Should You Be Teaching Artificial Intelligence?

John Favaro

Munich, circa 1991

Preface

Sometime in the early 1990s (I don't remember exactly when) I was invited to give a lecture in Munich to a congress of high school teachers. I decided to talk about a subject that was much in the news at the time – artificial intelligence – and consider the issue of whether this subject could or should have a place in a high school curriculum. I took the opportunity to explore some of the more interesting questions that were being debated both in engineering and philosophical circles at the time. The people and companies have changed since then (there are many today who don't know that Steve Jobs wasn't always the head of Apple!), but many of the issues are timeless and still worth thinking about.

Introduction

Over the last couple of years, no less a celebrity than John Sculley himself, the Chairman of the Board of Apple Computer, has been barnstorming around the countryside, with his vision of the future of computers in higher education. It is an inspiring vision indeed. His speech begins as follows:

> We are privileged to live during an extraordinary time. It is the turning of an era. The world is in passage from the industrial age to the information age. This is a time of profound changes, in which the key economic resources in the world will no longer be capital, labor, and raw materials, but rather knowledge, individual innovators, and information.
>
> We have an opportunity that is given only to few generations in history. I believe that if we respond with our best creative energies, we can unleash a new Renaissance of discovery and learning .. How will the organizations designed to thrive in the 19th and early 20th centuries learn to contribute to the 21st? Only by reinventing themselves through refocusing on *individuals*."

Let's see now how Sculley proposes to do this:

> The transformation I am calling for -- shifting focus from the institution to the individual -- has a close parallel in history. In medieval Europe people were subservient to the institutions of the church and feudal hierarchies. Then came the Renaissance, which redefined the individual as the epicenter of intellectual activity. The many forces which converged to bring about the Renaissance galvanized around one key technology: printing. The rise of printing led with astonishing speed to an explosion of literacy. Today, we are in need of a second Renaissance, which like the first can also be galvanized by technology. We are on the verge of creating new tools which, like the press, will empower individuals, unlock worlds of knowledge, and forge a new community of ideas. These core technologies and the tools they support will help create a new environment of lifetime learning.

John Favaro
Via Gamerra, 17
56123 Pisa - Italy

Tel +39 050 55 60 74
Fax +39 050 6143 1145
john@favaro.net

www.favaro.net/john

And there you have it: the core of Sculley's grandiose vision of the future of computing in education is a set of new core technologies that he believes is coming of age now. Let's take a closer look at what this technical vision consists of:

By the turn of the century, Sculley predicts personal computers will house the type of technology and interface design that will seem like descendants of today's UNIX and Macintosh systems. He calls this futuristic PC the *Knowledge Navigator* and has described what he feels are the five emerging core technologies that are crucial to its success:

1) The first feature is advanced telecommunications technology that can link processors and databases around the world, thus providing better vehicles and broader information pathways.

2) Secondly, he expects real-time, 3-D color animations to become commonplace as users rely more on graphic simulations. The purpose of these capabilities is to allow scientists to visualize their complex models as easily as we can now graph a spreadsheet.

3) He sees improved database technology as the key to creating intuitive and responsive information systems.

4) Fourth is hypermedia, which will give future PC users more intuitive ways of navigating through enormous collections of information while combining text, graphics, sound and motion.

5) Finally, rounding out the essentials is artificial intelligence technology -- which Sculley finds critical to the future vision of personal computing. AI will allow future users to create agents that can recognize and anticipate strategies and preferences as well as increase productivity."

What are these "agents" Sculley is talking about? He says "These agents will transform the nature of academic computing. Agents will be sent to prowl among remote databases and bring back the specific information and citations that the user requires."

So we are given a vision of a kind of brave new world of library research where artificial intelligence lights the way for us.

