History, motivations and core themes of AI

By Stan Franklin

Introduction

This chapter is aimed at introducing the reader to field of artificial intelligence (AI) in the context of its history and core themes. After a concise preamble introducing these themes, a brief and highly selective history will be presented. This history will be followed by a succinct introduction to the major research areas within AI. The chapter will continue with a description of currents trends in AI research, and will conclude with a discussion of the current situation with regard to the core themes. The current trends are best understood in terms of AI history, its core themes and its traditional research areas. My goal is to provide the reader with sufficient background context for understanding and appreciating the subsequent chapters in this volume.

Overview of Artificial Intelligence core themes

The history of artificial intelligence may be best understood in the context of its core themes and controversies. Below is a brief listing of such AI distinctions, issues, themes and controversies. It would be well to keep these in mind during your reading of the rest of this chapter. Each of the themes will be expanded upon and clarified as the chapter progresses. Many of these result from their being, to this day, no agreed up definition of intelligence within the AI community of researchers.

Smart Software vs. Cognitive Modeling

AI has always been a part of computer science, an engineering discipline aimed at creating smart computer programs, that is, intelligent software products to meet human needs. We'll see a number of examples of such smart software. AI also has its science side that's aimed at helping us understand human intelligence. This endeavor includes building software systems that "think" in human like ways, as well as producing computational models of aspects of human cognition. Such computational models provide hypotheses to cognitive scientists.

Symbolic AI vs. Neural Nets

From its very inception artificial intelligence was divided into two quite distinct research streams, symbolic AI and neural nets. Symbolic AI took the view that intelligence could be achieved by manipulating symbols within the computer according to rules. Neural nets, or connectionism as the cognitive scientists called it, instead attempted to create intelligent systems as networks of nodes each comprising a simplified model of a neuron. Basically, the difference was between a computer analogy and a brain analogy, between implementing AI

systems as traditional computer programs and modeling them after nervous systems.

Reasoning vs. Perception

Here the distinction is between intelligence as high-level reasoning for decision-making, say in machine chess or medical diagnosis, and the lower-level perceptual processing involved in, say machine vision, the understanding of images by identifying objects and their relationships.

Reasoning vs. Knowledge

Early symbolic AI researchers concentrated on understanding the mechanisms (algorithms) used for reasoning in the service of decision-making. The assumption was that understanding how such reasoning could be accomplished in a computer would be sufficient to build useful smart software. Later, they realized that, in order to scale up for real-world problems, they had to build significant amounts of knowledge into their systems. A medical diagnosis system had to know much about medicine, as well as being able to draw conclusions.

To Represent or Not

Such knowledge had to be represented somehow within the system, that is, the system had to somehow model its world. Such representation could take various forms, including rules. Later, a controversy arose as to how much of such modeling actually needed to be done. Some claimed that much could be accomplished without such internal modeling.

Brain in a Vat vs. Embodied AI

The early AI systems had humans entering input into the systems and acting on the output of the systems. Like a "brain in a vat" these systems could neither sense the world nor act on it. Later, AI researchers created embodied, or situated) AI systems that directly sensed their worlds and also acted on them directly. Real world robots are examples of embodied AI systems.

Narrow AI vs. Human Level Intelligence

In the early days of AI many researchers aimed at creating human-level intelligence in their machines, the so-called "strong AI." Later, as the extraordinary difficulty of such an endeavor became more evident, almost all AI researchers built systems that operated intelligently within some relatively narrow domain such as chess or medicine. Only recently has there been a move back in the direction of systems capable of a more general, human-level intelligence that could be applied broadly across diverse domains.

Some Key Moments in AI

McCulloch and Pitts

The neural nets branch of AI began with a very early paper by Warren McCulloch and Walter Pitts (1943). McCulloch, a professor at the University of Chicago, and Pitts, then an undergraduate student, developed a much-simplified model of a functioning neuron, a McCulloch-Pitts unit. They showed that networks of such units could perform any Boolean operation (and, or, not) and,

