KRF

Knowledge Representation Framework Project
Department of Computer and Information Science, Linköping University,
and Unit for Scientific Information and Learning, KTH, Stockholm

Erik Sandewall

The Structure of Artificial Intelligence Research

A Brief Introduction

This series contains technical reports and tutorial texts from the project on
the Knowledge Representation Framework (KRF).
The present report, PM-krf-023, can persistently be accessed as follows:

Project Memo URL:	http://www.ida.liu.se/ext/caisor/pm-archive/krf/023/
AIP (Article Index Page):	http://aip.name/se/Sandewall.Erik.-/2011/001/
Date of manuscript:	2011-01-04

Related information can also be obtained through the following www sites:

KRFwebsite:	http://www.ida.liu.se/ext/krf/
AIP naming scheme:	http://aip.name/info/
The author:	http://www.ida.liu.se/∼erisa/

Different people may have more or less different opinions about what are the subareas of Artificial Intelligence, and about which of these subareas are the most important ones or the most basic ones. The following represents my point of view, which I believe can safely be said to represent a strong majority opinion.

1 A list of sub-areas

As a starting point for the discussion I shall take the list of research areas in the call for papers of the 2011 AAAI conference (Association for the Advancement of Artificial Intelligence). It goes as follows.

```
A  Agent-based and multiagent systems
C  Cognitive modeling and human interaction
A  Commonsense reasoning
C  Computer vision
B  Constraint satisfaction, search, and optimization
C  Evolutionary computation
C  Game playing and interactive entertainment
C  Information retrieval, integration, and extraction
B  Knowledge acquisition and ontologies
A  Knowledge representation and reasoning
C  Machine learning and data mining
A  Model-based systems
C  Natural Language Processing
A  Planning and Scheduling
B  Probabilistic Reasoning
C  Robotics
C  Web and information systems
```

The list of topics for the International Joint Conference on Artificial Intelligence (IJCAI) is similar. - In this list I have marked the various areas with the letters A, B, and C, with the following meanings. The A and B groups include work with a similar methodology, based on core computer science (algorithms, etc), the use of formal logic, and applying these tools to the qualitative modelling of phenomena in the real world and reasoning based on the information in such models, as well as to the design of autonomous software agents that can make use of those models and that can perform the said type of reasoning.

The difference between the A and B groups is that the A-listed topics tend to be more basic, so that one should study them first before proceeding to the B-listed topic. [1] This borderline is not strict but it can serve as a guide.

The C-listed topics are by and large interdisciplinary, in the sense that each of them has strong connections to some other scientific discipline besides artificial intelligence, and in most cases outside computer science. The field

[1] One may ask why 'search' and 'optimization' are not listed in the most basic category, since these are mathematical techniques that are widely used in A.I. The answer is that from an A.I. point of view it is natural to first study the A listed topics in order to understand what the field is about, and then to proceed to the relatively more technical details of the algorithms being used. From a mathematical mindset one may prefer the reverse ordering.

of robotics contains a lot of work in mechanical engineering and in control engineering, for example; the field of natural language processing as seen from A.I. integrates seamlessly with work in computational linguistics and with the cognitive sciences in general, and so forth.

The cases of evolutionary computation and machine learning should also be mentioned separately. The actual literature that goes by the name of machine learning has a large component that makes heavy use of probability theory. This work is only remotely connected to the work in the A and B categories above. There are also some niche areas, such as case-based learning and inductive logic programming, that is well connected to the A and B categories, but these areas seem to be marginal in the contemporary research literature on machine learning.

With respect to evolutionary computation, there is a range of work that is represented at one end by the approach of 'artificial life', where one tries to develop computational life-forms using a simulated evolution process, and at the other end by work on generate-and-test techniques (i.e. systematic search techniques) that are applied to moderately complex objects that have a dynamic interpretation, for example simple scripts in a specialized language, or decision trees. In the former case the methodology is remote from what we have in the A and B listed areas; in the latter case there is a quite close connection.

This discussion implies that there is coherent point of view about artificial intelligence where, with simplification, the A listed areas are considered as the basic ones, the B listed areas are also included but are best studied in a second step, and where the C listed areas integrate work that is close to the A and B areas, with work that has an entirely different character.

Specifically, the C listed areas can be characterized in terms of three major groups, as follows

Bridge Areas

Most of the topics in the C category have the character of a bridge that combines Artificial Intelligence with one or more other disciplines. They form three natural groups, as follows.