[*Interruption from the perspective of early 2011: Sculley didn't do badly at all in his predictions of twenty years ago. The first four points have definitely come to pass – in particular, he anticipated the Web revolution with his mention of hypermedia. And it could be argued that the fifth point has been realized at least partially (think of all the automatic facilities for recognizing preferences and making recommendations that you see at Amazon.com, for example), although not always in the same form that he predicted.*]

The Fifth Generation

Now let's look at another vision of a brave new world, which came from a couple of researchers at Stanford University not too long ago and has had major implications for

the world-wide computer trade war. The basis of this vision was the book called *THE FIFTH GENERATION: Artificial Intelligence and Japan's Computer Challenge to the World*. The Japanese have begun a massive multi-year program bent on producing a new - a fifth - generation of computers that will outstrip anything built before, and the cornerstone of the great leap forward that they plan to take is artificial intelligence.

The authors wrote:

> The world is entering a new period. The wealth of nations, which depended upon land, labor, and capital during its agricultural and industrial phases -- depended upon natural resources, the accumulation of money, and even upon weaponry -- will come in the future to depend upon information, knowledge, and intelligence.

That is, knowledge, stored and manipulated by intelligent computers, would become "a saleable commodity like food and oil", a new wealth of nations.

This vision sounds similar to Sculley's, doesn't it? In each case, artificial intelligence technology (also known as "**A.I.**" in the branch) is at the forefront of an enormous expansion of knowledge.

In the case of the Fifth Generation, the systems in question are not Sculley's intelligent database "agents", they are the so-called *expert systems* that have become so popular recently.

Most of you have probably heard of the expert systems by now, maybe even seen some. You'll recall that the purpose of an expert system is to capture the knowledge that a human expert in some field has. They have been developed in fields where there are human experts with very specialized knowledge, like medicine, or law, or engineering.

As important as the technical vision of the future of computing represented by expert systems is the commercial vision of the future! For, they are the first artificial intelligence systems to have a great success on the *commercial* market.

Here is what a brochure says for one of the commercial expert systems on the market says: "Expert systems reduce waiting time, staffing requirements and bottle-necks caused by the limited availability of experts. Also, expert systems don't get sick, resign, or take early retirement."

Companies are being started to create and market expert systems, and young engineers are being hired out of school to staff these companies. In fact, you can hear complaints now and then about the shortage of young engineers with the proper training to help with the creation and maintenance of these complex, artificial intelligence-based systems.

What do these visions of John Sculley and the commercial success of expert systems and the Fifth Generation mean for you? They mean that, slowly but surely, artificial intelligence is becoming a branch of computer science that can no longer be ignored by educators charged with giving their students their first exposure to the world of computers.

But just exactly what are you to tell your students about artificial intelligence? Is the picture really as rosy as it is painted above? Anybody who has followed the history of artificial intelligence over the years knows that things are not always as they seem, and I'd like to illustrate that by reading a review that was made of that very same inspirational book *The Fifth Generation*.

A prominent AI researcher Johan de Kleer wrote in a scathing review that the book exaggerated the possibilities of expert systems, as though the shallow rules contained in such systems were all there is to intelligence. Above all, it is the attitude exhibited in the book that bothers him the most:

> The book is anti-intellectual and self-aggrandizing, giving the impression that this small corner of the AI world constitutes the whole field, and that the difficult problems of mechanizing intelligence are on the verge of begin solved.

Then de Kleer moves in for the kill:

> AI researchers have long been viewed as arrogant, and this book only exacerbates this view.

So where does this statement leave us? Does it mean that AI is nothing more than an exercise in arrogance with no real technical foundation in science? Is it therefore better left untouched, at least for now, by educators such as yourselves?

To answer that we need to get a better handle on what the purpose of introducing AI into the curriculum would be, and that must begin with some effort to get a technical definition of what AI is.

One of the best descriptions of what AI is all about has been written in the context of another application of artificial intelligence. This application was the Strategic Defense Initiative, which is of course also known as "Star Wars". This is the network of killer satellites and beam weapons that is supposed to set up the ultimate shield against nuclear weapons.