thus, any possible computation. Each of these units compared the weighted sum of its inputs to a threshold value to produce a binary output. Neural Nets AI, and also computational neuroscience, thus was born.
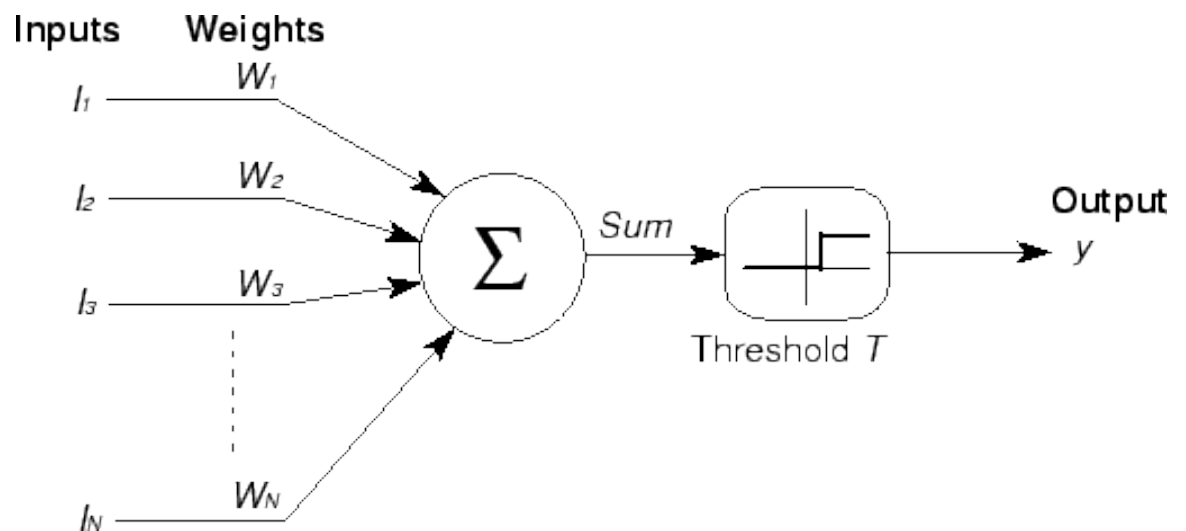
Formal neuron

Also called a linear threshold device or a threshold logic unit
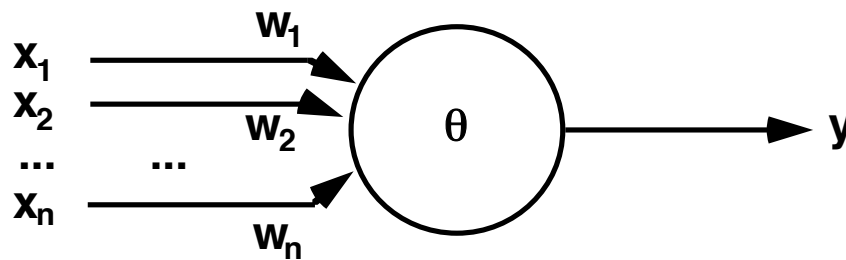
x_i -- the inputs

w_i -- the weights (synaptic strengths)

θ -- the threshold

y -- the output

$$y(t+1) = \begin{cases} 1 & \text{if } \sum_i w_i x_i(t) \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

Alan Turing

Alan Turing, a Cambridge mathematician of the first half of the twentieth century, can be considered the father of computing (its grandfather was Charles Babbage during the mid-nineteenth century) and the grandfather of artificial intelligence. During the Second World War in 1939-1994 Turing pitted his wits against the Enigma cipher machine, the key to German communications. He led in developing the British Bombe, an early computing machine that was used over and over to decode messages encoded using the Enigma.

During the early twentieth century Turing and others were interested in questions of computability. They wanted to formalize an answer to the question

of which problems can be solved computationally. Several people developed distinct such formalisms. Turing offered the Turing Machine (1936), Alonzo Church the Lambda Calculus (1936), and Emil Post the Production System (1943). These three apparently quite different formal systems soon proved to be logically equivalent in defining computability, that is, for specifying those problems that can be solved by a program running on a computer. The Turing machine proved to be the most useful formalization, and is the one most often used in theoretical computer science.

In 1950 Turing published the very first paper suggesting the possibility of artificial intelligence (1950). In it he first described what we now call the Turing test, and offered it as a sufficient condition for the existence of AI. The Turing test has human testers conversing in natural language without constraints via terminals with either a human or an AI natural language program, both hidden from view. If the testers can't reliably distinguish between the human and the program, intelligence is ascribed to the program. In 1991 Hugh Loebner established the Loebner Prize, which would award $100,000 to the first AI program to pass the Turing Test. As of this writing, the Loebner Prize has not been awarded.

Dartmouth Workshop

The Dartmouth Workshop served to bring researchers in this newly emerging field together to interact and to exchange ideas. Held during August of 1956, the workshop marks the birth of artificial intelligence. AI seems alone among disciplines in having a birthday. Its parents included John McCarthy, Marvin Minsky, Herbert Simon and Allen Newell. Other eventually prominent attendees were Claude Shannon of Information Theory fame, Oliver Selfridge, the developer of Pandemonium Theory, and Nathaniel Rochester, a major designer of the very early IBM 701 computer.

John McCarthy, on the Dartmouth faculty at the time of the Workshop, is credited with having coined the name Artificial Intelligence. He was also the inventor of LISP, the predominant AI programming language for a half century. McCarthy subsequently joined the MIT faculty and, later, moved to Stanford where he established their AI Lab. As of this writing he's still an active AI researcher.

Marvin Minsky helped to found the MIT AI Lab where he remains an active and influential AI researcher until the time of this writing.

