Classification in the KL-ONE Knowledge Representation System

James G. Schmolze

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, MA 02238 • USA

Thomas A. Lipkis

USC/Information Sciences Institute
4676 Admiralty Way
Marina Del Rey, CA 90291 - USA

Abstract

KL-ONE lets one define and use a class of descriptive terms called Concepts, where each Concept denotes a set of objects A subsumption relation between Concepts is defined which is related to set inclusion by way of a semantics for Concepts. This subsumption relation defines a partial order on Concepts, and KL-ONE organizes all Concepts into a *taxonomy* that reflects this partial order. Classification is a process that takes a new Concept and determines other Concepts that either subsume it or that it subsumes, thereby determining the location for the new Concept within a given taxonomy. We discuss these issues and demonstrate some uses of the classification algorithm.

KL-ONE is a knowledge representation system developed at Bolt Beranek and Newman over the past few years (see [Brachman *77,* Brachman 79, Schmolze *82.* Sidner 81]), that grew out of semantic network formalisms. The primary unit of information in KL-ONE is called a Concept, which denotes a set of objects. A Concept has a set of (syntactic) components, each denoting a property that must be true of each member of the set denoted by the Concept.[2] In particular, one type of component is a Roleset which is analogous to a "slot"' in a "frame-like" language. For example, we can construct a Concept denoting the set of all people where each person has a birth-date:

PERSON is a Concept and has a Roleset Birthdate [3]

Unlike early semantic networks, domain dependent relations are not represented as links, but as Concepts and Rolesets There is but a small number of types of links in KL-ONE, each corresponding to a Concept-forming or Roleset-forming operator. Due to space limitations, we cannot describe the language in detail and instead will offer examples of KL-ONE Concepts. We also sketch a semantics for KL-ONE The interested reader is referred to [Brachman 79] and [Schmolze 32].

This research was supported in part by the Defense Advanced Research Projects Agency under contracts N00014-77-C-0378 at BBN and MDA903-81 C-0335 at ISI Views and conclusions contained in this paper are the authors' and should not be interpreted as representing the official opinion or policy of DARPA. the U S Government or any person or agency connected with them

[2] We extend the usual notion of property to include situations where an object must stand in some relation to another object

In this paper, we will name Concepts after the elements of the set they denote All Concept names will be in upper-case, bold-face letters All Roleset names will be capitalized and bold-face

We define a relation between Concepts called subsumption such that Concept A subsumes Concept B only if the set denoted by A necessarily includes the set denoted by B. For example, if one wants to represent that all dogs are necessarily mammals then one specifies that the Concept denoting mammals subsumes the Concept denoting dogs. In a KL-ONE network.[4] we represent subsumption by a link between Concepts called (for historical reasons) SuperConcept

Since set inclusion is transitive, reflexive and anti-symmetric, we define subsumption correspondingly such that it determines a partial order on Concepts KL-ONE maintains a taxonomy of Concepts that explicitly represents this partial order. The taxonomy has a top; Concept named THING, which is defined to subsume every other Concept However, there is no corresponding bottom Concept

Subsumption is actually an out-growth of the "ISA" link of early semantic networks, and as with "ISA". Concepts inherit components from their subsumers By this we mean that each member of the set denoted by a Concept has the properties that are specified by the components of the Concepts subsumers (as well as the components of the Concept itself). Thus, our DOG Concept inherits all of the components 6f MAMMAL, which means that each dog has all of the properties specified for mammals

In our example, we stated explicitly that mammals necessarily include dogs However, there are cases where one can deduce that the set denoted by some Concept necessarily includes the set denoted by a second Concept, but where no subsumption relation between the Concepts was explicitly entered. Classification is a process that discovers these latter subsumption relations between Concepts and, in such cases, establishes SuperConcept links between them. Thus, we say that classification automatically places a Concept at its proper location in a KL-ONE taxonomy because it establishes the appropriate SuperConcept links between it and other Concepts

Some Uses of Classification

Many AI programs use taxonomically structured knowledge bases for modeling dynamic environments, and therefore require automatic classification of new knowledge as it is obtained. Automatic classification also provides a means of enforcing network semantics and checking consistency of descriptions, and is therefore a superior alternative to manual construction of static taxonomies.

We will use the words "network" and "taxonomy" interchangeably with respect to KL-ONE

Classification is also useful for generalized search. If one forms a search pattern into a Concept (call it PATTERN), classification will discover other Concepts that PATTERN subsumes. If the target of the search is also described by some Concept (call it TARGET), and if the pattern matches the target, then PATTERN will subsume TARGET. Hence, the first phase of a search process will be accomplished by using classification to restrict the search space of possible target descriptions.

The KL-ONE Language and a Semantics via some Examples

We now introduce the KL-ONE language and a semantics for it. For pedagogical reasons, and because of space limitations, we do so via simple examples. Note that this demonstrates only a fraction of the KL-ONE language.

