# FORTY-FIRST
# CONVOCATION ADDRESS

## *WHAT IS INFORMATION ENGINEERING?*

BY

**Professor Sir Michael Brady, FRS**
**Department of Engineering Science**
**University of Oxford**
**Oxford, U.K.**

07th March, 2007

INDIAN STATISTICAL INSTITUTE
KOLKATA

FORTY-FIRST
CONVOCATION ADDRESS

# *WHAT IS INFORMATION ENGINEERING?*

BY

**Professor Sir Michael Brady, FRS**
**Department of Engineering Science**
**University of Oxford**
**Oxford, U.K.**

07th March, 2007

INDIAN STATISTICAL INSTITUTE
KOLKATA

# What is Information Engineering?

## Preamble

**Professor M.G.K. Menon, President of the Institute, Professor Sankar K. Pal, Director of the Institute, distinguished Faculty Members and Guests, my colleagues and friends in the ISI, especially Professor D. Dutta Majumder and his colleagues, with whom I spent two very happy weeks in the early 1990s, and especially Dr. Dipti Prasad Mukherjee with whom I published a paper in the Proceedings of the Royal Society. I greet especially all the students of the ISI graduating today. It is both an honour and a privilege for me to participate in the Platinum Jubilee conference of the Computer & Communication Sciences Division of the Indian Statistical Institute and to address the 41st Convocation. Before I begin my remarks proper, I would like to take the opportunity to offer my heartiest congratulations to all of you, the students, who are graduating today. You have the opportunity to play important roles in the development of India – a point to which I will return at the end of my Convocation Address.**

**Reading through the extremely informative Annual Review of the ISI by the Director, Professor Sankar Kumar Pal, I am struck by the continuing breadth and depth of the world-class**

**research carried out in the Institute. I am sure that Prasanta Chandra Mahalanobis would have been proud of the development of his extraordinary vision when he founded the ISI seventy five years ago! Certainly, probability and statistics have been, and continue to be, important mathematical pillars of my own research. Indeed, at the International Conference on Computing: Theory and Applications (ICCTA), which is taking place at the ISI this week, I will be presenting a paper that reports results from joint work with my graduate student Niranjan Joshi, entitled "Non parametric mixture model based evolution of level sets", in which we propose a novel method for estimating probability density functions for use in image analysis, and we apply the method to colorectal cancer, a major disease in the West.**

**It has indeed been a pleasure to participate also in ICCTA, an apt choice for a Platinum Jubilee meeting organised by the ISI in view of the substantial research and teaching effort in the ISI dedicated to computing, and in view of the massive and growing importance of computing in every aspect of our lives. My chair in Oxford University was the first there in Information Engineering – it resulted from reassigning a chair from Civil Engineering to what the Department called "Information Engineering". I want to begin by discussing Information Engineering in general terms, then tell you a little about the research that I have done at**

Oxford. It will become clear that my work not only involves problems that span computing, mathematics, physics, and engineering, but also biology and medicine. This is quite a span, so I will spend a moment discussing how I think it can be possible to work across such a broad range of subjects! Finally, in a Coda, I want to return to India and to you – the students – since you are the most important part of today, and tomorrow.

## Information Engineering at Oxford University

The 1980s were a period of extraordinary growth in Information Technology (IT), around the world; but particularly in the UK and Europe. In a key development, the UK Government announced a major programme to stimulate academic research in IT, with ambitious plans for the training of both undergraduate and graduate students. The programme also aimed to stimulate industrial exploitation of the new possibilities afforded by IT and, crucially, to catalyse a web of collaborations between Academia and Industry. This plan, drawn up by Lord John Alvey, lately a senior executive of British Telecom, became known as the Alvey Programme. One goal of the Alvey Programme was the repatriation to the UK of IT specialists and scientists who had left the UK to work abroad, mostly in the USA. Concurrently, the European Union (EU) initiated a continuing series of Framework programmes to stimulate emerging areas of science that were judged to be of strategic importance for Europe: the first such was IT, not surprisingly the current focus is on issues such as the environment, healthcare, and proteomics.

