Winograd and Realism

Michael P. Frank
Stanford University
SSP 201, Senior Seminar
Final Paper - FIRST DRAFT

July 16, 1991

Abstract

Terry Winograd, famous for his contributions to, and later rejection of, traditional approaches to Artificial Intelligence, maintains many philosophical views that have not yet been widely accepted in what Winograd would call the 'rationalistic' community. One of the most controversial of these is his Heideggerian rejection of 'realism', the ontological doctrine that objects of perception and cognition exist objectively outside the mind. In the ongoing debate on this issue among the Symbolic Systems faculty at Stanford, there has been much opposition to Winograd, partly resulting from some basic misunderstandings about the exact nature of his rejection of realism, including its motivations and implications. In this paper, I will carefully explain Winograd's view on realism, and show why I think it is reasonable.

1 Introduction

Winograd's point of view on realism has a philosophical side, and a pragmatic side. First, philosophically, Winograd does not believe that we can show that any such things as objects, properties, facts, and states of affairs exist objectively, outside of our thinking about them. Second, pragmatically, Winograd does not believe that philosophically postulating the existence of objective things, even if it were correct, would be relevant to the scientific study of mind and language or to the design of computer systems. Instead, Winograd says "the appropriate guide is one based on provisional realism grounded in an understanding of discourse and conversation." ([6], p.14.) First, I will discuss the philosophical side, and try to shed some light on some of the points in it that seem to be generally less well understood. Second, I will go over the pragmatic issues, and explain why I agree with Winograd that the proper approach to the science of mind is not one that claims to have a handle on "the things that really exist" outside of any context of discourse. Following that, I will say a few things about what Winograd's alternative is to the realist approach to cognition, and finally, I will mention some promising future directions in AI that are compatible with Winograd's philosophy.

2 The Philosophical Side

Winograd's philosophical view on realism can be construed as a kind of rejection of ontology. Fred Dretske, for instance, says that Winograd is "just doing epistemology," [3] i.e., just considering what we can know, and not what's really out there. Winograd seems to agree with this:

> Well, I think there's an element of truth in what he says, which is, I'm making the claim that ultimately we can't do ontology free of epistemology, that's it's not doable, to separate those. So to the extent that he says I'm therefore giving up doing ontology, I'm giving up a certain sense of what it would be to do ontology. [4]

However, Dretske also suggests that Winograd is "confusing epistemology with ontology," [3] but that I think is inaccurate. Winograd says [4] that in his "Three Responses" paper [6] he is presenting an ontology, namely, that objects do not exist; but he says that he believes in "the facticity of the world." [4] Thus, Winograd's ontology can be seen as one in which there is one existent: the world (or to use Maturana's term, the "medium"). This has been a point of confusion for many people upon first hearing Winograd's views, including myself. When hearing that "objects don't exist," the natural reaction for someone educated with the traditional (what Winograd calls "rationalistic") background is to conclude, "well, if objects don't exist, then nothing exists, and this must all be an illusion." For instance, Dretske warns, in reaction to Winograd, "Don't mistake the limitations on our methods for an absence of what those methods are designed to tell us about. I may not have any way of telling what's in distant stars in outer space, I wouldn't conclude from that that there are no distant stars, or that they're hollow." [3] However, that's not what Winograd is saying at all. The world is real; he simply questions the ability of our logical notions of discrete "things" to capture the true nature of that reality. Sure, the stars aren't hollow, but our conception of "star" as an object, and "hollow" as a property, are simply useful classifications, distinctions, simplifications that we project onto the real world, which need not be organized according to human ontologies.

Although Winograd does not believe that philosophers' complex ontologies reflect objective truths, he does not discard them as useless for all purposes. Winograd says that "ontology as an enterprise is very useful." [4] It's something that philosophers and scientists in our culture do that helps us organize and understand the world. Even everyday people do a loose form of ontology, wondering about the existence of God, or whether life is all a dream. However, Winograd warns:

> So now, the question is, what do you think you're doing when you do ontology. And if what you think you're doing is getting a handle on "the things that really exist," aside from the fact that you and your culture are engaged in the process of doing ontology, that's what I say is wrong.

This is what Winograd means when he says he's "giving up a certain sense of what it would be to do ontology." He is asking us to realize that even when we are trying to think philosophically about what exists, to use reason to get to the truth, the terms of our thought are always still embedded in a background, in a context that is impossible to totally reveal, to totally decontextualize. Thus we can never really claim to have gotten to the truth about the world.

