AISB Quarterly No.123

The Newsletter of the Society for the Study of Artificial Intelligence and Simulation of Behaviour

Searching for the neural basis of action selection

The problem of action selection can be simply stated as: "how does an agent decide what to do next?" More specifically, we may ask: "what resolves the competition between functional units, with each requesting an action, for control of a set of effectors?" This problem of conflict resolution is equally relevant to situations as diverse as software agents controlling network traffic, robots navigating hazardous environments, and the behaviour of foraging animals. Thus, if a common, robust, action-selection mechanism can be found it would have many applications.

One promising method for finding such a mechanism is to see how the problem has already been solved by evolution—i.e. how animals embody neurally-based solutions to the action-selection problem. Neuroscientists have dedicated much effort to the basal ganglia, a group of nuclei spread amongst the front and centre of the vertebrate brain. In simulation, we have shown that the internal wiring of the basal ganglia creates a system capable of action selection and switching based on the urgency of those actions. Further, we have shown that co-ordinated, complex behaviour of a robot could be controlled by our basal ganglia model,[1] thereby demonstrating the potential general application of a basal-ganglia-based action-selection mechanism.

However, the basal ganglia is not a *necessary* component of the neural action-selection mechanism. Both decerebrate animals (with only brainstem intact) and altricial (helpless at birth) neonates lack a functioning basal ganglia but are capable of spontaneous behaviour and appropriate, co-ordinated responses to stimuli. Such animals clearly have some form of intact system for simple action selection. Is there, then, a brainstem substrate for action selection?

The cybernetics pioneer Warren McCulloch thought that there was. In 1969, he and his colleagues proposed that the medial reticular formation (mRF) (see Figure 1a) was configured to select the overall behavioural state of an animal. Their landmark computational model demonstrated signal selection in their interpretation of the mRF's anatomy. We thus tested this model as a robot controller (a long-held wish of the original authors) to evaluate its potential as an action-selection mechanism. We found that their model in its original form could not sustain action selection, but that, by evolving the model with a genetic algorithm, certain configurations could be found which did.[2]

Inevitably, given its age, aspects of the model were incorrect or omitted features known from more recent studies of the mRF. We thus turned to creating a new model, first establishing the *structure* of the mRF. Following an extensive review, we proposed that the mRF contained two neuron classes—projection and inter-neurons—which were arranged in intermingled clusters of cells (see Figure 1b).

By proposing a model which described this structure, we were able to assess the network properties of the mRF, treating the neurons and their inter-connections as nodes and links on a graph. We found that, to the extent the model does capture the mRF's anatomy, the mRF is a small-world network:[3] this is the first demonstration that such a network is formed by individual neurons of the vertebrate brain. A small-world network has two defining properties: its nodes are more clustered—more locally inter-connected—than would be expected if

Humphries and Prescott
University of Sheffield

Continued on p. 4

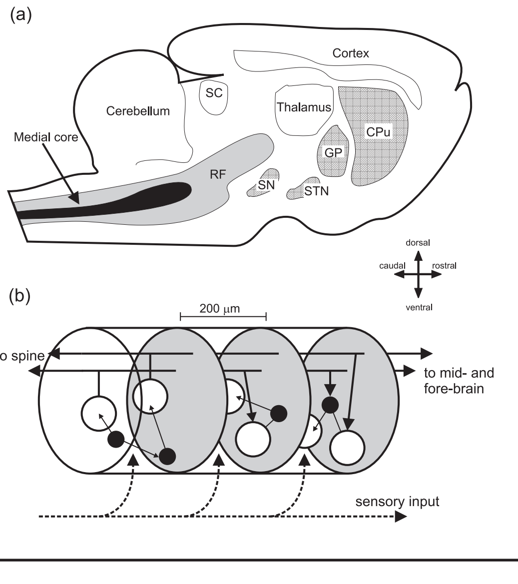


Figure 1. (a) Sagittal section (cut vertically front-to-back) of rat brain showing major nuclei including basal ganglia (CPu, GP, STN, SNr) and the medial reticular formation (RF). (b) Proposed cluster organisation of the medial RF. Projection neurons (large) and inter-neurons (small) are inter-mingled, inter-neurons project only within a cluster, projection neurons only out of a cluster. Cluster bounds (grey ovals) defined by first collateral of projection neurons' axons (solid line). The projection neuron axon collaterals connect the clusters together. This anatomy forms a small-world network.

In this issue:

Features

Society

Reviews

Father Hacker

Intelligent computation is more than is 1s and 0s

The **AISB Quarterly** is published by the Society for the Study of Artificial Intelligence and Simulation of Behaviour.

ISSN 0268-4179

Editor
Sunny Bains
EEE Department
Imperial College London
aisbq@aisb.org.uk
http://www.aisb.org.uk
and click on AISBQ.

Editorial Assistant
Stuart Barr

Page Layout
Freddy B-Apeagyei
Logic Media, London

Advertising and Administration
Therie Hendrey-Seabrook
aisbq@aisb.org.uk
School of Science and Technology
University of Sussex
Falmer, Brighton
UK, BN1 9QH
Tel: (01273) 678448
Fax: (01273) 671320

Much of what we have called 'Artificial Intelligence' in recent years has been carried out (on the implementation side, anyway) by programmers, computer scientists, and even mathematicians. However, as someone who is interested in embodied intelligent machines (robots), I see much of this work as ignoring or side-stepping the most important (and interesting) engineering problems: how intelligent creatures physically adapt to the outside world using their sensors.

I understand that there are good reasons for ignoring these issues if you are interested in modelling some kind of higher-level reasoning. However, I would like to argue that, when it comes to actually implementing a genuinely intelligent robot, sensor-processing will become one of the most difficult and critical issues. So critical, in fact, that the way you implement the sensor processing may well have a knock-on effect on the way you implement *all other processing* in the machine.

The basic argument is simple. To adapt intelligently to the environment, you need to be able to gather information about it. To adapt intelligently to an *unknown* environment and then perform unknown tasks, you must minimize the pre-conceptions you have about the structure of the information that you are collecting. Specifically, as you cannot know for sure what sensor resolution you will need for your unknown task in your unknown enviroment, it makes no sense to try to set this resolution in advance.

Most sensors are inherently analogue devices, their outputs made digital only by an a/d converter. This acts as a filter, throwing away information. Two signals that were not identical in analogue become so in digital. This information *may* not be useful, but we don't know for sure. And, after we convert, we'll never be able to analyze whether it was useful or not because we've already thrown it away.

In my work,[1] I have shown that not only is this information potentially useful, but that it was not eradicated by noise. This might seem to go against communication/information theory, but it doesn't: our *confidence* in the information is certainly eroded, but it turns out that we can manage our information-gathering process to overcome this problem. Essentially, we can listen harder, or longer. But none of this is possible if we have thrown the information away in advance.

Which is all a long way of saying that not only are sensors analogue, but it makes sense that the processing of the sensors be analogue too. At a Royal Society lecture, Christ Toumazou, an electrical engineer and now head of the Institiue of Biomedical Engineering here at Imperial College, summed up this concept for me. In order to interact with the world, every system had to have what he called an 'analogue shell': a way of taking information from the outside, analogue world and feeding it into the (presumably digital) inner world.