Oxford University responded to the opportunities afforded by the Alvey Programme, and Framework I, by bidding to establish a new research programme, to be led (from Engineering Science) by a newly-appointed Professor of Information Engineering, and an undergraduate degree *Engineering and Computing Science (ECS)*, to be run jointly between the Computing Laboratory and the Department of Engineering Science. The undergraduate programme proposal was established and led to the creation of 12 new academic posts: 4 in the Computing Laboratory (already well established under the leadership of Professor Sir Tony Hoare FRS), 4 in (general) engineering science, and 4 within the research area of the new professor. The story of how I came to be called for interview for the newly created chair could be the basis of a TV comedy, perhaps a Bollywood movie, but my 3 days of “interview” in Oxford ended with a meeting between me (dressed as usual in short sleeves, open-necked shirt) and four imposing looking figures in full academic dress (the Vice-Chancellor and his senior colleagues). They offered me the position, then asked me if there was anything I would like to ask. I thought for a

**moment and then, somewhat tongue in cheek, said "Yes, what is Information Engineering?" [hence the title of my talk] The Vice-Chancellor immediately snapped back: "Well, Dr. Brady, we rather hoped you would tell us!". I accepted the post and then really began to think about what Information Engineering meant – and still means – to me.**

**I believe that Information Engineering brings to Engineering Science the ideas and techniques of Computing Science. Engineering Science involves working on the scientific bases of problems that are important in practice. It explores fundamental relationships, independent of the specifics of the application. Of course, we all know that linear circuits (resistors, capacitors, inductances) are subject to the same conservation of energy second-order differential equations as the motions of damped spring masses. We find that the Navier-Stokes equations of viscous fluid flow also have application to the non-rigid displacements of organs, for example when a tumour grows in the liver. From this perspective, the themes that characterise Information Engineering, and that have characterised my research in Oxford, irrespective of the domain of application, are rooted in real problems. My work has progressed from the control of mobile robots, initially for application in manufacturing, to systems to help diagnose breast, liver, and colorectal cancer; mobilisation and development of ideas from Computing Science,**

particularly the ideas of representation and distributed computer architecture. This is particularly evident in my work on mobile robotics, but also in that on teleconferencing and medical image analysis. A particular theme is the application of Artificial Intelligence, in the quest to make systems "smart". A recent interest in distributed computation has to explore the potential of Grid technology for application to medical image analysis, building on a solid mathematical foundation. This permeates all of my work, whether it is a differential geometrical analysis of image surfaces, a model for x-ray image formation, infrared image analysis, a model of hypoxic tumour growth, or a model of defocussing blur in an image; and that the work is intrinsically inter-disciplinary: at Oxford I have worked with mathematicians, radiologists, clinical neurologists, psychiatrists, psychologists, surgeons, pathologists, clinical pharmacologists, and ancient historians, as well as with people from a range of industries, spanning defence to pharmaceutics.

I am struck by the similarity to Mahalanobis' views about statistics, as quoted in The Governor of West Bengal's 39th Convocation Address: "..statistics is not a branch of mathematics but is a technology which is essentially concerned with the contingent world of reality".

**When I arrived in Oxford, my first aim was to develop a mobile robotics project, for the reasons I will explain. As a result, and despite the enormous diversity in the work in the Laboratory, it is still called the Robotics Laboratory.**

**Apart from my developing conception of Information Engineering, there was a second guiding principle that underpinned the organisation of the Robotics Laboratory. Visiting any of the major IT laboratories in the USA, I was impressed, above all, less by the quality of the faculty (though this was uniformly high) and more by the large number, energy and creativity of the graduate students. I determined that the Robotics Laboratory should be built around a large number of graduate students, and that they would be equipped to enable them to realise their full potential: our students should be limited only by their own ideas and creativity, not by transients such as equipment, Our job, as academic advisors, should be to remove as many of the limitations on thinking and creativity that the students manifested! I believe this even more today than I did 21 years ago. For me:**

**"…the graduate students in a major laboratory such as ours have four major qualities:**
**they are smart;**
**they work incredibly hard;**
**they are fearless in tackling tough problems, and**

they leave after three years, to be replaced by another!

