7

POSSIBLE WORLDS AND COGNITION

Marie-Laure Ryan

Introduction

Possible worlds (PW) theory tells us that the contemplation of merely PWs – whether produced by counterfactual reasoning, by projections, or by the creative imagination – is as essential to cognition as forming representations of the actual world. This chapter focuses on the importance of the notion of PW for the question of fictionality. While one-world systems regard fiction as made of false statements, the many-worlds system that underlies PW theory allows propositions to be valued separately for each world. This allows the concept of truth to apply within fictional worlds. The chapter discusses the question of the completeness of fictional worlds, contrasts the cognitive implications of the rhetorical approach to fictionality with the PW approach, characterizes the user's cognitive stance as a recentering into the fictional world that allows immersion, addresses some of the criticisms that have been raised against the principle of minimal departure as a basis for imagining fictional worlds, and distinguishes three operations in the cognitive processing of fiction: imagining, rationalizing, and interpreting.

Literary studies has been marked in the past 20 years or so by a turn from language to world, which means a turn from asking what the essence of literary language is to investigating the ability of literature, and more particularly of narrative texts, to construct worlds that capture the imagination. This question extends beyond language-based works to all media and types of signs capable of narration, and beyond high culture to phenomena of popular culture such as the ability of certain story worlds to create a cult-like following and to inspire ever new expansions and modifications. The role of the critic, in this new paradigm, is no longer to act as the hero who reveals the deep meaning of literary texts to the masses of ordinary readers, as it was conceived in the heydays of deconstruction, but to capture the experience and mental activity that enable these ordinary readers to immerse themselves in the imaginary worlds created by narrative texts.

Possible Worlds Theory: An Overview

While not solely responsible for this turn, possible worlds (hence PW) theory is an important factor in the rise of the notion of world as focus of narrative experience. Though the theory has generated considerable activity among pure logicians, its philosophical barebones for those

interested in narrative applications are relatively few. The notion of PWs goes back to Leibniz, who famously argued that God contemplated all PWs and chose the best one to be actualized, though even the best one is not perfect because it contains suffering. Modern advocates of the concept of PWs deny being indebted to Leibniz, though they accept the idea that the actual world is unique, while non-actual PWs are many. The justification for this asymmetry is the intuitive belief that there is only one way things are, but there are many ways things could have been. The exact nature of actuality and the basis of the distinction between the actual world and merely possible ones is a matter of debate. Some philosophers regard the difference as ontological: the actual world differs in kind from other worlds because it is the only one that exists absolutely, independently of the human mind; merely PWs are non-existing objects produced by mental acts, such as imagining, dreaming, counterfactual reasoning, or producing fictions. As philosopher Nicholas Rescher (169) argued, in the case of the actual world, one can distinguish objectively obtaining facts ("dogs have tails") from the thought, or descriptive statement, that dogs have tails, so that even if the representation was never produced, the fact would remain, but in the case of non-actual PWs, this duality does not exist, and the state of affairs is entirely dependent on its mental representation.

An alternative position concerning the mode of existence of PWs is that of modal realism, defended by the philosopher David Lewis. According to Lewis, all PWs exist and exist in the same mode. The difference between the actual world and merely possible ones is a matter of indexicality, that is, of point of view:

> our actual world is only one world among others. We call it actual not because it differs in kind from all the rest but because it is the world we inhabit. The inhabitants of other worlds may truly call their own world actual, if they mean by 'actual' what we do; for the meaning we give to 'actual' is such that it refers at any world *i* to the world *i* itself. 'Actual' is indexical, like 'I' or 'here', or 'now': it depends for its reference on the circumstances of utterance, to wit the world where the utterance is located
>
> *("Truth in Fiction" 184)*

The opposition, or rather the complementarity, of these two conceptions of actuality has important consequences for the phenomenology of fictionality.

The concept of PWs was originally proposed by logicians as a tool for the semantic description of the modal operators of necessity and possibility. In a many-worlds model, propositions are assigned a separate truth value for each of the worlds of the system: a statement that is false in one world can be true in another world. A statement is possible if it is true in some but not all PWs; it is necessary when it is true in all PWs, and it is impossible when it is false in all of them. Modal logic is not limited to the operators of necessity, possibility, and impossibility, which constitute the so-called alethic system; it also comprises other systems involving reinterpretations of the basic operators. In an epistemic system, "necessary," "possible," and "impossible" are reinterpreted as "knowledge," "ignorance," and "belief"; in an axiological system, as "good," "bad," and "indifferent"; and in a deontic system, as "obligation," "prohibition," and "permission." In this last system, P is obligatory if it must be obeyed in all PWs; forbidden, if no world allows it; and permitted, if some worlds allow it and others do not. To be able to determine the truth value of every proposition, whether modalized or not, PWs cannot contain ontological gaps and areas of indeterminacy; they must be complete, maximal states of affairs.

For literary theory, or for cognitive studies, the notion of PWs has a much broader function than providing tools to logical semantics: the function of demonstrating and analyzing the importance for the human mind of constructing and contemplating imaginary worlds

that differ from the world we inhabit. Imagining non-existing states of affairs frees the human mind from the "here" and "now," this is to say, from immediately observable facts, and allows long-range planning, predictions, evaluation of past events, causal reasoning, innovation, discovery, and art. This ability distinguishes the human mind from the minds of other animals and accounts for its creativity.

