# AI in the Petroleum E&P Industry

## A talk given at the RCA Conwy on 21st September 2017 to the PESGB North Wales SIG by Barrie Wells, Geoscience Wales Limited

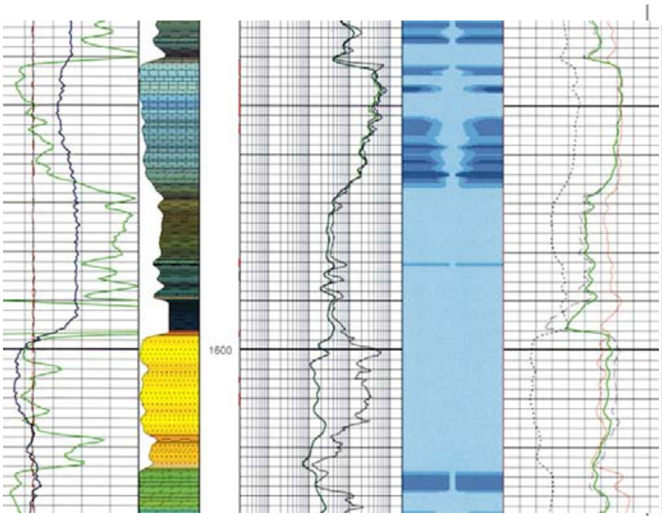
Clarke's First Law: "When a distinguished but elderly scientist states that something is possible he is almost certainly right. When he states that something is impossible, he is very probably wrong." (In the book Profiles of the Future: An Inquiry into the Limits of the Possible by Arthur C. Clarke). An example from E&P is (as quoted in the Daily Record recently) is "a senior BP person said he would drink every barrel of oil found in the North Sea." I am sure it was an Amoco executive; maybe the reporter was confused because BP bought Amoco. The context is supposedly that the exec thought the North sea was only gas-prone, not oil-prone, so his mistake is slightly more excusable and the reasoning more arcane than appears at first sight.

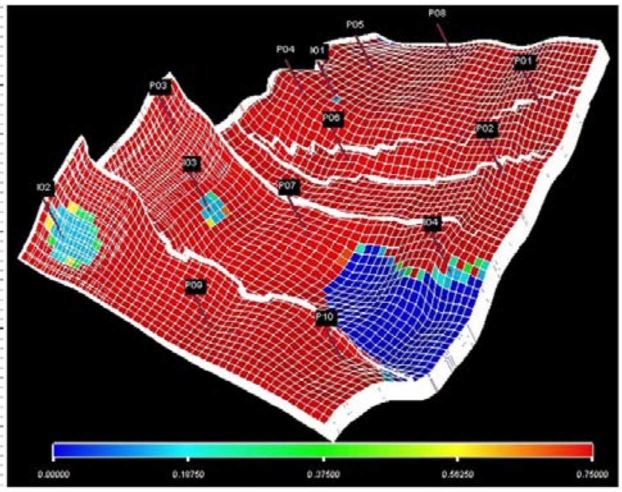
The brief history of AI has many examples of false prophecies of failure but also of false promises of success. In 1968, International Chess Master David Levy made a famous bet that no chess computer would be able to beat him within ten years. He won his bet in 1978 by beating Chess 4.7 (the strongest computer at the time) but it was only 3 years later when a computer first beat an International Master (Levy's strength) and gained that title itself. Chess was thought by early AI researchers to be a key indicator of intelligence and its mastery would demonstrate the intellectual superiority of computers. By contrast, we now realise that AI has taught us as much about the nature of intelligence as about the specific fields it has targeted. This was once characterised by saying that AI can beat humans at chess but cannot tie a shoelace; not an entirely accurate comparison but it has the right idea.

In 1990 the then-influential Byte magazine ran a special edition under the banner “Is AI dead?”. 27 years later I was motivated to give this talk by two main things: a PESGB London evening lecture in July 2016 taking the theme of AI in Stochastic Seismic Reservoir Characterisation, and a sub-set of events in the Finding Petroleum series, held at, although not under the auspices of, the Geological Society. As luck would have it, Theresa May gave AI a boost only 2 days before I was due to give this talk, by demanding that tech companies use AI to help counter the threats posed by terrorism.

To determine whether or not AI is (or was) dead, we need to know what it is. What is AI? "The primary goal of AI is to make computers smarter" ((Director of the MIT AI lab, 1987). This doesn't really help. What do we mean by "smarter"? Clarke's Third Law: "Any sufficiently advanced technology is indistinguishable from magic."

Both of these would have been considered “magic” less than 100 years ago, but why might one of them more reasonably be thought of as AI?




Viewing simulation results using ECLIPSE Office

In many ways, the reservoir simulation shown on the right is more advanced than the electrofacies on the left but I would contend that the wireline log interpretation is a better illustration of AI in the petroleum E&P industry: it involved rules (expert systems), data mining (analytics) and deduction (Machine Learning)

Any sufficiently advanced technology is indistinguishable from magic. Here the advanced technology is a combination of wireline logging and electronic computing. We can see how wireline logging works: the first logs were simple measures of resistivity. Neutron logs work on the billiard ball principle (an emitted neutron  will bounce off a larger object, pass through a smaller object and stop dead if it meets something its own size so, if we assume that the case of a smaller object simply delays the time when it will meet one of the other two, the number of same-sized objects determines how many neutrons bounce back. Hence we have a measure which Is a rough proxy for the amount of free radicals (i.e. water) in the formation). NMR logs rely on magnetic spin. And so on. There is no magic, it is simply applied physics.

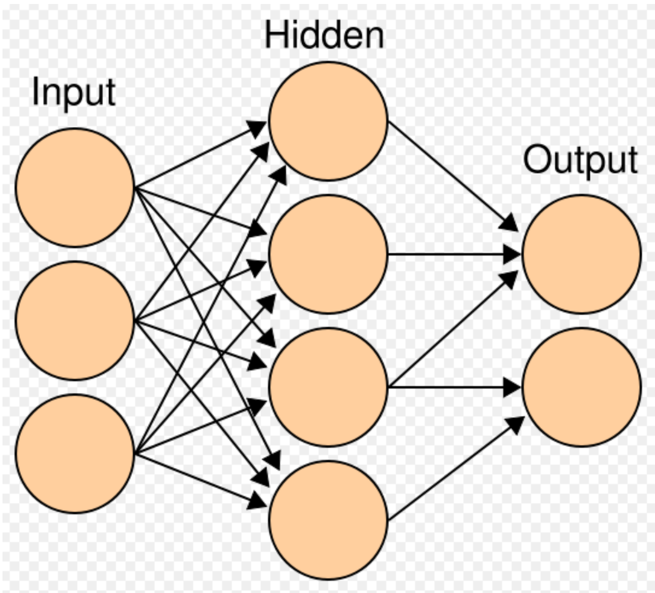
Archie could have formulated his ‘law’, and we could apply it, without the aid of computers, but they do make the task a lot simpler. Again, no magic. But as the task becomes more complicated, it can look like magic. However, the distinguishing feature of the log interpretation, in this context, is its reliance on rules; and a further claim to being AI is that the rules were derived both from data and from physics.

Initially, AI was synonymous with rules and the application of symbolic logic. Manipulating numbers pre-dated electronic computers: Babbage’s difference engines were commissioned by the Admiralty for calculating mathematical tables to be used in navigation, and many other calculating machines existed prior to the invention of the programmable electronic computer in the mid 20$^{th}$ century. The key breakthrough that led people to believe computers could be intelligent was not a faster way to do arithmetic but Turing’s exposition on symbolic logic, showing that computers could manipulate symbols as well as numbers.

Expert Systems and rule-based programming dominated applied AI in the 19070s and early 1980s. What happened to them? They became commonplace, embedded in everyday appliances: cars in particular are now dependent on them, as are production platforms. So AI is alive and well in the E&P industry, except that nowadays we don't call those old systems AI. This is the existential problem AI faces: once a "smarter" program has been implemented, it is just another piece of software, no longer AI. Someone at the first AAAI conference said "If we are still meeting in 10 years time, we will have failed", meaning they expected AI to move into the mainstream so it would no longer be a separate subject. The AAAI recently held their 47$^{th}$ conference, but AI hasn't failed, it has moved with the times, finding new ways to make computers smarter still, after each time they have just been made smarter.

