

What are System 1 processes like? Defeasible but logical perhaps?

KEITH STENNING

*Human Communication Research Centre
University of Edinburgh*

1) Background

Evans' dual process account of human reasoning grew out of the subfield known as the 'psychology of deduction' which is strongly influenced by a particular conception of the relation between logic and human reasoning. The resulting conception of dual processes identifies System 2 processes as 'logical' thinking and System 1 processes as non-logical (associative, pragmatic, . . .). Psychologically System 1 processes are fast, cheap, automatic, inaccessible, not-isolable from general knowledge, and primitive; whereas System 2 processes are slow, expensive of cognitive effort, requiring deliberate control, are isolated from knowledge, available to conscious reflection, and evolutionarily advanced, possibly uniquely human.

On many grounds, the conception of the relation between psychology and logic embodied in the 'psychology of deduction' is anachronistic. When a more adequate modern conception of the multiplicity of logics and their association with different reasoning tasks is adopted, the distinction of the two systems by the criterion of what is 'logical' is called into question. This paper builds on a program of research which assumes that the task of interpreting the experimental materials of 'the psychology of deduction' into suitable logics, is a substantial part of the problem facing subjects in these experiments, and that the logical interpretations which result are the chief determinants of subjects reasoning. This approach throws new empirical and theoretical light on the main experimental paradigms of the field (the selection task [Stenning and van Lambalgen(2004)], the suppression task [Stenning and van Lambalgen(2005)], syllogistic reasoning [Stenning and Cox(2006)]), in particular explaining what have been claimed to be 'content' effects as logical interpretation effects, and opening the way to explanations of the generality of human reasoning [Stenning and van Lambalgen(in press)]. An important range of logics required are defeasible nonmonotonic logics, and evidence accumulates that for most subjects these are likely to be a first interpretational resort. [van Lambalgen and Hamm(2004)] have shown how these same logics can be applied to a range of problems in temporal discourse semantics.

Here we will argue that once it is acknowledged that nonmonotonic logics are at the centre of human discourse interpretation, explorations of their tractability become highly relevant to the basis of the distinction between dual processes.

2) Dual logical systems.

For modelling the credulous process of interpretation, [Stenning and van Lambalgen(2005)] adopt a family of fast and efficient default logic based on logic programming, which formalise closed world reasoning. These logics are highly tractable, allowing the computation of unique 'intended models' of sets of assumptions and databases of 'general knowledge' conditionals in linear time. They show that this logic, given a suitable three-valued semantics, is susceptible to neural network implementations in spreading activation networks. The classical logic assumed to constitute logic by most of the psychological subfield (but see

[Evans(2003)] for an at least partial exception) is well known to be resistant to any such efficient treatment.

Unusually for a neural network treatment, there is a plausible outline available of the processes which create the working memory/long term memory networks which represent discourse interpretations. This outline draws on neuroscience studies of the different timescales of plasticity required [von der Malsburg(1981)]. Interestingly, it is the three-valued semantics necessary for the treatment of negation which points to the need for the process of retinotopic mapping which network construction involves [Stenning and van Lambalgen(2005)].

So our proposal is that it is possible to take seriously observations of dual processes with their gross psychological characteristics, but to see logics (of different kinds) underlying both Systems 1 and 2. The remainder of the paper is about the advantages of making this adjustment.

While the external characteristics of System 1 processes listed above have received much attention, their substance has remained remarkably little specified. Thinking of these processes as ‘associative’ raises fundamental problems. Much of natural language processing as studied in psycholinguistics has all the external properties of System 1 (fast, automatic, etc.), yet demolition of associative approaches to natural language processing was the first requirement and achievement of the cognitive revolution. Credulous language processing *is* System 1 but *is not* associative.

Basing human discourse interpretation on defeasible logics might be seen as a move from the frying pan into the fire: nonmonotonic logics have a bad reputation for intractability. However, logic programming is an exception, being both rather expressive and yet highly tractable, as befits its technological origins. Once a three valued semantics allows a more general treatment of negation, this family of logics provides a natural framework for natural language discourse processing.

