

### Gilbert and Directions of Correspondence

In this short paper, I will discuss J. David Velleman's paper *The Self As Narrator*. More specifically, I will question his claim that if a robot can maintain correspondence, he can maintain it in both directions.

To quote Velleman: "A robot that can maintain correspondence in one direction, by saying that he's locked in the closet when he is, should be able to maintain correspondence the other direction, by going into the closet when he has said that he will." (10) I take this to mean that being able to maintain correspondence in the past direction (being able to say he's locked in the closet when he is) is sufficient for being able to maintain correspondence in the other direction (being able to go in the closet when he has said he will). However, I don't see why this is necessarily the case. Certainly we can imagine a robot who is only able to narrate what is currently happening to him, or what has just happened to him (including relevant causes and effects), but is unable to make predictions about what he is going to do next, to "narrate ahead of himself." (10) So, he can say things like "I'm locked in the closet" but not things like "I'm about to get up and go to the closet."

What makes him unable to narrate ahead of himself? Maybe he doesn't know what he is about to do until he actually does it. This could happen for a number of reasons. Perhaps Gilbert finds himself propelled along, unable to control his actions (maybe a human with a remote controls him), and so unable to predict what he will do next. Or, perhaps what controls Gilbert's actions is internal to him, but he is still unable to narrate ahead of himself, because he lacks epistemic

access to his control center's decisions. There could be a simpler explanation: maybe Gilbert's program doesn't include a "narrating ahead" function (or a number of functions which together enable him to narrate ahead).

Alternatively, we can suppose that Gilbert can narrate past and current events, *and* that he is able to make predictions about himself. However, he is not able to fulfill those predictions. So, Gilbert can say things like "I'm going to go to the closet now," but he doesn't go to the closet; he never performs the actions which correspond to his predictions. This hardly sounds like Gilbert is able to narrate ahead of himself, since he never performs the actions he says he will perform. Again, we can ask why Gilbert is unable to fulfill his predictions about himself, and some of the above reasons apply. Perhaps he is externally controlled.

Whatever the reason, it is at least conceivable that Gilbert can make predictions about himself but never fulfill these predictions. Likewise, it is possible that Gilbert is able only to narrate current and past events, not future ones. There is no reason I see that a robot that can maintain correspondence in the past direction *should* be able to maintain correspondence in the other direction.

Now, it might be that the reasons I've suggested for Gilbert's inability to either make or fulfill predictions are problematic (they prohibit us from saying all that we could or would want to say of Gilbert). However, it seems plausible that his inabilities could be explained in many ways, not all of which are problematic. Or, perhaps it enough for us to be able to imagine Gilbert's inabilities, without requiring an explanation of them.