Help Me Help You: Bridging the Gaps in Human-Agent Collaboration — Papers from the AAAI 2011 Spring Symposium (SS-11-05)

Toward a Computational Model of ‘Context’

Wendelin Reich[*]
Fellow, *Swedish Collegium for Advanced Study* (SCAS)
wendelin.reich@swedishcollegium.se

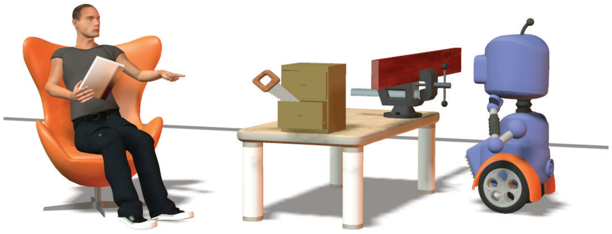
Introduction

According to a well-known statement by Donald Knuth, “Science is what we understand well enough to explain to a computer. Art is everything else we do.” If we use Knuth’s criterion, we must probably conclude that contemporary investigations of the ‘context’ of a social activity – as offered in fields such as linguistic pragmatics and social psychology – are still art rather than science (Reich 2010). The blunt truth is that social scientists have tended to treat context as a residual category that encompasses ‘everything else’ with regard to the object of interest (a communicative message, a social situation, and so on). The few explicit attempts to theorize context that do exist (e.g., Gumperz 1992; Sperber and Wilson 1995, chapter 3; Akman 2000; Dijk 2006) are general and exploratory, hence only marginally helpful for the purpose of specifying a precise, operational, implementable model.

Turning the elusive notion of context into such a model is a necessary task if researchers are to succeed in building artifical agents – be they robotic or virtual – that can interact and collaborate with humans on human terms (Bunt 2000). For instance, how is the robotic agent in figure 1 to distinguish between a pointing gesture that means (a) “Give me the tool” and one that means (b) “There is the tool you need”?

The purpose of this paper is to reinvigorate computational research on context by discussing three components of the context of extremely simple communicative acts in one specific social domain: *communicatively coordinated collaboration* (CCC). With this term I refer to a social situation in which two or more agents cooperate on a task by way of using *overtly intentional communicative acts* (which include not only speech acts but also nonverbal acts, indeed any behavior which is meant to influence a co-agent in an overt manner; Reich 2011). CCC can be contrasted with tacitly coordinated collaboration (TCC), which involves agents cooperating by watching and responding to each other’s pursuits without overtly intentional signals. Humans in all known societies engage frequently in CCC, and there is reason to expect that service-oriented artificial agents will have to share this ability if they are to gain acceptance by human users.

(a) `Imperative pointing: ``Give me ...''`



(b) `Declarative pointing: ``Here is ...''`

Figure 1: Context-sensitivity of communicative acts

The three components of context which I discuss below are derived from research on human social interaction in the cognitive and life sciences, especially evolutionary (social) psychology, paleoanthropology, cognitive neuroscience and primatology. Incidentally, the reader may want to know that this author’s primary affiliation is in the cognitive-scientific study of communicative interaction. But why should the reader, or any other researcher in AI and human-agent collaboration, be interested in what these fields have to say on

---

[*] The work reported here received generous support from Riksbankens Jubileumsfond (Bank of Sweden Tercentenary Fund). Valuable feedback was provided by Peter Gärdenfors and Margarita Martinez Mora.

Copyright © 2011, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

context in CCC? This issue will be discussed in the next section. The subsequent sections discuss the three components: action affordances (A), tool-mediated causal relationships (T) and social activities (S). Along the way I shall argue that agents which are able to process these three components in a sophisticated manner are essentially prepared to participate in basic CCC with humans. To keep things simple, I shall limit myself to 'first' communicative acts, i.e. ignore sequentiality.

Background: Three principles of cognitive evolution

Over the last decade or so, it has become increasingly clear that a large number of areas in the human brain are devoted to 'social' tasks, such as face recognition, 'Theory of Mind' (perspective taking), language, interaction competence, social emotions, and so forth (Frith 2007). As humans are, of course, evolved creatures, researchers within cognitive and neuroscience frequently draw on predictions from fields such as evolutionary biology and evolutionary psychology in order to develop hypotheses as to how areas within the 'social brain' might function.