Bridge to Robotics

This includes the following ones of the topics listed above:

```
A  Agent-based and multiagent systems
C  Computer vision
A  Knowledge representation and reasoning
C  Robotics
```

Bridge to Human-Machine Interaction

This includes:

```
A  Agent-based and multiagent systems
C  Cognitive modeling and human interaction
A  Commonsense reasoning
C  Game playing and interactive entertainment
A  Knowledge representation and reasoning
```

```
C  Natural Language Processing
```

Bridge to Large Knowledgebases and their Use

This includes:

```
C  Information retrieval, integration, and extraction
B  Knowledge acquisition and ontologies
A  Knowledge representation and reasoning
C  Web and information systems
```

2 Alternative Views of Artificial Intelligence

Our view of what Artificial Intelligence is about was described in the memo "The Goals of Artificial Intelligence Research. A Brief Introduction" which is part of the course literature. This memo begins with a discussion of concrete examples of the use of intelligence in small children and in adults, and takes it from there to an identification of what are the major problems and issues in A.I. This results in a view of A.I. where *agent intelligence* is emphasized, that is, one is concerned with the design of software systems that are able to *act intelligently* in some naturally occurring environment. This is the classical view of Artificial Intelligence, and it was expressed, for example, in the proposal for the design of an 'advice taker' by one of the field's pioneers, John McCarthy at Stanford University.

2.1 Systems with Specialized Intelligence

This classical view may be contrasted with approaches that work towards the design of quite *specialized varieties of intelligence.* One example of this is the work on programs that play specific games, for example chess-playing programs. These are sometimes seen as examples of artificial intelligence, in the sense that intelligence is considered as essential for being able to play chess well by humans. Today there exist advanced programs that play chess extremely well and that are able to match even the world's best human chess players. These programs are quite advanced and shall by no means be thought of as mere brute-force search programs. However, there are few connections between the techniques that are used in these programs and techniques that are of use A.I. in general: chess-playing program techniques are not used for A.I. in general, and vice versa, except for a few basic principles such as the importance of combinatorial search.

Another example of specialized intelligence is for automatic translation of natural language. It was previously thought that general-purpose artificial intelligence would be required for good-quality language translation, since in order for a person to translate a text he or she must understand its contents. However, at the present stage of development it turns out that better translation results are obtained using specialized, 'stupid' translation programs that use statistical methods and large databases of known translation patterns, and knowledge-based language translation systems are not able to compete with them at present. This is another example of a specialized task that seemingly requires general-purpose intelligence in humans, and where a specialized approach has been more successful so far in computers.

Given these observations one may wonder whether the most promising long-term strategy for intelligence in computers will be to build a number of specialized 'intelligences', like the ones just described, and to integrate them gradually. Some researchers have argued this position, sometimes under the name of the 'big switch' theory of intelligence. Time will tell, but it is fair to assume that the 'agent intelligence' approach that has been used for the present course will *at least* be one of the participating 'intelligences'. Whether it will just be one out of several, or whether it will also be the cohesive force that makes it possible to integrate the others, or whether it will eventually be the universal principle that includes the others, this remains to be seen.

2.2 Alternative Computational Infrastructures

The topics in the A and B categories of the topic list above are based on conventional computational technology: algorithms, datastructures, programming languages, software architectures, and the use of formal logic. This technology has been developed for, and is adapted to the use of the von Neumann computer design with a small number of processors that work with almost entirely passive memories.

This computational infrastructure is of course very different from the structure of the human brain with its extremely large number of concurrently active brain cells. Some branches of research propose that this is a serious handicap, and that a better strategy for the development of machine intelligence is to build computational systems that are more similar to the neural or neural-based system in people and in animals. This is the approach of *neural computation.*

A somewhat related view questions the precedence of the knowledge-based techniques in the A category above, and propose that it is better to first implement the counterparts of more elementary human behavior, including sensorimotoric behavior and other stimulus-response behavior. The proposal is that intelligence of the kinds discussed above – agent intelligence, chess-playing and language-translation capabilities, and so on – are based on these simpler behaviors and that they *emerge* from them, i.e. they are evolved from them. These simpler behaviors can then conceivably be implemented either using neural networks or other alternative hardware architectures, or using conventional programming-language technology.

Approaches using alternative computational infrastructures have produced good results concerning the implementation of lower-level behavior, below the level that is normally called 'intelligence'. However, their usefulness for the construction of agent-level intelligence or specialized intelligence of the kinds discussed above has not been demonstrated and there is no concrete indication that it will ever be a viable approach.

3 Representation of subareas in the present course

The contents of the present course can now be related to the structure of the field that was shown above. The following topics, among those listed

above, are the ones that have been represented to some reasonable extent in the present course.

```
A  Agent-based and multiagent systems
A  Commonsense reasoning
B  Constraint satisfaction, search, and optimization
B  Knowledge acquisition and ontologies
A  Knowledge representation and reasoning
A  Planning and Scheduling
```

A few additional areas have been touched upon very marginally, in particular:

```
C  Cognitive modeling and human interaction
A  Model-based systems
C  Web and information systems
```

This means in summary that the present course has concentrated the topics that are of basic importance, in particular for the design of agent intelligence, but also arguably for artificial intelligence in general.