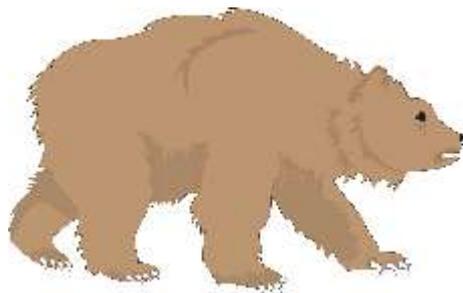_____

Ursus Philosophicus

Essays dedicated to Björn Haglund on his sixtieth birthday

_____

Truth in (of) Fiction

Staffan Carlshamre

Near the end of Aksel Sandemose's novel *The Werewolf*, the female protagonist Felicia Venhaug disappears. Her body is never found but it is taken for granted, in the novel, that she has been killed. Unfortunately, we, the readers, never get to know who killed her—at least, we are never told who did it, within the covers of the book.

I will use the question about who killed Felicia to discuss some issues concerning truth in fiction. Given that it is true in the novel that *someone* killed Felicia, is it also true of some specific person, in the fiction, that she or he did it? If the answer is yes, who is it?

There are two main problems involved, one epistemological and one semantic. The epistemological problem is about how we can possibly *know* who killed Felicia—what sorts of evidence are relevant and in what sorts of ways? The semantic problem concerns what would make a particular assertion of the type "X killed Felicia" *true* or *false*.

The best known theory about truth in fiction is in David Lewis' seminal paper "Truth in Fiction" from 1978.[1] I will argue that Lewis' theory, as well as a modification of it proposed by Gregory Currie, although it contains valuable insights, is fundamentally on the wrong track.[2]

1 Truth in (of) fiction

The Werewolf is about Erling Vik and Felicia Venhaug and some other people who are important in their lives. Erling and Felicia are lovers, but Felicia is married to Jan Venhaug, who knows and approves of the relation between Erling and his wife. The novel covers roughly the first half of the 20th century, from Erling's childhood to Felicia's death. The details of what happens to these people make up the most important part of what is true in the story.

But *The Werewolf* is not just a made-up story about some made-up characters. On another level it is about Norway, before, during and after the Second World War—not just about a made-up

---

[1] Reprinted in his *Philosophical Papers*, Vol. I (Oxford 1983).

[2] The basic statement of Currie's theory is in *The Nature of Fiction* (Cambridge 1990). I will also argue that Currie's theory *is* a modification of Lewis', i.e., that the difference between the two is smaller than it may at first sight appear to be.

Norway, but about the real one. On that level it could be praised for giving a true picture of Norway during the German occupation, or blamed for giving a false one.

On yet another level, suggested by the title to be the most important, the book is about envy, jealousy, evil and stupidity - everything that it gathers into the symbol of the Werewolf. On this level, too, what the novel "says" is not about the fictional world, but about the real world.

So we have two rough kinds of question regarding fiction and truth. We may contrast questions about *what is true in the fiction* with questions about *what the fiction says that's true*. Or we may avoid using the word "true" and distinguish questions concerning *what the story is* from questions about *what the story means*. The question about who killed Felicia is obviously of the first type, while questions regarding what *The Werewolf* says about Norway or evil or human nature belong to the second type. In the literature on truth in fiction, questions of the second type are usually mentioned just to be brushed aside. One point to be argued in the following, is that the two sorts of question cannot always, or usually, be quite so rashly isolated.[3]

The trivial nature of the examples generally used in this connection has helped, I think, to conceal this rather obvious point. There is not much meaning left in a Sherlock Holmes story once you know what the story is—but this is not the kind of material usually addressed by literary interpretation of a qualified kind. The question about who killed Felicia is, to the contrary, central to the interpretation of *The Werewolf* on all the levels that usually interest literary scholars and critics.

