of a set of works which are related in this stronger way to the actual world is interesting, or even necessary for critical discourse.

The constancy requirement. The requirement that a work of fiction be a consistent set of sentences has a further consequence for the
individuation of fictional worlds. We have already seen that the
constant intension of predicates is a necessary condition for a work's constituting a single work. And, although up until now, we have spoken mainly of historical names retaining their normal narration usual connotations in the constitution of a fictional universe, we can, taking a given fictional universe as our base, articulate another condition for a work's being considered a single work: The fictional names must retain their fictional usual connotation throughout the work. Ordinarily, this last condition is rather trivially satisfied by a given work. The fictional usual connotation of a name is set by its use in the novel in which it occurs. Conflicts are usually bridged by a compatibility hypothesis, e.g., a change of heart, a religious experience, etc. But this condition allows us to consider a work as distinct from a parody of the work, even if the parody satisfies the first two conditions, i.e., is logically consistent with the first work and is by the same author.

The individuation of fictional works is not ordinarily a serious problem, and I think the treatment of the question given here handles most of the specific difficulties which may arise. Although we now have criteria for individuating works, it may not yet be completely clear how such fictional worlds are constituted. Therefore I will give at least a provisional list of the kinds of semantically relevant expressions which may occur in a work of fiction. I will periodically speak of fictional "facts," instead of fictional sentences, since this will help to bring out the contrast between what we can claim to know with certainty about a fictional world, the "given" of the fictional world, and what we can claim to know about the actual world in this sense.
We are, of course, given what might be called in analogy to the actual world "atomic facts," where namelike expressions are concatenated with predicate expressions. Alternatively, we can say that we are told that a character is presented as having a certain characteristic.

An example of this is:

My Tomcat, Petronius the Arbiter, and I lived in an old farmhouse in Connecticut. Sometimes we are given such "atomic facts" connected by 'and', 'or', 'if...then', or prefixed by 'not'. We are often given quantified facts, e.g.,

No police were looking for me, nor any husbands, nor any process servers.

And sometimes, we are given more complex forms of quantified facts:

Nonetheless I was looking for the Door into Summer. Most of the ones I had checked lately had been swinging doors.

In the actual world, no doubt, we infer most of the quantified claims we claim to know from particular instances. But in a fictional world, such quantified sentences need not be inferred. They are incontrovertible facts of the fictional world. This case is to be distinguished from a matter that we will take up later, critical inferences of quantified sentences from the atomic sentences of such a work.

Sometimes we are given what would count as inductive generalizations or law statements in a description of the actual world, e.g.,

Cats have no sense of humor, they have terribly inflated egos, and they are very touchy.

Again, no matter what the epistemological status of such sentences in the actual world, they stand as incontrovertible facts of the fictional world. Cases where we are explicitly given such laws as facts must be
distinguished carefully from those which are included in a critical
description of a work on the basis of particular or quantified facts.
Sometimes we are given as a fact the constitutive logic of a universe.
In *The Door Into Summer*, for example, the narrator explicitly rejects
the branching time hypothesis. Sometimes, too, we are given facts of a
universe by an expression involving the verb 'knows' or forms of 'said
truly'. Occasionally, we are given what is presented as a work of
fiction within a work of fiction, sometimes only by title. We may ex­
pect to have to provide for such "nested fictional universes" in our
semantics of criticism.

No doubt there are other categories of given facts which might
demand separate listing. The above facts will all fall under one clause
of the semantics of criticism. If a sentence $s$ occurs in a consistent
set of sentences constituting a fictional work, then a sentence of
criticism corresponding to it will appear in a critical description of
the world. It is useful, however, to list the different categories of
facts. This is mainly because the extremely large variety of sentences
included in the class of given sentences might otherwise be overlooked,
and is extremely important when we come to the business of inferred
critical descriptive claims. Certainly the single most important general
feature brought out by such a listing is that we are very often given
as fictional facts the sort of fact which we could only infer about
the actual universe or propose as an explanatory theory about it.
I have argued that we should assign truth values to a certain basic group of sentences of descriptive criticism ultimately on the basis of how they match up with the expressions which occur in the work being talked about. I will now proceed to set up a procedure for assigning truth values to sentences of descriptive criticism on this basis, to set up a semantics of criticism. I am not here considering the question of the meaning of critical claims, only the method for assigning truth values. It is the way in which any critic must proceed practically, no matter what his view on the nature of fictional entities and the meaning of his claims.

There are a certain number of features that we would like our semantics of criticism to have. In fact, we can treat these desirable features as a set of adequacy conditions for a semantics of criticism.

(i) We take certain sentences, e.g., 'Sherlock Holmes was a detective' to be true, fictionally speaking; and certain sentences, e.g., 'Sherlock Holmes was eaten by the Hound of the Baskervilles' are taken to be false in the same sense. Our semantics should allow us to make such distinctions. This requirement rules out certain ways of assigning truth values to such sentences, at least for the purpose of a semantics of literary criticism. It rules out the view that all such sentences are false (or, for that matter, that they are all true) of fiction on the grounds that they contain a reference to a non-existent entity. And it rules out the view that they are all non-truth-valued because the existential presupposition attached to the subject expression is not fulfilled.¹⁰

In addition, there are certain conditions which we want the
assignment of truth values to fulfill. Let us say that if a sentence $p$ occurs in the narrative of a work of fiction, there may occur in a description of the fictional world a sentence $p^*$. Ordinarily, $p^*$ is produced by the critic by producing another token of the same sentence type of which the sentence in the work of fiction is also a token. Sometimes minor adjustments will be called for, however. $p^*$ for 'I went up on deck' in Moby Dick will not be 'I went up on deck', but rather 'Ishmael went up on deck'. I will say that the sentence of descriptive criticism formed in this way is the sentence corresponding to the sentence of fiction. The list of such corresponding sentences of criticism for a given work I will call a "basic critical state description" of the work. The semantics we choose should assign the value 'true' relative to the work to all the sentences of the basic critical state description.

(ii) A second requirement relates to those novels which contain namelike expressions which are also names in a description of the actual world. Take a novel in which, for example, 'Bastille' occurs. The novel also contains, say, 'The Bastille was stormed' and 'The Bastille was a prison', etc. Let us say in a preliminary way that such a novel and a total state description of the actual world have character overlap. It is clear that there need not be such overlap and that if there is, it need not be complete. In the case of The Fellowship of the Ring, there is no overlap. None of the namelike expressions of The Fellowship of the Ring occurs in a complete state description of the actual world and vice versa. But A Tale of Two Cities contains 'Bastille' and also 'Sidney Carton'. Here the overlap is partial. It is fairly clear from
chapter II and section 3.1 that 'Bastille' when it occurs in a
critical description of the novel does not name the same thing as
it names when it occurs in a state description of the actual world. As
a matter of fact, 'Bastille' as it occurs in a description of one
fictional work and 'Bastille' occurring in a description of another work
will not name the same thing. This does not mean, of course, that there
is no connection whatsoever between 'Bastille' as it occurs in a state
description of the actual world and 'Bastille' as used in a description
of a novel. Sometimes 'Napoleon' occurs in a novel, although we are
never given many sentences in which 'Napoleon' occurs. In such cases,
it is often clear that we are expected to fill in or construct those
sentences on the basis of the information we have about the historical
Napoleon. This we explained in 2.3 by saying that a historical name
retains its usual connotation when the expression is used to present a
fictional character. But note that such an expression retains its
usual connotation; it is not an orthographic accident that the same
expression is used to name the fictional entity and the actual prison.
If it were a mere orthographic accident, we could merely deplore the
ambiguity and regret that the author did not use 'George' instead of
'Bastille'. But this would deprive us of any reasonable basis for
filling in our information about the real Bastille into the novel. The
semantic account we adopt should enable us to mark clearly the distinction
between sentences containing 'Bastille' which are fictionally true and
those which are true of the actual world without our having to adopt
some trick for distinguishing the expression in its different uses.

(iii) A related requirement has to do with the relation between
descriptions of "fictional facts" and descriptions of the actual world. The "facts" of a fictional world are in general arbitrary in the sense that the sentences of a fictional world are the ones the author establishes. So long as he is consistent, he cannot "make a mistake." Since the "facts" of fiction are arbitrary in this way, set down by fiat, as it were, it is clearly wrong to hold that critical claims about a fictional world can challenge or support claims about the actual world, or that the reverse procedure is legitimate in any straightforwardly semantic sense. "I read it in a novel" is not per se good evidence for a claim about the actual world; and "That's the way things are in the world" is not per se good evidence for a claim about a fictional work. An adequate semantics should block direct inferences from critical claims true on the basis of the fictional world to claims about the actual world and vice versa. A somewhat less obvious requirement is that the semantic account should block inferences between fictional worlds. It should display the fact that a given sentence of descriptive criticism is true evaluated in the actual world because it is true of a fictional work. Clearly, this requirement is closely related to requirement (ii). It is because a fictional entity of some particular world, i.e., a part of someone's speech act, and because separate speech acts establish separate worlds that we cannot in general argue from what is true of one such world to what is true of another.

(iv) We know in advance what different sorts of critical claims could arise, and have some idea of what semantic solutions would be acceptable in each case. There are such statements as
(1) Sidney Carton helped Charley Darnay.
(2) Madame Defarge helped Charles Darnay.

We want the semantic account to be able to make a distinction between (1) and (2). We want to be able to "accept" (1) and "reject" (2), for (1) corresponds to the "facts" of the fictional world, and (2) misrepresents them. (1) is not about a person of the actual world of course; but we know perfectly well that it is true, just as (2) is false of the work in question. Moreover (1) is true of the fictional work as an object of the actual world. Our semantic account ought to express the fact that (1) is true of a fictional world and allow us to evaluate (1) and (2) in the actual world. We also have such critical statements as:

(3) Sidney Carton helped Anna Karenina.

(3) is different from (1) and (2) in that the "entities" referred to are entities of different fictional worlds. Our semantic account ought to display this peculiarity, and it ought to render (3) "unacceptable," but not necessarily in the way that (2) is.

(4) Lucy Manette had the green and gold gown that Madame Defarge knit.

presents still another kind of problem. Intuitively speaking, (4) goes beyond the sentences of the novel, and it contains a reference to an entity, i.e., an expression, 'the green and gold gown which Madame Defarge knit', which is not, so far as we know, in fiction at all. In this sense, (4) is different from (3), although like (3), it is "unacceptable."

Finally, there are such critical sentences as:
(5) Upper Swandom Lane was a street of London and was clean on Dec. 1, 1821.

'Upper Swandom Lane' is taken here to refer to a fictional entity. (5) we should like to pronounce "acceptable," with regard to the fictional work since by hypothesis it corresponds to the work in question. But although in (5), 'London' refers to a fictional entity, i.e., an expression, it could also be used in a description of the actual world. This is, of course, another facet of the overlap problem mentioned earlier. We do not want (5), if it is allowed to be true, to contribute to conclusions which conflict with our knowledge of historical fact. (This could happen, for instance, if there were very many fictional streets and we had conflicting information about only, say, two streets of London. Thus the assignment of the value 'true' to (5) should make it clear that it is "true-of-a-work-of-fiction," although it might well be false or non-truth valued with regard to the actual world.

(v) There are critical practices which are not straightforwardly a matter of getting a set of corresponding sentences for a work. Some of these are a matter of claiming that since \( D \) is the set of corresponding sentences for a fictional work and since \( D \) entails \( p \), \( p \) is in a derived sense true of that world. Here it becomes important to have a systematic semantics for criticism which furnishes a logic of critical argument. It is relatively uncommon for a competent critic to err in giving a basic description of a fictional world. But it is not unusual for such a critic to claim that a universally quantified sentence is true in the derived sense of a work, although if we check carefully, we find that his derived critical claim is incompatible with the basic description of
the world. We will therefore want a semantic account of criticism from which we can develop a notion of entailment such that we can tell which of the allegedly derived critical claims can in fact be logically derived from the basic state description.

(vi) As I have indicated from time to time, there are critical claims about works which are not part of the basic state description and are not logically derived from it, although they are alleged to be descriptive of the work. Some of these concern shrinking of inconsistent universes or expansion of a universe which is written from a limited point of view, e.g., What Maisie Knew. Some involve arguing that a work is included in a set of works which are possible relative to the actual world in some stronger sense than logical possibility, so that we can propose inductive generalizations about a work or complex interpretations of its elements. Sometimes we want to compare characters from different works. The semantics we choose should be such that we can extend it to allow us to assign truth values to more complex critical claims and to assess the logical structure and relative strength of arguments used to support such claims. Actually, we will not be dealing with such claims and arguments until chapters IV and V, but the fact that we will have to do so should be kept in mind.

This may not be a complete set of the requirements we should set for a semantics of criticism; but all those listed here seem to be necessary. An account which satisfies these requirements should be able to deal with future complications which arise.

The fact that fictional works have been defined as consistent with the predicate calculus with identity suggests that we might approach
the semantics of literary criticism from the point of view of possible world semantics. Corresponding to the usual possible world semantics notion of truth in some possible world will be the notion of being true of some fictional world. But the semantics for literary criticism will differ in interpretation from standard possible world systems in various ways. A fictional world is an entity of the actual world, i.e., a list of sentences. Where traditional model theory generally speaks of an ordered triple \(<G,K,R>\), where \(G\) is the actual world, \(K\) the set of all possible worlds, and \(R\) a binary relationship specified in terms of its logical properties, we shall take \(K\) to be the set of all actual fictional works related to the actual world by the relation of \(R\) being an entity of the actual world. For the sake of convenience, I shall take \(G\) to be, not the actual world, but a global description of the actual world. The domain associated with each fictional world \(H \in K\) will be "nominalistic," insofar as it consists of the name-like singular expressions which occur in \(H\).

Let us define as our model structure an ordered triple \(<G,K,R>\), where \(K\) is non-empty, \(G\) an element of \(K\), and \(R\) a binary relation on \(K\). Intuitively, \(K\) is the set composed of \(G\) (a list of sentences which comprise a global state description of the actual world) and all the fictional worlds accessible from the actual world by \(R\). A fictional world \(H\) is said to be accessible from \(G\) by \(R\) if the fictional world, i.e., the list of sentences, is an entity, a work of fiction, in the actual world. Let \(\emptyset\) be a valuation function which assigns the value \(T\) or \(F\) or \(X\) to all atomic sentences of criticism and to open and quantified sentences of criticism.
Of course, we want every sentence $S$ of criticism corresponding to a sentence of $H$ to come out true of $H$. The corresponding sentences for $H$ are not, as previously noted, obtained merely by copying out the sentences of the work. When the narrator says, for example, 'I lived in an old farmhouse', the corresponding sentence of criticism, given that the narrator's name is 'D. B. Davis', is 'D. B. Davis lived in an old farmhouse'. When one has gone through and replaced all personal pronouns which do not occur within quotation marks by the appropriate namelike expression, I will say that $H$ has been filled. Henceforth, when I speak of a given $H$, I shall consider that $H$ has been filled. We are now in a position to state the first clause of the truth definition for critical semantics:

1. For any sentence of criticism $S^*$, if the corresponding sentence $S$ occurs in a given fictional universe $H_1$, then for any $H_1$ and $H_j \in K$

$$\phi(S^*, H_1) = T,$$  
and for any universe $H_1 \in K$, $H_1 \neq H_j$,

$$\phi(S^*, H_j) = T_1.$$

Thus an $S^*$ will be true of the world for which it is a corresponding sentence, and will be fictionally true for any universe from which the world is accessible. (I) gives us the procedure for formulating a basic state description of a fictional world $H$. Any critical sentence in a basic state description of a fictional world will be true of that world.

We now want to begin to specify the truth conditions of the set of sentences, which, though they are not in the basic state description of a fictional world $H$, are members of a deductively derived state description of $H$. We first give a truth definition for the proposition-forming operators.
II. Take two sentences A and B such that $\psi(A, H)$ and $\psi(B, H)$ have been defined for every $H \in K$. Then:

1. $\psi(A \land B, H) = T$ if $\psi(A, H) = \psi(B, H) = T$;
   $\psi(A \land B, H) = F$ if $\psi(A, H) or \psi(B, H) = F$;
   $\psi(A \land B, H) = X$ otherwise.

2. $\psi(-A, H) = T$ if $\psi(A, H) = F$;
   $\psi(-A, H) = F$ if $\psi(-A, H) = T$;
   $\psi(-A, H) = X$ otherwise.

"L" presents some problems. Because of the way $R$ is specified, it will not do to define $\psi(A, H) = T$ if $\psi(A, H') = T$ for every $H' \in K$, such that $H' RH$. Intuitively, what formulas 'L(A)' are true in a world $H$ will depend on the given interpretation of the L-operator and upon what sentences are true in $H$ under the given interpretation of the L-operator. If L is interpreted as logical necessity, for example, then we can assume that for a given world 'L(A)' is true if it is assigned the value T by clause I or if it can be derived from the modal and non-modal truths of the basic state description by inferences valid for the logic of logical necessity. This is quite likely to be S5. So for the L-operator taken as logical necessity, let $\Gamma$ be the set of sentences

\[ \{ B : \psi(B, H) = T \text{ on clauses I and II} \} \]

Then:

3. $\psi(L(A), H) = T$ iff $\Gamma \models_{S5} L(A)$ and $\psi(L(A), H) = T_1$ iff $\psi(L(A), H_1) = T$ for $H_1 \in K, H_1 RH$, and $H_1 \neq H$;
   $\psi(L(A), H) = F$ if $\Gamma \models_{S5} -(L(A)) and \psi(L(A), H) = F_1$ iff $\psi(L(A), H_1) = F$ for some $H_1 \in K, H_1 RH, H_1 \neq H$;
   $\psi(L(A), H) = X$ otherwise.

The versions of clause (II 3) for temporal, epistemic, and imperative
interpretations of the L-operator are constructed in an analogous way. One obvious result of clause III is that inferences from 'S knows that p' (itself in a nonintentional context) to 'p' are allowed, since this inference must, presumably, be a theorem of any adequate epistemic logic (provided "know" is used in its strict sense).

Similarly, inferences from "S lied when he said that p" (itself in a nonintentional context) to "not-p" are legitimate. Any sentence of criticism assigned the value T on the basis of having been derived by successive applications of (II 1-3) will be said to be a member of a deductively derived state description of H. All members of such a description of H are true of the fictional world, i.e., they correspond to sentences of the fictional work or are deductively derivable from ones which do.

We also want to be able to deal with quantified critical claims which are not obtained directly from clause (I) \(^{15}\). In setting up a quantificational model, we associate with each world H \(\in K\) a "domain" consisting of all the "referring" expressions in H. Let \(\psi\) be the function which takes every H \(\in K\) onto the set \(\psi(H)\) of all the "referring expressions" in H. Let \(U = \bigcup_{H \in K} \psi(H)\). U is then the set of all the "referring expressions" in our whole library of extant fictional worlds plus the genuine referring expressions which refer to entities of the actual world. Let \(\langle G,K,R,\rangle\) be as before, and let \(\phi\), in addition to assigning values to all atomic formulae, assign to each predicate \(P^n\) an "extension" in every world H \(\in K\), such that \(\phi(P^n,H) \subseteq (\psi(H))^n\). Intuitively, \(\phi\) picks out for each world H and for an n-place predicate \(P^n\) the "names" or n-tuples of namelike expressions of the world H, which
when filled into \( P(x_1, \ldots, x_n) \) give a sentence true of the world. In setting up a quantificational model for criticism, we must remember that fictional works, intuitively speaking, have disjoint domains. Hence, the restrictions on the quantified sentences which will be true will be somewhat more stringent than they generally are for modal semantics systems. 16

III. For any assignment of elements \( a_1, \ldots, a_n \in U \) to the free variables \( x_1, \ldots, x_n \) of \( P^n \):

1. \( \phi(P(x_1, \ldots, x_n), H) = T \) if \( \langle a_1, \ldots, a_n \rangle \in \phi(P^n, H) \)
2. \( \phi(P(x_1, \ldots, x_n), H) = T_1 \) if \( \langle a_1, \ldots, a_n \rangle \notin \phi(P^n, H_1) \)
   \( H_1 \neq H, H_1 \in K, \) and \( H_1 RH \).
3. \( \phi(P(x_1, \ldots, x_n), H) = F_1 \) if \( \langle a_1, \ldots, a_n \rangle \notin \phi(P^n, H_1) \),
   where \( a_1, \ldots, a_n \in \psi(H_1), H_1 \notin H, H_1 \in K, H_1 RH \).
4. \( \phi(P(x_1, \ldots, x_n), H) = F \) if \( \langle a_1, \ldots, a_n \rangle \notin \phi(P^n, H) \),
   \( a_1, \ldots, a_n \in \psi(H) \cup \psi(H_1) \) for \( H_1 RH, H_1 \notin H, H_1 \in K \).
5. \( \phi((x)A(x, y_1, \ldots, y_n), H) = T \) if for every assignment of
   \( a, b_1, \ldots, b_n \in \psi(H) \) to \( x, y_1, \ldots, y_n \), \( \phi(A((x, y_1, \ldots, y_n), H) = T) \).
6. \( \phi((x)A(x, y_1, \ldots, y_n), H) = T_1 \) if for every assignment of
   \( a, b_1, \ldots, b_n \in \psi(H_1) \) to \( x, y_1, \ldots, y_n \), \( H_1 RH, H_1 \in K, \)
   \( H_1 \neq H, \phi(A(x, y_1, \ldots, y_n), H_1) = T) \).
7. \( \phi((x)A(x, y_1, \ldots, y_n), H) = F_1, \) if for some assignment of
   \( a, b_1, \ldots, b_n \in \psi(H_1) \) to \( x, y_1, \ldots, y_n \), \( H_1 RH, H_1 \neq H, H_1 \in K, \)
   \( \phi(A(x, y_1, \ldots, y_n), H_1) = F) \).
8. \( \phi((x)A(x, y_1, \ldots, y_n), H) = F \) if for some assignment of
   \( a, b_1, \ldots, b_n \in \psi(H) \) to \( x, y_1, \ldots, y_n \), \( \phi(A(x, y_1, \ldots, y_n), H) = F) \).
Let it be understood that no inference from premises which are true relative to a fictional world (assigned a subscripted truth value) lead deductively or inductively to a conclusion relative to some other world, and that there is to be no mixing of subscripts in quantification. Why this is so should be clear from the semantics. Sentences are unqualifiedly true relative to a specified domain, and the domains in question are disjoint.

