Bart Geurts*
Fictional Commitments

"One morning, as Gregor Samsa awoke from restless dreams, he found himself transformed in his bed into a monstrous vermin." Thus begins one of the modern classics of fiction, Franz Kafka's *Metamorphosis*. In his story, Kafka creates a fictional reality in which the protagonist struggles to come to grips with the fact that he has turned into an oversized beetle. What kind of reality is that? There are various ways of answering that question, but I am specifically interested in two of them. The first is that Kafka's story creates a reality of sorts in the minds of his readers; the second is that the reality in question exists somehow between the author and his readership.

In his target article, Emar Maier claims that we must choose between these two views, and opts for the first, which he calls "psychologistic." But why must we choose? The story of Gregor Samsa's plight is at once a form of interaction between Kafka and his readers, and one that affects the minds of the latter. Why rule out the possibility that it creates some sort of reality on the social as well as the psychological plane? No reason at all, as far as I can see. So let's adopt an ecumenical stance, and in that spirit consider fiction from an interpersonal perspective. Doing so is of considerable interest in itself, but will also help to contextualize and clarify Maier's project, which takes a personal point of view, namely, that of the reader.

This ecumenical approach recommends itself more generally. To give just one example, suppose Jack says: "I'll do the dishes." On the one hand, we may construe his utterance as expressing an intention on his part; this is the personal, psychologistic perspective. On the other hand, the same speech act may be viewed from a social stance, as committing Jack to a certain course of action. Both stances capture something worthwhile capturing, and neither seems reducible to the other: insincere promises are as binding as sincere ones, and speech acts can bring about unintended commitments. The social aspect of speech acts deserves our attention as much as the psychological aspect does, and this holds not only for fictional statements, but across the illocutionary board.

*Corresponding author: Bart Geurts, Department of Philosophy, Radboud University Nijmegen, Postbus 9103, 6500 HD Nijmegen, The Netherlands, E-mail: brtgrts@gmail.com

In modern times, storytelling has mostly become a drawn-out process, in which production and consumption are separated by long stretches of time, sometimes even millennia. Yet, fiction remains a form of communication, i.e. social interaction, and that's how I propose to consider it here. Following a variegated tradition that includes Peirce (1934), Brandom (1994), Walton and Krabbe (1995), and Krifka (2014), among others, I view communication as a way of negotiating commitments. Promising is the paradigm. If Jack promises Jill to walk the dog, his speech act causes him to become committed to Jill to make it true that he will walk the dog. Due to Jack's committing himself in this way, Jill becomes entitled to act on the premise that Jack will walk the dog, and thus Jack's commitment helps Jill to coordinate her actions with Jack's. That's what commitments are for: they are coordination devices.

Though promising is the poster child for most commitment-based accounts, the idea that utterances create commitments for the speaker applies equally to other speech act types. Most importantly, for our current purposes, it applies to statements, too. If Jill tells Jack that she loves endive, for instance, then she constrains her future actions in all sorts of ways, since her utterance entitles Jack to expect that Jill will manifest her fondness of endive if and when appropriate, will abstain from anti-endive behaviour and so on.

This view on communication is geared to action in more than one way. Not only is an utterance an action in its own right, it also serves to constrain the speaker's future actions, and thus enable action coordination between speaker and addressee. An important class of these future actions are themselves speech acts. Jack's promise, for instance, not only commits him to walk the dog, but also to not wonder whether he will walk the dog, to be willing to confirm that he will do so, and so on. In short, an utterance always implies a pattern of actions the speaker commits himself to, and part of that pattern will consist of speech acts he is expected to perform, if appropriate, or to refrain from performing.

As Bratman (1987) has emphasized, commitments must be persistent in order to serve their coordinative purpose. If Jack commits himself at 10 a.m. to walk the dog at noon, he remains so committed at least until noon. This is not to say that his commitment is immutable: there are all sorts of circumstances that may absolve him of his obligation to walk the dog at noon. It will be understood, for instance, that Jack's commitment becomes void if either he or the dog breaks a leg. Commitments are persistent, but only by default.

Commitment is a three-place relation between two individuals, a and b, and a propositional content φ, which I read as "a is committed to b to act in accordance with φ," without implying that φ is true. Hence, commitments may be seen as propositional attitudes, though as long as $a \neq b$, they are social relationships rather than mental states. (If $a = b$, a is committed to himself to act in accordance with φ).

