From an artificial intelligence to multiple subjective intelligences: Approaching the second level of subjectivity

Timo Honkela

Aalto University School of Science and Technology
Department of Information and Computer Science
Adaptive Informatics Research Centre
Computational Cognitive Systems research group
timo.honkela@tkk.fi

Abstract

In artificial intelligence and cognitive science research, it has been commonplace to concentrate on understanding and modeling the intelligence of a single agent as a "representative" of all agents as if they were all the same or at least similar enough. On the other hand, the researchers on distributed artificial intelligence have already for a long time paid attention to the social level of knowledge. A single agent may have knowledge of a particular part of some domain or it may be able to solve only particular kinds of problems. These agents then collaborate to solve more complex problems that exceed their individual skills or knowledge. We can call this distribution of skills and knowledge as the *first level of subjectivity*. However, in the classical AI research this consideration of subjectivity can be deemed to be superficial in a particular manner. Namely, it is usually assumed that the epistemological framework in which the knowledge is represented is shared by the individuals. In other words, the symbolic items in the knowledge representations are assumed to have fixed meanings, shared by all the agents. For instance, if one agent knows that mother(mary, john) and father(john, peter) and another agent knows that grandmother(X, Y) :- mother(X, Z) and father(Z, Y) these agents can together conclude that grandmother(mary, peter). It is notable that family relationships are an example of unusually clear relationships and the underlying concepts are crisp. The research on fuzzy sets has, however, already shown that many real world concepts that underlie their linguistic descriptions have imprecise boundaries. There are multiple methodological alternatives to represent this impreciseness apart from fuzzy sets but the underlying issue is that meanings should be considered as distributions rather than points or nodes in some discrete networks of connected items. The relationships are important but they are not sufficient alone in defining meaning. Let us assume, just for the sake of the argument, that the human mind stores a huge number of membership function value distributions related to a large number of concepts. The fuzzy set theory has some times been criticized for the question how does one define the membership functions in a proper manner as there are so many potential ways of doing it. But, actually this seems to be exactly the empirical situation in each human mind! Namely, each individual has a subjective way to assess the impreciseness of each concept. What is considered red, long, beautiful, democratic or lovely by one individual may not be same for another. The exact interpretation of, for instance, the concept of house, mother or computation depends on the subject and these interpretations are not shared in a manner that could be defined in a simple, logical way. Therefore, one cannot assume that some concept could be represented as a symbol

that then somehow miraculously represents its meaning. One may assume that one uses some prototypical meaning of the symbols but that does not solve the underlying problem and does not facilitate the proper contextual use of these symbols in intelligent systems. The meaning of a symbol needs to be grounded as has been pointed out by some researchers already for some time (Harnad 1990, Cangelosi et al. 2000). But this grounding has to be done for each individual separately (and this is typically neglected). This phenomenon of subjective meaning functions we call the *second level of subjectivity*. Methodologically, when we represent the meaning as distributions in high-dimensional spaces, grounded at perceptional, action-based and cultural level, the interpretations of symbols can be mathematically and computationally compared (Raitio et al. 2004). Moreover, steps towards the direction of a theory of representation and communication that approves the empirical fact of the second level of subjectivity can be taken (Honkela et al. 2008). In essence, this approach is based on the idea that the agents adapt themselves towards shared symbolic and linguistic representations through imitation in shared contexts that facilitate (partial) agreement of the use of the symbols (Lindh-Knuutila et al. 2006, 2009; see also Steels and Kaplan 2000). In general, considering the second level of subjectivity has many practical implications for the future AI and information systems research and development including, for instance, the possibility of building inter-operable systems that are able to adapt their understanding in changing contexts and communicative situations. Methodologically, sophisticated mathematical and computational methods are needed that match with the complexity of the underlying dynamic phenomena (consider, e.g., Ritter and Kohonen 1989, von Foerster 2002 and Hyötyniemi 2010, including philosophical discussions).

References

A. Cangelosi, A. Grec & S. Harnad (2000) From robotic toil to symbolic theft: grounding transfer from entry-level to higher-level categories. Connection Science 12(2): 143-62.

H. von Foerster (2002) Understanding understanding. Springer-Verlag.

S. Harnad (1990) The Symbol Grounding Problem. Physica D 42: 335-346.

T. Honkela, V. Könönen, T. Lindh-Knuutila & M.-S. Paukkeri (2008) Simulating processes of concept formation and communication. Journal of Economic Methodology, 15(3):245-259.

H. Hyötyniemi (2010) SAMPO Mills: Neocybernetic Grounding of Ontogenesis. Unpublished manuscript.

T. Lindh-Knuutila, T. Honkela & K. Lagus (2006) Simulating meaning negotiation using observational language games. Symbol Grounding and Beyond, Springer, pp. 168-179.

T. Lindh-Knuutila, J. Raitio, and T. Honkela (2009) Combining self-organized and Bayesian models of concept formation. In Proceedings of NCPW11.

J. Raitio, R. Vigário, J. Särelä & T. Honkela (2004) Assessing similarity of emergent representations based on unsupervised learning. Proceedings of IJCNN 2004, vol. 1, pp. 597-602.

H. Ritter & T. Kohonen (1989) Self-organizing semantic maps. Biological Cybernetics, 61: 241-254.

L. Steels and F. Kaplan (2000) Aibo's first words: the social learning of language and meaning, Evol. Commun., 4: 3–32.