# Metaphorically Simulating (Metaphorically Simulating) Self (ABSTRACT)

John Barnden
School of Computer Science
University of Birmingham
England, U.K.
http://www.cs.bham.ac.uk/~jab/

## 1 Metaphorical Self-Reflection as Option

One aspect of a complete agent, situated within anything like our world, must be the ability to reason about other complete agents and about itself as a complete agent. Now, as cognitive linguists and others have shown, much human discourse concerning agents is highly metaphorical. For example, we commonly talk about each others' minds as being physical containers, of each other as being made up of competing sub-people, of ideas as living creatures, etc. – all in perfectly mundane discourse, not (just) poetry and other literary art. A further tenet is that this discourse rests on metaphorical views that are crucial conceptual processing aids rather than just linguistic icing. Assuming that this is true, people's thinking, not just their discourse, about each other and about themselves is partly, and perhaps highly, metaphorical. It is therefore plausible to suggest that artificial complete agents should reason about each other and themselves in metaphorical terms. Note also that metaphor-based self-reasoning could include metaphor-based self-practical-reasoning, i.e. metaphor-based self-control.

## 2 Metaphorical Self-Reflection as Practical Necessity?

Indeed, it is even plausible to suggest that there is no practical alternative to doing substantial amounts of reasoning about mental states by means of metaphor, given

(a) the messiness of the world, other agents included,

(b) any complete agent's messiness even as perceived by its own internal reflective processing,

(c) the widely acknowledged ability of metaphor, when used appropriately, to provide more economical and precise description, and more effective reasoning, than is otherwise practical (or perhaps possible) in messy domains.

A further point is that

(d) a system may come gradually to have effective mechanisms of reflection by a process of learning partially based on self-observation; and this process could reasonably be expected to be related to the way we learn how to reason about other agents, a matter that could be partially mediated by existing discourse practices, including the use of metaphor.

## 3 Metaphorical Self-Reflection at One Remove

Moreover, even if an agent is not inclined to reason about itself on the basis of metaphor, it must have the ability to reason (in whatever way) about people reasoning about it in a metaphorical way. This is self-reasoning at one remove, so to speak.

## 4 A Relevant Implemented System

Part of the paper will describe a system called *ATT-Meta* that we have developed for conducting metaphor-based reasoning, and that has been applied largely to the special case of metaphor-based reasoning about mental states. For example, it can trace through implications of two ideas being "far apart" in a mind considered as a physical region. The techniques used in the system could also be used reflectively by a complete agent to reason about itself. The system can in addition be applied to reasoning about other agents' metaphorical reasoning, and in particular their metaphorical reasoning about the system itself. The system's metaphorical reasoning is done largely by a procedure akin to the mental-simulation procedure that is popular for reasoning about mental states. Indeed, mental simulation is ATT-Meta's main tool for reasoning non-metaphorically about agents' beliefs and reasoning.

## 5 Further Observations

Three further, linked, observations raised by the above considerations are as follows.

(1) Dominating Oneself with a Metaphor

If an agent's self-control is partly metaphor-based then, even if a particular metaphorical view that is involved in this self-control inaccurately describes the agent, the self-control may to some extent tend to make the system behave as if it were indeed accurately described by the metaphor. For example, if the metaphorical view fails to be sensitive to particular opportunities for external or internal actions by the agent, then self-control may be deprived of the opportunity for exploiting those possibilities, so that the agent does not in fact perform actions that it could in principle perform. (This means that it would be best to have a plethora of metaphorical views engaged in self-reflection/control, just as there are in ordinary discourse because of the diversity of mental life discoursed about.)

(2) Submitting Oneself to a Metaphor

If a complete agent is self-adaptive in the sense that aspects of it adjust to the activities of other aspects, then we can envisage the possibility of one aspect A of the agent "helpfully" becoming more like the way another aspect B perceives A as being like (much as people act stupidly if you convey that you think they're stupid). Thus, if B perceives A in a (partly) metaphor-based way, then A may adapt to become more like what the metaphorical view would expect. Thus, to continue the broad theme of (1), metaphor-based self-reflection could be to some extent a self-fulfilling prophecy.

(3) Metaphorical Qualia

Consider consciousness, in the maximally meaty sense of felt self-consciousness. If that self-consciousness in an agent partly rests on metaphorical views of itself, including its own mind, then, I claim, the self-consciousness could involve the agent feeling that its own mind is indeed as dictated by the metaphor. For instance, if one of the metaphorical views is MIND AS PHYSICAL CONTAINER, then (to some limited extent, and not all the time) the agent's mind could feel to the agent like a physical container. This feeling would be part of the reality of consciousness (because any feeling is automatically part of that reality), even though the metaphor would undoubtedly be misdescribing the underlying nature of the agent.

## References

http://www.cs.bham.ac.uk/ jab/ATT-Meta/