But this begs the question: where do you draw the line? Where does it make sense to put that analogue to digital converter? Potentially, the analogue shell could thicken to the point where not just everything outside is analogue, but everything inside too. Since our own neural systems are analogue (in terms of spike timing, at least) this does not seem so far fetched.

Realistically, I do not believe that we can do without *any* viable technology if we want to succeed in building embodied intelligence. We need to look at every task and choose the fastest, lightest, most-power-efficient solution for the job. This includes using clever mechanical engineering for compliant limbs, optics for dense neural interconnections, and adaptive analogue processing of sensory inputs. And it may also include higher-level reasoning implemented using the kind of heuristics we know and love.

What is interesting is that, potentially, analogue can do without digital. The same cannot be said the other way around. This is a fact that, so far, surprisingly few people in AI are even paying lip service to.

Sunny Bains
Imperial College London
Editor, AISB Quarterly (now retired!)
http://www.sunnybains.com

Reference

[1] See thesis and other publications at:
http://www.sunnybains.com/index.php?page=publications.htm

A final editorial

Having decided to step down as editor, this is my last issue of the AISB Quarterly. I would like first to thank my successor, Colin Johnson of the University of Kent at Canterbury, for taking on the position. I know he will do a wonderful job: his book reviews have been a great boone to the publication, as have the ideas he has sent in as a member of the Editorial Board.

I would also like to thank two members of the committee in particular for their help and support during my tenure: Geraint Wiggins (former Chair), and Louise Dennis, the current Secretary and Webmaster of the Society (who, incidentally, I think would make a fantastic future Chair!).

I wish the rest of you well, and hope you will help/encourage the committee to take the AISB from strength to strength. There are some great projects in the pipleine: perhaps most exciting on the publication side is the proposal to turn Father Hacker's writings into a book. I look forward to seeing all these exciting developments come to fruition. ***—SB***

Biologically-inspired image processing for a robotic grasping task

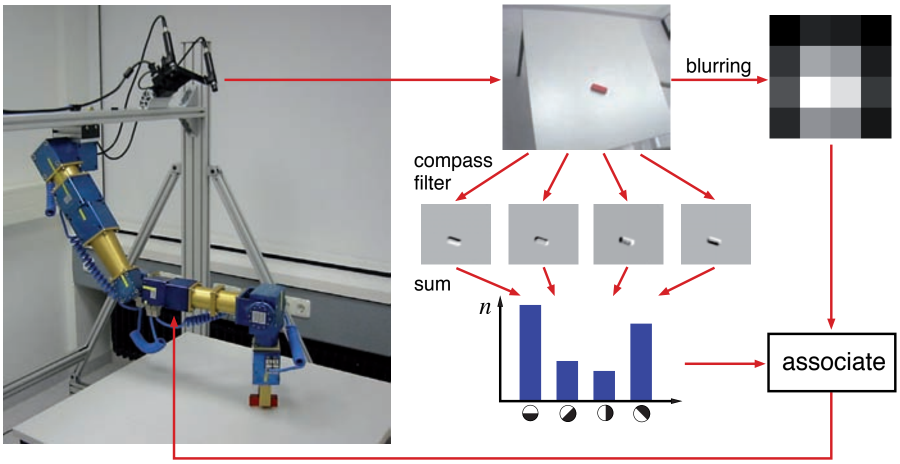


Figure 1. Shown is the information flow in the grasping task. The processing of the camera image is split into two parts. First, to extract position information, the image is blurred and sub-sampled. Second, to extract orientation information, four different compass filters (directional edge filters) extract edges in different directions. The sum of the white pixels in each of the four filtered images results in a histogram of edge distribution. This histogram, together with the blurred image, is associated with an arm posture that enables the robot to grasp the observed object.

Visual processing in mammals is adapted to their behavioral needs: likewise, in visually-guided robots, image processing needs to be suitable for a desired behavior. Thus, the function of the mammal brain may be a good guideline for choosing the right image-processing techniques for machines. In our work, we make robots learn through experience and thereby study which learning and image-processing techniques lead to a good performance for a given task.

Here, we describe a study in which our goal was to make a robot arm grasp an object presented visually.[1] The robot learned to associate the image of an object with an arm posture suitable for grasping. Learning an association means that there are no world coordinates and there is no tedious calibration of the vision system, instead, the robot learns by randomly exploring different arm postures and by observing the appearance of objects put on a table. Though the emphasis of our work is on learning techniques, here we will focus on the image processing.

We used a robot arm with six joints and a gripper: the vision system was a stereo camera head mounted on a pan-tilt unit (see Figure 1). This setup was located behind a table, which was the operational space and which was visible to the cameras. In training, the robot placed a red brick on the table in random positions and, for each position, recorded an image of the scene after removing the arm. Thus, the training set contains corresponding pairs of grasping postures and object images.

An image can be interpreted as a point in a high-dimensional space (with the number of dimensions equal to the number of pixels). A mapping from such a space to an arm posture suffers from the so-called 'curse of dimensionality': the distance between pair-wise different images is almost constant, and the orientation of the target gets lost under the dominance of the positional information.[2] Therefore, the image must be pre-processed.

The processing technique that was eventually successful was inspired by the function of the visual cortex. The image processing was split into two parts: one for the object's location and one for its orientation (see Figure 1). To decode the location, the image was first blurred and sub-sampled. Since here the target (the brick) was almost point-like within the camera image, the blurred image is like a population code of the brick's position. In a population code, many neurons encode a parameter: such a code for the retinal location of a stimulus exists also in the primary visual cortex.[3]

To decode the orientation, image filters were used to extract edges in different directions: for each, we counted the edge pixels within the image. This sum was invariant of the brick's position and was a measure of how close the brick was to a given orientation. Position invariance and orientation tuning are also properties of V1 complex cells.[4]

The resulting visual information could be used to first learn and then to recall the association with an appropriate arm posture for grasping (Figure 2). Specifically, the decomposition of the image process-

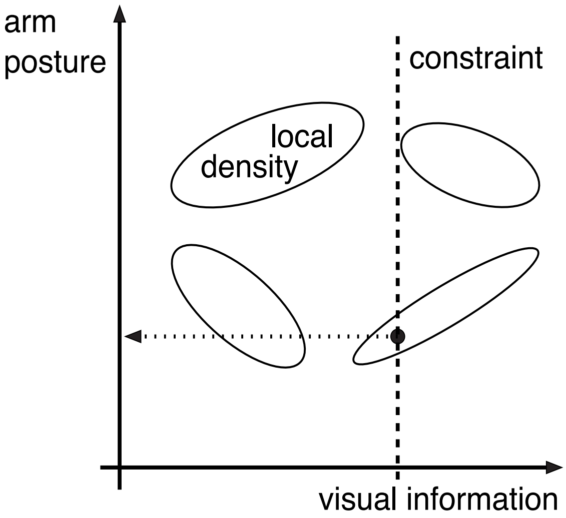


Figure 2. Pattern association. Training patterns lie in the product space of arm posture and visual information. The density of the pattern's distribution is modeled by a mixture of Gaussian functions (ellipses are iso-density curves). To map the visual information onto an arm posture, we define the output space as a constrained space anchored at the input. On this subspace, the highest local density gives the desired output.