2   Layers of meaning

I have already used the metaphor of *stratification* several times—talking of "levels" of interpretation and "layers" of meaning—and before I go on I want to elaborate that metaphor a bit. Look at the two "layers" I have mentioned so far—what the story *is* and what the story *means*. In a very extended sense they are related as "signifier" and "signified". When we interpret the novel for its lessons about Norway or human nature it is the *story* that we interpret—not, e.g., the *words* that are used to tell it. So, in a certain logical sense, we must have the story "first", in order to interpret it. This is not true in a temporal sense, of course—one of the points I want to stress is that an important criterion for determining what the story is, is that it shall make reasonably good sense on the "next" level.

But the story is not only the "signifier" for the higher levels, it is itself the "signified" for a (logically) prior act of interpretation—what is the "signifier" for that interpretation? A plausible

---

[3]Dave Davies argues the same conclusion in "Fictional Truth and Fictional Authors," *British Journal of Aesthetics* 36(1), January 1996

(and common) way to look at it is that we get the full story by "fleshing out" what is actually stated or given in the book—we may imagine a set of basic facts (or, rather, "ficts") from which we reason to a larger whole. The classical problem of truth in fiction is concerned with the nature of that reasoning—from the *given* ficts to *all* the ficts.

None of these three levels has anything particularly linguistic about it—we may imagine the same story and the same *sense moral* being derived from the film as from the book.[4]

3   Lewis' theory

The basic idea of Lewis' theory about truth in fiction is simple: what is true in the fictional world is what *would* be true if what is actually said in the novel had been true. Fictional truth is a species of counterfactual truth. If you bring in the details it gets a little more complicated, but not much. Lewis gives two different analyses—here's the first one:

> **Analysis 1:** A sentence of the form "In the fiction f, f" is non-vacuously true iff some world where f is told as known fact and f is true differs less from our actual world, on balance, than does any world where f is told as known fact and f is not true. It is vacuously true if there are no possible worlds where f is told as known fact.[5]

This is, of course, an application of Lewis' general theory of counterfactuals to the case of fiction. What is true in the novel is what is true in those possible worlds that are "closest" or "most similar" to the actual world, and where the text of the novel is told as "known truth".[6] Roughly speaking, we construct the fictional world by taking the actual world as background or default, and modify it just as much as we have to, in order to make what is actually said in the novel true. Constructing Sherlock Holmes' world we do not move Waterloo Station, because we do not have to, but if there

---

[4]I give a much more elaborate account of levels of meaning, in relation to literary works, in "Types of Types of Interpretation," in Staffan Carlshamre and Anders Pettersson, ed., *Types of Interpretation in the Aesthetic Disciplines* (Montreal & Kingston: McGill-Queen's University Press, 2003).

[5]David Lewis, "Truth in Fiction", p. 270.

[6]The part about worlds where the fiction is "told as known fact" may sound suggestive. Would it not be simpler just to settle for those worlds where the fiction is true? Has Lewis something exciting up his sleeve about narrators or narration, in relation to fictional truth? No he has not, the complication is there simply to exclude worlds where what is told in the fiction happens to be true without any connection to the story. Suppose, for example, that there actually was a person just like Sherlock Holmes, performing acts exactly similar to those that Conan Doyle ascribes to Holmes, but without Conan Doyle knowing it. (A natural question, that Lewis as far as I know neither raises nor answers, is who tells the story in those worlds where it is told as known truth? Is it Conan Doyle or Watson? The first alternative encounters the problem that worlds that houses Holmes may have no place for Conan Doyle—in fact, it belongs to the fiction that the stories are told by Watson. But the second alternative does not solve Lewis' problem about coincidental truth—suppose that, unbeknownst to Conan Doyle, there happens to be, in the actual world, a person just like Watson who tells stories exactly similar to the Sherlock Holmes stories, as "known fact"?)

is actually a bank at Baker Street 221, we have to evacuate it in order to make room for Holmes' lodgings.