There's another important point about the implications of Winograd's rejection of the objective existence of things in the world, a point that is often misunderstood. Dretske, and other philosophers to whom I've described the anti-realist view, seem to worry that if you deny the objective existence of objects and properties, then in a way, you are denying that world has any structure. Dretske: "If you take people away it [the world] doesn't become a kind of homogeneous bowl of jelly with no carving, no joints, as it were, what it is out there is it has exactly the joints we, given our conceptual scheme, take it to have." [3] To these people, if you claim the world has no objective boundary lines, then you seem to be claiming that the world is smooth, homogeneous, structureless. But that does not necessarily follow, and Winograd does not believe it. He says, "I agree, it's not jello," but he does not believe that the way we "carve up the jello with the kind of thinking we do" is objectively correct, that it gets at any truly existent things. To Winograd, the human mind may simply not be capable of capturing the true structure of reality:

> And see, in some sense, I guess one of my fundamental intuitive differences here from I think many philosophers is, I believe there's more out there than is understandable by the human mind... There is truth out there, right, in some sense whatever is, is; it's not imagination. Now, the notion that the structure of that is something which we can capture by having our set of properties, objects, and so on, is making a certain assumption about the capacity of the kind of reasoning and the kind of language we use to sort of get at what's "really there." And that's where I'm just not convinced, right? I think we're getting a good approximation to certain stuff which is practical, and so on, and I think ultimately there may be things which are just... you know, it's a lot bigger than we are.

3 The Pragmatic Side

Still, why doesn't Winograd think that realist ontologies are useful theories for CSLI researchers and other cognitive scientists to be studying? One might think, after all, even if one doesn't accept the objective truth of an ontology of objects and properties, or situations and predicates, it still seems like a very intuitive ontology, and perhaps one that people use internally when thinking about the world.

Actually, Winograd would agree with this; he sees situation theory and the like as perhaps a good formalization for dealing with natural language, for describing how people might think about the world. ([6], p.8). However, he says "the problem isn't with what Barwise and Perry are doing, but with what they say (and think) they are doing." To Winograd, Barwise and

Perry, in their work on situation theory, seem to set forth situation theory as a "true" ontology, something grounded in a posited "real world." Winograd would prefer a "theory of language in which meaning is grounded in the pattern of discourse itself." The ontology would be one of "provisional realism" that recognizes that although we have these everyday ontologies in which we take some things as objective, those things are really just that - taken to be objective, but not claimed, in the theory itself, to be truly objective. And indeed, this is reflected in the ontologies of everyday discourse, which, even though they take some things as objective, allow that "every element of that provisionally objective reality can be potentially open to re-articulation." ([6],p.14)

In other words, Winograd wants to draw this distinction between the ontology in the provisional realities that people adopt everyday, and the ontology that we believe as theoreticians to be true irregardless of people's everyday concerns. We can claim that the former is a profligate, open-ended ontology, like the parameterized formalisms in situation theory, reflecting the open-endedness of everyday discourse in which we move from one provisional reality to another. The latter ontology is something separate, and we should be careful not to confuse the two. To say that our reality is *really made up* of the indefinite set of possible distictions, the possible uniformities that we adopt in our everyday discourse, is a very strong additional claim. It is this this claim that Winograd not only rejects philosophically, as described above, but also rejects as irrelevant to our thinking about how people use their "provisional realisms" in their cognition and in their language.

Thus, when John Perry argues that there is a fundamental difference between a sentence's meaning really being true and it being false, and that this difference must be central to our theories of meaning, Winograd responds, "Of course, this distinction makes obvious sense. But so what? How is it relevant to anything we might care about as cognitive scientists or computer system designers?" ([4], p.11) For the objects of our study, the minds and the computers, can never get to the objective level at which meanings would formally "really be true" or not. Winograd is saying, yes, obviously "the way the world is" affects our perceptions, our possibilities for future action, and so on. But specifying this relationship, formalizing it in precise terms, is impossible, since we can never get to "the way the world is" as such; we can only work with provisional realities, every element of which is open to re-articulation. As an interesting illustration of this, we usually think of physics as having determined the objective truth of the reality of, say, subatomic particles. But Stephen Hawking, in [5], suggests an interesting alternative ontology, a re-articulation:

> We now know that Laplace's hopes of determinism cannot be realized, at least in the terms he had in mind. The uncertainty principle of quantum mechanics implies that certain pairs of quantities, such as the position and velocity of a particle, cannot both be predicted with complete accuracy.
>
> Quantum mechanics deals with this situation via a class of quantum theories in which particles don't have well-defined positions and velocities but are represented by a wave. These quantum theories are deterministic in the sense that they give laws for the evolution of the wave with time. Thus if one knows the wave at one time, one can calculate it at any other time. The unpredictable, random element comes in only when we try interpret the wave in terms of the positions and velocities of particles. *But maybe that is our mistake: maybe there are no particle positions and velocities, but only waves. It is just that we try to fit the waves to our preconceived ideas of positions and velocities.* The resulting mismatch is the cause of the apparent unpredictability. (pp.172-173)

(Emphasis added.) Winograd would go even farther than this, and say that there is no objectively correct interpretation that we can get to.