Simon and Newell brought the only running AI program, the logical theorist, to the Dartmouth Workshop. It operated by means-ends analysis, an AI planning algorithm. At each step it attempts to choose an operation (means) that moves the system closer to its goal (end). Herbert Simon and Allen Newell founded the AI research lab at Carnegie Mellon University. Newell passed away in 1992, and Simon in 2001.[i]

Samuel's Checker Player

Every computer scientist knows that a computer only executes an algorithm it was programmed to run. Hence, it can only do what its programmer told it to do. Therefore it cannot know anything its programmer didn't, nor do anything its programmer couldn't. This seemingly logical conclusion is, in fact, simply wrong

because it ignores the possibility of a computer being programmed to learn. Such machine learning, later to become a major subfield of AI, began with Arthur Samuel's checker playing program (1959). Though Samuel was initially able to beat his program, after a few months of learning it's said that he never won another game from it. Machine learning was born.

Minsky's Dissertation

In 1951, Marvin Minsky and Dean Edmonds build the SNARC, the first artificial neural network that simulated a rat running a maze. This work was the foundation of Minsky's Princeton dissertation (1954). Thus one of the founders and major players in symbolic AI was, initially, more interested in neural nets and set the stage for their computational implementation.

Perceptrons and the Neural Net Winter

Frank Rosenblatt's perceptron (1958) was among the earliest artificial neural nets. A two-layer neural net best thought of as a binary classifier system, a perceptron maps its input vector into a weighted sum subject to a threshold, yielding a yes or no answer. The attraction of the perceptron was due to a supervised learning algorithm, by means of which a perceptron could be taught to classify correctly. Thus neural nets contributed to machine learning.

Research on perceptrons came to an inglorious end with the publication of the Minsky and Pappert book (1969) in which they showed the perceptron incapable of learning to classify as true or false the inputs to such simple systems as the exclusive or (XOR – either A or B but not both). Minsky and Papert also conjectured that even mulit-layered perceptrons would prove to have similar limitations. Though this conjecture proved to be mostly false, the government agencies funding AI research took it seriously. Funding for neural net research dried up, leading to a neural net winter that didn't abate until the publishing of the Parallel Distributing Processing volumes (McClelland and Rumelhart 1986, Rumelhart and McClelland 1986).

The Genesis of Major Research Areas

Early in its history the emphasis of AI research was largely toward producing systems that could reason about high-level, relatively abstract, but artificial problems, problems that would require intelligence if attempted by a human. Among the first of such systems was Simon and Newell's general problem solver (Newell, Shaw, Simon 1959), which, like its predecessor the logical theorist, used means ends analysis to solve a variety of puzzles. Yet another early reasoning system was Gelernter's geometry theorem prover,

Another important subfield of AI is natural language processing, concerned with systems that understand. Among the first such was SHRDLU (Winograd 1972), named after the order of keys on a linotype machine. SHRDLU could understand and execute commands in English ordering it to manipulate wooden blocks, cones, spheres, etc. with a robot arm in what came to be known as a blocks world. SHRDLU was sufficiently sophisticated to be able to use the remembered context of a conversation to disambiguate references.

It wasn't long, however, before AI researchers realized that reasoning wasn't all there was to intelligence. In attempting to scale their systems up to deal with

real world problems, they ran squarely into the wall of the lack of knowledge. Real world problems demanded that the solver know something. So, knowledge based systems, often called expert systems, were born. The name came from the process of knowledge engineering, of having knowledge engineers laboriously extract information from human experts, and handcraft that knowledge into their expert systems.

Lead by chemist Joshua Lederberg, and AI researchers Edward Feigenbaum and Bruce Buchanan, the first such expert system, called Dendral was an expert in organic chemistry. DENDRAL helped to identify the molecular structure of organic molecules by analyzing data from a mass spectrometer and employing its knowledge of chemistry (Lindsay, Buchanan, Feigenbaum, and Lederberg. 1980). The designers of DENDRAL added knowledge to its underlying reasoning mechanism, an inference engine, to produce an expert system capable of dealing with a complex, real world problem.

A second such expert system, called Mycin (Davis, Buchanan and Shortliffe. 1977), helped physicians diagnose and treat infectious blood diseases and meningitis. Like DENDRAL, Mycin relied on both hand crafted expert knowledge and a rule based inference engine. The system was successful in that it could diagnose difficult cases as well as the most expert physicians, but unsuccessful in that it was never fielded. Inputting information into Mycin required about twenty minutes. A physician would spend at most five minutes on such a diagnosis.

Research During the Neural Net Winter

Beginning with the publication of *Perceptrons* (Minsky and Papert 1969), the neural net winter lasted almost twenty years. The book had mistakenly convinced government funding agencies that the neural net approach was unpromising. In spite of this appalling lack of funding, significant research continued to be performed around the world. Intrepid researchers who somehow managed to keep this important research going included Amari and Fukushima in Japan, Grossberg and Hopfield in the United States, Kohonen in Finland, and von der Malsberg in Germany. Much of this work concerned self-organization of neural nets, and learning therein. Much was also motivated by the backgrounds of these researchers in neuroscience.