Expert Systems were followed by fuzzy logic and data mining, which in turn have morphed into 'analytics'. Neural networks, one of the most recognisable AI tools, is embedded in tools such as Petrel and used routinely, whilst also leading to Genetic or Evolutionary Algorithms and Machine Learning..

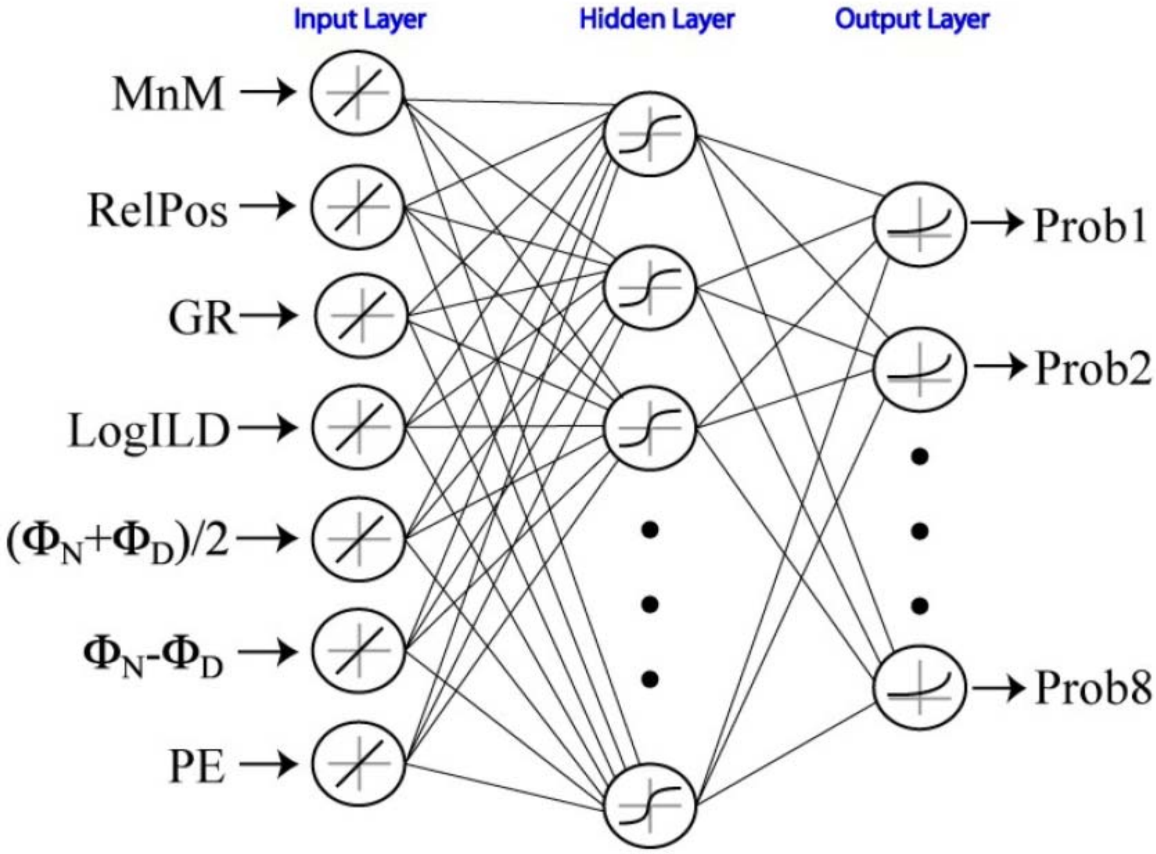
The standard diagram used to illustrate neural networks (NN) is



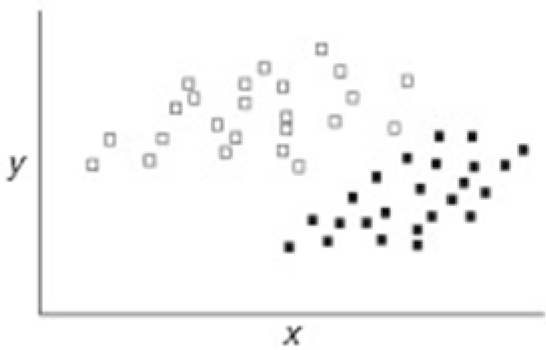
(Wikipedia)

which we could read as "we put data in, some magic occurs, and we get answers out". However, we are scientists, we know that magic doesn't exist, so we deduce that it's probably just sufficiently advanced technology (Deduction: "Daddy, there's a man at the door with a bill". "Don't worry, son, it's probably just a duck with a hat on").

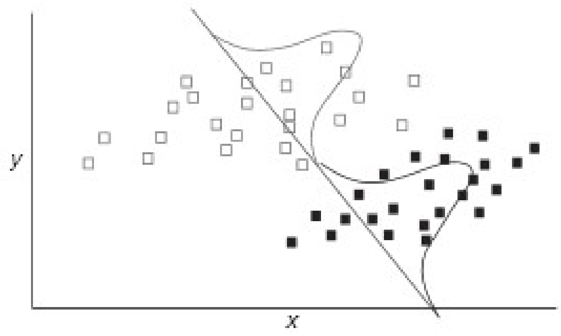
What is the technology? It is a variation on linear algebra: looking for relationships:




The software looks for linear relationships amongst the inputs and tries to find correlations and simplifications. For example, if we had this data:

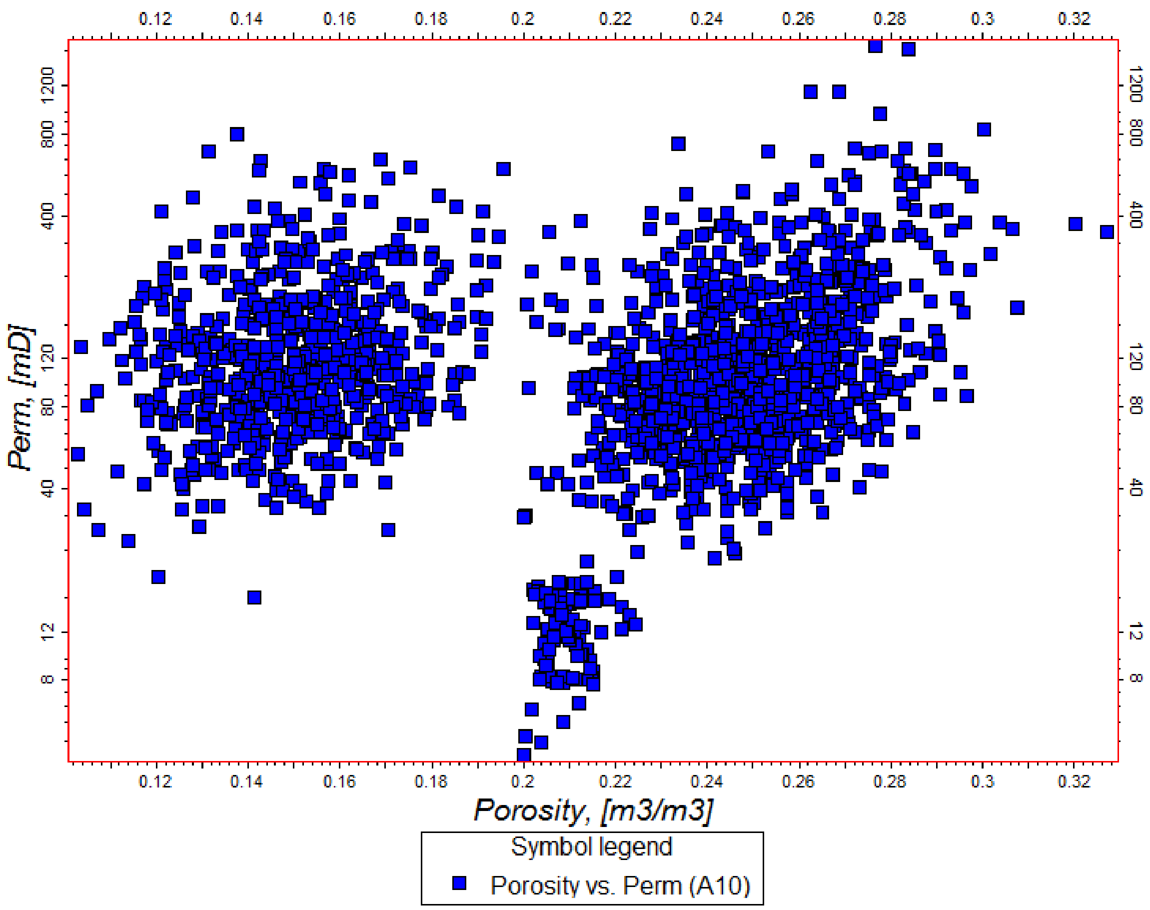



(where I have given the answer up-front by using two different symbols), the computer software could easily discriminate by changing the axes, a simple linear transformation:



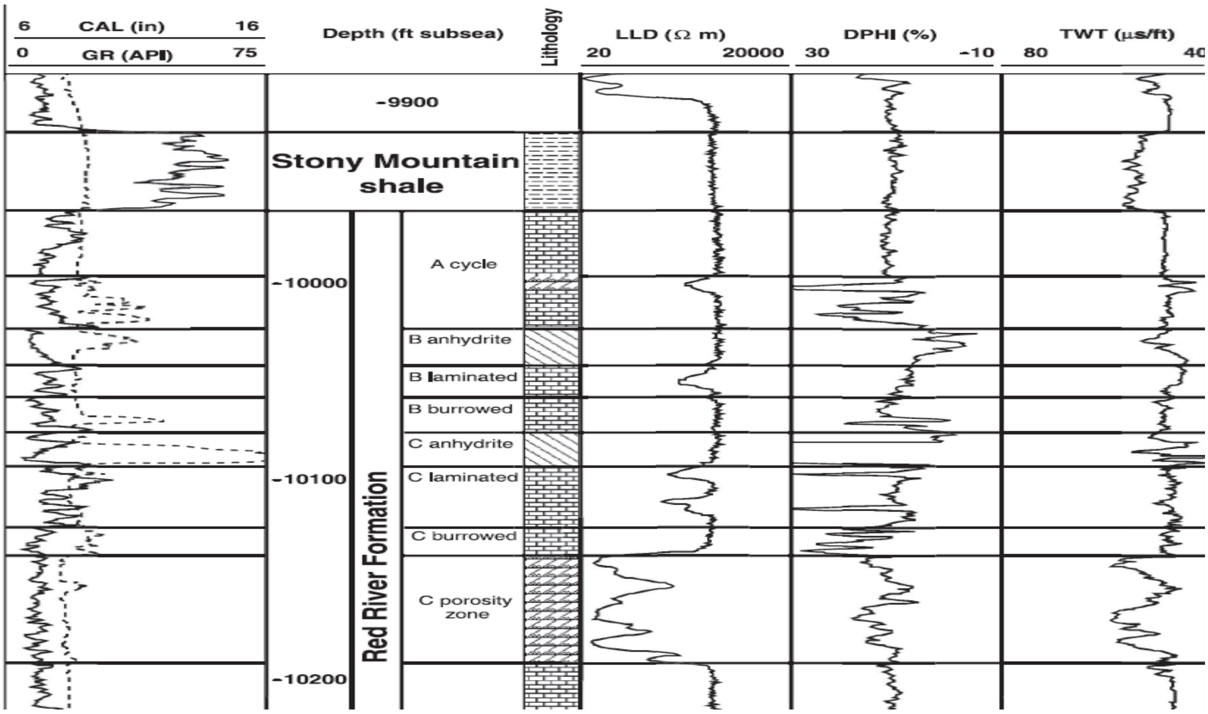

It could then find predictions in each region, using linear regression, thus reducing the complexity of the data, or reducing 50 pairs of numbers to one categorical variable and 4 constants. There will be some data loss, or error in the resultant predictions, because the linear predictors are not exact, but we can even quantify the error and say how good the predictions are. No magic, just maths.

Here I told the computer what the classification should be (in my opinion); we don't always have that luxury.

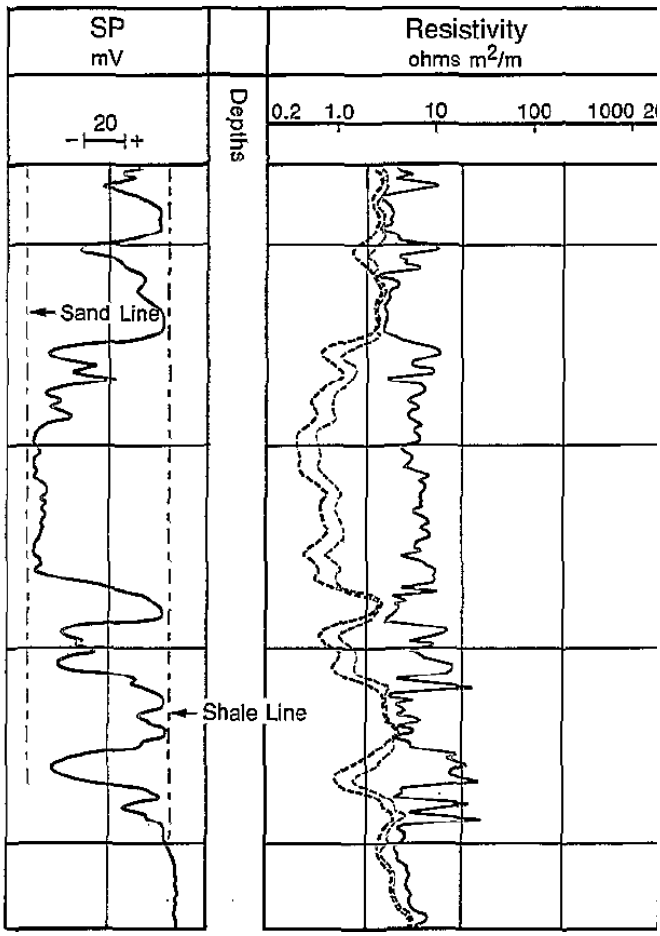



*(image courtesy Schlumberger, from Petrel)*

Here there could be two or three categories; or we could say there are hundreds, treating each point as its own category. Similarly, in wireline log examples we could provide either a set of categories or a stopping criterion:

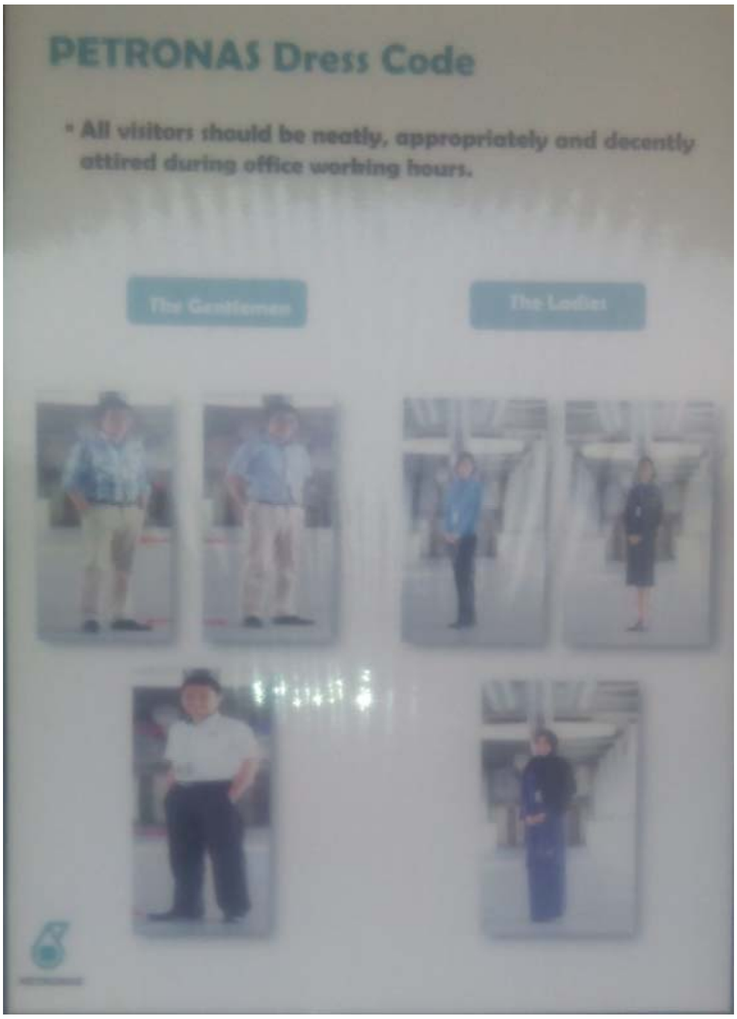



Here we have 4 categories (Shale, Anhydrite, Limestone, Dolomite); everything must be categorised as one of these four, and no other can be interpreted. We do not, however, have to impose our view, NN can work unconstrained:



In this case, if we want the computer to find rules for categorising wireline log traces in general, we need to state a stopping criterion, or else it will tell us that every reading is a separate class. This criterion doesn't have to be a fixed number, it could be based on a variance measure or, more sensibly, the rate of change of variance with number of categories, but we still have to impose our view on generalism versus specificity.