One of the essential technical strategies that has been proposed for the Battle Management Software of Star Wars has been the use of artificial intelligence and expert systems.

One of the most respected computer scientists, a professor named David Parnas, surprised everybody by resigning his rather prestigious membership in the so-called "SDI Panel on Computing in Support of Battle Management". This is what he wrote in his letter of resignation:

> I am resigning my membership in the panel. I do not believe that further work by the panel will be useful and I cannot, in good conscience, accept further payment for useless effort. My conclusions are not based on political or policy judgments. *My conclusions are purely technical.*

Let's take a look now at Parnas's "purely technical considerations". He begins by taking a little swipe at AI:

> Researchers in AI have often made big claims, and it is natural to believe that one should use this technology for a problem as difficult as SDI battle management.

Then comes the interesting part. He addresses the question: *What is artificial intelligence?* He notes: "Two quite different definitions of AI are in common use today.

AI-1: The use of computers to solve problems that previously could be solved only by applying human intelligence.

AI-2: The use of a specific set of programming techniques known as heuristic or rule-based programming. In this approach human experts are studied to determine what heuristics or rules of thumb they use in solving problems. In other words, the program is designed to solve a problem the way that humans seem to solve it.

Let's think about those definitions for a minute. The first definition, of AI-1, has a sliding meaning. In the Middle Ages, it was thought that arithmetic required intelligence. Now we recognize it as a mechanical act and we build pocket calculators to do it for us. Something can fit the definition of AI-1 today, but, once we see how the program works and understand the problem, we will not think of it as AI any more.

AI-2 more or less corresponds to the expert systems we talked about before, that mimic human beings.

Now let's see how Parnas judges each of these approaches:

> I have seen some outstanding AI-1 work. Unfortunately, I cannot identify a body of techniques or technology that is unique to this field. When one studies these AI-1 programs one finds that they use sound scientific approaches, approaches that are also used in work that is not called AI. People speak of AI as if it were some magic body of new ideas. There is good work in AI-1 but nothing so magic it will allow the solution of the SDI battle-management problem.

And now he discusses AI-2, and he doesn't mince words:

> I find the approaches taken in AI-2 to be dangerous and much of the work misleading. The rules that one obtains by studying people turn out to be inconsistent, incomplete, and inaccurate. Heuristic programs are developed by a trial and error process in which a new rule is added whenever one finds a case that is not handled by the old rules. This approach usually yields a program whose behavior is poorly understood and hard to predict. AI-2 researchers accept this evolutionary approach to programming as normal and proper. I trust such programs even less than I trust unstructured conventional programs. One never knows when the program will fail.
>
> On occasion I have had to examine closely the claims of a worker in AI-2. I have always been disappointed. On close examination the rules turned out to handle a small number of obvious cases but failed to work in general. The author was able to demonstrate spectacular behavior on the cases that the program handled correctly. He marked the other cases as extensions for future researchers. In fact, the techniques being used often do not generalize and the improved program never appears.

Then he finishes off with a scathing remark:

> Artificial intelligence has the same relation to intelligence as artificial flowers have to flowers. From a distance they may appear much alike, but when closely examined they are quite different. I don't think we can learn much about one by studying the other. AI offers no magic technology to solve our problem. Its techniques do not yield systems that one can trust.

Artificial Intelligence in Software Engineering

What you just heard was a technical appraisal of AI from an academic viewpoint. Now let me give you some thoughts from an industrial viewpoint:

When you teach these unpredictable AI techniques you are going against the grain of software engineering principles. Out in the industrial world there is no need for poorly understood programs whose behavior is hard to predict.

Let's look at some even more fundamental problems you're likely to encounter. AI systems require big computer systems. They require workstations like Sun. Often, they require special purpose hardware that was only built for AI purposes. There is little use for general purpose instructional IBM PCs or Macs in this world.

Good software engineering principles require modular, well-structured programming practices. There is a large body of knowledge and principles that have been built up in teaching good programming methodology -- AI techniques almost universally by-pass this body of knowledge. In most AI programming, it's still "every man for himself" today. If you teach AI you will be doing your students a disservice. AI is viewed with suspicion in the real world of industry.

