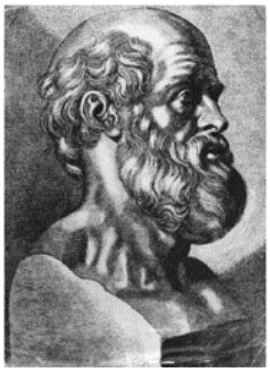
Hippocrates and the Beast of Business

Inaugural Lecture of Duncan Gillies
Presented on the 28th February 2007
Imperial College



Hippocrates (circa 460-370 BC)

The Beast of Business (circa 1960 AD)

Historical Introduction

One of the lessons I have learned in thirty years of lecturing to students is that if you take great care in preparing lectures, finding the best intuitions to explain a technique and refining the equations to their simplest form, then the students will learn a good deal, but will think of you as a little naive - "he only teaching the simple things" will be their thoughts. On the other hand, if you do the opposite, rush the preparation, put up slides of badly formulated obscure equations and wave your arms wildly to avoid giving a proper explanation then they will learn nothing but will leave the lecture thinking "what a clever fellow he is". In the light of this experience, I formulated the rule that to be obscure is to gain respect, and chose for my title "Hippocrates and the Beast of Business" to pose a puzzle, make you think that the lecture would contain insights into the ancient world and culture, and possibly tell a fable to entertain and give moral instruction. Had I called it "Computers in Medicine" you would have had a clear indication of its contents, but would doubtless have thought it a dull insipid topic unworthy of your attention. So, in order not to disappoint you I shall begin with Hippocrates.

Hippocrates of Cos, whose image, as conceived by Rembrandt, is seen in the picture above, lived from around 460BC to 370BC. He is widely regarded as the founder of modern medicine and is honoured by each newly graduating doctor taking the Hippocratic oath. In his teachings he rejected common superstitions that illness was either caused by an evil spirit or was a punishment inflicted by the gods. Instead he advocated that disease was an imbalance of the body's "humours" caused by a combination of environmental factors, diet and living habits. His therapies were based on the idea of restoring the correct balance of the humours by rest and restraint. He was generally opposed to intervention and avoided the use of drugs if possible. His methods are nowadays sometimes known as "holistic medicine" based on the principle of Aristotle that the whole human is more than the sum of the parts.

Hippocratic ideas were accepted for more than two thousand years after his death, but came under increasing attack during the nineteenth and twentieth centuries.

The doctors of the Victorian period increasingly began to take a reductionist approach to the subject. This involved isolating a single cause for a disease and treating it directly. The emergence of anaesthetics and antiseptics at the end of the nineteenth century paved the way for spectacular successes in the field of surgery. The discovery of antibiotics early in the twentieth century similarly revolutionised medicine. By the middle of the twentieth century it was a commonly held view that reductionism was key to the cure for all known diseases, and that it was simply a matter of time before the scourge of illness would be eliminated completely. However, that was not to be as bacteria mutated to escape the power of the antibiotics, and viruses adapted to find new ways to live off human cells. Hippocratic ideas are again becoming popular with practitioners, in the United Kingdom at least, advocating rest and good diet as a cure, and avoiding the use of antibiotics wherever possible.

Next to Hippocrates is a picture of the Beast of Business, as it appeared in 1960. This is the Control Data Corporation 1460 computer - then the fastest and most powerful computer ever built - designed by Seymour Cray one of the most innovative computer designers of the twentieth century. The Beast of Business was a nickname for the computer from that time reflecting its growing use and importance in commerce, though the 1460 was built as a military, not a business, machine. In the 1960s the computer successfully took over many clerical office duties. It would print the payroll, send out invoices, print envelopes for mailing lists, keep stock inventories and a whole host of other like activities, much faster and much more accurately than ever a clerk could. It was a reductionist machine par excellence, solving isolated tasks to perfection. Yet at that time the world of computing and the world of medicine had barely become acquainted. All that was to change dramatically in my forty years of working with the Beast.

It is easy to forget just how primitive the computer was in 1960. Typically computers could execute four thousand instructions a second - today they can execute four thousand million instructions each second. They were built then from single transistors and diodes - the idea of making a device containing more than one transistor had only emerged in 1958. Nowadays they are built from a few integrated circuit devices each containing several million transistors. Memories in the 1960s were made from magnetic beads, threaded with three wires and would typically store around four thousand instructions in a bulky and expensive module. Today, a cheap memory card holds a thousand million bits. The large tape drives for storage have been replaced by faster, bigger capacity cards the size of a coin. As the computer mutated to its present form so it increasingly found new tasks within its grasp.

Figure 2

In the nineteen sixties the computer, though crude, was driving change and ideas. Manufacturers were building systems that were smaller, cheaper and more applicable outside the business world. The Digital Equipment Corporation PdP-8 computer, shown figure 2, introduced in 1965, was intended for engineering applications, and was built with medium-scale integrated (MSI) circuits each containing hundreds of transistors. The arrival PdP-8 was an important moment for the computer in medicine because it opened the door for dedicated or embedded computing. A computer could be dedicated to a single engineering task, such as the monitoring and control of a process plant, or some diagnostic purpose. Software advances were driven by the possibilities offered by the new technology. It is hardly a coincidence that the fast Fourier transform algorithm, of vital importance in medical imaging, was published in 1965. The theory behind it, which exploits the symmetric properties of sine waves, was known in the time of Fourier. The world famous paper by Cooley and Tukey was fundamentally concerned, not with theory, but with how to implement the Fourier transform algorithm on a small computer. New thinking in hardware and software opened the door for the centuries most important advance in medical imaging - computed tomography (CT) - which was demonstrated in 1970, and to new approaches to analysing and characterising medical signals. In the ten short years from 1960-1970 the computer had become an essential element in medical practice. Its importance has continued to grow from these beginnings to the present day.

There are many ways in which one can describe and classify the tasks undertaken by the computer in medicine. I have grouped them into five categories: retrieval, prosthetics, diagnosis, simulation and analysis. The name prosthetics is being used here in a figurative rather than a literal sense to describe tasks in which the computer extends the capability of the practitioner to assess or treat the patient, for example by being able to view the internal structures of the body through medical imaging. Of these five tasks the first three are fundamentally reductionist in nature, and have been accomplished with astounding success. The last two, analysis and diagnosis, are more holistic in nature and have enjoyed some success.

