Dr. Robert I. Young, DSTL, Malvern, U.K.
and
Dr. Margaret J. Varga, QinetiQ, Malvern, U.K.

Knowledge Representation in Intelligence Applications

Knowledge representation is one of the central and in some ways the most familiar concept in artificial intelligence. There are many different definitions or interpretation of knowledge representation. Davis et al. describe five functions that knowledge representation performs.

1. A knowledge representation (KR) is most fundamentally a surrogate, a substitute for the thing itself, used to enable an entity to determine consequences by thinking rather than acting, i.e. by reasoning about the world rather than taking action in it.
2. It is a set of ontological commitments, i.e. an answer to the question: In what terms should I think about the world?
3. It is a fragmentary theory of intelligent reasoning, expressed in terms of three components: (i) the representation's fundamental conception of intelligent reasoning; (ii) the set of inferences the representation sanctions; and (iii) the set of inferences it recommends.
4. It is a medium for pragmatically efficient computation, i.e. the computational environment in which thinking is accomplished. One contribution to this pragmatic efficiency is supplied by the guidance a representation provides for organizing information so as to facilitate making the recommended inferences.
5. It is a medium of human expression, i.e. a language in which we say things about the world.

Each function serves different purposes and thus requires different properties to achieve the desired outcome, but together the five functions provide a framework useful for characterizing a wide variety of representations.

There are various knowledge representation techniques available, e.g. Bayesian belief nets, ontologies, graphs, case based reasoning, expert systems, neural networks, pattern recognition, AI planning, inductive learning, explanation based learning etc.

In recent years, the World Wide Web Consortium (W3C) have helped to re-use much research work on representing specialised domains. However, much remains unclear and unresolved. For example it is not clear how effectively the above techniques or their hybrids are able to manage an intelligence problem featuring the fusion of information contained within disparate datasets, nor is it clear how to evaluate this.

1. It is not clear whether there exist formal mappings of the KR techniques and algorithms onto such intelligence problem and tasks.
2. The practical effect of issues such as uncertainty, ignorance, data incest and data deluge is not clear, nor is the extent to which the techniques are able to manage imprecise scaling when observations are out of sequence or incomplete.
3. It is not clear exactly what benefits can be accrued; for example:
   - what can be automatically discovered in intelligence data;
   - can automated alerts be generated identifying events and activities together with confidence and priority measures?
   - can significant but sparse information be identified from large data sets?
   - can discovered data be effectively presented to and easily interrogated by the analyst, and in what form?
4. It is not clear whether there are metrics to evaluate the worth of the knowledge represented.
5. The complex nature of the intelligence problem implies the need for usability in related and different problems, but to what extent can this be achieved.

These are key questions for investigation, but so also are the more pragmatic issues of just what can be implemented with practical resources now and whether it is possible to handle the amount and variability of data in the form in which it is presently available.

Reference

[1] R. Davis, H. Shrobe, and P. Szolovits. What is a Knowledge Representation? *AI Magazine*, 14(1):17-33, 1993