On the Current Paradigm in Artificial Intelligence

Submitted by Nello Cristianini on Thu, 06/09/2012 - 12:24

Nello Cristianini - (Draft of article prepared for AIComm issue on History of AI)

The field of Artificial Intelligence (AI) has undergone many transformations, most recently the emergence of data-driven approaches centred on machine learning technology. The present article examines that paradigm shift by using the conceptual tools developed by Thomas Kuhn, and by analysing the contents of the longest running conference series in the field. A paradigm shift occurs when a new set of assumptions and values replaces the previous one within a given scientific community. These are often conveyed implicitly, by the choice of success stories that exemplify and define what a given field of research is about, demonstrating what kind of questions and answers are appropriate. The replacement of these exemplar stories corresponds to a shift in goals, methods, and expectations. We discuss the most recent such transition in the field of Artificial Intelligence, as well as commenting on some earlier ones.

Introduction

The expression “paradigm shift” was introduced by philosopher of science Thomas Kuhn in 1962, while analysing some common patterns found in the evolution of all scientific disciplines, and has since been abused in many ways, particularly within the language of marketing, which seems to describe any innovation in this way \cite{StructureScientificRevolutions}.

Instead, “paradigm shift” denotes a very specific phenomenon observed by Kuhn’s in the history of natural sciences. Rather than progressing by a steady accumulation of knowledge, all the sciences seem to proceed by alternating periods of “normal science” with abrupt shifts. Normal science is characterised by a puzzle-solving approach to research, whereby practitioners belonging to the same community apply or extend the current framework to various problems. They do not just share the same techniques and expectations: they share also the same tacit beliefs about the values, the direction and the goals of their field of research. These are encapsulated by the common acceptance of certain exemplar stories that illustrate how science should be done, in that area. These are called paradigms, and new problems are approached by looking for analogies with these paradigmatic examples.

Paradigm shifts – in Kuhn’s view - occur when a scientific community changes its values, goals, and methods, and this is reflected by the replacement of the success stories that are used to define the field. After a paradigm shift the consensus about what forms an interesting question, or a valid answer to a given question, may change. Scientists from the previous paradigm would not recognise the new problems and new solutions as part of what they consider as worthwhile or even legitimate science. Notice that these shifts do not need to be limited to major revolutions such as the Copernican one, but can involve much smaller scale transitions, as pointed out in the Postscript to the 1969 edition of “The Structure of Scientific Revolutions”.

A classical paradigm shift occurred – for example – when young molecular biologists started unravelling the genetic code in the 1950s and 1960s, bypassing a series of classical and widely taught priorities and methods in biochemistry. And another one came when physicists dropped deterministic laws to accept probabilistic statements from quantum physics. Neither shift pleased the previous generation, and it was not meant to.

Although Kuhn’s analysis was based mostly on the history of physical sciences, it has since been applied to many other disciplines, and I believe that it can also be applied to Artificial Intelligence, the quest for automation of intelligent behaviour. I will focus particularly on the latest of a series of paradigm shifts, and will try to describe the main assumptions behind the paradigm that has emerged in the past few years.

In AI we are now in a period of normal science, characterised by the widespread acceptance of a
On the Current Paradigm in Artificial Intelligence
Published on can you see a pattern? (http://www.see-a-pattern.org)

paradigm that can be called “data-driven” or “statistical AI” and can be exemplified by a set of very similar success stories ranging from statistical machine translation to computer vision. Pointing at those stories as examples to follow implies setting a research agenda, and defining the goals of the field. Current questions and answers in the field have the “puzzle-solving” flavour described by Kuhn.

The Current Paradigm of AI

Doing research in AI in 2012 is very different than just few decades ago. While the field was officially busy trying to create HAL 9000, a series of breakthroughs in applications created a strong business model for a part of AI. It is now possible to transcribe speech, translate text and recognize faces, to a sufficient extent that this can be used in real applications. The success stories of current AI do not involve sentient robots or Turing tests, but rather efficient processing of digital content, or the modelling of its consumers / producers. They do not involve the investigation of general cognitive capabilities but rather the reproduction of very specific behaviours.

