Statement of Interest

Amedeo Cappelli, Lorenzo Moretti
Knowledge Representation Group
Istituto di Linguistica Computazionale - CNR
Via della Faggiola, 32 56100 PISA Italy
Tel. 39 50 560481
Sistemi@ICNUCEVM.Bitnet

1989

The group has been working on KR since 1982, and is at present collaborating in the following projects: "Strategic Project for Artificial Intelligence" of the Italian National Research Council (CNR), area on Knowledge Representation. "Strategic Project for Natural Language Processing" of the Italian National Research Council (CNR), area on Knowledge-based Linguistic Tools. "Progetto Finalizzato Informatica" of the Italian National Research Council (CNR), area on New Tools for Expert Systems.

The main interests of the group are focussed on knowledge representation formalisms in the tradition of KL-ONE. Certain versions of "classical" KL-ONE have been implemented. Some theoretical issues concerning SI-Nets have been investigated both in the framework of general formalism to represent knowledge and in that of their relationships with natural language processing. In particular, attention has been devoted to the creation of natural and intuitive devices for interaction with an SI-Net KB. At present our aim is to study the relationship between epistemological formalisms and ontological phenomena. In particular, we are trying to build the tools suitable to make it possible to specify ontological theories in an SI-Net formalism. These issues are investigated in the perspective of the acquisition of knowledge from human operators or from structured knowledge sources such as dictionaries or encyclopedias (this goal is pursued in the framework of the
Italian "Progetto Finalizzato" concerning the design of new tools for expert systems). The main topics investigated by the group are described in the following.

**KL-Magma:**

Implementation of a knowledge representation language (KL-Magma) based on SI-Nets, which is similar to "classical" KL-ONE; the following versions are available:

- in Magma-Lisp on IBM mainframes;
- in Prolog2 on MS-DOS PCs;
- in C-Prolog on VAX and SUN;
- in NIP (New Implementation of PROLOG) on SUN;
- in LPA-MacProlog on Macintosh;
- in Allegro Common LISP on Macintosh.

The implementation in LPA MacProlog, which runs on Macintosh (PLUS, SE and II), is our last version; we give a brief description of it. KL-Magma allows the user to create and to use knowledge bases (KB) represented in the form of Structured Inheritance Networks (SI-Nets). The KB is accessed by built-in predicates which are described in: A. Cappelli, L. Moretti, KL-Magma: Reference Manual, Pisa, 1988. The user can interact with the system via Macintosh menus: an editor suggests the actions to be performed in order to build knowledge structures which are correct from the point of view of the SI-Net syntax. A classification mechanism is also available. The results of built-in predicates are visualized in a default window in the LPA MacProlog syntax. The user can access data by using special windows where data are shared on the basis of SI-Net data structures - concepts, roles, facets of roles, cables and wires, structural descriptions, and Hooks. A graphical editor is also provided by which the user interacts with the system by drawing on the screen bit-map ovals, squares and arrows. The graphical editor is a powerful support to the menu device previously mentioned.

**MetaDB:**

2
KL-Magma interacts with MetaDB, a tool which enables an SI-Net KB to interact with a relational DB. It consists in an interface which translates the data structures of a relational DB into those of an SI-Net KB. An abstract representation of relational structures has been realized in KL-Magma which controls the mapping between the two different environments.

KL-Conc:

KL-Conc is an experimental KR language which makes it possible to interact with SI-Nets in a non-structure-oriented way. The user is thus not required to have any knowledge of the SI-Net model; he uses quantification and association functions such as anyone, justone, many, test-configuration-of-properties, which are interpreted in a knowledge base in the form of SI-Nets. These functions are designed taking into account the means by which natural language conceptually manipulates real world objects, so that the user can create or retrieve objects from an SI-Net KB, by using his linguistic/conceptual intuition (the abstraction level is that of the linguistic manipulation of real world objects). A prototype of KL-Conc has been implemented in Magma-Lisp on IBM mainframes and in LPA MacProlog for Macintosh.

