Embodiment *vs.* Memetics: Is Building a Human getting Easier?

Joanna Bryson*
*Artificial models of natural Intelligence
University of Bath, United Kingdom
jjb@cs.bath.ac.uk

Abstract

This heretical article suggests that while embodiment was key to evolving human culture, and clearly affects our thinking and word choice now (as do many things in our environment), our culture may have evolved to such a point that a purely memetic AI beast could pass the Turing test. Though making something just like a human would clearly require both embodiment *and* memetics, if we were forced to choose one or the other, memetics might actually be easier. This short paper argues this point, and discusses what it would take to move beyond current semantic priming results to a human-like agent.

1 Embodiment

There is no doubt that embodiment is a key part of human and animal intelligence. Many of the behaviours attributed to intelligence are in fact a simple physical consequence of an animal's skeletal and muscular constraints (Port and van Gelder, 1995; Paul, 2004). Taking a learning or planning perspective, the body can be considered as bias, constraint or (in Bayesian terms) a prior for both perception and action which facilitates an animal's search for appropriate behaviour (Bryson, 2001).

This influence continues, arguably through all stages of reasoning (Chrisley and Ziemke, 2002; Lakoff and Johnson, 1999) but certainly at least sometimes to the level of semantics. For example, Glenberg and Kaschak (2002) have demonstrated the *action-sentence compatibility effect*. That is, subjects take longer to signal comprehension of a sentence with a gesture in the opposite direction as the motion indicated in the sentence than if the motion and sentence are compatible. For example, given a joystick to signal an understanding of 'open the drawer', it is easier to signal comprehension by pulling the joystick towards you than pushing it away. Boroditsky and Ramscar (2002) have shown that comprehension of ambiguous temporal events are strongly influenced by the hearer's physical situation with respect to current or imagined tasks and journeys.

These sorts of advances have lead some to suggest that the reason for the to-date rather unimpressive state of natural language comprehension and production in Artificially Intelligent (AI) systems is a consequence of their lack of embodiment (Harnad, 1990; Brooks and Stein, 2004; Roy and Reiter, 2005). The suggestion is that, in order to be meaningful, concepts must be grounded in the elements of intelligence that produce either action or perception salient to action.

The pursuit of embodied AI has lead us to understand resource-bounded reasoning which explains apparently suboptimal or inconsistent decision-making in humans (Chapman, 1987). It has also helped us to understand the extent to which agents can rely on the external world as a resource for cognition — that perception can replace or at least supplement long-term memory, reasoning and model building (Brooks, 1991; Clark, 1997; Ballard et al., 1997; Clark and Chalmers, 1998). However, despite impressive advances in the state of artificial embodiment (e.g. Chernova and Veloso, 2004; Schaal et al., 2003; Kortenkamp et al., 1998), there have been no clear examples of artificial natural language systems improved by embodiment.

I believe this is because embodiment, while necessary, is not a sufficient explanation of semantics. We *have* seen neat examples of the embodied acquisition of limited semantic systems (e.g Steels and Vogt, 1997; Steels and Kaplan, 1999; Roy, 1999; Billard and Dautenhahn, 2000; Sidnera et al., 2005). These systems show not only that semantics can be established between embodied agents, but also the relation between the developed lexicon and the agents' physical plants and perception. However, such examples give us little idea of how words like INFIN-

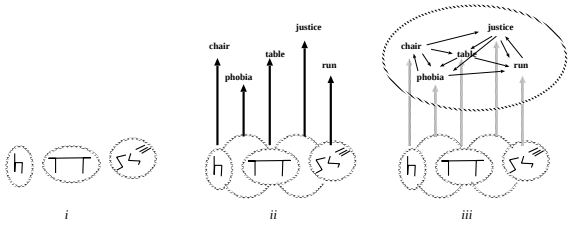
ITY, SOCIAL or REPRESENT might be represented. Further, they do not show the *necessity* of physical embodiment for a human-like level of comprehension of natural language semantics. On the other hand, it is possible that the semantic system underlying abstract words such as 'justice' may also be sufficient for terms originally referencing physical reality.

I do not contest the importance of understanding embodiment to understanding human intelligence as a whole. I *do* contest one of the prominent claims of the embodied intelligence movement — that embodiment is the only means of grounding semantics (Brooks and Stein, 2004). Roy and Reiter (2005) in fact *define* the term GROUNDED as 'embodied', which might be fine (compare with Harnad, 1990) if GROUNDED hadn't also come to be synonymous with MEANINGFUL. The central claim of this paper is that while embodiment may have been the origin of most semantic meaning, it is no longer the only source for accessing a great deal of it. Further, some words (including their meanings) may have evolved more or less *independently* of grounded experience.

