

## \* Massive Open Online Courses

The limitations of traditional actuarial education when it comes to the broader business analytics have been previously highlighted in this magazine by Adam Driussi of Quantum. Indeed, so called 'analytics' form a bewildering accretion of ideas developed over more than a century (e.g. scientific management, biostatistics, statistical quality control, operations research, financial economics, machine learning and so on). Further, it is quite demanding, though by no means impossible, to be conversant on such a wide range of topics and to appreciate the broad similarities.

While the actuarial profession has enjoyed not inconsiderable success in this area in Australia, it is important to realise that the Australian market has been fortuitously sheltered due to the absence of strong graduate programs (in the North American sense). Australian Masters almost exclusively cater to career changers and in the case of PhD degrees, there are no substantial advanced coursework requirements. Combined, these ensure that only a small percentage of students get a solid technical background outside of their immediate specialisation.

Can the actuarial profession develop an effective and globally competitive data science or analytics qualification? The topics in question easily span half a dozen university schools and departments (including statistics, finance, econometrics, applied maths, computer science, industrial engineering and others), making a substantive discussion of a suitable curriculum a challenge in itself.


The situation is, however, far from hopeless. The great benefit of modern communication technologies is that an individual learner is no longer held back by institutional limitations. Some superb materials from the leading universities have been made publicly available over the last 10 years and the pace is decidedly picking up.

There have been, perhaps, three main waves in the Internet facilitated open education to date. Back in mid 1990s, instructors in universities worldwide began making their notes available online. With increased bandwidth and proliferation of digital recording equipment, more and more courses have been made accessible in the form of video or audio recordings. Noteworthy initiatives include MIT Open Courseware, Stanford Engineering Everywhere and [webcast.berkeley.edu](http://webcast.berkeley.edu).

Over the last few years, drawing inspiration from the popularity of the MIT Open Courseware initiative, a new format of online courses has become increasingly popular. MOOCs go beyond videos and lecture notes and offer a somewhat interactive experience with discussion forums, multiple choice quizzes and automatically, or peer graded, assignments. Leading providers are Coursera and edX, with over a hundred universities participating.

While serious questions about accreditation and the overall MOOC business model remain, it is now possible to piece together a fairly comprehensive curriculum in analytics or data science that is taught by great researchers who also happen to be exceptional teachers.

Of course, textbooks and monographs have been available for centuries and it can be argued that the material was always accessible in some form to anyone within reach of a university library. There seem to be strong indications, however, that for someone engaged in self-study, additional informal insight provided by the best



Online courses could help develop a broad fundamental understanding of computing and mathematical modelling – which it could be argued will be core analytics skills for addressing a variety of future business problems.

professors can significantly boost the rate of success.

This is especially true in mathematical modelling and related topics, as textbook presentations tend to be overly formal and difficult to decipher without being given the right key. Indeed in my own experience the interactive features of the MOOCs have been easily overshadowed by the personality of the professor; yet the additional structure is not entirely without benefit and can provide an additional impetus towards completion.

Recognition of learning via open courses is likely to remain an unsolved problem for some time and I believe this is the main area where the Actuaries Institute could differentiate themselves. It is relatively straightforward to design independent assessment for the material contained in courses that are already publicly available and such an initiative has potential to attract students, especially if credit is given towards the Fellowship designation.

In what follows I give an outline of some particularly noteworthy graduate and undergraduate courses that could help develop a broad fundamental understanding of computing and mathematical modelling. These could be argued to be core analytics skills for addressing future business problems, with the increasing number of processes and low-level operational decisions subject to automation.

To access any of these, a quick Google search by course name will take you to the relevant pages.

## 1. ANALYTICS AT WEB COMPANIES

To get an impression of what the future might hold, it might be worthwhile to review some of the courses offered by people with experience implementing analytics solutions for the leading web companies. Examples include CS281B Scalable Machine Learning

at UC Berkeley [25] by Alex Smola (formerly of Yahoo) and Big Data, Large Scale Machine Learning at NYU [26] by Yan LeCunn (currently at Facebook). In particular the first course offers an interesting insight into the importance of understanding systems, numerical methods and statistics to develop analytics solutions at web scale.