PW theory has implications for two issues that have long been lumped together: the issues of fictionality and of narrativity.[1] The relevance of PW theory for narrativity lies in its explanatory power for the mechanics of plot. It consists, broadly speaking, of regarding fictional worlds, or more precisely, fictional universes, as autonomous modal systems centered on their own actual world, surrounded by the PWs of the characters' mental constructs, such as their knowledge, wishes, obligations, goals, and plans. The dynamics of plot is conceived as the attempt by characters to bring one or more of their private PWs in harmony with the actual world of the system or with the private worlds of other characters (Ryan *Possible Worlds*). But in this chapter, I will focus on the other issue, that of fictionality, because of its rich implications for cognitive studies.

Possible Worlds and Fictionality

PWs theory, as the name indicates, rests on an ontology made of many worlds and contrasts with one-world systems. In a one-world system, propositions are either true or false, depending on their value in the world, which is, by default, the actual world. In a many-worlds system, propositions are evaluated separately for each world, and what is false in the actual world can be true of a non-actual PW. This distinction provides a basis for a definition of fiction. According to the traditional one-world framework, fiction is conceived as being made of invention, this is to say, of false statements. Since Don Quixote and Emma Bovary are invented characters who do not exist in the actual world, and since reference presupposes existence, statements concerning these characters cannot refer to anything and therefore cannot be tested for their truth or falsity (which would be the position of Russell and Frege). From a collection of falsities, one cannot deduce any truth; therefore, in a one-world system, it is possible to speak of the truth *of* fiction, that is, of the message that fictional texts convey indirectly for the world, but not of truth *in* fiction.

One-world conceptions of fiction as lacking reference or as being made of invention still have wide currency in literary theory; the most recent manifestation of this approach is the "rhetorical theory of fictionality" (see Nielsen, Phelan, and Walsh, "Ten Theses"), to which I will return later. Yet one-world theories encounter a problem that philosopher John Woods, in the central argument of his book *The Logic of Fiction*, has described as "bet-sensitivity." If you bet that Emma Bovary was a rich aristocrat living in a castle or that she was faithful to her husband, you will lose your bet. On the other hand, if you bet that Emma was dissatisfied with her life as the wife of a country doctor, you will win, even though this statement appears nowhere in the text, so it cannot be reduced to the real-world statement, "Flaubert wrote that Emma was dissatisfied [etc.]." An efficient way to account for the bet-sensitivity of statements about fiction is to assume that a statement obviously false in the actual world (where there is no Emma) can be true in another world. This amounts to endorsing a many-worlds ontology and to recognizing the existence of fictional worlds, whether this existence is conceived as imaginative or objective, as it is by the advocates of modal realism. In this ontology, a given entity can exist in more than one world, with different or similar properties in each world, or be absent in some worlds and present in others. Thus, the proposition "Napoleon invaded Russia" is true in both the actual world and the world of *War and Peace* (1869); the proposition "Hitler lost World War II" is true in the real world but false in the world of Philip K. Dick's *The Man in the High Castle*

(1962), and the proposition "Emma Bovary was dissatisfied with her life" is true in the world of Flaubert's novel but false in the real world, since Emma does not exist in this world. (That there was a real-world model for Emma does not count.) We take all these judgments for granted, but it is the task of a theory of fiction to provide a justification for the seemingly obvious.

David Lewis proposes an account of fiction that provides an explanation for our intuitive attitude toward the statements listed above. He begins by observing that storytelling is pretense. For him, "the storyteller purports to be telling the truth about matters whereof he has knowledge. He purports to be talking about characters who are known to him and whom he refers to, typically by means of their ordinary proper names. But his story is fiction; he is not really doing these things" ("Truth in Fiction" 40). So, what are storytellers doing? They are playing a role, acting as if they were telling the truth, when in fact they are producing fiction. Narratologists will recognize this contrast of telling as fiction vs. telling as true as the opposition between the utterances of author and narrator. The author resides in the actual world, while the narrator resides in the fictional world: "Here at our world we have a fiction *f*, told in an act *a* of storytelling; at some other world we have an act a' of telling the truth about known matters of fact; the stories in a and a' match word for word, and the words have the same meaning" ("Truth in Fiction" 40). At the limit, an author could tell as fiction a text made almost entirely of true facts, since PWs can reside at any distance from the actual world. On the other hand, an author who presents as true in the actual world a text made of falsehoods would be guilty of deceit.

