THE SCOPE OF ARTIFICIAL INTELLIGENCE: Background document for a proposal for an artificial intelligence GCE/A-level syllabus.

Aaron Sloman
School of Computer Science
The University of Birmingham

What is Artificial Intelligence (AI)?

This document provides backround information about the scope of Artificial Intelligence for use in conjunction with a document proposing a new syllabus for teaching AI in schools, available here
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/courses/alevel-ai.html
I start with a high level summary definition of AI, and then elaborate on various aspects of AI to provide an indiciation of the scope of the field. This document cannot hope to keep up with the development of AI and will always be incomplete and out of date. Some additional sources of information are given at the end.

What is AI?

AI is a (badly named) field of enquiry with two closely related strands: science and engineering.

The scientific strand of AI attempts to provide understanding of the requirements for, and mechanisms enabling, intelligence of various kinds in humans, other animals and information processing machines and robots.

The engineering strand of AI attempts to apply such knowledge in designing useful new kinds of machines and helping us to deal more effectively with natural intelligence, e.g. in education and therapy.

AI is inherently highly interdisciplinary because all kinds of intelligence, whether natural or artificial are concerned with subject matters that are studied in other disciplines, and the explanatory models of natural intelligence have to take account of and be evaluated in the disciplines that study the natural forms.

NOTE

Like Alan Turing, in 'Computing machinery and intelligence', *Mind*, 59, pp. 433--460, 1950, I believe attempting to define 'intelligence' is a complete waste of time. We can collect many different examples of competences displayed by humans or other animals, and examples of challenging biologically-inspired behaviours required in future machines, and we can investigate requirements for modelling or replicating them without needing to draw any definite line between those that are and those that are not intelligent. We may find it useful to subdivide the cases in terms of either their capabilities, or the mechanisms required, or the kinds of information they use, or their potential usefulness in various contexts, and those divisions will be much more interesting and useful than any binary division based on a pre-theoretical concept like 'intelligence'.

Contents of the remainder of this document

The remainder of this document elaborates on the brief definition of AI given above. It is broadly based on the document on AI produced by the author for the QAA benchmarking panel on Computer Science degrees in 1999, available at
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/courses/ai.html
which was informed by consultation with university teachers and researchers in AI in the UK.

The aim is not to provide a syllabus for AI, but to provide a reminder of the scope of AI that can inform the more detailed design of a syllabus. However, many of these topics are too difficult to be included in a school syllabus, and are here merely for information.

The section on **AI tools and languages** gives some information about AI tools and languages, how they differ from some other programming languages and tools, and why they are needed. It is not intended to give a complete overview of AI tools or languages.

The final section **additional information** gives pointers to further information, including the ACM classification of areas of Computing, and information about some well known national and international AI Professional Organisations, some of which include information about the scope of AI in their own web pages. It is hoped that these pointers will help to compensate for any omissions or distortions found here.

ARTIFICIAL INTELLIGENCE: AN OVERVIEW

Not all practitioners will agree with the characterisation of AI given below, which is deliberately broad, encompassing both AI as science and AI as engineering. For instance, some people who work in the field regard it as a branch of engineering, and would not include the study of human intelligence, whereas that is a central focus of many AI researchers, including some of the founders of the discipline.

The overview lists a number of sub-fields of AI which are sometimes regarded as parts of AI, and sometimes simply as ordinary applications of computer science, e.g. image interpretation. It has often happened that a technique originally developed for AI purposes has been accepted as part of "mainstream" computer science or software engineering, and it is likely that this will continue to happen.

The scope of artificial intelligence

AI is a vast, and misleadingly named, multi-disciplinary field of research and teaching which grew up in parallel with computer science and software engineering, while also building on and overlapping with other subjects like linguistics, philosophy, psychology, biology, mathematics, and logic. There are some who think it also needs advances in quantum physics in order to make progress. Not only is it multi-disciplinary in its origins and contents: courses in AI are taught not only in computer science departments, but also in others, e.g. psychology departments. Likewise degree courses in AI may include components that would often be found in other degrees, e.g. courses in philosophy of mind or philosophy of science, courses in linguistic theory, courses in human perception, or development or other aspects of human psychology.

