

Do We Know What the User Knows, and Does It Matter? The Epistemics of User Modelling

Michael Ramscar^{1,3}, Helen Pain¹, and John Lee^{2,3}

- ¹ Department of Artificial Intelligence, University of Edinburgh, Scotland
² Human Communication Research Centre, University of Edinburgh, Scotland
³ Department of Architecture, University of Edinburgh, Scotland

Abstract. Whilst many user models can function perfectly adequately with a behavioural impression of the user, the provision of assistance in some task domains, notably design, requires a richer understanding, incorporating information about the user's knowledge and beliefs. This raises a number of important and difficult questions: How can we know what the user knows, and how can we know that we know? We present evidence that the psychological view of human conceptual knowledge that underpins typical approaches to these questions is flawed. We argue that user knowledge can be modelled, up to a point, but that to ask whether or not we can know what the user knows is to misunderstand the question.

1 What Do We Want to Know?

Many user models can function perfectly adequately with only a behavioural impression of the user: A user's actions can be sufficient input for a system to adapt in order to accommodate the particular needs of a particular user or user type. Even if we include a user's linguistic interaction with a system—in the form of text input—under the heading of “behaviour”, then providing that a model can anticipate the form and content of this behaviour, and have some notion of correct versus incorrect manifestations of the behaviour, still there are strategies for modelling such user behaviour in a generalised fashion. Good examples are “buggy” models, such as those developed by Brown and Burton (1978). In some domains (and perhaps in all domains at some times), it is impossible to anticipate the form and content of a user's behaviour. This is especially true of design, where the unpredictability of a designer's behaviour (in the widest possible sense) is perhaps an essential aspect of the activity. Modelling a user's conceptual knowledge raises important questions: How can we know what the user knows; how can we know that we know?

The basic problem facing models of users' conceptual knowledge can be best stated as one of relating the conceptual knowledge of a computer user to corresponding knowledge modelled within the computer (to this extent, modelling a user's concepts deviates little from other user modelling approaches). Since a user's knowledge must be expressed in some kind of external representation for communication with the computer, this is equivalent to the problem of reconciling multiple representations of concepts or objects. Two users might represent an event in their own different ways; the designer of a computer system might represent it a third way.

The weakness of the conceptual models in current techniques in ‘knowledge modelling’ (ad hoc “generalised” feature decompositions, Ramscar et al., 1996) is unsurprising, given the absence

of any convincing psychological model of conceptual categorisation. Development of a psychological model of conceptual categorisation has been fatally hampered by two important, erroneous assumptions: Firstly, categories have often been treated as a rigid, externally imposed phenomenon; secondly, researchers have concentrated upon category representation rather than on the process by which categorisation judgements are made.

This concentration upon category representation is surprising, given Rosch's (1978) theory of prototypical category representations—the basis for much of this work—which argues that: (a) prototypes are best seen metaphorically; what are really referred to are judgements of degrees of prototypicality; and (b) prototypes do not constitute a theory of representation of categories.

We have proposed (Ramscar et al., 1996) an approach to categorisation the focus of which is upon the categorical judgement process rather than on the representation of categories (or category prototypes); focussing on how representations are classified together, rather than specifying the representational form of categories. Research from the study of analogy (which concentrates upon modelling the process whereby two representations are considered analogous) is taken as a starting point for investigating categorical judgements. In doing this we have discarded the distinction between category membership and analogy.

A study by Ramscar and Pain (1996) has shown that where judgements are considered, this distinction does not hold at a cognitive level of description. The study involved giving subjects materials previously used in investigations into analogical similarity (Gentner et al., 1993) and asking them to perform categorisation tasks with them. Subjects' categorisation and analogical judgements utilised the same salient aspects of the representations within the materials, and their judgements of analogical and categorical similarity had a direct congruence, which contradicted previous distinctions between cognitive processes of analogy and categorisation. The removal of this distinction highlighted the parallels to be drawn between models of analogy and categorisation. It is becoming more widely accepted that structure plays a major role in category formation (Goldstone, 1994): Analogical reasoning research addresses a process which reasons amongst structural networks. Forbus et al. (1995) propose the following model of analogical reasoning: (a) Initial selection is dependent upon surface similarity. (b) Analogical similarity is determined by deeper structures.