This definition has important cognitive implications. By postulating two worlds, one in which the text is presented as fiction and the other in which it is told as true, it allows a double perspective on fiction. To put this differently, fictional worlds can be contemplated from the external perspective of the real world, as constructed by the author, or from the internal perspective of a member of the fictional world who learns about it through the narrator's discourse. The external perspective allows an aesthetic appreciation of the author's performance, while the internal perspective enables the user to be caught up in the story, to react to the temporal unfolding of the plot through suspense, curiosity, and surprise, to form mental imagery of the setting and narrated scenes, to simulate mentally the gestures and actions of the characters, and to experience emotions such as empathy, pity, admiration, or disgust for them. The internal perspective is the basis for immersion, an experience that has temporal, spatial, and emotional manifestations.[2] A full appreciation of a fictional text requires a game of in-and-out through which users alternate between these two perspectives. I have called the "in" phase a recentering into the fictional world (Ryan, *Possible Worlds* Chapter 1). Through recentering, users imaginatively make themselves at home in the fictional world. Though they know that there is only one actual world, they let the fictional world take the place of the actual world and become actual in make-believe. This substitution is facilitated by the fact that fictional discourse does not use the markers of irreality available in language, such as the conditional mode or *if. . . then* constructions. Insofar as fictional discourse takes the form of factual assertions (even though its style or content may mark it as fictional), it denies its own status as fiction: as Lewis puts it, fiction is told as true, despite being generally false.

This telling as true makes no distinction between statements that are objectively true, i.e., true in the real world, and statements that are objectively false: they all participate in the mental construction of the fictional world, and it is not necessary for the reader to sort out facts from inventions. This does not mean that all statements expressed in the text should be regarded as true of the fictional world: the phenomenon of unreliable narration has shown that it is not the case. But if we take "telling as truth" in an intentional sense, even unreliable narrators – who are always individuated first-person narrators – satisfy this condition since they tell to be believed

whether or not they are believable. In the case of impersonal third-person narration, "telling as truth" is guaranteed by the impersonality of the narrator. As Doležel observes (145–68), third-person narrators have "authentification" powers that first-person narrators do not possess because third-person narrators are not engaged in an imitation of real-world speech acts. Being deprived of individuality, of human substance, they have no reason for distorting the truth, and they never need to provide evidence for their knowledge of the narrative facts. This uncontestable authority of third-person narrators, as well as the irreducibility of their discourse to the forms and norms of real-world communication, explains why some theorists have proposed a "narrator-optional" theory that does away with a narratorial figure in third-person impersonal narration (Banfield, *Unspeakable Sentences*; Patron, *The Narrator*; Walsh, *The Rhetoric of Fictionality* Chapter 4).[3]

Fictional Worlds: Complete or Incomplete?

The dual perspective theory of fictionality allows a solution to a problem that has preoccupied literary scholars for a long time: the problem of the ontological completeness of fictional worlds and of their inhabitants. Ontological completeness means that every proposition is either true or false of a given entity. As we have seen above, it is a tenet of PW theory that PWs are complete; otherwise, they could not be used to determine whether any given proposition is possible, impossible, or necessary. But many theorists regard fictional worlds as incomplete, even some who are inspired by PW theory, such as Doležel, Ronen, and Margolin. These theorists argue that fictional worlds are created by texts, and texts, being finite, cannot describe all the possible properties of a given entity. To take a famous example: we will never know how many children Lady Macbeth had because the text does not specify this number, and consequently, as philosopher Nicholas Wolterstorff argues, "there is no such thing to know" (133). From the external perspective, the perspective of the real world, fictional entities are indeed incomplete objects created by a finite text, but from the imaginative perspective of the recentered reader, they are autonomous beings who exist independently of the text. In this perspective, Lady Macbeth is a human being like you and me, and she has a determinate number of children. The exact number is treated by the imagination as missing information, not as an ontological gap. Without this assumption of completeness, it is hard to imagine why readers would experience emotions for fictional characters or why they would assume that characters can live beyond the end of the story (an assumption fundamental to all sequels). It is worth noting, incidentally, that even in the case of the PWs of logic, their completeness is not a matter of being provable by going through the list of all propositions and specifying their truth value – an infinite and therefore impossible task – but a matter of assumption. So, if the PWs of logic can be theoretically assumed to be complete, why not extend this assumption to most of the PWs of fiction?[4]

Possible Worlds Theory vs. the Rhetorical Theory of Fictionality

The cognitive implications of the conception of fictionality sketched above stand in sharp contrast to those of the rhetorical theory, which at the time of this writing was widely promoted in narratology.[5] Here are the tenets of the theory and the responses of the PW approach, in a nutshell:

1 Fictionality is not restricted to "generic fiction" such as novels and short stories but appears in many kinds of discourse that do not refer to facts: "hypotheticals, counterfactuals, speculations. and other deviations from the actual" (Nielsen, Phelan, and Walsh 64). For instance,

the proponents of the theory regard as fictional those statements in Martin Luther King Jr.'s "I have a dream" speech that describe a hypothetical future, though the speech as a whole is (I assume) non-fictional. The PW approach responds that there is no pretense in most non-factual statements. While not limiting fiction to literature (it occurs in film, comics, theater, and computer games), PW restricts fictionality to playful, self-contained (i.e., text-length) manifestations of make-believe and regards non-factuals as statements about the actual world,[6] rather than as world-creating discourse.