***Heiko Hoffman**, Max Planck Institute for Human Cognitive and Brain Sciences*

Continued on p. 11

The Biologically Inspired Robotics Network (Biro-net)

A growing research community in the UK is successfully applying biological inspiration to robotics in research, industry and leisure. The observation of biological systems which deal effectively with complex problems that have taxed the efforts of robotic researchers has inspired the modelling of these systems to varying degrees. Some researchers choose to abstract the underlying biological details and use these models for inspiration to create more intelligent robots. Others take this further and apply these techniques to well-studied biological systems in order to build truly representative models. This requires a thorough understanding of the relevant biological field as well as the crucial and relevant aspects which will be abstracted into the robotic model. This requires close interaction with the relevant biological discipline, something which often leads to collaboration across many different disciplines, including computer science, artificial life, engineering, artificial intelligence, robotics, neuroscience, psychology, ethology, and related fields.

The Biologically Inspired Robotics Network (Biro-net) is an EPSRC-funded network (GR/S25340/01) which intends to draw together these diverse communities. Due to the nature of the subject, it is vital to facilitate interdisciplinary research and gain access to the knowledge and experience of researchers currently unaware of biologically inspired robotics. The network aimed to do this by; providing a dedicated forum for communication and exchange of ideas through a set of coordinated initiatives which include: organised symposia and workshops; providing a web-based resource centre; facilitating collaboration via symposia and lab visits; facilitating publications for special issues of internationally recognised journals; and advertising the field across a range of related research disciplines.

The network has now been running successfully for two and a half years and is due to complete at the end of April 2006. To date, the network has run four symposia hosted at different sites and in several formats. The first was an introductory meeting intended as a brainstorming session which was held at the University of the West of England. This was followed by a one day symposium held before TAROS '04 at the University of Essex, which hosted two established invited speakers and allowed young researchers to present their work. The third symposium, held at the University of Bath at Easter 2005, broadened the target audience with the scope defined as *Inner and Outer Space* and successfully attracted a more diverse audience. The final symposium run so far took place as a session within TAROS'05 (See http://biro-net.aber.ac.uk/events.php for details of the symposia series). Our final event will be a symposium run under the AISB umbrella which will run 3-4th April 2006. If you are interested in this area, watch the AISB pages for the biologically inspired robotics symposium (http://www.aisb.org.uk/convention/aisb06/).

A web site, newsletter, forum and e-mail service has been set up to disseminate information about the area throughout the community. For further details see http://biro-net.aber.ac.uk. You're invited to join the network (free!) and receive our e-mail updates on current events. Funding is still available to support inter-laboratory collaboration visits. See the web site for more details.

Myra Wilson
Department of Computer Science
University of Wales, Aberystwyth
E-mail mxw@aber.ac.uk

Searching for the neural basis of action selection

Continued from p. 1

the same number of total links were made at random; its nodes are also linked by shorter paths than would be expected if the same number of total links were made uniformly. These structural properties of a small-world network in turn imply dynamic properties—of rapid cross-network synchronisation, consistent stabilisation, and persistent activity—that may all be critical to the representation of actions, and to their co-ordination.

Our research is now focusing on a dynamic model of the mRF, in which the links between the model neurons are defined by the structural model. With this we are looking at the most likely form of action representation supported by the mRF. We are also able to look at how the basal ganglia and mRF may interact: whether they work together or whether the basal ganglia over-ride the mRF. Eventually we hope to obtain an understanding of the complete neural basis for action selection in the vertebrate brain.

This research was funded by the EPSRC grant GR/R95722/01.

Mark Humphries and Tony Prescott
Adaptive Behaviour Research Group
Department of Psychology, University of Sheffield
E-mail: {m.d.humphries, t.j.prescott}@shef.ac.uk

References

1. T. J. Prescott, F. M. Montes-Gonzalez, K. Gurney, M. D. Humphries and P. Redgrave, *A robot model of the basal ganglia: Behavior and intrinsic processing,* **Neural Netw.,** in press.
2. M. D. Humphries, K. Gurney and T. J. Prescott, *Is there an integrative center in the vertebrate brain-stem?,* **Adaptive Behavior 13**, p. 97, 2005.
3. M. D. Humphries, K. Gurney and T. J. Prescott, *The brainstem reticular formation is a small-world, not scale-free, network,* **Proc. R. Soc. Lond. B,** in press.

Automatic generation of constraint models: The quest for fully-automated solving of combinatorial problems

The dream of researchers in constraint programming is that one day computers will be able to automatically solve combinatorial problems stated naturally in some language that is easy to learn and to use. The aim of our research is to make this dream of automated problem solving a reality.

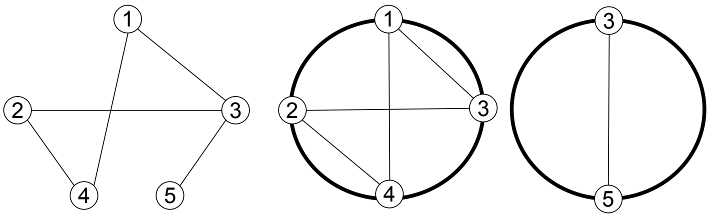


Figure 1. *An instance of the SONET fibre-optic communication problem: communication demand graph and solution on two rings. Ring capacity is four nodes.*

Though finite-domain constraint programming has proven to be an effective technology for solving a wide range of important combinatorial problems, including scheduling, timetabling, configuration and allocation problems, it does not provide fully-automated problem solving. In particular, solving a problem with constraint technology requires *modelling* it by:
1) a set of decision variables, whose values must be found, and
2) a set of constraints on the variables that characterise the solutions to the problem.

Effective modelling requires a great deal of expertise and is currently performed manually.

Our approach to developing an automated problem-solving system is to develop an abstract language in which combinatorial problems can be specified naturally, to develop a system that automatically generates an effective model from a problem specification, and finally to employ existing constraint solvers to find solutions to the model. Thus far, we have developed a prototype specification language, ESSENCE,[1] and a prototype model generation system, CONJURE.[2]

Combinatorial problems often require finding some complex combinatorial structure. Consider, for example, the Sonet fibre-optic communication problem[3] (see Figure 1), which requires finding an installation of client nodes onto one or more Synchronous Optical Network (SONET) rings such that any two nodes that need to communicate must share at least one common ring. Thus, a solution requires finding the set of nodes that is to be installed on each ring. Since there is a set of rings, the combinatorial structure that must be found is a set of sets of nodes satisfying the communication constraints. The key feature of ESSENCE is that it directly supports decision variables whose values are complex combinatorial structures. This enables problems to be stated directly and naturally. Without the use of decision variables with the appropriate type of value, the problem would have to be modelled by encoding the desired combinatorial object as a collection of constrained decision variables of some other type of value.

ESSENCE supports a wide range of type constructors, including sets, multisets, relations, functions and partitions. Uniquely among constraint languages, these constructors can be used to define types that are nested to an arbitrary depth; for example, a decision variable can be of type set of integers, multiset of sets of integers, set of multisets of sets of integers, and so forth.