It is clear that AI is still in its infancy: there have been many interesting theoretical developments and useful applications, but many hard problems remain unsolved, and the subject can be expected to evolve rapidly in coming years, especially as developments in the power and costs of computers both enable more effective research and also increase the need for AI.

AI has two main strands, a scientific strand and an engineering strand, which overlap considerably in their concepts, methods, and tools, though their objectives are very different.

AI as Science:

The scientific strand, which has motivated most of the pioneers and leaders in the field, is concerned with two main goals (a) attempting to understand and model the information processing capabilities of typical human minds, (b) attempting to understand the general principles for explaining and modelling intelligent systems, whether human, animal or artificial. This work is often inspired by research in philosophy, linguistics, psychology, neuroscience or social science. It can also lead to new theories and predictions in those fields.

AI as Engineering:

The engineering strand, which motivates most of the funding agencies and (consequently) younger researchers, is concerned with attempting to design new kinds of machines able to do things previously done only by humans and other animals and also new tasks that lie beyond human intelligence.

There is another nascent engineering application of AI: using the results of the scientific strand to help design machines and environments, and educational strategies that can help human beings. This may, but need not, including the production of intelligent machines. It could include the design of therapies or teaching strategies which engage more effectively with the information processing capabilities of patients or learners. Engineers also often need to take into account theories of natural intelligence when designing systems to interact usefully with humans.

Besides the contrast between Science and Engineering, there are other ways of dividing approaches to AI, e.g. according to the types of mechanisms or forms of representations thought to be most useful, or according to whether AI systems are explicitly designed or bootstrapped by some learning or evolutionary mechanism. New approaches often become fashionable for while and then are simply absorbed into the larger pattern of research and teaching.

Sub-areas of AI:

Sub-areas can be divided up in two main ways, according to the content of the study or according the tools and techniques used. These are expanded below.

(a) Content
Because humans and other animals, and equally robots and other intelligent artefacts, have a wide variety of capabilities, all of them very complex and hard to explain or model, both the scientific and engineering work in AI has spawned several sub-fields, dealing with particular aspects of intelligence.

(b) Techniques
Because the applications of AI are many and diverse some of the sub-fields are clustered around techniques relevant to a class of problems.

(a) Sub-fields based on content.

This list is not well structured, and the order could be improved.

- Perception, especially vision but also auditory and tactile perception, and more recently taste and smell. This breaks down into the study of different types of processes including physical transduction, analysis and pattern recognition, segmentation and "parsing" of complex sensory data, interpretation, and control of attention. This is a huge sub-field and could be broken down into more specialised sub-sub-fields according to the sensory modality, the kinds of things being perceived, the forms of representations used, whether perception is purely data-driven or includes top-down processes, the mechanisms used (e.g. neural or symbolic), larger architecture which contains the perceptual system, and the application domain.

- Natural language processing, including production and interpretation of spoken and written language, whether hand-written, printed, or electronic throughout (e.g. email).

- Learning and development, including symbolic learning processes (e.g. rule induction), the use of neural nets (sometimes described as sub-symbolic), the use of evolutionary algorithms, self-debugging systems, and various kinds of self-organisation.

- Planning, problem solving, automatic design: given a complex problem and a collection of resources, constraints and evaluation criteria create a solution which meets the constraints and does well or is optimal according to the criteria, or if that cannot be done propose some good alternatives.

- Varieties of reasoning: This includes studies both of informal common sense reasoning, and specialised expert reasoning. The former includes the study of analogical reasoning, defeasible inferences, case based reasoning. The latter includes logical and mathematical reasoning, including design of theorem proving and inference systems either with the intention of modelling various kinds of human mathematical and inferential capabilities, or for practical purposes, e.g. in symbolic algebra toolkits, or reasoning in robots or autonomous control systems.

- Study of representations: investigating the formal properties of various sorts of representations, the mechanisms required for them to operate, and the kinds of tasks for which they are good or bad. This can include the study of ontologies of various kinds. Some mechanisms are sometimes claimed not to use any representations (e.g. neural nets) while a counter-claim is that they are simply a particular type of representation, e.g. numerical and continuous as opposed to structural and discrete.

- Memory mechanisms and techniques: analysis of requirements for various kinds of memory, including very large knowledge stores containing diverse types of knowledge, either for modelling human knowledge or for use in various sorts of applications.

