REFLECTIONS ON THE KNOWLEDGE LEVEL

Allen Newell

School of Computer Science
Carnegie Mellon University, Pittsburgh, PA

30 April 1992

Reflections on the Knowledge Level

The Knowledge Level (Newell, 1982) describes the nature of *knowledge* as the medium of a systems level that lies above the *symbol* or *program level* in the hierarchy of levels of computational systems. The unique character of the knowledge level is that its central law of behavior, the principle that the system will taken whatever actions attain its *goals* given its knowledge, can be computed without knowing any internal structure of the system. Put another way, the knowledge level abstracts completely from representation, structure and process. That this is possible — that a system's behavior can be predicted based only on the *content* of its representations plus its knowledge of its goals — is the essence of being an *intelligent* system. The knowledge level finds it most common use in designing systems, where the designer specifies a component by positing its function (i.e., its goal) and the knowledge available to it.

Systems can be described at many levels. A system described at the knowledge level, can also be described at the symbol level, which is to say in terms of representations, data structures and processes. The symbol-level description shows how the knowledge-level behavior is attained, just as the register-transfer description shows how the symbol-level system is realized by architectural mechanisms, and so on down to descriptions in terms of physical concepts.

The paper claims this concept of knowledge and knowledge-level system is used in practice by the AI and computer science community, hence understood implicitly. The paper itself is just making explicit to the community what it already knows in practice.

The Genesis of the The Knowledge-Level

The Knowledge Level officially came about as a presidential address to the AAAI at their first scientific conference, which was held at Stanford University in the summer of 1979. The paper, published simultaneously in both *Artificial Intelligence* and the *AI Magazine*, is essentially a written version of the talk, and I always think of the two of them together.

The occasion of a presidential address, especially an initial one, has the potential to be precedent setting in a variety of ways and I gave considerable thought to what I wanted to do. Chief among these was to set the course for presidential addresses to be about our science and not about our society or the state of applications or any of the other myriad things that presidents of societies find interesting to talk about.

Giving a contentful talk on a ceremonial occasion would seem to add a strong constraint. But I had already developed a strategy of using ceremonial talks to present general views of the science, where the material was contentful but not highly technical. Herb Simon and I had employed this strategy in our Turing Award lecture (Newell & Simon, 1976), where we talked about symbols and search as the two great pillars of artificial intelligence. I had used the strategy in my talk on physical symbol systems, at the inaugural meeting of the Cognitive Science Society at La Jolla (Newell, 1980). Talking about the knowledge level fit this strategy exactly, it seemed to me.

This still might appear strange — the knowledge level was a new technical concept, which needed to be introduced to the scientific community in technical terms before it could be summarized or discussed in high-level terms. But that wasn't the way I saw it. Knowledge was, like two or three other fundamental concepts in computer science and AI, a concept that every one had, and had largely correctly, but a concept that had not been discovered or invented by any specific scientist — certainly not myself. Two other important examples are symbol systems and architectures. What I was doing in talking about the knowledge level, I believed, was simply giving voice to a scientific concept that we all had. Whether the knowledge level was exactly the right way to cast our common understanding of knowledge was perhaps a little problematic, but not the main understanding of the nature of knowledge.

There was a final ingredient in my orientation to the talk, and hence the paper. Although I believed that all of us in AI and CS had an essentially correct concept of knowledge, I was acutely aware that knowledge had not been conceptualized explicitly in the literature. The concept was used entirely implicitly in our practice. Indeed, the explicit use of the term *knowledge* still resided with the philosophers — who, from my perspective, didn't understand at all what knowledge was. In accordance with my own belief that I was not giving a scientific talk, only talking about scientific matters, I was not about to provide in my talk (and paper) a fully technical treatment of the knowledge level. So I viewed the talk (and paper) as launching a not-yet-robust explicit formulation, with the hope that others would pick it up, find out how to use it in technical ways, and thus begin the evolutionary process of creating a useful technical concept.

These are the four main ingredients that existed for me at the the time of creation of the knowledge-level talk — the honor of the first AAAI presidential address, the attempt to establish its culture to be scientific, the presentation of another noninvented but fundamental scientific concept, and the attempt to bring to the surface the concept of knowledge and launch its life as a technical notion. All of these ingredients were quite explicit, not only in my thinking, but in the talk, and some of them even survived into the published paper. There was, by the way, no prior work by me on the subject of knowledge, before being invited to be president and facing the challenge of what to do for the initial presidential address. After all, I had understood the concept since the late 1950s (Newell, 1962) and had felt no need to put its study on my scientific agenda.

