Stefan Meretz (September 1998)

Meaning Concepts used in Psychology and Computer Sciences

Abstract. This paper argues that the "meaning problem" cannot be treated within the cognitive framework. Many proposals using the cognitive approach only describe conventional worlds and presuppose the meaning which they allege to explain. The conventional world restriction arises from the assumption that the syntax-semantics relationship can be realized by explicitly assigning a meaning to a syntactical unit. Owing to a conflation of characters and meanings these approaches fail to explain meaning. Alternatively, an historical approach within the framework of German Critical Psychology is presented. It is shown that meaning must be grasped as a mediation concept. Generalized production of use-values leads to objectified societal meaning structures which mediate the life of the humans and provide a space of action possibilities. It is argued that all concepts that portray meaning as immediate determinants of action fail.

Although meaning is seldom treated explicitly, it is one of the most fundamental categories in psychology and computer science. Moreover, agreement about the meaning of meaning does not exist. Commonly the 'meaning problem' is treated as 'semantics of the mental' within the cognitive paradigm. I will try to show that most common definitions of meaning are untenable.

The key problem of meaning is how to join a syntactic unit like a cipher or a word with a distinct meaning. This problem was formulated in an exceptionally clear way by Kayser:

> The following thesis (T) is generally regarded as correct:
> > (T) Every sentence which is understandable has one (or, in case of ambiguity, several) meaning(s), and there exists a formalism - different from the sentence itself - which has the power to represent it (them).

> In more technical terms, the same thesis becomes:
> > (T) There exists a space S, called the space of meanings, and a mapping s which, for every element of the set E of understandable sentences, yields one ore more elements of S. (1984, p. 168)
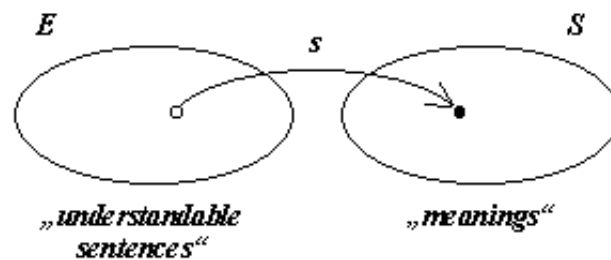

Figure 1.

Figure 1 shows thesis (T) in a graphical form. This thesis which I call the *assignment hypothesis* has a completely cyclic structure. It explains itself by defining a self-contained system of assumptions in which all conclusions are valid. This result can be shown by interpreting thesis (T) in two different ways.

Trivial interpretation

The set E contains sentences which are assigned meanings -- elements from the set S -- by using the mapping s. This assignment only makes sense when it is assumed that the sentences are meaningless before they got assigned their meaning. How, then, does s work; how can correct sentences be recognized; which gets 'its' meaning by assignment? These sentences can be recognized because they are understandable -- this is a basic assumption of the thesis. How, then, are sentences understandable? Because we know their meanings -- meaningless sentences are not understandable. By strict logic understandability implies that the meaning which is to be assigned by the mapping s is already present. Therefore, the mapping s is unnecessary. Accordingly, the sets E and S both contain meanings: their assignment is trivial. This conclusion is valid, too, if we assume that the sets contain different types of meanings, for example meanings of objects like the useable thing *chair* and symbolic meanings like the notion chair: A *chair* is a chair. The form may change, the meaning does not.

Conventional interpretation

Assuming that neither set E nor set S contains meanings in the sense of meanings of everyday life, it can be shown that the assignment thesis works well for conventional worlds. Conventions are artificial box worlds with strict rules defining the relations between signs and actions. Using a fragment of a computer program, the validity of the assignment hypothesis can be demonstrated. The clause print("Hello world") results in printing the text "Hello world" onto the screen during the program execution. Associating the character string to the set of sentences E and the resulting action to the set S, then we can state: One sentence of E 'means' one action from the set S. The assignment s of the string to the action is realized by the compiler or runtime system of the computer. In terms of computer sciences we would say that the relationship between text and action is described by the relationship between syntax and semantics. Following Newell's (1980) physical symbol system the computer 'interprets' an expression -- what he calls 'symbol' -- and performs the designated process. Newell describes conventional worlds too. The assignment hypothesis is only valid for artificial conventional worlds. The hypothesis is not valid, however, for the living world of humans.