Retrieval

One of the surprising things about the use of the computer for information retrieval was the length of time that elapsed before any effective patient records systems emerged. It is tempting to think that the business software of the 1960s could be easily adapted into hospital information systems, but this did not happen. Even as late as 1990, although there were many computer systems supporting records in specialist units in hospitals, there was no agreed health record system for general practice. The European union then funded a large research project, including 21 participant institutions, to define the prefect medical record, but in the end it was de-facto systems and standards that were finally applied. Very recently Fujitsu admitted that there were still major problems with their new patient records system for the NHS. The reasons for this are many. Health records are very diverse in the range and type of information they contain. Practitioners were slow to adopt new technology - the use of the keyboard for data entry was, for a long time, considered demeaning by some senior consultants. Computer scientists failed to recognise the difficulties, and were tempted to adopt the old-fashioned data processing approach of centralised computing, rather than distributed databases. Modern approaches have overcome many of these difficulties, with data being distributed widely and accessible through

the internet. Such systems are easy to implement and easy to access, and are becoming increasingly important in medical research as we shall see later.

Prosthetics

In contrast to retrieval, the prosthetic extensions offered to the clinician expanded dramatically in the 1970s, most notably in the field of imaging. Magnetic resonance imaging was demonstrated in 1976, and has been developing ever since. Some of the advances were gained by improved engineering of the magnets and coils, but much of the progress resulted from the development of computational techniques for reconstruction and visualisation of data. Other techniques of internal viewing were gaining ground at the same time - in particular flexible endoscopes were becoming widely used. A skilled endoscopist can insert a scope a distance of over one meter inside the body to examine the colon or the upper gastrointestinal tract, reaching as far as the caecum or the duodenum. However, there are difficulties with the technique. The endoscope itself is a difficult instrument to use, requiring the simultaneous control of two steering wheels and several buttons with the right hand while pushing or pulling the scope with the left hand. Initially it took two practitioners to perform an endoscopic examination. The unfavourable environment inside the human body complements the poor ergonomics of the scope. Even a well-prepared colon contains fluids that suddenly obscure the view, and is inclined to collapse in spasm at regular intervals. Its most awkward characteristic is its mobility. This makes advancing the endoscope difficult, and impossible in certain patients. Loops can form which cause paradoxical behaviour of the endoscope - pushing giving the appearance of moving the scope backwards - and have inherent danger in traumatising the colon wall. In 1984 I started working with consultant endoscopist Christopher Williams to see what the computer could do to circumvent some of these problems.

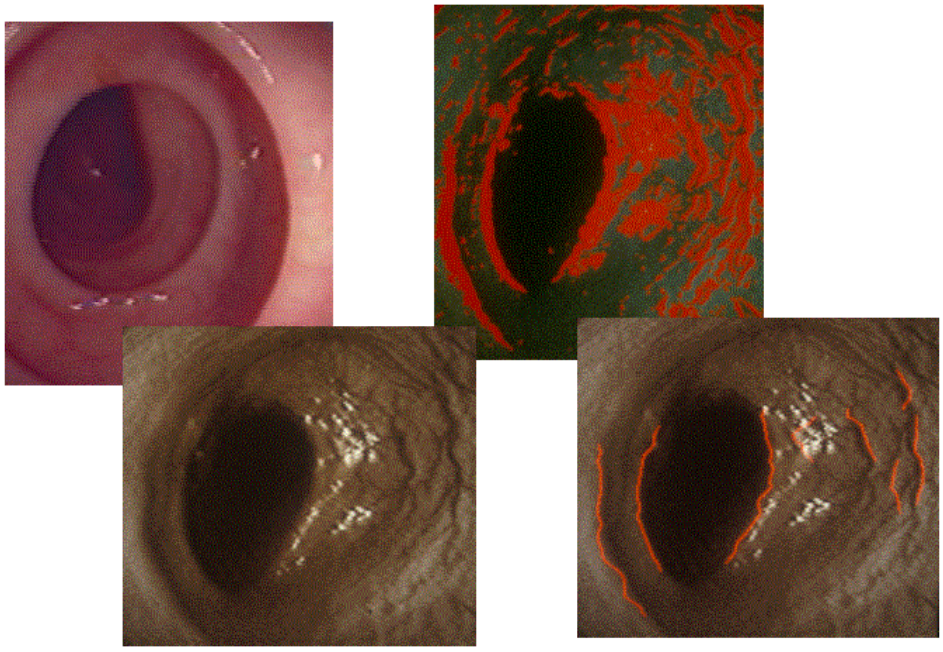


Figure 3

One of our main objectives was to design a system that would steer the endoscope automatically, freeing the consultant from the distracting task of using the thumb wheels. If we could design a system that would direct the endoscope towards the lumen, or centre line of the colon, the clinician would be left free to concentrate on the important task of observing the walls for early signs of cancer. Detecting the

lumen in an image of the colon is a computer vision problem, and like all problems of the sort is solved by identifying intrinsic visual characteristics of the object that we are seeking to detect. We made use of three characteristics of the lumen, which can be seen in the colon image on the top left corner of figure 3. First, it is darker than the rest of the image, being further from the light source. Secondly concentric rings formed by the muscles of the colon wall surround it, and thirdly, the normal vectors at the colon wall will on average meet at the lumen. This third property is more readily observed in the image on the bottom left hand side. Of these the detection of a large dark region proved the easiest and most effective method of steering the endoscope providing the lumen is in view. As this is not always the case, we still needed to make use of the other two characteristics. The detection of the concentric rings proved an interesting and difficult task, and is a good example of how generally applicable results can result from studying practical problems. The picture on the upper right hand side of figure 3 shows the output of a standard edge detector applied to a colon image. The contours are visible, because the light intensity changes across them, but there are so many other edge points - caused by a variety of discontinuities such as vein patterns and light reflections - that the computer could not isolate the contours. Together with Gul Khan, from whose PhD thesis figure 3 originated, we came up with a solution based on perceptual grouping. By this principal we grouped together edge points if their perceptual characteristics agreed. These characteristics could be, similarity in intensity, similarity in directionality, similarity of colour and so on. We also made use of continuity by grouping together edge points if they formed a continuous smooth line. The results are shown in the bottom right picture. The contour lines have been detected, and all but a few other edge points have been removed. By changing the perceptual criteria we can detect other image features that may be weak but perceptually significant.