Here I only mention three principles which have received significant corroboration and which are relevant to the discussion of context below. Firstly, the principle of functional continuity and cognitive parsimony. A large brain is a metabolically costly affair (Leonard et al. 2003), which means that evolution exerts perpetual pressures to the effect that new cognitive abilities (such as the ability to use stone tools or comprehend a summons gesture) be realized through maximal reuse of existing cerebral structures and minimal addition of new structures (Geary 2004; Tooby and Cosmides 2005). This tends to influence *how* the human brain realizes its many tricks – more often than not, it is by 'creative' rededication of cognitive abilities which humans already share with the great apes, and even much simpler animals. A consequence is that many seemingly 'complex' cognitive abilities are realized through surprisingly 'simple' computational mechanisms. A case in point is the innate ability of dogs – with a brainsize around $\frac{1}{10}$ of humans – to engage in basic CCC (to understand pointing gestures, to obey commands etc.). Various experiments have shown that this ability is not innately shared by dogs' closest relatives, wolves (Topal et al. 2009), which means that ca. 14.000 years of co-evolution with humans and minimal cerebral reorganization were sufficient to endow the dog brain with a sophisticated new ability.

Secondly, the principle of embodied cognition. In simple, evolutionarily ancient animals (e.g., nematodes), the nervous system is largely a device for motor control, hardwiring the organism to use movement in order to find food and to avoid threats. As predicted by the first principle, evolution built more complex cognition 'on top of' this ancient system for goal-directed foraging (Hills 2006). 'Embodied cognition' means that it did so by exploiting invariances in the animal's physiological layout and invariances in its (nonsocial and social) environment, which permitted keeping the complexity and cost of new cerebral structures to a minimum at each step (Pfeifer and Bongard 2006). One of the most publicized findings of neuroscience in recent years, the discovery of so-called 'mirror neurons' in various primates, fits well into this paradigm (Rizzolatti and Sinigaglia 2008). Mirror neurons are populations of neurons in the prefrontal cortex which fire both when the organism carries out a specific type of action and when it sees a conspecific carry out the same type (e.g., reach-for-object, grasp-object or ingest-food). However, the characteristic that renders mirror neurons relevant here is that they are selectively attuned to action-*goals* rather than mere movements. That is, when different movement trajectories are used by another organism to carry out the same action (e.g., reach-for-object), the same mirror neurons in the observer's brain fire, but when relatively similar movements are used to carry out different actions (e.g., grasping-for-placing vs. grasping-for-ingesting), different mirror neurons fire.

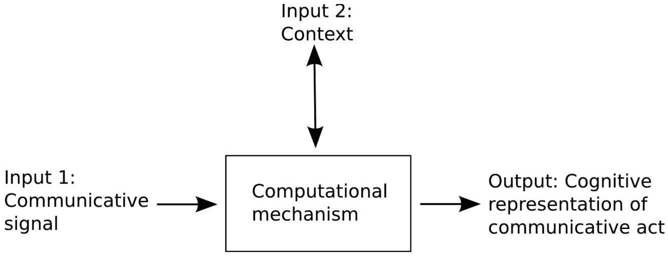
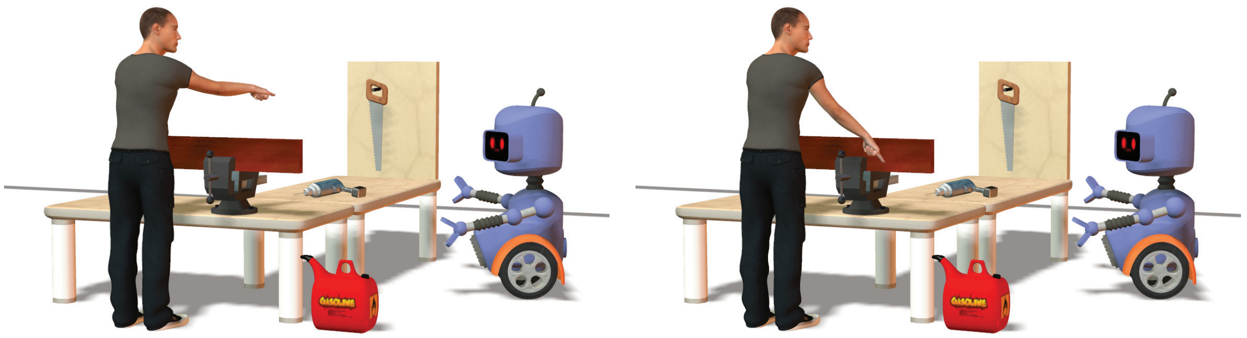