Mixing of characters and symbols

In cognitivist theories the origin of meaning is completely unclear. Cognitivists who want to solve this problem *must define* meaning as shown above (Kayser, 1984). These definitions treat meanings as properties of things. The 'thing' get the 'property' through an act of assignment. This view leads to the conflation of characters and symbols, because they can no more be distinguished. Boden illustrates this in a vivid way:

> Computers do not crunch numbers; they manipulate symbols. (...) A symbol is an inherently meaningless cipher that becomes meaningful by having assigned to it by a user, who therefore interprets it in a particular way. Examples of symbols include road signs, maps, graphs, badges, drawings, spoken and written words in natural language, hieroglyphics, alphabetic characters in various alphabets, 'deaf and dumb' signs, and numerical digits - whether the I, V, X, ... M of the Roman system, the Arabic 0, 1, 2, ... 9 of ordinary decimal arithmetic, or the 0 and 1 of the less familiar binary code. (1977, p. 15)

In the first two sentences 'symbols' are declared to be 'meaningless ciphers', which might also be called 'characters.' The following enumeration of examples includes many very meaningful things. Road signs, maps, graphs, badges etc. are created through societal labour with a generalized meaning, otherwise no one would understand them. Then Boden enumerates examples like characters and

numerals, which have their meaning only inside self-contained systems or as carriers of meanings assigned to them. This means that the latter are clearly characters and possible carriers of assigned meanings, while the former not. Therefore meaning cannot be treated as an act of assignment and interpretation as Boden does. How should interpreters recognize which meaning they are to assign to a 'cipher' if they do not already know that meaning when they see that 'thing'? It is a circle: What has to be constructed through assignment must be previously present. One can say that an 'a' shall be an abbreviation for a road sign, but one cannot say, that a 'road sign' shall be a 'rocket'. The meaning of societally produced things cannot be changed at will. Boden conflates characters and symbols, and both remain unclear. This conflation is characteristic not only of cognitivists but of all of computer science as well.

Cognitivist concepts of meaning

Smythe (1992) gave a clear overview of the main streams in cognitive theories. He showed that Newell's (1980) *physical symbol systems* represent computational control structures, the meanings of which are externally attributed by the user. Anderson's theory of *knowledge compilation* (1976, 1982, 1983) using *production systems*, and especially the *procedural semantics* approach of Johnson-Laird (1977, 1983, 1988) can generate procedural representations from declarative utterances, but this only dynamizes Newell's approach, which, however, stays inside computational syntactic operations without reaching meanings 'outside'. Smythe showed that procedures that pretend to provide interpretations are determined by the programmer, who implements his/her understanding of a mental model. The restriction to computer models could be overcome by the *coherence view* of interpretation by Pylyshyn (1984). In his view meaning is not an intrinsic function of computational systems, but is generated from extrinsic impositions. The theorist recognizes or declares some computer output as consistent with some meaningful interpretation, so he/she must have an appropriate interpretation for output patterns in mind. This implies, however, that coherence is only given by a convention which includes prior conditions of meanings as already described above. The *narrow content* approach by Block (1986) and Fodor (1981, 1983) explicitly defines such artificial boxes which I call conventional worlds. Their aim is to link mental representations with human practices. In order to avoid a logical circle the representation of the functions 'in the head' interpreting the inputs is itself declared as 'inexpressible'. These representations are called *language of thought* (Fodor, 1975, 1990). Fodor states:

> What we're all doing is really a kind of logical syntax (only psychologized); and we all very much hope that when we've got a reasonable internal language (a formalism for writing down canonical representations), someone very nice and very clever will turn up and show us how to interpret it; how to provide it with semantic. (1981, p. 223)

We are still waiting. All attempts to implement real-world meanings in convention worlds like *frames* (Minsky, 1975) or *skripts* (Schank, 1975) have failed.

Smythe (1992) concludes that meaning that is assumed to be explained is in fact presupposed. In his hermeneutic approach based on theories of Heidegger (1962) and Gadamer (1975) and popularized by Winograd & Flores (1986), however, Smythe remains in the static view that interpretation needs understanding and vice versa. This *hermeneutic circle* (a positive action-leading notion in the hermeneutic view) can be overcome in a historical perspective of development, as I will attempt to show. Meaning is not purely 'given' but historically developed. An historical approach sharpens the view on cognitivist theories.