Geologists refer to lumpers and splitters; computers have no such predispositions, we have to tell them when to stop. Which brings us to racist AI: when computers are programmed by humans, they are potentially subject to our prejudices. If I program a computer to solve the equations of fluid flow (i.e. construct a reservoir simulator), I have to make decisions or judgements: miscible or immiscible flow? 2 phase or multi-phase? Include corner point geometry? Flow lines or D'Arcy flow? Transmissibility or permeability? This is a fairly simple set of decisions, we can base them on the purpose of the simulator and the available technology. Other decisions are less obvious. At present, the BBC feels compelled to invite Nigel Lawson (a climate change denier) onto panels "for balance", but doesn't invite representatives of the tobacco lobby to state that the jury is still out on whether smoking cigarettes causes fatal diseases. Is this bias?




Bias can creep in without our noticing. If we were to use the PETRONAS dress code for men and women (above) as part of the input (i.e. training data) to a NN system for identifying employees arriving at the company's gates, the software could quite reasonably conclude that the best way to distinguish men and women (a perfectly reasonable first step towards identifying specific individuals) is that women have visible hands whereas men do not.

The racism charge picked up by the tabloids probably has a different origin: let us say, for the purposes of this argument, that young people drift into crime due to a feeling of alienation from society and an excess of spare time caused by unemployment and lack of opportunities. Further, assume that an historic racial bias has disproportionately alienated one particular racial grouping and discriminated against that group in employment and in opportunities to join local golf and polo clubs. Then that racial grouping will probably be disproportionately represented in crime statistics. Hence, using the best available data to train a NN could result in the conclusion that one particular racial grouping should be targetted for heavy-handed police scrutiny; i.e. existing prejudices will be reinforced. If I tend to call ratty sands "sandstone", then my interpretations will train a NN to predict better reservoirs than yours, if you prefer to call ratty sands "mudstone".

So far, so predictable and understandable. We can see how the systems work internally and we can separate magic from science; we can ask Expert Systems to justify, or at least explain, their reasoning. But as AI systems become more complex, they inevitably become more 'black box', until eventually we have to take them on trust. This can happen in two different ways: the datasets may be very large (data mining, data analytics) or the variables we are using may not be readily interpretable (this is arguably specific to AI in E&P).

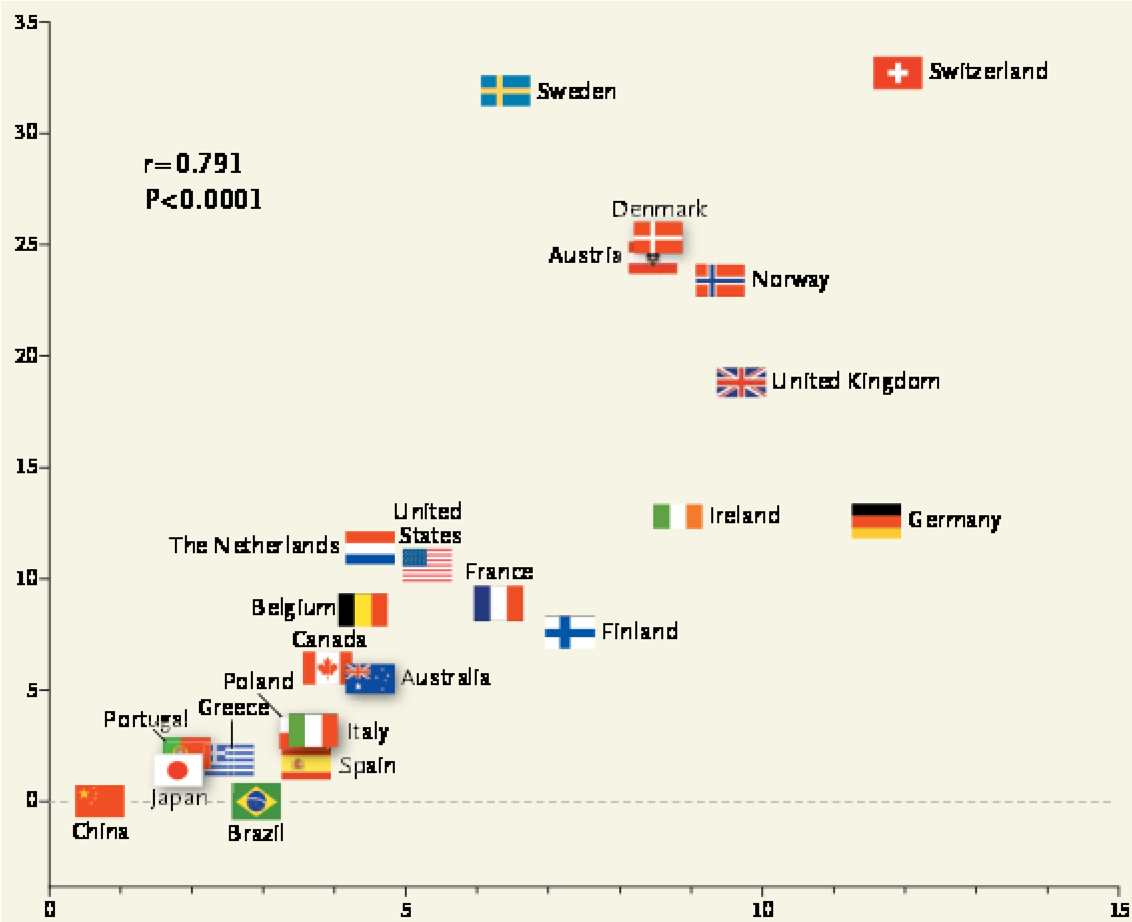
A salutary example of these phenomena was recently published in the New England Journal of Medicine. The health industry is well known for its penchant for throwing vast amounts of observed data into a statistical package and hoping something useful will emerge. The same tabloids that declare AI to be racist regularly proclaim that eating kale whilst standing on your head will cure warts whereas drinking water backwards cures hiccups. Such deductions are almost invariably the result of a large scale statistical study, which is what I presume motivated Messerli to publish this paper:

The NEW ENGLAND JOURNAL of MEDICINE

OCCASIONAL NOTES

**Chocolate Consumption, Cognitive Function, and Nobel Laureates**

Franz H. Messerli, M.D.




Horizontal axis: Chocolate consumption per 1000 population;
vertical axis: Number of Nobel laureates from that country.

The correlation (coefficient = 0.79) between chocolate consumption and number of Nobel laureates would only have occurred by chance about one time in ten thousand. If Sweden is treated as an outlier, the correlation coefficient increases to 0.86 (why might Sweden be an outlier?).