Analysis 1 gives us, according to Lewis, a possible and reasonable (and perhaps very common) way of reading a novel, but in some cases it leads to results that may seem odd or downright wrong, at least to some readers. A plausible test case is this: did Hamlet have an unconscious mind, in the Freudian sense? That question will divide those who read according to *Analysis 1* into two camps, according to what they think about the truth claims of psychoanalysis. If you think actual people have unconscious minds you will think that Hamlet has one as well; if you don't think real people have unconscious minds you won't think Hamlet had one either. But some readers will think it completely irrelevant whether actual people actually have unconscious minds or not. The important thing is not whether psychoanalysis is true but whether Shakespeare believed in it. Lewis tries to capture intuitions of this nature (but with a modification that I will come back to) in a second proposal:

> **Analysis 2:** A sentence of the form "In the fiction f, f" is non-vacuously true iff, whenever *w* is one of the collective belief worlds of the community of origin of f, then some world where f is told as known fact and f is true differs less from the world *w*, on balance, than does any world where f is told as known fact and f is not true. It is vacuously true iff there are no possible worlds where f is told as known fact.

Instead of using the actual world as a background for the construction of the fictional world, we take a certain "belief world" (or rather a set of belief worlds[7]) as default. Whose beliefs should we appeal to in this context? To my mind there are only two plausible alternatives—once we have decided not to use our own beliefs—both anchored to the real author of the fiction. Either we could use the author's own beliefs about the world, or we could use the author's beliefs about what the audience believes. But Lewis settles for a third alternative—the "collective belief world of the community of origin" of the fiction. This emphasis on community beliefs is, of course, very popular, but I find it quite unconvincing, and I will come back to it below, in connection with Currie's views.

It should be noted that none of Lewis' analyses singles out one unique world as *the* world of the fiction. This is obvious for *analysis 2*—if the relevant community neither believes nor disbelieves, say, that the highest building in Minsk is more than 30 meters, then it is neither true nor false, in the fiction, that it is. The fictional worlds of *analysis 1* are more completely specified—the actual

---

[7]The reason we have to use a set of worlds is that whatever subject we use to provide the beliefs, there will not be enough beliefs to single out just one possible world—instead we take the class of worlds compatible with what the subject believes. For simplicity I will continue to speak of "a" belief world in the singular. (We will also need a strategy for weeding out inconsistent beliefs, etc., but the details of that need not concern us now.)

world decides what is true in the fiction about every building in Minsk—but there are still "gaps" in them. Nothing in the actual world decides exactly how many hairs there are on Holmes' head when he first meets Watson.

There are three possible types of problem case for Lewis' theory.[8]

(1) **Too many ficts**. Some might feel, and some have felt, that both of Lewis' analyses are too generous in populating the fictional world with irrelevant ficts. Is it really plausible that Holmes' world contains ficts about the height of buildings in any number of exotic places that have nothing to do with the story?

My own feeling is that counterexamples of this sort do not tell against Lewis' theory. The actual world is full of irrelevant facts and I don't see what's wrong with fictional worlds being full of them as well. The next sort of problem case would be worse.

(2) **Too few ficts**. Do Lewis' criteria decide all the questions we in fact take as decided when reading a fictional story?

I will argue that the case of who killed Felicia is a counterexample of this type, and that this is fatal to Lewis' theory. For completeness I will mention the third type of problem case as well, though I will say nothing further about it here.

(3) **The wrong ficts**. It might be the case that Lewis' criteria decides some questions about the fictional world the "wrong" way, giving answers that contradict the answers we give as ordinary (or specialised) fiction readers.

4 Currie's theory

On the face of it, Currie's theory looks very different from Lewis'. Instead of asking us to work out what would be true if what is actually said in the novel were true, Currie wants us to work out what a certain "person" *believes* is true. But much of the seeming difference between the two approaches evaporates, once you realise that the "person", whose beliefs we are supposed to be interested in, is himself a fictional construct—Currie calls him "the fictional author."