Logical-analytical approaches (originally based on works of Frege and Wittgenstein) discuss meaning in an inner-language framework. Most of them treat language as an axiomatic system in which the truth value of lingual utterances can be determined. As Gödel (1931) showed for the axiomatic system of mathematics, however, any axiomatic system as a whole is not provable. This means that statements can only be proven under the conditions of the system but not for the general case exceeding the system borders. Related to the meaning problem, this conclusion indicates that neither the origin of meaning nor the relationship between meanings and characters can be clarified. The chains of the logical-analytical approach can be shaked off using the logical-historical approach presented here.

The origin of meaning

References of meanings to meanings are used daily. We deal with things, speak about them, read texts, follow graphical symbols etc. Without this trivial and permanent use of 'assignments' of meanings to meanings society would not work. It does not follow from this, however, that meaning is a Platonistic term which describes any affinity between the essence of a thing and a symbol (Katz, 1981, 1984). Meaning must be understood as a *mediation concept*. Meanings mediate between 'things' and us. Meaning is nothing absolute. It is neither a property of objects nor a property of individual perception. Meanings are relevant not only for humans but also for animals. Meanings emerge out of an evolutionary and historical process. The approach presented here is based on research work of German Critical Psychology (Holzkamp, 1973, 1983, 1992, Tolman & Maiers 1991, Tolman 1994, Lenz & Meretz 1995).

Meaning in the evolutionary process

Very simple organisms (e.g. cells) can discriminate their environments only in a binary way: for them it is either usable for metabolism or not usable and perhaps dangerous. Owing to the direct contact between the organism's surface and its surrounding substances, these substances have potentially two functions at the same time: they can be usable for metabolism and they can serve as signals. The separation of the metabolic and the signal functions of the substance through evolutionary development of the organism was a an important step forward. Non-metabolizable neutral substances acquired the role of a neutral and mediating "third" between the organism and environment. This mediating third did not appear out of nowhere: it was through a process of differentiating its functions that the organism came to be able to use these neutral substances as signals. Following Leontyev (1979), *psyche* is the life activity of the organism that has become mediated by signals. As a third, the signal references environmental aspects, it has a meaning for the organism. Realization of the meaning depends on the inner state, the need, of the organism. If signal and need are present the activity will be take place automatically. If one of them are not present then no activity follows and consequently no meaning is realized; or, more precisely, without activity the meaning does not exist for the organism. The meaning of a signal only exists during the activity. Realization of the meaning and activity are identical. Each species lives in an environment with its own specific meanings determining its activities.

Meaning in the historical process

Skipping over some stages of evolution -- e.g., development of orientational meanings, learning of meanings etc. -- we reach the field of transition between animals and human beings. While on the animal level of development the relationship between meaning and activity is *deterministic*, on the human level meanings acquire the character of action *possibilities*. Meanings are those objective aspects of the world that are used by humans, they are the bases for individual grounds for action.

Higher animals use means for single ends. They want to reach a goal, so they create tools. The meaning of a tool exists, however, only for the period of usage. The tool then looses its meaning and vanishes in the natural environment. Humans create tools for future use and for use by others. In this case the tool exists *before* any particular need to use it exists. The relationship between means and ends is reversed. This reversal of means and ends has consequences for the concept of meaning.

1. The meaning is conserved. While the meaning of tools used by animals is lost when the activity ends, the meaning of the tool manufactured by humans remains even when the tool is not in use. Due to the generalized production of tools (or others things), and due to the enduring meaning of those things, we can conclude that their meaning is an objective fact. This objective character of things created through human labour makes it possible for other humans to recognize their intended meaning and employment.

2. Meanings are societallyproduced. Animals individually find meanings in the environment determining their actions. Humans on the other hand create their living conditions by producing objects and their meanings in a societal manner. This implies that the relationship between humans and nature is not individual but societal. This leads to another advantage.

3. Societally produced meanings can be cumulated. Animals can pass on their practical experience of creating and using tools only directly to other individuals. If the animal population perished, the animal tradition of experience would be lost too. By contrast, human practical experience gets fixed with the objects which are produced. This societally fixed pool of experience is stable and cumulative. The knowledge concerns the tool and the use of the tool, because the application of a tool reflects the properties of the tool itself as well as of the object for which the tool is used.

4. Synthesis of meaning structures of entire society. The meanings produced in form of objects do not co-exist unrelated. Because societal production is generalized, every meaning is bound into a web of other meanings via the needs of the productive process. One tool is produced with the aid of another tool, which again itself is produced with other tools, or in other circumstances it can be used for other purposes, and so on. With an increasing and differentiated production of objects the associated meanings are integrated into a web of meanings which extends over the entire society.