The Rise of Connectionism

The end of the neural net winter was precipitated by the publication of the two Parallel Distributed Processing volumes (Rumelhart and McClelland 1986, McClelland and Rumelhart 1986). They were two massive, edited volumes with chapters authored by members of the PDP research group, then at the University of California, San Diego. These volumes gave rise to the application of artificial neural nets, soon to be called connectionism, to cognitive science. Whether connectionism was up to the job of explaining mind, rapidly became a hot topic of debate among philosophers, psychologists and AI researchers (Fodor and Pylyshyn 1988, Smolensky 1987, Chalmers 1990). The debate has died down with no declared winner, and with artificial neural nets becoming an established player in the current AI field.

In addition to its success in the guise of connectionism for cognitive modeling, artificial neural nets have found a host of practical applications. Most of

these involve pattern recognition. They include mutual fund investing, fraud detection, credit scoring, real estate appraisal, and a host of others. This wide applicability has been primarily the result of a widely used training algorithm called back propagation. Though subsequently traced to much earlier work, back propagation was rediscovered by the PDP research group, and constituted the preeminent tool for the research reported in the two PDP volumes.

The AI Winter

Due to what turned out to be an overstatement of the potential and timing of artificial intelligence, symbolic AI suffered its own winter. As an example, in 1965 Herbert Simon predicted "machines will be capable, within twenty years, of doing any work that a man can do." This and other such predictions did not come to pass. As a result, by the mid-nineteen-eighties government agency funding for AI began to dry up and commercial investment became almost non-existent. Artificial intelligence became a taboo word in the computing industry for a decade or more, in spite of the enormous success of expert systems (more below). The AI spring didn't arrive until the advent of the next "killer" application, video games (again more below).

Soft computing

The term "soft computing" refers to a motley assemblage of computational techniques designed to deal with imprecision, uncertainty, approximation, partial truths, etc. Its methods tend to be inductive rather than deductive. In addition to neural nets, which we've already discussed, soft computing includes evolutionary computation, fuzzy logic, and Bayesian networks. We'll describe each in turn.

Evolutionary computation began with a computational rendition of natural selection called genetic algorithms (Holland 1975). A population search algorithm, it typically begins with a population of artificial genotypes representing possible solutions to the problem at hand. The members of this population are subjected to mutation (random changes) and crossover (the intermixing of two genotypes). The resulting new genotypes are input to a fitness function that measures the quality of the genotype. The most successful of these genotypes constitute the next population, and the process repeats. If well designed, the genotypes in the population tend over time to become much alike, thus converging to a desired solution and completing the genetic algorithm. In addition, evolutionary computation also includes classifier systems, which combine rule-based and reinforcement ideas with genetic algorithms. Evolutionary computation also includes genetic programming, a method of using genetic algorithms to search for computer programs, typically in LISP, that will solve a given problem.

Derived from Zadeh's fuzzy set theory, in which degrees of set membership between 0 and 1 are assigned (1965), fuzzy logic has become a mainstay of soft computing. Using if then rules with fuzzy variables, fuzzy logic has been employed in a host of control applications including home appliances, elevators, automobile windows, cameras and video games. References are not given since these commercial applications are almost always proprietary.

A Bayesian network, with nodes representing situations, uses Bayes' theorem on conditional probability to associate a probability with each of its links.

Such Bayesian networks have been widely used for cognitive modeling, gene regulation networks, decision support systems, etc. They are an integral part of soft computing.

Recent Major Accomplishments

We'll conclude our brief history of AI with an account of some of its relatively recent major accomplishments. These include expert systems, chess players, theorem provers, and a new killer application. Each will be described in turn.

Knowledge based expert systems

Though knowledge based expert systems made their appearance relatively early in AI history, they became a major, economically significant, AI application somewhat later. Perhaps the earliest such commercially successful expert system was R1, later renamed XCON (McDermott 1980). XCON saved millions for DEC (Digitial Equipment Corporation) by effectively configuring their VAX computers before delivery, rather than having DEC engineers solve problems after their delivery. Other such applications followed, including diagnostic and maintenance systems for Campbell Soups' cookers and GE locomotives. A Ford Motor Company advertisement for a piece of production machinery stipulated that such a diagnostic and maintenance expert system be a part of every proposal. One book detailed 2500 fielded expert systems. Expert systems constituted the first AI killer application. It was not to be the last.