Note, however, that in clause III, we have considered only cases where \( a_1, \ldots, a_n \) are elements of \( U \), and specifically (III,4) only those cases where \( a_1, \ldots, a_n \) are members of domains accessible from the world of evaluation. Thus, we note that clause (III) does not provide an evaluation of such examples as:

\[
(4) \text{Lucy Manette had the green and gold gown knitted by Madame Defarge.}
\]

where 'the green and gold gown knitted by Madame Defarge' is not in \( U \), or for such sentences as 'Sidney Carton helped Anna Karenina', evaluated in either of the fictional worlds in question. We can remedy this problem by adding another clause.

IV. \( \emptyset(P(x_1, \ldots, x_n), H) = X \) for an assignment \( \langle a_1, \ldots, a_n \rangle \)

to \( x_1, \ldots, x_n \) if for some \( a_j \), \( 1 \leq j \leq n \), \( a_j \notin \emptyset (H) \) and

for all \( H_1 \in K, H_1 \neq H, \) and \( H_1 \models H, a_j \notin \emptyset(H_1) \).

\( X \) we interpret as a truth-value gap. By IV, (4) will be truth-valueless evaluated in \( G \), and a fortiori in any other universe of \( K \). And 'Sidney Carton helped Anna Karenina' will be truth-valueless when evaluated in The Tale of Two Cities (\( C \)), and in Anna Karenina, since Anna Karenina is not in the domain of any universe accessible from \( C \). But the sentence in question will be false when evaluated in \( G \) by clause (III,4), since
both Anna Karenina and Sidney Carton are in domains of worlds accessible from $G$.

Clause IV does raise one interesting question. Sometimes we are presented with a fictional world within a fictional world or with a reference to a fictional work about which we are given no other information, for example the references to the novels of Bergotte in Remembrances of Things Past. Such worlds I will call "indirectly accessible from $G$"; and I will say that $K$ includes all fictional works which are directly or indirectly accessible from $G$.

These four clauses, I think, give us a way of specifying completely the deductively derived critical description of a fictional universe. All the sentences which are members of a deductively derived description of a fictional world $H$ will be said to be true of $H$. It may be objected that clauses I–IV give us a way of constructing a good many critical sentences which are "irrelevant to critical purposes." This is, of course, true. On the basis of clause II, for example, and a definition of 'or', given one true disjunct, we can go on constructing ever longer disjunctions, all of which will be true in the derived sense of the work in question. But this is as it should be. Such an activity would be singularly useless, but all the items of the description so formed would be, strictly speaking, true of the fictional work.

The semantic basis set down in I–IV seems to fulfill the requirements set down in the first part of this section. We now have a semantic basis which allows us to distinguish sentences which are true of a fictional world from those which are false (or non-truth-valued). And clause I guarantees that every sentence of criticism which matches or corresponds to a sentence of a work will be true of that work (requirement
(1)). The account allows us to make a distinction between novel sentences containing 'Bastille' and sentences describing the actual world in which 'Bastille' occurs without tinkering with different senses of 'Bastille' (requirement (ii)). We now have blocked inferences from a novel to the actual world and vice versa, and inferences between novels blocked, at least on the simple semantic level (requirement (iii)). The truth values assigned to critical sentences follow the lines of our earlier intuitions about such sentences; and in addition, we now have a distinction between different kinds of unacceptability. Call the world of A Tale of Two Cities C. And the world of The Man With the Twisted Lip F, then we have

\[
\begin{align*}
(1a) & \quad \phi(Sidney Carton helped Charles Darnay, G) = T_c \\
(1b) & \quad \phi(Sidney Carton helped Charles Darnay, C) = T \\
(2a) & \quad \phi(Madame Defarge helped Charles Darnay, C) = F \\
(2b) & \quad \phi(Madame Defarge helped Charles Darnay, G) = F_c \\
(3a) & \quad \phi(Sidney Carton helped Anna Karenina, C) = X \\
(3b) & \quad \phi(Sidney Carton helped Anna Karenina, G) = F \\
(4a) & \quad \phi(Lucy Manette had the green and gold gown that Madame Defarge knit, C) = X \\
(4b) & \quad \phi(Lucy Manette had the green and gold gown that Madame Defarge knit, G) = X \\
(5a) & \quad \phi(Upper Swandom Road was a street of London, G) = T_f \\
(5b) & \quad \phi(Upper Swandom Road was a street of London, F) = T \\
\end{align*}
\]

(requirement (iv))

The semantic rules for constructing a deductively derived state description allow us to assign the value \( T \) to critical sentences which are derived logically from the basic state description. Finally, the account
gives rise to a concept of validity for critical inference. If we let each $\mathcal{H} \in K$ determine a set of logically possible worlds, then a sentence of criticism $A$ is fictionally valid if and only if $A$ is true for every valuation on every $\mathcal{H}$-model structure. We have already stipulated that every fictional world $\mathcal{H}$ is consistent with regard to the predicate calculus with identity. And the logic of logical necessity is $S5$. Thus critical arguments about what is true of a work on semantic grounds must be valid in quantified $S5$. In addition, once the tense structure, epistemic structure, etc., of a given world is determined, critical arguments about what tense or epistemic sentences are true of the work must be reached from the deductively derived state description of the work by inferences valid in the appropriate logic (requirement v). Henceforth, I shall say that a critical claim which can be justified as being a member of such a deductively derived description of a fictional work can be justified by factors internal to the work. In the next section, we will consider claims which cannot be so justified without further assumptions which are not logical in nature, or which cannot be so justified at all.

So far we have been carrying on a logical enterprise which is metaphysically neutral as to the status of fictional entities and the meaning of critical claims. The critic, even if he is a Platonist, will assign truth values to basic critical claims on the basis of what expressions and what ordering of expressions occur in a work (and his knowledge of English syntax $T_1$). For the assignment of truth values we have treated fictional entities, the domains of the elements of $\mathcal{H}$, as if they were expressions. I see no reason why this is not, in addition
to being a useful expedient for formal purposes, a perfectly adequate metaphysical position. In addition, I see no reason why we cannot let the truth conditions for critical sentences determine their actual meaning, saying simply that the critic who says that 'Fa' is true of a given work just is saying that in the novel 'F' is concatenated with 'a' or follows from critical claims arrived at in this way. This means that the critic in talking about the characters and events of works of fiction is in fact talking about parts of authors' speech acts. Both of these views, perhaps, sound counterintuitive. In the next two sections, I will attempt to dispel some of this counterintuitive air. In section 3.3, I will argue that this is by far the most plausible treatment of fictional entities and worlds. In section 3.4, I will argue that the view that the critic is talking about speech acts and their parts should not, when suitably explicated and qualified, do violence to the critics' own understanding of what he is doing. I am not at all sure that it is necessary to bring up either issue at this point. But both raise questions which have long dogged philosophers, and they are worth considering for this reason alone.
It seems important to maintain that in some sense there are fictional entities. Critics who talk about the characters are talking about something, it seems. In this section, I want to defend the view that fictional entities may be treated simply as parts of a speech act or of the complex product of a speech act in which they occur, i.e., in works of literature, that the character Emma just is a collection of expressions, e.g., "Emma", 'Miss Woodhouse', etc., used in a given way by Jane Austen and apprehended by her readers. In view of the fact that there are many copies of Emma and that 'Emma' occurs many times in each of them, I suppose we should more correctly talk of "expression types," for we want to say, I think, that readers who read different copies of a work nonetheless apprehend the same character, the one Jane Austen presents. If this view is correct, then the result is that we can claim that fictional characters exist, yet not claim that they are shadowy entities. A fictional person is not a person, and a fictional chair is not a chair, for fictional persons and chairs are literally parts of speech acts or their products. But as such, they are not shadowy, for they exist in a perfectly unexceptionable manner. On this view, the relation of the character Emma to the novel Emma is simply the relation of part to whole. The whole is the product of the author's speech act. (An alternative treatment would be to take the work as a whole as the author's actual speech act, rather than as the product of that act. This would make little difference in the most important features of the account. We would still, for example, have the restriction which prevents adding sentences to the work which are not those of the author. But since the author of a work is quite often
dead and gone, or at least otherwise occupied, when the work is read, this alternative is perhaps a little counterintuitive. And it might be argued that if we take the work to be the author's action, the question of the assignment of truth values is settled out of hand, since we do not assign truth values to actions. I do not consider this a good objection, but I am willing to define the work of literature so as to avoid it). The use of 'Emma' is merely a use of an expression which is part of the whole speech act. It is a non-referring use of an expression to present a character. The critic's talk of Emma refers to this part of the speech act or of its product. Performing and apprehending fictional speech acts is complicated business; and doubtless for both writer and reader, the work and the individual expressions have more or less vivid associated imagery. But there is, nonetheless, no "shadowy entity" or "imaginary entity" which is non-linguistic, and to which the discussion of the character (and presumably Jane Austen's use of the expression as well) refers. This view is in fact rather plausible, I think, despite its initial air of robbing us of some of our favorite entities. None of us, unless he is very much mislead about the status of a work of fiction, expects to run into Emma Woodhouse on the street or thinks that anyone ever did. Emma Woodhouse is a character of a novel, and to be a character of a novel is merely to be an expression type, although we typically associate certain imagery with a fictional character. And none of us holds, I think, that Emma Woodhouse has the quality of being brown haired in the sense that that quality daily meets the eye as sensibly apprehensible. Only real persons have that sort of quality. When we say that Emma Woodhouse had brown hair, we certainly
do not mean it in this sense. Of course, we would know a bad illustration or a bad bit of casting in a film of the novel if we saw one. That is because we know the meanings of the words used to "describe" the character, i.e., the expressions associated with 'Emma' by the fictional predication-analogue, fact-establishing. (By the same token I might describe what I am perfectly well aware is a hallucination as it it were a pink elephant with four feet. Just as I would not care to admit in that case that I was describing nothing at all, I should not care to claim that fictional entities were nothing at all. But one certainly does not hold that what one knows to be a hallucination "has four feet" in the literal sense and probably not that it "is pink" in the full usual sense). This distinction is reflected in the fact that, although critics sometimes talk as if Emma were a person with personal qualities, they also sometimes use locutions which are not of this kind. They might, for instance say, "Emma is the most vivid character Jane Austen ever created." Here I think that they are definitely talking in a way that indicates that Emma is a part of the author's speech act or the product of it. And this ontological distinction made by competent readers and critics should be taken into account in our metaphysics.

None of the alternatives to the view I am suggesting is very desirable. If we do not allow that the author's use of 'Emma' is an expression which is not used as a name and to which critics refer, then I think we have no choice but to say that 'Emma' is used both by the author and by the critic as a (perhaps unsuccessful) referring expression or name and that in both uses, the expression refers to or purports to refer to the same thing. Having said this, we may legitimately be asked
what it is that the author and the critic refer to when they use 'Emma'. It is clearly wrong, I think, to say that this object is the author's action or a part of that action or its product. We thus find ourselves committed to something beyond the linguistic phenomenon itself, something to which 'Emma' as used by the critic and by the author refers. There seem to be two plausible kinds of alternative accounts of this commitment. One has a very Meinongian ring; we can posit a realm of semi-existent or subsistent beings and events, or perhaps possible beings and events. The other alternative is to spell out the nature of fictional worlds and entities in terms of the mental states or objects, either of the author or of the readers.

The Meinongian solution is the least desirable. It is fairly clear that we do not want to admit a host of mind-independent subsistent entities to the furniture of the world if we can help it. My own solution shows, I think, that we can help it—if we are willing to give up the assumption that leads to this solution, i.e., that Jane Austen's use of 'Emma' is not a use of a referring expression, in particular not a use of a name. The other alternative seems prima facie more desirable. Most of us are ready to admit that at least mental states of the author (or brain states, if one likes) and the readers exist. But both versions of the second alternative involve us in major theoretical difficulties.

If, on the one hand, we attempt to say that fictional entities are mental states or mental objects of the reader, then we are immediately faced with the problem that there will be as many numerically (and perhaps qualitatively different) Emmas as there are readers of Emma. This is implausible in fact and undesirable theoretically. No doubt each reader
associates different characteristics with Emma Woodhouse, or at least emphasizes different characteristics. But critical discourse would bog down completely if we could not assume that careful readers were talking about the same thing in the strict sense. But what if, on the other hand, we say that 'Emma Woodhouse' refers to the mental state or object of the author. We can make some sense, I suppose, of this suggestion. We can see the author as dreaming up a complex image, then describing it for his readers. But in this case, we are definitely going to have some trouble individuating this sort of entity, and there will be some reader descriptions of Emma which we will be unable to evaluate without writing the author for further information. He may, after all, not have described his mental image fully in the novel. It is not clear that we want the author to be able to "complete" his work in this way—or "amend" it, should he discover that he should have said 'auburn', not 'brown'. These alternatives turn out to be very unsatisfactory in dealing with criticism. And it seems that the bit of plausibility they do have is captured equally well by conceding that the linguistic performance is an expression of or may be caused by the author's mental state, but claiming nonetheless that the fictional entity is a part of the author's speech act or of the product of the speech act. If we keep in mind the fact that a full psychological account of the author's state while writing the work and the reader's state while reading would no doubt show that both the author and the reader connect with the character a certain amount of associated imagery, then treating the fictional entity as the linguistic phenomenon seems to capture all that is plausible in the other suggested alternative.
The associated imagery would account for the often noted "real" quality of fictional characters. And taking the fictional entity as a linguistic entity avoids most of the theoretical problems posed by other alternative accounts. We need not include the associated imagery in our analysis of the semantics of criticism. For these purposes, it is what the author says and what the audience reads that counts.

The point, then, is not that by assuming that there is something beyond the linguistic phenomenon which is the fictional entity, we will annoy those whose metaphysical tastes run to "desert landscapes." These purists will no doubt be sufficiently irked by the fact that my solution commits me to expression types. The more important point is that treating fictional entities as something above and beyond the linguistic phenomenon seems to offer no theoretical advantage, and in some cases actually seems to generate problems that the account cannot solve satisfactorily.

Given that my solution is a plausible analysis of the metaphysical status of fictional entities, it supports the claim that there is no semantics of fictional language in the sense of a systematic assignment of truth values to sentences of fiction—for the simple reason that the sentences as used in a work of fiction containing an ostensible referring expression which is the "name" of a fictional entity do not actually contain names, and hence are not true or false of anything. It will follow from this claim that descriptive criticism strictly speaking must be construed as talk about linguistic phenomena, i.e., speech acts or their products. I will argue for the plausibility of this view in
section 3.4. This is a stronger claim than that advanced in section 3.2. There it was claimed that the basis of assigning truth values to sentences of criticism is what the author says, the expressions he uses and the order they occur in. The claim of section 3.4 is a thesis about the actual meaning of critical discourse.
In this section, I will defend the claim that it is plausible
to construe critics in their descriptions of fictional entities as
talking about linguistic phenomena, about those parts of speech acts
or their products I have called "expression types." Initially this
has a slightly counterintuitive ring. Very simply, most critics do
not seem to take themselves to be talking about linguistic phenomena
when they talk about the characters and incidents of a work of fiction.
And we are inclined to say that if the majority of speakers in a given
context do not take themselves to be doing something, then most likely
they are not doing it. Therefore any philosophical analysis which says
that they are is suspect. But in this case, as I will attempt to show,
there is a clear sense in which a critic who raises such an objection
quite literally does not know what he is talking about, i.e., he does
not know what it is that he is talking about. He has failed to notice
that his descriptions are peculiar in a significant and rather obvious
way. A good deal of the activity of such critics is legitimate, and a
good deal of what they say is true. Of course, the critic writes in
his own language, and he often writes as if fictional characters were
people, stood in actual relations with one another, etc. But if the
view defended in section 3-3 is correct, the literary critic who writes
about the characters of the novel is not writing about persons and
events, though he is writing about something which exists, i.e., speech
acts or their results and their parts. This view gains plausibility if
we note that such phenomena are widespread throughout criticism of
representational art. The art critic who writes about the figures of
a painting and their relations is writing about something which exists
in a perfectly unexceptionable sense, namely the colored shapes and
their arrangement. And when he says, speaking of "The Flight into
Egypt" that "Joseph is behind Mary," we understand what he means and
even concede that it is true. But strictly speaking, the predicate
'is behind' applies properly to three-dimensional situations. It is
certainly not literally true of the figures of the normal two-dimensional
painting, and it is even a bit odd to apply such predicates to two-
dimensional situations—perhaps odd enough that some would consider it
a category mistake. What is literally true in this case is that the
Joseph-figure is smaller than the Mary-figure and is located upwards
and to the left of it. Thus what the critic says is literally false;
but we are all prepared to admit that those familiar with the viewing
conventions for representational painting understand what he means and
that in that sense it is true. In this section I want to maintain that
the critic who says that Emma often resented her guardian is using
actual-world language in a similar non-literal sense. If Emma the
character is taken to be part of a speech act or of its result, then
what he says is literally false, and downright odd. But again, those
acquainted with the conventions will admit that what he means is true.
Faced with this situation, we have two alternative solutions. Either
we can say that, given the situation and the fact that such usage is
too frequent to be considered deviant, the critics' sentences simply
mean something quite different from what such sentences usually mean.
Or we can say that, though the critic speaks English, what is actually
said is a code for what is meant, a code which we understand if we
understand the context, i.e., the art form and the decyphering process
for what he says. In fact, the semantics for criticism proposed in
chapter III provide such a systematic decyphering process. Either of
these alternatives will lead to the result that the meaning of the
critic's claim is brought into line with its truth conditions and
divorced from the meaning of such sentences in their usual contexts.
Anyone interested in preserving the univocity of meaning will prefer the
second alternative, on which the critic is said to be speaking in code.

No doubt a complete examination of the language of criticism
would contain a far more detailed analysis of the language use in
question. Specifically, it would contain a precise account of why the
critic's "as if" language offers advantages over speaking literally.
For this, we should perhaps need an analogue to the much discussed
"picture space," which we might call "world space-time." One gets some
idea of what this would be like by noting that the physical order and
arrangement of the sentences of a work of fiction do not mirror the
"space" and "time" structure critics speak of a novel as having. I will
not attempt to do this here. Instead, I will consider two objections
which might be advanced by the critic to show that he is not talking
about linguistic phenomena when he speaks about the characters and in-
cidents of a work of fiction. The main outlines of the theory I have
in mind will, I hope, become obvious from the replies to these objections.

The first objection purports to show that our talk about fictional
worlds and entities is not reducible to talk about the expressions used
in the work, hence that we really are talking about something "beyond
the work" which is described and referred to by the expressions in the
work. The second objection purports to show that only if we assume that
works of literature describe and contain references to non-linguistic fictional entities, can we make sense of such sentences as 'Sherlock Holmes was a great detective' and 'Sherlock Holmes was (identical with) the detective who lived on Baker Street', sentences-of-criticism which we commonly and correctly acknowledge to be meaningful, non-trivial, and true (although not in a sense which would lead us to look for Sherlock Holmes' lodging in London or his name in the case files of Scotland Yard).