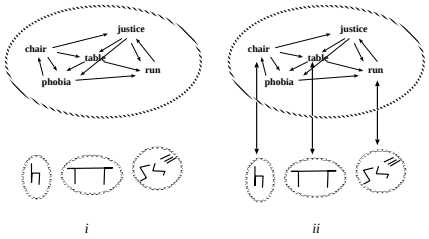
Figure 1: A two-dimensional projection of a semantic space, after Lowe (1997). The target words are taken from the experiments of Moss et al. (1995). Additional information on nearness is contained in the weights between locations in the 2-D space.

2 Memetics' role in development

We now know that humans could very well develop an interconnected web of words *independently* of the process of developing grounded concepts (See Figure 1). Grounding then becomes a process of associating *some* of these statistically acquired terms with embodiment-based concepts. Thus children can learn and even use the word JUSTICE without a referent. Gradually as they gain experience of complexity of conflicting social goals and notions of fairness develop a richer notion of both what the word means and how and when both the word and the grounded concept can be used in furthering their goals. But even before that, a relatively naive reference to the term could well accomplish something.



(a) Deacon's Theory



(b) Bryson's Hypothesis

Figure 2: In Deacon's theory, first concepts are learned *a(i)*, then labels for these concepts *a(ii)*, then a symbolic network somewhat like semantics *a(iii)*. I propose instead that grounded concepts and semantics are learned in parallel *b(i)*, then some semantic terms become understood *b(ii)*.

I want to be clear here: in my model, humans still acquire associations between these two representations, just as in the (Deacon, 1997) model that inspired it. What's different is the ordering. In my model, lexical semantics is learned in parallel with embodied categories and expressed behaviour. Subsequently, *some* words become grounded as connections are formed between the two representations (see Figure 2). Nevertheless, this model also leaves the door open to true memetics — perhaps *justice* is an evolved concept that has fundamental impact in our culture and institutions without anyone truly 'understanding' it in any deeply grounded way.

3 Building someone cheaply

The previous sections have talked about what composes current human intelligence. But let's change

the topic now to trying to build someone capable of a decent conversation, even of coming up with the occasional good idea apparently on their own. Someone that could pass the Turing test if you chatted to them at the bus stop for 20 minutes, assuming you couldn't see what they looked like.

Figure 2(b) implies that memetics can only give us half the story, but this is wrong on two counts. First, I do not think embodiment is necessary for concept formation. We develop concept for justice to go along with the label, and I expect this same process could go on for quite a lot of other words.

It is possible that we'd need to provide some preformed seed concepts to get the system rolling. This may be necessary for two reasons:

- Purely for bootstrapping the learning system. It's possible that all concepts formed from memetic experience *are* formed partially in relation to or contrast with established concepts, so our poor disembodied mind might need some good, rich precocial concepts to get started (see further Sloman and Chappell, 2005).

- Because our memetic culture might not carry knowledge *everyone* gets for free. Given that a huge amount of what it means to be human is embedded in our semantic assumptions, it is possible that the brain can fill in the gaps. Stroke and lesion patients sometimes recover enormous functionality deficits if they still have enough of their brain intact that they can use the existing bits. If sufficiently stimulated (the main point of therapy), these surviving parts can provide enough information about what the *missing* information should look like that the individual may recover some lost skills. However, it is possible that some concepts are so incredibly universal to human experience that there just isn't enough information in the culture to reconstruct them.

But in general, I still think it might be easier to program some concepts (or proto-concepts) by hand than to build and maintain a robot that is sufficiently robust and long-lived, and has a sufficiently rich motor and sensor capacities, that it could do a better job of learning such concepts from its embodied experience.

But the other reason Figure 2(b) is not showing us that memetics is half the story is because a very important part of the story is left out. Even if we had an agent with all the knowledge of a human (or say we had a search engine with all the knowledge any human has ever put on the web), if all that agent ever does is *learns*, it isn't very human-like. To build someone, we need not only basic capacities for perception and action (which in the meme machine's case is just language in and out) but also motivation and action selection (Bryson, 2001). Even the cheapest human-like agent would need to have a set of prioritised goals, probably some sort of emotional / temporally dependent state to oscillate appropriately between priorities, and a set of plans (in this case, syntax and dialog patterns) to order its actions in such a way that it can achieve those goals.

Fortunately, nearly everyone in AI who builds agents (even roboticists) builds this part of the system in software, so again, there is no driving reason to bring in embodiment. Of course, without a body these goals would have to be purely intellectual or social (find out about you, talk about me, figure out how to use new words appropriately) — many but not all human goals would be inaccessible to a disembodied meme machine.

4 Conclusion

This short paper argues that although embodiment is clearly involved in human thought and language usage, we have consequently evolved and developed a culture permeated with the knowledge we derive in our embodied existence, and as such a cheap but reasonably entertaining agent might be built with no embodiment at all. Of course AI has tried to do this for several decades, but I think they have come at it the wrong way, focusing on logic-based reasoning too much and case- or template-based reasoning too little. Humans however are imitation and case-learning machines — to such an extent that some of our wisdom / common sense may well have evolved memetically rather than ever having been fully understood or reasoned about by anyone.