Deep Blue beating Kasparov

Early AI researchers tended to work on problems that would require intelligence if attempted by a human. One such problem was playing chess. AI chess players appeared not long after Samuel's checker player. Among the most accomplished of these chess playing systems was IBM's Deep Blue, which in 1997 succeeded in defeating world champion Gary Kasparov in a six-game match, belatedly fulfilling another of Herbert Simon's early predictions. Though running on a specially built computer and provided with much chess knowledge, Deep Blue depended ultimately upon traditional AI game-playing algorithms. The match with Kasparov constituted an AI triumph.

Solution of the Robbins conjecture

Another, even greater, AI triumph was soon to follow. In a 1933 paper E.V. Huntington gave a new set of three axioms that characterized a Boolean algebra, a formal mathematical system important to theoretical computer science. The third of these axioms was so complex as to be essentially unusable. Thus motivated, Herbert Robbins soon replaced this third axiom with a simpler one, and conjectured that this new three-axiom set also characterized Boolean algebras. This Robbins conjecture remained one of a host of such in the mathematical literature until the prominent logician and mathematician Alfred Tarski called attention to it, turning it into a famous unsolved problem. After resisting the efforts of human mathematicians for over half a century, the Robbins conjecture finally succumbed to the banishments of a *general purpose* AI automatic theorem prover called EQP (EQuational Prover). Where humans

had failed, EQP succeeded in proving the Robbins conjecture to be true (McCune1997).

Games—the Killer App

Employing more AI practitioners than any other, the computer and video game industry is enjoying a screaming success. According to one reliable source, the Entertainment Software Association, 2004 sales topped seven billion dollars, with almost 250 million such games sold. AI's role in this astounding success is critical; its use is essential to producing the needed intelligent behavior on the part of the virtual characters who populate the games. Wikipedia has an entry entitled "game artificial intelligence" that includes a history of the ever increasing sophistication of AI techniques used in such games, as well as references to a half-dozen or so books on applying AI to games. At this writing there seems to be an unbounded demand for AI workers in the game industry. This highly successful commercial application is yet another triumph for AI.

Major AI Research Areas

There are almost a dozen distinct subfields of AI research each with its own specialized journals, conferences, workshops, etc. This section will provide a concise account of the research interests in each of these subfields.

Knowledge Representation

Every AI system, be it a classical AI system with humans providing input and using the output, or an autonomous agent (Franklin and Graesser 1997), must somehow translate input (stimuli) into information or knowledge to be used to select output (action). This information or knowledge must somehow be represented within the system so that it can be processed to help determine output or action. The problems raised by such representation constitute the subject matter of research in the AI subfield commonly referred to as knowledge representation.

In AI systems, one encounters knowledge represented using such logical formalisms such as propositional logic and first-order predicate calculus. One may also find network representations such as semantic nets whose nodes and links have labels providing semantic content. The underlying idea is that a concept, represented by a node, gains meaning via it relationships (links) to other concepts. More complex data structures such as production rules, frames, and fuzzy sets are also used. Each of these data structures has its own type of reasoning or decision-making apparatus, its inference engine.

The issue of to represent or not seems to have been implicitly settled, as the arguments have died down. Rodney Brooks of the MIT AI Lab seems to have made his point that more than was previously thought could be accomplished without representation (1991). His opponents, however, have carried the day, in that representations continue to be widely used. I believes that representations are critical for the process of deciding what action to take, and much less so for the process of executing the action. This seems to be the essence of the issue.

Heuristic Search

Search problems such as the traveling salesman problem have been studied in computer science almost since its inception. For example, find the most efficient route for a salesman to take to visit each of N cities exactly once. All known algorithms for finding optimal solutions to such a problem increase exponentially with N, meaning that for large numbers of cities no optimal solution can be found. However, good enough solutions can be found using heuristic search algorithms from AI. Such algorithms employ knowledge of the particular domain in the form of heuristics, rules of thumb, that are not guaranteed to find the best solution, but that most often find a good enough solution.

Such heuristic search algorithms are widely used for scheduling, for data mining (finding patterns in data), for constraint satisfaction problems, for games, for searching the web, and for many other such applications.

Planning

An AI planner is a system that automatically devises a sequence of actions leading from an initial real world state to a desired goal state. Planners may be used, for example, to schedule work on a shop floor, to find routes for package delivery, or to assign usage of the Hubble telescope. Research on such planning programs is a major subfield of AI. Fielded applications are involved in space exploration, military logistics, and plant operations and control.

Expert Systems

Knowledge based expert systems were discussed in the previous sections. As a subfield of AI expert systems researchers are concerned with reasoning (improving inference engines for their systems), knowledge representation (how to represent needed facts to their systems) and knowledge engineering (how to elicit knowledge that's sometimes implicit from experts. As we've seen above, their fielded applications are legion.

Machine Vision

Machine or computer vision is a subfield of AI devoted to the automated understanding of visual images, typically digital photographs. Among its many applications are product inspection, traffic surveillance and military intelligence[ii]. With images multiplying every few seconds from satellites, high-flying spy planes and autonomous drones, there aren't enough humans to interpret and index the objects in the images so that they can be understood and located. Research toward automating this process is just starting. AI research in machine vision is also beginning to be applied to security video cameras so as to understand scenes an alert humans when necessary.