Objection I: In chapter II we repeatedly encountered such expressions as 'character', 'fact of a fictional world', etc. This is an uneliminable part of our talk about works of fiction. And we cannot reduce talk about the character Emma Woodhouse to talk about the expression 'Emma Woodhouse'. Hence, if 'Emma' refers to anything at all, it must refer to something beyond the linguistic phenomenon of the novel itself, i.e., a non-linguistic fictional entity, to which 'Emma Woodhouse' as used in the novel presumably also refers and which we mean to talk about when we discuss the character Emma Woodhouse.

This is a fairly important objection; most of the objections which attempt to show that the view I am presenting is counterintuitive are of this sort. Against it, I will argue that there is some doubt about the truth of the premises and that, even if we concede for the sake of argument that the premises are true, it does not show what it purports to show.

First off, it is not exactly clear what is meant by claiming that certain kinds of expressions are "not eliminable" from our talk about works of fiction, that we "cannot reduce such talk to talk about
expressions." There are three possible versions of such a claim:

C1: We just do speak in terms of "the character Emma" and "the facts of a fictional world," and not in terms of expressions, etc.

C2: We do speak in terms of "the character Emma" and "the facts of the fictional world," and it would be incredibly cumbersome (although perhaps in principle possible) to do without them and talk about the expressions of the work instead.

C3: We do speak in terms of "the character Emma" and "the facts of the fictional world"; such discourse is legitimate, and we could not in principle restate our points in terms of the expressions of the work.

It is pretty clear that C1 is true, but it is not, I think particularly important. We cannot automatically read off our ontological commitments from the surface grammar of language we consider meaningful. C2 is also true, I think. But it is not clear what, if anything, follows from it, for we often use terms in discussion of complex matters simply for the sake of convenience. We speak about colors, for example, but we could rephrase what we say about colors in terms of light rays reflected from objects and the sensory and brain apparatus and activities of perceivers. Ordinarily we would not do so (and communication at a fashion show would bog down pretty badly if we did). But there are situations in which it would be desirable to do so. C3, if it were true, would be significant. If we could not in principle explain the truth of sentences of criticism about the character Emma Woodhouse in terms of the expression 'Emma Woodhouse', then we would, it seems, have to accept the fictional entity involved in some sense other than that of the expression. But there are serious doubts about the truth of C3. It seems to me that talk about the namelike expression in the language
of fiction, for example, is quite possibly a perfectly good substitute for talking about characters, so long as we are clear about the fact that the function of such expressions is parasitic on the function of genuine names in normal narration and that such an expression in use has vivid associated imagery for the author and the reader.

The claim made in C3, I think, has the same sort of *prima facie* plausibility as the claim that when speaking of "The Flight into Egypt," we cannot dispense with talk of "Joseph's being behind Mary" in favor of saying that the splotch of paint three inches from the top of the canvas is smaller than the splotch of paint two inches from the bottom. The above claim is no doubt true as it stands. But it seems that we can dispense with talk about "Joseph's being behind Mary" in favor of talk about the artist's utilization and the viewer's apprehension of colors and shapes in arrangements conventionally used to convey the impression of seeing a scene of a particular kind. To do so would be very cumbersome, but there might well be situations in which it was metaphysically or scientifically desirable to do so. If the parallel between the colors and shapes of the painting and the language of fiction holds, and I think that it does, then the strongest claim that the objector is entitled to is C2. And C2 is not significant in the way that C3 would be.

Even if C3 were true, however, the alleged consequence that in critical discourse we refer to something besides linguistic phenomena does not follow. To see this, let us suppose that the parallel claim about picture space and picture figures were true. Here I introduce some terminology to facilitate the discussion. Let us say that a portrait
of an individual, e.g., the "Mona Lisa," or a landscape or still life drawn from life is a piece of "representational art." Let us call works of art which are like such works except that they do not correspond in this way to the actual world "presentational art." It has been argued fairly convincingly that our understanding of representational and presentational art is not unschooled, that instead, we learn how to construe such and such a representation in terms of an internalization of the conventions employed in its construction. In fact, it seems likely that we understand a particular style of presentational art because we understand representational art in the same style (or at least that we could go about the task in this way). Suppose, for the sake of argument, that the story of the flight into Egypt, the whole New Testament in fact, were a hoax. We have supposed for the sake of argument that the graphic art analogue to C3 is true, that talk about "Joseph's being behind Mary" is in principle irreducible to talk about the size and arrangement of colors and shapes. Does it follow from this that the work in question, although by hypothesis it does not correspond to a state of the actual world, represents something, something beyond the painting to which the painting stands in a relation like that of a portrait to the person portrayed. Clearly not. Even if "Joseph's being behind Mary" is not eliminable from our talk about the painting, the expressions in question could still refer to a quality of the painting not reducible to its simpler qualities. But obviously on such an explanation, our expressions refer to the painting and its qualities, not to something beyond it which it represents. Of course, it may be true that if no painting in this style corresponded to the world, we would
be unable to interpret this one (we might, for example, say that if we had not learned the conventions for portraying distance using artistic perspective from comparing scenes and landscape representations of them, we would not in this case be able to construe the size and positions of the figures as "Joseph's being behind Mary.") But this is quite different from saying that there must be something beyond or corresponding to, each work of presentational art. That we certainly need not conclude.

Similarly, even if C3 were true, as I have argued it is not, it would not follow from this that our talk about the characters and facts of the fictional world commits us to something beyond the linguistic phenomenon involved in order to make sense of our discourse about such works. But the proponent of Objection I may advance a rather different kind of objection.

Objection II: Perhaps our talk about the characters of a work does not commit us ontologically to fictional entities beyond the name-like expressions of the work. But consider the sentence of criticism:

(1) Emma Woodhouse was wealthy.

which we commonly take to be true. If 'Emma Woodhouse' in this case is taken to refer to an expression, then we are in fact saying in the intuitively true (1) that an expression is wealthy. This is most likely false; and it is in any event certainly not what we meant to say.

The answer to Objection II is that, although we commonly take (1) to be true, we do not take it to be true in the sense in which

(2) Howard Hughes is wealthy.
is true. In the simplest case, (1) is true if 'Emma Woodhouse' is concatenated in the novel with 'was wealthy', just as 'Joseph is behind Mary' is true of "The Flight into Egypt" if the figures are arranged in a given way. In (1), we are not attributing a real predicate to an expression anymore than the art critic means that the Joseph-figure is literally behind the Mary-figure. We are instead saying something about what expressions go together in the novel. Thus (1) does not express a category error on the interpretation of 'Emma Woodhouse' I am proposing, although the critic in asserting (1) may on reflection be saying something a bit different from what he took himself to be saying. The complete analysis of (1) might be quite complex; no doubt using actual world language and talking about Emma as if it were history is more perspicuous than the analysis would ever be. And the critic may have been led astray by the fact that vividly presented characters become in a sense "real" for us—we talk of them as if they were people. But, of course, they are not people; and they do not have properties in the literal sense in which people have them. But the objector may try a slightly different approach.

Objection II: Consider a sentence of descriptive criticism like:

(3) Sherlock Holmes was (identical with) the greatest detective on Baker Street.

which we would no doubt consider true. If (3) is about 'Sherlock Holmes' and 'the greatest detective on Baker Street', then (3) is clearly false. For it says that the two expressions in question are identical, which they obviously are not (they have different shapes, etc.). Thus, if we want to hold that (3) is true, we are forced to concede that it is true
because the expressions in question refer to one and the same thing, i.e., a fictional entity which is not identical with either of the expressions.

I think we can answer Objection II' without conceding what it asks us to concede. When is a sentence such as (3) true? In the simple case, it is true just in case in the work 'Sherlock Holmes' is concatenated with 'was the greatest detective on Baker Street'. In a wider context, two expressions like those in (3) may be joined by the identity sign to get a sentence true of the work just in case they present the same character. They present the same character if they belong to the set of expressions which is such that every predicate expression which is concatenated with one of them in the novel may correctly be concatenated with the other. The conditions under which (3) is true give an important clue to what it means. To say that 'a=b' is true of a work of fiction is to say (i) that 'a' and 'b' are concatenated in the text by some stylistic variant of 'is identical with' or (ii) if 'a' and 'b' occur in the work and if for every sentence 'Fa' true of the work, there is a sentence 'possibly (Fb)' true of it, where 'possible' is interpreted as 'causally possible, relative to the constitutive laws of the universe in question'. Thus we see that to identify the characters of a fictional work with linguistic types is somewhat of an oversimplification. Often several expressions are used to present the same character. But the character is still nothing over and above the group of linguistic types. Ordinarily the fact that two expressions present the same character is clear from the text. The two expressions are at some point connected by a form of 'is identical with'. (One stylistic
variant of 'is identical with' is the appositive construction. e.g., 'my cat, Petronius the Arbiter!). Ordinarily, then, it is clause (i) of the fictional interpretation of 'a=b' which we rely on. It is hard to imagine a case where (ii) would be fulfilled. But (ii) helps explain how we sometimes suspect that an identity sentence is true of a novel long before (i) is fulfilled. In normal narration we discover that an identity sentence is true by discovering that two blocks of descriptions which back referring expressions determine the same referent (or by having someone tell us this). Sometimes the discovery of true identity sentences in normal narration takes a long time. In descriptive criticism, we decide that a given identity sentence is true of the work if we find that the expressions which are its terms present the same character. Sometimes, as in the case of Dr. Jekyll and Mr. Hyde, the discovery that the two are identical takes quite a while. Finally we are told this explicitly. But we suspect, long before we are told explicitly, that 'Dr. Jekyll = Mr. Hyde' is true of the novel. This is probably because we begin to realize that with every instance of 'F(Dr. Jekyll)', 'F(Mr. Hyde)' could be true as well, and vice versa. We never see Jekyll and Hyde in a relation that would require that they be distinct. And everywhere Hyde appears, we later discover that Jekyll could well have been, and vice versa.

The answers to Objections I, II, and II' make plausible the claim that sentences of criticism are about linguistic performances of their parts, even though they may not appear to be so. Sentences of criticism are true (or false) on the basis of what expressions are included in a work of fiction.
Notes

1 Willard van Orman Quine in his essay "On What There Is," From a Logical Point of View (New York: 1953), sets out an argument of this kind. Actually the chain of argument he sets out there starts not from any old claim about Pegasus which is, intuitively speaking, true, but from the "obvious" truth of the sentence 'Pegasus does not exist'. But, since he later refers to Russell's theory of descriptions as having "showed clearly how we might use seeming names without supposing that there be the entities allegedly named," it seems reasonable to suppose that he could just as well have started from 'Pegasus has wings'. Of course the air of paradox would be lost, but the point, it seems, would be the same.

2 Gottlob Frege quite possibly held a view similar to this in "On Sense and Reference," Translations from the Philosophical Writings of Gottlob Frege, ed., Geach and Black (Oxford: 1952), where he maintains that "the thought" of a sentence containing 'Oedipus' remains the same whether 'Odysseus' (and consequently for Frege, the entire sentence) has reference (for Frege a truth value) or not. Unfortunately Frege does not explicitly distinguish sentences of criticism from sentences of fiction, although at some points it seems clear that he has the latter in mind. In addition, he never makes it clear whether he thinks sentences of criticism lack truth values because their subject terms in fact happen to lack referents, or because they are used in a way which makes it wrong to assign truth values. Some of what he says, e.g., "The question of truth would cause us to abandon aesthetic delight for an attitude of scientific investigation," p. 63, point toward the latter reason; but the choice is never univocally clear.

This is the position I think is the most desirable. But there is an alternative. One may decide to assign truth values to such sentences in the usual manner, then deny that their truth values are an important feature of their use and even hold that it is misleading to call them "true" or "false." This is essentially the view taken by Wilfrid Sellars, "Presupposing," Philosophical Review LXIII (1966), 202. If one were to accept this view, then of course, the argument that the author's sentences are true by fiat is no longer acceptable, since some of them will now be true, others false, accordingly as they correspond to the way the world is. I see no good argument for doing this, but one might do so without disturbing the central points of the thesis I hold.

3 This sort of semantics for criticism is suggested by Bertrand Russell both in "On Descriptions," Introduction to Mathematical Philosophy, ch. XVI (London: 1919), and in "On Denoting," Mind XIV (1905), 491. Russell never carried out this program in detail, probably because he was not particularly interested in literary criticism.
It has been objected that in cases such as, for example, the Oz-series, where the initial author died and someone else carried on (presumably without the initial author's consent), we should certainly consider the whole series as a single fictional work, a single fictional world. I think this is unnecessary. In chapter V, techniques will become available to us for amalgamating separate universes, provided we make the separate origins of the work clear, and it seems to me that this treatment is in fact preferable. In addition, the gain of being able to handle the case in question costs a good deal, for it makes it much more difficult to handle cases in which we want to deny that two authors have not written parts of the same work.

Another problem is that of, say, ancient epics, where our data about authorship are sketchy, but points to multiple authors. Again, it seems that the amalgamation treatment is in no way inferior to the single work treatment, and that considerable clarity is gained by handling works in this way. (It is perhaps a good way to separate folk epics from literary epics).

If the majority of modal logicians are correct, the logic of logical necessity is the modal system S5. I am indebted to George Schumm for pointing this out and for help in determining the modal systems appropriate to different conceptions of time.


Tbid., p. 6

Tbid., p. 6

Tbid., p. 25.

Of course if someone makes what we at first take to be a critical claim, but it turns out that he thinks he is and intends to be describing the actual world, then we may be justified in applying one of these theories to his remarks. In this respect, I should agree with either Strawson or Russell. I am merely claiming that their theories are not useful at all in dealing with critical claims about works, since they lead to blurring almost all useful distinctions.

I will later argue (chapter V) that there are some acceptable arguments of this kind. They are based, as may be expected, on an author's (or a group of authors') customary mode of action. But they are inductive arguments about actions, not deductive arguments.

We might argue that since it is true that fifteen streets of London were clean on a given day in question (on the basis of fiction) whereas it is true that in fact two major streets were filthy on the day in question, that London was predominantly clean on that day. Of course, this could only happen if we did not differentiate the senses of 'true' involved, and this is exactly the distinction we want our semantic account to make clear.

I should here be taken not as strict identity, but as indicating the ascription of a semantic predicate.

Normally, the Hj we are interested in will be G, the actual world. But sometimes works of fiction are accessible from other fictional worlds, e.g., The Prioress' Tale is accessible from Canterbury Tales, and Ossian is accessible from The Sorrows of Young Werther.

I am indebted for this way of stating the quantification clauses to suggestions from William Lycan and George Schumm.

Among other things, this means he can distinguish which of the following uses of 'the' - constructions is parallel to the normal narration referring use:

(1) The Whale is the largest Mammal.
(2) The gladiator kicked the bucket.

and that he knows an intentional context of a modal or tense context when he sees one, and that he recognizes stylistic variants of quantified and identity constructions, etc.

cf. footnote 11.

Thus the account I am proposing differs from that proposed by Ryle and Braithwaite in "Symposium on Imaginary Objects," Journal of the Aristotelian Society, Suppl. XII (1933). Unlike Ryle, I hold that even if someone should turn up with all Mr. Micawber's characteristics (even being named 'Mr. Micawber'), the sentences used in David Copperfield to describe Mr. Micawber should not become true statements about him. They are not statements because they are not intended as statements, no matter who turns up. And contrary to Braithwaite, I hold that 'Emma' in 'Emma went to Bath' as used in the novel, is different in function from 'Jane Austen' in 'Jane Austen went to Bath' as it occurs in a literary history book.

No doubt, neither (i) nor (ii) is a necessary condition for the truth of 'a=b'. (i) alone is certainly sufficient, provided 'a=b' does not appear in a quote from one of the characters or in some other intentional context. (ii) alone is not sufficient if we consider only the sentences of the deductively derived state description. This is why there is no clause for identity sentences in the recursive truth definition. For 'John was in the library' and 'Tom was in the library' would go toward satisfying (ii), although Tom must be understood to have been there at t1 and John at t2. We must also take into account
sentences other than those available in the deductively derived state description, those we will cover in chapter V. (ii), it turns out, is incredibly strong, although if (i) fails to apply, it is the fulfillment of (ii) we work toward in criticism. We generally conclude the investigation considerably short of getting there.
It is interesting to note how sparse the deductively derived description of a fictional world is. It consists of (i) the corresponding sentences for the "facts" of the world, i.e., the basic description of the fictional world; and (ii) all those sentences inferred by inferences which are validated by quantified $S_5$, in short, of the logical entailments of the basic description. It is certainly true that in some cases arriving at a deductively derived description is itself very involved; for example, getting the "facts" of Balzac's *Human Comedy* is a major undertaking. But it also seems true that much of the most interesting literary criticism goes well beyond the critical claims we can validate by the procedure in chapter III. Indeed I will argue in this chapter that we are forced to go beyond the procedure in chapter III to give an account of kinds of criticism nearly everyone agrees are legitimate.

Despite the sparseness of the complete state description in chapter III, such a description is not without interest for literary criticism. For there is a popular critical approach, the so-called "new critical" or "formalist" approach on which legitimate criticism is restricted to those features which are explicitly in a work. And there have been attempts to extend this critical approach to the criticism of fiction as well. In section 4.1, I will attempt to show that this
position is misguided. I will proceed as follows. I will give a
formulation of the basic position of what I will call the "hypotheti-
cal formalist," construing such a position as generously as possible.
In doing so I will stress the typically formalist view that we are
restricted in the criticism of a work to elements internal to the
work. Such a procedure, of course, lifts the formalist position out of
its historical context and more or less ignores the fact that many
features of the position are stated as strongly as they are for polemi-
cal reasons. We shall also ignore many of the differences between
groups of flesh and blood formalist critics. But these, I think, are
not crucial oversights. For it is important to note that if we take
the basic tenet of the position at face value, then even on an extremely
generous formulation of the theory, there are certain kinds of criticism
which involve the formalist in difficulties if he accepts them as legi-
timate. If it should turn out that no formalist in practice adheres
to the strict statement of the position, so much the better for the
formalists. In section 4.2, I shall take up the case of an explicitly
formalist position and show that, if the critic in question is construed
as holding the position I have ascribed to the "hypothetical formalist,"
then he is in practice making assumptions in the interpretation of a
work which he is in no position to justify, though they could be de-
fended on other, explicitly non-formalist grounds. The point of 4.2
is to show that he must either allow such moves in theory or be open
to the charge of not applying his principles to his practice.
In this section, I will first list a group of critical activities which
seem to be thoroughly legitimate. I shall assume in any event
that they are. I will then argue that one of these is basic to all the
others and that the formalist cannot account for such criticism, since
it demands making interpretive choices he is in no position to make and
justify. As a result, I think it will become clear that we must give
up the formalist position as it is held by formalist critics, although
we may hope to salvage some of the features which make it so attractive.
The kinds of criticism that I propose to consider are the following.

(1) **Explanatory criticism.** In criticism we may be told what a
character says and does, where the claims are based on the procedure
of chapter III, and this alone may be very interesting. But the critic
may also give an analysis of novel situations based on an explanation
of the characters' psychology, their motives, and the kind of world
they live in. A good deal of the most interesting criticism, often
called "interpretive criticism" falls under this heading.3 (ii) **In-
competent narrator.** Fairly often we find ourselves in the case of
epistolatory novels or in novels with a deviant (lying, demented,
ignorant) narrator, speculating about what is "really going on." Or
we find ourselves in a center of consciousness novel reaching beyond
the interpretations the center of consciousness supplies, even of data
which are filtered through his consciousness.4 (iii) **Inconsistent
universe.** We are also faced with the problem of how to deal critically
with purported fictional universes whose descriptions seem to be inco-
herent. In such cases, the critic does not go beyond the facts
actually presented by the author. He commonsensically rejects some of
them. (iv) **Character comparison.** Some of the most interesting pieces of criticism are arrived at by comparing characters from different fictional works, and occasionally by using conclusions about a character or situation of another work to clarify a character or situation of another work. But on the basis of chapter III, no sentence of criticism connecting names of fictional characters from disjoint worlds is fictionally true. (v) **Poetic language devices.** We also find a large number of critical arguments based on the poetic language of works of fiction (simile, metaphor, symbol, etc.). Usually arguments about such uses of poetic devices occur within the larger frame of more wide-ranging arguments, which purport to show that a scene of an entire work has content which is not captured by a literal reading of the reports of the work or explanations of them. (vi) **Cognitive significance.**

Finally, we have the problem introduced in chapter I. Works of fiction, it is alleged, do have some function in expanding and clarifying our knowledge of the actual world. Most of the techniques developed in chapter III are directed at getting a description of the world of the work, not at explaining the relation of the work to our knowledge of the actual world. In addition, in chapter I, it was argued that works of literature do not have as their meaning propositions about the actual world. But there is something about them and their relation to our knowledge of the actual world which needs explaining.