> In the end, if you say, well, do Schroedinger probability waves *really* exist or not, right, then that's where I'm saying, well, you've been misled by the nature of the question. They certainly exist in the discourse by which we understand our world, no question about that; and to say, well, they don't exist, but photons do, is making this assumption that there's this sort of external notion of existence which is separate from that, which is not ultimately doable. [4]

So for Winograd, even if the world were to have a formalizable structure, we could never approach that structure, either as the cognitive beings that are being studied or as the philosophers who are doing the studying. Thus, Winograd says, “If we are trying to come up with causal explanations of mental processes, the claim of realism may not be wrong, but it is vacuous. It simply doesn’t make any relevant difference.” ([6], p.11)

4 The Alternative

So what does Winograd see as the alternative methodology for explaining mental processes? They must be connected with the world somehow. If we can’t describe language and thought as being based on “meanings” that refer to or connect up with objects in the actual world, then we’d better be able to come up with an alternative story, an explanation of how we come to have mental processes and behaviors that actually work, given that we are embedded in a world that we have to deal with. (Bearing in mind that any such explanation will be based on a provisional ontology, suitable for the scientific discourse at hand, but will never be an “objectively correct” breakdown of the interaction between mind and world.)

Winograd’s answer draws upon the writings of the biologist Maturana, and is discussed extensively in chapter 4 of [7]. Although it is beyond the scope of this paper to describe Maturana’s perspective in detail, in short, it is this:

Organisms develop ways to deal with world through a process of “structural coupling,” in which organisms come to have a structure that works effectively in harmony with the medium in which the organism is embedded, effectively enough to keep the organism alive. Evolution is the prime example of this: biological structures arise that are capable of dealing with the environment because those that don’t, fail to survive. There is no need to presume that the evolved structures represent objects and situations with which the organism can cope, they simply *work*. So much for biology. (Of course, this is an extremely cursory characterization over a complex process that we are just beginning to understand.)

For human cognition, another structural coupling takes place, this time during human development and learning. Exactly what happens during this kind of development is perhaps even less well understood than for the case of evolution. But the basic idea is that we come to have mental processes, patterns of reasoning and of ways of breaking down the world, that seem to be useful for our interactions with the world and with other humans. (Of course, the mind is not a totally blank slate at birth; our genes also play a role in guiding this development.) Note that in this view, the grounding for our mental processes is not in the "objective correctness" of our mental methodologies, (for instance, our breaking the world down into objects whose reality we provisionally take to be objective), but rather simply in that our processes do tend to work; the kinds of mental models we construct come to approximate the world in a useful way, while also being communicable through language. Our cognition works because the structure of our cognition has formed through a long process of evolutionary, developmental, and social "structural coupling." Of course, figuring out the details of this story is obviously a huge task, but the direction, the orientation, of this approach is at least fairly clear, and seems amenable to further scientific exploration.

Note, in particular, that this style of explanation for cognition is not grounded on any sort of "theory of meaning" in which thoughts or utterances reduce to meanings, or in which meanings refer to specific things in the world. Instead, the grounding lies in the process of structural coupling, in which cognition develops as a way of dealing with our embeddedness in a complex medium, and language evoloves as a way of interacting with our fellow humans.

I find it very interesting that Fred Dretske, die-hard realist, finds that traditional theories of meaning have difficulty making the objective facticity of a meaning in the world play a causal role in explaining behavior (discussed in [2]), and that his solution to this is to say that where the facticity, the world, plays a causal role in the efficacy of meanings, is exactly just in the learning, the evolution, which gives the meaning a correspondence with the thing in the world it represents. In other words, Dretske concludes, on his own terms, that these ontologies of "facts" alone don't explain anything about cognition, and the solution to this is to talk about how the cognitive processes develop as a way of dealing with the world, a perspective reminiscent of Maturana's structural coupling. In this light, the only difference between Dretske's view and the structural-coupling perspective is that Dretske maintains the ontology of objects and facts, while Winograd rejects that as implausible and unnecessary to the theory. An interesting experiment might be to try to translate the picture of the role of learning presented in [2] into a form whose only difference from the original is that it does not presume the objective existence of some "facts" in the world corresponding to meanings.

