A Classification of Associative and Formal Concepts

Uta Priss, Napier University

Researchers in psychology and linguistics (Sloman (1996), Pinker (1991) and others) have argued that there is evidence to support at least two co-existing systems of reasoning: an *associative* or sub-symbolic one and a *formal*, rule-based or symbolic one. This distinction occurs in many disciplines both in empirical models and in models that underly software implementations (Priss, 2001). In AI there is a divide between biologically inspired (i.e. associative) and logical-symbolical (i.e. formal) approaches. In cognitive science, a traditional Aristotelian formal model of concepts competes with fuzzy, prototype-based models that can be traced back to Wittgenstein (1953) and Rosch (1973) and with models of embodied cognition (Lakoff, 1987). In information science, traditional, formal approaches led to the construction of classification systems and web directories whereas associative approaches led to flexible, network-based information access systems.

This paper argues that associative and formal conceptual structures are combined in human cognition. Animal cognition involves only associative structures, whereas in human cognition a system of formal structures is set on top of associative structures. Cognitive development involves several stages, such as sub-symbolic stages, proto-language stages and stages with full language. Examples of proto-language are the linguistic abilities of children between 1 and 2 years of age and the linguistic abilities of some apes that have been taught linguistic symbols. Proto-language is characterized by a fairly small vocabulary of basic symbols and simple 1 or 2-word sentence structures. Full language is characterized by the existence of syntax which facilitates grammatical nesting and meta-language expressions. This paper claims that associative conceptual structures occur at all levels of cognitive development and form the basis of formal structures. But formal conceptual structures are a necessary and sufficient requirement of full language: formal concepts can only be expressed using symbols of a full language and the expression of full language requires the existence of formal concepts.

This paper describes a ten-fold classification of concepts that correspond to different stages of cognitive development. These stages do not form a linear progression but instead provide an explanation for a step-wise but modular or multi-path development of language. Examples from research in animal cognition provide evidence for the plausibility of the different stages. The classification itself, which is derived based on an analysis of the interactions between objects, concepts and signs, is similar in structure but not identical in content to Peirce's classification of signs. The classification provides a formal, philosophical framework that can serve as a common ground for integrating a variety of current theories of language development and cognition.

Conceptual structures that underly languages are not simply a union of more basic symbolic structures and sub-symbolic structures. Instead, to explain the full complexity of human cognition, dynamic interactions between associative and formal structures must be assumed. In analogy to Clark's (1997) notion of a *mangrove effect* which explains feedback loops that are initiated by a combination of external and internal representations, feedback loops in the interaction between sub-symbolic and symbolic structures can be identified (Priss, 2001). These feedback loops explain the exponential growth in cognitive abilities which arises from the combination of associative and formal structures.

1 The distinction between associative and formal

The divide between associative and formal structures occurs in many disciplines. While there may not be a precise definition of “associative” and “formal” that fits all these distinctions in different disciplines, a list of representative features can be compiled: Associative structures are usually fuzzy, complex, and emergent whereas formal structures are precise, defined or designed. Associative structures can be represented with words but also as maps, networks or other diagrams. The forms of the representations matter. For example, the associative content of poetry cannot easily be translated into other languages because of connotations. Formal structures can be represented using symbolic logic, rule-based knowledge systems, and conceptual graphs (Sowa, 1984). It is possible to translate between different formal representations because only the structure of representations matters (such as whether they are equivalent to first order logic) but not the form. In general, associative structures are grounded and depend on experiences, perception and observation. Formal structures on the other hand are often designed in a top-down manner and are theoretical. The main reasoning mechanisms of associative structures are analogy and recognition based on observation of similarity and co-occurrence; whereas the main reasoning mechanisms of formal structures are deduction, logical inferences and the establishment of causal explanations within a theory. Further details on the differences between associative and formal structures can be found in (Priss, 2001) and (Priss, 2002).

2 A classification of concepts

Deacon (1997) provides a convincing account of the differences between associative and formal structures and their involvement in the evolution of language and the human brain. One (minor) short-coming of his work is that he identifies “associative” with “indexical” and “formal” with “symbolic” (both in the Peircean sense). This can be misleading because, for example, human language, dog barking and communication among bacteria all have symbolic aspects but only human language can be considered “formal”. Thus the distinction between indexical and symbolic is necessary but not sufficient for distinguishing associative and formal concepts.