2 Fictionality consists of non-deceptive (i.e., signaled) invention. Since there are various non-invented facts in fiction, a text presented on the global level as fiction can contain non-fiction, while, conversely, globally factual texts can present islands of fictionality (the non-factuals mentioned under 1). A distinction is therefore made between global and local fictionality. The PW approach, by contrast, recognizes only global fictionality.[7] A text either describes the real world (whether accurately or not) or creates an alternate PW and does it for its own sake. In the second case, all the statements of the text contribute to the creation of this world, whether true or false in the actual world, and they are all part of the fiction.

3 Fictionality is not a world-creating device but a rhetorical resource for making statements about the real world. This resource consists of claiming truth through falsity, as do other rhetorical devices such as irony or metaphor. For instance, a fable invents a story about anthropomorphic animals to illustrate a moral that provides a useful guideline in real life. PW theory responds that popular narrative genres such as fantasy or romance, which are regarded by "high" literary theory as escapist, are no less fictional than didactic ones. While recognizing the possibility for fictional texts to make indirect claims about the real world, the PW approach regards the creation of a fictional world as inherently valuable. The contemplation of fictional worlds is a source of pleasure to the human mind, whether or not it conveys a message for the actual world.

Some Problems with Lewis' Definition of Fictionality

Like all theories of fiction, the PW-based theory has strengths and weaknesses. I hope to have suggested its strengths; as for the weaknesses, they have to do with the requirement "told as true" of another world. First, it may be objected that the formula limits fiction to verbal story-telling. Not all fictions tell stories; for instance, Beckett's novels or Dostoevsky's *Notes from the Underground* (1864) are ramblings of a fictional speaker rather than logically constructed narratives with a beginning, middle, and end. "Telling" must therefore be taken in a broad sense as "uttering." Second, not all fictions are language-based utterances. Media such as film, theater, comics, and computer games can be fictions, and they are based on showing rather than telling. To be extended to theater and film, Lewis' formula could be rewritten "here at our world we are watching actors impersonating characters; at another world we are witnesses to people living their life" (Ryan, *A New Anatomy* 44), but this formula does not work for language-based fiction, nor for computer games and comics. It would certainly be preferable to have a definition valid for all media. Third, how do we decide whether a text refers to the actual world or to an alternative possible one? If possible, worlds can be very much like our own world, and if fiction can be made in large part of true statements, fictionality is decided by an authorial intent signaled by a special framing, such as the label "novel," through which authors abdicate the responsibility of telling the truth about the actual world. But the genre of lyric poetry does not use fiction-indicating labels, or rather, the generic label "lyric poetry" is ambiguous with respect to fictionality: are poems fiction or statements concerning the real world, as the rhetorical theory suggests? For the rhetorical theory, as we have seen, real-world reference is

compatible with fictionality, but for the PW approach, it is not; therefore, deciding whether poetry is fictional is a matter of whether it creates its own world or describes reality. Poems like "Ode on a Grecian Urn" (1819) by Keats or "Les Chats" (1857) by Baudelaire could be seen as representations of actual objects. It is only when a poem creates imaginary personae, like the lovers Hatem and Suleika in Goethe's *West-Östlicher Diwan* (1819), that they fall squarely into the fictional category.[8]

Yet another problem with the "told as true" formula concerns fictions that cannot be told as true in any PW because they involve logical contradictions. The French New Novel, nonsense poetry, or postmodern texts such as Robert Coover's "The Babysitter" (1969) are classic examples. Yet these texts are obviously fictions. Could one say that they refer to "impossible possible worlds"? The expression is an oxymoron, and according to Lewis ("Truth in Fiction" 46), it should not be taken seriously. The alternative, suggested by Lewis in "Postscripts to Truth in Fiction" (279), is that contradictory texts represent fragments of different PWs. Lewis' formula "at some other world we have an act a' of telling the truth about known matters of fact" ("Truth in Fiction" 40) must therefore be replaced with "at other worlds we have acts a', a'' and a''' of telling the truth," each of these acts targeting a different world. A core set of statements – the intersection of all versions – will be true of all worlds. Taken as a whole, these texts do not project a fictional world but a semantic domain exploded into many mutually incompatible splinters. Narratologists such as Jan Alber (47–57) or myself (*A New Anatomy*) have pondered what readers actually do with such texts, and some strategies of naturalization have been proposed, by which readers work out contradictions in order to salvage a coherent world and a coherent plot.[9] I now believe that these efforts are doomed to fail in the case of intentional, obvious contradictions. Readers of experimental postmodern texts are not ordinary readers expecting a good plot or the immersive representation of a coherent world; they read out of interest for the experimentation itself, and they are on the lookout for self-referentiality. Engaging in a meta-reading, they find an inherent value in the subversion of traditional narrative pleasures, and they take contradictions for what they are rather than trying to resolve them.