What has Happened?

On the style of the AAAI presidential address, the results have been decidedly mixed. The presidential address has certainly not developed a tradition of dealing with science. There have been some such presidential addresses, but also many that dealt with more general matters — though no less presidential for that. Of course, it was naive of me to believe I could influence the successive leaders in AI, who, of all people, are independent agents.

On the other hand, I was quite successful in the strategy of using ceremonial talks to convey general scientific results without benefit of the hard research that should be there to back them up. *The Knowledge Level* has joined *Symbols and Search* and *Physical Symbol Systems* as some of my more well-known papers. However, some people have told me that *The Knowledge Level* is rather hard to understand, so these papers may be more technical than I realize (or just badly written — always a real possibility).

On the third hand, no one has taken seriously — or even been intrigued with — the proposition that the knowledge level was not invented or discovered by me (or anyone else). As well, this has not happened for Herb Simon and myself on the physical symbol system. In both cases, we are cited according to standard practice as the progenitors of these ideas.

Let me be clear, once again, about what is going on (although with no hope of communicating to anyone). The standard realist conception of science has it that nature is the way it is. Scientists, in attempting to understand nature, engage in deliberate activities of observation, measurement, and theorizing. They discover facts and regularities about nature and their theories sometimes successfully predict these fact and regularities. Science gives credit (and sometimes glory) to the scientists who make these discoveries. The social structure of science is such that the only way additions to the body of science get made is through the agency of human scientists.

But with computer science and AI, an alternate path has emerged. These fields build artifacts — computers, architectures, languages, software systems — which can embody scientific truths about the domain of information processing. The creators of these systems need not be aware of all that that they put into these artifacts. They can construct an artifact under one mental model that suffices to get the artifact built. But this same artifact can embody additional truths, which can be learned by the people that use it. Just as in mathematics, in computing there are many implications of a computing system that can emerge with investigation and use. It is obvious, for instance, that the engineers who designed and built the Rand Johnniac and the IBM704 did not envision list processing as something their machines could do. To this potential for learning must be added that AI and computer scientists spend great amounts of time using computers and programming systems, thereby learning from them. Someone who learns to use Lisp assimilates many of the essential concepts of physical symbol systems. It is irrelevant whether John McCarthy was thinking about physical symbol systems at the time he created Lisp. In fact, he wasn't quite. He was close in a number of ways, but conceptualized it somewhat differently.

People learn about the knowledge level because they must specify systems to be programmed. They do not specify these in terms of just inputs and outputs, but also in terms of what the program knows that it can use to compute the outputs from the inputs. So they induct from their practice a working notion of what it means for a system to have knowledge and what it means to supply the symbol-level mechanisms that encode that knowledge and extract it to make the program behave as desired. It does not require a Newton to conceive of the knowledge level, all that is required is to build general purpose computers which are large enough so that large and complex programs can be built for them. So we all read out of computers an essentially correct operational notion of knowledge, without having any particular person conceptualize computers in these terms. Thus, no scientist discovered either knowledge or the knowledge level in the sense of discovery that science uses and honors.

Now we turn to what has happened to the concept of the knowledge level since I gave the talk in 1979. The story here is very mixed and varies by community.

The knowledge-based systems community has picked up the concept and uses it extensively. They have not refined it much nor do they use it in a highly technical way. But it serves them well as a way of talking about what knowledge a system must have without regard to how it is to be represented. It is not clear that they need much more than this.

The logicist community will have nothing to do with the concept. It is not that they disagree with it or think it is wrong — at least so far as I can detect. It is that it doesn't help them in what they want to do and it doesn't interest them as a concept to be pursued. Hector Levesque was intrigued enough by the notion to use it to recast his thesis work (Levesque, 1984), but it has played no continuing role in his research, as far as I can see.