This paper suggests that in fact three dimensions are required for a distinction between associative and formal concepts. Each of the three dimensions has three classes. The three dimensions, which are explained in detail in the rest of this paper are the sign dimension, the dimension of internal representations and the object dimension. Normally three dimensions with three classes each would yield a direct product of 27 classes. But there are dependencies between the different dimensions which reduce the number of classes to 10. This classification is structurally equivalent to Peirce’s ten-fold classification of signs but the content of the classes is different. The structural equivalence may be due to the fact that in both cases the three dimensions contain three classes which are increasing in complexity i.e. what Peirce calls Firstness, Secondness and Thirdness.

2.1 The sign dimension

Figure 1 shows the three classes of the sign dimension, "sign = object", "iconic or indexical" and symbolic. This dimension refers to an external viewpoint of an observer who observes the relation between signs and objects within an associative context. Whether something is considered a sign or an object depends on the observer's viewpoint. Regier et al. (2001) state that an observer differentiates form and meaning by considering the focus of attention. Object (meaning) is whatever is the focus of attention and sign (form) is whatever is associated with an object but not the focus of attention. For example, if a mother teaches her child a new word, she will look at the object and utter a word. The word is thus a non-focal co-occurrent of the object. It should be emphasized that objects need not be physical objects but instead they can be activities, attributes, feelings and so on. Anything that a sign can refer to can be an object of that sign.

The first class, "sign = object", in this dimension refers to associative relations in which the sign and the object are essentially identical. Examples are contexts in which an observer views an object without any interpretation or intention and no communication is involved. No examples for this are provided in figure 1 because this dimension has no communicative value.

Iconic and indexical signs are grouped together in the next class because both involve a physical or causal relationship between a sign and an object which is grounded in an associative context. This corresponds to Deacon's (1997) interpretation who also considers the distinction between symbolic and the other two more significant than the distinction between iconic and indexical. An example for iconic similarity are shared features between object and sign, but this is also usually a physical relationship. Pointers establish causal, indexical relationships between signs and objects. Due to the physical or causal relationship, an observer needs no further information (such as linguistic knowledge) to identify a relation between object and sign but, of course, causality is observer-dependent.

Peirce's example of the sunflower which is an index for the location of the sun falls into this class. Reflexes in animals or humans are also usually indices of a perceived stimulus. Behaviors, such as running away, can be indexical of objects, such as "danger". Onomatopoeic words, such as "cuckoo" or "woof", are iconic of objects. A more complex example is how wolves use a communication system of pointing with their gaze and the direction of their muzzles during hunting activities (Coren, 2000). The wolves communicate relational objects such as "come", "sit down" and "go there" in that manner.

Symbolic signs are signs that are part of a conventional system, which can be a proto-language or language. A characteristic of symbolic signs is that they are habitual and cannot be understood without knowledge of the language. Even hormonal communication or communication among bacteria falls into this class because it cannot be understood by an observer who does not know the "code". Furthermore this communication is usually systematic. That means that the signs are used in a systematic manner and there may be a set of contrasting signs to represent contrasting objects. Complex