According to Currie, what is true in the fiction is what the *fictional author* believes is true, or, more precisely, what we, as readers, can reasonably infer that the fictional author believes from

---

[8]When I speak of "Lewis' theory" I refer to the specific analyses 1 and 2. It is, of course, possible to use Lewis' general framework—similarity relations on possible worlds—to formulate many other proposals. Indeed, the framework may be flexible enough to accommodate any possible theory at all.

what he says, i.e., from what is in the text, when we make-believe that the text is told as known truth.

To see the similarity to Lewis' theory we may formulate Currie's proposal like this: what is true in the story is what someone *would believe*, were he to tell the story as known truth. Among the things reported in a work of fiction are, usually, many speech acts. To understand the fiction you have to interpret these speech acts, as uttered in the fiction, i.e., to infer what intentions, attitudes and beliefs that are (fictionally) supposed to have given rise to them. In Lewis' theory this is done like any other "finding out" of what is true in the fiction—we reason as to what intentions etc. would have been present had the utterances actually been made, in the circumstances described. Currie's theory is that interpreting the fiction is like interpreting, Lewis' style, the speech act of telling the story as known truth.

It is important to know how Currie conceives the fictional author. In some ways the fictional author is quite similar to the real author. The fictional author of *Hamlet* is not Shakespeare, of course, but we must suppose him to come from generally the same circumstances as Shakespeare— to belong to the same epoch, to be "a member of the same community" as Shakespeare.[9] There are two important differences between the real Shakespeare and the fictional author of *Hamlet*.

(1) The fictional author is a *typical* member of his community—unlike the real Shakespeare he may not harbour any idiosyncratic beliefs or attitudes, at least not without expressing them clearly in his text.

This is Currie's counterpart of Lewis' use of the "collective belief worlds of the community of origin" of the fictional work as the anchor of *Analysis 2*, rather than the belief worlds of the author. Personally, I'm entirely unmoved by the considerations put forward in favour of this approach. Currie uses the example of how to envisage the function of the Ghost in *Hamlet*. Are we to think of the Ghost as real, or as, e.g., an hallucination? Off hand, it would seem that there are some very pertinent questions in this connection about whether Shakespeare believed in ghosts, or, perhaps, took the possibility of ghosts seriously, without actually believing in them, or expected his audience to believe in them, etc. But according to Currie (and Lewis) such questions are totally irrelevant— what matters is what Elizabethans in general believed. Referring to Hamlet's expression of doubts concerning the reality of the Ghost, Currie says:

> To see *Hamlet* as the product of a society prone to regard ghostly appearances as false representations rather than as dead souls is to be more receptive to the idea that the lines quoted were intended to be understood as an expression of genuine doubt, and

---

[9]Currie, *The Nature of Fiction*, p. 78.

written with the expectation that such doubts would find ready response in the audience. It would not matter if Shakespeare actually had no such intentions and expected no such response (he might have been out of step with contemporary thought).[10]

Surely, this must be putting the cart in front of the horse. In most cases, of course, it is entirely reasonable to go by the community beliefs—on the assumption that the author, being a member of the community, can be expected to share them. And often, as in the case of Shakespeare, there will simply be no other evidence as to what the author believes. But what could be the *rationale* of sticking to the community beliefs even in a case where we would have positive evidence that the (real) author neither shared them, nor expected his audience to share them? Why would we expect a set of beliefs totally unconnected to the actual "production" of the text to throw light on how it is to be understood? The second important difference between the fictional and the real author is also familiar from Lewis' theory:

(2) The fictional author tells the story not as a fiction but as known fact.

There are some questions about how this should be understood, but at the very least it must be taken to imply that the fictional author is not to be construed as issuing a novel or a play, or a work belonging to any other fictional *genre*. And this, in itself, is enough for Currie's theory to be affected by the sort of counterexample which I will now proceed to describe.