The means-end reversal lead to a societal production of generally useful objects. This process causes a change in cognitive capabilities of pre-humans too. Thus, on the individual level we can add the following consequences.

5. The purpose of use is anticipated. As tools are produced for the general case of a future use, the purpose of the tool has to be ideally anticipated. This anticipation is not yet thinking! One can imagine these anticipations as a kind of presentiment of future meanings to be produced without presence of speech.

6. Essential and unessential properties of tools are distinguished. Anticipating the work result and its final purpose of use, it is necessary to distinguish between essential and unessential properties of the tool during production. These abstractions are not a conscious act of thinking, but again a kind of presentiment of future meanings of the tool.

Both aspects, anticipations and abstractions during the working process, are *preconditions* and *not results* of speech. Anticipations and abstractions are prelingual symbolic representations of requirements of generalized production. These *practical symbols* are meanings which are not bound to the presence of an object. Due to communication requirements during cooperative work, practical symbols develop into acoustic-, picture-, and character-based representations of symbolic meanings. Symbolic meanings create their own relatively independent web structures of meanings, like logic,

mathematics etc. They are always back-referenced, however, to the meanings of societally produced objects.

Language and meaning

Language is a powerful carrier of symbolic meanings. Notions are successors of practical symbols used by hominids. Like practical symbols, they represent the main object dimensions of purpose and property. Notions are symbol meanings. An object or a thing or matter is always perceived in the form of its notion. Perception is recognition of the general in the special. When I search for the notion "chair" then I search for the object meaning "to-sit-on", because a chair was made for this general purpose.

Symbol meanings can have different carriers. The character shape of a notion can change. Therefore, *notion* as content and *word* as form have to be sharply separated from one another. The notion "chair" and the word "chair" are two different things. The object meaning "to-sit-on" was created through societal production, because the produced thing fulfilled the anticipated purpose. Hence, the meaning is not changeable and cannot be defined by agreement. On the other hand, the forms in which the meaning is transported, can be replaced: "chair", "Stuhl" or "cadeira" are different word forms for the same meaning. Translation is based on the possibility of changing the transport forms of meanings. Of course, references to shifted meaning in other languages may occur, but this only shows different developments of society and their reflection in lingual forms.

Characters cannot *be* symbols, characters are meaningless. They can be sensual covers of a meaning, they can carry a meaning. Symbols do not *have* a meaning, they *are* a meaning. The word "symbol" has a greek origin and can be translated to a "picture which makes sense". The sense (or meaning) can be carried by the picture, and the picture can be changed, the sense does not. Thus, a meaningless symbol is a contradiction by itself. A "picture which makes sense" without any sense does not make sense!

Conclusions

German Critical Psychology provides a metatheoretical framework for research on both psychological and computational tasks. One important part of this is the logical-historical development of the meaning category. It is shown that meaning is nothing absolute but objective. Meaning is neither a property of things nor only present as an imagination of cognition. Thus, meanings cannot be 'defined' or 'assigned' as commonly thought. Meanings arise from societal production of use-values. The objectified meanings have to be recognized by the society on average. They can, however, remain unrecognized individually or be perceived in a deviating sense, because the generalized production does not determine the individual use. Thus, the relationship between the objective character of meaning through societal production and subjective perception and use is one that is only determined by the individual based on his or her individual grounds for action. In contrast to animals, humans live in a possibility relation to the world. They can choose, they can make objective conditions/meanings into premises for their action, but they can also ignore them or act in still some other way. Consequences of this for psychology are far-reaching and have been discussed elsewhere (for example with a special focus on *learning* in Holzkamp, 1993).

The meaning category sketched in this paper can provide sharp criteria for computer sciences too. In this perspective characters and symbols (or syntactic units and semantics) -- mostly mixed -- can be clearly distinguished from one another. This leads, however, to a complete rejection of all cognitivist attempts to represent meanings of the human living world by computers like in 'artificial intelligence' or 'connectionist' approaches (for more details, see Lenz & Meretz, 1995). It enables the researcher to

reconstruct the subject 'outside' the computer instead trying to map it onto formal structure inside the computer. Moreover, it gives new criteria for a subject-oriented software development process, but this must be the topic of another paper.

References

Anderson, J. R. (1976). Language, Memory, and Thought. New York: John Wiley.