- Search mechanisms and techniques: finding a solution to some problem in a space of possibilities is a recurring theme in AI. Many different forms of search have been studied, related to different forms of representation, different problem domains and different requirements (e.g. Must the solution be optimal, or is satisficing the target? If satisficing is not enough, and optimality is too hard could it suffice to find a solution that is guaranteed to be close to optimal, within some limit?)

- Multi agent systems: the study of various kinds of communication (linguistic and non-linguistic, explicit and implicit, intentional and unintended), types of cooperation and conflict, recognition of plans and intentions in others, etc. Some studies of multi-agent systems are concerned with understanding human social interactions, whereas others are concerned with the design of applications involving multiple robots or multiple concurrent software systems. Some multi-agent systems are proposed as an architecture for a single complex intelligent agent.

- Affective mechanisms: during the late 1990s there has been a growing interest in the role of motivation and emotions in intelligence. This is sometimes studied as a topic in its own right, and sometimes as part of the study of complete architectures for autonomous intelligent systems. A general theory would need to account for a wide variety of affective states and processes, including desires, preferences, dislikes, pleasures, pains, long term goals, intentions, ideals, values, attitudes, moods, and more. One of the current debates concerns whether emotions are required for intelligence, or whether they are merely side-effects or emergent features of mechanisms that are required for other functions.

- Robotics: one of the oldest sub-fields of AI. This is sometimes studied for the purpose of producing new kinds of useful machines (which is why this heading is also in the tools section below), and sometimes because designing complete working robots provides a test bed for integrating theories and techniques from various sub-areas of AI, e.g. perception, learning, memory, motor control, planning, etc. I.e. it is a context for exploring ideas about **complete** systems. Sometimes robot designers attempt to show that certain sorts of mechanisms are **not** required in systems with a certain type of intelligence, e.g. showing what robots using no planning or deliberative capabilities can do.

- Architectures for complete systems. This was mentioned above in connection with *Affective mechanisms.*
  Up to the mid 1980s most of the work in AI was concerned with specific forms of representation and specific algorithms for performing some task. Since then there has been a growing realisation of the importance of an *architecture* in which many different mechanisms are combined to provide a system with many different sorts of functionality, often concurrently active mechanisms. (There were some earlier discussions of architectures in AI.)

- Ontologies have been receiving considerable attention as a result of a growing realisation that it is not enough to specify the forms of representation that an intelligent system uses. It is important also to investigate what sorts of things should be represented. An ontology specification is a specification of what kinds of things are taken to exist: two individuals that share an ontology may, however, disagree as to which particular things allowed in that ontology actually exist, or which laws govern their behaviour. (This topic is closely related to old philosophical theories about what

exists or what can exist.) Development of an ontology as a result of interacting with some environment is an important kind of learning.
NOTE:

> The word ’ontology’ is often used to refer to what is here described as the *specification of an ontology.* That specification could be a particular document in a particular formal language. However, the same ontology (the same set of entities, properties, relations, states, processes, etc.) could be specified using different notations, in many different documents. So it is less confusing to use the word ’ontology’ to refer to the content not the vehicle, i.e. to what is specified rather than the means of specifying it.

- Development of languages and tools listed below in the tools section.

NB. The above is not intended to be a complete list.
There are many other sub-fields which could be listed. For a more comprehensive survey follow the pointers given in the final section, below.

(b) Sub-fields of application of AI

There is a very open-ended set of fields of application of AI. The following are merely examples, not a complete list:

- AI in medicine, including interpretation of medical images, diagnosis, expert systems to aid GPs, monitoring and control in intensive care units, design of prosthetics, design of drugs, intelligent tutoring systems for various aspects of medicine.

- AI in robotics: including vision, motor control, learning, planning, linguistic communication, cooperative behaviour

- AI in many aspects of engineering: fault diagnosis, intelligent control systems, intelligent manufacturing systems, intelligent design aids, integrated systems for sales, design, production, maintenance, expert configuration tools (e.g. ensuring sales staff don’t sell system that won’t work). AI in software engineering includes work on program synthesis, verification, debugging, testing and monitoring of software.