AI is programmed with exotic languages that are outside of the mainstream. Languages like LISP and PROLOG. LISP can be extremely difficult to learn, and depends entirely on such advanced concepts as recursion that are generally only found in more advanced programming courses. It is totally unrelated to the kinds of languages you are most likely to introduce students to. PROLOG, which means literally "Programming in Logic", is just that. It requires an understanding of theory that advanced university programs only offer. The mainstream languages of computer science, both in industry and education, are languages such as Pascal and C, especially Pascal. You won't find these in AI programming. In fact, the AI practitioners actually even look down their noses at them.

But the reverse is also very true. Once I was interviewing for a job in Berkeley, California. The system they were building was a very standard application, in networking, I believe, and certainly had nothing to do with AI. The head of the company told me that they had programmed a certain part of their system in LISP for certain technical reasons, but that they were keeping it a secret from the outside world, because LISP had a reputation in industry as being the programming language of those crazy AI people. They actually had to hide their use of AI techniques from the outside world.

The software systems that are currently in use for creating AI programs are enormously complex, such as the expert system shells. The likelihood that your students will understand them is small. In short, you will be educating your students for an industrial world for which they will be ill-prepared and unwelcome.

Why should you teach AI, from a technical point of view?

Given this grim picture I just painted of the technical arguments against artificial intelligence, is there any technical justification at all for introducing the subject?

AI programs are all about symbol manipulation -- as opposed to number crunching. Symbol manipulation is what computers are all about. It is what makes computers the most powerful, most exciting invention ever devised by man. You can do no better favor to your students than to make them aware of this. A personal story in this regard: From the time I was a boy I had always assumed that I would go into computers -- and this has a lot to say considering the time that was -- the 50's, when computers were also in their infancy. Then when I took my very first computer course, we were presented with accounting applications, calculating mathematical functions like square roots, real number crunching applications. I hated it. I was plunged into a personal crisis -- all my life I had been planning to go into computers, and I discovered I hated them. Was this all there was to it? Were computers nothing more than giant accounting machines, spewing out endless reams of mind-numbing calculations? I dropped out of the course -- my course load was too heavy at the time anyway, and I was only a freshman, so I put off worrying about the whole thing until the next year. It turned out to be a fateful, lucky decision, at least in my own life.

The next year I tried again. It just so happened that during that year the computer science department had had a change of personnel, and they had brought in Alan Perlis, one of the pioneers of computer science, as head of the new computer science department. It was this man who now taught the introductory computer science course, and you would not have recognized it from the year before. In his course he brought forth the great problems of computer science. Indeed, there was, quickly, one numerical problem, calculating some square roots. But we quickly moved on to symbolic problems, pattern matching, searching and sorting, language processing. Problems that elucidated the revolutionary nature of computers as symbol processors. No branch of computer science represents this better than AI; no branch has more potential for inspiration - I'll come back to this later.

For all the bad reputation suffered by AI among so-called "true" engineering branches of computer science, some of the most significant innovations in computer science have come out of AI research, from the lowest to the highest level. The early development of the LISP language brought us the first real treatment of the dynamic management of computer storage (also known as garbage collection) that is at the very foundation of most sophisticated applications you'll see today like text processors, *Hypercard*, window systems. Speaking of window systems, even these were pioneered by AI researchers working on the LISP systems at Xerox research in Palo Alto. These same LISP researchers gave birth to the so-called integrated programming environments that you see today in the form of systems like *Turbo Pascal* that young students like to use. In some sense, you can say that AI workers are our best and our brightest. They are attracted by the anything-goes atmosphere and sense of the great unknown, and they are the antithesis of what we might think of as the stolid engineer.