Machine Learning

The AI subfield of machine learning[iii] is concerned with algorithms that allow AI systems to learn (see Samuel's checker player above). Though machine learning is as old as AI itself, its importance has increased as more and more AI systems, especially autonomous agents (see below), are operating in progressively more complex and dynamically changing domains. Much of machine learning is supervised learning in which the system is instructed using training data. Unsupervised, or self-organizing systems, as mentioned above, are

becoming common. Reinforcement learning, accomplished with artificial rewards, is typical for learning new tasks. There is even a new subfield of machine learning devoted to developmental robotics, robots that go through a rapid early learning phase as do human children.

Natural Language Processing

The AI subfield of natural language processing includes both the generation and the understanding of natural language, usually text. It's history dates back to the Turing test (see above). Today it's a flourishing field of research into machine translation, question answering, automatic summarization, speech recognition and other areas. Machine translators, though typically only 90% or so accurate, can increase the productivity of human translators fourfold. Text recognition systems are being developed for the automatic input of medical histories. Voice recognition enables spoken commands to a computer and even dictation.

Software agents

An autonomous agent is defined to be a system situated in an environment, and a part of that environment, that senses the environment and acts on it, over time, in pursuit of its own agenda, in such a way that its actions can influence what it later senses (Franklin and Graesser 1997). Artificial autonomous agents include software agents and some robots. Autonomous software agents come in several varieties. Some like the author's IDA "live" in an environment including databases and the internet, and autonomously perform a specified task such as assigning new jobs for sailors at the end of a tour of duty. Others, sometimes called avatars, have virtual faces or bodies displaying on monitors that allows them to interact more naturally with humans, often providing information. Still others, called conversational virtual agents, simulate humans, and interact conversationally with them in chat rooms, some so realistically as to be mistaken for human[iv]. Finally, there are virtual agents as characters in computer and video games.

Intelligent Tutoring Systems

Intelligent tutoring systems are AI systems, typically software agents, whose task it is to tutor students interactively one on one, much as a human tutor would. Results from early efforts in this direction were disappointing. Later systems were more successful in domains such as mathematics that lend themselves to short answers from the student. More recently intelligent tutoring systems like AutoTutor have been developed that can deal appropriately with full paragraphs written by the student. Today the major bottleneck in this research is getting domain knowledge into the tutoring systems. As a result, research in various authoring tools has flourished.

Robotics

In its early days robotics was a subfield of mechanical engineering with most research being devoted to developing robots capable of executing particular actions, such as grasping, walking, etc. Their control systems were purely algorithmic, with no AI components. As robots became more capable, the need for more intelligent control structures became apparent, and cognitive robotics

research involving AI-based control structures was born. Today, robotics and AI research have a significant and important overlap (more below).

Recent Trends

As 2007 began, artificial intelligence has not only emerged from its AI winter into an AI spring, but that spring has morphed into a full-fledged AI summer with its luxuriant growth of fruit. Flourishing recent trends include soft computing, agent based AI, cognitive computing, developmental robotics, and artificial general intelligence. Let's look at each of these in turn.

Soft computing

In addition to the components described earlier, namely neural nets, evolutionary computing and fuzzy logic, soft computing is expanding into hybrid systems merging symbolic and connectionist AI. Prime examples of such hybrid systems are ACT-R, CLARION, and the author's LIDA. Most such hybrid systems, including the three examples, were intended as cognitive models. Some of them underlie the computational architectures of practical AI programs. So=oft computing now also includes artificial immune systems with their significant contributions to computer security as well as applications to optimization and to protein structure prediction.

AI for data mining

Along with statistics, AI provides indispensable tools for data mining, the process of searching large databases for useful patterns of data. Many of these tools have been derived from research in machine learning. As databases rapidly increase in content, data mining become more and more useful, leading to a trend toward researching AI tools for data mining.

Agent based AI

The situated, or embodied, cognition movement (Varela Thompson and Rosch 1991), in the form of agent based AI, has clearly carried the day in AI research. Today, most newly fielded AI systems are autonomous agents of some sort. The dominant AI textbook (Russell and Norvig 2002), used in over 1000 universities world wide, is the leading text partially because its first edition was the first agent based AI textbook. Applications of AI agents abound. Some were mentioned in the section on software agents above.

Cognitive computing

Perhaps the newest, and certainly among the most insistent, current trends in AI research is what has come to be called cognitive computing[v]. Cognitive computing includes cognitive robotics, development robotics, self-aware computing systems, autonomic computing systems and artificial general intelligence. We'll briefly describe each in turn.