The logicist community would seem to be a natural group to take up the nature of knowledge and the knowledge level. They do deal with knowledge, but in their own way. Two factors seem to me to have worked against the knowledge level becoming a topic for logicist research. First, and most important, the vehicle for research within the logicist community is a logic. There are dozens of logics — monotonic, nonmonotonic, temporal and otherwise. A logic is a particular symbol-level representational system. No matter how general a logic is, it is not at the knowledge level. The knowledge level is, in a certain sense, formless and it is not clear what to do with it. Second, the logicists pay attention to the philosophy of epistemology and take seriously its orientation on knowledge and what the proper questions are. Neither of these lead to taking the knowledge level seriously. Interestingly, there exists a well-developed axiomatic formulation of knowledge (Hintikka, 1962), but as far as I know it has never caught the attention of the AI logicist community either, except for the small foray by Levesque mentioned above.

The machine-learning community picked up the concept through an important article by Dietterich, *Learning at the Knowledge Level* (Dietterich, 1986). *Knowledge-level learning* is where the system acquires new knowledge. In contrast, *symbol-level learning* is where the system learns only to evoke the knowledge it already has faster and more reliably. Systems that take in information from the outside world are engaged in knowledge-level learning. Systems that cache are engaged in symbol-level learning — they run faster after caching but they don't know anything they didn't before. It is a nice distinction and very useful, and established the use of the knowledge level in the machine-learning community. The work is limited by its restriction that the knowledge of a knowledge-level system must be representable as the deductive closure of a set of statements in a logic. This reinforced a common error in understanding the knowledge level, since there are many other ways to represent the knowledge in a system.

For the main stream AI community, concerned with constructing intelligent systems (which includes the machine-learning community), a key implication of the knowledge level is its provision of a definition of *intelligence* (Newell, 1990 Chapter 2). The knowledge-level description of a system is the ideal point — any system that can be fully described at the knowledge level is a system of perfect intelligence, for its symbol-level mechanisms bring to bear all that can be brought to bear to solve its problems. Systems that cannot be so described (which are most systems) are intelligent to the degree they approximate a knowledge-level system — which is to say, to the degree that they can bring the knowledge in the system to bear on the system's goals. I detect very little interest in this definition of intelligence within the AI community and, in fact, very little interest at all in defining intelligence. I think this arises because a definition of intelligence does not seem to play any operational role in their research. I can tie this to my original assessment that the knowledge level had not yet been developed into an appropriate technical concept. That is still missing here. There is no way, given the current formulation, to assess the degree of approximation, hence to be able to get some leverage on how systems could be made more intelligent or any number of other questions. Dietterich's effort, described above, started this process for machine learning. No one has done it for the nature of intelligence. Until then, I can't blame the AI community for ignoring the concept.

The cognitive-science philosophy community does not pay the concept of the knowledge level any attention. There are two good reasons for this. First, philosophy essentially owned the concept of knowledge for all history up to the coming of computer science. So they believe they already have a line on the notion. Its focus was and remains on questions of certainty, and its style is epitomized by trying to establish by argument that "knowledge is justified true belief". Second, a concept of *intentional stance* and *intentional system* was introduced by Dan Dennett (Dennett, 1969, Dennett, 1988a), which is in most ways identical to the knowledge level. I believe that it puts the emphasis in the wrong place, namely on the outside observer who takes the intentional stance vis a vis an agent. The concept conveys the view that whether an agent can be described as an intentional system (i.e., at the knowledge level) is in the eye of the beholder — which is false without a fine capability of hallucination. But my attempts to convince Dennett have proved basically futile (Newell, 1986, Dennett, 1986, Dennett, 1988b, Newell, 1988). In any event, the intentional stance serves to keep the knowledge level alive within philosophy. My impression is that it plays a minor role there — something that Dennett introduced that has to be acknowledged on occasion.

The last community is a very small one, namely, the Soar community, a collection of about a hundred scientists around the world who use the Soar architecture (Rosenbloom, Laird, Newell & McCarl, 1991). Here the knowledge-level has been deeply integrated into the research on Soar. We take the nature of intelligence to be how well a system delivers the knowledge that can be ascribed to it, and find this a very useful formulation (Rosenbloom, Newell & Laird, 1991).