Chunking System:

This is a mechanism which keeps trace of all the paths followed through the network when association functions of KL-Conc are triggered. Imagine that you have a function (Test-Configuration-of-Properties x y z) which searches in the network to check whether a relationship between x and z exists, optionally passing through y, and returning the piece of knowledge connecting x to z. This function allows an 'intuitive' behavior since it makes it possible to declare a 'vague' association between two or three terms without specifying their "deep" relationships. For instance: (Test-Configuration-of-Properties house color) can correspond to the relationships: (house...base...color) or (house...front...color). If no "semantic" links are specified, all paths between the terms of the association have to be returned (from a natural language processing viewpoint, this roughly corresponds to accounting for ambiguity between components of a NP). This is a time-consuming task when performed on a large knowledge base where there may be more than one conceptual path between entities,
each of which can cover a different conceptual process. However, time-
consuming is drastically reduced when partial (or entire) paths between them 
have already been followed and stored in a working memory (WM). In this 
case, tasks can be immediately satisfied without searching in the network. 
Paths followed between entities are stored in a WM as chunks. As a conse-
quence, we obtain a partial reconfiguration of the network into chunks, where 
every element of a chunk can be an access point to the SI-Net (considered as 
the Long Term Memory). So, satisfying a goal means searching through the 
two memory areas in accordance with the following procedure:

\[
\text{search}(<x>, y, z) :- \text{search-on-chunks}(x, y, z), !; 
\text{search-on-LTM}(x, y, z), !. 
\]

\[
\text{search-on-LTM}(x, x, <x>). 
\]

\[
\text{search-on-LTM}(x, y, z) :- \text{successors}(x, s), 
\text{search}(s, y, z1), z=<x;z1>, \text{store-new-chunks}(x, z). 
\]

Furthermore, it makes it possible to increase the interface capabilities of 
KL-Conc: experiences which have been stored can be considered as "preferences" and can be used to solve conflicts during the navigation through 
the network. The chunks chosen by the user (which are the conceptual paths 
corresponding to his "vague" associations) are given a preference in the WM: 
they are retrieved first from the WM; they constitute the context to solve 
further goals characterized by a sort of ambiguity. An experiment carried 
out on a knowledge base of about 250 concepts (80 generic and the rest indi-
vidual) concerning the World Cup Football Championship in Mexico gave 
an increase in the task performance in accordance with an exponential law 
\( T = B N - a \). We have tested the system taking into account all possible 
questions which can be expressed via complex associations such as, for in-
stance: "How many goals did Maradona score in the final", "Who was the 
referee of the final" etc. The following versions of Chunking are available: 1) 
in Magma Lisp on IBM mainframes; 2) in LPA Prolog on Macintosh; 3) in 
AllegroCL on Macintosh.

**SI-Nets and Natural Quantification:**

SI-Nets formalism has been used to describe natural quantification. In 
particular, an experimental system has been implemented which allows the 
user to create and retrieve knowledge stored in two different functional areas 
(an SI-Net KB and a DB) by using natural quantifiers. The KB contains
the generic domain knowledge while the DB contains all information relative to individuals. Questions can return objects or truth values interpreting the two areas. A model of a question-answer interface has been designed which specifies a model of the question (containing natural quantifiers), a model of the answer and a mapping between the two. Types of questions are mapped onto types of answers: every type is semantically described in terms of the operations it triggers in the two areas and in terms of objects it returns from them. The system is implemented in NIP Prolog on SUN, using INFORMIX for the representation of the DB.

SI-Nets and Natural Language Processing:

SI-Nets have been the reference model of many linguistic studies concerning the cognitive representation of the lexicon, the relationship between an ATN parser and a domain knowledge in the form of a semantic network, and the acquisition of lexical knowledge. A system has been implemented which translates meteorological bulletins into KL- Magma. The system is composed of a parser written in Prolog and a resident KB where descriptions of meteorological facts and of the geography of the Mediterranean Sea are given. The mapping between the syntactic output and the KB is performed by continuously testing domain knowledge.

Lexical Knowledge Acquisition:

SI-Nets formalism has been used to model the behavior of a lexicographer in acquiring lexical definitions from dictionaries. KL-Magma is used to represent lexicographic definitions and also the abstract model by which lexicographers define linguistic terms. Types of definitions and parts of definitions receive their own descriptions in terms of concepts and roles. Lexicographers interact with the system via this abstract model which suggests operations to be done in order to create definitions which are correct with reference to the model. Let us imagine, for instance, that the model requires that every definition might be constituted by a genus and some differentia (which roughly correspond to superconcept and role in SI- Nets): the representation of such notions in terms of SI-Net objects makes it possible to use these specific concepts of lexicographers’ knowledge as ”natural” media of communication.

