Looking at the Roots of Cooperation Through the Brain, And What We Find There

In most animal species cooperation is either limited to small groups or absent. All humans societies, in contrast, are based on cooperation with unrelated people (Hammerstein, 2003). Does an evolutionary-adaptationist perspective help us understand the roots of human compliance to norms of cooperation?

Given what we know about the neurobiology of human social-behaviour (e.g. Fehr&Camerer 2007), the thesis defended is twofold: At present, whereas an evolutionary-adaptationist perspective (see e.g. Barkow, Cosmides, & Tooby, 1992; Pinker, 1997; Gigerenzer & Selten, 2001) does not provide clear answers about cooperation, a neurocomputational perspective (Dayan, 1994; Churchland & Sejnowski, 1992) may enable us to make some steps towards a better understanding of the roots of cooperation.

Even though the two approaches are compatible with each other, if we don't have solid grounds to believe that a trait is an adaptation, it is wise to focus on the "how" questions of how certain bits of the brain work, instead of ultimate "why" questions. The answers to those "how" questions will help us gain insight into the *real* evolutionary roots of those circuits and the behaviour they support. In this case, an adaptationist perspective would be worthless.

The most solid case for an adapted neurocognitive mechanism is a module dedicated to reasoning about cooperation for mutual benefit (Cosmides & Tooby, 2005). The existence of this mechanism at the neurobiological level seems to be supported by one neurological study (Stone et al 2002). After having considered this study in some detail, I argue that it does not provide sufficient grounds for a cheater-detector neurocognitive adaptation. On the one hand, the evidence indicates that the patient in question had a general, not domain-specific, difficulty in representing the social domain; on the other, his performance can be explained without appealing to adaptations. In particular, I show that reinforcement learning (RL) ideas (Sutton & Barto, 1998) from computational neuroscience can be used to give a plausible explanation of the deficit of that patient.

The RL framework explains how brains have found a way to be efficient systems. An efficient system makes reliable predictions about the amount of reward that the outcome of a certain action will yield, and thereby pursues the action with the highest return. Having food, drinking water and having sex are primary rewards since they are fitness-enhancing. I argue that the efficiency of our brains is maintained by computational systems, which enable cooperative behaviour, but whose computational role is valuation. Although RL doesn't explain why biologically-arbitrary actions such as cooperation are valued more than others, it does lend support to the view that human social behaviour is largely determined by *culture* – a system of representations transformed by neurocognitive systems which are likely to have evolved to facilitate efficient neurocomputation.

An objection to this thesis is that the neurocomputational mechanisms underlying valuation are common to most animal species. Why then only humans society is based on cooperation?

I conclude the paper with a tentative answer by relying on the notion of script and rule-like representatioin (Bicchieri, 2006, ch.2).

References

Barkow, J., Cosmides, L. & Tooby, J. (Eds.) (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Bicchieri, C. (2006). *The Grammar of Society: the Nature and Dynamics of Social Norms*, Cambridge University Press.

Churchland, P. S., & Sejnowski, T. J. (1992). *The Computational brain*. Cambridge, MA: MIT Press.

Cosmides, L. & Tooby, J. (2005). "Neurocognitive adaptations designed for social exchange." In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology* (pp. 584-627). Hoboken, NJ: Wiley.

Dayan, P. (1994). "Computational modelling". *Current Opinion in Neurobiology*, 4:212-217.

Fehr, E., & Camerer, C. (2007). "Social Neuroeconomics – The Neural Circuitry of Social Preferences." *Trends in Cognitive Sciences 11*, 419-427.

Gigerenzer, G. & Selten, R. (Eds.). (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.

Hammerstein, P. (2003). *The Genetic and Cultural Evolution of Cooperation*. Cambridge, MA: MIT Press.

Pinker, S. (1997). *How the Mind Works*. New York: Norton.

Stone, V., Cosmides, L., Tooby, J., Kroll, N. & Knight, R. (2002). "Selective Impairment of Reasoning About Social Exchange in a Patient with Bilateral Limbic System Damage. *Proceedings of the National Academy of Sciences*, 99(17), 11531-11536.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.