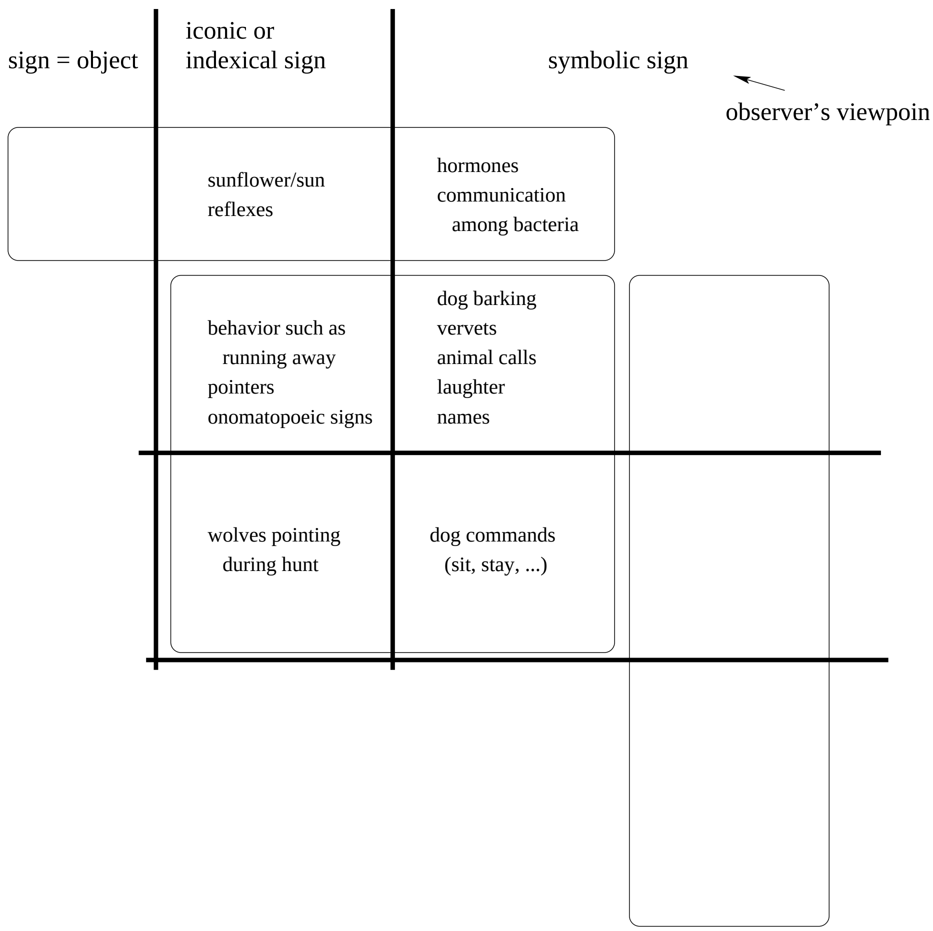

Fig. 1. The sign dimension

systems of animal calls, such as the different barks a dog can produce (Coren, 2000), also fall into this class for the same reasons. Deacon points out that humans have retained a few of such calls, such as laughter and sobbing. The degree to which these calls are innate or learned is irrelevant for the purpose of this classification because that does not affect the nature of the signs. For example, children who have been raised by wolves exhibit wolf behavior which is certainly not innate but may be identical to innate wolf behavior in its sign character. There are obvious differences between dog barking and human language because dogs cannot communicate abstract objects but humans can. But this first dimension is not sufficient to characterize this difference.

2.2 The dimension of internal representations

The second dimension pertains to the internal representations or conceptualizations that mediate between the perception or contemplation of objects and the production of signs. "Internal representation" in this paper refers to the existence of an internal brain-like or higher-order neural representation within the sign producer. This dimension does thus not pertain to the viewpoint of an observer but instead to the viewpoint of a sign producer. But observers often have some limited means of determining the existence of internal representations of other sign producers based on certain clues. The three classes in this dimension are represented in figure 2 as follows: the first class is the top row, the second class is the square of four classes in the center, the third class is the column on the right.

In the first class, there is no internal representation. Examples of the lack of internal representations are a sunflower turning to the sun, communication among bacteria or hormonal communication. These processes are entirely deterministic or of the stimulus/response type without an opportunity for choices. Learning can only occur at the system level through evolution but individuals cannot learn during their life-time and cannot change their behavior. In this class an external object as input to an agent is directly (although possibly with temporal delay) followed by the output of a predetermined sign. This sign can be iconic/indexical, for example, it can involve a direct causal relationship such as fear/sweat or reflexes, or can be symbolic such as in the case of hormones.

In the second class, the internal representations are opaque from the sign producer's viewpoint. From an observer's viewpoint there is evidence for the existence of internal representations provided by the fact that the sign producer appears to have choices. The sign producer does not appear to react according to simple stimulus/response mechanisms or deterministic input/output processes. Instead the sign producer's behavior is influenced by subtle contextual changes in a complex manner. But there is no evidence that the sign producers at this level can reason about their internal representations. Cognitive abilities that can be achieved in this class are stimulus generalizations (i.e., objects can be grouped into categories) but the category boundaries are fuzzy and based on prototypes. For example, current artificial neural networks can learn to categorize simple and relational objects but the networks cannot also output the reasons why they categorize in a certain manner. The symbolic signs produced by opaque internal representations are limited to one-word statements, such as produced by 1-2 year old children and most animal calls.