Knowledge Acquisition:
The group is now working on the framework of knowledge acquisition with the aim of specifying certain tools for guiding a human operator in the structuring of different types of domains. Several modules cooperate in helping the user to consult various knowledge bases (not necessarily entirely structured), and to structure the knowledge of a specific domain while linking it to other "naturally" related domains. The user navigates through domains by using a general representation of the universe, created on the basis of concepts used in dictionaries or thesauri for describing terms (about 2,000 concepts). This KB is linked to the terms of a dictionary. While the former can be intended as a general reference schema for linking different world, the latter can be considered as the source (partially structured), where one can get the basic information about the knowledge of any domain. Our hypothesis requires that both the general representation of the universe and that of any domain chosen by the user have to be represented in the form of an extended SI-Nets formalism. This extension requires the inclusion of ontological descriptions in the terminology thus allowing interpretations of knowledge in order to establish links between different worlds, as described in the following.

Epistemology and Ontology:

LOT (Theory-Oriented Language), is a knowledge representation language which combines epistemological and ontological aspects in a conceptual description. LOT extends a KR language based on Structured Inheritance Networks, by modifying classification, individuation and aggregation mechanisms. Classification and aggregation are restricted by using ontological constraints. These constraints do not necessarily depend on a unique ontological system but can belong to different systems interacting with one another. Such systems correspond to worlds or theories. By using several ontological theories, it is possible to create partitions of knowledge in order to increase the efficiency of the deductive system. A theory is the union between the structural aspects of the world objects (epistemology) and a set of rules for the organization of the world (ontology). Classification, individuation and aggregation are controlled by an ontological module - which gives the right weight and meanings to the construction of taxonomies. In particular, this means giving a meaning to the role from an ontological point of view so that properties relations or individuation principles are interpreted
on the basis of the principles deriving from the specific nature of certain
types of concepts. In such a "pluralistic" universe, the reasoning processes
must i) manage pieces of knowledge inside one specific ontological theory or
ii) relate ontologies. i) can be seen as one of the declarative aspects of the
classical interpretation of viewpoints and of the relevance/irrelevance theory.
The introduction of such processes opens new perspectives in the computa-
tion of knowledge: one can for instance design partial evaluation mechanisms
which use partitions of a KB into theories in order to disregard the irrelevant
aspects of a situation by excluding them from the current evaluation. ii) im-
plies overcoming a restricted ontological aspect and putting together pieces
of different ontological modules. This allows us to create some associations
that are not strictly connected with principles deriving from first-order logic.
LOT is based on the KLONE grammar with some additional functionalities.
It makes it possible, first of all, to embed the contexts or the theories into the
epistemological structures, and secondly, to use ontological modules in order
to manage them. In other words, the epistemological structure becomes the
data structure used by the real interpreter of the system, i.e. the ontological
module. While the first functionality may be reduced to a syntactic structu-
turing mechanism, the ontological modules cannot be reduced to a first order
logic system but each module is composed of assertions derived from semantic (conceptual) aspects of knowledge. For instance: (1) if x is no longer a
dog then x does not exist (2) if x is no longer a puppy then x is a dog are
two simple rules which model the concept "dog" according to the theory of
substance-concept. These rules are an example of the kind of assertions to be
introduced into an ontological module. The general architecture is composed
as follows: a module which covers the syntactic-epistemological functionalities; several ontological modules, each of which contains sets of rules for the
interpretation of the entities in the world; an Intermediate Inferential Module
which controls the communications between ontological modules and, among
others, contains the classification and the chunking mechanisms used by the
different theories.

List of Publications:

A. Cappelli, L. Moretti, Introduzione al KLONE, ILC- NLP-1982-1, Pisa,
1982.

S.A. Cerri, A. Cappelli, A Review of current Artificial Intelligence Re-
search in Italy, AISB Quaterly, 44 (1982).


A. Zampolli, A. Cappelli (eds.), The Possibilities and Limits of the Computer in Producing and Publishing Dictionaries: Proceedings of the European Science Foundation Workshop, Pisa 1981,


A.


A.


A.M. Bianchimani, Manuale TPY:KL-Magma: Shell di Tipicalità per