Acknowledgements

I'd like to thank Push Singh for getting me to think about common sense, even though we never agreed what it meant. I'd like to thank Will Lowe for teaching me about statistical semantics, even though we still don't agree what the implications are.

References

Dana H. Ballard, Mary M. Hayhoe, Polly K. Pook, and Rajesh P. N. Rao. Deictic codes for the embodiment of cognition. *Brain and Behavioral Sciences*, 20(4), December 1997.

A. Billard and K. Dautenhahn. Experiments in social robotics: grounding and use of communication in autonomous agents. *Adap. Behav.*, 7(3/4), 2000.

Lera Boroditsky and Michael Ramscar. The roles of body and mind in abstract thought. *Psychological Science*, 13(2):185–188, 2002.

Rodney A. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.

Rodney A. Brooks and Lynn A. Stein. Building brains for bodies. *Aut. Robots*, 1(1):7–25, 2004.

Joanna J. Bryson. *Intelligence by Design: Principles of Modularity and Coordination for Engineering Complex Adaptive Agents*. PhD thesis, MIT, Department of EECS, Cambridge, MA, June 2001. AI Technical Report 2001-003.

David Chapman. Planning for conjunctive goals. *Artificial Intelligence*, 32:333–378, 1987.

S. Chernova and M. Veloso. Learning and using models of kicking motions for legged robots. In *The Proc. of ICRA-2004*, New Orleans, May 2004.

Ronald Chrisley and Tom Ziemke. Embodiment. In *Encyclopedia of Cognitive Science*, pages 1102–1108. Macmillan, 2002.

Andy Clark. *Being There: Putting Brain, Body and World Together Again*. MIT Press, Cambridge, MA, 1997.

Andy Clark and David Chalmers. The extended mind. *Analysis*, 58(1):7–19, 1998.

Terrence Deacon. *The Symbolic Species: The coevolution of language and the human brain*. W. W. Norton & Company, New York, 1997.

Arthur M. Glenberg and Michael P. Kaschak. Grounding language in action. *Psychonomic Bulletin & Review*, 9(3):558–565, 2002.

Stevan Harnad. The symbol grounding problem. *Physica D*, 42(1–3):335–346, 1990.

David Kortenkamp, R. Peter Bonasso, and Robin Murphy, editors. *Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems*. MIT Press, Cambridge, MA, 1998.

George Lakoff and Mark Johnson. *Philosophy in the Flesh: The embodied mind and its challenge to Western thought*. Basic Books, New York, 1999.

Will Lowe. Semantic representation and priming in a self-organizing lexicon. In J. A. Bullinaria, D. W. Glasspool, and G. Houghton, editors, *Proc. of the Fourth Neural Computation and Psychology Workshop: Connectionist Representations (NCPW4)*, pages 227–239, London, 1997. Springer.

H. E. Moss, R. K. Ostrin, L. K. Tyler, and W. D. Marslen-Wilson. Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21:863–883, 1995.

Chandana Paul. *Investigation of Morphology and Control in Biped Locomotion*. PhD thesis, Department of Computer Science, University of Zurich, Switzerland, 2004.

Robert F. Port and Timothy van Gelder, editors. *Mind as Motion: Explorations in the Dynamics of Cognition*. MIT Press, Cambridge, MA, 1995.

D. Roy and E. Reiter. Connecting language to the world. *Art. Intel.*, 167(1–2):1–12, 2005.

Deb Kumar Roy. *Learning from Sights and Sounds: A Computational Model*. PhD thesis, MIT, Media Laboratory, sep 1999.

Stefan Schaal, Auke Ijspeert, and Aude Billard. Computational approaches to motor learning by imitation. *Philosophical Transactions of the Royal Society of London, B*, 358(1431):537–547, 2003.

Candace L. Sidnera, Christopher Leea, Cory D. Kiddb, Neal Lesha, and Charles Richa. Explorations in engagement for humans and robots. *Artificial Intelligence*, 166(1–2):140–164, 2005.

Aaron Sloman and Jackie Chappell. The altricial-precocial spectrum for robots. In L. Pack Kaelbling and A. Saffiotti, editor, *Proceedings of the Nineteenth International Conference on Artificial Intelligence (IJCAI-05)*, pages 1187–1194, Edinburgh, August 2005. Professional Book Center.

Luc Steels and Frederic Kaplan. Bootstrapping grounded word semantics. In T. Briscoe, editor, *Linguistic evolution through language acquisition: formal and computational models*. Cambridge University Press., 1999.

Luc Steels and Paul Vogt. Grounding adaptive language games in robotic agents. In C. Husbands and I. Harvey, editors, *Proceedings of the Fourth European Conference on Artificial Life (ECAL97)*, London, 1997. MIT Press.