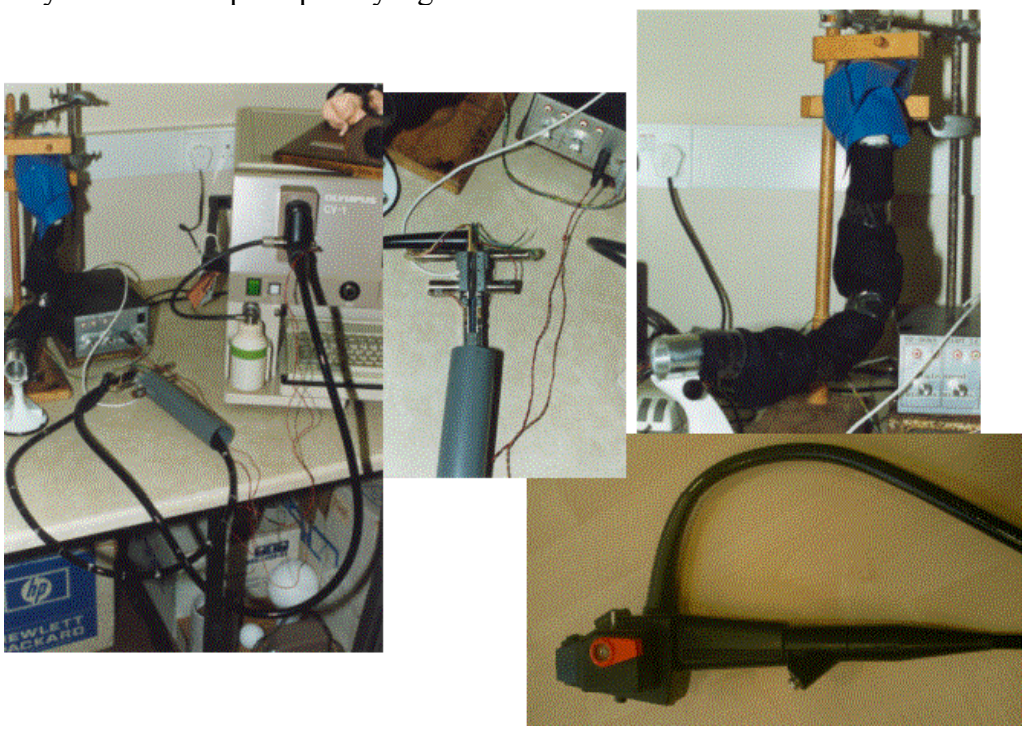


Figure 4
Figure 4 shows some of the test rig that we built to try out the self steering endoscope, and on the bottom left hand side is a servo driven endoscope (no wheels to turn) built by the Olympus Optical company for this work.

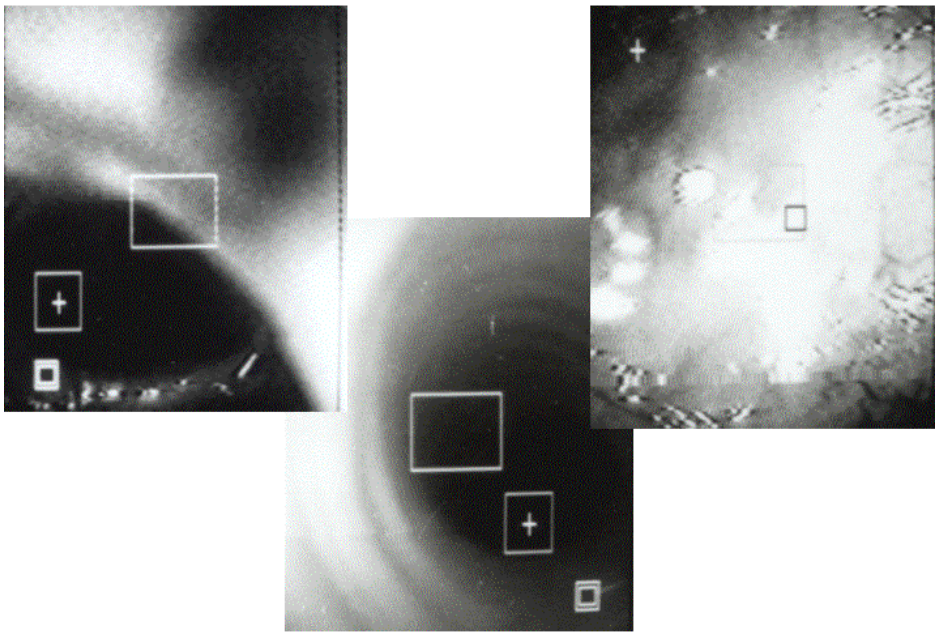


Figure 5

There is a problem in having several different algorithms to compute the same property, and that is they don't always agree. Figure 5, from Enrique Sucar's PhD thesis, shows some examples of this, taken from our experimental endoscopy system. On the left and in the centre are images where the lumen is in view. The large square indicates the direction that the endoscope points in, the plus shows the estimate based on the detection of a large dark region, and the small square the detection based on the perpendiculars to the colon wall. In the first two examples the agree, more or less, with the inaccuracies in the perpendicular algorithm resulting from the large areas of specular reflection which lack any reconstruction information. However in the third image, the two methods disagree completely, and not surprisingly since a human could not easily interpret this image. This exposes the problem of how to fuse the results of different algorithms together to come up with the best single estimate of the lumen position. It is the point where diagnosis and prognostics merge. In the third image, the clinician can observe two possible candidates for the lumen, but knows that only one can be correct, since the colon is an un-branching tube. Making this decision requires visual diagnosis. There is also another possible interpretation of the image that one or both of the black regions may be diverticula - or pockets in the colon wall. In certain circumstances diverticula can look very similar to the colon lumen. Clearly we need to be able to diagnose diverticula disease from the visual characteristics of the colon images and avoid pushing the endoscope into a diverticulum, which could well result in a fatality.