Current constraint solvers can handle decision variables whose values are only atomic elements (and perhaps some limited extensions). Therefore, the central task in modelling most combinatorial problems is to translate, or *refine*, a problem conceived as a set of constraints on complex decision variables into a set of constraints on atomic variables. We have formalised this process by a set of recursive refinement rules, which are embedded in CONJURE. In formulating the refinement rules, we have overcome two primary difficulties and many secondary ones. The first difficulty arises because expressions, particularly decision variables of non-atomic types, can usually, if not always, be refined in multiple ways. Furthermore, the refinement of an operator depends on how its operands are refined. The second major difficulty arises from arbitrary nesting of types. Suppose that A and B are sets of some deeply-nested type and we wish to refine the constraint $A = B$. Such a constraint would involve all components of both A and B. Furthermore, we wish to produce refinements in which A and B are not refined in the same way.

Our current implementation of CONJURE generates models for a small but useful fragment of ESSENCE. In tests on seven problems from the literature, including the SONET problem, CONJURE was able to generate a set of correct models for each problem including many of those produced by human experts.

Besides completing our implementation of CONJURE, many other challenges remain. Principal among these is the need to develop heuristics that enable CONJURE to select among the set of models that it generates.

Ian Miguel is supported by a Royal Academy of Engineering/EPSRC Research Fellowship.

Alan M. Frisch and Ian Miguel*
Artificial Intelligence Group
Department of Computer Science
University of York.
E-mail: frisch@cs.york.ac.uk
*School of Computer Science
University of St Andrews
E-mail: ianm@dcs.st-and.ac.uk
http://www.cs.york.ac.uk/aig/constraints/AutoModel/

References

1. A.M. Frisch, M. Grum, C. Jefferson, B. Martinez-Hernandez and I. Miguel, *The Essence of ESSENCE: A Constraint Language for Specifying Combinatorial Problems,* **Proc. 4th Int’l Workshop on Modelling and Reformulating Constraint Satisfaction Problems,** pp. 73-88, 2005.
2. A.M. Frisch, C. Jefferson, B. Martinez-Hernandez and I. Miguel. *The Rules of Constraint Modelling*, **Proc. 19th Int’l Joint Conf. on Artificial Intelligence (IJCAI),** pp. 109-116, 2005.
3. H.D. Sherali, J.C. Smith and Y. Lee. *A Branch-and-cut Algorithm for Solving an Intra-ring Synchronous Optical Network Design Problem Allowing Demand Splitting*, **INFORMS J. on Computing 12,** pp. 284-298, 2000.

AI and stylometric analysis

Stylometric analysis is the name given to the attribution of authorship to literary works by the use of (primarily) statistical methods. The question of authorship of a text is equivalent to a classification problem.

There is no absolute technique for authorship attribution, although some feature sets have produced more success than others. It requires several variables in order to separate out authors in a sufficiently large feature space. Figure 1 shows a function word-based discriminant analysis of three groups of letters by Thomas Jefferson. Group 3 consists of letters written later in time than groups 1 and 2 and is separated out along the first discriminant.

AI techniques can be applied to authorship attribution in three possible ways: in making the choice of feature set, in classification, and in extracting features from text.

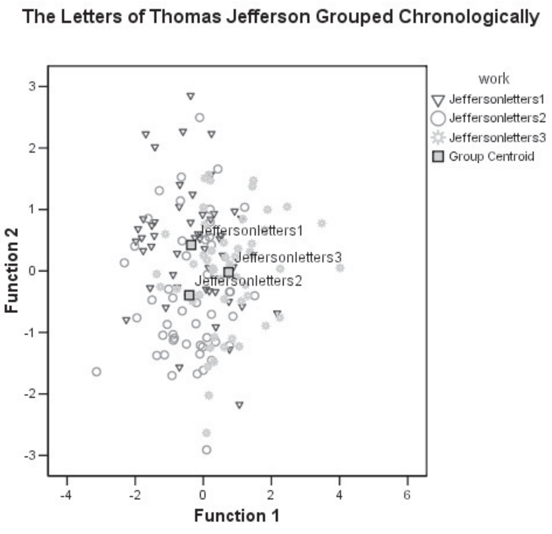
Choice of feature set

Features may be chosen from many different linguistic levels. For example, spoken language is analysed at the phonetic level in forensic cases of speaker identification. A forensic phonetician will carry out a detailed investigation of an audio recording of a reference and disputed transmission of speech. Some stylometric analysis techniques have made use of letter frequencies or letter collocations, but by far the most common linguistic level for feature sets is at the word level. Stylometrists have not stopped there, but have recently started to examine recurrent grammatical features and it is in this area that natural language processing can prove invaluable.

Word-level features are easy to understand and to automate, but not so easy to get right. Document classification may involve matching document vectors produced from lists of key words and applying a similarity metric, however the words that are of interest to stylometrists are very different. They are primarily interested in a small closed set of function words. These words include *and, the, in,* and so forth. However, no general agreement has been reached as to what exactly that closed set should include.

Additionally, questions remain about the approach, not the least of which is, why it works at all. Experiments by psycholinguists have suggested that although we have conscious control over our use of lexical (content) words, we seem to process function words differently and in a more automatic way. A second, not unrelated question is: why should we expect the distribution of function words to be different from one individual to the next?

To explore this, we need to look more deeply into the role that function words play in language. If we examine a word such as "that", by referring to a comprehensive grammar text (see Reference 1) we can easily discover 40 or more uses for the word. Indeed in Reference 2 we identified over 700 distinct uses for around 50 function words. Alternatively we might turn the question around completely and ask: "with the huge variety of usage for a limited number of function words, why should we expect everyone to use them in a totally uniform manner?" What is certainly true is that authors appear to have consistent patterns of function-word usage that hold over large quantities of text.



When we move on to consider function words in their grammatical context we are faced with a much larger feature set, along with the challenge of grammar-based feature extraction. In these cases, AI search methods can be very useful in helping to identify the correct feature set to employ and natural language processing can help with automated grammatical analysis of texts.

Classification of texts

Classification techniques have included the use of neural networks and naïve Bayesian classifiers as well as more traditional statistical methods such as discriminant analysis or principal components analysis.

Grammatical feature extraction

Up until recently, stylometrists have stopped short of using grammatical patterns, but recent work suggests both the need and a promise of success, for example.[2] A stylometrist might be interested in the pattern <noun> and <noun>, for example and conventional part-of-speech taggers can be helpful. But there is also an increasing need for far more sophisticated grammatical analysis in order to extract features. For example, in Reference 2 we identified a high frequency of usage of subordinate clauses as subject complement, by one speaker in a forensic text. It is these and other features that any natural language processing tool will need to be capable of identifying to be of use to a stylometrist. This is the challenge that lies ahead for stylometric analysis.

Peter W.H. Smith
Department of Computing
City University, London, UK
E-mail: peters@soi.city.ac.uk

References

1. R. Quirk, S. Greenbaum, G. Leech and J. Svartvik, **A Comprehensive Grammar of the English Language,** Longman, London. 1988.
2. P. W. H. Smith and G. De Jong, *Speaker Identification: Function Words and Beyond,* **Int'l Association of Forensic Linguistics Conf.,** Cardiff, 2005.

