

AI article

Description

Extract from Coyne, R., *Designing Information Technology in the Postmodern Age: From Method to Metaphor*, Cambridge, Mass.: MIT Press, 1995. Chapter 1.

The Theoretical Orientation to Computer Systems Design

The theorist of AI, Roger Schank (1946-), identifies five elements of human intelligence.^[1] According to Schank, the first is the ability to communicate.^[2] For artificial intelligence, communication is commonly understood in terms of the passage of information from one agent (or human being) to another. The second constituent of intelligence is "internal knowledge."

We expect intelligent entities to have some knowledge about themselves. They should know when they need something; they should know what they think about something; and, they should know that they know it.^[3]

Third, is "world knowledge." Intelligence involves knowing about the "outside world" and being able to retrieve information about it. Memory is the capacity to encode past experience and use it to guide us through new experiences. The fourth constituent of intelligence is the ability to formulate goals and plans.

Goal-driven behaviour means knowing when one wants something and knowing a plan to get what one wants.^[4]

The fifth constituent of intelligence is creativity. This is the ability to formulate new plans, look at things in a new way, adapt to changes in the environment, and learn from experience.

Schank's description of intelligence borrows of the tradition of thinking known as Rationalism, broadly attributed to René Descartes (1596-1650), Gottfried Leibniz (1646-1716), and Baruch Spinoza (1632-1677). Schank's description of intelligence highlights the rationalistic assumption that there is an inside and an outside to human cognitive experience (ie of thinking). There is an inside world of knowledge that involves a self knowing about itself, and there is an outside world about which that self can also know. Sometimes the former is called "subjective knowledge," the latter "objective knowledge." The inside world is that of the subject, the outside is that of the object. For Schank these distinctions are unproblematic. In artificial intelligence research as presented by Schank the world of the subject and the object are not seen merely as metaphors. Nor is their identification dependent on some particular context of discussion. There is no agonising over the "constitution of the subject" prevalent in *critical theory* for example.^[5] In assuming the certainty of subject and object, communication is largely a matter of the passage of information from one subject to another through the medium of the "external world." Communication gives us access to each other's subjectivities.

Rationalism also assumes the priority of purposeful behaviour in human affairs. The subject (the thinking self) identifies problems and goals, much in the manner of Descartes's method of reason, then sets about achieving those goals. According to Schank, intelligent action is purpose driven. We are constantly forming and reforming goals in response to problem situations, and then carrying out plans to accomplish those goals. We have *ends*, which are our goals, and we develop *means* to achieve these goals. The means are plans.

According to Schank the ability to engage in purposeful activity involves *knowledge*. According to his presentation of artificial intelligence, knowledge can be made explicit as procedures, rules, frames or semantic networks.^[6] The "internal" knowledge representations are often described as "cognitive models."^[7] Habitual, unreflective activities, such as being engrossed in a hobby or moving about one's office, or the practice of skills, do not appear to involve goals or plans explicitly. However, Schank suggests that in such habitual activity goals, plans and knowledge are "internalised," "compiled,"^[8] or handled by a kind of "default reasoning."

Behaviour not readily accommodated by the idea of representable knowledge is also addressed by the term "creativity." Creativity is a major preoccupation of artificial intelligence. Creativity is evident where we are not merely mapping goals and plans to situations through readily articulated knowledge. Artificial intelligence theorists recognise that "creative knowledge" is extremely difficult to capture.

Researchers outside the area of automated intelligence are also interested in models similar to those proposed by Schank. According to some researchers general computer systems design should involve a consideration of cognitive models. A computer system embodies a representation of the goals and plans of the user. In turn, the user of the system has a model of the computer, and also of the situation domain they are working in. These researchers consider that, irrespective of whether the system is to be "intelligent," good computer systems design takes account of cognitive models.^[9]