The most interesting thing we have done with the knowledge level is to use it in formulating the *problem space computational model.* Gradually it has become apparent that between the knowledge-level description of a Soar system and the symbol-level description (the one that talks about recognition memory, working memory, the decide process, impasses, and chunking) there is an organization in terms of problem spaces which in many ways is like another computational model. It talks only about operators, states, desired states, selection knowledge for operators etc. This must be a symbol-level organization (there is no way to have a genuine system level between the symbol level and the knowledge level), but different from the one we call Soar5.2 (the recognition memory, etc.). In searching for the nature of this independence-appearing organization, we have formulated a mutually recursive structure in which knowledge-level systems fill in the details of the problem spaces (such as selecting operators and implementing operators) and problem spaces are the symbol-level organizations that realize the knowledge-level systems. Finally, of course, there is immediately available knowledge to realize a knowledge-level system for some aspect of a problem space, and no further decomposition into subproblem spaces is required (Newell, Yost, Laird, Rosenbloom & Altmann, 1991). Thus the knowledge level is beginning to play a technical role in our attempts to define Soar as having a problem-space computational model as an intermediate level.

What will Happen?

Nothing has changed my mind about the existence of the knowledge level or the nature of knowledge in the decade since the paper was written. After all, nothing has happened in the meantime to change the way the computational community uses the concept of knowledge, which is the grounding of the paper. I think the acceptance of the concept will rise very gradually. The logicists will never give it a thought, no matter what. For the rest of us, everything will depend on some scientists turning the concept to specific technical uses, such as the assessment of the intelligence of a system, which will lock the concept into the technical fabric of our field. This, of course, is the same thought which ended the original paper.

References

Dennett, D. C. (1969). *Content and Consciousness.* London: Routledge and Kegan Paul.

Dennett, D. C. (1986). Is there an autonomous 'knowledge level'? In Pylyshyn, Z. W., & Demopoulos, W. (Eds.), *Meaning and Cognitive Structure: Issues in the computational theory of mind.* Norwood, NJ: Ablex. (commentary on Newell, same volume).

Dennett, D. C. (1988). *The Intentional Stance.* Cambridge, MA: Bradford Books: MIT Press.

Dennett, D. C. (1988). Precis of the *The Intentional Stance. Behavioral and Brain Sciences* , *11*(3), 495-546. (with commentary by 25 authors and final response by Dennett).

Dietterich, T. G. (1986). Learning at the knowledge level. *Machine Learning, 1*, 287-316.

Hintikka, J. (1962). *Knowledge and Belief.* Ithaca NY: Cornell University Press.

Levesque, H. J. (1984). Foundations of a functional approach to knowledge representation. *Artificial Intelligence, 23*, 155-212.

Newell, A. (1962). Some problems of basic organization in problem-solving programs. In Yovits, M. C., Jacobi, G. T. & Goldstein, G. D. (Eds.), *Self Organizing Systems.* Washington, DC: Spartan.

Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4*, 135-183.

Newell, A. (1982). The knowledge level. *Artificial Intelligence, 18*, 87-127. (Also in *AI Magazine, Vol 2,* Summer 1981, pp 1-20; and as CMU CSD Technical Report, Aug. 1981).

Newell, A. (1986). The knowledge level and the symbol level. In Pylyshyn, Z. W., & Demopoulos, W. (Eds.), *Meaning and Cognitive Structure: Issues in the computational theory of mind.* Norwood, NJ: Ablex.

Newell, A. (1988). The intentional stance and the knowledge level: Comments on Daniel Dennett, *The Intentional Stance. Behavioral and Brain Sciences* , *11*(3), 520-522.

Newell, A. (1990). *Unified Theories of Cognition.* Cambridge, MA: Harvard University Press.

Newell, A. & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM, 19*(3), 113-126.

Newell, A., Yost, G., Laird, J. E., Rosenbloom, P. S., & Altmann, E. (1991). Formulating the problem space computational model. In Rashid, R. F. (Ed.), *Carnegie-Mellon Computer Science: A 25-Year Commemorative.* Reading, MA: ACM-Press: Addison-Wesley.

Rosenbloom, P. S., Laird, J. E., Newell, A., & McCarl, R. (1991). A preliminary analysis of the Soar architecture as a basis for general intelligence. *Artificial Intelligence, 47*, 289-325.

Rosenbloom, P. S., Newell, A., & Laird, J. E. (1991). Towards the knowledge level in Soar: The role of the architecture in the use of knowledge. In VanLehn, K. (Ed.), *Architectures for Intelligence.* Hillsdale, NJ: Erlbaum.