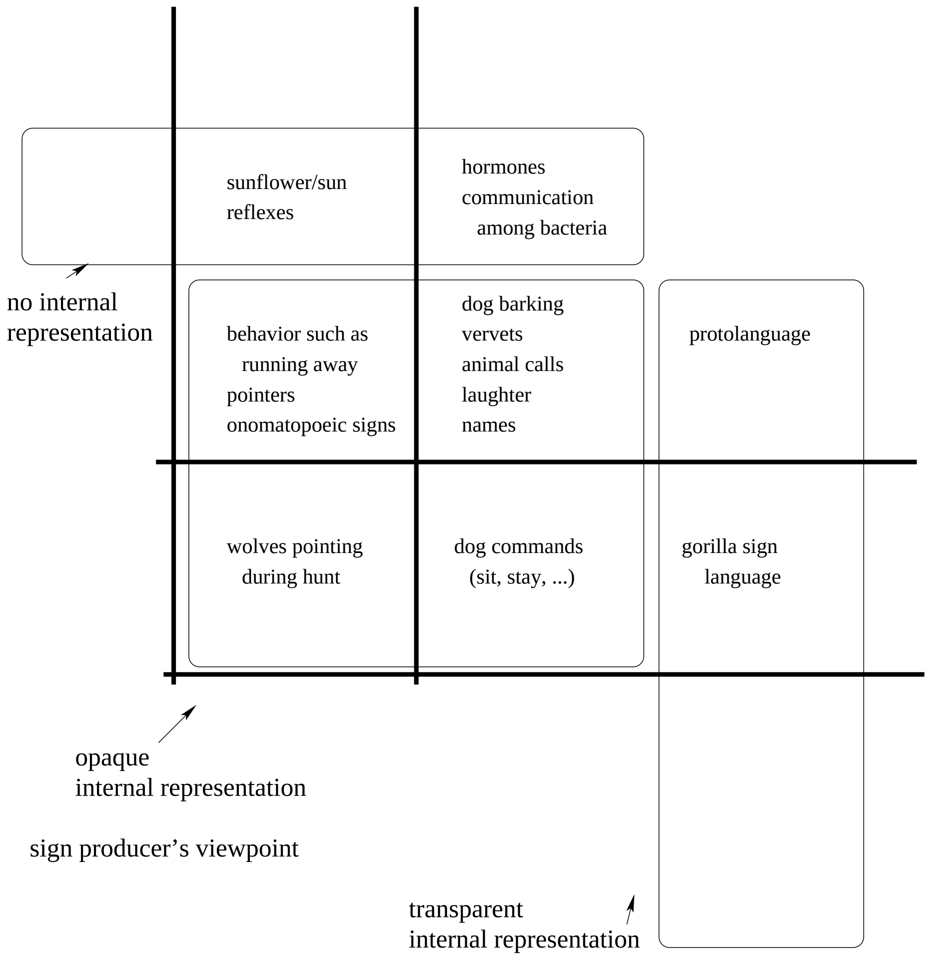

Fig. 2. The dimension of internal representations

The last class contains transparent internal representations which means that the sign producer appears to have some insight into her internal representations. From an observer's viewpoint, the evidence for this is the fact that the sign producer can build simple syntactic combinations of signs in the form of object/attribute (object HAS attribute) or object/class (object ISA class) associations. Instead of simply associating objects and signs, the sign producer can thus express some reasons why objects and signs are associated. This stage represents proto-language (Devlin, 2000) and an example is the use of sign language by gorillas (PBS, 2001). These are simple sentences of subject-verb structure but without nesting. The notion of "transparency" is not meant to imply that at this level all concepts are fully transparent to sign producers or that they can be consistently defined. Concepts at this level may only exist in the sense of prototype theory (Rosch, 1973). But "transparency" means that humans are capable of contemplating at this level about what their concepts are made of. Transparent internal representations require the use of symbolic representations because symbols are required to express the intensions. Full human language also falls into this class but the difference between full language and proto-language cannot be explained without the third dimension.

2.3 The object dimension

The three classes of the object dimension are simple objects, relational objects and abstract objects. This dimension refers again to an external viewpoint of an observer who observes these objects within an associative context. The classes are thus not intrinsic features of objects but based on the judgments by an observer. Figure 3 shows all three dimensions with each of their three classes.

