



## Ontology Engineering by Fuzzy Cognitive Maps

*Abstract:* We present work in progress at Oracle's Warsaw Mobile and Wireless Center of Expertise to develop and manage knowledge in order to be able to store, search and make reasoning with it. We present why we strongly see the problem of content management as part of a much broader field of knowledge management where the representation falls in the domain of the Ontology and the computation is carried on by a Fuzzy engine related to the approach derived from that of Fuzzy Cognitive Maps (FCM).

### 1. Content management and ontology

It has been observed [1] that three major paradigm shifts have been occurring in AI research community: (1) from process-centered to Information-centered, (2) from computer-centered to human-centered, and (3) from form-centered to content-centered. Not only but Content Management is identified with a key role in the Knowledge society of our century.

There are several approaches to Content Management; however the common root is that, in general, Content Management deals with the creation, collection, storage and refinement of structured content and media assets in a managed collaborative environment. In a sense, it facilitates assembly, analysis, processing and re-use of content.

Our goal is that of being able to make "reasoning" about the content, and the target would be the one of passing the Turing Machine test. Surely the goal is ambitious, but, others are working similar directions. In particular Lenat and Guha [2] are attempting to construct a massive knowledge base – where the knowledge is not too domain specific - containing millions of encoded facts, categories, relations, and so on, with the intent that the finished knowledge base will define our consensus reality -- will capture the basic knowledge required to comprehend, for example, a desk top encyclopedia. This effort is the enCYClopedia project. CYC is not a dictionary, but a huge common sense knowledge base whose upper level structure is an ontology.

Trying to clarify ontology and knowledge is a very debatable subject as there is no clear boundary between the two. Starting from definitions The American Heritage Dictionary defines Ontology as "the branch of the metaphysics that deals with the nature of being" a

multitude of different perspectives is grown. The term ontology as been adopted by the AI community to refer to a set of concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. In the context of AI, the ontology of a program is described by defining a set of representational terms. In such ontology, definitions associate the names of entities in the universe of discourse (e.g., classes, relations, functions, or other objects) with human-readable text describing what the names mean, and formal axioms that constrain the interpretation and well-formed use of these terms.

In information technology, ontology is the working model of entities and interactions in some particular domain of knowledge or practices, such as electronic commerce or "the activity of planning.

We could continue in listing different point of view, but the fact is that no overall and generally accepted definition of ontology exists. That's the reason why we are taking a pragmatic approach and consider an ontology as something that needs to be constructed. There are multiple examples of these constructions: glossaries, terminology databases, encyclopedias, knowledge bases, etc.

When these pragmatic constructions (most often domain-specific – the only exception being CYC) are built, there is also another set of problems that should be considered:

- What is the impact of the designer (typically a "knowledge engineer") on his/her design?
- What is the coverage of this knowledge?
- What is its expressivity?
- Which sort of computation can be done on top of it?

### 2. Computational Aspects

Our objective is to explicit the relationship with computational semantics in order to be able to carry on reasoning. In this direction [2] proposes 3 levels of ontology:

- Level 1: which is a structured collection of terms defining the concepts of the hierarchy. E.g. Yahoo taxonomy.
- Level 2: which is the definition of concepts and relationships.

- Level 3: which is the executable part of the ontology: answers can be answered.

There are several notations for the top level ontologies (e.g. CYC or Conceptual Graphs [4]) however there is general consensus on a certain number of points [5]:

- There are objects in the world
- Objects have properties/attributes that can take values
- Objects can exist in various relations with each other
- Properties and relations can change over time
- There are events that occur at different time instants
- There are processes in which objects participate and that occur over time
- The world and its objects can be in different states
- Objects can have parts

We will return of this points later during the presentation of where our work fits.

### 3. The nature of the problem

By perception and by computation or measurements, as indicated above, is the way humans carry on everyday activities. Perception is very imprecise, which explains the reason why constructing a knowledge base using rule-based system can be so complex (if possible at all). However, even if perception is very imprecise, everybody is able to carry on an activity and to make intelligent decisions in everyday situations.

We do not believe that human spend time to solve equations or to build probability matrix when they decide, for instance, to stop for a break and have a coffee. Our observation is that the problem is solved because anyone has learnt a certain number of things and that her/his cabling is done in that certain specific way. Our problem is in fact a series of subtasks:

1. The question that we are here to ask is: „can we explain what the criteria are that decided us to make a certain move?“
2. If we could explain that particular process, then the next step would be to find if we can find a mechanism to capture and represent that process in such a way that a machine could perform the same task.
3. If we can model and represent the process then the next question we need to ask and try to answer is if multiple individuals can get consensus about different pieces of knowledge or, simply, average the knowledge and create bigger sets of knowledge so that the whole is bigger than any single part.