This is precisely the kind of shift that Kuhn talks about and it has far reaching consequences. Success stories tend to act as templates for future research, to define for new researchers which new puzzles should be solved and how. The adoption of new stories changes the definition itself of the scope of a research field. And - importantly - all these breakthroughs were achieved by the use of the same set of techniques and by the same overall approach: that of statistical AI.

From spelling correction to face recognition, including question answering, machine translation, information retrieval, a series of problems in machine intelligence were (partly) conquered over the past decade by the deployment of data intensive methods. For example the simple use of word frequencies in a massive corpus can deliver very good spelling correction, without the need for any knowledge of morphology or other linguistic theory. Domain knowledge is hard to obtain and encode, while data is often freely available “in the wild”.

So, what are the exemplars, the expectations, the goals of the present paradigm of Artificial Intelligence? What do new students expect to be doing in their career? How do they expect that new challenges will be conquered?

There is a beautiful account of what it means to be doing research in this new paradigm. It can be found in the article “The unreasonable effectiveness of data” published in 2009 by three leading scientists working within Google Research \cite{UnreasonableData}. It reads like the perfect manifesto of a scientific paradigm in Kuhn’s definition, containing success stories, recommendations and directions of expansion. I am sure that I am not the only professor recommending this reading to his students.

“The biggest successes in natural-language-related machine learning - the authors say - have been statistical speech recognition and statistical machine translation. The reason for these successes is not that these tasks are easier than other tasks; (...) a large training set of the input-output behavior that we seek to automate is available to us in the wild.”

They continue: “Similar observations have been made in every other application of machine learning to Web data (...) for example (...) the task of scene completion: removing an unwanted, unsightly automobile or ex-spouse from a photograph.”

In this article (a reference to the celebrated article by Eugene Wigner on “The Unreasonable Effectiveness of Mathematics in the Natural Sciences” \cite{UnreasonableWigner}) the authors make a series of implicit and explicit points, by discussing those paradigmatic examples. One point - implicit - is that the fundamental goal of their research is to generate a desired behaviour in a system, not the way in which that behaviour is obtained. Another – explicit - is that useful behaviour can be generated by gathering vast amounts of data, and training an adaptive system to emulate the desired behaviour. Having shown this to work in the exemplars, the authors suggest this should be the approach to solve a series of other problems.

Important statements found in that article include: “invariably simple models and a lot of data trump more elaborate models based on less data”, “memorisation is a good policy if you have a lot of data” and “simple n-gram models or linear classifiers based on millions of specific features perform better
than elaborate models that try to discover general rules”. These are of course important experimental findings, but they are also a statement of methodology. Think for example how that article would have been read by researchers belonging to the previous paradigm, you may call it “knowledge-driven”, committed to the notion of generating behaviour by means of deduction from a set of axioms, and see how far the field has moved from that vision (we will expand on this in the next section).

The authors focus on how one could get more of the precious data: an important and natural question within this framework, but not one that would be recognised as mission critical by practitioners of the previous approach. If the solution of AI challenges can be achieved by squeezing vast masses of data, then the obvious problem becomes how to secure that kind of resource. But if the same behaviour had to be generated by deduction from a knowledge base, then the obvious next step would be to create increasingly powerful knowledge bases, for the inference engines to crunch (and this was indeed done, see \cite{CYC} ). The choice of paradigm does affect the list of open problems for a field.

This leads the authors to discuss some open problems. One involves automatically linking existing databases: “Of course, we’ll find immense opportunities to create interesting data sets if we can automatically combine data from multiple tables in this collection. This is an area of active research”.

They give for granted that it is the collection of the raw data, not the invention of new algorithms, or the collection of axioms, or abstract rules, or knowledge bases, which will deliver the intelligent behaviour that the field is pursuing. The article concludes with a statement of purpose for the field: “So, follow the data. Choose a representation that can use unsupervised learning on unlabeled data, which is so much more plentiful than labelled data. Represent all the data with a nonparametric model rather than trying to summarize it with a parametric model, because with very large data sources, the data holds a lot of detail.” (...) “Now go out and gather some data, and see what it can do”.