Diagnosis

Reasoning about medical data always involves a degree of uncertainty, and many other computer vision problems are also classic examples of reasoning under uncertainty. The most obvious approach to deal with uncertainty is to use some form of probabilistic inference. Thus, to guide our endoscope, and at the same time diagnose divertuicula disease we turned to the work of the Rev. Thomas Bayes whose picture is seen below. Probability theory is very old. It has been studied as long as man has had an interest in games of chance, which means as long as civilisation has existed. Bayes' theorem is the fundamental theorem of probability and it expresses the relation between conditional and unconditional events. Proving Bayes' theorem is

very simple. Suppose we wish to know the probability of two events occurring which we denote S and D. We write:

$$P(S\&D) = P(S)\,P(D|S)$$

that is the probability of both events occurring is the probability of S occurring, times the probability of D occurring given that we know S has occurred. We could equally well write:

$$P(S\&D) = P(D)\,P(S|D)$$

and equating the two forms we get:

$$P(S)\,P(D|S) = P(D)\,P(S|D)$$

or

$$P(D|S) = P(D)\,P(S|D)/P(S).$$

(Bayes theorem)

Now read disease for D and symptoms for S and immediately we see the utility of the equation for probabilistic inference. The left hand side is the probability of a disease given the symptoms - exactly what we want for diagnosis - and the right hand side is made up of probabilities we can measure from case histories: The probability of the disease and the symptoms, and the probability of the symptoms occurring given the disease. Most importantly the right hand side can be broken into two parts. One is the probability of the disease divided by the probability of the symptoms. This expresses established or prior knowledge about the domain of interest. The other part P(S|D) expresses a causal link. We say that the disease causes the symptoms with a certain probability. Both parts are crucial for inference. For example there is a very high probability that a patient who suffers from tuberculosis will have coughing as a symptom. However, a doctor observing a patient coughing will not immediately diagnose tuberculosis - unless of course the doctor is watching an Italian opera. This is because the prior probability of the disease is very low.

Bayes' theorem gives us a simple fundamental mechanism for diagnostic reasoning. However, in practice its application is far from simple since diseases have many symptoms that are normally linked together through a complex causal chain. Until the 1980s causal chains were expressed through sets of equations, which were difficult to understand and to manipulate, and as a consequence probabilistic inference was hardly ever used in medical applications. Instead computer scientists adopted *ad-hoc* ways of expressing uncertainty in diagnosis. In the 1980's Judea Pearl of UCLA developed a graphical representation of causal chains, with probability propagation rules based on Bayes' theorem, which was called a Bayesian network. This was a major breakthrough since these networks had both the expressive power and the strong theoretical foundation that were required for medical inference problems.

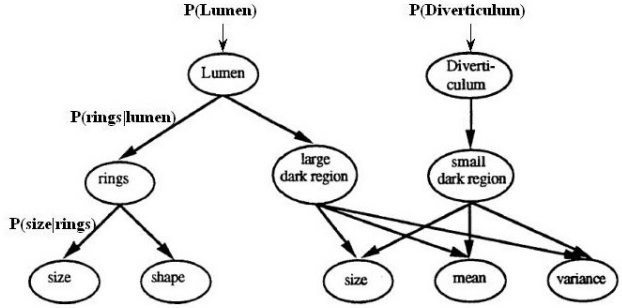


Figure 6

Working with Enrique Sucar and my brother Donald, we designed a Bayesian network for reasoning about colonoscopy images, part of which is seen in figure 6. Each arc represents a cause, or in probabilistic terms a conditional probability. We can calculate probabilities by applying Bayes' theorem independently at each node. The expressive power of the diagram is obvious. If we had written this out as a set of equations it would have been very difficult, if not impossible to understand its structure or reason about it with the domain experts. However, the simplicity is gained at the cost of some underlying assumptions that we must be aware of. When testing the system based on this network we made an interesting observation, namely that if we deleted either the node "mean" or the node "variance" the performance accuracy increased. This may look paradoxical at first, because we are removing information from the network, yet gaining performance, but the reason is quite simple. In a video image there is a relationship between mean and variance. If the endoscopist increases the brightness during an examination all the pixels intensities increase, hence the mean of any large dark region increases. But also, the range over which the intensity changes also increases, and thus the measured variance increases. Thus there is a relationship between mean and variance that we have not included in this diagram, and to make the network correct it is necessary to include another arc expressing this causal chain. This is bad news computationally since Pearl's elegant propagation equations cannot, in general, cope with loops. Other algorithms can, but all other methods become computationally infeasible as the number of arcs gets larger and larger. Considerable progress has been made into computational methods for solving large multipy connected Bayesian networks. However there is a further problem as the number of arcs in a network increases. If we consider a slightly more complex network, and add to it all the causal links that exist, we end up with a structure that is not only computationally infeasible, but also intractable to any form of interpretation or reasoning.

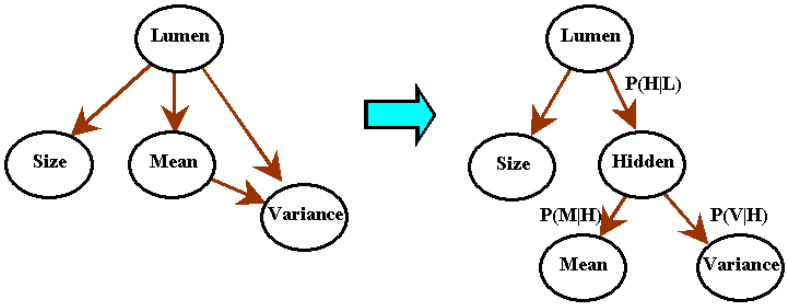
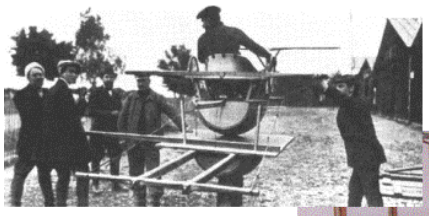