5   Straight intentionalism

I do not want to say, of course, that counterfactual reasoning plays no role in finding out the ficts of a fictional story—I even agree that this sort of reasoning is the most important mechanism involved. But I want to argue that we cannot take that mechanism to *define* what it is to be a fict, to be true in the fiction. As a contrast, I will use a theory that I will call *straight intentionalism*—the fictional world just is what its author takes it to be. [11]

In the last section of "Fictional Truth and Fictional Authors", Dave Davies discusses various intentionalist analyses of truth in fiction. His first suggestion—the "most obvious candidate" though "clearly inadequate"—may be paraphrased like this:

P is true in the fiction S just in case the real author of S intends the reader to make believe that P.

---

[10]Currie, *The Nature of Fiction*, p. 83.

[11]There is an additional problem, of course, about fleshing out what the author explicitly *thinks* about the fictional world, but I will ignore that here.

The reason that Davies takes this to be inadeaquate is that it "does not allow for the failure of the real author to produce a work in which her intention is realized." But what does it mean to say that the intention is realized in the work? Barring any legislation on just which devices an author can use to get her intentions across to the reader, it seems that the most natural suggestion would be that the intention is realized if the reader realizes it. The failure would then be a failure on the part of the author to make it known to the reader which story he intends, not a failure to tell the story he intends. To my mind, this seems perfectly adequate. It also makes natural sense of the idea that a story may be a success for some readers, and a failure for others—there would, of course, be many cases where we would more naturally blame the reader than the writer, and indeterminate cases as well. If one needs a less relative notion of success or failure one would, most reasonably, invoke the intended reader.[12]

If straight intentionalism is true, as an account of what it *is* to be true in the fiction, we still have the epistemological problem—how do we find out what the author intends? This is, more or less, the converse of the author's problem—how does he *make known* what he intends. The author has to give us *clues* to his intention. The clues are in the text of the fictional work, and interpreting the clues will often involve counterfactual reasoning of the Lewis type. But I will use the example of *The Werewolf* to show that these are not *all* the clues and that we need some other account of what they are clues of. [13]

Maybe, at this stage, I should something about what a theory of truth in fiction is a theory of. What makes the theory true or false? I think it clear that Lewis thinks of his theory as a piece of conceptual analysis, which is ultimately about the reading habits of fiction readers. He explicitly says that analyses 1 and 2 may both be correct for different groups of readers. My counterexample will be given from the same point of view. I will show you a way of reading that does not fit the counterfactual model, and I hope you will agree that it is not unique or awkward but of a fairly standard sort.

Towards the end of the paper I will leave the purely descriptive viewpoint, however, and speculate a bit about what makes it *reasonable* to read according to straight intentionalism, on this

---

[12]Davies' argument against straight intentionalism may be taken as a variant of a general argument against "actual intentionalism" in interpretation, namely that it does not exclude "Humpty Dumpty Semantics" (cf. Paisley Livingston, "Intentionalism in Aesthetics", *New Literary History* 1998, p. 831, and Noël Carroll, "Interpretation and Intention," in Margolis and Rockmore, editors, *The Philosophy of Interpretation*, Oxford: Blackwell 2000, pp. 75-6). I think that this argument confuses a constraint on how it is reasonable for speakers to speak (namely so that they will have a good chance of being understood) with a constraint limiting the application of actual intentionalism. If we really had reason to believe that some expression was used in a Humpty Dumpty fashion, in a certain text, few of us would, I think, go on to insist on a normalizing reading.

[13]Whether all the clues are "in" the text is to some extent a terminological question. Certainly, much of the information needed to interpret the clues is not in the text but in the "context"—information about genre, about relations to other works, by the same and other authors, etc. And what is a relevant context? In this case: a context intended by the author.

level—why should we let the author decide, so to speak, what the ficts are of his story? And my hypothesis will be that we want to *agree* with the author about what the story is, in order to be able to meaningfully *disagree* about what the story means—to be able to have an interesting controversy on the next level.

6   A good story

Let us suppose that one of the counterfactual analyses of "truth in fiction" were correct and try to apply it to *The Werewolf.* (I will for, simplicity, stick to Lewis' *analysis 1*, but, *mutatis mutandis*, everything I say will hold for *analysis 2* and for Currie's theory as well.)