In doing criticism of types (i) - (vi), we certainly do not feel that rigor has gone on a holiday. An account of literary criticism needs to take account of all of them. It is important to note that explanatory criticism (type (i)) is basic to all the other kinds of
criticism listed. At present I will merely point out in general why this is so; I will argue the point in detail in 5.1, where I take up (i) - (vi) individually.

Criticism of type (i) provides us with information about the kind of world we have to deal with, the motivations of the characters, the causal connections between events, etc. Usually, in the case of incompetent narrator criticism (type (ii)) we are dependent upon arguing that a narrator is reporting on a world of a given kind. Otherwise we could not very well claim that he is incompetent; we would have nothing to weigh his reports against. In such cases, we are obviously depending upon extensive criticism of type (i), using it to decide what the world is like and what the narrator is like so that we can decide whether the narrator is competent, and, if he is not, in what sense he fails. In the case of an apparently inconsistent universe (type (iii)), we have to decide whether the universe is inconsistent per se or whether the inconsistency is produced by our own assumptions about what the world is like. In the latter case, we are dependent upon criticism of type (i) to decide whether our assumptions are warranted. Criticism concerning poetic devices is not so obviously dependent upon explanatory criticism, but in section 5.1 I will argue that this is the case nonetheless. The case for criticism involving the cognitive significance of a work is obvious. We have to know what the fictional world is like in order to explain how it functions in influencing our knowledge of the actual world.

Since explanatory criticism seems to be so central, let us first turn our attention to it and see if the formalist critic can give a
satisfactory account of it, bearing in mind that if he cannot, he will also be unable to handle the kinds of criticism which are dependent upon it.

The most important conclusions which we customarily draw about fictional works and which go beyond a description of the work and its logical implications are inductive generalizations about the way the fictional world is and analyses of particular characters and situations. All these critical activities fall under what I have called "explanatory criticism." As in the description of the actual world, such activities presuppose other, methodologically more basic assumptions about the law structure or constitution of the universe in question. (Which assumptions are presupposed may vary with the kind of conclusions we are interested in.) Of course, we are more or less aware that many of the higher level descriptive claims we make about the actual world are "theory-laden" or "theory-dependent." The terms we use make sense and the claims we make are correct, given the tenability of certain basic theories about the way the world is. (We would be hard pressed to talk about personal interaction, if we could not rely on certain correlations between behavior and intentions, etc.). What I wish to emphasize is that our situation with respect to a fictional world is similar, that to achieve a comparable level of sophistication in our descriptions of a fictional world, we rely on assumptions about the lawlike constitution of the fictional world, about the sorts of laws which govern it, e.g., whether it is a "magical" or "naturalistic" universe. And we commonly claim that such assumptions are justified.
Can the formalist account for this procedure? The position of the enlightened formalist is that in describing the facts of a work, we are restricted to the explicit facts of a work and their deductive consequences, and to those descriptions which we obtain by analyzing the meanings of the expressions of a work in isolation and in the juxtaposition in which they are used. This is no simple task. The procedures outlined in chapter III are complex. And their implementation demands as a minimum a very competent grasp of the syntax of the language of the work. In addition, the meanings of expressions are not always accessible to the common reader. Often, we may have to search through the whole dictionary account of an expression to get a sense of the full meaning of the expression, its special nuances and connotations. And often we must read a combination of expressions again and again in order to apprehend the full force of their combinations, combinations which may be structurally complex and significantly recurrent, their special shadings and complex interaction. But when the critic has finished this undertaking, his job is finished according to the hypothetical formalist. He has dealt with the literary object qua object, and no more can be asked of him. He need have no recourse to the circumstances surrounding the creation of the work, what the author intended, what motivated the author, the time in which it was written, or other works of the period. His attention is riveted to those features of the work which are apprehensible, although sometimes not obvious.

The appeal of such a position is obvious. Some poems, we are inclined to say, are ironic or satiric, even if the author protests that they are not. And we do not need to know the biography of Dickens to
note that his heroines are either simpering featherheads or figures of serene wisdom and strength. But the question is whether this position so strongly formulated can handle the question of explanatory criticism. For there is a marked difference between fictional worlds and the actual world. In the actual world, we suppose ourselves to be at least theoretically in a position to justify our lawlike basic assumptions in a given case in some objective way, perhaps by an appeal to the authority of the appropriate science or by reflecting on our past experience of the actual world. In the case of a fictional world this particular avenue is never open to us. We cannot very well appeal to science about the actual world or our experience of it to justify our decisions about the lawlike constitution of a fictional world. To be sure, we must have such assumptions even to be able to read the work intelligently. We must assume that certain non-logical laws or relations of predicates which hold in the actual universe also hold in the fictional universe, or that they are replaced by laws which do not hold in the actual universe. The question is how we argue for such assumptions in the case of a fictional world.

There is no point in making the naive claim that we just begin reading, then learn from the atomic and quantified facts of the fictional world that certain laws hold, as we do in the actual world. We cannot just learn from a fictional world the way we do from the actual world. For many of our common procedures for dealing with the world, for example simple induction, depend upon prior assumptions about the constitution of the world, in this case causal regularity. And in the case of a work of fiction, it is those assumptions that we
are arguing about. Nor will it do to maintain that some set of assumptions about the constitution of the fictional world gives us a feeling of appropriateness, "just fits." This simple approach neglects the fact that there may well, probably always are, more than one set of basic assumptions which would satisfactorily and completely explain all the "data" obtained by the procedures outlined in chapter III. Any of these would presumably give us that comfortable feeling of fittingness.

It will not do to argue that when we know by the procedures of chapter III that certain laws hold in the fictional universe, we are entitled to claim that others hold also, that our assumptions about the law structure of the work are therefore based on its internal features and what follows from those features. For the relationships between laws are in general non-deductive. And they do not seem to be analytic or synthetic a priori either. And it is only such inferences to which we are entitled. A conclusion about laws drawn in the proposed way would be based both upon what laws we know to hold in the fictional world and the connections we believe to hold between laws in the actual world. In such a case, we would have to argue that the fictional world was enough like the actual world that normal connections between laws still obtained. And it is about the nature of the assumptions we are allowed to make that we are arguing. Without such assumptions we can never get to the sophisticated descriptions of the fictional world which even the formalist critic agrees are legitimate.

But the formalist has presumably not yet had his say. He may attempt to claim that he is entitled to the assumptions he makes
because the assumptions are true by virtue of the meanings of the expressions which constitute them. Let us allow the formalist to make any assumptions about legitimate inferences that seem to fall under the heading of knowing the syntax of the language. We have to allow these to anyone who reads the work. We will also allow what we have assumed in the discussion so far, that fictional expressions are analogous to normal naming and individuation, that predication and identity constructions, direct and indirect quotation, propositional attitude constructions and metalinguistic constructions, and tense, modal, epistemic, and deontic expressions, all have the functions in fiction which are analogous to their function in normal narration. Perhaps these also come under the heading of knowing the syntax of the language; at any rate, any competent reader of the language will understand such constructions. In addition, to forestall any arguments, let us concede to the formalist those connections of predicates usually pronounced "analytic" or "synthetic a priori," which have often been held to be true because of the meanings of the predicates, e.g., 'A bachelor is an unmarried man' and 'Everything colored is extended'.

Can the formalist now give a satisfactory description of a fictional world? Among the inferences he wants, presumably, will be ones like that from:

(1) $S$ knows that $p$

to:

(2) $S$ believes that $p$

or, for example, from what Roger Chillingworth says to Hester in the prison scene to:
(3) Roger Chillingworth wanted revenge on Hester's partner in adultery.

Or perhaps he will, for example, want to go from a description of Roger Chillingworth's actions and from what he says to the conclusion that Roger Chillingworth had always lacked a goal structure and therefore jumped at the chance to adopt one, even one which would destroy everyone involved. Can such inferences be justified on the basis of the given data, syntax, and analytic and synthetic a priori connections? I think not. In normal narration, we would base inferences of this kind on deeply embedded generalizations about the actual world which we originally learned from experience, but which play an important role in our subsequent acquisition of knowledge and form a foundation of knowledge, e.g., that persons' intentions and personality traits are evinced by what they say and do, that causality is in some sense uniform, that there is a comprehensive and uniform spatio-temporal framework, etc. We also use some common rules of non-deductive inference, e.g., enumerative induction. But these in turn depend on certain assumptions about the way the actual world is, e.g., that it is causally regular. To make the above inferences, the formalist must rely on parallel assumptions and analogous procedures which he claims are legitimate in the fictional world. And in a fictional world such assumptions are not always warranted. Or at least, we cannot automatically assume that they are. Let us suppose, for example, that Balzac assigns two different dates to the same event. If something of this sort occurred in a history book, we would know immediately that one or both of the dates was incorrect. We should, of course, be very surprised if a Balzac universe were not causally normal relative to the actual
universe, even if it contained "reports" which indicated that an event had two distinct dates. Nonetheless, we could argue in such a case that Balzac was, contrary to popular opinion, writing science fiction. No doubt, this would involve us in extensive re-explanation of all the other "facts" of the world. The explanation that Balzac simply lost track of his earlier report, that the universe is normal after all, is simpler. But both theories account for the facts. And it is by no means clear that the simpler theory is automatically preferable. This view has consequences for some works that few critics would want to accept. In a science fiction work with a first person narrator, it would generally be simpler on a commonsense level to assume that the narrator has taken an overdose of something and hallucinates the purported events than to work out the complex tense logic involved and decide whether it is consistent. There is also the question of whether in aesthetics in particular the simplest set of assumptions should be the set of assumptions needed to ground an interpretation which is to be preferred.

The formalist might, of course, object at this point that in a problem case we need not argue that some set of assumptions leads to the correct interpretation. It must may be, he might contend, that problematic works just are ambiguous. The critic's job in such a case is simply to delineate the alternatives which would explain the facts. But the consequences of this move are more far reaching than the formalist realizes. It turns out that every work is in principle a "problematic" work.

Let us take *The Scarlet Letter* as a novel which supposedly poses
none of the sort of interpretive problems we are worried about, no
double dating, etc. Yet in the case described in chapter I, we
assume, though we are not told, that Chillingworth's face contorts
in horror for some reason or other—people generally do not do so un-
provoked. And we presume that it is what he has just seen which
causes the contortion. People do recoil in horror from what they see.
We regard Chillingworth's subsequent behavior and the account he gives
of his behavior as establishing this hypothesis beyond the shadow of
a doubt. In the contrasting case, we do not causally connect the
sombre weather with the breeze stirring the clouds overhead, although
one report follows the other. Sombre weather just does not cause breezes
to stir clouds. In such cases, we are in fact assuming that the causal
and psychological laws of The Scarlet Letter are the same as the cor-
responding laws of the actual world, and that The Scarlet Letter is,
in the terms of chapter III, one of a group of worlds which form an
important subset of \( K \), those which are at least partially causally
normal relative to \( G \). In such worlds, we can conclude on the basis of
\[ \phi(L(A), G) = T' \]
where 'L' is interpreted as 'causal necessity', that
\[ \phi(L(A), H) = T' \]
where \( G \) is the actual world and \( H \) is any member of the
subset of worlds of \( K \) which are causally possible relative to \( G \). And
it is only by making such an assumption that we can arrive at the "facts"
of a fictional world.

But even in such a case as this, no doubt a sufficiently ingenious
critic could construct alternative sets of assumptions which would
account for the explicit facts and which did not rely on the assumption
that the world of *The Scarlet Letter* is causally normal relative to the actual world. Some of the alternative theories might be even simpler than the one generally adopted. The price of the formalist's move in the "problematic case" is thus extremely high. Any work may be a problem case. We must grant, if his move is allowed, that all works are interpretively ambiguous.

Thus the formalist finds himself presented with a dilemma. He can claim that all criticism of type (i) (and all criticism dependent upon it) is illegitimate. In this case, we will be unable to get any criticism beyond the criticism we could get through the procedures in chapter III and that which is based on the meanings of the expressions of the work. Once the hidden presuppositions of our normal reading procedures are exposed, it turns out that on this alternative, we lose most of criticism. Or the critic can allow criticism of type (i), thereby granting that all works are interpretively ambiguous. This dilemma, I think, constitutes a reductio ad absurdum of the formalist position. The obvious solution is to give up the objectional portions of the formalist position. Henceforth, I will assume that the formalist position has been given up and explore the possibilities which are open to us in explanatory criticism, once it is given up.

Assuming that the formalist position has been given up, we are in a much better position to deal with explanatory criticism. We can argue extremely plausibly, I think, that nearly all of Balzac's novels are plausibly construed as causally normal relative to the actual world, as are nearly all novels written before the 20th century. In one case, we depend upon an induction about a person's customary mode of action, in
the other on a mode of behavior common at a historical point. The first argument has the form of an inductive argument that an agent A ordinarily performs an action of type B when he is performing an action of type C. In Balzac's case, we might argue in addition that the work in question is a part of a series, The Human Comedy, all of which are plausibly construed as causally normal. And in some cases, we may argue from the artist's intentions, e.g., from Balzac's prefaces or his literary-theory incursions into his fiction—so long as we keep in mind that it is a general truth about actions that their performers can misdescribe them. I prefer to restrict 'the artist's intention' to the descriptions and interpretations an author actually does give of his work and to those things which follow from what he says, i.e., to what he would or should assent to, given what he actually says. Other external arguments based on circumstances surrounding the creation of a work, I prefer to call "genetic arguments." It is perfectly conceivable, I think, that an author might give an aesthetic description of a work or an interpretation of it which was not a good explanation of the facts presented in the work or which was at variance with a set of assumptions which did deal more adequately with the presented "facts."

In such a case, we would say quite appropriately that he had misdescribed his own (quite complex) speech act. Of course we can also be misled by genetic arguments. We might, for example, conclude from the date of a work that it was a romantic work and proceed to interpret it as if it had the special symbolic vocabulary of such works, then discover that such an interpretation could not account for the "presented facts." But in the absence of a counterargument from features internal to the work,
an argument based on the author's intentions or the circumstances surrounding the creation of a work is admissible as an argument about the nature of the law structure of the work, i.e., about the subset of $\mathcal{K}$ to which the work belongs. In the present case, none of Balzac's individual novels gives evidence that he has misdescribed his action in telling us that he is constructing a normal universe.

So far we have considered only arguments to show that a work does have a particular law structure. But there are times when it is important to argue that we are not justified in using laws or connections of predicates which hold in the actual world to map out the geography of a fictional world. The clearest cases in which we are not allowed to do so are those in which we have direct or internal evidence that some one of the generalizations we normally rely on does not hold, that the normal assumption cannot explain the data. But external or combined arguments may also be important here. In a work by Kafka, for example, or a work in the genre of science fiction, caution is a rule of thumb. But the same features which influence our reading procedures may be used to defend a particular interpretation. In a science fiction novel with a personal narrator, we could, of course, always argue that the universe is normal and that the narrator has taken an overdose. Such interpretations become extremely implausible, however, in the face of appeals to the author's other works or to works by authors writing in the same tradition or the same artistic circle which are also plausibly explained by the assumption of abnormal law structures. We can also have recourse to the author's explicitly stated intentions or, for example, to the fact that he was discursively considering some problem
in his diaries or non-fictional works and proposed to write a
fictional work along the same lines. If such considerations support
an explanation which is already conceded to be explanatorily adequate,
then that interpretation is probably the best one.

It will no doubt be objected at this point that the theory I
propose is excessively intentionalistic. And certainly it is what I
should prefer to call "externalistic." One answer to such a criticism
is that the alternative is universal interpretive ambiguity, not with
regard to the specifically aesthetic properties of the works considered
or our evaluations of them, but with regard to the "facts" of the work.
Furthermore, it is not clear that the theory I propose is vulnerable to
the stock objections to intentionalistic theories. It is certainly true
that some arguments based on the author's intentions turn out on my
account to be good ones. But the author does not thereby become the
final arbiter of the nature of his fictional speech act, any more than
any user of language is the ultimate arbiter of the nature of the speech
act he performs. The author of a work of fiction may, as I have pointed
out, ascribe to the work aesthetic qualities which it in fact lacks.
Or he may advance an interpretation which is not explanatorily adequate.
Or he may make a value judgment about the work as a work of a given
kind which we do not accept. Very likely, he knows better than the
rest of us what his intentions were—although even this is open to de-
bate. But he may not know whether the work embodies or fulfills his
intentions. If I am right, we cannot conclusively decide on the inter-
pretation of a work on "internal grounds." Nevertheless, internal
evidence may rule out some interpretations. It is also true that on
the theory I propose, some genetic arguments, e.g., those based on an
author's customary mode of action or the literary circle to which he
belonged, turn out to be good ones. We have already pointed out that
the author's explicitly stated intentions may be discounted on "in-
ternal grounds." They might also be discounted on the basis of broader
genetic arguments. (A trivial example of this, perhaps, is that we
might discount the author's statements because we knew that he had
undergone a massive personality change between the writing of the work
and the expressing of the opinion or that he had disavowed at the time
of stating his critical opinion the worthwhileness of a genre we have
good reason to believe the work belonged to).

Even lacking such specific information, we can sometimes make in-
terpretive decisions which are not totally internal. A parallel with
graphic art is useful here. If we had no information about Monet,
we could base our interpretive arguments about his work upon other works
similar to his in their internal features, works about which we did have
specific information. If we had no historical knowledge at all about
the impressionists, but did have a large body of their works and in
addition some photographs of the scenes they represented (paperclipped
to the backs of one or two of the pictures, let us say), then we could
begin to theorize about the "facts" of the works via our conclusions
about the conventions of the kind of representation involved. If we had
no information at all about the circumstances of creation, however, we
would not be able to decide whether a group of works was representational,
let alone what was represented (if anything). But this conclusion, far
from being undesirable, is clearly what common sense dictates. The
parallel for works of fiction, which are representational in a slightly different sense, is obvious. Here again we are relying on external considerations, but I cannot see that we are relying on the author's intention in the narrow sense. The theory I am proposing is externalistic. But I cannot see that it is vulnerable to the stock criticisms of intentionalistic theories. If it could be shown that all external arguments were reducible to intention in my narrow sense, or that my distinction between intentional and genetic arguments is untenable or unhelpful, then the theory I propose would, of course, properly be described as "intentionalist." But I think the notion of "intention" would then have been broadened enough to escape the most obvious criticisms of intentionalistic theories.¹⁰

Of course, the formalist is right in claiming that internal evidence is primary, that the data given in the work is the initial testing ground for any explanatory theory about the work and for descriptions which rest on such explanatory theories.¹¹ What I have tried to point out is that categorically restricting criticism to the primary data cripples criticism in a rather unexpected way.

What is the status of descriptive claims about a work which rest on the assumption of some law as holding for the work? Is it true of The Scarlet Letter that Roger Chillingworth recoiled in horror because he saw Hester on the platform? If a given theory is taken for granted as part of the lawlike constitution of the actual world, then we do not ordinarily qualify statements which are dependent upon it by a reference to the theory. The descriptions of a work of fiction dependent upon explanatory assumptions are not unqualifiedly true of the work as are
those which come out true on the basis of the procedure of chapter III. I think we may say that such statements are true of a work, given that they are derived from true bodies of data and from warranted assumptions about the constitution of the universe. If the assumptions are uncontroversial, we need not qualify the descriptions by a reference to the interpretive theory. In cases where the assumptions are contested or where we are just not sure that the interpretive assumptions are the best explanations of the data of the work, we should include a reference to the interpretive assumptions upon which the higher level descriptions rest.
The formalist theory I have selected for consideration is advanced by William Handy in "Toward a Formalist Criticism of Fiction." This formalist theory is by no means the only theoretical possibility for a formalist theory of fiction. But it is typical of one particular school of formalist critics. And one of the interesting features of such positions is that they are not in general vulnerable to what seems to be the most powerful line of criticism of formalist critics such as Beardsley. It may be helpful to review this line of objection to Beardsley-type theories.