### 4. Artificial Intelligence (AI)

In the past years there have been multiple and different approaches to the representation of large

chunks of knowledge, the guarantee of their consistency and the effective use of such knowledge. All that is part of the AI (Artificial Intelligence) domain. We need to provide a common understanding of the main points of the different approaches. Let's quickly review some basic concepts.

#### 4.1. Expert Systems (ES)

For these purposes Expert Systems were proposed in AI to manage knowledge processing. An expert system is simply a program that, on a narrow domain, performs tasks normally done by human experts. It is based on the idea of “eliciting” or “extracting” knowledge from specialists in a domain, express it in some representation and obtain a result that is similar to what the human expert would have reached.

One of the more complex tasks with expert systems is not only to build (elicit and represent) this knowledge but also to have agreement of different experts on the same topic.

In order to mediate and acquire the knowledge, the figure of the Knowledge Engineer (KE) is in place. The KE spends a lot of time and effort to build and debug the knowledge that is normally stored in what's called Knowledge Base (KB).

Many different ways of representing knowledge are in use. Traditionally expert systems generally employ the so-called IF-THEN rules to represent this information. An example that is very often given is the now famous MYCIN system [8] that diagnosis microbial diseases of blood. MYCIN rules have normally the following form:

```
IF <condition> THEN <conclusion> (<confidence>)
```

The *condition* is typically a logical expression which links some variables whose value can be inferred, measured or entered by the user. The *conclusion* determines the new value of some variables and the probabilistic nature of the rule is captured by the *confidence* factor in the rule.

Behind this approach there is the belief that is possible to understand and represent knowledge by simply translating it into the appropriate language or formalism. If we take the example of a black box, an expert pretends to build a black box, by specifying exactly the rules that, internally to the black box, determine how system inputs related to system outputs.

Another big problem with the expert system approach is that these knowledge representations are similar to „trees” with „jumps”. Hence the major issues:

- complexity to build them
- dramatic degradation of the results when operated with uncomplete knowledge or on the borderline

- very small tolerance to the errors (both in the representation and in the acquisition of the data).

## 4.2. Neural Networks (NN)

Neural Networks are computational models that have been inspired by neurophysiology. They have been introduced as Connectionist Expert Systems by Gallant [9] and the knowledge base is implemented by a neural network (instead than by a rule-based system).

A neural network consists of a multitude of nodes (simple units) called *neurons*. The neurons are densely interconnected and the interconnection carries a numerical weight that is used to „store” the knowledge.

So if the neuron  $j$  collect its real inputs from the outputs  $y_i$  where  $i$  are all the incident neurons connected to  $j$ .

Typically the connections between  $j$  and the  $i$ s are indicated by the real number  $w_{ji}$ . If  $w_{ji}$  denotes the *bias* of the neuron  $j$  when the input  $y_0$  and is 1, then the so called excitation level  $\xi_j$  of the neuron  $j$  is computed as the weighted sum of its inputs:

$$\xi_j = w_{j0} + \sum_i w_{ji} y_i \quad (1)$$

The *state* (i.e. the output) is determined from its excitation level  $\xi_j$  by applying an *activation function*  $\sigma$  as follows:

$$y_j = \sigma(\xi_j) \quad (2)$$

where:

$$\sigma(\xi) = \begin{cases} -1 & \text{for } \xi < 0 \\ 1 & \text{for } \xi \geq 0 \end{cases} \quad (3)$$

Starting from the above idea multiple architecture have been derived (e.g. feedforward networks[10]).

Clearly the function:

$$\vec{y}(\vec{w}) : \mathfrak{R}^n \rightarrow \mathfrak{R}^m \quad (4)$$

computed by the neural network is parametrized by the vector  $\vec{w}$  of all its weights. Neural networks learn this function from data provided as examples (training patterns). Learning can be done in a *supervised* or *unsupervised* manner. Details can be found in [10].

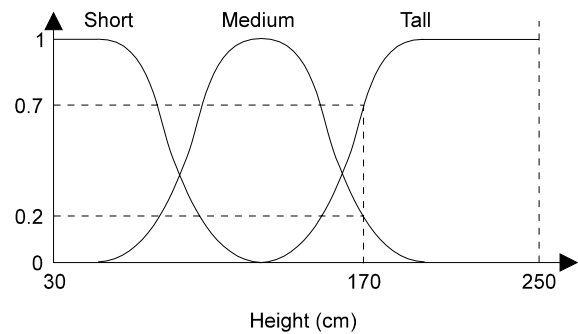
The capability of a neural network to memorize information depends upon the number of nodes and on the architecture (topology).