- AI in interfaces and "help" systems: as computers are used for more and more applications that involve interaction with human beings, there are ever growing pressures to make the machines easier for non-experts to use. One approach to this is to give machines more intelligence so that they can guide, or advise users. This type of application can include many of the sub-fields listed above in section C.4.(a)

- AI in education: including various kinds of intelligent tutoring systems and student management systems. Particular applications might include diagnosis of a student’s knowledge gaps, various kinds of drill and practice tutors, automatic marking of programming assignments and essays, etc.

- AI in information management: this includes the use of AI in data mining, web crawling, email filtering, etc.

- AI in mathematics: design of tools to help with various kinds of mathematical tasks, now used so widely that they are not recognised as AI products.

- AI in entertainment: increasingly AI is being used in computer games and in systems for generating and controlling synthetic characters either for textual interaction or generating films with cartoon characters or interactive avatars in virtual worlds.

- AI in biology: there are many hard problems in biology where more or less intelligent computer-based systems are being developed, e.g. analysis of DNA, prediction of folded structure of complex molecules, prediction, modelling many biological processes, evolution, development of embryos, behaviours of various organisms.

- AI in Law: e.g. expert systems to help lawyers, or systems to give legal advice and help to non-lawyers.

- AI in architecture, urban design, traffic management: tools to help solve design problems involving multiple constraints, helping to predict the behaviours of people in new environments, tools to analyse patterns in observed phenomena.

- AI in literature, art, music: identification of authors, modelling of processes of generation and appreciation, teaching applications.

- AI in crime prevention and detection: e.g. detection of forgeries, learning to detect evidence of crooked police officers, software to monitor internet transactions, helping to plan police operations, searching police databases for evidence that crimes are committed by the same person, etc.

- AI in commerce: a fast growing application area of AI enabled by the internet involves commerce, especially electronic commerce and the use of software agents of various sorts to provide, search for, analyse or interpret information, take decisions, negotiate with other agents, etc.

- AI in space: control of space vehicles and autonomous robots too far from earth to be directly manipulated by humans on earth, because of transmission delays.

- AI in military activities. This may be the area in which most funds have been spent. It is also not easy to learn about the details.

Again this is not intended to be a complete list.
It is also not claimed that there are no other useful ways of dividing AI into possible course topics or research areas. Many of the above sub-headings were chosen simply because there are groups of researchers, conferences, journals, or books that focus on the topics given.

Far more information about types of AI applications can be found at http://www.aaai.org/aitopics

---

AI tools and languages

AI researchers have often found that existing software tools and programming languages, and standard design and development methodologies, were too restrictive, and therefore developed new special purpose versions. However many of the ideas have been taken up more generally, so that the distinctions are no longer clear. For example, many early AI systems were more concerned with manipulation of symbols and symbolic structures than manipulation of numbers, and they therefore included many constructs for non-numeric programming. However this is now a standard requirement for software engineering. Similarly early AI systemts included automatic 'garbage collection' mechanisms for working out which data structures were still in use and which were not after a series of complex manipulations. Now automatic garbage collection mechanisms have been added to other software development environments.

Some of the reasons for the special requirements of AI languages are

- Many AI problems are not well defined initially: in such cases the process of developing and testing software helps to clarify the problem.
- Often the tasks require development of very complex designs which are hard to get right. This requires powerful interactive testing facilities, without having to start the program again and run it with a separate debugging tool.
- AI problems often require different kinds of knowledge to be represented and different kinds of inference mechanisms to be used. This requires AI languages to support mixed-paradigm programming, and to be extendable so as to allow constructs suited to new problems to be added easily.
- Intelligent systems may need to modify themselves at run time, including modifying the code they run. E.g. a result of learning could be the development of a new techniques or skill requiring new programs to be used. This meant that AI development systems were often based on interpreted or incrementally compiled programming languages.
- AI systems often need to construct complex networks of temporary structures, where different portions have different life-spans.
- AI systems often need to improve their own performance either by learning or by simulating biological evolution.
- AI systems often have to perform complex *searches* for solutions to problems, e.g. searching for a plan, searching for an interpretation of an image, searching for a proof of a theorem, or searching for the best generalisation to summarise a collection of learning experiences.

These requirements have led to the development of languages which

- Are extendable using macros and other facilities
- Include support for automatic garbage collection (once described by a well known computer scientist as a luxury for the idle rich).
- Use interpreters or incremental compilers
- Have built in inference mechanisms or libraries providing inference mechanisms.
- Allow modules to be modified or replaced at run time

- Support multiple programming paradigms
- Provide run-time type-checking instead of compile-time checking based on syntactic types.
- Include editors and other development tools as part of the run-time system.
- Provide mechanisms for building neural nets or evolutionary computation systems.
- Provide mechanisms to support controlled searching possibly using different mechanisms.