As we have seen in chapter I, critics like Beardsley are driven to the position that the work of literature contains an alleged proposition about the actual world in its poetic or symbolic meanings, and that the work must be construed as consisting of or containing such propositions or statements. Beardsley is a fairly representative critic of this type. The objection which he himself acknowledges to be the most powerful one against his theory is based on the fact that he must claim that certain elements of the work come to have quasi-allegorical significance. The main objection to this is not that it involves a blatant misclassification of works of fiction, but rather that if such a view is correct, then either every literal element of the narrative must have some allegorical significance, or we must be provided by the formalist critic with some non ad hoc way of deciding which ones do and which do not. The first alternative quickly becomes ludicrous; and Beardsley in effect admits that he is in no position to provide the criterion demanded if he holds to a strict intranscendent position. Nor is he willing to make one of the standard retreats by pointing to
certain archetypal patterns or plots which are common in literature
and which furnish a guide to which elements of a work are symbolically
significant. The obvious reason for this is that the move in question
involves an appeal to other works of literature (or to the way the
world is) to explain which of the features of a work explain its
meaning, whether or not the appeal is made explicit.

Handy's theory is an explicitly formalist theory which is not sub-
ject to the criticisms which count against Beardsley-type theories.
Handy argues, proposing an interpretation of Ransom's view, that a work
of fiction should be considered, not with regard to its "argument" or
"plot," but with regard to its "texture," composed of "devices or forms
through which a concrete presentation of meaning is achieved." He is
proposing, then, that we treat the work of fiction, not as a fictional
world composed of "facts" in imitation of the actual world, but as a
complex aesthetic surface which has structure, somewhat like a picture.
He seems to mean by this that we are to see the work as extended in
time, much as we see a painting extended in space, when we ignore its
grosser features sometimes described as "representational" and con-
centrate instead on its more abstract qualities of composition (logic)
and texture. The basic unit of poetry seen in this way is the image;
the corresponding unit element in fiction is the scene. The scene
becomes the literary artist's means of "transmuting ideas into sen-
sations." I suppose that what Handy means by this is that when, for
example, the artist wishes to convey the atmosphere of hopelessness,
he may write a scene or sequence of dialogue and description consisting
of terse, colorless statements which, so to speak, "fall dead," have
no beauty or poetic sonority. And the succession of scenes in a
fictional work, Handy claims, presents a meaningful pattern because
the scenes in the framework of composition provided by the work's
"logic" constitute not merely a representation of experience, but a
judgment about experience. In some sense, then, the arrangement of
scenes of the work impose an ordering on experience which we experience
as a judgment about the way the world is, or perhaps the way to see the
world. The poem in this sense educates us to the way to see the world
in which we live. We seem prima facie to have here the claim that the
scenes of the work of fiction are so arranged as to constitute or con­
tain a proposition or propositions about the actual world or about our
experience of the actual world.

It is readily seen that such a theory escapes the main pitfalls of
other types of formalist theories. Let us concede for the sake of
argument that we have a clear notion of what constitutes a scene of a
work of fiction and of how to individuate them. When asked which of the
elements of the work are relevant to determining its meaning, Handy can
reply, "All of them." In addition, given what Handy says about
"judgments," his theory will perhaps escape some of the stronger argu­
ments against Beardsley's position, those based on Beardsley's inability
to provide elements which serve as identifiable subjects and predicates
of the propositions he claims are the meaning of the work. For on
closer inspection of Handy's view, we find that he maintains only that
the scenes of a work of fiction are individually expressive of an ab­
straction, hence that they vividly exemplify an idea, and that their
ordering imposes on the more chaotic blurry run of daily experience an
order.\textsuperscript{17} Let us concede then that "judgment" was an ill chosen term, that Handy has in mind something weaker.

It must also be conceded that Handy has picked for his example a novel for which his case should be very strong, The Sound and the Fury. Due to the fact that the narrator is an adult idiot, we have a case where the orderly presentation of events, the "plot" or "argument" as Handy would call it, seems singularly unimportant. Handy's claim that we are here presented with a "conglomeration" of Benjy's "disparate experiences," a "confused welter of kaleidoscopic experiences,"\textsuperscript{18} is more or less plausible. With such a work, it is at least initially plausible to say that we are presented with a composition of "textured" scenes, out of which we are invited to draw by means of their arrangement or "logic" some central principle of the ordering of experience. This principle, he says, is that though Caddie is to outward appearances "a fallen woman, driven by misshapen values to a life of prostitution," she is "also a saint."\textsuperscript{19} And Handy would certainly approve of our applying this to the actual world, ordering the welter of our experience by the "judgment" that profound forms of sanctity come in disturbingly perverse outer garments.

Thus Handy's approach is a singularly strong and interesting presentation of a formalist approach to fiction. But what we want to know is whether in practice he can arrive at his interpretation using only the tools of the hypothetical formalist. Of course, Handy is not here explicitly defending the formalist view that nothing from outside a work is needed for its interpretation. Nevertheless, we may suppose he holds this view, in that he calls his theory "formalist," and places
his theory in a tradition whose members explicitly hold this view (although, of course, whether they would allow themselves in practice to be held to the strong, literal position of our "hypothetical formalist" has been left open).

Has Handy succeeded in arriving at the alleged "judgment" on purely formalist grounds? Basic to Handy's interpretation is the claim that the reader becomes "aware of Benjy's pathetic situation as an adult idiot who possesses a touching, though limited, sensibility." Where do we get this information? Presumably we can see from the text that Benjy does not make many connections between the events he related, and that his central preoccupation is his sister. But where do we get the information that Benjy is an incompetent observer, and that his interpretations are not in accord with the "real facts" that he is presumed to be reporting on? Why not simply assume that Benjy is fully competent, that the world he reports on is chaotic? This assumption would explain the "data", and it is a good deal simpler than the assumption Handy makes. Handy's assumption must rest on the assumption that the most plausible expansion of the data given by Benjy in his disconnected way is a causally normal universe, one in fact, which has the socio-economic characteristics of the changing South. From this assumption we can argue that we must in fact be dealing with a truncated universe with an incompetent narrator--for Benjy's reports taken as "true" would not lead us to a causally normal universe. The basis for this assumption about the proper expansion of the data is, I suspect, that Faulkner nearly always creates a causally normal universe of such a kind. We know this from his earlier work. We use our knowledge of the
actual world and the conclusion that Faulkner is creating a universe of the same kind to ground our assumptions about the nature of the world Benjy is reporting on. And we then use these assumptions to draw conclusions about Benjy's competence as a narrator. It is our knowledge of the way the actual world is which shows us the respects in which the narrator is limited and the respects in which he can be trusted. And it is our knowledge of how the adult mind functions and how a child's mind functions which leads us to an exact estimate of the nature of his limitation. But again, there is probably more than one theory which explains Benjy's limitations. Why not hypothesize that he is an extremely clever madman masquerading as an idiot? But we know that Faulkner habitually experiments with limited narrators who are genuinely incompetent—and that he habitually invests those narrators with deep metaphysical insight. Thus in certain respects, we rely on Benjy. We rely on his basic truthfulness, although he habitually misinterprets or underinterprets the data. And we rely on his moral vision. Benjy sees his sister as good; we search through the facts he reports for a basis for this judgment and come up with the conclusion that Caddie has been able through her love to respond to Benjy as a human being, that she is a saint. But to arrive at this conclusion, we have had to rely on our knowledge of personal interaction in the actual world and upon our knowledge of Faulkner's usual mode of composition.

In short, the techniques employed in reading the novel are precisely those of assuming the world's causal normalcy, deciding that we have an incompetent narrator. We can then use the extensive psychological parallels between the world of the work and our actual world and
socio-economic parallels with the changing South, to estimate the limitation of the narrator, and finally to go on and draw complex conclusions about the world of the work. Contrary to what Handy seems to think, we cannot afford to ignore the "plot," in this case the course of affairs which make up the expanded universe, in drawing our conclusions about the work. For it is the world of the work which is the basis of the assumptions Handy starts with. In assuming that we have got the world of the work figured out beforehand, Handy ignores the immense amount of information which we must import into the work even to be able to read it. If Handy merely means that we can at our present state of criticism afford to take what we bring to the work for granted, well and good. I have merely given a clear procedure for making explicit what we may in practice take for granted. If he means, as I suspect he does, that such procedures are unnecessary, then I believe I have shown that they are not.

Thus, if Handy's position really is that of the hypothetical formalist, then it has been shown to have the characteristic weaknesses of that position. He would then have to be construed as denying the central importance of "what happens," with the result that his own assumptions are unjustified. In addition, once we have gone through the procedure of assuming certain similarities between the work and the actual world, we have available a way of explaining its cognitive import. We need not do so in terms of some special kind of poetic knowledge. We need merely point out that the work is a rich source of hypotheses about psychological truths and moral judgments about similar cases in the actual world. The alleged "special poetic knowledge" turns
out to be a superfluous entity. Of course, we might sometimes for aesthetic reasons take a step back and attempt to see the work "kaleidoscopically," as Handy suggests.

It is certainly true that the formalist theory advocated by Handy is simpler than the kind of theory I propose. It has only one thing to account for, at least superficially, i.e., the presence of a "judgment" or proposition in a fictional work. The theory I propose in chapters III and IV has two things to account for, i.e., both conclusions about the world of the work and the effect of the work of fiction on our beliefs about the actual world. This latter question I will take up in some detail in chapter V. The theory I am proposing is what has sometimes been called a "mimetic theory"; and mimetic theories are more complicated than their formalist counterparts. What I hope to have shown is that no formalist account which attempts to draw a judgment about the actual world out of a work of fiction can do so without relying on techniques such as those outlined in chapters III and IV. The position of the hypothetical formalist is simple, but only because it is incomplete.
Notes

1 I shall be dealing here with a formulation of a formalist position which says that in the interpretation or description of a work of fiction, we are restricted to elements present in the work itself. It should be noted that the theory as originally proposed by Wimsatt and Beardsley in "The Intentional Fallacy," reprinted in The Verbal Icon, ed. Wimsatt and Beardsley (Kentucky: University of Kentucky Press, 1964) 3-18, was stated as a theory about the evaluation of a work of fiction. "The design or intention of the author is neither available nor desirable as a standard for judging the success of literary art." (underlining mine). Later the theory is broadened from the consideration of the intentions of the author to include considerations related to the circumstances of the creation of the work, p. 10. But Beardsley in Aesthetics, 17-34, and in his other works holds the theory as an explicit theory about interpretation.

2 I will make some minor amplifications of the theory when it is explicitly stated in section 4.1.

3 A mere description of Roger Chillingworth's behavior is, of course, fairly interesting. This we can get from the basic description of the work. But the more interesting critical claim that in his personality, desire has been replaced completely by hatred is a critical claim which is not logically entailed by the "facts" of the work.

4 We ask, for example, whether Milly Theale at the end of The Wings of a Dove has really done one more deed of sweetness and light. All of the centers of consciousness assume this; but then, they are all a bit limited in their metaphysical vision. There is the intriguing possibility that the surrounding decadence has produced in her a really remarkable unconscious flair for the perfect act of permanent revenge, that for the other characters the sheltering wings of the dove will be an enduring curse.

5 By clause I of the truth definition in chapter III and by the definition of \( \phi(P_n,H) \) and clause (II,4).

6 We must assume, for example, that the critic understands and can recognize referring and non-referring uses of 'the'-constructions, e.g., 'The man on tenth street is ill' and 'The man kicked the bucket', respectively.

7 Some theorists would no doubt claim that we always appeal ultimately to observation. Others, no doubt, would claim that we can arrive at at least some assumptions by reflecting upon the structure of our experience. The point is that we assume we can objectively justify the assumptions by an appeal to other features of the world we live in or our experience of it.
8. We might argue, for example, that no one at the time the work was written had even considered a non-standard time conception. This sort of argument is allowed by even such strict anti-intentionalist theorists as Wimsatt in "History and Criticism" reprinted in The Verbal Icon, pp. 252-265.

9. This is somewhat similar to justifying the claim that someone is carving a figurehead (although it might otherwise just as well be a free-standing statue) because he is known to be working on what appears to be and what he claims is a Viking ship in his basement.

10. It has recently been argued on a much broader scale by Kendall Walton, "Categories of Art," Philosophical Review (1970), 334-367, and by William Lycan and Peter Machamer in an unpublished (to date) manuscript entitled "A Theory of Critical Reasons," that the formalist position is incorrect in the more general context of aesthetics. The basis of their claim is the argument that what distinctive aesthetic proper ties a work is said to have is a function both of its non-aesthetic properties (tones, colors, etc.) and of the category or genre we see the work as belonging to. I am in agreement with this position. Neither article makes explicit a point which I wish to make about epic literature and which is probably equally plausible for representative painting, that what we take to be the "facts" of the work are dependent on the kind of work we take it to be. (For example, a dark foreground means something quite different in a Flemish landscape and an impressionist painting). Both articles stress the importance of the circumstances of a work's origin in deciding what kind of work it is, and a fortiori, how its non-aesthetic properties are to be construed. It seems to me that both articles are more insistently than I would be upon placing the work in a specific genre, e.g., "representationalist," "surrealist," etc. I am more interested in the kind of work usually produced by the given artist or his circle or at a given time.

11. A theory which is explanatorily adequate need not necessarily explain all the critical sentences obtained by the procedures in chapter III as literally true. If the work contained a large amount of poetic language, particularly metaphors, this requirement would obviously be too strong. But if it cannot explain all the sentences of the deductively derived description as literally true, then an explanatorily adequate theory must furnish a systematic basis for deciding which of the sentences of the description are taken literally and which are not. For example, if a universe is taken to be normal in its logical constitution, then most metaphorical sentences will have to be taken non-literally. This problem, I suppose, could be used to pose one more unpleasant difficulty for the formalist, who simply has to assume that he always knows a metaphor when he sees one. Given the possibility of Kafkaesque worlds, it is difficult to see how he can justify such an assumption.

12 Ibid., pp. 95-96.

13 Ibid., p. 96.

14 Ibid., p. 97.

15 Ibid., p. 97.

16 Ibid., p. 97.

17 Ibid., p. 97. Thus we have here a cognitivist formalist theory of fiction which may not be, strictly speaking, a version of the proposition theory, despite Handy's use of the term "judgment." In this sense, the theory is unlike that advocated by Beardsley, as treated in chapter I, and like those of Eliott and Vivas. The sort of theory Handy proposes is usually proposed in the context of the criticism of poetry, but it would have to be shown that Handy really is guilty of assimilating fiction to lyric poetry and that this move is in principle objectionable.

18 Ibid., p. 101.

19 Ibid., p. 101.

Chapter V

In section 4.1, I gave a list of those types of criticism which I argued briefly are dependent upon a basic kind of criticism (type (i)). I attempted to show there that the formalist, construed as the "hypothetical formalist" is in no position to carry out this basic kind of criticism. If I am right about this, he is a fortiori in no position to deal with those kinds of criticism dependent upon explanatory criticism. I then sketched a non-formalist approach which, it seems, can account for the justification of explanatory criticism. In section 5.1, I will extend this non-formalist approach in detail to the other kinds of criticism listed in chapter IV. In section 5.2, I will turn to a specific dispute between non-formalist critics. I will attempt to show that the arguments used are seen to be valid, once their structure is clarified by means of the schematic account advanced in section 5.1. Once we have ascertained the structure of the arguments and decided upon their validity, we are free to turn our attention to the truth of the premises involved in each. It will emerge that the arguments on one side are basically sound, and that the arguments on the other side are fundamentally unsound. The techniques proposed in chapters III and IV in section 5.1 thus provide us with a fairly powerful means of interpreting and evaluating critical arguments. I do not claim to have given a decision procedure for evaluating critical claims about works of fiction.
A good deal will always rest on the truth of the premises, and in some cases we will no doubt simply have to admit that two interpretations of a given work are equally defensible. I am content with the less impressive accomplishment of having given a clear and comprehensive account of the "logic" of critical arguments so that we are in a far better position for evaluating them for validity and soundness than we would otherwise be. There may perhaps be kinds of criticism which are not taken into account here. If such examples are proposed, we will simply have to see if they can be fit into the schemata proposed here, or whether they perhaps demand some new kind of treatment. This task, however, is beyond the scope of the present chapter.
We now turn to an examination of the kinds of criticism which are dependent upon explanatory criticism (type (i)). In each case I shall try to indicate the respect in which such criticism is dependent upon explanatory criticism and give a schematic account of how argumentation should proceed in such a case, given that we now need not restrict ourselves to the arguments available to the hypothetical formalist.

(ii) Incompetent narrator. Let us take an extremely simple case of a truncated universe, one in which we somehow want to go beyond the narrator's presentation of the facts of the world, or the "facts" as filtered through the "center of consciousness." If we can give an account of how literary criticism should proceed in such a case, in the analysis and description of the work, then we will have a basic schematic account of critical argumentation for a wide range of works, including criticism of epistolatory works in general, novels with a deviant first person narrator, and nested narrative novels, and center and stream of consciousness novels. For despite the superficial disparity of these kinds of works, they have a common characteristic: Our access to the universe if filtered at least through the point of view of a character or characters, whether it is merely filtered through his apprehension or in addition through what he chooses to tell us, or whether the information is filtered through one character or through more than one.

I suggest that the general schematic form of our procedure is as follows. First, we decide whether to assume that the universe is causally normal relative to the actual universe G. The arguments for this are familiar. There will be some internal evidence for our claim;
we must at least be able to explain all the data. The explanation need not take all the sentences derived from the procedure in chapter III as literally true, but if some of them are taken as metaphorical, then we must be furnished with some effective way of differentiating literal and non-literal truths. In most novels, assumptions about the lawlike constitution of the work will rule out some sentences obtained by the procedure in chapter III as literally true. A fuller treatment of these critical claims will be given later in this section. In the case of Lady Susan, the dating of the letters follows an orderly sequence. And in this case, the short third-person section at the end confirms our assumptions that the letter writers are distinct characters related by the events they describe, that normal causal laws are certainly not obviously suspended. There is also strong external evidence that Lady Susan is causally normal relative to the actual universe. Jane Austen always constructs a universe of this kind, and the novel in question was written in a period when novel universes were, with a few exceptions, normal. In addition, we know that Jane Austen almost never "cheats," that her characters tend to be consistent and possessed of normal psychology. In choosing these assumptions we are obviously heavily dependent upon criticism of type (i). All these preliminary assumptions lead up to the crucial step in interpreting a truncated universe. We have to decide which of the sets of reports to trust—and how far. Having made certain minimal assumptions about the law structure of such a universe, we then construct a series of worlds $W_1\ldots W_n$, such that each $W_k$ contains the truncated universe $W$ and also a set of non-letter facts which form a "backdrop" for the letters. In the present case, our task
is simplified by our knowledge that at the time the novel was written, the general practice in epistolatory novels was to invest at least one character with the authority, at least as regards the veracity of his reports. We can thus set up the series as follows. One such universe, say \( W_1 \), is formed by taking "everything which every character says as true," by constructing for every letter-sentence \( a \) of the truncated universe \( W \) a non-letter sentence \( a' \) of \( W_1 \), such that \( a' \) is related to \( a \) by a relation \( R \), which, intuitively speaking, relates each \( a \) to an \( a' \) which is just like \( a \), but is not in an intentional context. In other words, in \( W_1 \), every character is taken as "telling the truth." We thus take \( W_1 = \{ p : p \in W \land (\exists s)(\exists s')(s \in R(s', s)) \} \). In another \( W_1 \), say \( W_2 \), we take all that Lady Susan's siter-in-law, Madame DeCourcey, says as "true." In \( W_3 \), we take all that Lady Susan says as "true," and so on. We then argue that one of the \( W_1 \)'s so constructed is the most plausible expansion of \( W \). Generally, this will be on the grounds that the expansion allows us to explain in the simplest and most coherent way all the data in \( W \). Again in the case of Lady Susan, our task is simplified. There are only two such sets of reports which are comprehensive, i.e., such that no one else reports on an event which is not reported on in the set in question. We thus have to choose between Lady Susan and Madame DeCourcey, for in the present case, if we were to choose \( W_1 \), we should have to take account of the fact that \( W_2 \) and \( W_3 \), which it contains, are at least superficially inconsistent. For example, Lady Susan's daughter is alleged by her to be stubborn, willful, and bad-tempered, whereas Madame DeCourcey claims the opposite. If we choose \( W_1 \), either the most plausible expansion of \( W \) is inconsistent, or
we must adopt some compatibility hypothesis, say that the daughter is schizophrenic. (In fact, we will need many such hypotheses, for the reports differ on many counts). Jane Austen usually does not create inconsistent universes, and her characters are ordinarily normal in psychology. And the set of compatibility hypotheses make \( W_1 \), if it is consistent, extremely baroque. If simplicity is ever a criterion of preference at all, it counts against \( W_1 \). On all these grounds we reject \( W_1 \) and search for some authority, a single set of reports which is consistent and complete. What about \( W_2 \)? Lady Susan is no doubt the most clever of the letter writers. But if we assume that she always tells the truth, we will have to contend with the fact that she on occasion contradicts herself, that she readily counsels her correspondents to lie, and that she at least once admits having lied. A paradox threatens. In addition, it is bad practice to trust a liar when there is an alternative. On these grounds, \( W_2 \) is clearly unacceptable. It is clear that \( W_2 \) is the best choice. We in effect choose to rely on the sister-in-law as the reliable reporter. It is the only other extensive set of letter-reports, invites no paradoxes, requires no extensive compatibility hypotheses.