This perspective (which – by the way – I adopt within much of my own research) is a remarkable departure from a previous season. There was a time when it was the collection of domain knowledge, or the creation of new heuristics, that was meant to deliver us the holy grail of an intelligent machine. No more. All the hopes of this approach are on finding sufficiently large sets of data, on which relatively simple statistics can work their magic.

While I believe that statistical and data-driven approaches to AI are an incredibly promising area for AI, I am however painfully aware that this is just one of the possible paradigms, and as such it suffers from the slightly self-fulfilling property of emphasising the importance of the kind of problems that it can solve, while being sometimes blind to the importance of the kind of things that it cannot do. And I do see this as a possible risk for our field.

The Previous Paradigm of AI

As already observed, the present way of doing AI is very different from that of earlier days. One of the most prominent previous paradigms in this field – that we call “knowledge driven” for convenience – saw intelligent behaviour as the result of symbolic reasoning, where reasoning was in turn framed as a search problem. The earliest exemplar story defining the approach was that of the “Logic Theorist”. This was a program written in 1955 by Newell, Simon and Shaw, designed to deduce theorems of mathematics from a small set of axioms. Its success in re-proving many theorems from the Principia Mathematica was often cited as evidence that machines can perform “intelligent tasks” \cite{LogicTheorist}.

Theorems were proven by symbolic manipulations that gradually transformed the set of axioms into a desired proposition. The search was directed by “heuristics”, a term borrowed from the influential book of George Polya on the solution of mathematical problems. Of course just the very comparison of real world problems with the solution of mathematical puzzles contained a strong bias about the nature of intelligence, and the methods required to replicate it. Other programs were created, based on this general idea: that actionable knowledge should be deduced from general axioms by a deductive process. Within that perspective, it made perfect sense to interpret any failures as the result of insufficient axioms or imperfect heuristics. Research was
therefore naturally aimed at addressing those problems, as well as at solving other applied problems with the same template.

Based on that perspective (other classic examples included programs to play chess and similar games) it was natural to try and cast a diversity of other problems into the familiar template. The puzzle-solving activity of normal science was aimed at framing any sort of problems as cases of the general “reasoning as search” paradigm.

After all intelligence was a matter of reasoning, and reasoning was a matter of symbolic manipulation, as was made explicit by Herbert Simon, and more important implicit by the choice of success stories to transmit via textbooks. This approach did not deliver the kind of results that today we would consider to be successful. For example it did not deliver translation, nor summarisation, nor viable robot navigation. And it is not that this was not tried.

We should not be surprised to find a trace of this earlier season in the proceedings of IJCAI, the longest running conference series in this field. In the first few editions, starting from 1969, we can see articles such as: “Application of Theorem Proving to Problem Solving”, “PLANNER: A Language for Proving Theorems in Robots”, “Theorem Provers as Question Answerers”, “Robot Planning System Based on Problem Solvers” and “Problem Solving Approach In Data Management”.

A very diverse set of problems were shoe-horned into this template, despite the fact that many of them resisted the attempt. This is exactly what “normal science” looks like. But then – Kuhn points out – things always change. Some puzzles refuse to be solved, and some questions start looking less interesting. Maybe the general cultural climate changes, maybe the funding landscape does too. When success failed to materialise, the obvious analysis was that the set of axioms or the set of heuristics, were insufficient. And if general purpose systems of this type failed to materialise, maybe the goal of AI was to develop domain specific ones. Changing goals may be easier than changing methods.

Since the mid 1960s Herbert Simon’s student Ed Feigenbaum led the way towards one possible solution: adding increasing amounts of domain knowledge to symbolic deductive systems, to make their behaviour relevant to specific (if narrow) domains of reality. These were called expert systems. The approach worked in sub-domains of chemistry and medicine (highly influential and cited exemplar stories were the Dendral project and the Mycin system, \cite{dendral}). Dendral was able to solve some problems relative to spectroscopic data, and Mycin was able to recommend which antibiotic to give to a patient based on their symptoms. This second system was reported to give the correct advice 69% of the times, beating some human experts, and this fact was often repeated as a crucial example of the power of knowledge-driven expert systems.