Simple objects are gestalt-like structures or patterns in an external world. Examples are "stone", "hot", "yellow", and "three". Simple objects are similar to what Lakoff (1987) calls "basic-level structures". He states that basic-level structures arise "as a result of our capacities for gestalt perception, mental imagery, and motor movement" (p. 302). Even though "hot" and "yellow" are fuzzy when used in language and vary among different speakers and situations, they correspond to simple physiological gestalts: "hot" corresponds to an unpleasant heat sensation and "yellow" to one of three color receptors in the human eye. Devlin (2000) explains that the numbers one, two and three are perceived in an immediate manner by many animals and by humans and do not require a counting ability.

Following Lakoff & Johnson's (1999) argument about the embodiment of cognition, it is conceivable that gestalt perception is a deterministic property of an external world constrained only by the physical, bodily properties of perception. That means that beings with similar bodies and perceptive mechanisms are capable of perceiving similar simple objects. The notion of "objects in an external world" is to be understood in this manner. Principles of gestalt perception have been established by psychologists and can be simulated using artificial neural networks. The more challenging aspect of gestalt perception is not to form gestalts but to select the ones which are interesting for or relevant to an individual in a situation. (This is one of the main challenges for data mining applications.)

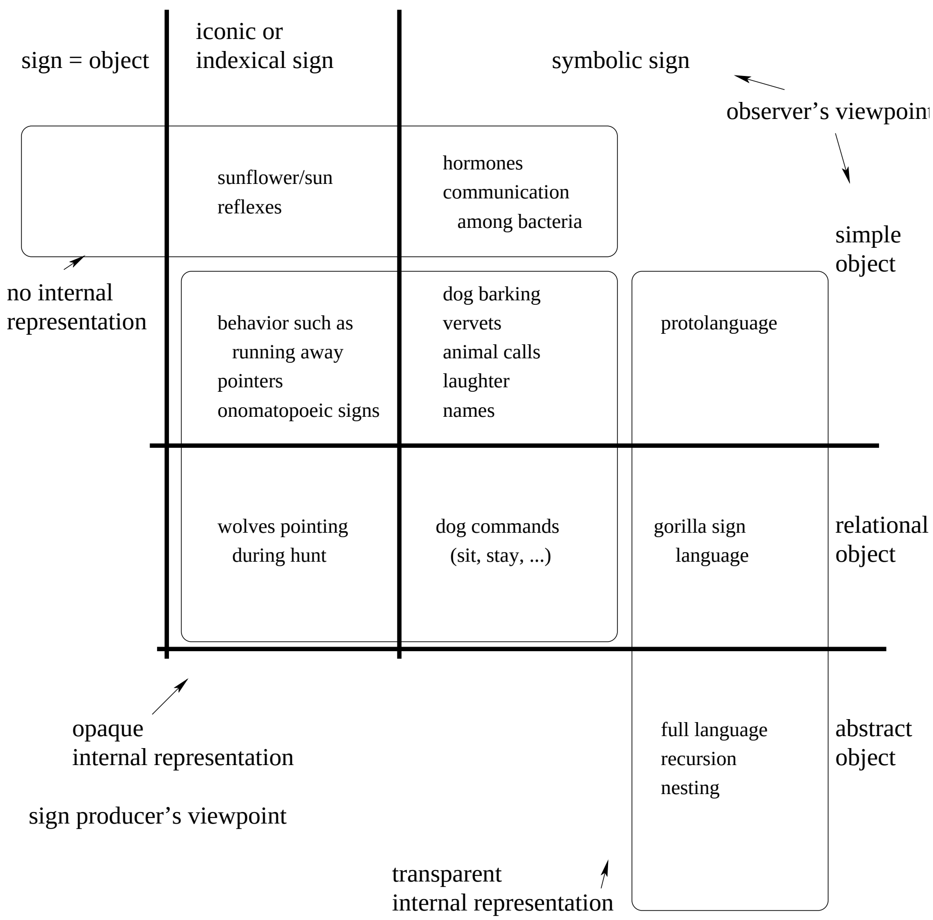

Fig. 3. All three dimensions

Relational objects are objects that consist of relations among objects. Examples of relational objects are part-whole relations, many prepositions and verbs. For example, “over” is a relational object that consists of a relation between two simple objects. These relations are usually identified with respect to an “external world”. To recognize such relations, some kind of internal representation (or conceptualization) is required, thus relational objects cannot occur in combination with the “sign = object” class of the sign dimension.