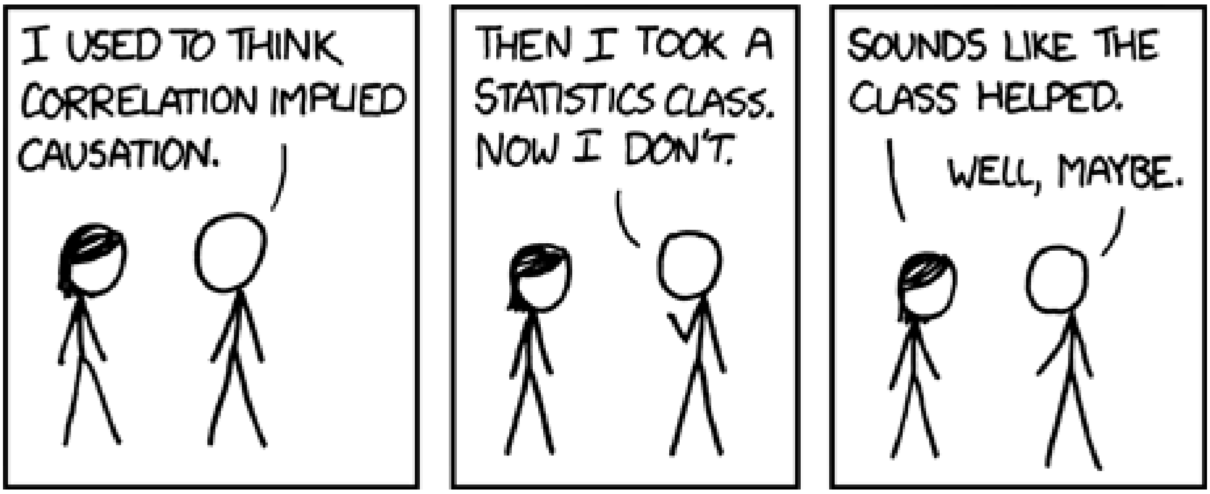
Here we have an obvious correlation but it does not seem to have a cause. When this occurs, it is often because there is a hidden third variable to which both parameters are correlated. In this case, it is likely to be GDP or some other measure of prosperity: If a country can afford to eat a lot of non-essential or luxury food (e.g. chocolate) then it can also afford to spend heavily on education and send a large proportion of its people to well-equipped universities where they have an increased chance of winning a Nobel prize.

An AI engine would spot this correlation and use it; the Daily Express would run a headline “Eating chocolate makes you brainy”. The problem is that, because the correlation is not actually between these two variables, it cannot necessarily be expected to continue. For instance, a well-educated country might *en masse* give up eating chocolate because of the health dangers of eating refined sugars, in which case the prediction would break down, but the AI engine would take a long time to find out and to correct its prediction, making very bad predictions in the meantime.

Alternatively, we may mistake the direction of causality. Here we drew chocolate on the horizontal axis, implying its independence, but why not posit that Nobel prize awards may increase chocolate consumption in a country, perhaps as a form of celebration? Better examples of this Reverse Causation are available. the advice to 'Avoid coffee when pregnant' assumes that caffeine is bad for the baby, but it arises from statistics rather than an expected cause, as there is no medical evidence that caffeine is bad for either an unborn child or an expectant mother. Rather, the advice is based on an observed correlation between those expectant mothers who drink coffee and those who have bad pregnancies. However, maybe it is reverse causality in the statistics. Nausea is related to good pregnancies: in a good pregnancy, the mother is likely to be nauseous, and nausea puts people off drinking coffee. So maybe the correlation between those mothers who drink coffee and those who have bad pregnancies is that a bad pregnancy doesn't discourage coffee consumption.

A simpler example can be found in statistics for accidents involving left-handed people. In English, slang words for ‘left-handed’ are often synonymous with ‘awkward’ or ‘clumsy’, and hence error-prone. However, people who are left-handed are not necessarily any more clumsy than the norm, although they may have more accidents; the higher incidence of accidents is almost certainly because the world is designed for right-handed people, the majority.

“Birthdays are good for you. Statistics show that those who have the most, live longest”




The publicity blurb for the “Finding Petroleum” series on “Solving E&P problems with Machine Learning & Analytics” includes the claim that “scatter plot and regression loyalists are falling way behind those that are applying advanced analysis techniques to identify hidden signals in their data” but doesn’t address the problem that these hidden signals might be misleading.

Another danger arises from not understanding what the parameters represent. This doesn't just happen to AI engines, people frequently make the same mistake. A peer-reviewed published paper proudly reported the prediction of porosity thickness from seismic attributes:

| Correlation coefficient | Attribute |
|---|---|
| -0.755 | Slope spectral frequency |
| -0.719 | Slope of reflection strength |
| +0.700 | Ratio of positive to negative samples |
| -0.683 | Average reflection strength |
| -0.683 | Average peak amplitude |
| -0.683 | Average absolute amplitude |
| +0.608 | First dominant frequency |
| +0.584 | Second dominant frequency |
| -0.575 | Root mean square amplitude |
| etc., for a all 20 attributes | |

An absolute value of correlation coefficient of 0.755 or greater would only occur 3% of the time, just at random, so at first sight this looks quite good, perhaps something we could use for prediction. But if we roll the dice 20 times, the chances of an unlikely event increase, and here we have correlated porosity thickness with 20 seismic attributes, so the chance of a 0.755 cc, just by chance, with at least one of the attributes, increases from 3% to 46% (i.e. probability increases from 0.03 to 0.46, or $(1 – (1\text{-}0.03)^{20})$). In other words, it was almost evens that at least one of the attributes would have a cc as high as 0.755, even if there were no correlation at all, i.e. if the data were purely random. So this isn't really anything to rush into print about, and certainly not the sort of evidence on which to base a multi-million dollar exploration project.

It is easy to set up an experiment in Excel to demonstrate this phenomenon:

| | ϕ.h | Y1 | ϕ.h | Y2 | ϕ.h | Y3 | ϕ.h | Y4 | ϕ.h | Y5 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.53 | 0.24 | 0.53 | 0.79 | 0.53 | 0.59 | 0.53 | 0.10 | 0.53 | 0.35 |
| 2 | 0.59 | 0.51 | 0.59 | 0.81 | 0.59 | 0.98 | 0.59 | 0.04 | 0.59 | 0.42 |
| 3 | 0.15 | 0.73 | 0.15 | 0.97 | 0.15 | 0.93 | 0.15 | 0.49 | 0.15 | 0.28 |
| 4 | 0.82 | 0.99 | 0.82 | 0.27 | 0.82 | 0.49 | 0.82 | 0.21 | 0.82 | 0.13 |
| 5 | 0.34 | 0.27 | 0.34 | 0.89 | 0.34 | 0.38 | 0.34 | 0.47 | 0.34 | 0.33 |
| 6 | 0.25 | 0.97 | 0.25 | 0.59 | 0.25 | 0.43 | 0.25 | 0.21 | 0.25 | 0.21 |
| 7 | 0.87 | 0.13 | 0.87 | 0.23 | 0.87 | 0.71 | 0.87 | 0.49 | 0.87 | 0.11 |
| 8 | 0.47 | 0.84 | 0.47 | 0.72 | 0.47 | 0.47 | 0.47 | 0.26 | 0.47 | 0.30 |
| Correlation | | -0.25 | | -0.80 | | 0.01 | | -0.18 | | -0.46 |

| | ϕ.h | Y11 | ϕ.h | Y12 | ϕ.h | Y13 | ϕ.h | Y14 | ϕ.h | Y15 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.53 | 0.29 | 0.53 | 0.88 | 0.53 | 0.58 | 0.53 | 0.61 | 0.53 | 0.92 |
| 2 | 0.59 | 0.27 | 0.59 | 0.55 | 0.59 | 0.25 | 0.59 | 0.41 | 0.59 | 0.58 |
| 3 | 0.15 | 0.81 | 0.15 | 0.33 | 0.15 | 0.37 | 0.15 | 0.64 | 0.15 | 0.09 |
| 4 | 0.82 | 0.85 | 0.82 | 0.80 | 0.82 | 0.32 | 0.82 | 0.28 | 0.82 | 0.08 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 0.34 | 0.49 | 0.34 | 0.01 | 0.34 | 0.01 | 0.34 | 0.90 | 0.34 | 0.51 |
| 6 | 0.25 | 0.69 | 0.25 | 0.16 | 0.25 | 0.55 | 0.25 | 0.93 | 0.25 | 0.51 |
| 7 | 0.87 | 0.84 | 0.87 | 0.99 | 0.87 | 0.83 | 0.87 | 0.79 | 0.87 | 0.88 |
| 8 | 0.47 | 0.75 | 0.47 | 0.76 | 0.47 | 0.01 | 0.47 | 0.77 | 0.47 | 0.72 |
| Correlation | | 0.11 | | 0.79 | | 0.34 | | -0.48 | | 0.28 |