In order to complete our example, we re-specify the Concept PERSON such that PERSON is a primitive Concept denoting the set of all persons. Primitive Concepts are interpreted as having essentially incomplete definitions, and thus, all Concepts denoting "natural kinds" (e.g., people, elephants, chairs) are primitive.

PERSON is a primitive Concept, is subsumed by MAMMAL, and has a Roleset Birthdate with:
a number restriction of exactly one. and
a value description of DATE.

By this specification, we mean that each person is a mammal and has exactly one birth-date that must be a date (we let MAMMAL and DATE be primitive Concepts). When specifying that a Roleset is a component of a Concept, one must also state:

- the number restriction for the Roleset, which specifies constraints on the number of fillers,
- and the value description, which specifies constraints on the type of each filler.

Non-primitive Concepts are interpreted as being completely defined. An example is PARENT.

PARENT is a non-primitive Concept, is subsumed by PERSON, and has a Roleset Child with:
a number restriction of one or more, and
a value description of PERSON.

This specification defines a parent to be a person who has at least one child who is a person.

We specify our semantics for KL-ONE Concepts by a mapping from KL-ONE Concepts into the language of informal set theory. Others have chosen a different specification (particularly in [Israel 82]), but ours suffices for demonstrating the properties needed to explain the behavior of the classifier. As stated earlier, each Concept denotes a set. Each Roleset denotes a set of ordered pairs whose domain and co-domain include the sets denoted by the Concept with which the Roleset is associated and the value description, respectively.

The set denoted by PARENT has been defined to be just:

{x | x is a person and
(By) (y is a child of x and y is a person)}

Our last example is that of GRANDPARENT.

GRANDPARENT is a non-primitive Concept, is subsumed by PERSON, and has a Roleset Child with:
a number restriction of one or more, and
a value description of PARENT.

This defines a grandparent as a person who has at least one child who is a parent. The set denoted by GRANDPARENT is defined to be:

{x | x is a person and
(By) (y is a child of x and y is a parent)}

Note that the specification of GRANDPARENT did not include PARENT as a subsumer. however, we can deduce that every grandparent must also be a parent. Thus, we have discovered an additional subsumer of GRANDPARENT, namely. PARENT. This is a simple example of a deduction made during classification

The Algorithm for Classification

The primary component of the classifier is a function, called SubsumesP, that compares two Concepts and decides whether or not the first subsumes the second. Since a KL-ONE taxonomy is organized by the partial ordering of Concepts under subsumption, the classifier can find the proper location for a Concept by using SubsumesP to compare it to all other Concepts in the taxonomy, and deduce subsumption relationships (the issue of completeness is discussed in the next section). Actually, it usually needs to search only a small fraction of the taxonomy.

We only have space to sketch the portions of the SubsumesP algorithm that apply to our example. A complete discussion of the classification algorithm is presented in [Lipkis 82].

Given Concepts A and B. we wish to know whether A subsumes B. Our test performs a piece-by-piece comparison of the components of A with those of B. including inherited components. As an example, we will show the results of testing SubsumesP(PARENT,GRANDPARENT) SubsumesP(A.B) is true if and only if:

- All primitive Concepts that subsume A also subsume B (Both PERSON and GRANDPARENT have the same primitive subsumers, namely, PERSON and MAMMAL.)
- For each Roleset of A, some Roleset of B denotes the same relation (both PARENT and GRANDPARENT have just 2 Rolesets, Birthdate and Child), and for those corresponding Rolesets:
    * The number restriction for As Roleset includes that of B's. (For the Birthdate Roleset, both PARENT and GRANDPARENT have the same number restriction, namely, exactly one. The same applies to Child, namely, each has a least one.)
    * The value description of As Roleset subsumes that of B's. (For the Birthdate Roleset. both PARENT and GRANDPARENT have the same value description, namely. DATE. For Child. PARENT'S value description subsumes GRANDPARENT'S, *i.e.*. PERSON subsumes PARENT.)

Thus, SubsumesP(PARENT,GRANDPARENT) is true.

Properties of SubsumesP

We are interested in whether SubsumesP has the following properties, where A and B are Concepts:

- soundness: SubsumesP(A.B) implies that the set denoted by A necessarily includes the set denoted by B
- completeness: That the set denoted by A necessarily includes the set denoted by B implies that SubsumesP(A.B).
- totality: SubsumesP always terminates (and returns either TRUE or FAIL).

If we could show that SubsumesP satisfied these properties, then Concept subsumption would be wholly decidable In other words, we could then state that if SubsumesP(A.B) returns TRUE, then As set necessarily includes B s set. and if SubsumesP(A.B) returns FAIL, then As set does not necessarily include B s set Thus. FAIL would be equivalent to false.

