Representational Content and Agent-Environment Interaction

Tibor Bosse (tbosse@cs.vu.nl)

Vrije Universiteit Amsterdam, Department of Artificial Intelligence
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

Catholijn M. Jonker (jonker@cs.vu.nl)

Vrije Universiteit Amsterdam, Department of Artificial Intelligence
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

Jan Treur (treur@cs.vu.nl)

Universiteit Utrecht, Department of Philosophy
Heidelberglaan 8, 3584 CS Utrecht, The Netherlands

Classical approaches to representational content are based on correlations between an agent's internal state properties and external state properties; cf. (Kim, 1996), pp. 191-193. For example, the presence of a horse in the field is correlated to an internal state property that plays the role of a percept for this horse. However, for embodied agents that have an extensive interaction with their environment, this classical correlational approach does not suffice. In particular, an internal state in such an agent does not depend on just one state property of the external world, but is affected both by external aspects of the world and by internal aspects of the agent itself and the way in which these aspects are interwoven during the (ongoing) interaction process.

Given this problem, it is under debate among several authors whether adequate alternative notions of representational content exist for such an embodied agent's internal states. Some authors claim that for at least part of the internal states it makes no sense to consider them as conceptual or as having representational content; e.g., (Sun, 2000, 2002). Other authors claim that some notions of representational content can be defined, but these strongly deviate from the classical correlational approach; e.g. (Bickhard, 1993).

In (Bosse, Jonker & Treur, 2003), for a number of notions of representational content it was explored in a case study how they work out, and, especially, how they can be formalised. This case study involves the processes to unlock a front door that sticks. Between the moment that the door is reached and the moment that the door unlocks the following reciprocal interaction takes place:

- the agent puts rotating pressure on the key
- the door lock generates resistance in the interplay
- the agent notices the resistance and increases the rotating pressure
- the door increases the resistance
- and so on, without any result
- finally, after noticing the impasse the agent changes the strategy by at the same time pulling the door and turning the key, after which the door unlocks

This example shows different elements. The first part of the process is described in terms of (Sun, 2000, 2002)'s subconceptual level, whereas the last part of the process is viewed in terms of the conceptual level. For both parts of the process the notion of representational content has been discussed and formalised by identifying executable local dynamic properties for the basic dynamics. On the basis of these local properties a simulation model has been made. The specifications of the representational content of the internal state properties have been validated by automatically checking them on the traces generated by the simulation model. Moreover, by mathematical proof it was shown how these specifications are entailed by the basic local properties. This shows that the internal state properties indeed fulfil the representational content specification.

In work in progress, the approach described above is applied to the process of conditioning. As a case study, dynamic properties for the basic dynamics of the neural conditioning mechanism of the sea hare *Aplysia* have been identified, and a simulation model has been made. It is analysed whether the traces generated by this simulation model validate the formal specifications of the representational content of internal state properties created by conditioning.

References

Bickhard, M.H. (1993). *Representational Content in Humans and Machines*. Journal of Experimental and Theoretical Artificial Intelligence, 5, 1993.

Bosse, T., Jonker, C.M., & Treur, J. (2003). *Representational Content and the Reciprocal Interplay of Agent and Environment*. Proceedings of the IJCAI 2003 Workshop on Cognitive Modeling of Agents and Multi-Agent Interactions.

Kim, J. (1996). *Philosophy of Mind*. Westview Press.

Sun, R. (2000). *Symbol grounding: a new look at an old idea*. Philosophical Psychology, Vol.13, No.2, 2000.

Sun, R. (2002). *Duality of the Mind*. Lawrence Erlbaum Associates.