Such private commitments may result from self-directed speech acts, or "self talk"; they lie outside our current purview. See Geurts 2016 for discussion.)

As a rule, commitments are supposed to be shared: if a commits himself to b to act in accordance with φ, then ceteris paribus b is expected to reciprocate a's commitment, and will do so, ceteris paribus. Joint commitment may be defined in terms of shared commitment, as follows: a and b have a joint commitment to act in accordance with φ iff they:

(1) share a commitment to act in accordance with φ,
(2) share a commitment to act in accordance with (1),
(3) share a commitment to act in accordance with (2),
     and so on.

Within a commitment-based framework, joint commitment suggests itself as a natural candidate for defining the notion of common ground, which is widely agreed to be a cornerstone of communication, if not social interaction at large (Stalnaker 1973, 2002; Clark 1996). As a first stab, we may equate the common ground between interlocutors a and b, at a given time point t, with a and b's joint commitments at t. As a and b toss all manner of speech acts at one another, their common ground expands, while at the same time serving as the basis on which a and b perform their speech acts. For instance, if Jack promises to do the dishes, and Jill volunteers to help, then Jill's offer is predicated on Jack's commitment, which just entered their common ground.

This way of defining common ground is closely related to more established definitions in terms of mutual belief or mutual knowledge (Lewis 1969; Fagin et al. 1995; Stalnaker 2002), but there is an important difference, which has to do with the fact that commitment doesn't entail belief. If it is mutual belief (or knowledge) between a and b that φ is the case, then a and b must believe that φ. Contrariwise, a and b can be jointly committed to φ without believing φ to be true. Hence, on a commitment-based definition, the notion of common ground is doxastically neutral.

To explain why this may be a desirable feature, consider the following scenario. Due to heavy rainfall, the river will burst its banks within the next 48 hours. Given its situation, it is inevitable that the local school will be flooded, and therefore the school director and her staff are convening to draw up an emergency plan. If we call this scenario S_1, scenario S_2 is a variant on S_1, in which the school is safely situated on a hilltop, and in no danger of being flooded either now or ever. In S_2, director and staff are planning for that contingency by way of a team-building exercise; they are engaged in pretend play. Despite the differences between S_1 and S_2, it doesn't seem too far-fetched

to imagine that the team's deliberations are the same, word by word, in both scenarios, and that in each case they agree to raise furniture to a higher level, block drains, turn off water, gas, and electricity and so on. That is to say, in both scenarios, the team enter into the same commitments for the duration of the planning session. Specifically, for as long as their meeting lasts, they are committed to the truth of the proposition that the school will be flooded. But whereas in S_2 their commitments expire with the end of the planning session, in S_1 they are effectuated afterwards. Hence, whereas in S_1 commitment entails belief, in S_2 it doesn't.

There is a stark contrast between S_1 and S_2, which, drawing inspiration from Maier's distinction between regular and fictional statements, we might express by distinguishing between regular commitments and merely fictional ones. Very roughly, the suggestion is that, unlike their regular siblings, fictional commitments only hold provisionally. There are various ways of fleshing out this idea, but in any event the upshot will be that fictional commitments cause a split in the common ground. For, speakers who enter into fictional commitments do not thereby suspend all their regular commitments, but only some. In S_2, the interlocutors' commitments are largely the same as in S_1, except for the ones hanging on the counterfactual assumption that the school will be flooded. The space of fictional commitments in S_2 is demarcated from the regular commitment space, but the two spaces are not shut off from one another: apart from its location, the school is the same in both, as are the pupils and teachers, the laws of physics, last year's inflation rate, and so on. In brief, regular and fictional commitment spaces are distinct, but strongly connected, especially by virtue of the fact that they share inhabitants.

Whereas the school team's pretend play in S_2 is predominantly realistic, children's pretend play is further removed from reality, and literary fiction allows itself still greater degrees of freedom, sometimes to the point of logical inconsistency (Priest 2005). Literary fiction is special also in that it reduces turn taking to a minimum: the author makes all his commitments in advance, and reading is an act of consummation in which the reader comes to share the writer's commitments, one after the other. Still, whether literary or not, fiction is a social practice, a form of interaction between speakers and hearers, or writers and their readers, which is not fundamentally different from other types of discourse.