Figure 2: Schematic overview of the comprehension process

Thirdly, the hypothesis that human social skills were evolutionarily selected for in order to permit humans to engage in, and profit from, direct and indirect cooperation (Cosmides and Tooby 2005). Although many social skills *appear* to be altruistic, closer inspection and a large number of models and experiments in human biology have shown that they generally cohere with the Darwinian principle of the (unfortunately named) 'selfish gene'. Human communication and language skills constitute no exception to this rule. Various aspects of the human fossil record suggest that they first evolved ca. 2-3 million years ago when humans became obligate meat-eaters and obligate cooperators, allowing them to coordinate their hunting, foraging, nest-building, toolmaking, parenting and teaching activities (Reich 2011). Interestingly – especially for researchers of human-agent collaboration – this means that human interaction competence likely *originated* from CCC.[1] Presumably, it did so (i) in face-to-face settings (ii) where objects or other agents had to be manipulated and (iii) where this activity could be coordinated by use of primitive, pre-linguistic signals, such as pointing gestures, 'yes'/'no' vocalizations, 'come here' gestures, and so forth (see again figure 1).

---

[1] Which partially explains why a person such as this author has become interested in human-agent collaboration. Indeed, I believe that more cooperation between cognitive/life scientists on one hand and computer scientists/engineers on the other could and should demonstrate the *central* position of h.-a. c. in the broader study of communicative interaction in general, and in the study of human-computer/human-robot interaction in particular.



(a) `Context 1: pointing at saw` (b) `Context 2: pointing at gasoline`

Figure 3: Exploiting action affordances and tool-mediated causal relationships in CCC

What is context?

Which hypotheses about the structure of context in CCC can be derived from the three mentioned principles? To consider this issue, let us first look at a simplified overview of the computational architecture of communicative comprehension in figure 2. The end result, the cognitive representation of the communicative act made by the other interactant, depends on the communicative signal, on 'context' and on the cognitive mechanism through which the two are integrated. The integration of context (Input 2) after the perception of the signal (Input 1) is an active process in which contextual information can indicate the necessity to gather further contextual information (e.g., through active perception, such as turning one's head). It is important to keep in mind that the mechanism itself necessarily introduces "assumptions" about the end result. For instance, a primitive, evolutionarily ancient version of this mechanism could "assume" that an interactant who makes a pointing gesture – see again figure 1 – *either* wants something from the addressee (imperative pointing) *or* wants to help the addressee (declarative pointing), and that no third alternative exists under normal circumstances. However – and this is important yet easily overlooked – the mechanism cannot introduce information *sensu strictu*, that is, it cannot introduce *variation* into the output (see Shannon 1963). Signal and context are thus the only two sources of information that can account for differences in the end product, and because signals tend to be easily identifiable and their nature not subject to fundamental scientific debate, we can define context and mechanism against each other: 'Context' refers to explanantia of the meaning of a given signal that can vary between situations, 'mechanism' refers to explanantia that cannot vary in this manner. To demonstrate the utility of this conceptual differentiation, consider that it allows us to assign everything that has to do with social meta-reasoning (such as the kind of recursive social inference that is commonly emphasized in Gricean theories of comprehension; Grice 1957; Bara 2010) to mechanism rather than context. Whatever the other interactant may be thinking, intending, planning etc. is hence not part of context because it is not perceivable ("thoughts do not travel"; Sperber and Wilson 1995, p. 1). Consequently, a theory (or model) of context proper must be 'externalist'; it can only comprise sources of information which are perceivable or memorizable for the addressee in the current circumstances. In the following, I discuss three such sources of information which, I suggest, allow a collaborative agent who participates in CCC to comprehend a simple but realistic set of evolutionarily 'basic' communicative acts.