Or are they really our best and our brightest? Here is an extract from an article written by a Yale University AI researcher about the members of his profession. He entitled the article *Artificial Intelligence Meets Natural Stupidity*:

> As a field, artificial intelligence has always been on the border of respectability, and therefore on the border of crack-pottery. Many critics have urged that we are over the border. We have been very defensive toward this charge, drawing ourselves up with dignity when it is made and folding the cloak of Science about us. On the other hand, in private we have been justifiably proud of our willingness to explore weird ideas, because pursuing them is the only way to make progress.

> ... In a young field, self-discipline is not necessarily a virtue, but we are not getting any younger. In the past few years, our tolerance for sloppy thinking has led us to repeat many mistakes over and over. If we are to retain any credibility, this should stop.
>
> As AI progresses (at least in terms of money spent), this malady gets worse. We have lived so long with the conviction that robots are possible, even just around the corner, that we can't help hastening their arrival with magic incantations.

A Rational or Irrational Being?

When I took my AI course in college, the professor gave me a project that illustrated beautifully the difference between AI-1 and AI-2. It was the following: remember taking trigonometry in high school, where you tried to prove trigonometric equations by substituting known formulas, like tangent = sine/cosine. This is a classic case of AI-2, a set of heuristic rules for trying to solve a problem, without a guarantee of success.

It just so happens that there exists a straightforward mathematical solution to the problem (in case you're interested, it has to do with the fact that sine and cosine can be represented in the complex numbers -- don't worry, I didn't know about it either until he told me). He then gave me both of these tasks to program.

So I dutifully set about programming the straightforward, mathematical solution; and then I got out my old trigonometry book from high school and began building in all those rules of thumb that we had learned in Mr. Ryerson's class. When I finished, I watched in fascination as my AI-2 solution tried valiantly to solve the trigonometric formulas like a person, sometimes backing itself hopelessly into a corner and throwing up its arms in despair. Then my mathematical, AI-1 program would calmly take the same formula and churn out a definitive answer in a short and guaranteed time like a robot.

I've never forgotten the lessons learned from that project -- they are the fundamental lessons of the computer revolution; indeed they lie at the heart of all scientific inquiry. For, somewhere deep inside the mind of man lies the question of his true nature, as a rational or irrational being.

A book has recently appeared, whose title in French is *Histoire des Moeurs* (Volume 1). In English we loosely translate that word *moeurs* into customs, manners or morals. The official English translation is *A History of Human Behavior*. It is the attempt to give a rational answer, outside the teachings of religion, to the question "What is man?"

A recent review of this book discussed its approach, only to reveal those very same principles that I studied in my humble little trigonometry project - on the one hand my scientific, mathematical AI-1 solution; and on the other, my "human", "intelligent", "imaginative" AI-2 solution.

The book deals with man as a cultural animal, organizing himself in relation to his environment. Many contributors in the book discuss various communities in various kinds of settings throughout the world. They show in much detail how each community has a particular appreciation of the terrain it occupies. In every area of their physical life they have preferences or taboos, which have evolved throughout history but which remain different from the others.

However, an important point that is reiterated again and again is that these customs relating to the physical world, while governed to a large extent by material circumstances, are ultimately regulated by the imagination.

Whereas the non-human animals are strictly bound by their immediate instincts and their habitat, man has no settled nature or particular setting. He is nowhere or everywhere at home, thanks to his imagination, which adapts the world to his purposes and clothes it with transcendental meanings.

The reviewer points out the irony of this, by the way. Rational analysis leads to the conclusion that man is primarily non-rational. The author of the book suggests the *sense of the sacred* as the universal human-defining factor here. But it is pointed out that ever since human civilization began with the Neolithic revolution around 10,000 years ago, the basic trend has always been - despite many violent throwbacks - towards the replacement of the sacred or gratuitous accounts of phenomena by scientific reasoning.