This approach to fiction is quite different from Maier's. Though our points of departure coincide in that we agree that an account of fiction must be embedded in a theory of speech acts, Maier's treatment of fictional speech acts is rather perfunctory. He begins by stipulating that fictional statements are "pretend assertions" rather than "regular" ones:

> Unlike regular assertions, fictional statements do not express information about the way the world is but rather invite the reader to imagine a certain state of affairs. (p. 4)

Adopting Walton's (1990) phrase, Maier characterizes fictional statements as "prescriptions to imagine," but instead of embarking on an in-depth analysis of this species of speech act, Maier concentrates his attention on how they affect the mind of the addressee:

> Intuitively, interpreting a fictional text is just not a matter of updating some abstract intersubjective common ground between speaker/writer and hearer/reader. To make sense of the philosophically compelling intuition that fictional statements are prescriptions to imagine, we need a framework that captures what happens in the mind of the reader when she interprets a text. (p. 6)

On the face of it, there isn't much that our accounts have in common. To mention only some of our differences, whereas Maier presupposes that fictional statements are a special kind of speech act, I would argue that they are continuous with regular statements. Also pace Maier, I claim that fiction is very much a matter of building common ground between speakers and writers, on the one hand, and their audiences, on the other. But most importantly, whereas I see fiction as a form of social interaction, Maier's principal concern is with the effects of fiction on its consumers' mental representations.

These are substantial differences, for sure, but most of them are overcome by embracing the ecumenical view I suggested at the beginning, and viewing Maier's enterprise and my own as complementary rather than contradictory. We can either view communication as social interaction or home in on the psychology of speakers and hearers, and there is no need to suppose that the two projects are irreconcilable (cf. Geurts and Rubio-Fernández 2015). Indeed, if both projects are worth pursuing, it behoves us to show how they can be reconciled. Specifically, we need to understand how the two stances hang together, and how explanatory labour is to be divided between them. Clearly, answering these questions is too tall an order for the present occasion (not to mention the present author), and in the remainder of this note I will confine myself to drawing out only some of the ramifications of the ecumenical view.

Maier's starting point is that a work of fiction invites its readership to imagine the world as being so-and-so. Like hopes and beliefs, imaginings are mental states, and they are connected with other mental states, which may include mental states of the same type; i.e. imaginings may be connected with each other. For instance, while reading Kafka's story, I imagine a reality in which a sales representative named "Gregor Samsa" finds himself transformed into a beetle. Pondering Gregor Samsa's sad fate, I imagine a happier course of

events in which the same man turned into a world-class DJ instead. Thus, I arrive at two imaginings, the first of which was prompted by Kafka's story, while the second is "parasitical" on the first, as Maier puts it. With some minor variations, Maier would represent this situation as follows:

(1)

IMG	x
	name(x, Gregor)
	turn.into.beetle(x)

IMG	
	turn.into.dj(x)
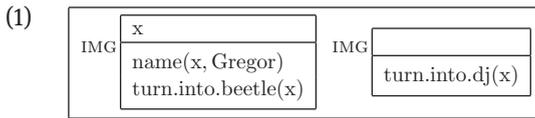
Citing Kamp (2015) and myself (1999) in support, Maier argues that such structures must be viewed as mental representations. That's the view I used to hold, too, but never entirely without misgivings. To explain my worry, and at the same time illustrate the generality of the issue, consider that hackneyed example:

(2)　　If a farmer owns a donkey, he beats it.

Within the framework of Kamp's (1981) discourse representation theory (DRT), the most natural reading of this sentence is represented as follows:

(3)

x y	
farmer(x)	
donkey(y)	$\Rightarrow$
own(x, y)	

	beat(x, y)
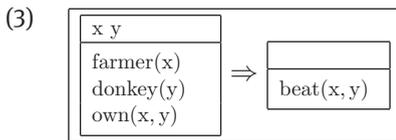
The structure of this representation is identical to that of (1) in two key respects: it separates its content into two boxes, which are connected by discourse referents that are introduced in one box and reused in the other. The main challenge of a dynamic semantics like DRT is to ensure that structures with these features are interpreted in a way that captures their intuitive meaning. Kamp (1981) showed how to do this for the original version of DRT; Maier does the same for a rather more expressive version.

　　It can hardly be denied that, like (1), the structure in (3) *may* be viewed as a mental representation, but the claim under discussion is that it *must* be viewed that way in order to account for the anaphora in (2). What has always worried me about this claim is that it seems to be too strong. Even if an utterance of a sentence like (2) affects the hearer's mental state, and even if it makes sense to ask how the hearer's mental state is affected, it seemed puzzling to me that we should have to refer to mental representations in order to account for the anaphoric links between "a farmer" and "a donkey," on the one hand, and the pronouns "he" and "it," on the

other. Armed with the insight that communication can and must be understood on more than one level, I believe I can justify my puzzlement.