The latter point is the essential driver of creativity in the laboratory.

The academic achievement I am proudest of is *not* my election to bodies such as the Royal Society but the fact that I will have graduated over 100 PhDs by the time I retire. To date, my students have come from over 20 countries. Not surprisingly, the country that is the main source has been Great Britain; but it accounts for less than one half. After Britain comes China with 22. I have to date had relatively few from India, though I have supervised several UK students of Indian descent. My current crop of 19 DPhil students and 6 postdoctoral researches come from 11 countries – interdisciplinary, international, and incredibly bright!

## Robotics and image analysis

The first research focus of the Robotics Laboratory was the development of robot vehicles, also known as Autonomous Guided Vehicles (AGVs). Their development represents a tough challenge for robotics. To perceive their surroundings, particularly if they are to operate outdoors and contend with the vagaries of British weather, AGVs need to be equipped with a range of sensors, since no single sensor can be relied upon in all situations. In practice, they also need to maintain a representation of their nominal environment so that unexpected obstacles, and unplanned changes to the environment can be detected and dealt with. Practical applications demand that the vehicle operate in real time, and to provide the required bandwidth and fail-safe operation, implies a distributed computer architecture. AGVs involve not only the design and implementation of complex mechanical and electrical structures; but also that of complex, multi-scale computer architectures.

Uncertainty is rife: signal and imaging sensors are inevitably noisy, and the real world cannot be predicted with complete certainty. Objects in the robot's environment get moved; but this should not stop the robot completing its task, in the same way that while roadworks may frustrate us, they generally don't stop us getting to our destination. We may note the cause of the roadworks, and, if not told explicitly how long they will disrupt our route, we may plan an alternative route for tomorrow. Sometime later, we will probably try the original, preferred, route again. Sensor fusion and the integration of sensory data with possibly incorrect maps is a fundamental challenge of engineering, and of AI, and it requires the mobilisation of probability and statistics! I like to think that Mahalanobis would have worked on mobile robots!

This is not the place to summarise our work on mobile robotics; but I do want to mention one particular outcome from it. From the start of our work in 1986, we had worked in close collaboration with industry, in this case GEC Electrical Products, specifically the mobile robotics team headed by Dr. Malcolm Roberts. In 1990, GEC closed down their mobile robotics group following a takeover by CEG. Malcolm, together with two of his colleagues and myself, decided to launch our own company, which is now called Guidance (http://www.gcsltd.co.uk/ ). Starting from a hardware/software platform to enable multiple mobile robots to locate their positions and headings accurately, using retro-reflecting targets illuminated by on-board scanning infrared lasers, and the LICA (locally-intelligent control architecture) technology we had developed in Oxford University, Guidance has come to dominate the navigations sector of the mobile robotics market in the USA, with nearly 2000 AGVs using our technology at the time of writing. Most recently, we have extended the sensing and control technology, as well as the algorithms to measure and reduce uncertainty, to the navigation of ships as they attempt to dock alongside off-shore platforms which are typically heaving under rough seas.

Guidance was my first, but by no means my last, entrepreneurial venture based on my research. I believe that there can be a symbiosis between academic research and commercial exploitation: technology-based companies evidently need a stream of new developments to keep at the forefront; what is less evident, but none the less true, is that the challenges which industry confronts pose some of the deepest intellectual problems for us to work on. I encourage my students to have close relationships with companies – many of them have started their own companies, or worked in those I have set up.