> No one now believes, as did the Aztecs, that the movement of the sun depends on human sacrifice. And in Europe, since the development of modern knowledge from the Renaissance onwards, the sacred has always, in the long run, had to accommodate itself intellectually to the scientific, not vice versa.
>
> Every time the Pope rides in an airplane, he vindicates Galileo against the Church, and he would have been killed in the assassination attempt of a few years ago if it hadn't been for the hated *libido sciendi* approach to the human body that underlies modern surgery.

The review ends by noting

> This is not to say that scientific reasoning can replace the play of the imagination. Man is a rational being only in short spurts, from time to time. Besides, each new generation is born ignorant of the processes of reason, which have to be learned. The vital problem is how to ensure the healthy functioning of the imagination, our main mode of being. We need a rational assessment, in the light of history, of the proper role of the imagination in the scientific, post-industrial culture of today.

But that is exactly the point of the AI-1 versus AI-2 controversy. With the entry of computers, the exclusive domain of the imagination has been invaded. But the advent of the computer and artificial intelligence has accelerated the process of blurring the distinction between pure products of the human imagination and mechanical processes.

The AI-1 versus AI-2 controversy has also inflamed the philosophical world in its attempts to come to grips with the problem. The most famous exponents of the anti-rationalist argument are two brothers who are philosophy professors at the University of California at Berkeley, Hubert and Stuart Dreyfus. Let's hear what they say about whether thinking can be reduced to computation:

> The mistaken view that human skill and expertise is produced by complicated inferences drawn from masses of facts and beliefs is rooted in the thought of Socrates and Plato ... and has flowered in the era of the computer and artificial intelligence. It now permeates our society. Expert legal testimony must take the form of facts and inferences rather than experience-based intuitive judgments, although the latter are superior. Environmental impacts must be modeled, not intuitively assessed based on prior observations. If school children cannot explain why they know something, they are accused of guessing or cheating. Doctors, once trusted and admired for their wisdom, now attempt to rationalize their diagnoses and recommended therapies. Politicians impress voters, not be a record of sound judgment, but with factual knowledge and debating prowess. The list is endless.

> Intuitive expertise, acquired through concrete experience, is an endangered species. We must resist the temptation to exalt calculative reason as personified by the computer. Instead, we must recognize that facts, rules and logic alone can produce neither common sense, the ability to go to the heart of a problem, or intuition, our capacity to do what works without necessarily knowing why. Only if we recognize and appreciate the unique ways in which human beings transcend any reasoning device will machines *that* think become what they rightfully are, subservient aides and assistants to people *who* think.

Is it really that simple? I'd like to take up two threads of their statement. The rationalizing of the doctor's diagnosis and the expert legal testimony. Philosophy is one thing -- the complications of the real world is another.

The MYCIN program is the most famous expert system of them all. It diagnoses meningitis and blood infections.

The MYCIN program has been acknowledged by human experts, to be able to make medical diagnoses -- in its area of specialty -- better than any humans. Yet, the MYCIN system -- in spite of being acknowledged as better than any human being -- is still not used in one single place. Why is that?

It is because of the matter of legal responsibility. Nobody has been willing to take on the legal responsibility for diagnoses made by a computer program. No doctor has been willing to place his reputation behind the program. Nor have lawyers been willing to back up the statements made by legal expert systems in a court of law.

As AI begins to move out of the laboratory and into our lives, the political and moral implications of thinking machines will become as important as the scientific questions. When we live in a world where artificial systems make decisions, who will take responsibility for bad judgments. Can a computer program be held legally and morally accountable? This is what the philosopher Margaret Boden of the University of Sussex once had to say about it:

> Hackers and laymen alike constantly refer to programs ... in psychological terms. We speak of their **reasoning**, **judgments**, **evidence**, **knowledge**, **ignorance**, and **mistakes**. We speak of what they are trying to do,. and what priorities are guiding their decisions. It this simply sentimentality, a sloppy way of speaking which can and should be avoided -- above all, in the law courts? If it is not, if people as a matter of fact do not or cannot avoid using such terms in conceptualizing AI systems, then what implications follow? If we are allowed to use some psychological words when describing AI programs, why not all? If we use the language of **knowledge** and **inference**, and even **choice**, then why not the language of **purpose**, **effort** -- and even ***blame***?