In contrast to simple and relational objects, abstract objects are always culturally determined. They are defined as objects that are under no circumstances directly emergent from an external world but have components that are culturally created and require interpretation. Examples are the abstract notions of ”mathematics” and “democracy”. Typically it is possible to represent simple and relational objects in an iconic or indexical manner. But abstract objects must be represented symbolically using the symbols of a language. Abstract objects are thus a defining characteristic of full language: only full language can represent abstract objects.

Animals, which have similar body size and perceptive abilities as humans, are most likely capable of perceiving the same simple and relational objects as humans. But, for example, only if it were possible to explain a series of abstract concepts, such as “democracy”, to a gorilla, there would be an indication that gorillas are capable of full language not just proto-language. Explaining abstract objects does not mean simply training an animal to perform some activity when it encounters the sign of an abstract object. Instead it involves establishing relationships between abstract objects. Deacon (1997) explains this difference between processing animal calls and concepts that are expressed by full language in detail.

Further features of formal concepts are the facilitation of off-line thinking, i.e. thinking about objects that are not necessarily part of the immediate physical environment (Devlin, 2000, p. 219); the expression of meta-level statements and recursion. It is not possible, for example, to express “the word *word*” using associative concepts. Formal concepts facilitate nesting, such as “I believe that ...” or “John says that ...”, and hypothetical statements.

3 Abstract objects define formal concepts and full language

Only abstract objects require transparent internal representations and full language (as opposed to protolanguage). And full language always requires some abstract objects. Figure 4 shows the distribution of associative and formal concepts over the ten classes. This distribution does not mean that human cognition is restricted to the realm of formal concepts. On the contrary, human cognition heavily involves associative structures. All non-linguistic cognition falls into the associative areas. Thus human emotions, reflexes and instincts may be very similar to those of animals. But only humans are capable of thinking in formal concepts and of using full language.

Concepts can change their nature and migrate from formal to associative. For example, unicorns are originally formal because they do not exist and have been invented by humans. But they can obtain a virtual existence. For example, humans might associate a certain shape, color, smell, and texture with unicorns which is stored in the brain in exactly the same manner as similar associations of horses. Devlin (2000) argues that mathematicians think about mathematical objects in exactly the same manner as other people think about physical objects.

Associative concepts are embedded into interactions with an external world because simple and relational objects are elements of an external world. Abstract objects, on the other hand, can be imaginative or formally constructed from other objects. Thus formal concepts do not directly refer to external objects but instead they refer to other formal concepts. Deacon (1997, p. 83) states that “Words also represent other words. In fact, they are incorporated into quite specific individual relationships to all other words of a language.” and “This referential relationship between the words [...] forms a system of higher-order relationships.” As a system, formal concepts are grounded into a system of associative concepts. But not every individual concept has a direct relationship with associative structures. This implies that a search for universals or “primitive” concepts may be futile.

The system of formal concepts is also somewhat independent of a language system. For example, when subjects are questioned about a text they read, they can usually recall the conceptual content of the text but not the exact wording. Bilingual speakers may remember the content of expressions but not in which language these were uttered. Thus conceptual knowledge and language form independent systems, which relate to and depend on each other but are not mapped in a one-to-one relation.

4 Conclusion

There has been a recent increase in interest in associative and formal concepts, which have been identified as a dichotomy in several disciplines. This paper provides a classification of associative and formal concepts using ten classes and three dimensions. An understanding of the ways in which formal and associative concepts combine in human cognition can have major impact on the development of artificial intelligent devices. While it is fairly well understood how to implement associative and formal structures, for example, as artificial neural networks and formal logic, it is not yet well understood how to combine them. This may be one of the reasons why artificial intelligence cannot yet simulate human intelligence at a very high level.

References

1. Clark, Andy 1997. *Being There. Putting Brain, Body, and World Together.* MIT Press.
2. Coren, Stanley 2000. *How to Speak Dog: Mastering the art of dog-human communication.* The Free Press, 2000, New York.
3. Deacon, Terrence 1997. *The Symbolic Species. The Co-Evolution of Language and the Human Brain.* Penguin Press, London.
4. Devlin, Keith. 2000. *The Maths Gene. Why everyone has it, but most people don’t use it.* Weidenfeld & Nicolson, UK.
5. Lakoff, George. 1987. *Women, Fire, and Dangerous Things.* Chicago, IL: University of Chicago Press
6. Lakoff, George and Johnson, Mark 1999. *Philosophy in the Flesh.* Basic Books, New York.