## Medical image analysis

In 1989, at the height of our mobile robotics work, I started to work on medical image analysis, more specifically on mammography – x-ray imaging of the breast. I had wanted to get involved with breast cancer ever since my beloved mother-in-law died of the disease at far too young an age. I hope you will not be surprised that I started by teaming up with a talented graduate student, Ralph Highnam, and by seeking out collaborations with Drs. Shepstone and English of the Breast Clinic of the nearby John Radcliffe Hospital. Building on Ralph's DPhil, I have supervised many others on breast cancer, patented some of our methods, and launched a company both to exploit the results commercially and to have them used by clinicians. Nowadays, it is happily the norm for medical image analysis research to be done in close collaboration with teams of clinicians. However, industry should be a third partner in such collaborations, ideally from the outset. Research may lead to prototype software (or whatever) that can be tested in small numbers of

**patients by collaborating clinicians (this is known as translational medicine). However, if it is successful, and clinicians want to use such a technology regularly, then they need to acquire it from industry, because medicine is necessarily highly regulated and so products have to be submitted for regulatory approval. This observation, coupled with my earlier point that the problems faced by industry often pose the toughest intellectual challenges, is why we have worked for over a decade as part of a three-party collaboration: university researchers, clinicians, and industry.**

**Ever since Röntgen's discovery of x-rays in 1895, and the development of MRI in the 1970s, doctors have been able to look inside the body without physically opening it up. Until the mid 1990s the information provided by imaging was primarily of the person's anatomy. However, clinicians are interested in detecting, diagnosing, and treating illness, which is fundamentally related to physiological abnormalities and for which anatomical signs provide, at best, indirect evidence. Increasingly, however, particularly using contrast agents, we can image physiology – the body in action. More recently, we have begun to be able to image the functioning brain, and with the advent of PET we can now image metabolic activity. We have progressed from anatomy to physiology and now to function. The opportunities seem limitless – we just lack the bright young minds to explore those**

opportunities! At the same time, there is increasing need for quantification: measurements and measurement error: how big is this tumour? How much has it grown over the past 3 months? Should it be re-staged? By how much has it responded to therapy? Is surgery a viable option? These are the questions that clinicians increasingly pose. The mathematics and image analysis required to answer them definitively is not straightforward, and solutions are not to hand.

Some facts that underline the importance of image analysis applied to cancer. 10% of Western women develop breast cancer at some point in their lives, and it accounts for 19% of cancer deaths and 24% of cancer cases among women. Approximately 500,000 cases are treated annually in the EU and USA. Similarly, colorectal cancer (CRC), one of the fastest growing cancers in the West, accounts for 500,000 deaths annually, with 700,000 new cases in the USA and EU (and rapidly rising). It is the second most common form of malignancy in the West. It is certain that as India develops strongly over the next 2-3 decades that these figures for the USA and EU will be replicated in India. Note that in all cases of cancer, early diagnosis massively improves prognosis.

This is not the place either to summarise our medical image analysis work, which has overwhelmingly been concerned with cancer, but which has also included work on functional

**MRI of the brain at the fMRIB centre in Oxford (http://www.fmrib.ox.ac.uk/). However, there is one fundamental point I wish to make about medical image analysis, since it has informed all of our work for the past 16 years. On the one hand, medical images offer a very difficult challenge: their signal-to-noise ratio is significantly worse than that from modern CCD cameras; the noise is rarely independent identically-distributed Gaussian (for example, the noise in MRI is better modelled using a Rayleigh distribution); the images vary a great deal for the simple reason that people's anatomy and physiology varies a great deal; and disease processes are often subtle, requiring medical expertise to detect anomalies and interpret them. For all these reasons, medical image analysis pushes image analysis to its limits. On the other hand, to be useful (and used) in clinical practice, medical image analysis algorithms need to be able to work successfully over 99% of the time on hundreds of images, day in and day out. How can it be possible to achieve such a stringent performance requirement on such complex imagery? For me, a fundamental part of the answer simply has to be to develop mathematical models – of the physics of image formation, of the growth of a tumour, of the take-up of a drug, etc – and mobilise such a model in the analysis of images. Ralph Highnam and I began our work on mammography by developing – over a ten year period – a model of the passage of x-rays through breast tissue,**

including scatter, attenuation, and glare, to produce the image that the radiologist views. Recently, another graduate student, Chris Tromans has developed a refined and extended version of the model, and we are now extending it to the new opportunities afforded by tomosynthesis, in which a small number of low dosage x-ray images taken over a limited angular range (eg 25 degrees) enable the reconstruction (based on statistical methods) of the three-dimensional structure of the breast, with a substantial improvement to both sensitivity and specificity.