So System 1 processes must be logical (rather than associative) and could be logical (in some defeasible logic). We will argue that these observations can greatly enrich our conceptions of dual process theories, and point us at insightful questions.

The paper concludes with a discussion of implications for the relation between System 1 and Systems 2 in the development of communication and learning, and in evolutionary accounts of how the modern human cognitive architecture arose. Defeasible logics related to that proposed for System 1, are known as ‘planning logics’ and underlie low-level motor planning in some approaches to robotics [Shanahan(1997)]. This accords with an idea proposed by several strands of research that the continuity of human language with ancestral cognition is to be sought in planning rather than communication capacities [Greenfield(1991)] [Arbib and Rizzolatti(1997)]. On this view recursion, far from being the critical innovation of human language evolution [Hauser et al.(2002)Hauser, Chomsky, and Fitch], is already a feature of our ancestors’ cognition, but its employment in social interaction is the innovation that led to language.

Our ancestors’ working memories were capable of holding their model of their current situation and goals, as supported by their long term representations of environmental regularities and their current perceptual input. However, they had quite severe limitations in the ‘semantic distance’ possible between this model and the current situation. System 1 processes continually defeasibly updated this model as information arrived or goals changed. In humans, the possible semantic distance

between working memory model and current situation is hugely extended, and this capacity has to be supported by sophisticated interpretational machinery which can accommodate the fact that the interpretation of working memory often has highly indirect implications for action. We propose that System 2 processes arose by the development and extension of interpretative supervisory systems driven by the need to repair breakdowns in interpretation arising in System 1. Such breakdowns inevitably occur both in individual cognition and an social communication. The balance of forces between these two modes is of course one of the major unsolved problems of human cognitive evolution.

References

- [Arbib and Rizzolatti(1997)] M. A. Arbib and G. Rizzolatti. Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition*, 29:393–424., 1997.
- [Evans(2003)] J.St.B.T. Evans. In two minds: dual-process accounts of reasoning. *TRENDS in Cognitive Sciences*, 7(10):454–459, 2003.
- [Greenfield(1991)] P.M. Greenfield. Language, tools and the brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and brain sciences*, 14:531–595, 1991.
- [Hauser et al.(2002)Hauser, Chomsky, and Fitch] M. Hauser, n. Chomsky, and T. Fitch. The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298:1569–79, 2002.
- [Shanahan(1997)] M.P. Shanahan. *Solving the frame problem*. MIT Press, Cambridge MA, 1997.
- [Stenning and Cox(2006)] K. Stenning and R. Cox. Rethinking deductive tasks: relating interpretation and reasoning through individual differences. *Quarterly Journal of Experimental Psychology*, 59:??–??, 2006.<http://www.hcrc.ed.ac.uk/~keith/InterpretationandReasoning/stenningandcox.pdf>.
- [Stenning and van Lambalgen(2004)] K. Stenning and M. van Lambalgen. A little logic goes a long way: basing experiment on semantic theory in the cognitive science of conditional reasoning. *Cognitive Science*, 28(4):481–530, 2004.
- [Stenning and van Lambalgen(2005)] K. Stenning and M. van Lambalgen. Semantic interpretation as reasoning in nonmonotonic logic: the real meaning of the suppression task. *Cognitive Science*, 29(6):919–960, 2005.
- [Stenning and van Lambalgen(in press)] K. Stenning and M. van Lambalgen. Explaining the domain generality of human cognition. In M. J. Roberts, editor, *Integrating the Mind*. Psychology Press, in press.
- [van Lambalgen and Hamm(2004)] M. van Lambalgen and F. Hamm. *The proper treatment of events*. Blackwell, Oxford and Boston, 2004.
- [von der Malsburg(1981)] C. von der Malsburg. The correlation theory of brain function. Internal Report 81-2, Dept. of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, Goettingen, Germany, 1981. Reprinted in: E. Domany, J.L. van Hemmen and K. Schulten (eds.) *Models of neural networks II*, Springer Verlag, 1994.