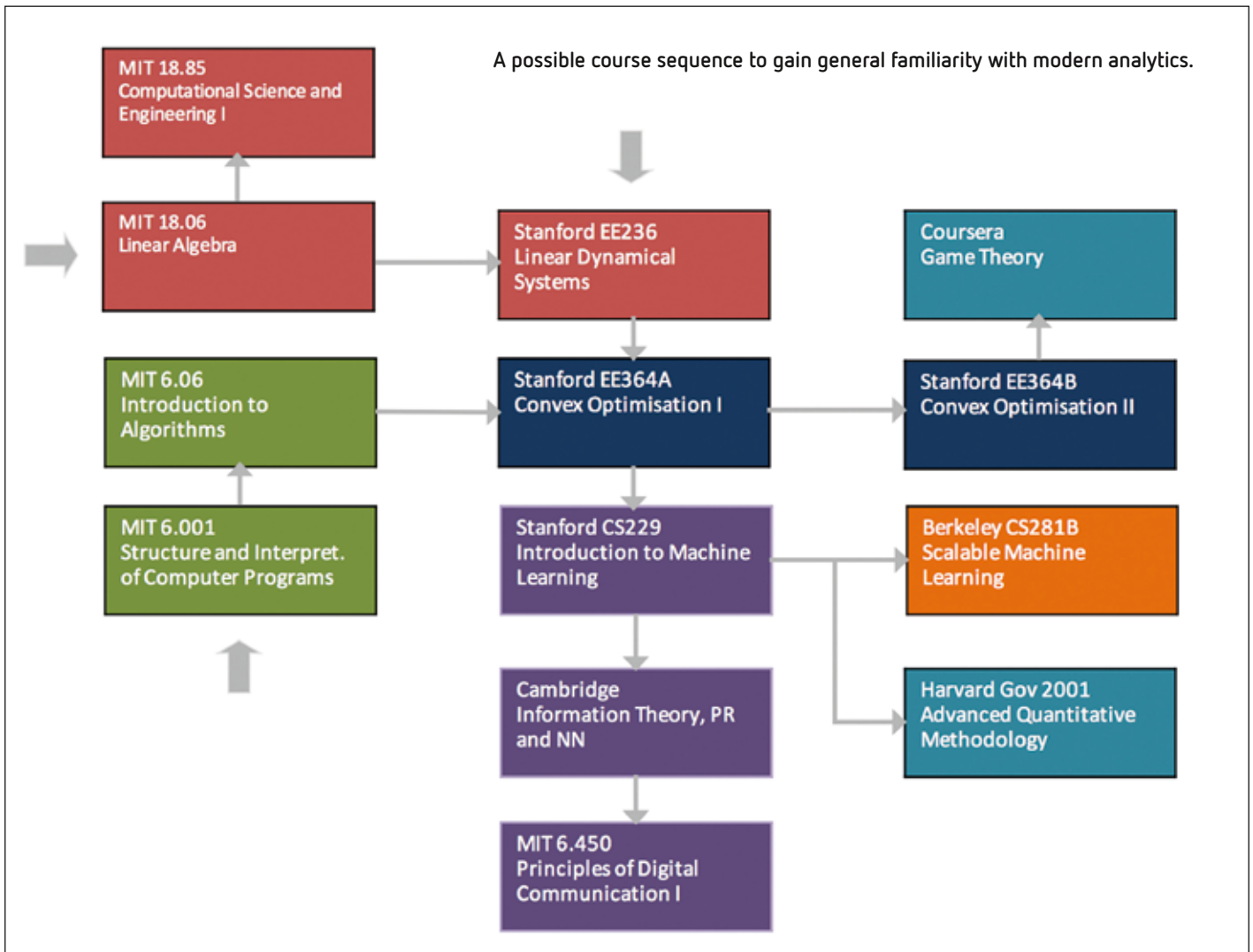
Prerequisites for these broadly include linear algebra, basic probability and statistics and, ideally, convex optimisation and an introduction to machine learning, as discussed next.