This became known as the “knowledge principle”: that in order for a system to function in the real world, it needed large amounts of symbolic knowledge about the world. And the more complex its world, the more knowledge it was needed. “The power is in the knowledge” was the mantra of that research community. This led to an explosion in the need for increasing amounts of symbolic knowledge to be encoded to make these systems relevant.

The main lesson exemplified by the paradigm of Dendral and Mycin was that the power of AI algorithms (for a while AI was identified with expert systems) was not in the reasoning method they used but in the knowledge base they had. This also implied the problem of finding application domains where the same trick could be repeated: sufficiently small and formalised to allow for their basic knowledge to be represented symbolically. Searching for amenable problems is a classic marker of a period of normal science.

The official open problem however was how to assemble vast knowledge bases. The project Cyc was based on just that vision: started by Douglas Lenat in the 1980s had the goal to codify, in machine usable form, millions of pieces of knowledge that comprise human common sense. It is still under way, well outlasting the season of expert systems \cite{CYC}.

A paper \cite{dendral} summarising the Dendral project states a vision of AI that stands in stark contrast with that outlined in the “Unreasonable Effectiveness” paper discussed in Section 2: “The knowledge principle (...) is: A system exhibits intelligent understanding and action at a high level of
on the current paradigm in artificial intelligence

Published on can you see a pattern? (http://www.see-a-pattern.org)

competence primarily because of the specific knowledge that it can bring to bear: the concepts, facts, representations, methods, models, metaphors, and heuristics about its domain of endeavor. ” It was another statement of paradigm.

Then things moved, as we have seen in Section 1. Emphasis moved from knowledge, logic and reasoning to data, learning and statistics. Success stories like the ones mentioned in the “Unreasonable Effectiveness” paper started being passed on, from teachers to students, more often than stories involving rules, theorems and logical deductions.

Data became central in the new narrative, intelligent behaviour was acquired by the system through automatic analysis of vast amounts of data. This is how AI conquered a multitude of problems in computer vision, text processing, speech recognition, and so forth. And because of this centrality of data, machine learning became the lingua franca of AI researchers, and increasingly this was true for the specific version of learning that had enabled all those success stories: statistical learning. The languages of statistics and of optimisation took hold of discourse in the AI literature, replacing that of logic. After all, all the success stories seemed to be based on modelling the problem of intelligent inference as an inverse problem, solved by maximisation of some probabilistic quantity.

Visualising the Paradigm Shift

We can see a historical record of this transition from knowledge-driven to learning-driven AI by analysing the proceedings of IJCAI. By simply counting the frequency of the keywords contained in the articles presented in the 22 conferences that span the 44 years from 1969 to 2011, we see some obvious trends.

In order to select the keywords in objective way, I followed a method that I used in the past with time series of gene expression data: I measured the relative frequency of each word that appeared in the titles more than once, creating a time series of 22 time points for each such word. Then I ranked them by measuring the variance of the time series, so to identify those terms whose frequency varied the most in that period. I also repeated the same process by using the entropy of each time series, as a sanity check.

It turns out that the words that have the largest variance in this period are: “learning”, “knowledge” and “reasoning”, the first growing very fast, the other two declining. If we use the entropy as a method to rank keywords, we see the most active terms are: “representation”, “understanding”, “natural”, “reasoning” and “learning”, all marked by a decreasing pattern, except the last one. These results are shown respectively in Figure 1 and Figure 2, along with many other words which have been changing their relative frequency in IJCAI titles.
In Figure 3 I present also the time series for a set of words that I have hand-picked just based on my own curiosity. The trend is clear, in my opinion, with a complete set of concepts fading away: those related to reasoning, logic, problem solving, heuristic search. And another set of terms strongly emerging: those related to data, statistics and learning.
So we can see some “archaeological” evidence for the claim that a paradigm shift has taken place. Many of us will think that this evidence was not needed, as we are all familiar with the discrepancy between the AI we studied (and sometimes even teach) and the AI that we practice. But I could not resist the habit of having a data-driven component in my article. I also add here two word clouds, to illustrate the difference in zeitgeist at two IJCAI meetings that are 30 years apart: 1981 and 2011. The word clouds for all editions are available at "patterns.enm.bris.ac.uk/AI-history".
Conclusions