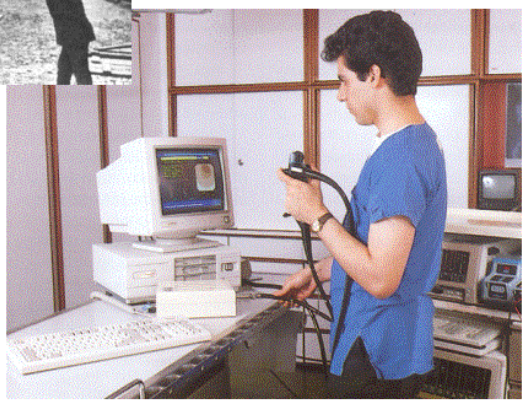
Figure 7

I was then working with Chee Keong Kwoh, and we adopted another approach to this problem. The idea was to change the structure of the network to incorporate the troublesome dependencies in a different way. Using our previous example, we explicitly model the relation between mean and variance in video images by adding a new node, which is called a hidden node. The process is shown in figure 7. The name "hidden" indicates that it is not one of our measured variables and doesn't appear in any data set. The problem we tackled was how do we find the conditional probabilities of the hidden node without any data. The solution was to use an optimisation procedure. We devised an algorithm based on gradient descent, and out method was found to provide high computational accuracy at a fast speed. The method generalises from the simple case shown to any arbitrarily connected network, and works well in medical diagnosis where usually we can rely on a lot of data to estimate the conditional probabilities.

Simulation

We turn now to a different problem - simulation of biological systems- but still with endoscopy as an exemplar. Aircraft pilots have been trained using simulators since 1920, and with computer based simulators since the 1960s, so it was an appealing idea in the 1980s to use the computer to train doctors to carry out difficult procedures.





There is an underlying problem with this idea and that is that the computer must replicate the medical procedure accurately. If it does not then the training may well be counter productive. This makes the problem quite different from a computer game. An aircraft behaves in a way that is governed by fluid mechanics - a field that has been studied extensively and is well understood. It responds to controls in a uniform and predictable way, subject to some random displacements that can be simulated easily. By contrast, as we have seen already, an endoscope moves through the human colon in a complex way. Its behaviour is determined by the shape of the colon, the friction at the walls, the reactive forces caused by squashing the small intestine, which surrounds the colon and the configuration of the endoscope. The biggest difficulty of all is that the endoscope deforms the colon as the procedure is carried out. There is little or no data available on the mechanical properties of living tissue and what is known implies that the behaviour is complex non linear with time dependent properties.

However, there are two features of colonoscopy that make it amenable to computer simulation. Firstly it is a keyhole procedure. The consultant views the world through a video monitor, or an eyepiece, which is similar. Consequently, we don't need any of the paraphernalia of virtual reality systems to create an animation. We can interact through a normal computer monitor. Secondly the consultant never sees the deformation he is causing. He only looks at the view in front of the endoscope. I was working with Angelo Haritsis on the simulation. We abandoned the idea of creating a mechanical model to describe how the endoscope moves through the colon, and instead adopted a behavioural model. The behavioural model works by simulating what the trainee endoscopist sees rather than what he does. If the scope forms a loop, it behaves in a paradoxical fashion. It becomes harder to insert, and pushing it expands the loop and causes the tip to move backwards. In our simulation, once the endoscope is inserted beyond a critical distance we can choose to put it into a different mode of behaviour whereby we apply the brake a little and cause the tip to

move backwards as the trainee pushes in. The trainee then has to follow the correct procedure to un-loop the endoscope. All modes of behaviour encountered during colonoscopy can be encoded in the same way, and can be invoked with certain probabilities to create easy or difficult cases.

All the time I worked in this field the beast of business was evolving, getting faster. In 1984 the best we could achieve was a simple line drawing representing the colon. Five years later we could produce reasonable realistic shaded colons. We knew how to create visually realistic images, but couldn't do it in animation. Modern endoscopy simulation systems can now achieve this. The graphics has become more beautiful, but little further progress has been made on the modelling.

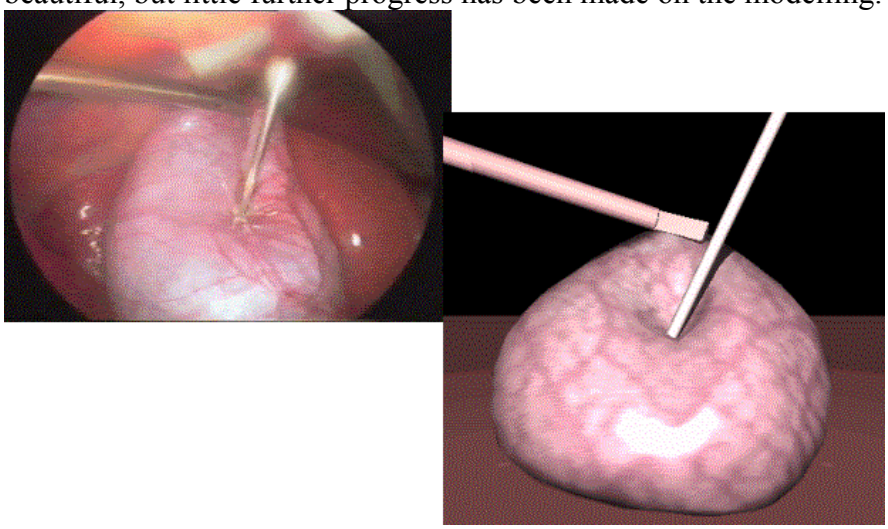


Figure 8

Simulation of virtually all other medical procedures require that the computer shows the deformations accurately, and even where this is possible the computational demands are high. Figure 8, from Dino Moutsopoulos' thesis, shows some work that we did in simulating laparoscopic cholesysectomy. Here we have a complex object, the gall bladder, partly filled with fluid, which is grabbed and pulled from side to side and finally cut off during the procedure. Our simulation computed the behaviour of an elastic membrane, and the deformations look plausible, yet we know that the real system is more complex. Cutting the gall bladder cannot be reproduced accurately. Moreover, the creation of tactile feedback is an enormous problem. At the time we were working on this simulation it was not possible to compute even a simple elastic model fast enough for animation, and that remains true today if accurate simulation is required.

We see that behind simulation for medical training there is the difficult problem of modelling the underlying biological system. There are other important uses for biomedical data modelling that were becoming apparent at that time. One of these was the area of pre-operative prediction of difficulties that might occur in medical procedures. Here we have the leeway to compute the model accurately as we no longer have the time constraint imposed by creating an animation.