Having made the choice of a plausible expansion of \( W \), we are in a position to use \( W_2 \) to draw some interesting conclusions about the limitations of the other letter writers. Now we have something to measure their reports against, i.e., \( W_2 \), which has been ruled the most plausible expansion of \( W \). Why do Reginald's and Lady Susan's reports fall to square with \( W_2 \)? We know from Madame DeCourcey's reports that her brother Reginald is not a liar. He errs in judgment, then; and he
errs systematically whenever he encounters the charming Lady Susan.
We can explain his distorted grasp of the facts by the hypothesis that he is young and fatuous, something of a sap. Similarly, Lady Susan's reports deviate systematically from \( W \), whenever the facts place her actions in an unpleasant light. The whole novel affords some interesting insights into the wide range of descriptions which the characters can give of a single state of affairs, and into the extent to which characters' motivations are opaque to themselves.

It should be obvious from the example and from the general discussion that in this crucial step of the interpretation of a truncated universe, we are extremely dependent upon explanatory criticism (type \((i)\)), both to decide what sort of world the characters are likely to be reporting on and to decide who is to be trusted. It may be objected that this procedure is far more complex than the way we in fact proceed. But this apparent over-complication is due mainly to the extremely simple nature of the example. In a complex novel, we often do go through these steps explicitly, although in the simple cases we no doubt short-cut them and seem to ourselves almost to proceed without making the decisions involved. But even in the simpler case, if we are to defend an interpretation, we have to proceed in this way. This way of presenting the procedure for criticism of truncated universes has the distinct advantage that it moves the discussion from the amorphous plane often found in literary criticism to arguments which are much easier to evaluate. And even the simple example brings out the three crucial areas in which we may expect to find critical disagreement about a complex work. First, there may be questions about the kind of universe we
are entitled to assume we are dealing with; in this case, we have argued for various reasons that the universe is causally normal relative to G. Secondly, there may be a good deal of question about the interpretation of the data of the truncated universe. In the case of Lady Susan, there is no great problem in this respect, due to special features of the case. And finally, there may be considerable disagreement about which of the W_i's constitutes the most plausible and effective expansion of the universe. With a more complex work, any or all of these may be open to debate. But the procedure proposed has the advantage of showing that not all truncated universes are significantly interpretively ambiguous. In many cases, there is one interpretation which is clearly the best one.

None of these debates is yet directly about the aesthetic value of the work of fiction or about its role in contributing to our knowledge of the actual world. We are still at this point very much involved in theorizing about the constitution and the higher level descriptions of the universes in question. The important point to note is that, even at this point, we have already gone far beyond the strict "internalistic" justification of critical claims. We are not claiming that such critical claims are true because they are members of the set of sentences assigned the value "true" on the basis of the recursive truth definition of chapter III or on the basis of the "meanings" of the expressions. In criticism of type (i) and even more radically in criticism of type (ii), we are dependent upon factors external to the work for our critical judgments about it.

(iii) Inconsistent Universe. We have now treated fairly extensively
the problem of expanding a truncated fictional universe. A related problem is the critical treatment of universes which are or appear to be inconsistent. There are several ways in which a fictional work may appear to be inconsistent, and the critical treatment of the work varies accordingly.

**Case A:** In some cases, the basic or deductively derived description of the universe is logically inconsistent relative to quantified S5.

**Case B:** In some cases, if we assume that certain laws hold in the fictional universe, then the laws together with the deductively derived description are inconsistent or lead to inconsistency. In some such cases, the assumptions about the constitution of the fictional world are clearly warranted for some reason or other. Faulkner, for example, sometimes gives two different dates to the same event or presents a single object as being in two places at the same time. But the assumption of normal laws of time and space and physics in such a universe is clearly warranted.

**Case C:** In some cases an inconsistency results from the data plus assumptions which are pretty clearly not warranted. We generally assume, for example, that the sequence of moments of time is like a number line, such that the temporal moments follow in sequence, and no moment state is (partially) determined by its own future (or, what amounts to the same thing, that no moment has (a part of) its history as its past). This is to assume, in formal terms, that the temporal model is linear, reflexive, transitive, and non-clustered. Let us term these "the normalcy assumptions." But on this kind of assumptions, Heinlein's *The Door into Summer* would become inconsistent, for in this novel, there
are intervals of time where, because of time travel, the future influences what the past had to be. In this case, the tensed sentences of the novel determine a tense model in which, at some points, time becomes circular. The normal logic of linear time (S4,3) is weaker than the logic of circular time (S5); the stronger system validates all the inferences the weaker one does, and some others besides. Hence we will not get any inferences we do not want by assuming that the weaker logic is the logic of time throughout, although we shall be unable to get at some of the inferences we are entitled to over the circular interval. But as soon as we actually make explicit the normalcy assumption, that time is linear throughout, then a straightforward incompatibility will arise. A slightly different kind of case is exemplified by Heinlein's Farnham's Freehold. Time in this novel is pretty clearly branched at two points. Some of the characters are thrown forward in time. Their appearance in a possible future state of the world $f_0$ (which we presume was destined to be actualized) changes the states following $f_0$, so that the possible future state $f_1$, which would otherwise have been actualized, does not become actual. Instead, another state $f_2$ is actualized. In $f_2$, the time travellers are sent back in time to a point slightly before their original point of departure. Let us call the point to which they are sent back $P_0$. But their arrival at this point diverts the actual future of the world from the possible course of events which would otherwise have occurred (that leading to $f_0$) to another possible course of events (which, we presume, will not lead to $f_0$). Thus at certain points, $P_0$ and $f_0$, time branches; at such a point, the future may go one way, or it may go
another. The logic determined by the branching time model is $S^4$, and this logic is weaker than what seems to be the normal logic or garden-variety unbranched time, i.e., $S^3$. If we were to make the normalcy assumption here, then the novel would become straightforwardly inconsistent.  

In each of these cases, we have to decide how to treat the fictional universe. We have to determine whether we really have an instance of case (A), that we are not, for example, dealing with an apparent incompatibility due to a shift in point of view. (If this were the case, we would construct our criticism along the lines of the truncated universe lines). If we really do have an instance of case (A), then I think we must proceed as follows. We must first decide whether we should treat the work as a work of fiction in our sense. If a work is not consistent, then the theory I propose simply is not applicable. We may argue for treating the work as a work of fiction on internal grounds. We may argue, for example, that the inconsistency is minor, that the rest of the work is consistent. Or we may argue in favor of treating the work as fiction on external grounds, for example, that we know that the author intended to write a work of fiction of this sort, that the work is the third member of a trilogy whose earlier members are logically consistent, that there may be an inaccuracy in the transmitted text, that the work was written at a time when a subtle error in logic would not have been noticed. If we decide to treat the work as fiction, then we must come to grips with the fact that consistency is one of the necessary conditions of the work's being a work of fiction. To be able to handle the work, we must adopt a shrinking procedure which
is the reverse of the universe expansion procedure. We set up a
series of shrunken universes $W_1, \ldots, W_n$, where each $W_i$ constitutes a
proper subset of $W$ and is consistent. We then argue that some of the
$W_i$'s is the most plausible shrinking of $W$. Such arguments will rest
on familiar sorts of considerations; that $W_i$ will be the most plausible
shrinking of $W$ which involves the least reduction of $W$ and which is
itself susceptible to the most plausible explanation. If one of the in-
consistent sentences seems to be relatively independent of the rest of
the work, whereas another of them figures in a crucial incident, then
it stands to reason that it is the first of them which must go. In
some cases, we may find that although $W$ itself is hopeless, it has an
interesting, consistent subset, say a subplot. In such a case, we can
merely make the subplot an honorary member of $K$ and operate upon it as
if it were a genuine, complete fictional work. In such cases, we are not
somehow arriving at the "real" fictional universe, which the author has
misdescribed. The shrunken universes are members of $K$ only by courtesy,
although they are not, strictly speaking, related to $G$ by $R$. This fact
will be relevant when we evaluate the success of the author's actual
fictional speech act; strictly speaking, he has failed. But the shrinking
procedure enables us to explain a fairly common critical practice, i.e.,
"disregarding" minor inconsistencies. And this kind of procedure is not
anomalous in the general context of speech acts. It is rather like
that situation in which a speaker, strictly speaking, fails to refer,
but nonetheless succeeds in picking out an individual for his hearer—
because the hearer corrects for the error. In such a case we would not
say unqualifiedly that the speaker had succeeded in referring. But what
he is doing has been understood, despite his incompetence.

In cases which fall under case B, we have supposed that our assumption of normal laws is justified on some grounds or other. This phase of criticism is clearly dependent on explanatory criticism. We have no reason, for example, to suppose that Faulkner has nonstandard time logic. The events of the novel for the most part determine a standard model. His other novels have a normal linear time structure for objective time. In such a case, we make our normal assumption about the logic of time and conclude that the resulting world description is inconsistent. We are thus dependent upon explanatory criticism to decide whether such a work is plausibly construed as inconsistent. We then proceed to apply the universe shrinking procedure in the manner already discussed.

In cases which fall under case C, the inconsistency results from our assuming a normal tense logic in a situation where this assumption is unwarranted. If we assume in The Door into Summer that time is not circular at any point, we leave a good deal of the data unaccounted for. And in Farnham's Freehold, we are not allowed to make the assumption that time is linear. In both cases, the events of the novel determine a tense model which is not linear and non-circular, i.e., not the normal conception of time. And these models determine tense-interpreted modal systems which are themselves consistent. (We may, of course, complain that the novel stretches our credibility to the breaking point. But neither universe is logically inconsistent—although both are distorted). There is also external evidence for disallowing the normal time assumptions. Heinlein has written a series of novels which are plausibly
construed as having nonstandard tense logics, and he openly designates his work "science fiction." These are additional reasons for considering the universes as distorted, not as inconsistent.⁴

This schematic account shows, I think, the way in which we should proceed critically when presented with a fictional universe which appears to be inconsistent.

(iii) Character comparison. A question arises about the treatment of critical claims such as:

(1) Hester Prynne was wiser than Anna Karenina.

in which we compare or contrast characters from different fictional worlds. On an intuitive level, there is a clear sense in which (1) might be claimed to be true. One could say that there are certain similarities between the histories of Hester Prynne and Anna Karenina. Both desert a husband for an unreliable lover, etc. Hester's greater wisdom, it might be said, lies in recognizing that the main significance of her action is that it establishes her individuality, whereas Anna does not realize this. Yet we note that on the basis of the semantics outlined in section 3.2, sentences like (1), if we try to assign them a truth value at all, are always false. The valuation function \( \phi \) assigns an extension to \( P^n \) such that for every assignment \( a_1, \ldots, a_n \), \( \phi(P^n, H) \subseteq (\Psi(H))^n \) and \( \phi(P(a_1, \ldots, a_n), H) = T \) only if \( \langle a_1, \ldots, a_n \rangle \in (\Psi(H))^n \) and \( T_i \) only if \( \langle a_1, \ldots, a_n \rangle \in (\Psi(H_i))^n \) for some \( H_i \in K \), \( H_i \neq H \), \( H_i \not\in H \). \( \phi(P(a_1, \ldots, a_n), H) \) will never come out true or fictionally true if \( a_1, \ldots, a_n \) are not in the domain of some single \( H \in K \). The restriction seems in order. The domains of fictional worlds are disjoint, and it makes no sense to say that their members are related to one another by a
predicate such as 'helped'. It seems quite correct that \( \phi(\text{Sidney Carton helped Anna Karenina}, G) = F \). But to claim that sentences like (1) are always false if they are evaluated at all is to deprive ourselves of a good deal of useful criticism. The kind of predicate in (1) seems to give (1) special status, and it seems if we could once get over the separate-world problem, (1) would be both significant and true. The solution to the problem thus seems to lie partly in the analysis of sentences containing a comparative predicate. This much we can borrow from normal narration. In addition, we will need some technique for dealing with two fictional worlds at once, so that (1) will not automatically be assigned the value "false." This technique I will call "amalgamation."

We first consider how we assess the truth of comparative sentences in normal narration. Although no treatment of such sentences at present seems conclusively correct, some features of such expressions point to a possible approach. In the first place, it seems clear that there is a group of predicates, e.g., 'large', which are significantly applied to a thing only relative to an explicit or implicit specification of a class of things to which the thing in question is compared. Such predicates have been called "relative modifiers." This peculiarity is marked by the fact that the inference from:

(3) Tom is a large featherweight.

to:

(4) Tom is large.

is illegitimate, and by the fact that (4) alone does not tell us much about Tom in the absence of knowledge of whether Tom is, say, a baby or
an adult. To say that Tom is large, it seems, is always to say that he is large with respect to a class of comparison $\mathbb{C}$. This I will express by 'Tom is large$_c$'. Perhaps the characteristics of relative modifiers can be used to approach comparatives in general. In this terminology,

(5) Tom is larger than John.

means simply that Tom is large with respect to some class of comparison to a greater degree than John is large with respect to that class of comparison, (5) can be rewritten as:

(5') Tom is large$_c$ to a greater degree than John is large$_c$. Of course, 'to a greater degree' is itself a relative expression, but it can presumably be spelled out by ' $>$ '. ' $>$ ' in turn can be explicated so as not to involve a comparative. This suggests that we might try a similar analysis for predicates whose positive degree is not clearly one of the relative modifiers.

(6) Tom is wiser than John.

would become on analysis:

(6') Tom is wise$_c$ to a degree greater than John is wise$_c$. For modifiers like 'wise' at least, Tom and John must be taken as both being in $\mathbb{C}$. (We cannot very well allow someone to pick the class of ants for a class of comparison, and then, because both Tom and John are wiser than all the ants, conclude that (6) is false). The sense of (6'), then, is given by:

(6'') Tom is wise-for-a-$\mathbb{C}$ to a greater degree than John is wise-for-a-$\mathbb{C}$.

Such an approach depends upon our being able to find a way in which the predicate in question can be considered quantitatively. This may well be a fatal weakness. In addition, the notion of "being-wise-for-a-$\mathbb{C}$" may
merely locate the difficulty without solving it, but it will do as an example for the next step in the analysis of (1).

For even on this analysis, (1) still presents a problem. The rewriting of (1) will be:

(1') Hester is wise-for-a-C to a greater degree than Anna is wise-for-a-C.

Intuitively, (1') might be true if Hester turns out to be very wise for a woman, and Anna not very wise at all for a woman. The problem is that on the basis of the semantics of chapter III, even if we determine the degree of wiseness for each character with respect to the novel in which she occurs, (1') will still come out trivially false, since (1) is about characters from different worlds. There were good reasons for this restriction, but now, it seems, so long as we make the different origins of the characters explicit, we should be able to assign a truth value to (1). In fact, what we will do, heuristically speaking, is to combine the two worlds, making the combination explicit. This procedure I shall call "amalgamation."

We proceed in the amalgamation of $H_1$ and $H_j$ as follows. We first combine the domains $(H_1)$ and $(H_j)$ to get a new domain $H_k = H_1 \cup H_j$. We then consider the set of sentences $H_k = H_1 \cup H_j$ as a world. In some cases, we will have to apply the shrinking procedures already discussed to render $H_k$ consistent, for although $H_k$ is not, strictly speaking, a fictional world, it must function as a world, i.e., it must be consistent. In the present case, there is no shrinking to be done. It is hard to see how a conflict in the non-modal sentences of $H_1$ and $H_j$ could occur. And $H_1$ and $H_j$ belong to the subset of $K$ whose members are causally possible relative to $G$. They have the same constitutive laws.
The most difficult cases to consider will be those worlds which conflict with respect to their constitutive laws. We can, it seems, amalgamate nearly any two universes, even if they have radically different law structures. We can either relax the laws of the stricter universe (e.g., we can allow magic in *Madame Bovary*) or impose stricter laws on the looser universe (e.g., we can make *The Turn of the Screw* naturalistic). Sometimes there is a special trick by which we can amalgamate the universes, e.g., since the hero of *Nausea* allows that a Kafkaesque world is a possible world relative to his own and predicts how he would act in such a universe, we can take the Kafka world as a possible world relative to his and see whether his predictions about his own and human action in general are borne out in it. The most important danger is that in changing the structure of a world, certain significant attributes of the characters disappear, making a comparison with a character in the changed world useless with regard to the character as originally presented. This problem is most acute if the characteristic which disappears is important for the specific comparison which we want to make. If, for example, we were to make the world of *Dr. Jekyll and Mr. Hyde* naturalistic, it is hard to see how we could assess Jekyll's moral guilt, for we should just have eliminated the possibility of his offense.

We will not encounter this difficulty often. In most of the cases we are interested in, it is precisely the similarities of the universes which lead us to make comparisons of the characters. We want to compare, say, the courses of action that the characters elect to follow in situations which are similar in relevant respects.
The gain from the amalgamation procedure is in being able to do for characters in different universes what we can do with two characters of the same universe, compare and contrast them with regard to predicates which are significant. Once the amalgamation is carried out, we can consider (1) to have the value \( T_k = T_1 \cup j \). It is interesting to note that predicates such as 'helped' usually fare no differently in amalgamated universes than on the account of chapter III. (The only exceptions are probably amalgamations of worlds with overlapping domains which are the works of non-co-operating authors, e.g., the Oz Books). We do not add sentences in amalgamation, and there will be no sentence of \( H_k \) linking characters from previously non-overlapping domains by such predicates as 'helped' (with the exception noted). One way to look at this situation is to say that \( \emptyset('helped', H_k) = \emptyset('helped', H_1) \cup \emptyset('helped', H_j) \). If \( a \in \psi(H_1) \) and \( b \in \psi(H_j) \), then 'a helped b' will not be in \( H_k \) unless it is in \( H_1 \) or \( H_j \).

Once the amalgamation has been completed, we can very well compare two individuals from \( H_k \) with respect to the degree to which they are large or wise and assign a truth value to the sentence of criticism in which this is done.

(iv) Poetic language devices. Criticism concerned with poetic language devices, i.e., with the use of simile, metaphor, symbol, and archetypal patterns, is probably the area of criticism in which there is the greatest methodological confusion. The aesthetics of the work of fiction is not my primary concern here. Hence my treatment of this problem will be brief. I will here be dealing with claims about the uses of poetic language devices in a work of fiction which have to do
with the work of fiction as the semantic basis of criticism. I will be dealing specifically with three kinds of poetic device: explicit simile, explicit metaphor, and functional metaphor or symbol.

In the following passage, we have explicit similes (underlined portions):

She sank down on the couch with an easy motion, vaguely suggestive of a flutter and preening as when a bird touches a bough. Her hair was yellowish, like metal, with grey in it now, but the grey was metallic too, like spun metal woven and coiled into the yellow.

The simile does not present us with much of a problem semantically speaking. What we are told is that some object or event A is like some other object or event B in some respect C (which may or may not be partially specified). This information we may take directly and literally into a state description of a work under clause I of the truth definition in chapter III. Although we might be troubled by the report that Jack's mother was exactly like a bird, there is no such problem with being told that she was like a bird in some respects. The one problem with the interpretation of the information conveyed by a simile is that it is underdetermined in the following sense. When we are told that Jack's mother's hair was yellowish, like metal, we are still not quite sure what quality is being attributed to the color of the hair. Perhaps it is the tone, although here we still have our choice between a dull and bright finish. Similes often remain partially indeterminate in this sense. Of course, the simile in question might be made determinate by the inclusion of a report to the effect that her hair sparkled placed later in the passage. In this sense, the determinate meaning of a simile is subordinate to the non-simile facts of the work, though it
may add detail and vividness to those facts, once they are established.

Explicit metaphors do present us with a problem semantically.

Let us assume that we have decided that Jack, the first person narrator, is to be trusted at least with the veracity of his reports, and that he says that his wife Lois is a cross between a Georgia peach and a carnivorous bivalve. Given that this is a report, it should go into the basic description of the world under clause I of the truth definition, although it need not be taken in literally. Taken literally, the report is odd. In such a case, I would maintain, we have a choice. We can take the report literally. We could say that Lois managed to pull off the marriage in her woman state, but that Jack never quite knows what he will come home to, to Georgia peach or a Venus flytrap. Or we may take the report metaphorically. If we choose this latter alternative, I suggest that for semantic purposes, we reduce the metaphor to a simile and say that Lois is in some respects like a Georgia peach and in some respects like a Venus flytrap. In making the choice between a literal and a metaphorical reading of the report in question, our decision about the kind of universe we are dealing with is important. If the universe in question were Kafkaesque, the literal reading might be quite acceptable. But All The King's Men is excessively, depressingly normal. And for this reason, we will no doubt decide to read the report metaphorically, i.e., as a compressed simile. The determination of the simile will depend on the interpretation of the non-simile facts which form its context.