There are perhaps one hundred different kinds of cancer, with different characteristics, and whose characteristics vary markedly as the disease progresses. An inevitable consequence is that no single imaging modality (x-ray and CT, MRI, ultrasound, …) provides definitive information in all cases. Clinicians are well aware of this and so they routinely take several images of different types in order to frame a diagnosis and treatment option. The integration of different images of nominally the same cancer (or organ) into a single coordinate frame is known as fusion. In 1999, we recognised the growing need for image fusion, especially as combined imaging devices (eg CT-PET) were becoming available. This led us to form a company that became Mirada Solutions. Initially, Mirada aimed to work both on mammography and on fusion; but it quickly became clear that the market desperately wanted fusion and so we threw our entire resource into gaining regulatory approval for *Fusion7D*, a software package that started out life in our laboratory and is now used routinely in hundreds of hospitals around the world. Mirada Solutions was so successful that it was acquired in 2003 by CTI Molecular Imaging, and then in 2005 CTI was acquired by Siemens, so our little company is now the Advanced Applications Laboratory of Siemens Molecular Imaging (for a web page, Google on "Siemens molecular imaging")

I noted above that our work in medical imaging has been characterised from the outset by the development of models. For the first fourteen years those models were essentially rooted in mathematics, physics, and signal/image analysis – my comfort zone! However, four years ago it became clear that there is vastly more scope for modelling; but it required me, toward the end of my academic career, finally to learn some biochemistry! We are at the threshold of a major development in medicine and pharmaceutical research, in which the insights gained over the past decades through molecular biology and medicine will be integrated with novel imaging devices and image analysis techniques. This will herald a significant paradigm shift, for example in oncological image analysis and in cancer biology, with analogous developments in neurology and brain science, endocrinology, etc. More precisely, research will centre on the

**visual representation, characterisation, and quantification of biological processes at the cellular and subcellular levels within intact living subjects. Molecular Imaging necessitates the close collaboration of scientists from a broad range of disciplines. The name for this new synthesis is *Molecular Imaging*. Whereas conventional imaging techniques rely on (non-specific) physical or physiological characteristics of tissues to obtain contrast, *molecular imaging aims to exploit (specific) molecules as the source of image contrast*. We have begun our work on molecular imaging by studying: the oxygenation status of tumours; the way tumours can control the acidity of their local environment to advantage their growth; the ways multiply drug-resistant tumours defend themselves against attack by chemotherapy drugs; and the opportunities afforded by dynamic PET. I believe that molecular imaging – the integration of molecular biology with image analysis and novel imaging technologies – is the great breaking wave. If only, I could be 18 again!!**

**You may ask, how can we work on problems such as tumour hypoxia which span computing, physics, engineering science, and biology. I suppose that the conventional approach would be to assemble a large and diverse team of experts, and then appoint a manager. Fortunately, there is an alternative approach, which I hope by now will not surprise you: it is to train young minds - undergraduate and**

**graduate students - to be driven by a set of key scientific challenges and to disregard subject barriers which are put in place by universities. I have learned quite a bit about the biochemistry of cancer by learning from my graduate students who trained in biology and who I have, in return, taught some mathematics. The scientists who have most inspired me - Newton, Gauss, Einstein, Feynman, Crick - paid no heed at all to subject boundaries; they just tackled tough problems and learned whatever was necessary to solve them, or they discovered the necessary science for themselves. At Oxford University we have been running a Doctoral Training Programme (http://www.lsi.ox.ac.uk/) for 5 years which is a four year programme that takes students from a range of disciplines – mathematics, physics, computer science, engineering, biology, … - and trains them in subjects such as genetics, mathematical modelling, programming, image and signal analysis, … so that they can work at the interface between the Physical and Life Sciences. The graduate students I have supervised from this programme have taught me a great deal.**