As mentioned above, robotics in its early days was primarily concerned with *how* to perform actions, and was mostly a mechanical engineering discipline. More recently this emphasis is shifting to action selection, that is, to deciding *what* action to perform. *Cognitive robotics*, the endowing of robots with more cognitive capabilities, was born, and is becoming an active subfield of AI.

Another closely related new AI research discipline, *developmental robotics*, combines robotics, machine learning and developmental psychology. The idea is enable robots to learn continually as humans do. Such learning should allow cognitive robots to operate in environments too complex and too dynamic for all contingencies to be hand crafted into the robot. This new discipline is supported by the IEEE Technical Committee on Autonomous Mental Development.

Government agencies are investing in cognitive computing in the form of *self-aware computing systems*. DARPA, the Defense Advanced Research Programs Agency sponsored the Workshop on Self-aware Computer Systems. Ron Brachman, then director of the DARPA IPTO program office, and now the president of AAAI, the American Association for Artificial Intelligence, spelled it out thusly:

> "A truly cognitive system would be able to ... explain what it was doing and why it was doing it. It would be reflective enough to know when it was heading down a blind alley or when it needed to ask for information that it simply couldn't get to by further reasoning. And using these capabilities, a cognitive system would be robust in the face of surprises. It would be able to cope much more maturely with unanticipated circumstances than any current machine can."

DARPA is currently supporting research on such biologically inspired cognitive systems.

IBM Research is offering commercially oriented support for cognitive computing through what it refers to as *autonomic computing*. The primary interest here is in self-configuring, self-diagnosing and self-healing systems.

A very recent and not yet fully developed trend in AI research is the move toward systems exhibiting a more human-like general intelligence, beginning to be called *artificial general intelligence* (AGI). The development of this AGI trend can be traced through a sequence of special tracks, special sessions, symposia and workshops:

- AAAI'04 Fall Symposium entitled Achieving Human-Level Intelligence through Integrated Systems and Research
- AAAI'06 Special Track on Integrated Intelligent Capabilities
- WCCI'06 special session entitled A Roadmap to Human-Level Intelligence
- CogSci'06 symposium on Building and Evaluating Models of Human-Level Intelligence
- AAAI'06 Spring Symposium entitled Between a Rock and a Hard Place: Cognitive Science Principles Meet AI-Hard Problems
- Workshop on Artificial General Intelligence

Such AGI systems being developed include LIDA, Joshua Blue, and Novamente.

AI and Cognitive Science
The science side of AI is devoted primarily to modeling human cognition. Its application is to provide hopefully testable hypotheses for cognitive scientists and cognitive neuroscientists. In addition to cognitive models with more limited theoretical ambition, integrated models of large portions of cognition have been developed. These include SOAR, ACT-R, CLARION, and LIDA. Some of them have been implemented computationally as software agents, becoming part of

embodied cognition. One of them, LIDA, implements several different psychological theories, including global workspace theory, working memory, perception by affordances and transient episodic memory. The importance of this cognitive modeling subfield of AI has be recognized by a few computer science department offering degree programs in cognitive science.

The core themes — where do they stand now?

Smart Software vs. Cognitive Modeling

As throughout AI history, both pursuits are still active in AI research, the engineering side and the science side. Currently, both are moving toward a more general approach. Smart software is beginning to including AGI. Cognitive modeling is moving toward more integrated hybrid models such as ACT-R, CLARION and LIDA, in addition to its traditional interest in more specialized models. Another major push on the smart software side is toward more autonomous software agent systems.

Symbolic AI vs. Neural Nets

Both symbolic AI and neural nets have survived their respective winters and are now flourishing. Neither side of the controversy has won out. Both continue to be quite useful. They are even coming together in such hybrid systems as ACT-R, CLARION and LIDA. ACT-R melds symbolic and neural net features. CLARION consists of a neural net module interconnected with a symbolic module. LIDA incorporates passing activation throughout an otherwise symbolic system making it also quite neural-net like.

Reasoning vs. Perception

Research into AI reasoning continues unabated in such subfields as search, planning and expert systems. Fielded practical applications are legion. Perception has come into its own in machine vision, agent based computing and cognitive robotics. Note that they come together in the last two, as well as in integrated cognitive modeling and AGI.

Reasoning vs. Knowledge

In addition to reasoning, knowledge plays a critical role in expert systems, and in agent based computing, self-aware computing and autonomic computing also. Again both are alive and flourishing, with the importance of adding knowledge to practical system ever more apparent. Data-mining has become another way of acquiring such knowledge..

To Represent or Not

Without representation, Brooks' subsumption architecture accords each layer its own senses and ability to choose and perform it's single act. A higher level can, when appropriate, subsume the action of the next lower level. With this subsumption architecture controlling robots, Brooks successfully made his point that much could and should be done with little or no representation. Still, representation is almost ubiquitous in AI systems as they become able to more intelligently deal with ever more complex, dynamic environments. It would seem that representation is critical to the process of action selection in AI systems, but

much less so to the execution of these actions. The argument over whether to represent seems to have simply died away.