Component A: Action affordances

The first and most basic component of context that I shall discuss concerns the way in which agents who collaborate via face-to-face interaction share information about their physical environment. In line with principles one and two from the earlier section on 'Background', human (indeed, mammalian) object-recognition and scene-understanding have been shown to be organized in terms of action *affordances* (Gibson 1977; Rizzolatti and Sinigaglia 2008, p. 34f). That is, instead of simply perceiving features of their physical and social environment as abstract categories ('tool', 'apple', 'sibling' etc.), humans perceive them to a significant extent in terms of the sets of action opportunities 'afforded' by them ('graspable', 'liftable', 'edible', 'available-for-interaction' etc.). The aforementioned mirror neurons and their attunement to goals appear to be integrated into the perception of affordances in social settings.

What has largely gone unnoticed, however, is that action affordances are heavily exploited in CCC. To understand this claim, take a look at figure 3, which shows two minimal variations of figure 1a. The communicative sign is identical in both cases – the human agent is extending his arm in a pointing gesture. The referents vary: the affordances furnished by the saw in figure 3a render it integrable into the human agent's current activity (woodworking), whereas the affordances of the gasoline canister in 3b do not provide connection points to this activity. The robotic agent should thus be able to interpret the communicative act from 3a as meaning something like "Give me the saw" and the one from 3b as "Take away that canister", or something similar.

In other words, the communicative meaning of one and the same sign varies due to properties of physical objects in the interactants' vicinity – and, in the example given, *only* due to these properties. No complex social meta-reasoning along the lines of *'He knows that I know that he needs...'* is required to make the aforementioned distinction (*pace* an influential doctrine in linguistics, see only Bara 2010). Unsurprisingly for anyone with an interest in cognitive evolution, a seemingly complex ability (context-sensitive recognition of a pointing gesture) is built 'on top of' an older system (for action-oriented perception), allowing it to stay computationally lean.

Affordances are a relatively new concept for AI researchers, but a number of roboticists have recently proposed formalizations and/or implemented systems for recognizing and learning them (Cos, Canamero, and Hayes 2010; Ugur and Sahin 2010). Especially interesting is the formalization proposed by Sahin et al. (2007, p. 462), which uses the following structure (<> denote equivalence classes, which need not concern us here):

(<*effect*>, (<*agent*>, <(*entity*, *behavior*)>)).

Implemented on an embodied artificial agent, an affordance-oriented vision system might thus supply lists of action affordances for each recognized entity to higher cognitive-behavioral systems, instead of mere classifications for each entity. In the situation depicted in figure 3, this would result in definitions such as:

(<*in-hand*>, (<*robot*>, <(*saw*, *grasp*)>))
(<*in-hand*>, (<*robot*>, <(*canister*, *grasp*)>))
(<*emptied*>, (<*robot*>, <(*canister*, *pour*)>))
(<*in-hand*>, (<*human*>, <(*saw*, *grasp*)>))
(<*cut-board*>, (<*human*>, <(*saw*, *saw*)>))
...

As can be gathered from this list, affordance data can then be used by a comprehension system in order to compute the context-dependent meaning of an signal. This is especially obvious for the following types of (evolutionarily basic) communicative acts.

Imperative and declarative pointing. Conveys a request or proposal for the addressee to realize a salient action affordance of the pointed-at entity.

Object offer. Conveys a proposal to take control over the offered object and/or one of its salient affordances.

Object request. Conveys a request to turn over control of an object and/or one of its salient affordances to the requester.

It is especially noteworthy that such meanings can – in some cases, easily – be computed when affordance data are made available to comprehension algorithms which are based on the notion of *plan-recognition* (a notion which has attracted significant attention during recent years, see Carberry 2001; Geib and Goldman 2009). Affordances are also relevant for comprehending other types of communicative acts, but in order to see which, let us first look at two additional components of context.

Component T: Tool-mediated causal relationships

While action affordances are important as an initial input to the comprehension process, they still *under*determine the meaning of a communicative signal in CCC. The reason is that the causal effects they highlight may stand in potentially complex relationships to the goals of the signaling agent. Both the saw and the canister in figure 3 furnish more than one affordance, for instance. How is the comprehending robot to determine which of these is singled out by the human's pointing gesture?