For something to be blamed, it must, in some sense, be considered a *person*. Over the years, societies have extended personhood to previously excluded groups such as blacks and women. Some people would like to see it granted to fetuses. In many ways corporations are considered persons in the eyes of the law. So why not intelligent machines as well? As attorney Marshal Willick pointed out, whether a computer is responsible for its actions is as much a legal as a philosophical question, and ironically, for all the philosophical questions surrounding artificial intelligence, the question of whether a program is intelligent might be decided by a judge and a jury, as the result of a lawsuit filed by someone who feels damaged by a negligent machine. This might happen sooner than we think, Willick said. After all, for a machine to be held

responsible in a court of law, it probably wouldn't have to be of superhuman intelligence.

In fact, proving that a computer should be legally considered a person might simply be a matter of showing that it is not dead:

"When a person dies, he loses all of his rights", Willick explained. "The law in this area tends to set an over-inclusive minimum, so that any human but one who can be shown to have died tends to be defined as 'alive'. Given the recent emergence of 'brain death' as a critical factor, and since many computers today can exhibit far more 'intelligent' behavior than that of comatose human beings (who do enjoy legal recognition), a legal minimum standard test of personality could probably be satisfied by a computer system in the proper circumstances.

"The emergence of the modern corporation provides the most subtle means by which computer systems might achieve legal recognition," Willick continued. "Corporations have names, can buy and sell property, and can commit crimes, but they cannot be drafted, be married, or vote. They are persons, but they are owned."

Corporations, Willick explained, demonstrate the concept of *partial personality*.

> "This concept is applied in many ways in modern society. Minors, for example, slowly acquire rights and obligations as they grow older ... while rights are removed from the retarded and the insane ... The legal system is thus equipped with a variety of approaches with which to decide the extent and variety of rights that should be given to computers that are recognized as persons.
>
> "Computer systems that perform increasingly complicated tasks in an increasingly competent manner will be thrust onto these drifting sands of constitutional presumptions, tests, and standards. Since there does not seem to be an analytically sound test of 'personality' that will exclude computer systems which behave intelligently, the question of legal recognition will remain one of 'when' and not 'if', until and unless some absolute limitations on the abilities of such machines can be demonstrated. Once computer systems can satisfy established legal tests of personality, either a valid ground of distinction between them and humans will have to be found, or the distinction will have to be abandoned as mere prejudice."

Conclusions

So, should you be teaching AI?

YES: you will be introducing your students to one of the key technologies of the future, which has already begun to permeate every aspect of our lives. No one will remain untouched in the years to come. Driven by the most brilliant and farsighted researchers, artificial intelligence will ride at the crest of the next wave of the computer revolution. The engineers that are born in your classes will no longer be mere software or electronic engineers - they will be knowledge engineers, dealing in the currency of the coming industrial struggle between Europe, America and Japan. Your students must be prepared, both technically and philosophically. They must understand the great moral, ethical and legal issues that will be thrown up at us as artificial intelligence takes its part alongside human intelligence in society.

NO: AI is an overrated collection of toys, produced by an overrated collection of egotistical zealots who entered the field of artificial intelligence because of greed and a fear of dealing with something with a sound scientific basis. The techniques and tools of AI are at odds with everything your students will need in their future relationship to computers. You would be giving them a fantastical, Buck Rogers video-game vision of the computer world whereas the industry requires a clear-eyed, sober assessment of how the computer fits in as a tool in mankind's world.

I hope I have succeeded in dividing your loyalties in this talk. That was certainly my intention. There will probably never be a right answer to the questions raised here, at least not in our lifetime. But one thing is for sure: artificial intelligence cannot be ignored anymore. I hope I have given you some interesting things to think about, but in the final analysis, it is up to you as educators to make up your own minds as to where you stand, and to act accordingly to present the issues to your students in the best way you know how.