SELF-KNOWLEDGE AND SELF-REPRESENTATION

John Perry

Department of Philosophy and CSLI, Stanford

ABSTRACT

In this paper I introduce a contrast between homomorphic and non-homomorphic ascriptions of informational content to representations. In the former case there is a mapping from the parts of the representation onto the constituents of the content. In the latter case, there is not; some of the constituents of the content are settled bj background factors. I contrast this distinction with that between context dependent and context independent ascriptions of content. I note that in cases where the ascriber of content shares the background with the agent, one is inclined to ascribe homomorphic content of a sort that does not have a fixed truth-value to a representation. This leads to the notion of relative information. Some uses for relative information are noted. Finally, the distinctions developed are used to distinguish three types of self-knowledge and account for their relations.

I HOMOMORPHIC AND NON-HOMOMORPHIC REPRESENTATIONS

Philosophers of mind and language and researchers in artificial intelligence must confront the question of the content of representations. Philosophers want to understand what it is to know or believe or say something; researchers in AI want to build things that do know, believe and say things. In fact, most researchers in AI also share the philosophers goal; indeed, as far as I can tell, most of them are philosophers. But even if this were not so, understanding the nature of knowing, believing and saying ought to be helpful in learning how to construct things that have these capacities.

At one time, the paradigm for representations was the utterance of an eternal sentence, with a content fixed by meaning alone. More recently, we have realized the importance of context. The meaning *of* a representation typically fixes content only relative to context. In this paper I discuss a further factor, background, that interacts with context in interesting ways.

Consider my utterance of SAM WAS SLEEPING. This utterance has the informational content that Sam was sleeping. (I use informational content for that which is information if it gets things right, and call it "content" for short.) This content has, intuitively, three constituents: the individual Sam, the past time in question, and the property of being asleep. The sentence I used has a number of constituents: SAM, WAS, SLEEPING, and WAS SLEEPING. Among these are three that identify the constituents of the content. So, there is a homomorphism from the constituents of the sentence onto the constituents of the content.

This might seem inevitable. After all, how could constituents make their way into a proposition that an utterance expresses, unless some constituents of the uttered expression identified them (at least in a suitably generous notion of consitutents of an expression)?

But now consider another case. George sticks his left hand out the window of his car, thereby signalling that he will turn left at the next corner. The content has as constituents George, the next corner, and the relation of turning left at. We might reasonably say that the particular way he holds his arm stands for the relational activity of turning left at. But the way he holds his arm does not have constituents or aspects that identify the corner at which he is going to turn. Which corner is not, so to speak, a question that the structure of the signal needs to resolve. Once the angle of his arm

identifies right or left, we have all the information we need. The rest of the propoeitional constituents are not supplied by aspects of the signalling movement, but by what I shall call the background.

The background is typically determined by the context. We are imagining the signalling to occur in the U.S. in the last half of the twentieth century, where there is an institution of signalling for turns. This institution provides a propositional function for each context. The function takes directions (left, right) as arguments and returns a proposition: that the driver involved will turn (left, right) at the next corner.

This suggests the following picture. Communication takes place against a background, determined by context. A background provides a propositional function, taking some sort of entities as arguments. The job of the representation is to provide the necessary arguments to get from the background to a proposition. The case of uttering a complete declarative sentence "in vaccuo", so to speak, is the limiting case. The background is null: it provides only the identity function, from propositions to propositions. The constituents of the proposition are supplied by the background and the representation; homomorphic representation is just a special case, where the background is null.

Backgrounds of this sort may be established in various ways. In the case of the driver's signal, there is an institution that establishes a background for each driver in the vicinity of a corner. We may think of questions as providing backgrounds. The question "Who shot Liberty Valence?" provides a function from individuals and types of individuals to propositions. An answer need only specify an individual (Jack Palance, say) or a type of individual (some man with a large gun).

II BACKGROUND AND CONTEXT

The distinction between homomorphic and non-homomorphic representations cuts across that between eternal and context sensitive representations.

For example, if you ask me the square of 2 and I say "Four", my answer is eternal. "Four" stands for four in any context (or so we

may assume). But the representation is non-homomorphic, because the informational content, that the square of two is four, picks up constituents from the background.

When I say "I am sitting", the representation is homomorphic but context sensitive. If I say "John Perry is sitting on May 14, 1985", the representation is homomorphic and (arguably) sternal. If you ask "Who wants a chocolats milkshake?" and I reply "Me", my answer is both context sensitive and non-homomorphic.