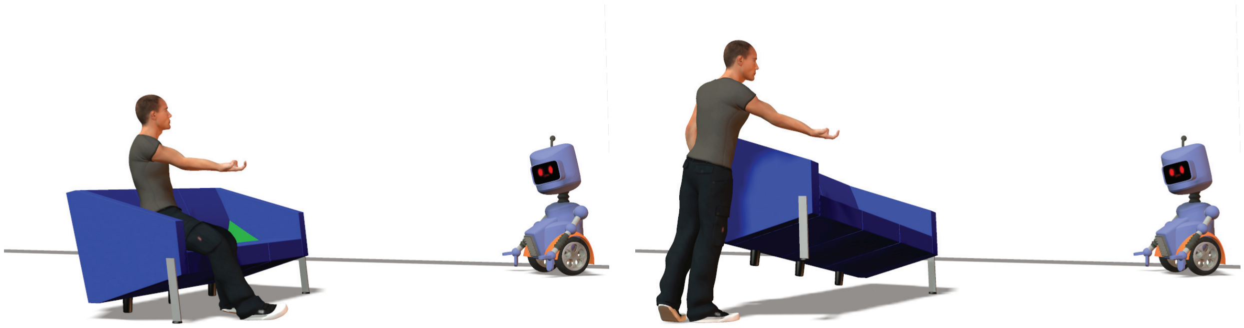
In order to solve this problem, the robot needs to employ causal reasoning. Such reasoning – which normally occurs at an automatic, pre-conscious level – is a hallmark of modern humans, which far exceed the great apes in this respect (Johnson-Frey 2003). Humans are also adapted to reason about causal relationships which are mediated by tools; specialized cortical structures are dedicated to this task (Frey 2008). Tools here are a highly general phenomenon, refering to any object or entity which is used intentionally to manipulate another entity (as opposed to using one's body to manipulate this entity directly). The type of role occupied by a tool in a directed, acyclic graph of causes and effects is potentially complex, and modeling it is a task which is currently not yet automatizable in a comprehensive manner (Rehnmark et al. 2005, p. 34).

An important distinction to consider when attempting to model tool-mediated causal relationships – e.g., through causal Bayesian networks (Pearl 2009) – is that between technical tools and agent tools. The former refer to 'tools' in the conventional sense of the term, while the latter refer to the use of other individuals as 'tools' in TCC or CCC. Among the primates, overtly intentional use of conspecifics as voluntary agent tools is unique to humans, who likely evolved this ability in order to cooperate more effectively and flexibly (recall the 3rd principle from the section on 'Background'). Various aspects of the human fossil record suggest that this characteristic evolved at about the same time as the use of prepared stone tools. A working assumption of current research in cognitive science is thus that technical tools and agent tools are perceived, reasoned about and used in virtue of overlapping cortical structures (Ferrari, Rozzi, and Fogassi 2005; Frey 2008).

By definition, all CCC involves use of (agent) tools in one way or another (Reich 2011). Tool-mediated causal relationships thus form a requisite part of the context of CCC. Considering once again figure 3b, the robot can use causal reasoning to discount the interpretation 'Give me the [graspable!] canister' for the reason that the canister is already in the reachable vicinity of the human, rendering it unnecessary to solicit help from the robot. Of course, its action affordances are also incompatible with the activity the human is currently engaged in – which brings me to the final component of context to be discussed here.

Component S: Social activities

Like recognition of action affordances, but unlike that of complex tool-mediated causal relationships, recognition of



(a) `Context 1: activity ‘relaxing on couch’` (b) `Context 2: activity ‘moving couch’`

Figure 4: Exploiting knowledge about the current social activity in CCC

ongoing social activities in one’s primary group is an ability we share with our primate cousins. In this way, groups are frequently able to share activities such as foraging, hunting (which chimps do), fleeing, playing or grooming. CCC is not another activity type but a coordinative mechanism that can be used in any of the aforementioned activity types, and many additional ones that require cooperation.