| | ϕ.h | Y6 | ϕ.h | Y7 | ϕ.h | Y8 | ϕ.h | Y9 | ϕ.h | Y10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.53 | 0.83 | 0.53 | 0.66 | 0.53 | 0.71 | 0.53 | 0.22 | 0.53 | 0.58 |
| 2 | 0.59 | 0.80 | 0.59 | 0.68 | 0.59 | 0.80 | 0.59 | 0.45 | 0.59 | 0.30 |
| 3 | 0.15 | 0.42 | 0.15 | 0.28 | 0.15 | 0.59 | 0.15 | 0.03 | 0.15 | 0.26 |
| 4 | 0.82 | 0.59 | 0.82 | 0.31 | 0.82 | 0.93 | 0.82 | 0.78 | 0.82 | 0.73 |
| 5 | 0.34 | 0.76 | 0.34 | 0.44 | 0.34 | 0.67 | 0.34 | 0.22 | 0.34 | 0.42 |
| 6 | 0.25 | 0.07 | 0.25 | 0.80 | 0.25 | 0.23 | 0.25 | 0.36 | 0.25 | 0.28 |
| 7 | 0.87 | 0.92 | 0.87 | 0.14 | 0.87 | 0.18 | 0.87 | 0.37 | 0.87 | 0.43 |
| 8 | 0.47 | 0.56 | 0.47 | 0.06 | 0.47 | 0.87 | 0.47 | 0.74 | 0.47 | 0.01 |
| Correlation | | 0.63 | | -0.31 | | 0.10 | | 0.58 | | 0.51 |

| | ϕ.h | Y16 | ϕ.h | Y17 | ϕ.h | Y18 | ϕ.h | Y19 | ϕ.h | Y20 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.53 | 0.92 | 0.53 | 0.78 | 0.53 | 0.63 | 0.53 | 0.06 | 0.53 | 0.12 |
| 2 | 0.59 | 0.40 | 0.59 | 0.73 | 0.59 | 0.92 | 0.59 | 0.90 | 0.59 | 0.17 |
| 3 | 0.15 | 0.96 | 0.15 | 0.42 | 0.15 | 0.52 | 0.15 | 0.06 | 0.15 | 0.54 |
| 4 | 0.82 | 0.29 | 0.82 | 0.64 | 0.82 | 0.96 | 0.82 | 0.84 | 0.82 | 0.25 |
| 5 | 0.34 | 0.34 | 0.34 | 0.01 | 0.34 | 0.19 | 0.34 | 0.59 | 0.34 | 0.54 |
| 6 | 0.25 | 0.08 | 0.25 | 0.33 | 0.25 | 0.43 | 0.25 | 0.19 | 0.25 | 0.79 |
| 7 | 0.87 | 0.41 | 0.87 | 0.62 | 0.87 | 0.09 | 0.87 | 0.23 | 0.87 | 0.39 |
| 8 | 0.47 | 0.12 | 0.47 | 0.04 | 0.47 | 0.69 | 0.47 | 0.56 | 0.47 | 0.93 |
| Correlation | | -0.20 | | 0.52 | | 0.15 | | 0.41 | | -0.51 |

Max. Abs. correlation coeff. = 0.80

Mean correlation coeff. = 0.06

Whatever anyone might tell you, and try to blind you with algorithms to convince you they have harnessed magic in the service of data analysis, this search for patterns and relationships is the fundamental basis for data analytics: If we ask a computer to find patterns, it will find patterns. Whether they are meaningful, whether they arise for the reasons we assume, the computer has no way to judge, it is performing statistical analyses in an unintelligent way because that is the only way it has.

analytics (noun)

the systematic computational analysis of data or statistics.

information resulting from the systematic analysis of data or statistics.

We are therefore probably quite safe, in terms of losing our jobs to an AI bot, at least for the time being, provided we use analytics sensibly and provided we know enough statistical theory to outsmart the bots (and those humans who don't understand the nature of probability). Claims to "pull out of the many potential influences those that have the biggest impact on production; enable vast amounts of data …. to be understood" will remain just that, only claims, until we better understand any physical bases for apparent relationships.

So where are we likely to get the next wave of benefits from AI? The most promising area, and one that is already in widespread use, is solving inverse problems; and perhaps the most important of this kind of problem occur in seismic geophysics: inversion and interpretation.

What is an inverse problem? It is the inverse of a regular problem. If we think of a regular problem as being of the kind "What is 2 x 2?", an inverse problem might be "I know the answer is 4, but what was the question?" With no more information, there is an infinite number of answers, but hopefully we have some examples, or additional input.

In general: I know the answer, now what was the question?

In particular I see a log response, what rock type gave rise to it?

Example: What is the missing number (i.e. what rule generates the 2nd number from the 1st):

32 → 21

14 → 12

20 → 15

6 →

An algorithmic approach might be to find a pattern (all results are smaller than their inputs), use pre-programmed rules to narrow down the possibilities (e.g. numbers get smaller from subtraction and division, bigger from addition and multiplication), try a few possibilities (i.e. find an initial position from which to perturb), make perturbations in different directions to see which ones improve the starting guess and which make it worse, and so home-in on the solution, which is add 10 then divide by 2, so 6 → 8. Note that had we been given this as an input, the solution might have been easier because now we have some bigger and some smaller outputs, and where the output is bigger than the input, the input is small. Very useful information.

Other answers are possible, but this is the simplest. Of course, a real situation might be:

32 → 21.10

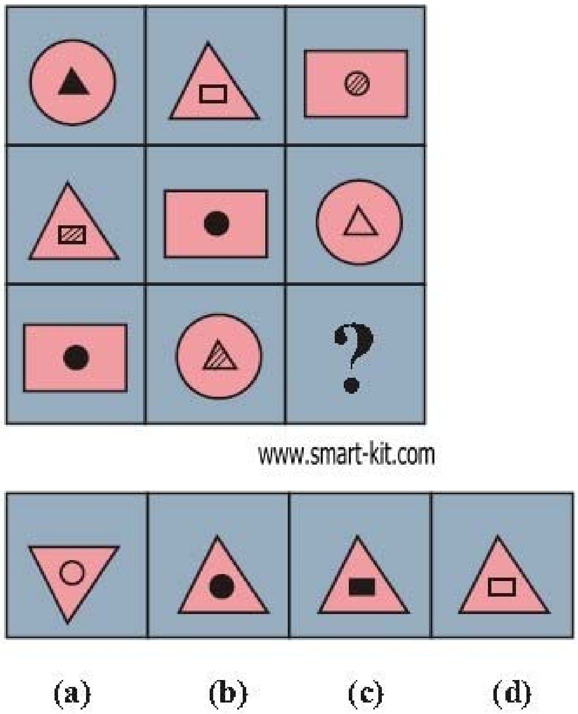
14 → 11.92

20 → 15.05

and the answer is still "add 10 then divide by 2".

Remembering that the foundation of AI was Turing's insight that computers were machines that could manipulate logic, we could also look at a non-numeric example:

Which figure (a), (b), (c), (d) best completes the matrix?




1. There are 3 rectangles, 3 circles, and 2 triangles => the missing large shape is a triangle.
2. Each large shape contains a specific type of inner shape (circles contain triangles, rectangles contain circles) => missing shape (triangle) contains a rectangle..
3. There are 3 black inner shapes, 3 shaded inner shapes, and 2 open inner shapes showing => the missing shape has an open inner shape.
4. Answer: d

These are examples of inverse problems. In seismic inversion, we have the "answer" of a series of wiggle traces and we seek the question: what lithological successions could have given rise to this outcome?