Mosaic World: Evolving a-life agents in an ecologically-relevant coloured environment

It is well known that the relationship between the image on the eye (visual stimulus) and its real-world source is not just complex, it is indeterminate. The reason is that any stimulus attribute is derived simultaneously from multiple attributes of the world (reflectance and illumination in the case of a stimulus' spectra, or size and distance in the case of visual projected angle, etc.). Consequentially, a given stimulus could represent an infinite number of possible combinations of sources. While this physical fact has historically been underestimated by the machine vision research community, overcoming stimulus ambiguity is the basic challenge that faces any natural or artificial sensory system.[1]

Though the mechanisms and computational principles for resolving this challenge remain unknown, recent neuroscience research suggests natural systems have evolved a strategy of encoding the probabilistic relationship between spectral stimuli and their sources in past experience.[1-3] If correct, then a necessarily corollary of this view is that obtaining a complete understanding of vision will require quantitative descriptions of an agent's neural architecture, as well as its history of sensory experience. This, however, creates a significant practical problem for rationalising natural visual systems, especially for humans in which the necessary information is unknown, if not unknowable. Research on artificial systems may, therefore, provide the necessary platform for explaining the 'probabilistic/empirical' basis of vision. Here we briefly introduce one such artificial system called Mosaic World (see Figure 1).

Mosaic World is a virtual space composed of 'coloured' surfaces, illuminants and a-life 'critters.[4,5] Critter behaviour is determined by the output of their multilayered, 3D artificial neural network, all attributes of which are evolvable, including its internal architecture and the spatial location and spectral sensitivity of its input receptors. In evolving the networks, there is no explicit fitness function. The system is *fully open-ended*: once the critter's network is initiated in a random configuration there is no further intervention by the experimenter. If a critter is to survive, it must evolve the ability to move (fast-to-slow), reproduce (sexually or asexually) and eat (with small-to-big bites).

A critter's metabolic rate is directly proportional to its activity and the size of its neural network. The higher the rate, the more resources a critter must obtain from its environment. In Mosaic World, resources are surfaces described by a reflectance efficiency function between 400nm and 700nm. The surfaces themselves are under multiple and dynamic illuminants, each of which is itself described by intensity function between 400nm and 700nm. As in nature, however, critters have no *direct* access to information about either the surface or illuminants. Instead, the only information they receive about their environment is the *product* of a reflectance and illumination. In this way Mosaic World preserves the fundamental ecological challenge that natural systems evolved to resolve.

We have now shown that—through evolution—these virtual agents can not only learn to overcome stimulus ambiguity, they also learn to balance reproduction, movement and resource consumption, which is necessary for survival.[4,5] By analysing the resulting network structures, information processing and visual behaviours of these artificial agents, our current aim is to use this system to better understand how natural systems (and future robotic systems could) resolve stimulus ambiguity, as a necessary steps towards building more robust visually guided robotic systems.

Ehud Schlessinger
and R. Beau Lotto
Inst. of Ophthalmology
University College London
E-mail: lotto@ucl.ac.uk
http://www.lottolab.org

References

1. D. Purves and R. B. Lotto, **Why We See What We Do: An Empirical Theory of Vision,** Sinauer Associates, Sunderland, MA, 2003.
2. R. B. Lotto and D. Purves, *The effects of color on brightness.* **Nat. Neurosci. 2,** pp. 1010-1014, 1999.
3. R. B. Lotto and D. Purves, *An empirical explanation of color contrast*, **PNAS 97**, pp. 12834-12839, 2000.
4. E. Schlessinger, P. J. Bentley and R. B. Lotto, *Evolving Visually Guided Agents in an Ambiguous Virtual World,* **Proc. Genetic and Evolutionary Computation Conf.,** June 25-29, 2005.
5. E. Schlessinger, P. J. Bentley and R. B. Lotto, *Analysing the Evolvability of Neural Network Agents through Structural Mutations,* **Proc. European Conf. on Artificial Life,** September 5-9, 2005.

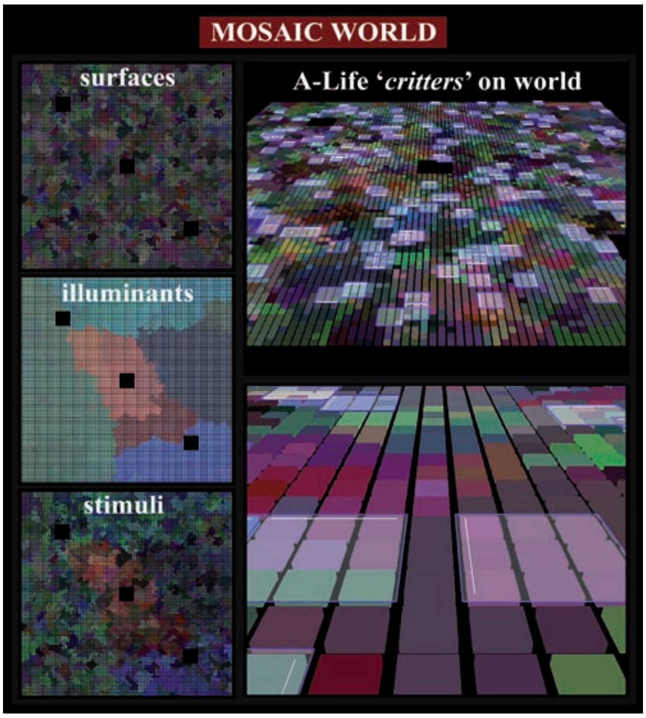


Figure 1. *A 'critter's-eye view' of Mosaic World: 'surfaces' in the upper left box are surfaces in Mosaic World under a dim, uniform white light. 'Illuminants' shows an example of a distribution illumination across the space at a given time-step. The 'stimuli' are the resultant combination of 'surfaces' and 'illuminants', which are presented to the agents. The image at the upper-right of the figure shows a population of 'critters' (which look like pieces of frosted glass), and are shown again at a higher 'magnification' in the lower right box. The black region surrounding Mosaic World, and the large three black regions running diagonally across its surface are 'holes'. (See colour version in online edition).*

How does a-life inform the mind-body problem?

Thought experimentation remains a popular method amongst some philosophers for investigating the relationship between mind and body. Consider for example: if a mad scientist sneaks into my bedroom at night and replaces my brain with the *Digi-Brain 2005*, do I still experience the world as *me* in the morning? How about if the mad scientist removes Jen's brain and deposits it into Jan's body? Does Jen still feel like herself, even in a new body, or does she now feel like Jan?

If your intuitions are not exactly clear in these cases, you are not alone. Metaphysical thought experiments like these are commonly criticized for being too far removed from reality, rendering any inferences drawn from them exceedingly suspect. But there is one problem with thought experiments of this kind that generally escapes criticism; namely that they presuppose the plausibility of mind-body dualism. That is, without the assumption that it is coherent to talk about the mind's being separated from the body, thought experiments like these cannot even get off the ground. Yet this assumption is wholly objectionable to those seeking a naturalistic account of mind, one that is committed to explaining the mind by appealing to nothing but the natural principles of life.