What is an activity type? There is no general answer to that, but we may not need one in order to make hypotheses about the role of shared social activities in determining the meaning of a communicative act. For illustration, consider the two situations depicted in figure 4. Both portray a human agent who solicits an artificial agent by way of a summons gesture. In 4a, the human agent appears to be ‘relaxing on the couch’, or be engaged in an otherwise unspecified/generic activity. A parsimonious comprehension mechanism should therefore conclude that he merely wants the artificial agent to approach. The activity in 4b is more specific – the human appears to be trying to move the couch. Here, the interpretation process should determine that the artificial agent is requested to drive toward the other end of the couch and lift it as well. This shows that the computational role of a current activity – as part of ‘context’ – is to *prune* the *relevant* computational ‘search space’ of affordances and causal relationships for recognizing collaborative actions. The physical space right before the human affords ‘standing on’, but the space before the other end of the couch affords the more specific, and expectably desirable, action of ‘standing on and helping to move the couch’.

I hypothesize that components A, T and S together permit an observing agent to make sense of a sizeable number of evolutionarily basic communicative acts, including not only **imperative-/declarative-pointing**, **object-offer/-request** and **summoning**, but also: **ego-attention-request** (i.e., getting the other to look at signaler), **relocation-request** (directing the other to a different location), **sending-away**, **get-down-/-up-request**, **forestalling/prohibition**, **slow-down-request**, **hush**, **selection-offer**, **request-for-affirmation**, **affirmation**, **rejection/refusal**, **approval** and **disapproval** (Reich 2010). For instance, approval and forestalling/prohibition presuppose that the two agents share knowledge of the activity which the addressee is currently engaged in, which often includes interaction with objects or even tools. How the comprehension mechanism integrates such information in the process of interpreting a newly perceived communicative act (see again figure 2) should be an interesting issue for interdisciplinary research.

Here, I only mention that a number of suprisingly robust systems for automatic recognition of current activities have been deployed (Bao and Intille 2004; Wojek, Nickel, and Stiefelhagen 2006; Gu et al. 2009). A reason behind this success seems to be the rich availability of (behavioral) feature-data for differentiating among different activities, which allows these systems to make use of statistical techniques for pattern recognition (Kim, Helal, and Cook 2010). However, when it comes to recognizing social activities as part of the context of a communicative act, the set of behavioral features will often be leaner, making the current, statistical approach more difficult. For instance, the human agent in figure 4a can change his activity to the one of figure 4b by simply getting up and lifting the couch. Such changes may turn out to be difficult to detect without a more explicit structural/causal representation of the concept of ‘activity’ within the comprehending agent.

Conclusion

The hypotheses offered in this paper are of a conceptual nature, and they are compatible with different technical approaches for implementing a system for communicative comprehension within human-agent collaboration. Action affordances, as perhaps the most primitive part of the context-model that I have argued for, have previously been formalized in a way that makes them integrable in planning and plan-recognition systems (Sahin et al. 2007), but I do not wish to imply that other types of architectures can be discarded. At any rate, the position offered here is certainly just a starting point, but hopefully one that indicates which kinds of contextual information system-builders could focus on in order to allow an artificial agent to participate in CCC.

References

Akman, V. 2000. Rethinking context as a social construct. *Journal of Pragmatics* 32:743–759.

Bao, L., and Intille, S. S. 2004. Activity recognition from user-annotated acceleration data. In Ferscha, A., and Mattern, F., eds., *PERVASIVE 2004, LNCS 3001*. Springer. 1–17.

Bara, B. G. 2010. *Cognitive Pragmatics: The Mental Processes of Communication*. MIT Press.

Bunt, H. 2000. Dialogue pragmatics and context specification. In Bunt, H., and Black, W., eds., *Abduction, Belief and Context in Dialogue: Studies in Computational Pragmatics*. Amsterdam: Benjamins. 81–150.

Carberry, S. 2001. Techniques for plan recognition. *User Modeling and User-Adapted Interaction* 11(1):31–48.

Cos, I.; Canamero, L.; and Hayes, G. M. 2010. Learning affordances of consummatory behaviors: Motivation-driven adaptive perception. *Adaptive Behavior* 18(3-4):285–314.

Cosmides, L., and Tooby, J. 2005. Neurocognitive adaptations designed for social exchange. In Buss, D. M., ed., *The Handbook of Evolutionary Psychology*. Chichester: Wiley. 584–627.

Dijk, T. A. v. 2006. Discourse, context and cognition. *Discourse Studies* 8:159–177.

Ferrari, P. F.; Rozzi, S.; and Fogassi, L. 2005. Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *Journal of Cognitive Neuroscience* 17(2):212–226.