5 Non-Realism and AI

Most researchers in AI are still pursuing a very traditional approach that takes some formalized domain as the primary reality within which the intelligent entity operates, thus presuming that the everyday world in which true general-purpose intelligent machines will live and work will ultimately be expressible as such a formalized domain. There is usually no provision for open-ended re-articulation of the very terms in which the world is described. There is generally no process of coupling with the environment, in which feedback allows the mental processes to evolve. However, certain new approaches such as connectionism and other machine-learning techniques are beginning to change these traditions and reveal new possibilities, although the techniques, and the learning processes they are trying to emulate, are still poorly understood.

Possibly one of the most radical departures within AI away from the traditional, realism-based philosophies can be found in the robotics work of Rodney Brooks at MIT. Brooks argues that "the *symbol system hypothesis* upon which *classical AI* is based is fundamentally flawed," (emphasis in original) and proposes "an alternative view... based on the *physical grounding hypothesis*." [1] The traditional "symbol system hypothesis" says cognition can be defined in terms of the processing of symbols, with traditional realist meanings and referents in the world. The "physical grounding hypothesis" bases cognition on interaction with the physical world. This idea is remarkably close to Winograd's view (derived from Heidegger and Maturana) that embeddedness in the world is the primary reality, and that cognition is based mainly upon that, and not on formal representations. There are many interesting things to say about Brooks' style of "non-representational AI," too many to present here, but suffice it to say that Brooks' work has had some notable successes, and that the similarity of Brooks' perspective to Winograd's thus lends credence to Winograd's claim that non-realist philosophy is a reasonable basis for understanding intelligence, although its sufficiency of course remains unproven, and will likely continue so for some time.

6 Conclusion

In this paper I have tried to explain exactly what Winograd's rejection of realism entails. Some common misconceptions about Winograd's view are hopefully eradicated, namely the myths that (1) Winograd doesn't believe the world is real, (2) non-realism implies the world is homogeneous, and (3) Winograd rejects ontology for all purposes, not just cognitive philosophy. In section 3, I have tried to explain why Winograd rejects realism for purposes of explaining the mind, and in section 4, I have gone over the alternative orientation for dealing with cognition. Section 5 mentions some future directions in AI consistent with Winograd's philosophy, and this conclusion briefly sums up all the sections without doing much else. As a personal note, I must admit that Winograd's perspective often seems weird and difficult at first, but whenever I think about these issues more and more, his view always seems to make more and more sense, while more traditional philosophy seems to almost become more confusing than before. I think Winograd may be right when he claims in [7] that Heideggerian philosophy seems to tap into deeper intuitions, in its basis in everyday experience, than the abstract, removed philosophies of the "rationalistic tradition." Of course, this is just my personal impression. Before I came to Stanford, I was somewhat biased against the methods of "traditional AI," despite wanting to pursue the goal of general-purpose AI. But, I was also biased against Winograd, since I saw him as an opponent of this goal. Since coming here, I have had more contact with traditionalists than with Winograd's ilk. However, the traditionalists have failed to impress me much with the classical methods, whereas Winograd has convinced me of many points that I would have vehemently disagreed with a few years ago. When thinking about philosophy of cognition on my own, I have often found that I could easily think of better arguments on Winograd's side of an issue, despite having had far more classes involving traditional philosophy. Could it be because Winograd is right? Whatever the answer, it's an important question for one who is planning a career in Artificial Intelligence research. Perhaps I should move from CS to Philosophy. Either way, I'm glad I took Symbolic Systems. (Thanks, Helen!) This sentence will be the last sentence in this paper until I write a final draft with a new conclusion (and other revisions).

References

[1] Rodney A. Brooks. Elephants don't play chess. *Robotics and Autonomous Systems*, 6:3–15, 1990.

[2] Fred Dretske. Putting information to work. chapter 4, pages 112–140.

[3] Fred Dretske and Michael P. Frank. Dialogue on realism. Unpublished transcript of personal discussion, March 1991.

[4] Michael P. Frank and Terry Winograd. Dialogue on realism. Unpublished transcript of personal discussion, March 1991.

[5] Stephen W. Hawking. *A Brief History of Time*. Bantam Books, 1988.

[6] Terry Winograd. Three responses to situation theory. Technical Report CSLI-87-106, Stanford University, Stanford, CA, 1987.

[7] Terry Winograd and Fernando Flores. *Understanding Computers and Cognition: A New Foundation for Design*. Addison-Wesley, 1986.