We are supposed to treat what is said in *The Werewolf* as a true report, relative to a possible world where it is *not* told as fiction but as true. What would we then know about who killed Felicia? Practically nothing, unfortunately. We have some very circumstantial evidence—mainly a mysterious letter, and some possible, but not very strong, motives (the intimate circle have perfectly good alibis). We seem to be in the "how many hairs on Holmes' head?" situation—the novel, told as true, simply does not give us enough to go by, to decide what is the case. And so, according to Lewis' theory, there will presumably be *no* true answer to the question about who did it—and accordingly the effort (made by every thoughtful reader of the book, I suspect) to discover the secret will be simply futile. But suppose that we do not take the story as if it were told as known truth, but instead take it as a novel, written by Aksel Sandemose—as it in reality is. What do we then know about who killed Felicia? A great deal, it seems. Here are some general, but helpful, "default" assumptions that might get us started:

1. The motive for the murder is connected to the symbol of the Werewolf. It has to do with *envy*, *jealousy*, perhaps a certain form of *stupidity*. (How do we know? The murder is perhaps the most important single event in the book—its importance being marked, e.g., by its position at the very end of the story. The symbol of the Werewolf gives the book its title—it would be strange indeed if it were unrelated to its very *finale*.)

2. The killer is "known" to the reader, from the story. This is a striking departure from the case of a factual report—nothing at all, besides the conventions of fictional storytelling, entitles us to disregard potential killers outside those who have figured in the rest of the story.

3. Who killed Felicia affects the overall significance or meaning of the story, the *sense moral*. Expressed in another way: the question about who killed Felicia is *important*.

In short, there is a huge difference between having something told as *known truth* and having it told as a *good story*, and ultimately as a good *fictional* story— a difference that has to do with the requirement that a good story shall have integrity and a point.

7   Who did it?

And so—who did it? I will discuss two answers to that question, to the best of my knowledge the only answers that have actually been proposed. My goal is not, of course, to decide which answer is correct, but to look at the arguments that have been presented for and against them.

According to the standard hypothesis the killer is Gulnare. She was Erling's first love in their youth, but, with time, she has hardened to a bitter old woman, through a series of events that are important to the story of the book, but that I will not go into here. During the war, Felicia killed Gulnare's husband, a Nazi collaborator, as an act of revenge for the death of Felicia's two brothers; this would, perhaps, have given Gulnare a reason to kill Felicia, had she known about it—but she does not. So, the motive has to be taken in a more general fashion—jealousy, envy, unspecified desire for revenge on Erling, general disgust at the way of life at Venhaug.

What specific reason is there to suppose that Gulnare is the killer? Mainly this. On the train to Venhaug, on the day before the murder, Erling has seen a woman (or, more precisely, the back of a woman) who has induced in him an unpleasant and uncanny feeling. He does not recognize her, but when he tells the incident at Venhaug, Jan (Felicia's husband) suggests that it was Gulnare—a suggestion somewhat angrily dismissed by Felicia.

Clearly, that piece of evidence would not do much for the case of the prosecution in a court of law. But the court of literature is a different thing. What other reason can there be for Sandemose to recount the incident on the train, but to implicate Gulnare in the murder? The important thing is not the episode in itself, but the fact that Gulnare as a killer seems to make the novel a reasonably satisfying whole, weaving together many threads of the narrative in a way that connects with the main overall theme. The incident on the train is Sandemose's way to suggest that this is the ending he has in mind—whether it counterfactually closes the case, so to speak, has nothing to do with it.

As I said, this is the standard opinion among scholars and critics. And it seemed to receive conclusive support from Sandemose himself when, a few years after the publication of *The Werewolf*, he issued a sequel to the novel, *Felicia's Wedding*, where he seems to state, quite explicitly, that Gulnare is the killer.

Nevertheless there is one dissenting voice among the critics, and an alternative hypothesis about who killed Felicia. Ruth Jørholt has argued that the killer is Felicia herself—it's not a murder but

a suicide.[14] I will look at some of her arguments, and then come back to the question what these arguments reveal about an underlying notion of fictional truth.