## Grid

**One of the novel technologies that is emerging concurrently with molecular imaging is Grid computing. In essence, Grid computing strives**

**to develop a set of standards, known as middleware, that will enable new kinds of distributed computing. It is likely to develop alongside the service-oriented architecture of mobile phones. In one incarnation of Grid computing, a user may wish to set off a complex, computationally intensive algorithm, for example a non-steady flow calculation or the dynamic creation of an atlas of a human organ (eg the brain) from several thousand images that need to be co-registered. A Grid realisation would distribute this computation to many, perhaps thousands, of computers that are perhaps geographically widely spaced and without worrying whether they are running Windows, Linux, or whatever. As you may imagine, the more you pay, the faster you will get the results! Of course, there are already middleware standards to support Grid computing, and they are getting better and more widespread each year. I am sure that 5 years from now most PCs will have middleware built in, just like local area network software/ hardware is now.**

**There have been several pioneering Grid projects, including iXi, CancerGrid, BIRN and the EU HealthGrid projects. We have worked on two such: eDiamond and MammoGrid. I will mention eDiamond briefly because it is relevant to the ISI. In the UK, over 2 million mammograms are taken in 93 breast imaging centres (and in mobile mammography machines). The hospitals taking those**

mammograms are legally required to curate those images. Suppose now that we want to develop either a pattern classifier to distinguish different classes tumours, or we want to build a data mining system. Some years ago, I realised that there is a problem of scale: on the average, there are (happily) only about 6 tumours per thousand screening mammograms, and on the average, an imaging centre sees about 20,000 images. That is, each centre sees, on the average, only 120 tumour instances – but there are many kinds of tumour – and so there is insufficient statistical power in any given centre from which to develop our pattern classifier. Imagine, however, that you could connect *all* of the imaging centres to form a large federated database – federated because, of course, the original mammograms need to remain in the centre where they were taken. If we could create such a federated database, then we would expect to see 12,000 tumour instances annually, and we would have enough statistical power! Grid computing enables us to realise the vision of a federated database, because the middleware standards can be the basis of secure transmission of the mammogram images (which, being medical records, are subject to strict legal constraints). The eDiamond project, involving several universities, hospitals, and two companies (IBM, Mirada) developed the basis for such a federated database. Now we wait for technology to catch up! Grid computing will transform late stage clinical trials of drugs; see yet another company I am involved in, and

**which was a start-up from UCL and Imperial: http://www.ixico.com/**

## Information Engineering and the ISI

**I hope I have communicated to you my understanding of what Information Engineering is; it has taken me just 21 years to answer the question I posed at my Oxford University interview! I hope also that I have convinced you that statistics is a technology that permeates Information Engineering, so you are all, in this sense, information engineers! I encourage the researchers of the ISI to apply their deep theoretical insights about uncertainty and statistics to the tough problems that confront especially India. Our work has always been done within the context of a specific application, with application-specific criteria for success. We have worked hard to translate the fruits of our research, not just to small scale clinical trials but to commercialisation. Writing research papers is fun; but it just doesn't compare to the thrill of watching your science being used to diagnose and help treat a patient, or control a mobile robot navigating around a factory! I hope that the great Prasanta Chandra Mahalanobis would have approved!**

**Science and Engineering are as international as much of modern business is. India has the massive intellectual resources not to be a**

**feeding ground for great universities in the USA and Europe; but to build itself as a powerhouse that will attract the brightest minds *from* the USA and Europe, who will also be attracted by your cultural diversity, the colours, the sounds, and the history. I like to think that over the next few years India has the opportunity to create a hugely scaled up version of what Oxford University, the UK, and Europe developed in response to the Information Engineering explosion in the 1980s. You, the graduating students, have your lives ahead of you, so go and realise your full potentials, and never accept academic or other unnecessary barriers to progress.**