The Principle of Minimal Departure

I am keeping for the end of this chapter the most significant cognitive implication of the PW approach to fiction: the question of truth in fiction. Many philosophers and critics recognize that what is "fictionally true" of a text goes far beyond what is explicitly stated or logically implied in the text, or even that the author can anticipate: as Umberto Eco puts it, "texts are lazy machines that ask someone to do part of their job" (214). Moreover, the imagination does not work from a blank state but rather relies on life experience to produce mental imagery and to make inferences about fictional worlds. These two observations combined lead to a need for a principle based on what Stacie Friend calls a "reality assumption" (29). Formal versions of this intuitive idea have been proposed by Kendall Walton[10] and David Lewis. Here, I will concentrate on Lewis' proposal, called Analysis 1 and 2, which I have rebranded as "the principle of minimal departure" (*Possible Worlds*, Chapter 3; hence MD). Analysis 1 reads: "a sentence of the form 'in the fiction f, Φ' is non-vacuously true iff some world where f is told as known fact and Φ is true differs less, on balance, from our actual world than does any world where f is told as known fact and Φ is not true" ("Truth in Fiction" 42). This does not mean that fictional worlds must be realistic, but rather, that in imagining fictional worlds, we do not make gratuitous changes from the actual world. For instance, we do not imagine that Charles Bovary is one-legged because we know that the vast majority of people have two legs. If Charles were one-legged, the text would say so. Peter Rabinowitz captures this idea in

informal terms: "a basic rule of reading" is that "all fiction, even the most fantastic, is realistic, except when it signals to the contrary" (342). Lewis' Analysis 2 revises Analysis 1 (at the cost of a much more complicated formulation) by specifying that the version of the actual world from which information is imported into the fictional world should be "one of the collective belief worlds of the community of origin" ("Truth in Fiction" 45). This means, practically, that "computers exist in Little Red Riding Hood" is false because the story originated in a community that preceded the invention of computers.

To see what is at stake with MD, consider the question, "Does Bordeaux exist in the world of *Madame Bovary*?" (The real-world towns of Rouen and Tostes are mentioned but not Bordeaux.) Reader A, an advocate of MD, would say, "The world of Madame Bovary is the world of nineteenth-century France; Bordeaux exists in this world; therefore, if it did not exist in *Madame Bovary*, this would be a needless departure." Reader B, a partisan of the rhetorical approach to fictionality, would say, "The text does not mention Bordeaux and no relevant interpretation can be built on the existence of Bordeaux. Therefore, Bordeaux does not exist in the world of *Madame Bovary*."

This opposition is dramatized in an episode of *Don Quixote* (1605 and 1615). An innkeeper asks Don Quixote if he has money to pay for a night at the inn. Don Quixote answers that he has none because he has never read in stories about knights errant that they carry money. The innkeeper replies that the need to pay for services is not mentioned in chivalric romances because it is too obvious; therefore, knights errant do carry money, and Don Quixote must pay if he wants to spend a night at the inn. Who is right? First, notice that in the world of the novel, patterned after sixteenth-century Spain, there are inns, and there is a madman who thinks he is a knight errant. By virtue of MD, inns in this world are like those of the real world; you need to pay to spend a night, and therefore, there is money in the world of the novel. For Don Quixote, who confuses the real world with the world of chivalric romances, there are neither inns nor money in chivalric romances; hence, he does not need money to pay the innkeeper. He may be right about chivalric romances, but he is, of course, wrong in transferring this assumption to the real world of the novel. As for the innkeeper, his reasoning sticks closely to MD. He is right to assume that if a novel has inns, readers can take it for granted that guests need money to pay; if all inns were free of charge in a fictional world, the text would say so. But in assuming that there are inns in chivalric romances, the innkeeper may be taking MD too literally: Chivalric romances typically do not focus on the needs of the body, unless knights are hurt in combat, and they mention neither inns nor money. This points to a difficulty of MD with existential statements. Yes, if there are inns in a fictional world, they must be like the inns of the real world unless otherwise specified, which means that there must also be money in this world to pay innkeepers; but should one automatically assume that because the real world has inns, all fictional worlds do too?

If one extends MD to existential statements ("there are x's"), this can lead to a proliferation of unwanted objects. It is this proliferation that Lewis Analysis 2 tries to stop: one cannot assume that there are computers in chivalric romances. But Analysis 2 does not exclude inns and money from chivalric romances. One way to solve this problem is to assume that the mental construction of fictional worlds not only depends on our experience of the real world but also on intertextual considerations, such as generic landscapes built on the basis of many texts or on explicit reference to other texts. An example of a generic landscape is that of fairy tales: readers will assume that there can be fairies, dragons, and witches; that fairies have magical power; that dragons spit fire and are evil; that some animals (but not all) are able to talk; and that love between heroes and princesses lasts forever. By the same reasoning, the generic landscape of chivalric romances would include some supernatural beings but exclude inns and money. An

example of intertextual reference is the case of transfictionality, a widespread phenomenon in contemporary culture (Saint-Gelais, 2011). Whether transfictionality consists of the expansion or modification of a preexisting fictional world, readers construct the transfictional world as the closest possible to the original, which itself is constructed as the closest to our actual world. For instance, it is true of the world of episode 7 of the *Star Wars* franchise, *The Force Awakens* (2015), that the Luke Skywalker of the movie is the son of Darth Vader, as revealed in episode 6, though no real-world knowledge nor episode 7 itself can establish this truth.