Embodied cognition research in a-life comprises a radically-new method for running nondualist thought experiments on the mind-body connection. Generally speaking, thought experiments test our intuitions about x by imagining a world like ours wherein x holds to try and uncover any lurking contradictions. The intuitive framework I advocate for thought experimentation on the mind-body connection is a particularly strong version of naturalism called the *strong continuity thesis* (SCT) of life and mind, first introduced in the a-life literature.[1] Weaker forms of naturalism prescribe some version of reducing the mind to the brain, leaving open the questions of how the brain relates to the body and to the world; SCT dissolves such questions by extending the ontological continuity further, maintaining that the very same organizational principles will be central to an understanding of both life and mind. It should be noted that SCT is meant to be not only a philosophical thesis, but also an empirical one. For example, physical explanations could be developed for how the natural process of morphogenesis directs in *similar* ways the physical growth of the brain and the rest of the body, but in *different* ways the storage of genetic information therein. Such an account would reflect the real philosophical import of SCT, i.e., that life and mind differ in degree and not kind, each a qualitatively different result from the same quantitative process. Herein the contribution of a-life would not be thought experimental, but empirical research on morphogenesis using computer simulation.

Keeping within the philosophical framework of SCT, my account extends the key notions in work on the nature of life to an analysis of mind. The related concepts of *autopoeisis*[2] and *supple adaptation*[3] emphasize the fact that an organism maintains itself as distinct from its environment despite continual, dynamic interaction with its environment. I adopt the related concept of a *self-preserving process* to characterize life and, crucially, extend this notion to the mind. If life and mind differ only in degree and not kind, then *the mind is a sophisticated facet of life's preserving itself*. Conceptualizing mind in this way constitutes a radical philosophical shift—while most current definitions of life (in biology) and mind (in philosophy) typically presuppose the phenomena to be things with attributes, SCT by contrast forces us to think about 'simple' life and 'minded' life as processes with characteristic natures.

Determining just what these characteristic natures are has been the crux of Alife and AI efforts. Robotics offers a framework in which to run thought experiments that test our intuitions about the nature of the mind-body connection. The philosophical significance of biologically-inspired robots is the challenge they pose to assumptions about the mind's being a special kind of entity that sets us apart from other forms of life. According to SCT, all organisms are self-preserving processes, and humans are distinct only in the sense that, for example, while an earthworm has only to dig, eat, excrete, and reproduce, the multifarious environment that is home to humans is demanding in much more complex and complicated ways that necessitate a flexible and adaptive mind. The reason dynamic systems theory (DST) has recently become more attractive to philosophers of mind and cognitive scientists in favor of naturalizing the mind is that DST recognizes that brain, body, and environment are equally important in explicating human cognition; and as a result, several philosophical puzzles that arise from attempts to distinguish the mind as something fundamentally different from the natural world are dissolved.

So far, research with biologically-inspired robotics has not borne out any contradictions in the intuition that mind is continuous with life, i.e., that being cognitive (in some sense) is nothing more than the attuned dynamic emerging from embodiment and embeddedness in the world. It is clear that at least some simple cognitive tasks that facilitate survival in one's environment do not require anything like what was postulated in the older AI paradigm and dualist-type philosophy of mind. But much work remains to be done, both in Alife and in philosophy, in confirming the intuition that higher-level, *human-like* cognitive functions result from just more of the same or, in other words, that 'minded' human life is nothing more than an enriched version of 'simple' life.

Liz Stillwaggon
Philosophy Department
University of South Carolina
Columbia, SC, USA
E-mail: stillwag@mailbox.sc.edu

References

1. P. Godfrey-Smith, *Spencer and Dewey on Life and Mind,* in M. A. Boden (Ed.), **The Philosophy of Artificial Life,** pp. 314-331, Oxford University Press, Oxford, 1996.
2. H. Maturana and F. Varela, **Autopoiesis and Cognition,** Reidel Press, The Netherlands, 1980.
3. M. Bedau, *The Nature of Life,* in M. A. Boden (Ed.), **The Philosophy of Artificial Life,** pp. 332-357, Oxford University Press, Oxford, 1996.

Modelling the emergence of compositionality in a Talking Heads simulation

Studies on language evolution have received an increasing interest from computer modellers over the past decade or so. This is primarily because computer models can be used effectively to investigate many aspects of language evolution and its origins in a way that non-computational studies cannot. One aspect that has attracted much attention is the emergence of compositionality in language. Compositionality refers to expressions of which distinct parts relate to distinct parts of its meaning, and the way these are combined. In contrast, holistic expressions have no parts that relate to a part of its whole meaning.

In computational studies on language evolution, two models appear most prominently. First, there is the *iterated learning model* (ILM) in which language evolves by iteratively transmitting the language 'vertically' from one generation to the next.[1] Second, there is the *language game model* (LG) in which a language develops from scratch 'horizontally' within one generation.[2] ILMs have been used to demonstrate how initially unstructured 'proto-languages' can change into compositional languages if the language is transmitted through a bottleneck (i.e. agents only learn from a small subset of the language). LGs have been used to show how shared languages can arise through self-organising properties in which a (potentially large) population tries to optimise communicative success. Of course, both models have their limitations: ILMs usually have the semantics in the language predefined, while these are acquired from scratch in most LG models. The LGs, however, typically do not have a generational turnover, but ILMs often have small populations of one agent per generation.

In order to deal with these problems, the strengths of both models have been combined in a simulation of the Talking Heads experiment.[2] In this model, implemented in THSim,[3] agents engage in language games where speakers try to produce expressions, which hearers try to interpret. While doing so, the agents develop their own meanings and grammar. Speakers can invent new words and hearers can acquire ones expressed by speakers. While a hearer acquires a new word-meaning pair, it tries to find co-variations in both meaning and signal with respect to previously learnt associations. This allows the hearer to generalise its language by forming compositional structures. For instance, if the hearer learns the association between meaning *01* and word *ac*, while it previously acquired the holistic association between *00* and *ab*, then it can induce that *a** means *0** (where * is a wildcard), **b* means **0* and **c* means **1*. In case such a regularity cannot be found, the hearer adopts the word holistically.[4]

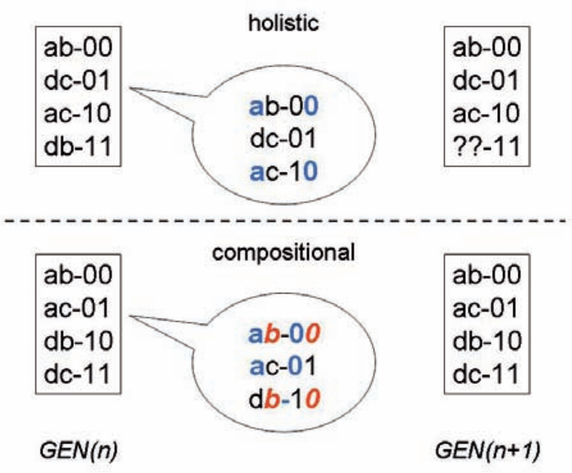


***Figure 1.** Holistic languages cannot be maintained when transmitted through a bottleneck, because the learner of GEN(n+1) cannot acquire the entire language of GEN(n) (top). If the language is compositional the learner can (bottom).*

Simulations with this model confirm the findings of ILMs that compositional structures emerge when the language is transmitted from one generation to the next through a bottleneck. The reason for this emergence is that when new adults start to speak, they may encounter meanings they have never seen before because of the bottleneck. If their language is compositional, they may nevertheless be able to produce an expression; whereas if their language is holistic, they need to invent new words (see Figure 1). However, simulations using different conditions than those typically used in ILMs reveal that compositionality can emerge in the absence of a bottleneck as well.