We now come to the case of the functional metaphor, or what is sometimes called "the constitutive symbol," i.e., one or more persons or events or objects which have a legitimate place in the narrative, but
which, it is claimed, also embody or symbolize certain features of
the work. Consider the following two passages from *All the King's Men*:

"She had slid down a little in the seat, with her head lying on the top of the back cushion so that now she wasn't looking toward the horizon, but up into the sky with the moonlight pouring down on her face to make it look smooth as marble in the moonlight. And how her hands lay supine on her lap, the fingers curling a little as though to receive a gift."

"With me is my wife, Anne Stanton, and the old man who was once married to my mother... Anne and I shall never live here again, not in the house or at the landing. (She doesn't want to live here any more than I do.) She has let her place go to the Children's Home she was interested in. The gift of the house was finally her gift to the ghost of Adam, a poor gift humbly offered, like the handful of wheat or a painted pot in the tomb, to comfort the ghost."

No doubt the reports are part of the literal data of the work. But it might be claimed that the passages in question have another function. In the first passage, Anne is described in terms which suggest the balanced repose and untouchableness of classical statuary (a context in which Jack has earlier placed Anne explicitly). Anne Stanton is at this point untouchable and immeasurably removed from Jack, or at least he sees her in this way. In the second passage, she is no longer at a distance, and her repose and untouchableness are gone. The juxtaposition of the scenes, it might be claimed, reinforces an important central theme of the book, that the downward path to wisdom involves a messy, final dropoff. We have plenty of non-poetic evidence for this generalization. Nearly everyone in the book who progresses at all toward understanding comes downward with a nasty crack. But the different descriptions of Anne reinforce this theme powerfully.

Doubtless there are better examples of functional metaphors. But this example and its provisional interpretation display the main feature
of such devices. We first have to decide whether the alleged functional metaphor is to be read metaphorically or whether it is to be read merely literally. If we do decide upon a metaphorical reading in addition to the literal reading, then, I suggest, we should treat it as a simile for semantic purposes. In the above case, for example, we might say: In the earlier episode, Anne is like a classical statue for Jack, in the later scene, she is like a worshipper before such a statue. We will then have to decide in what respect she is like what she is compared to. Functional metaphors are thus indeterminate in a twofold way. They are a result of a prior decision to read the element metaphorically, and the sense in which the element is like what it is supposed to resemble is indeterminate. Their determination is effected by a consideration of the rest of the work, although they may convey information in addition to that conveyed by the rest of the work. A minimal condition for the interpretation of such a device is that it should not conflict with what is taken on the basis of clause I to be basic data in the literal sense. If Jack had said elsewhere that in the earlier episode Anne seemed to him contemptible, then the above interpretation of the whole work and the interpretation of its functional metaphors would be incorrect. In addition, the interpretation of the whole work and the interpretation of the functional metaphors should be mutually reinforcing. This has two consequences. First, if an already plausible interpretation is supported by an independently plausible reading of a functional metaphor, this is an additional point in favor of the interpretation. Second, if a reading of an alleged functional metaphor conflicts with an extremely plausible interpretation of a work,
it is probably better to construe the alleged functional metaphor literally or to find another reading. This inequality arises from the fact that similes, and hence the metaphors which are in effect reduced to them, are ordinarily indeterminate in a way that interpretations are not.

Critical claims that a work exemplifies an archetypal pattern or has a traditional mythic plot often have the same kind of force. Of course we can conclude on our own from the "facts" of Bulgakov's *The Master and Margarita* that Margarita is an inessential moment of the master's passion for truth and an essential moment of his eventual salvation, but the fact that we have here a clear version of the Faust legend reinforces these conclusions.

(v) **Cognitive significance claims.** We now turn to those critical claims which we called "critical significance claims." These have the form 'It is (part of) the meaning of (moral of, theme of) the work that $p'$, where $p$ is some claim about the actual world. This is the problem with which we began in chapter I. And if we can settle the question of the basis from which such critical significance claims are made, their logical relationship to the semantics of criticism, and how they can be defended, then we shall finally be in a position to explain the sense in which novelists "tell the truth."

On the surface, the problem is this. The truth values of sentences of criticism arrived at by a complete semantic description of the work as set out in section 3.2 are subscripted to, or true of the work itself, not of the actual world, if they are true at all. No matter how many connections of predicates and lawlike regularities we have transferred
from the actual world to the fictional world on the strength of the assumption that they are in some respects alike in structure, the conclusions we draw by induction about the fictional world, the expansions and shrinkings we adopt, and the descriptions and conclusions we support by an appeal to poetic language of the work, still tell us something about the work, not about the actual world. On the other hand, the expansion and clarification of knowledge we ultimately claim to have gained from a work of fiction is related to the actual world.

A second problem arises out of the fact that what the novelist does, no matter how much he relies on the features of the actual world, is to create a fictional world. On the other hand, when we say that the novelist "tells the truth" (or "is wrong" about the way things are), we clearly mean that he has conveyed information or misinformation about the actual world. And a good many novelists, although they class themselves as artists, not as orators, i.e., admit in effect that what they do is to create fictional worlds, take themselves to be instrumental in influencing their readers' knowledge of the actual world. Indeed, in many cases, they allege that that is why they write.

This was the problem of chapter I. But now, I think, we are rather near to the answer to the problem. We have noted that certain groups of works of fiction, "causally normal universes," have significant parts of their law structure in common with the actual world, and that certain works, "distorted universes," involve departures from the laws we take to characterize the actual universe. In addition, certain novels contain fictional names and definite descriptions which have "usual connotations" in the actual world. And some have plots which are imposed either by
the facts of history in the actual world or by what is supposed to be an archetypal representation of the real pattern of human affairs or the cosmos. In reading a work, we have to note the similarities and dissimilarities between the fictional work and the actual world even to be able to read the works in question competently. Once having made our decision about the laws which govern a fictional universe, we go on to draw conclusions about the fictional universe much as in the actual world, conclusions about the motivations of characters, about the way a kind of character will act and interact with other characters, conclusions about the effects of the social structure or structures portrayed, in some cases conclusions about the nature of evil, the nature of art or beauty, etc. In reaching such conclusions about novel universes, we have no choice but to rely on relationships between the fictional universe and our own. Having noted this, we are inclined to reason that if in a universe governed by a particular set of "rules," certain conclusions can be drawn from a body of "data," related conclusions might be drawn from or might explain more complex bodies of data in situations of the actual world which are somehow related to the fictional universe in a specified way. In a sense we reverse our reasoning in figuring out a fictional universe, and "transfer" to the actual world the conclusions we draw about the novel universe as hypotheses for explaining complex bodies of data in the actual universe.

The value of a work of fiction in expanding and clarifying our knowledge of the actual world lies precisely in the fact that certain conclusions reached easily in essentially simpler fictional universes are treated as hypotheses about the actual world. Our actual application of conclusions
about the fictional world to the actual world is a matter of psychology, not logic. The question is why the conclusions we draw about fiction are so important in shaping our beliefs about the actual world. The answer, I think, is that the work of fiction serves as a very clear illustration, almost as a thought experiment, about "what would happen if such and such were the case."

In some cases, it is the similarity of the novel universe and our own which makes the difference. In Bleak House, the slow moving wheels of probate "justice" bring everyone who is touched by the legal process to grief; no such person remains unscathed. By simple enumerative induction we can draw the conclusion that in the world of Bleak House the institution of probate justice harms those it touches. Then, because the world of Bleak House is causally normal relative to the actual world and shares its socio-economic features with England of the 19th century, we can pose as a hypothesis about England of the 1800's the claim that England's system of probate justice generally harmed everyone concerned by it. The fact that the characters who come to grief are of widely divergent positions in society reinforces the induction. The entire work functions as a kind of thought experiment.

But with some works we have no such string of cases. We draw a conclusion, it seems, from a single character. Isabel Archer, for example, is the one character in The Portrait of a Lady who is "given" practically unlimited freedom of choice and the wherewithal to implement her choices. What she does, of course, is to mistake the appearance of spontaneous choice for freedom, thereby losing her freedom. But the hypothesis we pose about the actual world is not a hypothesis about
Isabel Archer, or even the weak hypothesis that sometimes people throw their freedom away. We pose the more general hypothesis that it is likely that someone in a position like Isabel's will throw freedom away. Certainly we do not infer this general conclusion by induction from the world of the work; or at least we should not, for induction from a single case is not in general a good practice. I should maintain that the conclusion we draw about the novel itself is a conclusion about the central figure, that we do not invalidly draw the general conclusion about the world of the work from a single case. But the hypothesis we transfer to the actual world is a general one. Why do we feel free to pose a general hypothesis about the actual world on the strength of one fictional case? The answer, I think, is that Isabel and the course of her life serve as a slightly different kind of "thought experiment." We are clearer about Isabel's history than we generally are about people in the actual world. For instance, we "have access" to most of her private conversations. We are much surer about her thoughts and motivations than we would ever be about those of an actual person, for we are given the information in question as a matter of unquestionable fact, whereas in the actual world, probably in our own case, we would have to arrive at such information as a conclusion from behavior, introspection, etc. Moreover, the information we have to deal with is limited by the "data" of the novel and whatever explanatory hypotheses we have adopted. It is as if we had said, "Let us take a clear-cut case. What would we say in such a case?" Our conclusion about Isabel Archer has the force that the conclusion of such an experiment would have. We pose a general hypothesis which we then proceed to test as we would test any other
general hypothesis.

The author has "told the truth" in the novelist's sense if the novel suggests hypotheses which are validated by their efficacy in dealing with the data of the actual world when they are posed as hypotheses about the actual world. He has not, of course, told us directly any of the conclusions we draw about the work, i.e., he has not made a statement. He has merely furnished a universe in which the careful reader can hardly help drawing certain conclusions, and has made the application to our own universe obvious and tempting. The selection of novel facts and the similarity of the actual world and the novel universe are the novelist's tools for manipulating the beliefs of his readers.

The manner in which we decide what hypothesis to transfer back to the actual world depends also upon whether the fictional universe in question is normal or distorted and upon the sense in which the universe is distorted. So far we have dealt with cases in which the fictional world in question is very much like the actual world. It would be quite reasonable for a critic to respond at this point that although it is easy to see how we might readily use fictional universes like ours as a kind of thought experiment about the way the world is, there is no reason to suppose we would feel impelled to do this, if the fictional universe were radically unlike ours, if the universe were distorted.

The situation in a "distorted" universe is somewhat different. These universes are taken to be unlike ours in their underlying constitutive laws. Distorted universes may expand and clarify our knowledge of the actual world in several ways.

We may encounter a universe which is distorted in some sense, but
is to a large extent normal. Heinlein's *The Door into Summer* is such a novel. Great pains are taken to show that the universe is causally normal in most respects. The biochemistry introduced in explaining "the long sleep," the process by which a person is frozen, stored in a frozen state, then thawed at a later date; the computer science involved in explaining how the household helper robots function; and the detailed examinations of some of the characters' psychology, exhibit the extent to which the fictional world and the actual world have common laws. There is just one distortion—people can drop back in time, and sometimes their doing so is a necessary causal antecedent of the future state from which they dropped back. (There is the hint that this happens periodically, that one Leonard Vincent dropped back from about 1970 to become Leonardo de Vinci.) The interesting feature of the novel is that the distortion involves no manifest conceptual impasse in a world in which most of our general assumptions about the actual world hold. When we transfer the hypothesis about time structure back to the actual world as a hypothesis, we find, as expected, that it is compatible with most of our general assumptions about the actual world, and that with a few conceptual adjustments, we could accommodate the new hypothesis. The new hypothesis would have some explanatory power with regard to geniuses who are said colloquially to have been "ahead of their time." Here, there is a clear, although perhaps playful, suggestion that a hypothesis which we have heretofore considered outlandish may well be true. Such distorted universes may be called "revisionistic." There is at least a jesting recommendation that we incorporate a new hypothesis into our world view, and a thought-experiment demonstration that we could do so at
no great expense and with some small gain.

But not all distorted universes are revisionistic. Consider Kafka's *Metamorphosis*. In the novel, Gregor Samsa wakes up one fine morning to find that he has somehow acquired the body of a bug, though his mind, his memories, etc., have remained intact. In the course of the work, the bug's body gradually makes inroads upon the originally human mind. Gregor begins to like the sort of filth bugs eat and to have the urge to crawl the walls incessantly. As those around him react to him increasingly as a bug, he loses the will to resist these inroads. He dies a bug, with no shred of human dignity intact. Certainly the hypothesis we transfer to the actual world is not that if someone becomes a bug in body, he will eventually become a bug clean through. Such a hypothesis would be extremely uninteresting, since no one, we presume, ever does become a bug overnight. Nor, I think, need we take the *Metamorphosis* as offering a dismal revision of our thinking about the world. We need not entertain seriously the dismal possibility that we may wake up bugs tomorrow; no one ever does. The point is rather that if someone were to acquire a bug's body, he would eventually cease to be human at all, due to the inroads of his changed body upon his mental processes and the inability of others to respond to him as anything as a bug. Or we may state the point positively as a lawlike generalization: Possessing a human body and receiving interpersonal recognition are necessary conditions for being a human being. I shall say that such a universe gives rise to a hypothesis "counterfactually."

But not all counterfactual distorted universes function in this way. In *Farnham's Freehold*, it becomes clear that if we accept the law
that time is branched, the result is that another law no longer holds, the law that experience is experience of the actual. Time branches because the characters experience a future of the world which, because they experience it, will never be actualized. Surely we cannot accept the branching time hypothesis as a revision; the cost is too great.

The counterfactual claim, that if the universe were structured in this way, experience would not be restricted to the actual, is interesting in its own way. But there is something else of importance. The universe of *Farnham's Freehold*, though it is perhaps only logically possible, stretches credibility to the utmost. Nonetheless, in that universe, provided we assume that normal laws of human psychology obtain, it is true that power corrupts in predictable ways. Thus, we are able to say that, even if the universe were radically different from the way it is, if normal laws of psychology held, it would nonetheless be true that power over other human beings corrupts the person who wields the power, that this truth of human nature is independent of the metaphysical constitution of the universe to a great extent. In this case, what we have been "told" is that a certain connection of truths is independent of other factors, something which we might not otherwise have suspected. The universe is counterfactual, but it is "eliminatively counterfactual," i.e., it shows that even if things were different, certain truths would still hold.

In revisionistic universes, the author may be said to have "told the truth" in the novelists' sense, when we conclude upon examination or reflection that the "distortion" might be adopted without loss of most of the laws we generally assume, and with some gain in explanatory power.
In the counterfactual universe, he has "told the truth" if the connection he displayed and which we formulate counterfactually proves to be tenable as a law upon examination. In the first case, the novelist has "misled us" if we discover that the suggested revision has hidden costs when posed as a hypothesis about the actual universe. In the counterfactual case, he has "misled us" if we discover that the connection does not hold, if we discover, for example, that normal laws of psychology could not hold in the universe Heinlein describes.

This, I think, exhausts the sense in which a critical significance claim is tempting. We in effect short-cut the complicated process in which we first conclude certain things about the fictional universe, then transfer them back to the actual universe.

Since the connection between the fictional world and our beliefs about the actual world is at least to some extent psychological, not evidential (unless we construe "evidential" so broadly as to include thought experiments under it), it may be asked whether my position differs from the emotivist position described in section 1.2. The answer is that it certainly does differ from that position in significant ways, although there are some similarities between the two accounts.

First, the emotivist position as defended by Richards and Stevenson offers no systematic way of delineating claims true about the work itself. Indeed theses (E₂) flatly denies that this is possible. The proponent of (E₂), it will be remembered, maintains that discourse about a fictional work is itself expressive language. He does not mean by this merely that critical descriptions and interpretations typically have the special function of guiding or directing our apprehension of a work
or that such discourse sometimes conceals an evaluation of a work. I would agree with both of these weaker claims. But (E2) is instead the very strong claim that in description and interpretation of a work of fiction, there is no objective fact of the matter, i.e., that we cannot say of any critical description or interpretation that it is true or false of a work, even if it appears wildly irrelevant or perverse. We can only devoutly hope that such a critic will be shouted down.

I think I have shown in chapter III that for a small segment of critical discourse, there is a perfectly good way of defending the truth of critical claims. We can show that they match the sentences of the work or that they may be inferred validly from sentences which do match them. In addition, I think I have shown in chapter IV and in the present chapter that with regard to a great many interpretive claims which are not of this kind there is a perfectly good "fact of the matter." A critical interpretive claim is acceptable on my account if it can be shown to be part of an acceptable theory about the world of the work and unacceptable if it conflicts with all such theories. And if one of these theories is decisively the best one, an individual interpretive claim which conflicts with it is just wrong. In chapter I, I said that I saw no good reason to accept (E2). I should now claim that we have every reason not to, for it conflicts with what seems to be a good comprehensive account of the justification of descriptive and interpretive critical claims.

Second, the denial of (E2) puts my account in a far better position with regard to what I have called "critical significance claims." On my analysis, those claims advanced as critical significance claims
are not legitimate if they are advanced (as they usually are) as claims about the meaning of a work of fiction, where that meaning is taken to be a (general) proposition about the actual world. But my theory can account for the force of such claims. For they can be handled as what I have called "cognitive significance claims." As such they are not acceptable arbitrarily, for they must satisfy at least a minimum criterion of acceptability: the claim which is transferred must be true of the fictional world (either on the basis of the procedure of chapter III or because it is a part of or due to some plausible interpretation of the work as this notion is explicated in chapters IV and V) or legitimately related to such a claim. A claim is legitimately related to a true or acceptable interpretive claim about a work if it is a generalization from a single case, where the work itself contains no counterinstances, or if it is a counterfactual claim derived from a positive description of the world. This rules out "irrelevant" cases, e.g., the belief that the sea is salty supposedly acquired by reading *The Scarlet Letter*.

It will be obvious from chapter II that to some extent I agree with the emotivist theses (*E*₁). I join the emotivist at least in maintaining that it is wrong to assign truth values to sentences of fiction—wrong precisely because these sentences have a distinct use context which makes the assignment of truth values incorrect, not just misleading. If this is what the emotivist means when he says in (*E*₁) that the language of fiction is expressive language, then we are in agreement. My criticism of (*E*₁) as advanced by the emotivists was that as a causal explanation it was inadequate for explaining the nature and diversity of the
function of works of literature in shaping our beliefs about the actual world. The proponent of \(E_1\) merely says that they do have such an effect. The account I have given in this chapter of the transfer of what is learned from fictional worlds to the actual world is at least a preliminary sketch of a more adequate theory. A complete explanation would account in much greater detail for the role of thought experiments in shaping our beliefs about the actual world. It may very well turn out that in the light of further exploration of this topic, the relation of fictional works to our beliefs about the actual world is at least quasi-evidential. Insofar as this turns out to be the case, my endorsement of \(E_1\) will be more restrained. As it stands, my theory has the status of a much more highly articulated causal theory than that offered by the proponents of \(E_1\).

This explanation, I think, clarifies the respect in which the theory I propose differs from the emotivist position and the respects in which, though similar to it, it offers a significant expansion of it. My theory offers a way to accept the view that works of epic literature expand and clarify our knowledge of the actual world without endorsing some version of the proposition theory. That view has, I think, been shown to be theoretically inadequate, as well as presenting a distorted view of what the author is doing. My theory also presents an analysis of the sense in which the novelist may be said to "tell the truth about the actual world," while not confusing novels with, say, speeches or encyclopedias.

The mistake of the proposition theorists in chapter I is in ignoring the fact that the novelist is in the first instance constructing a universe, hence that any account of his contribution to our knowledge of
the actual world must be phrased in terms of this primary activity. The novelist is not an allegorist. But the pessimism occasioned by the failure of versions of the proposition theory is likewise unwarranted. For a fictional universe may, by virtue of its relation to the actual universe, be very important in contributing to our knowledge of the actual world. We have said and will say very little about the importance of the cognitive value of a work of fiction in determining its absolute value as a work of art. I am content now with the claim that it is one extremely important feature of our final evaluation of the novel. For such predicates as 'trite' and 'unlightening' count heavily against a work as a work of art; and they are assigned on the basis of the cognitive usefulness of the work.