The PESGB London evening lecture in July 2016 took the theme of AI in Stochastic Seismic Reservoir Characterisation. The inverse problem used as illustration, i.e. the analogy with inversion, was reconstructing a snippet of language from an incomplete print-out. So, assume that the answer (what we can see, the data available) was

"N_w i_ t__ w__ter    ou_ di_co__eqt"

and now we want to establish the question: what text might have this been originally? Equivalently, what happens if if remove some letters from, and mis-read some other letters in:

"Now is the winter of our discontent"

A possible progression, from bad copy to best reconstruction of original text, was presented:

| | | |
|---|---|---|
| 1 | Now is the winter of our discontent | model |
| 2 | N w i t w ter ou di co eqt | data |
| 3 | Nyw iq t,e w?fter pi oug dixcow!eqt | random letters |
| 4 | New if tie waiter pi out disconnect | random words |
| 5 | New is the waiter at our discotheque | words plus grammar |
| 6 | New as the winter of our discontent | 16c. words plus grammar |
| 7 | New is the winter of our discontent | 16c. words plus grammar |

The approach outlined above is similar to that being used by Google and others for speech recognition, language translation, and similar problems of language and cognition. In the early days of AI, a problem which was deemed to be a showstopper, and one which would truly show the worth of AI if it could be solved, was to distinguish meaning between:

Time flies like an arrow

Fruit flies like a banana

The Google approach is not to try to understand nor to conduct a grammatical analysis, but to scan vast amounts of written material and find patterns: instances of ‘flies’ with ‘banana’ (and ‘apple’, ‘orange’, etc.) as opposed to instances of ‘flies’ with ‘arrow’ (and ‘axe’, ‘sword’, etc.). This is an approach derived from ‘The Wisdom of the Crowds’, also known as "17 quadrillion flies can't be wrong, eat dung" ( where the re-use of 'flies' is accidental), or perhaps now better characterised as "63 million Americans can't be wrong, Trump is  the best candidate for USA president". It is made possible by the availability of vast amounts of written language, scanned or otherwise posted onto the internet. Provenance is unknown, Shakespeare’s words receive equal weighting to Donald Trump’s. This is the age of Wikipedia, where everyone can be an expert.

The first point to note here is that we have the same problem as previously, that patterns will be found, whether or not there is one. But now there is a second interesting issue: context. Patrick Connolly’s choice of illustration is itself an illustration of the importance of context. Having successfully inverted this problem and found the question was “`Now is the winter of our discontent`”, what do we do with it?  If  Middle Ages GCHQ had intercepted the message “Now is the winter of our discontent” they might reasonably have deduced that the King was in the middle of a winter of discontentedness, and they might therefore have decided that the time was right for invading Iraq, which would have had disastrous consequences, because he hadn’t even paused for breath before continuing “made glorious summer”. Patrick Connolly’s analogy was that sentences or phrases such as “`Now is the winter of our discontent`” are analogous to “short vertical geological profiles”. But if we take these and build a case for war without placing them in context, we’ll be in the Chilcot before you can say “by this sun of York”.

"Now is the winter of our discontent" does not mean that it is winter now, nor that we are discontent. In modern language, Act 1 Scene 1 of Richard III might be written:

Not a good #discontent winter.Been improved bigly #BroTed.Big victories.Now it's #partytime,not me btw #hunchback.Clarence,is all fake news.  Traitor #lockhimup.

or, as Shakespeare more nearly said: "Since I can’t amuse myself by being a lover, I’ve decided to become a villain. I’ve set dangerous plans in motion, using lies, drunken prophecies, and stories about dreams to set my brother Clarence and the king against each other." This illustrates a well-known problem in AI: if we have limited context, we may make wrong decisions. A classic illustration is:

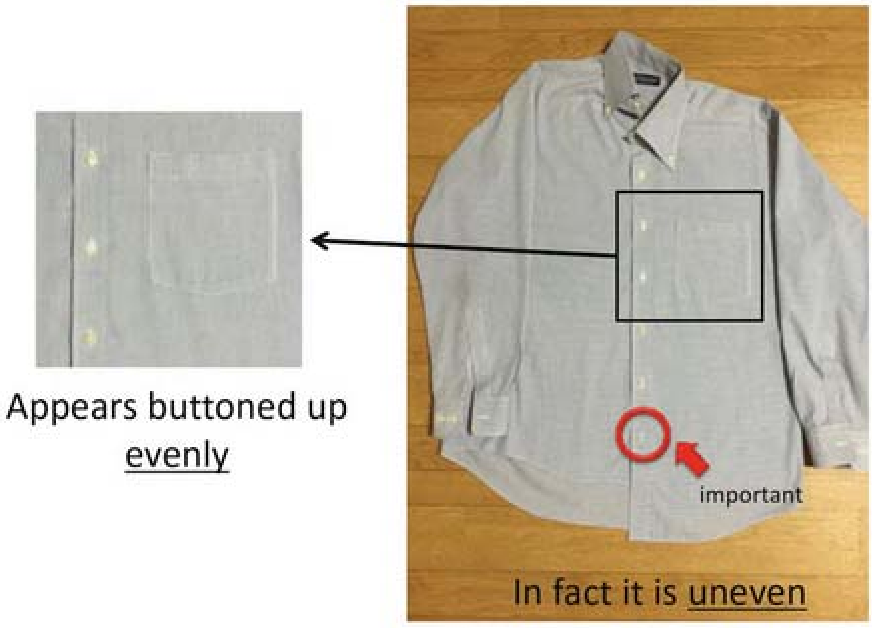


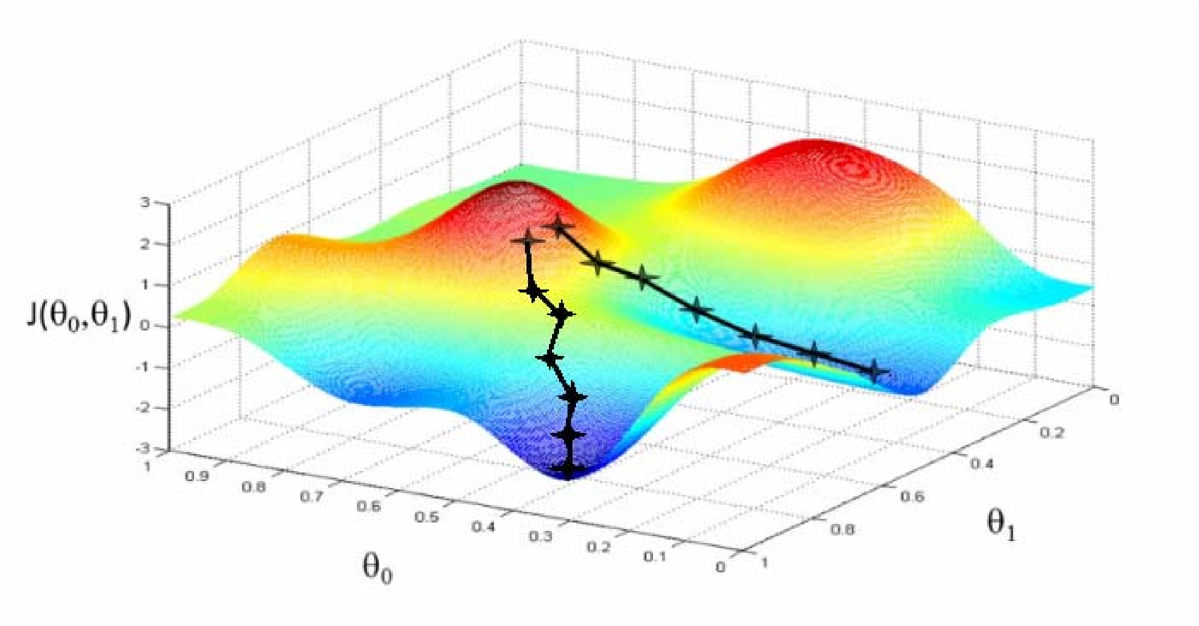

**Fig. 3** Illustration of uneven buttoning (Published with kind permission of © Takuya Hasegawa 2015. All rights reserved)

Or, looking at the actual problem (using pseudo-wells):you can probably insert your own example of poor interpretation resulting from limited contextual information.

We should expect to see significant benefits from AI in E&P in the solving of inverse problems, possibly from use of Genetic Algorithms:

‘The genetic algorithm is a method for solving both constrained and unconstrained optimization problems that is based on natural selection, the process that drives biological evolution. The genetic algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm selects individuals at random from the current population to be parents and uses them to produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution.’