In 1995 I began working with Peter Charters on preoperative prediction of patients that would present difficulties to the anaesthetists. One of the most common procedures carried out by anaesthetists is laryngoscopy in which the anaesthetist inserts a tube into the larynx through which a mixture of oxygen and anaesthetic gas can be administered to the patient keeping him or her safely asleep during the operation. To do this it is necessary to compress and displace the tongue so that the vocal chords and larynx can be seen. As this is an unpleasant process the patient is

sent to sleep first by injecting a small amount of muscle relaxant. Figure 9 shows the procedure in progress, the view of the vocal chords and the centre of the human airway in an MRI image.

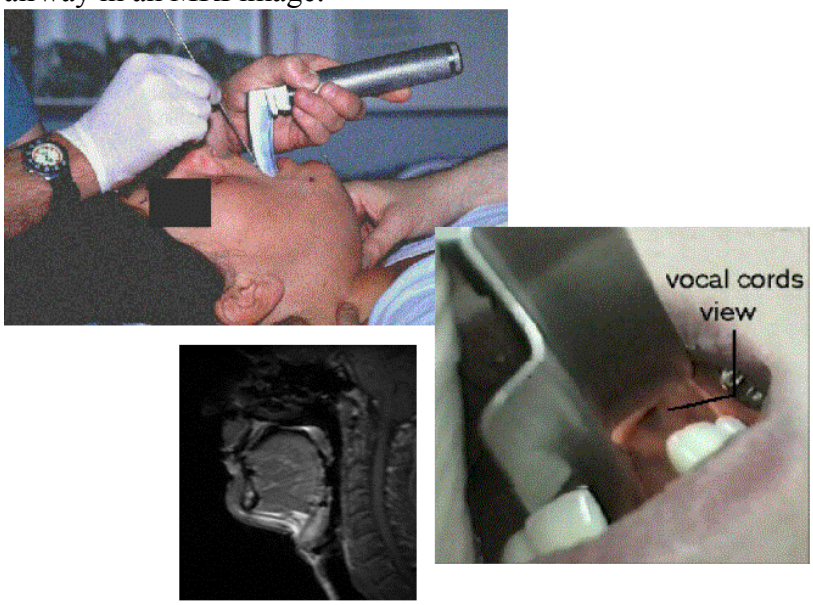


Figure 9

Laryngoscopy sounds quite simple and straightforward, but the process carries the biggest risk associated with anaesthetics. Depending on the shape of the human airway it is perfectly possible that the larynx cannot be seen and if the anaesthetist is struggling to compress the tongue there is a risk that the incisor teeth may be broken. Worse still, but fortunately extremely rarely the tongue may, in its relaxed state block the airway and the lack of oxygen cause brain damage or even a fatality. For these reasons there is a considerable interest among anaesthetists in what is called “the difficult airway”.

Working with, Andreia Rodrigues, we began an investigation into whether we could simulate how the tongue would be deformed during laryngoscopy. The method we used was a very well established technique in mechanical engineering called finite element analysis. If the shape of an object is too complex to solve directly for its mechanical behaviour it can be broken up into small simple elements, and the deformations of those parts can be computed individually. There is an inherent danger in taking an engineering technique and applying it to a biological system since the biological system is always much ore complex than one can possibly imagine. Engineering materials can be made to be uniform and to conform to well known laws, such as Hooke’s law of elastic deformation. Biological tissue isn’t uniform and behaves in a complex, non-linear time dependent way. Moreover, very little relevant data on the mechanical properties of living tissue is available. So we are left to make a well-informed guess as to how the deformation will occur, and then validate the results experimentally. By taking this approach we were able to build a plausible simulation, and to set the parameters of the simulation so that it replicates a real procedure. However, the results are preliminary and require much more validation before we can confidently make a prediction.

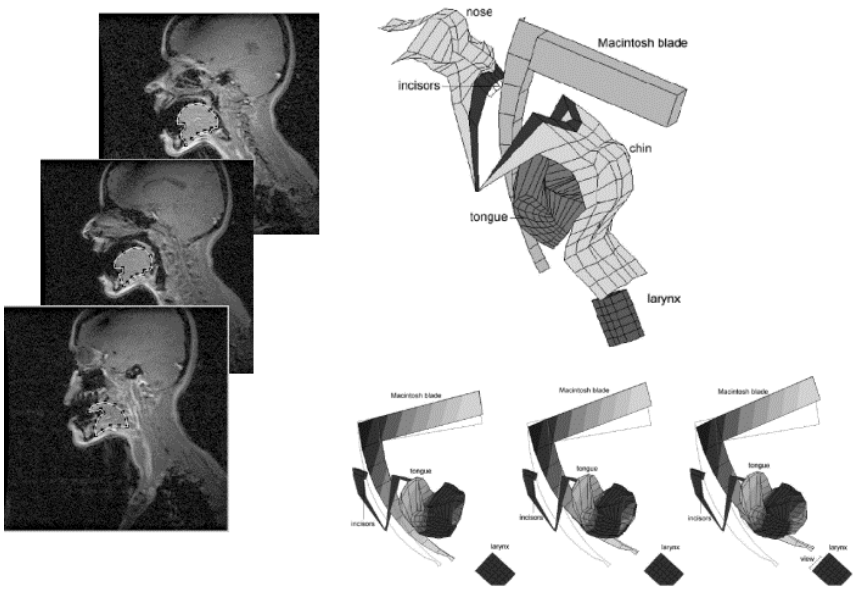


Figure 10

The process of simulating a single patient's airway is very difficult and time consuming. It is illustrated in figure 10, which was extracted from Andreia Rodrigues' PhD thesis. The patient must be scanned first. Then the individual 2D image slices must be manually segmented to find the shape of the tongue and the mandible, which limits the deformation, the position of the incisor teeth, which occlude the view, and the position of the larynx whose visibility we are trying to determine. Once all that data is assembled we need to create the finite element model, decide the procedure for applying force from the laryngoscope blade and the element properties, and only then can we do a simulation. A skilled technician might be able to achieve all this in two or three weeks. So unless we can devise a way of automating the process the technique will be too expensive to apply in practice. Our approach was to try to create a generic model of the upper airway.