As I suggest above, Schank's characterisation of cognition embodies the rationalism of Descartes and Leibniz. His view of cognition includes the self-evident notion of the thinking subject (*res cogitans*) set apart from an independent and measurable spatial object world (*res extensa*). The essential aspects of thought can be described or reproduced as formulas, production rules and axioms in predicate calculus, able to be processed through context-independent and unprejudiced reason. Reason can be removed from culturally situated human agency and run through computers as programs.^[10] The assumptions Schank makes about human reasoning are evident in a tradition of well articulated researches dating back to Alan Turing (1912-1954) and Alonzo Church (1903-1995), including the work of Donald Michie (1923-2007), Marvin Minsky (1927-), Herbert Simon (1916-2001), and Daniel Dennett (1942-). They owe much to the notions of formal reasoning developed by Gottlob Frege (1848-1925) and Bertrand Russell (1872-1970), and many of their theories are developed (and critiqued) within the tradition of Analytic Philosophy with its emphasis on language, truth and logic.^[11]

Some criticisms of AI are presented by researchers Terry Winograd and Fernando Flores.^[12] The critics point to the difficulties in identifying what is subjective knowledge and what is objective knowledge independently of a particular situation. We are always involved in a context, even when trying to be "objective." This has led some philosophers to assert that the basic Cartesian (ie pertaining to Descartes) premise of identifying a subject opposed to an object world is fundamentally flawed. The articulation of human knowledge in a representational formalism is also problematic. There is never a basic, underlying formalisation that accounts for everything we know in a particular situation. Such formalisations are themselves artefacts, inventions of the moment to fit a particular purpose.

Finally, the quest for goals appears elusive. We readily construct our intentions after the event. We know about planning as a tool of social mobilisation, and we live and work in a culture that uses blueprints, methods and procedures. We are so used to talking about goals that we readily identify plan formulation as a basic cognitive function. According to the critics, rationalism largely ignores the social and prejudicial nature of its own understandings. Winograd and Flores appeal to studies in language and hermeneutics to develop these and other criticisms.

These criticisms of rationalism find support from philosophers such as Friedrich Nietzsche (1844-1900), Martin Heidegger (1889-1976) and the phenomenologists, who have pointed to the inadequacy of the Cartesian focus on the *cogito* (the thinking self) rather than *Being*.^[13] The elevation of *ontology* (the study of being) over *epistemology* (the study of knowledge) leads to a different line of philosophical inquiry than that offered by the tradition of rationalism. The assumptions of Schank and others regarding internal and external knowledge, and the importance of representations, goals and plans are criticised by writers such as Winograd, Flores and Hubert Dreyfus from the perspective of phenomenology and hermeneutics.^[14] Also, the school of thought commonly known as *critical theory* presents criticisms of rationalism from the point of view of social justice. In appropriating aspects of Heidegger, along with Karl Marx (1818-1883) and Sigmund Freud (1856-1939), proponents of *critical theory* commonly focus on the unquestioning acceptance and appropriation of *reason* in Cartesian theories.^[15] According to *critical theory*, rationalism pays no regard to, yet surreptitiously promotes, entrenched power structures and regimes of injustice. Joseph Weizenbaum (1923-2008), a computer scientist and former artificial intelligence researcher, in railing against the "imperialism of instrumental reason," has been influential in bringing the critical perspective to the attention of the community of computer systems designers.^[16]

[1] See Schank, R. (1990) What is AI anyway? in *The Foundations of Artificial Intelligence: A Source Book*, D. Partridge and Y. Wilks (eds), Cambridge University Press, Cambridge, pp.3-13.

[2] Ibid., p.4.

[3] Ibid., p.5.

[4] Ibid., p.5.

[5] See for example Zavarzadeh, M, ud and Morton, D. (1986-87) Theory pedagogy politics: the crisis of the subject in the humanities, *Boundary 2*, 15, pp.1-22.

[6] See Barr, A. and Feigenbaum, E.A. (eds) (1981) *The Handbook of Artificial Intelligence* Volume 1, William Kaufmann, Los Angeles, California.

[7] For a discussion of the uses of the term "model" in artificial intelligence and cognitive science see Wilks, Y. (1990) One small head: models and theories, in *The Foundations of Artificial Intelligence: A Source Book*, D. Partridge and Y. Wilks (eds), Cambridge University Press, Cambridge, pp.121-134;

and Snodgrass, A.B. and Coyne, R.D. (1992) Models, metaphors and the hermeneutics of designing, *Design Issues*, Vol.9, No.1, pp.56-74. For a discussion of models in science see Hesse, M. (1966) *Models And Analogies in Science*, University of Notre Dame Press; and Turbayne, C.M. (1970) *The Myth of Metaphor*, University of South Carolina Press, Columbia, South Carolina.