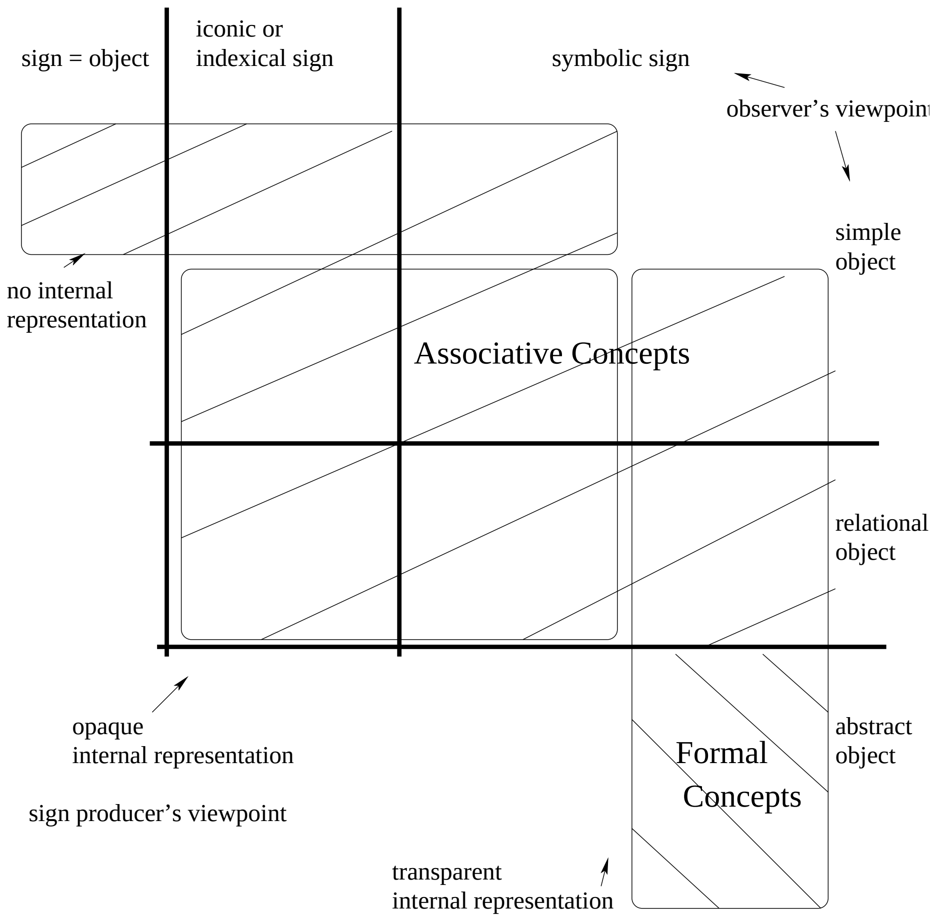

Fig. 4. Associative and formal concepts

7. PBS Nature. 2001. *A Conversation with Koko.* On-line available at http://www.pbs.org/wnet/nature/koko/index.html.
8. Pinker, Steven. 1991. *Rules of language*. Science, 253, p. 530-535.
9. Priss, Uta. 2001. *Multilevel Approaches to Concepts and Formal Ontologies*. In: Proceedings of the 12th ASIST SIG/CR Classification Research Workshop.
10. Priss, Uta. 2002. *Associative and Formal Structures in AI.* In: Proceedings of the 13th Midwest Artificial Intelligence and Cognitive Science Conference, Chicago, 2002, p. 36-42.
11. Regier, T., Corrigan, B., Cabasaan, R., Woodward, A., Gasser, M. and Smith, L. 2001. *The Emergence of Words.* Proceedings of the 23rd Annual Meeting of the Cognitive Science Society, p. 815-820.
12. Rosch, Eleanor. 1973. *Natural categories*. Cognitive Psychology, Vol 4.
13. Sloman, S. A. 1996. *The empirical case for two systems of reasoning.* Psychological Bulletin, 119, p. 3-22.
14. Sowa, John F. 1984. *Conceptual Structures: Information Processing in Mind and Machine.* Addison-Wesley, Reading, MA.
15. Wittgenstein, Ludwig. 1953. *Philosophical Investigations.* Basil Blackwell, Oxford.