Which background is relevant is generally a matter of context , but it need not be. We can imagine a convention of language, for example, that placing one's hand vigorously over one's heart supplies the answer "yes" to the question "Is America the home of the brave and the land of the free?" no matter who does it and when. In this case, we imagine the background being supplied by the meaning of the representation, and not its contsxt. (For some purposes, however, it is useful to think of the language being used as a contsxtual fact.)

Thus, the context of a representation is just the larger situation of which its use is a part. The context may contain all sorts of facts relevant to the content of the representation. The background for a representation is not determined by the meaning of any part of the representation, but by the meaning of the whole. The meaning may determine the background "eternally" as in our patriotic example. More commonly the meaning determines the background only relative to context. In the signalling case, the system of signalling determines a relation from contexts to backgrounds. For a signaller A approaching a corner C, the background is the partial function that takes a direction as argument and returns the proposition that A will turn in that direction at C.

III RELATIVE INFORMATION

When a background is supplied by context, the background may be more or less sensitive to shifts of context. For example, the statement "It's four o'clock" said by me now, sitting in Palo Alto, has the informational content that it is four o'clock Pacific Coast Time. We can imagine that "It" refers to a time. Then there really is not a property of times, being four o'clock, but only

a relation between times and zones. So my representation is non-homomorphic. The zone supplied as background does not vary over wide geographical regions, although it changes suddenly when it does.

For people in a single time zone talking to one another, there is no point in mentioning the time-zone. Indeed, people who grow up in one time-zone and do little traveling and don't watch TV, may function perfectly well without ever realizing that there are time-zones, and that being four o'clock is really a relation and not a property. They don't need the concept of a time-zone at all to make use of information about the time of day and to detect and communicate such information themselves.

In general, the meaning of a representation, a sentence of English, for example, can be seen as a relation between the sorts of situations in which it occurs and the content of those situations. Thus we can take the meaning of I AM SITTING to be that relation between utterances and contents that obtains if the content requires the speaker to be sitting. Given this picture, contents provide an equivalence relation among utterances that employ sentences with different meanings. So, the utterance where I utter I AM SITTING and the one where you utter YOU ARE SITTING while talking to me have the same content. This equivalence relation is important in understanding the flow of information, where informational content is preserved across changes of context. It is also important for understanding such psychological notions as continuing to believe the same thing (Barwise and Perry, 1983; Perry, 1980; Perry. 1985).

The phenomenon of non-homomorhic representation suggests that we need to broaden the notion of informational content, to include not just "propositions", that are true or false absolutely, but also various types of relative informational content, that are true and false relative to a background.

Consider communication about the time of day. So long as this takes place within a time-zone, there is little need to worry about the propositions expressed. If I ask Ingrid what time it is and she says "It's four o'clock", I do not need to think "She is in the Pacific Zone, so that means that it is fou: o'clock Pacific Coast Time." I just think ■It's four o'clock". We can say that she

communicated the information that it's four o'clock, where this is relative information, information that is true or false only relative to a background. The transaction is information-preserving only if the participants share a background. If Ingrid is talking to me long-distance from New York City and I don't realize it, the transaction will not be information preserving.

We might suppose (and I did for a long time) that we do not need to recognize relative information, since meanings, relationally conceived, give us entities that are true or false relative to context. But there are transactions in which meanings change systematically to preserve relative information. Suppose Ingrid says "It's four o'clock by my watch," and I tell you "It's four o'clock by Ingrid's watch." Meaning has not been preserved, and the shift in meaning does not automatically preserve non-relative information. What is preserved is information relative to a time-zone. We seem to need the notion of relative information to think about this transaction, and to be able to carefully characterize just what goes wrong when she made the original remark long-distance from New York, and my remark to you was false

IV SELF-KNOWLEDGE AND SELF-REPRESENTATION

I think the notions of non-homomorphic representation and relative information are crucial in a number of areas of epistemology and semantics. For example, Jon Barwise (Barwise. 1985) has recently analyzed conditionals as providing relative information about the three place relation among types of situations, T involves T' relative to T". T" is supplied by the background, and Barwise shows how some of the puzzles about conditionals involve inferences that are only valid when the background is kept fixed. I suspect these notions will be important in dealing with subjunctives, unbounded dependencies of various sorts, and other troublesome topics.

I want to end the paper by focussing on a particular topic, however, that will suggest a more or less deep reason why relative information and non-homomorphic representations are so important.