## 2. MATHEMATICAL BACKGROUND AND NUMERICAL COMPUTING

In some sense, the lynchpin of applied mathematics is linear algebra. The majority of computational procedures for solving mathematical models ultimately reduce to iteratively solving systems of linear equations.

An excellent introductory treatment of linear algebra is given by Gilbert Strang in MIT 18.06 [2]. The material is further developed in MIT 18.085 and 18.086 [4], demonstrating a very broad range of applications across engineering subfields. The observation that the differential operator can be discretised as e.g. a tri-diagonal matrix (the so called 'finite differences' method) is the key connection between linear algebra, traditional calculus (in the form of integral and differential equations) and computing.

Another take on the material is given in Stanford EE263 taught by Stephen Boyd - in addition to basic linear algebra, the course gives highly intuitive exposition to least squares regression, regularisation, singular value decomposition and linear dynamical systems (which can be viewed as a generalisation of a wide class of time-series models in the Part I syllabus). Somewhat



off the usual track, the material above should provide sufficient background to appreciate the technology behind modern robotics platforms, such as those formerly developed at Boston Dynamics, now part of Google (MIT 6.832 Underactuated Robotics [12]).

Finally, the Fourier transform is one of the most famous special cases of a linear operation – an intuitive introduction to the subject and its multitude of applications, including the Central Limit Theorem, is given in Stanford EE261 [17].

### 3. OPTIMISATION

Beyond differential equations, one of the main applications of linear algebra is in mathematical optimisation or mathematical programming. Optimisation based models are pervasive in analytics, whether it be maximum likelihood estimation, empirical risk minimisation, Neyman-Pearson hypothesis testing, optimal control, Markowitz portfolio theory or option pricing.

Prof. Stephen Boyd's course EE364A Convex Optimisation [19,20] not only gives a solid grounding

in the theory but also considers many of the above-mentioned examples. Convex optimisation is widely seen as the foundation of modern statistics, machine learning and signal processing. Familiarity with theory and algorithms will enable the practitioner to identify and implement solutions to a very wide range of problems across industries.

There is also an interesting connection between mathematical optimisation and classical algorithms studied in undergraduate computer science courses (e.g. MIT 6.06 [7]) – many of the problems such as sorting, shortest path, max flow &c. turn out to be special cases of linear programming (itself a special case of convex optimisation). The follow up course EE364B [21] provides more detailed background on scalable and distributed optimisation as well as the clearest introduction to the General Equilibrium theory of microeconomics you are likely to find.

The background for these courses is limited to linear algebra, MIT 18.06 [2] and EE263 [18] – Introduction to Linear Dynamical Systems, and basics of multivariable calculus (gradient, Hessian) – MIT 18.02 [1].

#### 4. PROBABILITY, STATISTICS, MACHINE LEARNING, INFORMATION THEORY

There are few unequivocally great introductory probability and statistics courses publically available, at least at the moment. MIT 6.041 [9] is a useful probability refresher. A worthwhile follow up is MIT 6.262 [10] Discrete Stochastic Processes.

When it comes to statistics, or at least a take on the topic that is more attuned to analytics applications, Stanford Statistical Learning [21] is a solid introduction from the authors of the well-known book. A closely related subject area is machine learning, with the introductory course by Andrew Ng, CS229 Machine Learning [23], and a much more in depth treatment by Alex Smola, CMU 10-701 Introduction to Machine Learning [24]. So called 'deep networks' are a recent 'hot' topic in machine learning, providing state of the art performance for many recognition tasks. This material is covered in NYU Deep Learning [27].

Information theory provides perhaps one of the most successful and widely used applications of probability. There are also important connections to statistics and machine learning (as efficient compression requires effective conditional probability estimation). MIT 6.450 Principles of Digital Communications I [11] is an excellent course by the pioneer of digital communications Rob Gallager, who invented one of the most effective known coding schemes and was a founding engineer at Qualcomm where he designed the first 9600 baud modem. Information theory is an essential foundation of all digital information processing technology.