Taking again the analogy of a black box, it appears that in the neural network case, there is no desire to model the black box internal rules. On the other hand, what is of interest is the mapping of inputs with outputs without any interest to the rules that are inside the black box. This means that no artificial rules are elaborated (in conjunction with an expert) and the network simply adapts to the examples presented to it during the training period.

While neural networks looks better from the perspective of error degradation, they have the similar difficulties as rule-based systems in building the knowledge base. In fact, in the neural networks case, a meaningful training set is required. The main advantage is that the knowledge base is automatically created.

## 4.3. Fuzzy Systems

Since 1965 Zadeh paper on Fuzzy Sets [11], the basic ideas to model the uncertainty of the Natural Language were set. A fuzzy set has a graphical description that expresses how the transition from one to another takes place. This graphical description is called a membership function.



Membership function for the concept of „height” is presented below:

Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth (i.e. truth values between "completely true" and "completely false").

According to Zadeh’s „Extension Principle” instead than regarding fuzzy theory as a single theory, we should regard the process of “fuzzification” as a methodology to generalize ANY specific theory from a crisp (discrete) to a continuous (fuzzy) form. Thus recently researchers have also introduced "fuzzy calculus", "fuzzy differential equations", and so on (see [12]).

Fuzziness introduce the new idea, common in the human behaviours, that concepts flows *gradually* from membership to non membership because most knowledge in the real world is not measurement-based (i.e.: „it is 5:34 o’clock”, „there are 24 students”,...) but is perception based (i.e.: „it is almost 6 o’clock”, „there about 20 students”,...).

## 4.4. Fuzzy Cognitive Maps (FCM)

Fuzzy Cognitive Maps has been introduced by Bart Kosko in 1986 [6][7]. However already in 1948, Tolman presents the key concept of the “cognitive maps” to describe complex topological memorizing behaviours in the rats [14]. In the Seventies, Axelrod describes the “cognitive maps” in the shape of directed, interconnected, bilevel-valued graphs, and uses them in

decision theory applied to the politico-economic field [15].

FCM and are considered as a combination of fuzzy logic and artificial neural networks. Since then some interest is grown around this idea but we have the feeling that fuzzy cognitive maps are used little in research, and nothing in the industry.

A FCM is a dynamical system and has forward chaining ability only. It can answer the question “What’s happen if...?”, but not the question “Why...?” because of non-linearity. FCM help the prediction of the evolution of a system (behavior simulation) and can be augmented with capacities of hebbian learning as proposed by Kosko and Dickerson [16].

It is interesting to note an important difference between a Fuzzy Cognitive Map and a Neural Network: all the nodes of a Fuzzy Cognitive Map graph have a strong semantic that is defined by the modeling of the concepts. On the contrary input and output nodes of the graph of a NN have a weak semantic, only defined by mathematical relations.

Fuzzy Cognitive Maps are fuzzy signed directed graphs with feedbacks, and they model the world as a collection of causal concepts and relations between concepts. The relationships are directed arcs between the nodes. Each arc has a weight that defines the type of causal relationship between the nodes, i.e. positive or negative causal relation between the two concepts-nodes. Nodes stand for fuzzy sets and arcs stand for fuzzy rules.

Nodes are named by concepts forming the set of concepts  $C = \{C_1, \dots, C_n\}$ . Arcs or edges  $(C_i, C_j)$  are oriented and represent causal links between concepts; i.e. how concept  $C_i$  causes concept  $C_j$ . Weights of the arcs are associated with the link matrix and  $L_{ij} \in M_n(K)$  where  $K$  is  $Z$  or  $R$ ; if  $(C_i, C_j) \notin A$  then  $L_{ij} = 0$  else excitation (vs. inhibition) link from concept  $C_i$  to concept  $C_j$  gives  $L_{ij} > 0$  (vs.  $L_{ij} < 0$ ). So, a FCM with  $n$  nodes has  $n^2$  edges and because nodes are fuzzy sets they can take values  $[0,1]$ . So the state of a FCM at an „instant”  $t$  is a point in a fuzzy hypercube represented by the vector  $C(t) = (C_1(t), \dots, C_n(t))$  and the path from a state to another is a trajectory in the same hypercube. Given non-linearities we can have three possible outcomes: attraction to a point, to a stable cicle and to a chaotic attractor.