Frey, S. H. 2008. Tool use, communicative gesture and cerebral asymmetries in the modern human brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363(1499):1951–1957.

Frith, C. D. 2007. The social brain? *Phil. Trans. R. Soc. B* 362:671–678.

Geary, D. C. 2004. *The Origin of Mind: Evolution of Brain, Cognition, and General Intelligence*. Washington, DC: American Psychological Association.

Geib, C. W., and Goldman, R. P. 2009. A probabilistic plan recognition algorithm based on plan tree grammars. *Artificial Intelligence* 173(11):1101–1132.

Gibson, J. J. 1977. The theory of affordances. In Shaw, R., and Bransford, J., eds., *Perceiving, Acting, and Knowing*. Lawrence Erlbaum. 67–82.

Grice, H. P. 1957. Meaning. *Philosophical Review* 66(3):377–388.

Gu, T.; Wu, Z.; Tao, X.; Pung, H. K.; and Lu, J. 2009. epSICAR: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition. In *IEEE International Conference on Pervasive Computing and Communications*. IEEE Computer Society. 1–9.

Gumperz, J. J. 1992. Contextualization and understanding. In Goodwin, C., and Duranti, A., eds., *Rethinking Context*. Cambridge: Cambridge Univ. Press. 229–252.

Hills, T. T. 2006. Animal foraging and the evolution of goal-directed cognition. *Cognitive Science* 30(1):3–41.

Johnson-Frey, S. H. 2003. What's so special about human tool use? *Neuron* 39:201–204.

Kim, E.; Helal, S.; and Cook, D. 2010. Human activity recognition and pattern discovery. *Pervasive Computing* 9:48–53.

Leonard, W. R.; Robertson, M. L.; Snodgrass, J. J.; and Kuzawa, C. W. 2003. Metabolic correlates of hominid brain evolution. *Comparative Biochemistry and Physiology - Part A* 136(1):5–15.

Pearl, J. 2009. *Causality: Models, Reasoning and Inference, 2nd edition*. Cambridge Univ. Press.

Pfeifer, R., and Bongard, J. 2006. *How the Body Shapes the Way We Think: A New View of Intelligence*. MIT Press.

Rehnmark, F.; Bluethmann, W.; Mehling, J.; Ambrose, R. O.; Diftler, M.; Chu, M.; and Necessary, R. 2005. Robonaut: The 'short list' of technology hurdles. *Computer* 38:28–37.

Reich, W. 2010. Basic communicative acts (BCAs): A strategy for implementing context-sensitive dialogue in social robots. In *Dialog with Robots: Papers from the AAAI Fall Symposium*, 169–170.

Reich, W. 2011. The cooperative nature of communicative acts. *Journal of Pragmatics* (in press).

Rizzolatti, G., and Sinigaglia, C. 2008. *Mirrors in the Brain: How Our Minds Share Actions and Emotions*. Oxford: Oxford Univ. Press.

Sahin, E.; Cakmak, M.; Dogar, M. R.; Ugur, E.; and Ucoluk, G. 2007. To afford or not to afford: A new formalization of affordances toward affordance-based robot control. *Adaptive Behavior* 15(4):447–472.

Shannon, C. E. 1963. The mathematical theory of communication. In Shannon, C. E., and Weaver, W., eds., *The Mathematical Theory of Communication*. Urbana/Chicago: Univ. of Illinois Press. 29–125.

Sperber, D., and Wilson, D. 1995. *Relevance: Communication and Cognition*. Oxford: Blackwell, 2 edition.

Tooby, J., and Cosmides, L. 2005. Conceptual foundations of evolutionary psychology. In Buss, D. M., ed., *The Handbook of Evolutionary Psychology*. Chichester: Wiley. 5–67.

Topal, J.; Gergely, G.; Erdohegyi, A.; Csibra, G.; and Miklosi, A. 2009. Differential sensitivity to human communication in dogs, wolves, and human infants. *Science* 325(5945):1269–1272.

Ugur, E., and Sahin, E. 2010. Traversability: A case study for learning and perceiving affordances in robots. *Adaptive Behavior* 18(3-4):258–284.

Wojek, C.; Nickel, K.; and Stiefelhagen, R. 2006. Activity recognition and room-level tracking in an office environment. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*. Heidelberg. 25–30.