To get her own alternative into the race Jørholt first has to argue against the Gulnare hypothesis. Gulnare has no reasonable motive to kill Felicia—none of those that have been suggested will do. Does she hate Felicia for killing her Husband? No, she does not know that Felicia killed him. Is jealousy against Felicia the motive? (For Felicia's having Erling, presumably, and a happy life in general.) Jørholt admits that there is room for envy, perhaps, but Gulnare can't blame any of her own misfortunes on Felicia, and she does not have anything to gain from Felicia's death. Does she hate Erling? That love can turn to hate is a favourite theme with Sandemose, and Jørholt admits that some passages in the book point in this direction, but she thinks that these hints are "a bit too obvious" and should rather be taken as misleading clues, in detective story fashion.

The first two counterarguments seems to fit the counterfactual model well, but with the last one we are on different territory—the crucial premise is the author's wish to create a particular kind of story, with a certain kind of suspense and related to certain expectations on part of the audience. The same sort of argument would, presumably, be adduced against the importance of the uncanny encounter on the train—"a bit too obvious".

What about the statement in *Felicia's Wedding* that Gulnare is the killer? In the novel, this assertion is made by Erling Vik, and it is transferred to Sandemose only by way of the rather obvious fact that Erling is largely an *alter ego* for the author. Jørholt only has to argue that this transference is illicit, and that it belongs to the fiction that Erling is mistaken. In principle, of course, there would be another possibility open to her—she might just declare Sandemose's own opinion irrelevant to what the ficts are. On Lewis' theory (and Currie's) that is precisely what she should do—for on that theory the opinions of the (real) author are irrelevant. But it is plain that Jørholt never considers that possibility—she simply takes it for granted that we seek Sandemose's real intentions, and so she must argue that Erling's views about the killer are not the author's.

I will not dwell much on Jørholt's direct arguments in favour of the suicide hypothesis, most of which makes good sense from the point of view of either theory. She argues that Felicia has reasonable motives to commit suicide, and that she acts reasonably for a person who has decided to kill herself—she has made precautions not to harm those she cares about, Jan and the children, and she sends a symbolic message to Jan on her way out. More telling, from my point of view, is an argument that the suicide interpretation gives reasonable meaning to the *style* and *tone* of the relevant chapters. The use of stylistic clues are not easily accommodated into Lewis' theory, which concentrates on what is said but leaves no role for how it is said.

---

[14]Ruth Jørholt, "Felicia försvann, kan någon säga hur?". In Dupont, Sørensen and Væth (eds.), *Atlanten har så mange mil: Streiflys over Sandemose og hans forfatterskab* (Oslo 1986).

I shall end my synopsis of Jørholt's argument, with a curious but interesting response that she gives, to what otherwise would seem to be a devastating objection to her interpretation. Felicia is drowned in a river, her body is never recovered. But at the point where she leaves the road and goes off towards the river there is some blood in the snow, and there is a cork from a bottle—clear indications, it seems, of another person and a struggle. Now, how do you explain that away?

Jørholt's solution is that the blood and the cork shall not be interpreted realistically but *symbolically*. As a symbol, the blood is connected to the symbolism of the Werewolf, while the bottle and the cork is supposed to link Felicia's sexual obsessions to Erling's alcoholism. She gives two arguments against a realistic interpretation of the blood. First, the blood is never mentioned in the subsequent police investigation. It is "implausible" that this should just be an oversight on Sandemose's part—for this, he is too well acquainted both with police procedure and with the conventions of detective stories. Second, there is just a little blood—enough to carry the symbolic meaning, but not enough to indicate a real murder. Whatever you think of this response, it seems plain that counterfactual theories cannot make much sense of it.