Another excellent discussion of information theory is given in the course taught by David MacKay at Cambridge [36] – Information Theory, Pattern Recognition and Neural Networks, bringing together topics from coding theory, statistics and machine learning.

Convex optimisation provides a very helpful background for the courses in this section even if it is not explicitly alluded to.


#### 5. PROGRAMMING

There exists a very wide range of high quality introductory programming courses. Perhaps the Stanford sequence deserves a particular mention, CS106B Programming Abstractions and CS107 Programming Paradigms [15,16]. Alternatives include the introductory courses at MIT, MIT 6.00 and MIT 6.01 [5,8], and YouTube videos of UNSW

COMP1917 Higher Computing [13] (Richard Buckland is an ex-actuary and has won multiple teaching awards).

MIT 6.001 [6] (now superseded) is the most celebrated introductory programming course of all, with the textbook *Structure and Interpretation of Computer Programs* used in dozens of top universities.

While Scheme, the language that it uses for teaching programming concepts, has for a long time been considered less than practical, over the recent years there has been a dramatic resurgence of popularity of the related body of ideas called functional programming, underpinning many of the latest big data technologies. Beyond the introductory courses, Programming Paradigms [17] gives a useful overview of design choices behind a variety of programming languages and University



Over the recent years there has been a dramatic resurgence of popularity of the related body of ideas called functional programming, underpinning many of the latest big data technologies.

of Washington – Programming Languages [31] offered on Coursera by the University of Washington, provides a more advanced grounding in the functional programming paradigm.

An introduction to Scala, an increasingly popular compatible replacement of Java, is available from its creator on Coursera by searching 'EPFL – Principles of Functional Programming in Scala' [32].

No such list would be complete without an algorithms class – MIT 6.06 [7]. Conceptual links with optimisation or mathematical programming offer a connection back to the material in the earlier sections.

#### 6. FINANCE, ECONOMICS AND SOCIAL SCIENCE

While the exact relation between actuarial pricing and financial economics is not clearly set out in the Part I curriculum in Australia, it has been understood in the academic literature for some time as the so called 'incomplete markets' setting. An introductory

discussion of the modern theory of finance (CAPM, option pricing &c.) from this more advanced point of view is given in John Cochrane's (University of Chicago) class Asset Pricing on Coursera [30].

A useful generalisation of the concept of an optimisation problem (see e.g. Stanford EE364A and CVX101) [19,20] is offered by game theory. Instead of considering a central planning problem where all the decisions are taken by a single agent, game theory looks at situations where there are multiple self-interested parties involved. Coursera classes, Stanford/UBC – Game Theory and Game Theory II: Advanced Applications [28,29] provide an introduction to a range of topics, including auctions and mechanism design. Applications of game theoretic methods to the study of social insurance, optimal taxation and related ideas are given in the Harvard course Public Economics [34].

Problems addressed by business analytics are not dissimilar to those found in the social sciences, especially when it comes to identifying what is sometimes called actionable insights – a social scientist may instead talk about policy targets. While causal attribution is often times not necessary, it is important to be aware of limitations of analyses carried out purely on observational data. One example in social science where large-scale experiments have been possible is development economics. The course MIT 14.73 [35] offers an in depth discussion of considerations that go into designing a convincing experimental study. A broad introduction to the design of quantitative methods that are directly applicable to the question being studied is given in Gary King's excellent methodology course at Harvard, Gov 2001 Quantitative Research Methodology [33].

#### CONCLUSION

The courses referred to above are only a relatively small selection as higher education has never been more accessible than at present. The more individual members of the profession that become familiar with the existing modes of thinking in business analytics the easier it will be to collectively respond to the future challenges presented by the rapidly shifting technology landscape – both inside and outside existing practice areas. **A**

For a listing of references go to <http://www.actuaries.asn.au/Library/AAArticles/2014/MOOCsReferences.pdf>