This is strikingly similar to Medin and Ortony's (1989) proposed knowledge representation scheme for categorisation: (a) The identification procedure is based upon surface features. (b) Classification is determined by deeper structures.

Where research into analogy differs from research into categorisation is in the richness of its process models. A number of detailed, plausible, computationally implementable models of the analogical process exist; the same cannot be said of categorisation. Research in analogy has been far more successful because it has focused upon the analogical process, and the interplay between this process and the representations of analogues (rather than simply determining conditions by which representations might be considered to be analogous). We are now using Gentner's analogy theory to develop an implementable model of the categorisation process.

2 Does It Matter That We May Not Know What We Want to Know?

We have outlined the state of psychological research into categorisation ("conceptualisation") and suggested a characterisation of human conceptual judgement that fits the evidence avail-

able. Moreover, like Gentner's theory of analogy, it is amenable to computational implementation. Our model makes no distinction between analogical and "literal" conceptual judgements at the cognitive level: Understanding this is essential to understanding human conceptual judgements. Following Wittgenstein (1953), we argue that it is conceptual judgements ("use"), rather than definitive representations which provide the key. Wittgenstein, having demolished the idea of necessary and sufficient conditions for category membership, offers up compelling arguments to believe that "categories" exist insofar as they are used by people, and that they are bounded to the extent to which individuals in a community can agree to boundaries. What is important is not individual representation of "categories", but the alignment, in terms of judgement agreements, between individuals' category representations. It is this alignment between representations that our model aims to capture. According to our model, there is no definitive answer as to whether a given representation is an instance of this or that category. Rather, one must look for an empirical answer: agreements between individuals in a linguistic (conceptual) community. We can show that much of the basis for this agreement stems from structural alignments between individuals' conceptual representations. Our hypothesis is that system representations need not be definitive but rather must function pragmatically. In so far as these representations sufficiently approximate the kind of structure and content representations used by individuals in a given community, they will be able to map and model those human expressions of conceptual knowledge. Such modelling wouldn't enable a system to know what the user knows, nor could it enable it to know that it knew. It might, however, enable it to judge that if some external representations of the user aligned structurally with some of its stored representations, then it could pragmatically attribute certain conceptual knowledge to a user, much as humans do (Wittgenstein, 1953). Having done so, it might then be in a position to respond or adapt to that user knowledge in some useful way.

References

- Brown, J. S., and Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science* 2:155–192
- Forbus K., Gentner, D., and Law, K. (1995). MAC/FAC: A model of similarity based retrieval. *Cognitive Science*, 19:2:141–205
- Gentner, D., Ratterman, M., and Forbus, K. (1993). The roles of similarity in transfer. *Cognitive Psychology* 25: 524–575
- Goldstone, R. L. (1994). The role of similarity in categorization. *Cognition* 52:125–157.
- Medin, D., and Ortony, A. (1989). What is psychological essentialism? In Vosniardou, S., and Ortony, A., eds., *Similarity and Analogical Reasoning*. Cambridge University Press.
- Rosch, E. (1978). *Cognition and Categorisation*. Erlbaum.
- Ramscar, M., and Pain, H. (1996). Can a real distinction be made between cognitive theories of analogy and categorisation? *Proceedings of the 18th Conference of the Cognitive Science Society*. Erlbaum.
- Ramscar, M., Lee, J., and Pain, H. (1996). A cognitively based approach to computer integration for design systems. *Design Studies* 17:4:465–483.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell. Translated by Anscombe, G. E. M.