Critiques of Minimal Departure

Though MD and its relatives have been generally well received in narratology, they have not encountered universal acceptance. One objection is that MD reduces all fictional worlds to a realist paradigm and cannot therefore account for unnatural worlds. Some critics have tried to counter MD with a principle of maximal departure. In *Fictional Worlds*, Thomas Pavel invokes such a principle for texts that defy our ideas of human psychology, such as Georg Büchner's *Woyzeck* (1837) and Georges Bataille's *Madame Edwarda* (1941): "when confronted with fictions of this kind, we seem to anticipate a maximal departure, and look anxiously for its signs" (93). Following Pavel's suggestion, W. Michelle Wang writes that "under the principle of maximal departure, the fantastic, strange, and nonsensical (to borrow from J.R.R. Tolkien) are not anomalous but default modes for the fictional world" (134). I believe, however, that the notion of maximal departure rests on a false symmetry with MD. Both Pavel and Wang conceive maximal departure as an anticipation or as a retrospective judgment that relates to certain works or genres. But MD is much more than a text-specific judgment: it is a principle, applicable to *all* fictions,[11] that specifies what is true, what is false, and what is acceptable of a fictional text. It operates in fantastic as well as in realistic texts; for instance, when applied to "The Metamorphosis" (1915) by Kafka, it tells us that Gregor Samsa, now transformed into an insect, has six legs and can use them to climb walls. We may conclude after reading "The Metamorphosis" that its world differs considerably from the real world, but nevertheless, we imagine this world according to MD and make only the changes specified by the text. For instance, we do not infer that insects are able to talk. Moreover, while MD offers concrete guidelines to the imagination (use your knowledge of insects to imagine Gregor's body), maximal departure does not. What is maximally different from a human or an insect: a cow, a rainbow, a dragon, or the square root of -1? The worlds of fantastic works like Tolkien's *Lord of the Rings* (1954) or of Pavel's Büchner and Bataille examples have lots in common with the real world; otherwise, they would not be understandable. Even if we could define a maximality of departure, maximal and minimal departure would be birds of different feathers: one is a genre description, applicable to some texts; the other is a universal principle for assessing truth in fiction.[12]

Another objection to MD is that it fills fictional worlds with irrelevant clutter. Adopting the second position in my Bordeaux examples and defending the thesis of the fundamental incompleteness of fictional worlds, critics of MD such as Richard Walsh (*Rhetoric* and "Fictionality") and Sten Wistrand argue that there is no point in imagining fictional worlds beyond the sphere explicitly shown by the text. As Walsh writes:

> The horizon of the reader's encounter with a fiction is determined not by what it is possible to infer, but by what is worth inferring. The reader will not pursue inferential reasoning beyond the point at which it ceases to seem relevant to the particulars of the narrative, in a specific context of interpretation. (*Rhetoric* 18)

AU: Please check whether the formatting of the extract has been set correctly.

Elsewhere, he writes that "the desire to treat fictions as ontological wholes" turns the act of interpretation into a "runaway engine of world construction," "whether or not [this act is] formalized under a principle of minimal departure" ("Fictionality" 401). Walsh proposes to replace MD with a "principle of relevance" borrowed from linguistic pragmatics. According to this principle, verbal utterances are interpreted by the participants in the communicative act in a way that "maximizes relevance" with respect to the particular context. For instance, if somebody asks me, "can you pass the salt" during a dinner, I will understand that this person is requesting me to pass the salt, rather than asking about my physical abilities to do so. But while relevance in face-to-face communication is defined with respect to a specific, concrete situation, what is the standard for relevance in fictional communication? How are readers, in other words, going to establish the limits of relevance? In the absence of definable criteria (and Walsh does not provide any),[13] relevance is tautological: an interpretation is relevant when it is relevant.

I assume that a relevant inference makes it possible to explain the behavior of characters, and it does so by assuming that the characters are able to think and have the same reasons for acting as we do: they want to satisfy their interests (sometimes indirectly, as in the case of sacrifice), and they try to avoid undesirable situations. Even Martians, trolls, elves, and zombies act in that way.[14] To infer such motivations is to apply MD to non-human characters: they differ from humans through all sorts of criteria, and they may have different goals and ideals, but they think basically as we do. It is therefore necessary to apply MD to non-human characters to understand their motivation.

But how can one prevent MD from filling fictional worlds with undesirable objects and properties? My answer is that the principle operates in a fundamentally negative way. Inferences that violate the principle are automatically invalid: one cannot build an interpretation of *Madame Bovary* on the idea that Charles has only one leg or that Bordeaux is not part of France. But the principle does not tell us which inferences are relevant to a specific interpretation and which ones are not. Moreover, it remains uncommitted with respect to certain inferences: the world where Emma has blue eyes is as close to the actual world as the world where she has brown eyes. To adopt the concepts of alethic logic, some inferences are necessary, others are forbidden, and still others are permitted. Contrary to Walsh's claim, who wants to limit the act of imagination to those inferences that produce relevant interpretations so as to prevent the "engine of world construction" ("Fictionality" 401) from running out of control, MD recognizes the pleasure taken in world construction and welcomes imaginings but makes a distinction between those that are legitimate and those that are not. Rejecting MD or any of the variations of the "reality assumption" (Friend 29) would mean that the imagination does not work from life experience but constructs representations of fictional entities *ex nihilo*, an untenable conclusion.