In this chapter, I have attempted to give an account of how the basic account developed in chapter III can be extended to explain the assignment of truth values, or at least probability values, to more complex critical claims and the logical role of arguments which rest on considerations "external to the work." In many cases, the prejudice against such arguments has been due to the failure to see what precisely they are arguments for. The present chapter shows, I think, that we can make sense of a good deal of criticism without making the troublesome metaphysical assumptions which, it has often been assumed, would automatically follow from doing this. In the following section, I will try to show that certain critical arguments can be analyzed as having the form I have claimed they have, and that such analysis puts us in a good position to evaluate them.
If purely formalist theories are not appropriate, we are left with the sort of eclectic theory I have attempted to lay out schematically and defend in chapters IV and V. In chapter IV, I think I have shown that in practice even the formalist presupposes or uses such procedures. It remains to show that critics who consciously espouse such an eclectic theory of criticism actually do use arguments which can be shown to have the forms I have sketched and that at least some of their disputes can be arbitrated by a clear, systematic setting out of the arguments in the terms I have proposed. It may be asked why we need to worry about the way critics in fact proceed (beyond showing that in certain cases their practice belies their theoretical position). The answer, I think, is that in giving the logic of criticism, one is giving the logic of a pre-existent form of discourse. If it turned out that no one in fact proceeded in the suggested way, or that on this logical account of such discourse intuitively acceptable arguments regularly came out invalid, I think we would conclude that the logic was inadequate. We would at least be suspicious of it. And if it could be shown that such a logical analysis added nothing to an examination of such discourse, we should be inclined to conclude that it was wasted effort. In fact, neither of these is the case, as I hope to show. There are intuitively acceptable arguments in literary criticism which have the logical structure I claim they should have. There are arguments which are similar, intuitively speaking, which turn out to have the same form. And the analysis of these arguments in the terms I propose adds considerably to our ability to state clearly where the critical disagreement is focused, and in some cases to arbitrate it.
The example I have selected is a dispute between Marius Bewley and F. R. Leavis over Henry James' *The Turn of the Screw.* This dispute is ideal for our purposes for several reasons. Bewley and Leavis are pretty much in agreement on critical methodology. They allow a wide range of arguments. In particular, neither Bewley nor Leavis denies out of hand that the critic may argue from the fact that some set of an author's works have certain characteristics in common that they have others in common, that ambiguities and unclarities in one work may be elucidated by reference to another, more articulate work of the author which is similar in relevant respects. In addition, neither of them denies out of hand that circumstances surrounding the creation of the work or its relation to the actual world may serve as the basis for arguments about the correct interpretation of the work. In the terminology of section 5.1, the dispute involves arguments about the lawlike structure of the worlds, about expansions of truncated universes, arguments based on poetic devices, arguments about the comparison of characters.

Superficially Bewley is arguing that *The Turn of the Screw* and *What Maisie Knew* are similar and that this similarity can be exploited to throw light on "the facts" of the former work. Leavis argues that the dissimilarity of the works rules out such a procedure and in addition that Bewley misinterprets both works. I shall be concerned mainly with the dispute over the facts of *The Turn of the Screw* and with the dispute over the relation between the two works.

The dispute over *The Turn of the Screw* is a complex one. There is disagreement over Bewley's final conclusion about the "fact of the matter," that the children are innocent and victimized and that it is the
governess who is in touch with the demons, which are "the objective symbols of the governess' distorted moral sense." The demons are, he indicates, evoked by her own overwrought imagination. It is the governess who is possessed, not the children. In addition, there is disagreement over the arguments which allegedly lead to the conclusions. In The Turn of the Screw, Bewley maintains, there is an initially unsolvable separation of appearance and reality. Thus, he alleges, we cannot locate the focus of the evil without turning to What Maisie Knew. Leavis argues against this that there is no ambiguity in The Turn of the Screw to start with and in addition that there is no basis for the comparison of the events and characters of the two works.

Of course in dealing with any work of Henry James, there is a third critic to reckon with, James himself. Both Bewley and Leavis agree that they would entertain seriously an appeal to James' explicit intentions and spend some time trying to assess James' comments in the preface. But James' own remarks are inconclusive, and so both turn to other considerations. In addition, Bewley attempts to place The Turn of the Screw in a long American tradition of novelists who have left "the fact of the matter" indeterminate, largely by using limited or deceptive narrators. He argues further that James, in works which led up to The Turn of the Screw, was experimenting with deviant narrators. We would certainly have to acknowledge the force of such arguments if they supported an interpretation of The Turn of the Screw which was explanatorily adequate. Leavis, I think, clearly recognizes this. Both critics thus turn to what they consider "internal considerations." This terminology is somewhat misleading. They mean merely that they have
turned away from James's stated intentions and facts about his biography to test the adequacy of the interpretation against the data.

Bewley attempts to support his interpretation by showing that there is an unsolvable ambiguity in the book between "appearance and reality." Leavis correctly brings the dispute down from the amorphous level of "appearance versus reality" to the concrete question of whether we can trust the governess as a narrator. The answer to this question, he maintains, is unambiguously "no." The disagreement can be construed as an argument of a form which was dealt with in section 5.1, those dealing with the deviant narrator. Both critics may be said to concede that the universe is truncated, that we must argue to a plausible expansion, to a set of facts "beyond" the governess' reports, which either do, as Leavis maintains, or do not, as Bewley maintains, correspond to her reports. The question of whether we can trust the governess rests on two factual questions. (1) Does the governess see the ghosts, or does she evoke them? (2) Are the children innocent, or are they guilty as the governess concludes? The answer to (2) depends in part on the answer to (1), for if the governess evokes the demons, then she is demented and cannot be relied on. It depends in part upon whether the governess, even given that she is sane, reaches warranted conclusions about the children, given her evidence, i.e., upon whether the governess is competent. Bewley, to make his final interpretive claims plausible, must show that the governess-narrator is deviant, either that she is demented or that she is incompetent—preferably both. It is more important to show that she is demented, since his later claims rest on the claim that the universe is not genuinely haunted, but the product of her sick imagination.
In the argument which follows, Leavis and Bewley both make normalcy assumptions about the lawlike constitution of the fictional world. They differ in how far they think the normalcy extends. Bewley needs to show that the parallel is complete. Leavis argues that it is only partial, that *The Turn of the Screw* is to some extent normal, but is in addition haunted, i.e., contains genuine ghosts, exhibits demon possession, etc.

Bewley first argues that the governess "evokes" the ghosts. I think we must construe this as the claim that she hallucinates them. He must argue this way in order to show that *The Turn of the Screw* is causally normal. His argument for this is that she is, on the basis of the presented data, the only one who ever sees them. He also alludes to the constant tinge of hysteria in her reports. This evidence, given certain similarities between the novel universe and the actual universe, would be good evidence for Bewley's conclusions in the absence of counter-evidence. Leavis rather overstates his case when he maintains that it is quite clear that the governess does see what she claims to see. But Leavis has a crushing argument against Bewley, it seems, for after the second apparition, the governess describes the ghost of Peter Quint (whom she has never seen in life) in great detail to the housekeeper; and the housekeeper unhesitatingly identifies the ghost as that of Quint. As Leavis correctly points out, Bewley does not give a coherent account of the ghosts; he avoids doing so because of this fatal bit of evidence and claims that no coherent account can be given. But in this context it is no good to point out, as Bewley does, that Hawthorne's demons are commonly accepted as ambiguous. In the case of Hawthorne, nothing rests
on reaching a definite decision. In this case, something, namely the reliability of the narrator, and hence the "factual expansion," are utterly dependent upon reaching a decision. Despite the ambiguity of many tests, Leavis' objection is decisive. The governess, whatever her emotional state, is not demented. She really sees the demons, whether or not anyone else sees them. Bewley's expansion does not explain the data, and Leavis' expansion does. Leavis' expansion is the better one on this count.

Bewley has rather obviously failed to show that the universe is not haunted, and so has failed to show that the narrator is demented. He sees well enough what he has to show, and he goes about it in the right way, but the interpretation he offers leaves a good deal of the data unaccounted for. It is not explanatorily adequate, and so it makes no difference what external arguments he adduces to support it.

He tries, in addition, to show that the governess is incompetent, that her conclusion that the children are guilty is unwarranted. Bewley argues that she concludes that Miles is vicious on the basis of an "extremely ambiguous letter" which says simply that his school will not accept him back as a student. Bewley's point is clearly that her evidence by no means warrants the conclusion she draws from it. Leavis argues that her conclusion is warranted, that "no English headmaster would have dared to expel a boy--and a boy belonging to a family of distinguished 'County' standing--without being able to substantiate against him as grave a charge as the governess divines from the letter," i.e., that he is an injury to others. Leavis is appealing, then to the assumption that the world of the novel is similar with regard to its
social structure to the England of James' day. His assumption is based on what we should term external grounds, that "James, when he wrote the novel was Englishman enough to know the force of such a letter." He might also argue that we need such an assumption throughout the novel in order to understand the relative positions and authority of the different characters. Given this consideration, and supposing Leavis to be a competent reporter on English social customs, his argument has considerable force. Leavis might also have gone on to argue that the boy Miles finally confesses to an offense which fits the specifications, to telling "those he likes things which were too bad to write home in the letter."¹⁹ This is a fact which Bewley leaves out of consideration, and one which his theory of Miles' innocence would have a difficult time accounting for.

Bewley has again failed to account for some of the data. Leavis¹ account does account for the data, and it is in addition supported by acceptable external arguments. I think we can conclude that Bewley has failed to make his case against the governess with respect to Miles.

The governess' case against Flora is not quite so strong; but, as Leavis point out, the housekeeper bears full testimony to Flora's "speaking horrors" after the scene at the lake. Bewley's explanation of this fact, i.e., that Flora has suddenly acquired the governess' moral consciousness and has come to see her former associations in this distorted light, is particularly weak. For from the housekeeper's account, what Flora does is not to reveal her past associations in a new and damning light, but to use particularly disgusting language to revile the governess.²⁰ No doubt this fact could be explained away more
competently than Bewley does. But the hypothesis obviously open to Leavis, since he has successfully defended the claim that the universe is haunted, is that the demon in possession of the child, after hearing itself named, is shrieking obscenities through the lips of the possessed child. Again, Leavis' interpretation would explain all the data easily, whereas Bewley's does not (although perhaps we could improve on the explanation Bewley gives). In addition, the interpretation open to Leavis fits into an interpretational scheme which is already superior on two other counts.

Leavis' conclusion, then, is that we can trust the governess, that she is both sane and competent. He has argued convincingly that the most plausible expansion of the truncated universe is a "haunted universe," one in which demons appear and possess children, etc. And this vindicates the governess. At least this is the underlying form of his arguments. Bewley can be construed as holding that we have here a normal universe with a demented and incompetent narrator. There is little doubt, I think, that Leavis is right. He can account for several pieces of data that Bewley cannot account for, i.e., the housekeeper's recognition of Peter Quint's ghost from the governess' description and Miles' confession. His interpretation fits the totality of the data about Miles better than Bewley's, given the strength of his external argument about the significance of the letter and Miles' subsequent confession. Leavis' theory explains the data about Flora at least as well as Bewley's. And Leavis' account accounts better for other data which are not mentioned by either critic, e.g., the children's precocity, Miles' sudden death upon mentioning Quint's name, etc. Systematizing the arguments in this
way helps to locate the exact deficiency in Bewley's arguments. He and Leavis are both following correct argumentative procedures. They are arguing about the narrator's competence on the basis of what the world must be taken to be like. The problem is that Bewley has not got a coherent explanation of the data to begin with.

The second point of disagreement between Leavis and Bewley is the conclusion about the characters based on the alleged similarity between What Maisie Knew and The Turn of the Screw. The dispute is again not a general methodological one. Both Leavis and Bewley explicitly allow arguments from one work to another, provided such arguments are based upon the author's customary mode of action and that the works are similar in their lawlike constitution. Leavis claims merely that this specific argument is a bad one, because one of its premises is false. His main contention may be expressed in the terms of 5.1 as follows: Given the correct interpretation of The Turn of the Screw, the two universes are profoundly different in their lawlike constitution, hence arguments from one to the other are not legitimate, simpliciter, and this leads to there being no basis for comparison of characters from the two works. The Turn of the Screw, as he points out repeatedly, is a haunted universe; What Maisie Knew is an overwhelmingly naturalistic universe.23 This can be construed simply as the claim that the two works belong to different subsets of K on the basis of their constitutive logic.

No doubt Leavis is right about this. But it does not follow immediately that character comparison between the two universes is fruitless. The question is whether the two worlds can be amalgamated in such a way as to make the character comparison interesting and profitable.
Leavis is not much help here. He merely reiterates his claim that such a comparison is useless and never actually gives his argument. But perhaps on the basis of section 5.1, we can give an account of what he might mean, or at least of what he should have said.

Bewley is trying to argue that if Maisie is not vicious, then neither are Flora and Miles. For this to be convincing, he needs to build on a general similarity between the two universes, at least as far as the respective predicaments of the children are concerned. Leavis might have attempted to deny that there is any such similarity by arguing in the following way. If we try to amalgamate the worlds by construing The Turn of the Screw as non-haunted (as Bewley no doubt intends us to), then Flora and Miles do not in the book face the same alternatives as Maisie faces. The main crime they are accused of, i.e., being possessed, cannot now occur in the non-haunted world. In addition, we know nothing about their thoughts, motives, etc., whereas we know everything about Maisie. We have nothing to go on to establish the requisite similarity and no reason to think that there is any. If, on the other hand, we allow the amalgamated universe to be a haunted one, the result is even worse. Maisie is not possessed in the amalgamated universe; she is responsible for her actions. Miles and Flora are possessed; they are not responsible for their actions—if indeed they may be said to perform actions. We can contrast the characters. Maisie is innocent. Miles and Flora are guilty and depraved. But we certainly cannot argue on the basis of similarity between Maisie and the haunted twosome that we can make similar judgments in the two cases. The characters are not similar in the relevant respect. This may be the kind of argument that Leavis has
in mind. At any rate, it is a good argument. Given that Bewley has failed to show that the universe of *The Turn of the Screw* is a normal universe or that the narrator is incompetent, and given that the amalgamation of the two fundamentally different universes is fruitless for his purposes, it seems clear that Bewley loses his argument against Leavis. The argument I have suggested in connection with amalgamating the two universes could very well be what Leavis means to accomplish by his repeated and rather vague claims that Maisie cannot be compared to the children of *The Turn of the Screw* because she is "an actual child" confronted with the problem of a "squalid world," not with "evil, horror, and sexual depravity." In such passages, he seems to be trying to make some point distinct from his original claim that Bewley cannot show the universes to be significantly similar.

Leavis' position is convincing, particularly if he has in mind the argument I have attributed to him. He is ultimately right about the "facts" of *The Turn of the Screw*, even if they are not so unequivocally clear as he makes them out to be. He is right about the fact that the two universes are fundamentally different in constitution. And he is right about the fruitlessness of comparing Maisie with Flora and Miles. He has made his point without claiming that arguments of the kind used by Bewley are *per se* objectionable. I think I have shown that both Bewley and Leavis argue in the way I have sketched in chapter 4. Both demand an explanation of the data of the novel which depends on constructing a plausible expansion of the universe in question; and both fall back on external arguments, appeals to the way the actual world is, to the author's intentions, to other works by the same author, or to
other works which may be presumed to have the same constitutive logic.\textsuperscript{25}

I think I have shown, then, that they argue in the way I have suggested in 4.2, and that once the arguments are seen in this light, a decision may be reached as to who is in the right. This is fairly important, since Bowley's argument looks plausible until it is examined in detail. At most it looks like a case of overreading, which is more or less permissible, provided the interpretation is adequate and accounts for the data. It is only when it is realized that he simply cannot account for part of the data that he appears obviously wrong.

Leavis' judgment about the cognitive value of \textit{The Turn of the Screw} is that it is an insignificant little thriller loaded with a purposeless atmosphere of ambiguity. His judgment is certainly based in part on his view that the world of \textit{The Turn of the Screw} is too unlike our own to yield hypotheses which are directly transferrable to the actual world. No doubt he also thinks the work is too ill-defined and too shallow to produce interesting hypotheses by what we have called "the counter-factual method." Bowley certainly does not approve of the ambiguity he claims to find in \textit{The Turn of the Screw}; but he clearly thinks that the work, considered in the light of \textit{What Maisie Knew} as naturalistic, provides an interesting hypothesis about the corruption of innocence through the imposition of "moral sensibility." Leavis is more nearly right, I think; but he does ignore one interesting possibility. There is a theme, a typically Jamesian theme, which \textit{The Turn of the Screw} could be construed as conveying counterfactually, namely, that evil generates terror and horror and preserves its aura of magnificence precisely so long as it is unfocused. This, to judge from James'
preface, may well be what he had in mind. Seen in this light, the novel remains a "thriller," but not completely a "non-significant thriller," as Leavis alleges.

This concludes the specific demonstration I proposed to give in this chapter that the critical approach I sketch in chapters III and IV does in fact provide a logic of criticism which can be used to arbitrate both methodological disputes and disputes about "the fact of the matter" in specific cases. I think it is also clear from the example, that in some cases the result of a decision about the correct interpretation of a work provides a premise which is important in leading to an aesthetic judgment about the work. I certainly do not hold that the cognitive value of a work of epic literature ought to be or even always is decisive in our total evaluation of the work as a work of art. Nonetheless, certain cognitive features of the work will be important in our overall evaluation of a work. Banality and triteness count against a work. And a work which is genuinely and significantly ambiguous, one which is susceptible to various interpretations which produce interesting hypotheses about the actual world, will be likely to be judged "rich" or "profound." I will not attempt to deal any more completely with the question of the evaluation of works of epic literature. It is not a question I set out to deal with, although it is one about which something, at least, must be said. I think I have indicated the lines along which further investigations of the question might proceed.
Notes

1For example, for Fredrieca to be the owner of the stock, in 1003, it is necessary for the main character to go back 30 years in time to sign it over to her properly.

2I am indebted to George Schum for help in identifying the systems for the models determined by these novels.

3Whether such conceptions of time are metaphysically comprehensible is another matter. The point is that if we make the proper assumptions, they are logically consistent. It will turn out later in the section that the price which must be paid for such tinkering is rather high.

4Of course, there may be "distorted universes" which even with distortion assumptions are inconsistent in the way in which cases which fall under case (B) are inconsistent. For example, if all but one tense sentence of a work determines one tense model, and one sentence does not fit the model, then we simply declare the universe inconsistent in the case (B) sense, then proceed to shrink the universe to consistency.

5Robert Penn Warn, All The King's Men (New York: 1946), 112.

6Ibid., p. 275

7Ibid., pp. 436, 38


9Bewley, "Appearance and Reality," 102. Leavis agrees that the procedure per se is legitimate and himself argues in this way about the proper interpretation of What Maisie Knew.

10Actually, Bewley has in mind a wider range of works than this, including Henry James' The Liar and The Golden Bowl.


12Ibid., p. 112.

13Ibid., p. 103.


274

16Ibid., p. 113. Of course, Bewley might claim that the incident is an amazing coincidence, but as he is committed to the view that the work is essentially causally normal, this would certainly be a weakness in his theory.

17Ibid., p. 107.


20Ibid., p. 117.

21Bewley could, for example, argue that Flora is a normal child who has learned to apply a lot of foul language appropriately from her previous association with Quint. This would not fit Bewley's overall strategy very well, but it would eliminate the only putatively decisive evidence that Flora is depraved.

22What Leavis charitably does not point out is that Bewley reports the data inaccurately on several occasions.

23This, I think, is a concise statement of Leavis' objection to the comparison. He goes on to group the work with other works by James and with works by Dickens, all of which are "naturalistic" in this sense, not "haunted." "What Maisie Knew," 123.

24Ibid., p. 123.

25In addition, both Bewley and Leavis fall back on supporting from poetic imagery to reinforce claims they have already drawn on other grounds. In form both the arguments of Bewley and those of Leavis are correct. Bewley's have the disadvantage of supporting a fundamentally incorrect interpretation. He argues that Flora's innocence is symbolized by her appearance with the withered fern. And Leavis argues at one point that Maisie's affirmation of Sir Claude's claim with her mouth full of bread and butter conveys "the strength of positive values" and "the concept of essential human goodness." These are both examples of arguments based on fuctional metaphors.
BIBLIOGRAPHY

Books


Articles


Harman, Gilbert. "Deep Structure as Logical Form," paper delivered at a colloquium at The Ohio State University, 1972.


"The Logic of Art: A Rejoinder to Dr. Hoekstra," Philosophy and Phenomenological Research, V (1945), 365-78.