The idea of a generic model is to encode the shape of the airway in a computer model such that by setting a small number of parameters we can generate a model that matches any specific patient. When we wish to simulate a particular patient we make a small number of critical measurements and adjust the parameters of the generic model accordingly. Once the generic model matches the specific patient we carry out the simulation to determine whether he or she will be difficult to anaesthetise. This is a simple idea, but a difficult task. Working with Krista Lam we were able to create a generic model of the mandible bone. Interestingly most of the shape variation can be described in just four modes. The biggest variation comes from size changes, and after that shape differences between male and female. We found that we could reconstruct an unseen mandible with sufficient accuracy using just seven measurements which can be made with simple callipers. The next stage of the project is to determine whether we can similarly reconstruct the tongue and put together these two components into an evolving upper airway model.

This is a good point to reflect on how the role of the computer in the world of medicine has been evolving. The earlier projects that I have described are reductionist in nature. They make small but useful contributions to medical practice such as steering an endoscope or making a simulation of the observed behaviour during colonoscopy. In trying to create a generic airway we are now aiming at simulating a whole system. It has several quite different structures that interact - the tongue, the mandible the incisor teeth, the hyoid bone, the epiglottis, the soft palate and divers

muscles and ligaments. Several modelling methods are needed including shape and geometric models, finite elements and kinematics. All of this cannot be done without the computer, and we know that we will not be able to predict difficult laryngoscopy without the computer model, because the difficulty is a result of many complex interacting factors. Equally well there are doubts about the underlying computer modelling that we cannot easily resolve without considering the behaviour of the whole system and comparing it to the real procedure. The role of the computer is changing from reductionist to holistic.

Analysis

We come now to a current topic that I'm working on in conjunction with Georgia Chan, Alok Mishra and Dave Thornley. One of the most researched subjects of modern medicine is genetics, and the work in this field now cannot be undertaken without the computer.

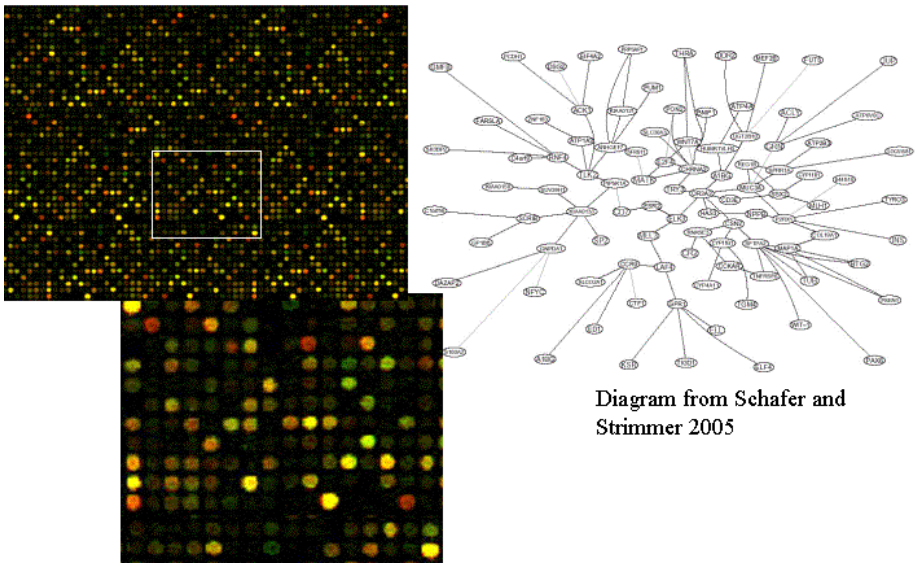


Diagram from Schafer and Strimmer 2005

Figure 11

Figure 11 shows a picture of a micro-array. This is a glass slide on which there is an array of dots, each dot measuring the activity of a particular gene. Genetic material is extracted from a sample of interest and washed over the array. The genes that are active in the sample react with one particular spot. This example is from a differential experiment in which two different samples are colour coded, one red and one green. Spots that show up as red correspond to genes that are active only in one sample, those that show as green correspond to genes that are active only the other sample and those that show as yellow indicate genes that are active in both samples. The technique could be used, for example, to compare cancerous and normal cells. On the right hand side is a dependency diagram extracted from a set of micro-array experiments. Each circle represents one gene and each arc represents some dependency. The dependency diagram is one way of expressing our understanding of how the genes interact. It is a starting point from which higher level descriptions of gene behaviour could be discovered.

Micro-array analysis is a classic example of what is referred to as the small sample size problem. We have a very large number of variables, the genes, whose behaviour we wish to study, but only a small number of micro-array experiments to use as data. Typically we may have fifty micro-array experiments and several thousand genes in each. Worse than that, the experiments are full of variance due to

experimental error, and the dependency diagram is complex, as can be seen in the example of figure 11, where the number of genes considered is only a small subset of the active genes in the experiment. One standard statistical approach to this sort of problem is to increase the sample size, but unfortunately this does not work well. The experiments are difficult to carry out and expensive, and mixing the results of different micro-array experiments destroys significant information.

Between the micro-array and the diagram is an involved set of computer processes. It is necessary to find the dots, and measure their optical density, compensate for flaws in the slides and variations in the reaction conditions and finally come up with a number for each gene activity. Only then can statisticians begin to build the dependency diagram. There are many ways of expressing the dependency and the classic one using the covariance matrix in which the diagonal elements represent the degree to which each gene varies in its activity, and the off diagonal elements indicate how each gene pair vary together. In the case of small sample size problems the co-variance matrix is poorly estimated. If we blindly extract a dependency diagram from it then it will not be a robust result. Perturbing the data to the degree we expect to happen through error in an experiment will change the diagram dramatically. This means that we must stabilise the result by trying to reduce the effect of the less certain data. A standard statistical means of doing this is the technique known as shrinkage. For example, we may believe that, in the co-variance matrix, the diagonal data is better estimated than the off diagonal data. This would not be an unreasonable assumption for micro-arrays, since the diagonal elements of the covariance matrix simply tell us how much each individual gene varies. Co-variance estimates, in contrast, may be confused by time and rate factors in the underlying biological behaviour. We can therefore make our estimate more reliable by shrinking the matrix towards the diagonal. We do this by reducing the off diagonal elements and increasing the on diagonal elements until we obtain a robust result.