The cognitive processing of fictional texts comprises at least three operations. First comes a stage of *imagination*, that is, of mental imagery formation, through which readers form multisensory, but mostly visual, representations of the setting, the characters, their movements, and their bodily interaction with the environment. This type of experience focuses on the situation presently shown by the text, and it is a major factor of immersion. Concurrently with imagination comes an operation of *rationalization*, through which readers construct the logic of the story. They do so by establishing the facts of the fictional world, by constructing the characters' desires, obligations, knowledge, fears, goals, and plans, and by assessing the consequences of their actions, or of accidental events, for both the fictionally actual world and for the private PWs of characters. Insofar as this form of cognition involves a predictive attitude (What will characters do next? What will be the result of their actions? How will they react to new information?), it is responsible for the fundamental narrative experiences of suspense, curiosity, and surprise (see Sternberg). The conjunction of imagination and

rationalization forms the precondition to emotional reactions to the fate of characters. Third comes a mostly retrospective operation of *interpretation*, through which readers evaluate the significance of the text with respect to some real-world issue either strongly suggested by the text or corresponding to the reader's personal interests, such as social and political issues, philosophical and ethical dilemmas, gender issues, human psychology, environmental concerns, or aesthetic considerations. Of these three cognitive operations, the first may be more or less extensive depending on the reader's particular inclinations; the third is optional and highly variable among readers,[15] but the second is mandatory and relatively fixed – most readers will arrive at similar rationalizations. Walsh's idea of relevance and MD concern different aspects of this layered process: MD affects the act of imagination by allowing readers to turn the necessarily incomplete representations of the text into rich mental images, and Walsh's idea of relevance concerns mainly the stage of interpretation. But both principles operate on the level of rationalization: MD, because readers assume that characters have the same reason for acting as we do and that objects have the same properties as in the real world, unless otherwise specified, and relevance, because rationalizations must respect narrative information. The importance of MD is that it is a guide for the imagination and not a procedure for extracting the meaning of texts. It deploys a fan of permissible inferences, but it is up to the reader to focus on some of them and to ignore others. There is no algorithm for the production of relevant interpretations, and this is fortunate, for otherwise human readers could be replaced by computers.

Conclusion

Cognitive approaches to literature or narrative have recently focused on affect, emotions, sense of embodiment, sense of identity, and subjective experience of the world. If there is one contribution of PW theory to cognitive studies, it is to remind us that logical reasoning is an essential part of the human mind, that truth judgments matter in both living our lives and in the processing of narrative texts, that these judgments should be based on knowledge rather than emotion-supported belief, as they are in the so-called post-truth era, and above all, that the contemplation of merely PWs – whether produced by counterfactual reasoning, by projections, or by the creative imagination – is as essential to cognition as forming representations of the actual world.

Notes

1 The landmarks in the application of PW theory to literary or narrative matters are *Truth in Fiction*, by David Lewis, *The Role of the Reader* by Umberto Eco, *Fictional Worlds* by Thomas Pavel, *Possible Worlds, Artificial Intelligence and Narrative Theory* by Marie-Laure Ryan, *Possible Worlds in Literary Theory* by Ruth Ronen, and *Heterocosmica* by Lubomír Doležel, though Pavel's, Ryan's, and Doležel's books are based on earlier articles. For a survey of the development of the movement, see the introduction to Bell and Ryan to *Possible Worlds Theory*.
2 See the chapter by Federico Pianzola in this companion.
3 It could be argued that Lewis' formula reduces the utterance of the fictional narrator to an imitation of real-world factual discourse, but I disagree with this conclusion. I interpret the formula as saying: anything that can be told as true fact of a fictional world can be told as fiction in the real world. What can be told as true fact of a fictional world is not limited to what is reliably tellable in the real world since third-person narrators, by virtue of their authentification power, can represent aspects of reality that are inaccessible to the members of the real world, such as the content of other minds.
4 I do not want to exclude incompleteness, especially in the case of absurdist texts built on contradictions.
5 The foundational text of the rhetorical approach is by Nielsen, Phelan and Walsh from 2015. The volume *Fictionality and Literature*, edited by Gammelgaard et al., ambitions to redefine these core

concepts in the light of the rhetorical theory. In addition, an entire issue of *Style* from 2019 is devoted to Richard Walsh's exposition of the theory, and to responses by leading scholars.

6 When used in fiction, nonfactuals refer to the fictionally actual world.

7 Fictional segments can be inserted in a globally non-fictional text, for instance a parable in a sermon, if it is clearly delimited and self-sufficient.

8 While poetry presents difficulties for the PW approach, it also does for the rhetorical approach, whose criterion of fictionality is "invention": is invention synonymous with "creativity," "originality of expression," or "attracting attention to itself" (all highly subjective judgments), or does it require reference to entities that do not exist in the real world? Then many poems fail this criterion.

9 For instance, mentalism (the contradictions are the products of a confused mind), virtualization (the contradictions represent different worlds), and do it yourself (using the text as a construction kit to build your own stories). Other proposed strategies are self-referential rather than naturalizing, such as reading allegorically and figural interpretation (see Ryan, *A New Anatomy* 159 and Alber, *Unnatural* 47–48).