Science is a human enterprise. Now more than ever it is also a social enterprise, and its activities are...
shaped by social forces, much like other forms of culture. The difference between science and those other forms is that science should also be subject to other forces, first of all the scientific method with its emphasis on repeatability, un-ambiguity, peer-review, and all those other constraints that we know and cherish.

I have already described elsewhere how progress in our field has enabled us to solve problems that were not solvable by the previous generation \cite{NelloDataDrivenAI2010}. Now computers can translate, they can recognise handwriting and speech, sometimes also faces. The power of scientific method has obviously enabled this field to efficiently harness the key technologies to solve these problems. What the scientific method cannot do is tell us what are the important questions and the acceptable types of answers. This must be the result of social agreement. In this, science is not different than other forms of culture.

To what an extent is the new AI a redefinition of goals, rather than an achievement of older goals?

In his famous response to Wigner's article, Richard Hamming \cite{UnreasonableHamming} observed that our awe at the effectiveness of mathematics is not entirely well placed. "Almost all of our experience of the world - he said - does not fall under the domain of science or mathematics". His point was that obviously mathematics is powerful, but only at doing the kind of things that mathematics can do. And they are not that many. A problem with paradigms is their slight circularity. The boundaries of physics have been traced also based on this criterion: over the years various parts of the world were ‘expelled’ from the domain of physics precisely because they resisted mathematisation. Scientists are very good at restricting the universe of their discourse to what they can formalise. When they stop seeing the rest of the world, they might convince themselves that they can formalise everything. Could we convince ourselves that AI boils down to those problems that we can solve by data-driven approaches?

Just like our predecessors bent on casting all intelligent behaviour as the result of rule-based deduction, we are at risk of moving some interesting problem into a collective blind spot. How many problems are we going to shoe-horn into the current data-driven paradigm? And how many more reluctant problems are we going to overlook, just because we cannot shoe-horn them? How many of the current open problems will end up one day looking like some of the previous hubristic quests, just with data sets replacing knowledge bases and statistics replacing logic? This too will be a natural consequence of the paradigm shift we have just completed. So as we celebrate the amazing achievements of statistical data-driven AI, let us not turn our eyes away from those aspects of the field that have not been - and possibly cannot be - conquered by the sheer power of big-data. Or else we will end up working on a restricted version of our world, again, as when our field tried to reduce the entire AI enterprise to a rule-based problem-solving heuristic-driven search in symbolic space.

Is statistical data-driven AI a new solution to an older set of problems, or is it a shift in our focus, towards a different class of problems? It is of course a bit of both, and could not be otherwise. But while we should celebrate the first, the conquest of certain long-standing questions, we also should be wary of the second: our natural inclination to increase our focus on the kind of tasks that we can do with this paradigm, and to ignore those that defy its power. These will not go away, they will just wait for the next paradigm shift.

<table>
<thead>
<tr>
<th>Attachment</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>ecai-figure-1.png</td>
<td>135.2 KB</td>
</tr>
<tr>
<td>ecai-figure-2.png</td>
<td>127.87 KB</td>
</tr>
<tr>
<td>ecai-figure-3.png</td>
<td>126.7 KB</td>
</tr>
<tr>
<td>ecai-figure-4.JPG</td>
<td>72.58 KB</td>
</tr>
<tr>
<td>ecai-figure-5.JPG</td>
<td>98.06 KB</td>
</tr>
</tbody>
</table>

Source URL: http://www.see-a-pattern.org/?q=content/current-paradigm-artificial-intelligence