For the current implementation, we have informally shown (though not here) that SubsumesP is both sound and total However, we have not shown it to be complete Thus, if SubsumesP(A.B) returns FAIL, we are not certain whether or not A subsumes B In our studies, however, we have not discovered any case where SubsumesP(A.B) will return FAIL when, in fact, TRUE would have been correct *(i.e..* As set necessarily included B's set). In our application programs, we have treated FAIL as equivalent to false without any ill effects. In the near future, we hope to show that SubsumesP actually is complete and also to show soundness and totality formally

Certain characteristics of KL-ONE allow for the possibility of such a decision procedure For example, cancellation, for which no clear semantic account has been offered (see [Brachman 80. Israel 81]), is not allowed. Also, the language of KL-ONE has less expressive power than a standard first-order language, where no decision procedure exists Moreover, for this first phase of our research on the completeness of SubsumesP we have not studied cycles, *i.e..* sets of concepts that are specified in terms of each other. (Note that the language of KL-ONE allows cyclic specifications and we believe that the classification algorithm, which is not completely described in this paper, finds appropriate subsumption relations involving such concepts.)

Conclusion

The KL-ONE system and the classifier have been used for several years as representational and inferential components of several artificial intelligence systems (see [Brachman 79, Sidner 81, Mark 81. Zdybel 81]). Recently, we have attempted a formal specification of the semantics of KL-ONE, and of the relation of the classifier to these semantics. The classifier determines subsumption relationships between Concepts, where subsumption denotes necessary set inclusion between the sets denoted by the Concepts. We are confident that the classifier is sound and totally defined, and we hope to show it is complete.

Similar work that combines a knowledge representation system with a decision procedure is being done by Brachman. Fikes and Levesque [Brachman, Fikes and Levesque 83] with their KRYPTON system. We are currently re-designing the KL-ONE system, and to some extent, the language of KL-ONE, An important consideration has been the decidability of classification. In fact, our decision to include certain new language constructs hinges on showing that they do not preclude decidability.

Acknowledgments

There are many other contributors to this research, including R.J. Brachman. R.J. Bobrow, R. Fikes. D.J. Israel. H. Levesque. W. Mark, and W.A. Woods. Many, many thanks go to David Israel for his efforts and ideas regarding this paper.

References

[Brachman 77] Brachman, R.J., *A Structural Paradigm for Representing Knowledge,* Ph.D. thesis. Harvard University, May 1977 Also. BBN Report No.3605, Bolt Beranek and Newman Inc., May 1978

[Brachman 79] Brachman, R.J., Bobrow, R.J., Cohen. P.R., Klovstad, J.W., Webber. B.L., Woods, W.A., *Research in Natural Language Understanding - Annual Report: 1 Sept 78 - 31 Aug 79,* Bolt Beranek and Newman Inc.. BBN Report No. 4274, August 1979.

[Brachman 80] Brachman. R.J, I Lied About the Trees 1980. Unpublished manuscript

[Brachman. Fikes and Levesque 83] Brachman. R.J.. Fikes. R.E., and Levesque. H.J.. KRYPTON. A Functional Approach to Knowledge Representation, 1983 (submitted to IEEE Computer special issue on knowledge representation - 1983)

(Israel 81] Israel. D.J.. and Brachman. R.J., "Distinctions and Confusions. A Catalogue Raisonne." in *Proc UCAl-81.* pp. 452-459, Vancouver, British Columbia. 1981.

[Israel 82] Israel. D.J., *On Interpreting Semantic Network Formalisms.* Bolt Beranek and Newman Inc.. technical report 5117. September 1982. (Also to appear in a 1983 issue of International Journal of Computers and Mathematics.)

[Lipkis82] Lipkis. T.A., "A KL-ONE Classifier," in *Proceedings of the 1981 KL-ONE Workshop,* pp. 128-145, BBN. June 1982.

[Mark 81] Mark. W., "Representation and Inference in the Consul System." in *Proc. UCAl-81,* pp. 375-381, Vancouver, British Columbia, August 1981.

[Schmolze 82] Schmolze, J.G., *Proceedings of the 1981 KL-ONE Workshop.* Bolt Beranek and Newman Inc.. BBN Report No. 4842. June 1982.

[Sidner 81] Sidner, C.L.. Bates. M., Bobrow, R.J., Brachman, R.J.. Cohen, P.R., Israel. D., Schmolze, J., Webber, B.L. and Woods, W.A., *Research in Knowledge Representation for Natural Language Understanding, Annual Report: 1 September 1980 - 3 August 1981.* Bolt Beranek and Newman Inc., Cambridge, MA, BBN Report No. 4785, 1981.

[Zdybel 81] Zdybel, F., Greenfeld, N., Yonke, M., Gibbons, J., "An Information Presentation System," in *Proc. UCAl-81,* pp. 978-984, Vancouver, British Columbia, August 1981.