This holds out the hope of overcoming one of the main pitfalls in current methods available for solving inverse problems (e.g. Steepest Descent, Simulated Annealing), which is that the choice of starting point can influence the solution obtained, and in particular a local minimum may be obtained instead of the desire global minimum:




Genetic Algorithms could allow controlled 'mistakes' to help us work our way out of the wrong hole. However, to do so they need to be different in kind to the controlled errors introduced by stochastic methods and yet be sufficiently small as to distinguish the method from a purely random "let's try something and see if it works". i.e. there needs to be a balance between learning from previous trials, which in its purest form would be iteration, and randomly starting again. This is the area Gas ned to address if they are to be useful in E&P.

Evolution is based on making the best use of mistakes (mutations) - mistakes are key to evolution. Without random mistakes, there would be no evolution; natural selection would have no raw material from which to create higher ("fitter") forms. However, Clarke's Second Law was: "The only way of discovering the limits of the possible is to venture a little way past them into the impossible." This goes against our natural thinking process. Sherlock Homes said "When we have ruled out the impossible….", i.e. he did not want to even consider the impossible. But perhaps our intelligence is based in part on making mistakes; sometimes it is a mistake that inadvertently leads us to the right conclusion.

The IMDB.com plot synopsis of the film "The Day of the Jackal" (1973) explains:

.... because of the Jackal's anonymity from even his own employers, no one has a clue where to start looking for him; the key to finding and stopping the Jackal thus is to establish his identity.

The British police come up with a suspicious character named Charles Calthrop ..... a search on Calthrop's home reveals he is absent, but his passport is there, so they assume he is travelling on a false passport. A review of all passport applications reveals that one was made for one Paul Oliver Duggan, who had died at the age of two. (This is the clue that eventually leads to the thwarting of the plot.)

After the failed assassination in Paris, Charles Calthrop appears at his flat in London. The British authorities take him in for questioning, but conclude that he had nothing at all to do with the Jackal and close the matter (also forswearing any responsibility regarding the Jackal, since he was an Englishman but also a Dane and a Frenchman).

I am sorry, I haven't read the book so I do not know how important this was to the original plot, but in the film it is clear that the clue that led to the thwarting of the assassination attempt was a serendipitous mistake; Calthrop had nothing to do with it. The discovery of penicillin is said to have had similar assistance, and evolution relies on mistakes. We still do not know how to harness this effect and in particular we have little idea as to how it might impact the strong versus weak AI debate.

What is the strong versus weak AI debate? Since the advent of AI we have been concerned with the possibility that computers will one day be more intelligent than us. They already beat us at every known game of skill (Go was the latest and arguably the last), they can out-perform us at any mathematical trick or skill (how would you go about verifying a number with more than 22 million digits is prime?), so what do we mean by "more intelligent than us"?

Wikipedia has its own idea: *Weak AI* is defined in contrast to either *strong AI* (a machine with consciousness, sentience and mind) or artificial general intelligence (a machine with the ability to apply intelligence to any problem, rather than just one specific problem). ... *Weak* or "narrow" *AI*, in contrast, is a present-day reality.

*https://en.wikipedia.org/wiki/Weak_AI*

This is subtly but significantly different to the accepted definition in the AI community:

> **Strong AI** holds that if we continue to improve a system in terms of its computational power and its available data (which may include both pre-defined rules and a capability to find its own data), there will come a point at which it achieves an

autonomy of action, an ability to generalise, and be deemed capable of independent thought.

> **Weak AI** holds that there is a difference in kind between a programmed machine and a sentient being, so that no such point can exist.

Or, quoting http://www.math.nyu.edu:

> There are two major ways to think about the future and current utilization and power of artificial intelligence. The weak AI hypothesis states that a machine running a program is at most only capable of simulating real human behaviour and consciousness (or understanding, if you prefer). Strong AI, on the other hand, purports that the correctly written program running on a machine actually *is* a mind -- that is, there is no essential difference between a (yet to be written) piece of software exactly emulating the actions of the brain, and the actions of a human being, including their understanding and consciousness. I'm sure many people would like to argue the finer points of these rough definitions, but the general ideas are universally accepted (not to be correct truths, but as valid definitions). Some supporters of weak AI prefer to call it cautious AI.

The terms were originally invoked by Searle:

The terms *strong* and *weak* don't actually refer to processing, or optimization power, or any interpretation leading to "strong AI" being *stronger* than "weak AI".

- AI hypothesis, strong form: an AI system can *think* and have a *mind* (in the philosophical definition of the term);
- AI hypothesis, weak form: an AI system can only *act* like it thinks and has a mind.

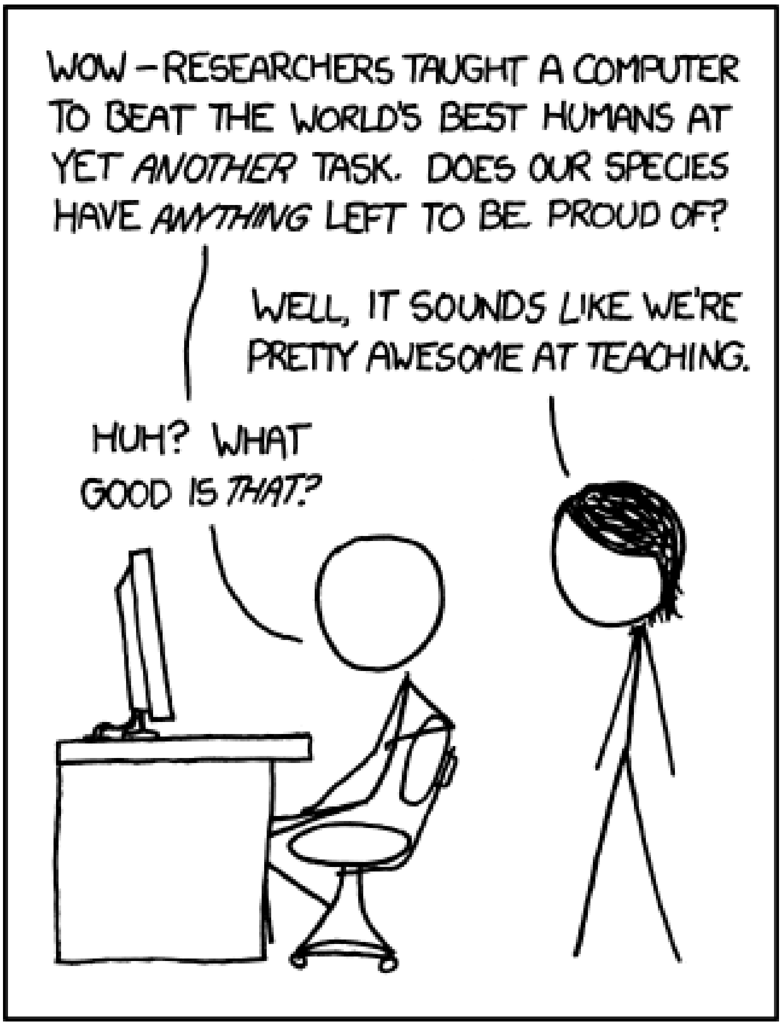
The problem with these definitions is that they're fuzzy. For example, AlphaGo is an example of weak AI, but is "strong" by Go-playing standards. A hypothetical AI replicating a human baby would be a strong AI, while being "weak" at most tasks.

However, this is an arcane argument which is still considered open within the AI community, even if it considered closed by the Daily Express ("AI will kill us all"). What is germane to AI in the E&P industry is the reliance on public domain and unverified data sources: Wikipedia has a definition that disagrees with academic opinion; academics are more likely to build and program the AI machines; but the data sources on which they rely, if subject to the Wisdom of the Crowds, could disagree with them. So maybe, instead of rising up against us, the AI machines will disappear in a cloud of logic as they try to reconcile such paradoxes, but we have a more immediate responsibility: to ensure our companies don't make exploration decisions based on information supplied by the bloke down the pub at 11 o'clock on a Saturday night..

Should we be worried about job loss? Maybe I should, as it seems my branch of petrography may imminently see humans replaced by computers:

Zheng Ma and Shichen Gao (2017) "Image analysis of rock thin section based on machine learning" International Geophysical Conference, Qingdao, China, 17-20 April 2017: pp. 844-847
https://doi.org/10.1190/IGC2017-213)

but in general I think we need to be kept on by E&P employers, to keep an eye on the computers. As Randall Munroe put it in his online comic (xkcd.com):




Barrie Wells, September 2017.