8   Concluding remarks

Whatever we think about who "actually" killed Felicia, it should be obvious that both parties to the discussion reviewed take it for granted that they have a much wider array of arguments at their disposal than Lewis' or Currie's theories would allow. This means that the constraints imposed by counterfactual reasoning are not *all* the constraints on the fictional world—the fictional world is, so to speak, more determined than counterfactualism allows. And this, in turn, means that counterfactual reasoning from information given in the text cannot be taken to *define* what it is to be true in the fictional world.

The only reasonable alternative to counterfactualism, I think, is some form of intentionalism: the author decides what the fictional world is like, and then makes his or her decision known to us by a variety of clues. Counterfactual reasoning, from information given in the text, supplies one important type of clue. Among other kinds of clue, some are tied to more or less specific literary *genres,* some are stylistic, some may fall under the general heading of the symbolic, some may presuppose familiarity with the habits of the author—many authors, as they get older, count on an established friendship with a group of sympathetic readers—and so on.

This concludes my case against Lewis and Currie. I want to end, however, with offering a suggestion why a particular kind of counterfactual reasoning is particularly important for a particular kind of literature—namely literature that is expected to carry the sorts of extended "moral" meaning that I tried to indicate for Sandemose's novel, at the beginning of the paper. We can describe

the entire fictional story as a "signifier" for this moral "signified."

The relevant sort of semantic function has a prominent place in classical rhetorical theory—it is the *exemplum* or *paradigma* of Aristotle's *Rhetoric*.[15] Now, an important feature of this sort of signification is that it is not *arbitrary* but *motivated*. What sort of motivation is involved here?

Suppose that I want to let someone know that some other person is greedy. I can do that by simply saying so—I will have stated my opinion, and if my interlocutor trusts me, she will perhaps believe that I am right. But I can also use the more indirect method of telling an anecdote, a story, *displaying* the person's greed. Why should I do that? There are at least two advantages to the indirect method. First, I will not only have stated my opinion, I will also have given some *evidence* that it is true. The semantic function of *exemplification* is closely linked to the epistemological function of *induction*. From a logical point of view, the evidence is, of course, very weak—a single incident, chosen with malicious intent. But it will make it harder for my interlocutor to resist my view. It is not just my opinion any more, she will have to refute my evidence—by denying that the story is true, or that it is characteristic of the person described, or by offering another interpretation of it.

The second advantage to the method of the anecdote is that I do not have to *state* my opinion at all. If all goes well, the hearer will grasp my point, of course, and grasp that it *is* my point. But if something goes wrong, I have a way out—"oh no, I didn't mean to say that, it's just a funny story!"[16]

The standard objection to reading any sort of "message" or "moral meaning" into a work of fiction is the question why the author should go to the trouble of writing a story to convey that message, and not just *state* it, in so many words. Part of the answer to that objection is in the link between exemplification and induction. The story is not meant to appear as a vehicle of the author's opinion, it should *force* its conclusion on the reader.

In reading a novel like *The Werewolf* we are not just reading for the story, we want to know what the story means. And we readily distinguish between what Sandemose *wants* his story to mean, and what it *really* means. It is largely in order to have that discussion with the author, and among ourselves, that we read this sort of book at all. But to be able to meaningfully disagree about what the story means, we have to agree about what the story is—and so we let the author decide that.

But to make his own case as persuasive as he can, the author cannot make up any old story and convey it to us by any old means. If the story is too evidently made up to convey the message, if it is too "obvious" or "fabricated" we will lose interest and not go along. It has to be "real." And

---

[15]A modern version is Nelson Goodman's notion of "exemplification", in *Languages of Art* (Indianapolis 1968).

[16]Roland Barthes discusses this sort of semantic function, of the sign that does not admit to being a sign, in many places—most extensively in "Le myth, aujourd'hui," printed as a part of *Mythologies* (Paris 1957) .

one way for the story to be real, to achieve "the reality effect," is to cog into what we otherwise believe about the world—which means, among other things, that it should allow us to reason counterfactually from it, using our own beliefs about the world.

Staffan Carlshamre
Stockholm University
SE-106 91 Stockholm
Sweden
staffan.carlshamre@philosophy.su.se