Brain in a Vat vs. Embodied AI
For once we seem to have a winner. Embodied, or situated, AI has simply taken over, as most of the new research into AI systems is agent based. Perusal of the titles of talks at any of the general AI conferences like AAAI or IJCAI makes this abundantly clear.

Narrow AI vs. Human Level Intelligence
Narrow AI continues to flourish unabated, while the pursuit of human level intelligence in machines is gaining momentum via AGI.

Except for the strong move of AI research toward embodiment, each side of every issue continues to be strongly represented in today's AI research. Research into artificial intelligence is thriving as never before, and promises continuing contributions, both practical to engineering and theoretical to science.

References

Brooks, R.A., 1991. Intelligence without representation, Artificial Intelligence 47, 139–159.

Chalmers, D. 1990. Why Fodor and Pylyshyn Were Wrong: The Simplest Refutation. *Proceedings of the 12th Annual Conference of the Cognitive Science Society,* Hillsdale, NJ: Lawrence Erlbaum.

Church, A. 1936. An unsolvable problem of elementary number theory. *Amer. J. Math.* 58:345-363.

Davis, R., B. G. Buchanan and E. H. Shortliffe. 1977. Production Rules as a Representation for a Knowledge-Based Consultation Program. *Artificial Intelligence* **8**: 15-45

Fodor, J.A. and Pylyshyn, Z. 1988. Connectionism and cognitive architecture. *Cognition*, 28, 3-71.

Franklin, S., and A. C. Graesser. 1997. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In *Intelligent Agents III*. Berlin: Springer Verlag.

Holland, John H. 1975. *Adaptation in Natural and Artificial Systems*, Ann Arbor: University of Michigan Press.

Lindsay, Robert K., Bruce G. Buchanan, Edward A. Feigenbaum, and Joshua Lederberg. 1980. *Applications of Artificial Intelligence for Organic Chemistry: The Dendral Project.* Columbus, OH: McGraw-Hill,

McCulloch, W. S., and W. H. Pitts. 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5:115-133.

McClelland, J. L., Rumelhart, D. E., and the PDP research group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume II.* Cambridge, MA: MIT Press.

McCune. W. 1997. Solution of the Robbins problem. *J. Automated Reasoning*, 19. 263--276.

McDermott, John P. 1980. R1: The Formative Years. AI Magazine. 2. 21-29.

Minsky, M. 1954. Neural Nets and the Brain Model Problem. Ph.D. Dissertation. Princeton University.

Minsky, M. 1985. *The Society of Mind*. New York: Simon and Schuster.

Minsky, M., and S. Papert. 1969. *Perceptrons*. Cambridge, MA: MIT Press.

Newell, A., Shaw, J.C., Simon, H.A. 1959. Report on a general problem-solving program. *Proceedings of the International Conference on Information Processing.* pp. 256-264.

Post, E. L. 1943. Formal Reduction of the General Combinatorial Decision Problem. *Amer. J. Math.* 65:197-215.

Rosenblatt, Frank. 1958. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory, Psychological Review, v65, No. 6, pp. 386-408

Rumelhart, D. E., McClelland, J. L., and the PDP research group. (1986). Parallel distributed processing: Explorations in the microstructure of cognition. Volume I. Cambridge, MA: MIT Press.

Samuel, A. L. 1959. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Develop.* 3:210-229.

Smolensky, P. 1987. The constituent structure of connectionist mental states: A reply to Fodor and Pylyshyn. *Southern Journal of Philosophy*, 26 (Supplement), 137-63. [Reprinted in T. Horgan & J. Tienson (Eds.), 1991), *Connectionism and the Philosophy of Mind*, Dordrecht: Kluwer Academic. 281-308].

Turing, A. M. 1936. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 42:230-265.

Turing, A. 1950. *Computing Machinery and Intelligence. Mind, 59:434-60. Reprinted in: Computers and Thought, eds. E. Feigenbaum and J. Feldmans (1963)*. New York: McGraw-Hill.

Varela, F. J., E. Thompson, and E. Rosch. 1991. *The Embodied Mind*. Cambridge, MA: MIT Press.

Winograd, T. 1972. *Understanding Natural Language*. San Diego: Academic Press.

Zadeh L.A., 1965. Fuzzy Sets, *Information and Control*, 8. 338-353.

---

[i] While still a pure mathematician, your author spent some years on the Carnegie Mellon faculty where he knew both Simon and Newell. He learned no AI from them, a wasted opportunity.

[ii] Notice the lack of the expected citation here.

[iii] Searching Google with the key words "machine learning" yielded this message: "Google is looking for Engineering experts to join our team. Apply!"

[iv] One such, called Julia interacted so realistically that young men would hit on her.

[v] The author heads the Cognitive Computing Research Group at the University of Memphis.