Let's return to our time-zone bounded

folk. We saw that they can communicate perfectly well about the time of day without having any words for or even concepts of time-zones. Because they live within a certain constant background, they have no need to worry about it. In particular their thought about the time of day can be keyed to certain perceptions and actions in a perfectly workable way, even though they lack such concepts. They look at the clock and think "It's five o'clock⁻, and so they close up shop and go home for supper.

Even those of us with a well-developed conceptual apparatus for dealing with time-zones operate, for the most part, in ways that allow our perceptions of the time of day and the actions we take in virtue of the time of day to ignore time-zones. Like the folk above, we go home when the clock shows five o'clock (or so).

I think these facts about time-zones are suggestive about certain basic facts of the human condition. The information we pick up through perception is always relative information. When I see a terminal before me, there need be nothing in my perceptual state that is indicative of its being me that the terminal is in front of. Not only is there not a constituent of the perception that "eternally[11] stands for me, there need not even be a constituent that picks me out in context. I am the background for my perceptions, and you are the background for yours. When you are in the very same perceptual state I am now in, you know that there is a terminal in front of you, not that there is one in front of me.

Imagine now a somewhat simpler organism than myself, perceiving a potato rather than a terminal in front of it. We can imagine this perception leading to a cognitive state, that in turn leads to the action of seizing the potato and jamming it into one's mouth. We can ascribe content to the action, in terms of its result. Each such organism, taking the action in question, will jam the potato into its own mouth, not someone else's. So we can think of actions as having non-homomorphically determined content, or homomorphically determined relative content. These organisms do not need to have any concept of themselves. They surely do not need any "eternal" idea of themselves, but they also do not need any internal indexical either. Since they are always in the background of their perceptions

and actions, they need not be represented in the cognitions that intervene between them.

We can suppose, then, that relative information is systematically connected with types of perception and action Equally important, we cannot imagine that non-relative information is systematically tied to them. That Elwood has a potato in front of him can't be something that all our little organisms know on the basis of being in the potato-in-front percpetual state, and can't be something the cognition of which leads them all to grab and shove. At any rate, if things worked this way, Elwood is the only member of the group that wouldn't go hungry.

These reflections suggest the following picture of our cognitive make-up. At the "bottom" level, we have cognitions that have no representation of ourselves (or the present moment), which are tied pretty directly to cognition and action. This gives us self-knolwedge of a sort: we know the world from our perspective. At the "top" level we have representations that are not systematically tied to perception and action (or at least not to the same sorts of perception and action), in virtue of which we have relatively context insensitive cognitions that homomorphically determine propositions about ourselves. This is self-knowledge of another sort. I have it if I read a note "John Perry must call home," where "John Perry" designates me. Note that I would have it even if I had forgotten my name, and didn't realize, as I might put it, that I was to call home. The real purpose of indexicals is to mediate between these levels, yielding full-blown self-knowledge (although not yet Socratic self-knowledge). When I read a note, "John Perry please call home," I think "I must call home" and then go into a state that we might express with "must call home". The step from the top level to the intermediate level varies from person to person; the step from tne intermediate level to the bottom level, and from it to action (modulo procrastination) is universal. The purpose of indexicals is to align the homomorphic representations we get through language and other forms of communication, at some kinds of memory, with a more basic, selfless, cognitive system.

One who has the bottom and top levels correctly linked knows who he is (where he is, what time it is). This is still oversimple, in

a variety of ways. We would need more levels, for example, to get at what happens when one realizes that it is four o'clock Eastern Time, hence one o'clock real time, hence one o'clock here, hence time to go to lunch. But I hope 1 have said enough to suggest that the topics of non-homomorphic representation and relative information are worth careful thought, whether or not the thought 1 have provided is careful enough.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Barwise, J. "The Situation in Logic~II: Conditionals and Conditional Information". CSLI Report No. 21, 1985.

[2]   Barwise, J.   and   J. Perry, Situations and Attitudes.   Bradford   Books,   Cambridge, Massachusetts: MIT Press, 1983.

[3]   Kaplan, D.   "On   the   Logic of Demonstrat- ives", in Perspectives   in the   Philosophy of Language, ed. French, P. A., T. E. Uehling,   Jr.,   and   H. K.   Wettstein. Minneapolis, 1979, pp. 383-412.

[4]   Perry, J.   "A Problem about   Continued Belief".   Pacific   Philosophical Quarterly 61, 1980, pp. 317-332.

[5]   —, "Perception, Action, and The Structure of   Believe-   in Grandy & Warner: Philosophic*;   Grounds   of   Rationality, Oxford. Oxford University Press, 1985, pp. 330-359.