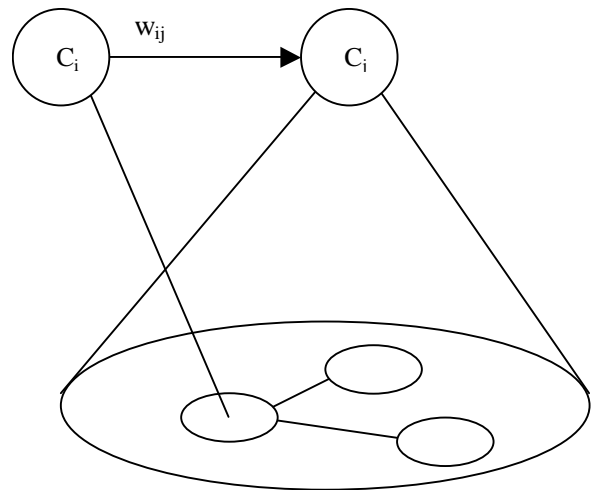
## 5. System Description

In our approach to the problem of representing an ontology makes the assumption that domain knowledge can be represented by two dimensions:

- Theoretical. It represents the concept in its abstraction and the (causal) links with other concepts.
- Practical. It represents all the information associate to a concept. It is a cluster of data.

We say that a person is knowledgeable when he has mastership over the two dimensions.

By taking advantage of Fuzzy Cognitive Maps representation we map Theoretical knowledge into the nodes (concept) . Concepts can be recursive and have links that go beyond the FCM they belong to. In this sense we introduce the concept of Fuzzy Cognitive Hyperspace.



Documents (e.g. Multimedia data) are part of each node and represent what somebody knows about the concept  $C_i$ .

Knowledge elicitation is done by building spaces (or hyperspaces) of nodes in a declarative way. The mechanism behind fuzzy cognitive maps will take care of averaging these sets of knowledge dumps.

## 6. Conclusions and Future work

We presented the ongoing work at the Mobile and Wireless Center of Expertise in Warsaw on Automatic construction of Fuzzy Ontologies.

The main assumption we made is that our ontology can be represented as a set of concepts and causal links, hence mapping into the FCM approach.

The challenge that is high in our list of things to do is the automatic creation of these nodes from the corpus of associated data.

## BIBLIOGRAPHY

- [1] Riichiro Mizoguchi, "A Step Towards Ontological Engineering", 12th National Conference on AI of JSAI, pp.24-31, June, 1998
- [2] D.Lenat and RV.Guha  
<http://www.cyc.com/public.html>
- [3] T.R.Gruber. "A translation approach to portable ontology specifications". Knowledge Acquisition, 5(2), 1993.
- [4] J.Sowa, "Conceptual Graphs"
- [5] B.H.Far, "Advanced Lectures on Knowledge Engineering – ontology engineering", Saitama University, 2000
- [6] B.Kosko. Fuzzy cognitive maps. International Journal of Man-Machine Studies, 24(1986) 65.75.
- [7] B.Kosko, "Neural Networks and Fuzzy Systems", Prentice-Hall, Englewood Clis, 1992.
- [8] B.Buchanan, E.Shortliffe, "Rule-based Expert systems", 1984, Reading, MA: Addison-Wesley
- [9] S.I.Gallant, "Connectionist Expert Systems", Communications of the ACM, 31(2):152-169
- [10] D.E.Rumelhart, G.E.Hinton, R.J.Williams, "Parallel Distributed Processing: Explorations in the Microstructure of the Cognition", 1986, Cambridge, MA, The MIT Press.
- [11] L.A.Zadeh, "Fuzzy sets", Inf. Control 8, 338-353, 1965.
- [12] L.A.Zadeh, "The Calculus of Fuzzy Restrictions", in Fuzzy Sets and Applications to Cognitive and Decision Making Processes, edited by L.A.Zadeh et. al., Academic Press, New York, 1975, pages 1-39.
- [13] B.Kosko, "Fuzzy Thinking", Flamingo, 1994.
- [14] E.C.Tolman, "Cognitive Maps in Rats and Men", Psychological Review, 42, 55, 189-208, 1948.
- [15] Axelrod R., Structure of Decision, Princeton University Press, Princeton, New Jersey, 1976.
- [16] J.A.Dickerson, B.Kosko B., Virtual Worlds as Fuzzy Cognitive Maps, Presence, 3(2):173-189, MIT Press, 1994.