Anderson, J. R. (1982). Acquisition of Cognitive Skill. Psychological Review, 89, 369-406.

Anderson, J. R. (1983). The Architecture of Cognition. Cambridge/MA: Harvard University Press.

Block, N. (1986). Advertisement for a Semantics for Psychology. In P. A. French, T. E. Uehling, & H. K. Wettstein (Eds.), Midwest Studies in Philosophy: Vol. 10. Studies in the Philosophy of Mind. Minneapolis: University of Minnesota Press.

Boden, M. A. (1977). What Is Artificial Intelligence? In M. A. Boden (Ed.), Artificial Intelligence and Natural Man, New York: Basic Books.

Fodor, J. A. (1975). The Language of Thought. New York: Crowell.

Fodor, J. A. (1981). Representations. Philosophical Essays on the Foundations of Cognitive Science. Brighton: Harvester Press.

Fodor, J. A. (1983). Modularity of Mind: An Essay on Faculty Psychology. Cambridge/MA: MIT Press.

Fodor, J. A. (1990). Why there STILL Has to Be a Language of Thought. In D. Partridge & Y. Wilks (Eds.), The Foundations of Artificial Intelligence - a Sourcebook. Cambridge: Cambridge University Press.

Gadamer, H.-G. (1975). Truth and Method (G. Barden & J. Cumming, Trans.). London: Sheed & Ward.

Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. Monatshefte für Mathematik und Physik, 38, 173-198.

Heidegger, M. (1962). Being and Time. (J. Macquarrie & E. Robinson, Trans.). New York: Harper & Row.

Holzkamp, K. (1973). Sinnliche Erkenntnis. Historischer Ursprung und gesellschaftliche Funktion der Wahrnehmung. Frankfurt/Main: Athenäum.

Holzkamp, K. (1983). Grundlegung der Psychologie. Frankfurt/Main: Campus.

Holzkamp, K. (1992). On Doing Psychology Critically. Theory and Psychology, 2(2), 193-204.

Holzkamp, K. (1993). Lernen. Subjektwissenschaftliche Grundlegung. Frankfurt/Main: Campus.

Johnson-Laird, P. N. (1977). Procedural Semantics. Cognition, 5, 189-214.

Johnson-Laird, P. N. (1983). Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness. Cambridge/MA: Cambridge University Press.

Johnson-Laird, P. N. (1988). The Computer and the Mind: An Introduction to Cognitive Science. Cambridge/MA: Harvard University Press.

Katz, J. J. (1981). Language and Other Abstract Objects. Totowa/NJ: Rowman & Littlefield.

Katz, J. J. (1984). An Outline of Platonist Grammar. In T. G. Bever, J. M. Carrol, & L. A. Miller (Eds.), Taking Minds: The Study of Language in the Cognitive Sciences. Cambridge/MA: MIT Press.

Kayser, D. (1984). A computer scientist's view of meaning. In Torrance, S. B. (Ed.), The Mind and the Machine. Philosophical Aspects of Artificial Intelligence. Chichester: Ellis Horwood.

Lenz, A., and Meretz, S. (1995). Neuronale Netze und Subjektivität. Braunschweig/Wies-baden: Vieweg.

Leontyev, A. N. (1979), Problems of the Development of the Mind. Moscow: Progress Publishers.

Minsky, M. (1975). A Framework for Representing Knowledge. In P. H. Winston (Ed.), The Psychology of Computer Vision, New York: McGraw-Hill.

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

Pylyshyn, Z. W. (1984). Computation and Cognition: Toward a Foundation for Cognitive Science. Cambridge/MA: MIT Press.

Schank, R. C. (1975). Using Knowledge to Understand. In Association for Computational Linguistics (Ed.), Theoretical Issues in Natural Language Processing, Cambridge/MA.

Smythe, W. E. (1992). Conceptions of Interpretation in Cognitive Theories of Representation. Theory & Psychology, 2(3), 339-362.

Tolman, C. W. (1994). Psychology, Society, and Subjectivity. An Introduction to German Critical Psychology. New York: Routledge.

Tolman, C. W., and Maiers, W. (1991) (eds). Critical Psychology: Contributions to an Historical Science of the Subject. New York: Cambridge University Press.

Winograd, T., & Flores, F. (1986). Understanding Computer and Cognition: A New Foundation for Design. Norwood/New Jersey: Ablex Publishers.