Briefly, the argument goes as follows. The principal problem that DRT is designed to solve is to capture the content distributed over separate compartments connected by recurrent discourse referents. But as we have seen, commitment spaces present us with the exact same problem, and if the DRT framework is *essentially* mentalistic, that is, if structures like (1) and (3) *must* be viewed as mental representations, then that precludes a unified treatment of anaphora across the social and psychological levels; which is unsatisfactory, because the problem appears to be the same for both levels.

Let me unpack this argument with the help of another classic example:

(4)   Pedro owns a donkey. He beats it.

In this case, the problem is to explain the anaphoric links connecting "Pedro" with "he" and "a donkey" with "it," and I maintain that that problem presents itself both on the psychological and on the social level. In psychological terms, the hearer first has to deal with the speaker's belief in the proposition that Pedro owns a donkey, and then with her belief in the proposition that x beats y, where x and y are somehow connected to the content of the speaker's first belief. In social terms, the speaker commits herself first to the truth of the proposition that Pedro owns a donkey, and then to the truth of the proposition that x beats y, where x and y are somehow connected to the content of her first commitment. Since the two sets of facts are structurally identical, and we are dealing with propositional attitudes in both cases, one should expect the same account to apply to both. The claim that DRT's representational structures are essentially psychological stands in the way of such a unified account.

It appears, therefore, that the standard interpretation of DRT is overly concrete. It may seem natural to interpret its structures as mental representations, and on the psychological level such a construal may be defensible, but on the social level it isn't. A more abstract interpretation is called for, which allows for the possibility that the same structures are realized at both levels. Speakers and writers alter representations in the minds of their addressees, but at the same time their linguistic output alters social relationships between themselves and their audience, and those relationships have representational structure, too. Moreover, the representations emerging at the psychological and the social levels are similar both contentwise and structurally, even if their substrates are fundamentally different. This is what permits us to say that there is a sense in which Gregor Samsa exists in the minds of Kafka's readers, but there is also a sense in which the selfsame Gregor Samsa exists between Kafka and his readers.

References

Brandom, Robert. 1994. *Making it explicit*. Cambridge, MA: Harvard University Press.

Bratman, Michael E. 1987. *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.

Clark, Herbert. H. 1996. *Using language*. Cambridge: Cambridge University Press.

Fagin, Ronald, Joseph Y. Halpern, Yoram Moses & Moshe Y. Vardi. 1995. *Reasoning about knowledge*. Cambridge, MA: MIT Press.

Geurts, Bart. 1999. *Presuppositions and pronouns*. Oxford: Elsevier.

Geurts, Bart. 2016. Making sense of self talk. In *Review of Psychology and Philosophy* to appear.

Geurts, Bart & Paula Rubio-Fernández. 2015. Pragmatics and processing. *Ratio* 28. 446–469.

Kamp, Hans. 1981. A theory of truth and semantic representation. In J. Groenendijk et al. (eds.), *Formal methods in the study of language*, 277–322. Amsterdam: Mathematical Centre Tracts 135.

Kamp, Hans. 2015. Using proper names as intermediaries between labelled entity representations. *Erkenntnis* 80. 263–312.

Krifka, Manfred. 2014. Embedding illocutionary acts. In T. Roeper & M. Speas (eds.), *Recursion: Complexity in cognition*, 59–89. Cham: Springer.

Lewis, David K. 1969. *Convention*. Cambridge, MA: Harvard University Press.

Peirce, Charles S. 1934. Judgment and assertion. In C. Hartshorne & P. Weiss (eds.) *Collected papers of Charles Sanders Peirce*, Vol. 5, 293–312. Boston: Harvard University Press.

Priest, Graham. 2005. *Towards non-being: The logic and metaphysics of intentionality*. Oxford: Oxford University Press.

Stalnaker, Robert C. 1973. Presuppositions. *Journal of philosophical logic* 2. 447–457.

Stalnaker, Robert C. 2002. Common ground. *Linguistics and philosophy* 25. 701–721.

Walton, Doug N. & Erik C. Krabbe. 1995. *Commitment in dialogue: Basic concepts of interpersonal reasoning*. Albany: State University of New York Press.

Walton, Kendall. 1990. *Mimesis as make-believe: On the foundations of the representational arts*. Cambridge, MA: Harvard University Press.