In ILMs, language is transmitted vertically, but in human society, children start speaking when they are very young. In simulations where speakers included children and hearers included adults (similar to the LG), it appeared that no bottleneck needs to be imposed by the experimenter. Instead, it has been shown that children, when starting to speak, may need to communicate about objects they never encountered before, so they face the consequences of the bottleneck earlier in life. Interestingly, where traditional ILMs have had difficulties in scaling population size, the combined model allows the emergence of compositionality to improve when the population size increases.

The current model has already provided many new insights as to how compositionality can emerge; for instance, the development of meanings have both a positive and a negative effect on the emergence of compositionality. However, there are still many open questions, some of which are addressed in the New Ties project funded by the EU FP6-IST Program in which a large cultural society is to evolve by combining evolutionary adaptation, individual learning and cultural evolution.[5]

Paul Vogt
Language Evolution and Computation Research Unit
School of Philosophy Psychology and Language Sciences
University of Edinburgh
E-mail: paulv@ling.ed.ac.uk
http://www.ling.ed.ac.uk/~paulv

References

1. S. Kirby, K. Smith and H. Brighton, *From UG to Universals: linguistic adaptation through iterated learning,* **Studies in Language 28** (3), pp. 587-607, 2004.
2. L. Steels, F. Kaplan, A. McIntyre and J. Van Looveren, *Crucial factors in the origins of word-meaning,* in Wray, A. (Ed.), **The Transition to Language,** Oxford University Press, Oxford, UK, 2002.
3. http://www.ling.ed.ac.uk/~paulv/thsim.html
4. P. Vogt, *The emergence of compositional structures in perceptually grounded language games,* **Artificial Intelligence 167** (1-2), pp. 206-242, 2005.
5. http://www.new-ties.org

BOOK REVIEW

Selfish Routing and the Price of Anarchy

Tim Roughgarden

Publisher: MIT Press **ISBN:** 0262182432
Hardcover: published June 2005, 240pp, £22.95

A common assumption of contemporary political and economic discourse is that individual self-optimising behaviour will lead to overall results that are optimal for everyone. For those of us familiar with game theory, examples such as the Prisoner's Dilemma show that this is not always the case: we can each make decisions which are rational, yet this can give rise to a global result which is worse for every individual.

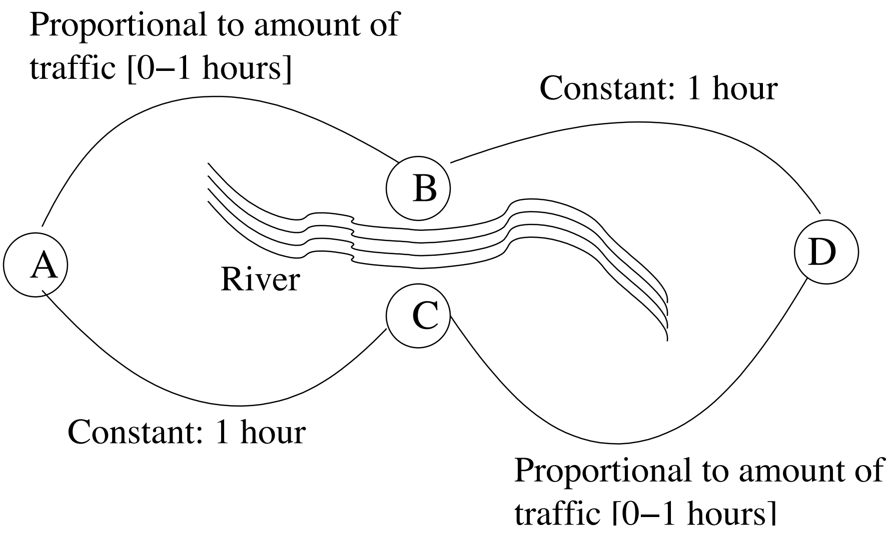
Nonetheless, such self-optimising processes seem to perform competently in many situations. This book aims to quantify this issue by asking: "how bad can the difference between globally optimal behaviour and the overall result of individuals acting 'selfishly' get". Most of the examples used in this book are drawn from road traffic routing.

The book begins with two classic scenarios: those of Pigou[1] and Braess.[2] Braess's example will illustrate the type of problem. In Figure 1 we can see a road layout: cars are travelling from A to D via two possible routes, annotated with the time cost of travelling along that road. Some roads are wide, and time cost is effectively constant; with others, the cost is proportional to the amount of traffic. There is no rational basis upon which to choose between the two routes, so traffic is spread evenly between them: total time cost one-and-a-half hours. B and C are assumed to be near each other and separated by a river. To improve traffic, a bridge (of negligible time cost) is built over the river. As a result the individual optimum route is now A,B,C,D: at no point will taking one of the constant-time roads benefit a particular individual. However as a result both of the proportional-time roads operate at maximum capacity, taking one hour to traverse, leading to a total travel time of two hours. Nonetheless is it not rational for any individual to switch.

The author also presents some illuminating examples drawn from other fields of study. One of the most interesting examples discussed in the book concerns weights and springs. Consider the network of weights and springs in Figure 2(a). If the string were to be cut in the middle; one would probably imagine that removing a constraint would mean that the weights would fall. However surprisingly the weights rise: the removal of the constraint means that the two springs are able to act in parallel, each lifting an individual weight. The role of the string is similar to the role of the bridge in the above scenario.

The main contribution of this book is to quantify these differences. Instead of only pointing to examples where selfish choice-making leads to inferior overall behaviour for all individuals, a ratio called the 'price of anarchy' is defined and analysed in the context of traffic-flow problems such as that discussed above. This price is the ratio of the cost if every individual acts selfishly (i.e. the Nash equilibrium) to the cost that can be obtained by a global optimisation process. For some examples, this turns out to be quite acceptable: for example with linear cost-functions, the price of anarchy is bounded by 4/3; so Braess's example above is an example of the worst-case behaviour.

Figure 1. *Braess's example of the difference between globally optimal behaviour and the overall result of individuals acting 'selfishly'.*



Once this core analysis has been carried out, a number of extensions and further examples are given. The book expands on its original points by considering cases where the decision-makers do not have perfect information, and cases where there is an absolute upper bound on the amount of traffic that can travel down a particular route. In this latter example, the differences in overall system performance between the optimal routing and the selfish routing scenarios can be huge. The book also explores other interesting topics including attempting to predict when networks will suffer from these problems and redesigning them so that the selfish behaviour gives rise to optimal or near-optimal behaviour. It also examines whether centralised control of a small proportion of the overall traffic, whilst letting the remainder behave selfishly, can give rise to globally near-optimal behaviour.