[8] See Simon, H. (1969) *The Sciences of the Artificial*, MIT Press, Cambridge, Massachusetts.

[9] See Gardiner, M.M. and Christie, B. (1987) *Applying Cognitive Psychology to User-Interface Design*, Wiley, Chichester; Norman, D.A. (1988) *The Psychology of Everyday Things*, Basic Books, New York; Eberts, R.E. and Eberts, C.G. (1989) Four approaches to human computer interaction, *Intelligent Interfaces: Theory, Research and Design*, P.A Hancock and M.H. Chignell (eds), Elsevier, Amsterdam, pp.69-127; Laurel, B. (ed.) (1990) *The Art of Human-Computer Interface Design*, Addison Wesley, Reading Massachusetts; and Falzon, P. (ed.) (1990) *Cognitive Ergonomics: Understanding, Learning and Designing Human-Computer Interaction*, Academic Press, London.

[10] See Schank, R. C. *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*, Cambridge, Cambridge University Press, 1982; and Schank, R.C. and Abelson, R. (1977) *Scripts, Plans, Goals and Understanding*, Lawrence Erlbaum, Hillside, New jersey.

[11] Strictly speaking, analytical philosophers such as Ayer would maintain that the question 'can machines think?' is undecidable and therefore meaningless. But such philosophers demonstrate their allegiance to the calculative (and Cartesian) view of reason with comments such as the following.

'the questions with which philosophy is concerned are purely logical questions; and although people do in fact dispute about logical questions, such disputes are always unwarranted. we may be sure that one party in the dispute has been guilty of a miscalculation which a sufficiently close scrutiny of reason will enable us to detect.' (Ayer, A.J. (1990) *Language, Truth and Logic*, Penguin, London [first published in 1936], p.144)

See also Ayer, A.J. (1976) *The Central Questions of Philosophy*, Pelican, Harmondsworth, Middlesex. For a brief summary of the history of artificial intelligence see Dietrich, E. (1990) Programs in the search for intelligent machines: the mistaken foundations of AI, in *The Foundations of Artificial Intelligence: A Source Book*, D. Partridge and Y. Wilks (eds), Cambridge University Press, Cambridge, pp.223-233.

[12] See Winograd, T. and Flores, F. (1986) *Understanding Computers and Cognition: A New Foundation for Design*, Addison Wesley, Reading Massachusetts; and Winograd, T. (1990) Thinking machines: can there be? are we?, in *The Foundations of Artificial Intelligence: A Source Book*, D. Partridge and Y. Wilks (eds), Cambridge University Press, Cambridge, pp.167-189.

[13] See Nietzsche, F. (1966) *Beyond Good and Evil: Prelude to a Philosophy of the Future*, Trans. W. Kaufmann, Vintage Books, New York, originally published in 1886; and Heidegger, M. (1962) *Being and Time*, trans. J. Macquarrie and E. Robinson, Basil Blackwell, Oxford, originally published in German in 1927. For a summary of Heidegger's complex thought see Dreyfus, H. (1991) *Being-in-the-world: A Commentary on Heidegger's Being and Time, Division I*, MIT Press, Cambridge, Massachusetts.

[14] See Dreyfus, H.L. (1972) *What Computers Can't Do: The Limits of Artificial Intelligence*, Harper and Row, New York; Dreyfus, H.L. and Dreyfus, S.E. (1985) *Mind Over Machine*, MacMillan/The Free Press, New York; and Winograd and Flores, *Understanding Computers and Cognition*, op. cit.

[15] See Marcuse, H., (1991) *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*, London: Routledge.

[16] See Weizenbaum, J. (1984) *Computer Power and Human Reason: From Judgement to Calculation*, Penguin, Harmondsworth, Middlesex.

Date Created

September 18, 2011

Author

rcoyne99

default watermark