Some of the specialised languages developed to support AI research and applications include Prolog, Scheme, Smalltalk, OPS-5 and other production system interpreters, several varieties of Lisp including Common Lisp and its main precursors Maclisp and Interlisp, Pop2 and its derivatives, e.g. Pop-11, hybrid systems supporting more than one language, e.g. Loglisp, Poplog.

Using these and other languages, various tools have been developed to support knowledge acquisition and testing, theorem provers, planners, problem solvers, parsers and other forms of software for manipulating natural language, neural net toolkits, image processing tools, robot development tools, tools for designing and testing cognitively rich agents, tools for developing multi-agent systems, rule induction and learning systems, automatic program generating and testing tools, tools for doing experiments in artificial life, and tools for supporting evolutionary computation.

Some of the tools are closely related to particular theories, or intended to support particular types of techniques, e.g. constraint-manipulation toolkits, tools for building cognitive models based on SOAR, or ACT-R.

There have been some experiments in designing new forms of hardware to support AI, e.g. hardware tailored to playing chess, hardware for vision, hardware for implementing AI languages like Lisp or Prolog, hardware for neural computation, in addition to robots and robot components. In future there may be AI models or applications using entirely new forms of computers, e.g. quantum computers or DNA computers.

Many of the tools required for AI as engineering overlap with those required for AI as science, since the task of producing intelligent applied systems has much in common with the task of producing models of natural intelligent systems. Of course there are some AI researchers who would dispute that.

Should there be a separate section on AI techniques and representations?

Precursors

This document is based to a large extent on the overview of AI produced for the QAA CS Benchmarking panel in 1999
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/courses/ai.html
The version of this document that existed around December 1999 was published as
A.Sloman, Artificial Intelligence, an Illustrative Overview, in *AISB Quarterly* Winter 2000, no 103, pages 31-35, with a brief introduction by the editor, Blay Whitby.

Acknowledgements

Useful comments on the precursor of this document were made by John Barnden (Birmingham University), Ann Blandford (Middlesex University), Benedict du Boulay (Sussex University), Inman Harvey (Sussex University), Bob Hendley (Birmingham University), Phil Husbands (Sussex University), Peter Ross (Edinburgh University).

Additional information

The Association for Computing Machinery has a list of sub-fields of computer science, which changes from time to time. Several versions are available at http://www.acm.org/class/

It is not updated often, so it is now out of date and should be treated with caution.

AI Organisations

There are two main UK AI organisations.

The Society for the Study of Artificial Intelligence and Simulation of Behaviour, claims to be the oldest AI society in the world. See http://www.aisb.org.uk/

The British Computer Society Specialist Group on Artificial Intelligence (BCS-SGAI) has a more applied focus, though its seminars and annual conferences are very wide ranging. See http://www.bcs-sgai.org/

The main European AI organisation is ECCAI (European Coordinating Committee on AI), to which national AI organisations are affiliated. See http://www.eccai.org

The largest AI organisation is the Association for the Advancement of Artificial Intelligence (AAAI). Information about it is at http://www.aaai.org . It includes AITOPICS, a collection of Web pages (under continual development) that attempt to characterise the scope of AI, and provide a steady stream of news about AI and its applications.

The major regular international conference on AI is The International Joint Conference on AI (IJCAI) held every two years since 1969. See http://www.ijcai.org

There are also many national Artificial Intelligence Societies, which organise conferences, and other activities.

There is a growing collection of web sites providing information about AI, some compiled by individuals and some by firms or organisations.

A document on AI for School Careers Advisers

A document was put together in 1998 for a conference of school careers advisers in England giving a more discursive overview of AI: http://www.cs.bham.ac.uk/~axs/misc/aiforschools.html

This includes pointers to several other sources of information about AI, including summaries by some of the founders of the field (e.g. Minsky, McCarthy), some textbooks, and UK universities known to be offering undergraduate degrees with AI in the title in May 1998.

This file maintained by Aaron Sloman
Last updated: 4 Mar 2007