The book uses tools from a number of different disciplines including game theory, network flow analysis, and computational complexity theory. This attempt to combine tools from mathematical economics and theoretical computer science is largely successful: for example there are analyses not just of whether a particular strategy is possible, but whether it is possible to find such a strategy on a tractable timescale. This does however mean that the book requires the reader to posses an eclectic technical background, or a willingness to do a little interdisciplinary reading to obtain a full understanding of all of its detailed arguments. Nonetheless, this combination of the theoretical aspects of computing

Roughgarden
Continued on p. 11

Selfish Routing and the Price of Anarchy

Continued from p. 10

with economic ideas seems promising, and there is much potential for applying these ideas in other areas.

Overall this book works at a number of levels. The basic arguments and ideas will be interesting and accessible to many. In particular the idea of quantifying the difference between the outcomes of selfish and globally optimal behaviours is an interesting idea which could see application in other areas. The more detailed mathematical sections require a lot of work to understand, due to the unusual mixture of techniques used. However, for someone with an interest in the intersection of game theory, social theory, modern economics, and self-optimising processes, there is a lot to be learned from the concepts developed in this book.

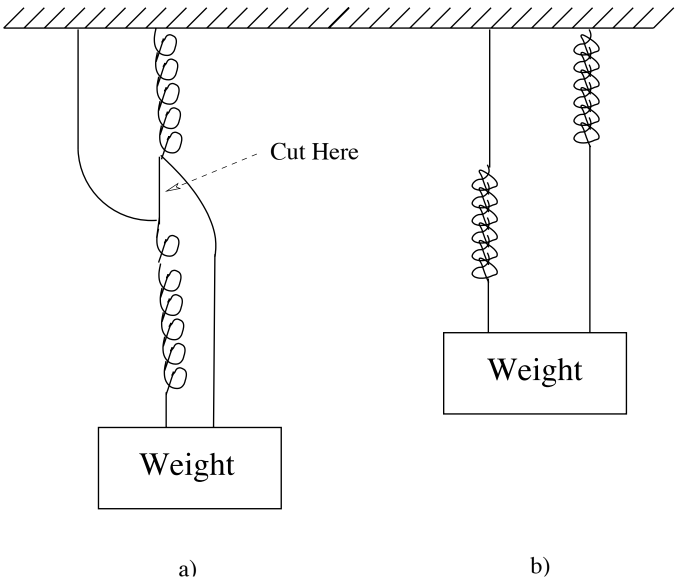


Figure 2. *A second example from Roughgarden using weights and springs.*

Colin G. Johnson

Colin Johnson is senior lecturer in Computer Science at the University of Kent at Canterbury. His main research interests are in the application of computational and mathematical techniques in biology and medicine, and in the use of ideas inspired by the natural sciences to devise novel computing techniques.

References

1. A. C. Pigou, **The Economics of Welfare,** Macmillan, 1920.
2. D. Braess, *Uber ein Paradoxon aus der Verkehrsplanung,* **Unternehmenforschung 12,** pp. 258-268, 1968.

Biologically-inspired image processing for a robotic grasping task

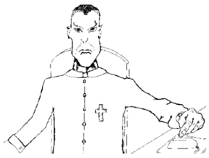
Continued from p. 10

ing into two parts and the use of population codes kept the grasping errors low.[1,2] This robot experiment demonstrated that brain functions can provide guidelines for robotic control, but also robots can help us to understand the brain. This is done by first demonstrating that certain (often hypothetical) functions actually work and then showing the advantages of certain data-processing techniques in a behavioral context.

Heiko Hoffmann
Max Planck Institute for Human Cognitive and Brain Sciences
Munich, Germany
E-mail: mpi@heikohoffmann.de
http://www.heikohoffmann.de/publications.html

References

1. H. Hoffmann, W. Schenck and R. Möller, *Learning visuomotor transformations for gaze-control and grasping,* **Biological Cybernetics 93,** pp. 119-130, 2005.
2. H. Hoffmann, **Unsupervised Learning of Visuomotor Associations,** Vol. 11 of MPI Series in Biological Cybernetics, Logos Verlag, Berlin, 2005.
3. D. Jancke, W. Erlhagen, H. R. Dinse, A. C. Akhavan, M. Giese, A. Steinhage and G. Schöner, *Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex,* **J. of Neuroscience 19** (20), pp. 9016-9028, 1999.
4. T. W. Kjaer, T. J. Gawne, J. A. Hertz, and B. J. Richmond, *Insensitivity of V1 complex cell responses to small shifts in the retinal image of complex patterns,* **J. of Neurophysiology 78,** pp. 3187-3197, 1997.



The Life of A. Hacker

by Fr. Aloysius Hacker

Cognitive Divinity
Programme
Institute of Applied
Epistemology

Episode 2: Formative Years

Disowned by my absent parents, in 1950, as a twelve year old functional orphan, I entered the Academy for Belief in and the Upholding of Spiritual Education (ABUSE™) based in Manchester. By their example, the ABUSE™ Brothers taught me the principles of authority, respect and honesty that were to guide me throughout my life: that great respect bestows the power of authority that makes honesty irrelevant. Moreover, nothing bestows respect more than religious office, a role that I planned to adopt myself.

I lost no time in joining the local amateur engineers' society: the Manchester Association for Geniuses, Inventors and Constructors (MAGIC™). As the youngest member of the group, I was immediately nicknamed 'Baby'. But I was also the most popular member. I still had my father's hoard of electronic components and, in this period of post-war austerity, these were in short supply. No group were more anxious to befriend me than the University team who had just completed the construction of the Manchester Mark I computer, which burnt out valves at an alarming rate. In recognition of my contribution to their pioneering work, they renamed their prototype Small-Scale Experimental Machine, the World's first stored-program computer, in my honour: the Manchester Baby.

It was through my University friends that I first met the great Alan Chewring. He and I enjoyed many late night sessions. Our favourite topic of conversation was what I christened 'Artificial Intelligence', although Alan was reluctant to adopt this term. What he did avidly adopt was the test for machine intelligence that came to be called The Chewring Test and which Alan later publicised in his famous *Mind* article. I initially resented the lack of acknowledgement of my idea, but I later realised that I had learnt another valuable life-lesson about the value of trust and loyalty: there is no value in trust and loyalty.

During the next two years, Alan and I invented the automated game-playing technique of mini-max. I subsequently used mini-max to implement a chess-playing program to run on the Mark I, to which I now had regular access. I called my program CHESS™ (for 'CHESS™ is a Heuristic Expert-System Solver') thus, in passing, anticipating, by several decades, both expert systems and recursive acronyms.

Given my stellar research productivity during the early 50s, it is no surprise that I attracted the attention of some senior AI researchers world-wide. By 1955 I was engaged in discussions with John McCarthief about the first AI conference, which he planned to hold at Dartmouth in 1956. John adopted my suggestion of 'Artificial Intelligence' as the name for this new field, although I again noted that senior researchers did not always give due credit to the ideas of their juniors. Maybe this was the key to becoming a successful researcher. I resolved to emulate such role models.

By 1956 I had established BOOTLEG™ (Blackboxes, Other Outfits and Thingamajigs, Likewise Electronic Gizmos), a thriving business supplying apparatus and components to the hobbyist and inventor community. It was time to strike out on my own. I resolved to say goodbye and good riddance to the ABUSE™ Brothers. What better way to mark my independence than to spend the profits of BOOTLEG™ to fund my attendance at the first ever AI conference. It was time to start my adventures in the USA.