Shrinking towards the diagonal is a crude way of solving the problem. What we would prefer to do it to invoke the aid of Thomas Bayes again, and shrink our experimental observations towards some expression of prior knowledge. Although this has not been possible until recently there are now emerging “curated” data bases which store information on known gene interactions. Information in these data bases is obtained from specific experiments targeting hypotheses about gene activity rather than the broad unfocussed approach of the micro-array. An approach of this kind creates many problems for the computer scientist such as how to resolve conflicting data and how to characterise the context within which a dependency can be assumed to hold. Finding and encoding the prior knowledge will be the first step in a new approach to describing gene dependency. I recently worked with Carlos Thomaz on small sample size problems. Carlos made the astute observation that shrinkage methods, even those that use prior information will always destroy some vital information about a process. Significant parts of the data from an experiment will be diluted to the same degree as the unreliable data whose influence we seek to reduce. Instead of shrinking Carlos developed a technique in which we maximise the information content, or entropy, of the co-variance matrix. This works by changing the representation of the data such that we can identify the strongest dependencies expressed in an experiment and also those expressed in the prior knowledge. We then simply select the strongest patterns of dependency from each to form our co-variance estimate. In his thesis Carlos demonstrated that, for biometric recognition, the maximum entropy method was considerably faster to compute than all previous methods of estimating co-variance, and in almost all cases was more accurate. If we

can find a way of characterising the prior knowledge then this idea offers real potential for extracting robust and accurate dependency structures from micro-array data.

Prognosis

In summing up I would like to draw attention to one of the most important pieces of research in the history of medicine. The gentleman in uniform is Ronald Ross an army doctor who served his queen and country in India. While there he discovered the mechanism by which malaria was spread. He did this by dissecting mosquitoes that had been fed on the blood of malaria patients using his own dissection kit and a microscope with a cracked lens. He found the conclusive evidence that the malaria parasite lodged itself and grew in the stomach of the mosquito and was able to transmit from human to human through the feeding of the insect. It took him nearly two years of failure before he finally found the conclusive evidence. Ronald Ross was able to achieve his results largely by himself and through his own tenacity. His work in many respects should be looked on as scientific, rather than medical. The idea that insects might transmit malaria was not originally his. He reviewed the research literature, chose a hypothesis and set out experiments to test it. He provided one major link in a holistic solution to the problem. Things in the modern world have changed. No amount of staring down a microscope at a micro-array slide will reveal a jot of information. Progress can only be made through the collaboration of different experts drawn from medicine, science, statistics and computing. Nowadays the volume and complexity of the data we analyse is beyond the understanding of a single individual, and we must rely far more on the work of each other on different components of the same problem. Thus more than ever we need to follow the example of Ronald Ross and look at the whole system behaviour to enable us to assess the validity of each part. The ability of the computer to store and retrieve information is vital in creating the common framework in which this can be achieved.

In the last forty years the beast of business has changed from being an adjunct to medicine to being an essential component, and its ever expanding role has brought to light many difficult and challenging research problems. Now in addition to its reductionist prowess it is providing the means whereby we can take an holistic approach to medical research and follow in the footsteps of Hippocrates.

Acknowledgements

Throughout my career I have worked with many original and innovative thinkers. To list them all would be very difficult. In an attempt to do this, the list below is made up of those with whom I have worked directly either in writing a paper, a grant proposal or a doctoral thesis.

Hassan Amin, John Baillie, Jung Wook Bang, Fernando Bello, David Booth, David Borislaw, Dan Boswell, Michael Bourmpos, Peter Burger, Georgia Chan, Peter Charters, Jung Tai Chen, John Collinge, Peter Cotton, Don Cowell, Ara Darzi, Vivek Datta, Abhinandan Deb, Laura Dempre-Marco, Tellis Dosis, Eddie Edwards, Raul Fietosa, Donald Gillies, Peter Groom, Chris Guy, David Hansell, Angelo Haritsis, Daniel Heesh, John Hoare, Alan Howsan, Cho Chi Huang, David Hykin, Imdad Ali Ismaili, Paul Jowell, Agnes Kaposi, Gul Khan, Jieun Kim, Chee Keong Kwoh, Krista Lam, Sean Mackay, Jay Malhotra, Richard Marks, Hugh Markus, Salman Marvasti, Farok Marvasti, Andrew McKinnon, Jacey Lynn Minoi, Alok Mishra, John Mizon, Dino Moutsopoulos, Bob Nicholls, Maja Pantic, Alex Pappas, Andrew Poon, Anil Rao, Haroon Rashid, Ortwin Rave, Tony Reno, Shah Jehan Riaz, Simon Robertshaw, Andreia Rodrigues, Daniel Rueckert, Steffan Rueger, Brian Saunders, Bruce Sayers, Hock Soon Seah, Azidah Shafebegli, Mel Slater, Enrique Sucar, Chris Sumner, Tarkan Tahseen, Eldos Then, Surapa Thiemjarus, Carlos Thomaz, David Thornley, Dave Weston, Christopher Williams , Guang Zhong Yang, Alexei Yavlinsky

There are many who have supported me outside the workplace who deserve equal thanks. Prominent among these is my family, over three generations, parents, siblings and children and especially my wife Glenda.

Lastly, working in the Department of Computing at Imperial College has been a constant pleasure. My colleagues have been a source of help, support and encouragement and inspiration.

Thank you all.

References

Khan G N and Gillies D F "Extracting Contours by Perceptual Grouping" Image and Vision Computing **10** (2), (1992), 77-87.

Gillies D F, Haritsis A and Williams C B "Computer Simulation for Teaching Endoscopic Procedures Endoscopy **24**(Suppl. 5), (1992).

Sucar L E, Gillies D F, Gillies D A, "Objective Probabilities in Expert Systems" Artificial Intelligence **61** 187-208 (1993).

Kwoh C K and Gillies D F "Using Hidden Nodes in Bayesian Networks" Artificial Intelligence **88** 1-38 1996.

Rodrigues M A F, Gillies D F and Charters P "A biomechanical Model of the Upper Airways for Simulating Laryngoscopy" Computer Methods in Biomechanics and Biomedical engineering **4**(2) 127-148 (2001)

Thomaz C. E., Gillies D. F. and Feitosa R. Q. ”A New Covariance Estimate for Bayesian Classifiers in Biometric Recognition” IEEE Transactions on Circuits and Systems for Video Technology Special Issue on Biometrics 14(2) 214-223 2004.