10 Here is Walton's formula: "The fictionality of r1, . . ., rn (whether generated directly or indirectly) prima facie implies the fictionality of q if, and only if, were it the case that r1, . . ., rn, it would be the case that q." (147). He then refines this formula with a "principle of mutual belief" by adding "if and only if it is mutually believed in the artist's society that were r1, . . ., rn true, q would be true" (152).

11 Except those based on nonsense (see endnote 12).

12 The only fictions that render MD inoperative are those that are systematically built on contradiction, such as nonsense rhymes. Lewis' Analysis 1 says that any given statement of the form "In the fiction f, Φ" is "vacuously true iff there are no possible worlds where f is told as known fact" (42). In other words, everything and its opposite is true in nonsense texts.

13 An exception is the claim that fiction's relevance concerns the actual world (see Nielsen, Phelan, and Walsh).

14 But not the heroes of the theater of the absurd, such as Vladimir and Estragon in Beckett's *Waiting for Godot* (1953). This is precisely why this theater is absurdist.

15 As Catharine Abell writes in *Fiction: A Philosophical Analysis*, "ordinary audience members often arrive at an adequate understanding of works of fiction without succeeding in interpreting them" (15).

Works Cited

Abell, Catharine. *Fiction: A Philosophical Analysis*. Oxford UP, 2020.

Alber, Jan. *Unnatural Narrative: Impossible Worlds in Fiction and Drama*. U of Nebraska P, 2016.

Banfield, Ann. *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. 1982, Routledge, 2014.

Bell, Alice, and Marie-Laure Ryan, editors. *Possible Worlds Theory and Contemporary Narratology*. U of Nebraska P, 2019.

Doležel, Lubomír. *Heterocosmica: Fiction and Possible Worlds*. Johns Hopkins UP, 1998.

Eco, Umberto. *The Role of the Reader: Explorations in the Semiotics of Texts*. Indiana UP, 1984.

Friend, Stacie. "The Real Foundations of Fictional Worlds." *Australasian Journal of Philosophy*, vol. 95, no 1, 2017, pp. 29–42.

Gammelgaard, Lasse R., et al., editors. *Fictionality and Literature: Core Concepts Revisited*. The Ohio State UP, 2022.

Lewis, David. "Postface to Truth in Fiction." *Philosophical Papers I*. Oxford UP, 1983, pp. 276–79.

Lewis, David. "Truth in Fiction." *American Philosophical Quarterly*, vol. 15, 1978, pp. 37–46.

Margolin, Uri. "Individuals in Narrative Worlds: An Ontological Perspective." *Poetics Today*, vol. 11, no. 4, 1990, pp. 843–71.

Nielsen, Henrik Skov, James Phelan, and Richard Walsh. "Ten Theses about Fictionality." *Narrative*, vol. 23, no 1, 2015, pp. 61–73.

Patron, Sylvie. *The Narrator: A Problem in Narrative Theory*. 2009, U of Nebraska P, 2023.

Pavel, Thomas G. *Fictional Worlds*. Harvard UP, 1986.

Rabinowitz, Peter. *Before Readings: Narrative Conventions and the Politics of Interpretation*. Cornell UP, 1987.

Rescher, Nicolas. "The Ontology of the Possible." *The Possible and the Actual: Readings in the Metaphysics of Modality*. 1973, edited by Michael J. Loux, Cornell UP, 1979, pp. 166–81.

Ronen, Ruth. *Possible Worlds in Literary Theory*. Cambridge UP, 1994.

Ryan, Marie Laure. *A New Anatomy of Storyworlds: What Is, What If, As If*. Ohio State UP, 2022.

Ryan, Marie Laure. *Possible Worlds, Artificial Intelligence and Narrative Theory*. U of Indiana P, 1991.

Saint-Gelais, Richard. *Fictions transfuges: La transfictionalité et ses enjeux*. Seuil, 2011.

Sternberg, Meir. "Telling in Time: Chronology, Telelology, Narrativity." *Poetics Today*, vol. 13, no. 4, 1992, pp. 463–541.

Walsh, Richard. "Fictionality as Rhetoric: A Distinct Research Paradigm." *Style*, vol. 53, no. 4, 2019, pp. 397–425.

Walsh, Richard. *The Rhetoric of Fictionality: Narrative Theory and the Idea of Fiction*. Ohio State UP, 2007.

Walton, Kendall. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Harvard UP, 1990.

Wang, W. Michelle. "Postmodern Play with Worlds: The Case of *At Swim-Two-Birds*." *Possible Worlds Theory and Contemporary Narratology*, edited by Alice Bell and Marie-Laure Ryan, U of Nebraska P, 2019, pp. 132–56.

Wistrand, Sten. "Time for Departure? The Principle of Minimal Departure—A Critical Examination." *Disputable Core Concepts of Narrative Theory,* edited by Göran Rossholm and Christer Johansson, Peter Lang, 2012, pp. 15–44.

Wolterstorff, Nicholas. *Works and Worlds of Art*. Oxford UP, 1980.

Woods, John. *The Logic of Fiction*. Mouton, 1974.