Proceedings of the 18th Conference on Behavior Representation in Modeling and Simulation, Sundance, UT, 31 March - 2 April 2009
[paper: 09-BRIMS-005, pp. 9-14]

Changing our Brains: Systemic Causality in Complex Human Systems

David F. Batten
Commonwealth Scientific and Industrial Research Organisation
107-121 Station Street
Aspendale, Vic. 3195
AUSTRALIA
+61 3 9239 4420
david.batten@csiro.au

Keywords:
brain dynamics, causality, complex systems, frames, memes, neural networks, worldviews

ABSTRACT: When we change our mind about something, does it change our brains? According to George Lakoff, the way we view the world is conditioned by frames, metaphors and narratives that, over time, become hard-wired in our brains. It is not easy to overcome established patterns of thought (or our instincts) when trying to make sense of new information. Is the brain organized to enhance or suppress specific types of thought or behaviour? Can we provide a biological explanation for differences between individuals and groups in the behaviours they enhance or suppress? Daniel Levine claims that insights from social neuroscience, brain imaging, and neural network theory suggest that certain cortical-subcortical interactions serve to enhance or suppress classes of behaviour. This paper reviews some of the literature on the brain behaviour suppression issue. It concludes by suggesting that a coupling of the rich dynamics of neural networks with Dawkins' evolutionary notion of selfish memes could provide a framework for understanding the roots of complex human behaviour – especially instinct and personality.

1. Introduction

Complex patterns of human behaviour are difficult to capture in multi-agent simulation models of social and economic behaviour. Knowing each agent's strategy at one point in time may not help when trying to predict the collective behaviour of certain systems – e.g. if it is in each agent's best interest to do the opposite of most other agents. In self-defeating situations like these, the collective population may exhibit a panorama of simple or complex behaviour. The self-defeating nature of particular social situations and resource management problems has been studied previously and will not be discussed further here (see, e.g. Arthur, 1994; Batten, 2000; 2007).

On the other hand, knowing how an agent has behaved in the past is crucial in other human collectives, like in politics or in our defence forces. According to George Lakoff, Professor of Cognitive Science and Linguistics at the University of California at Berkeley, how we view the world is conditioned by frames, metaphors and narratives that, over time, become hard-wired in our brains. Although we can overcome established patterns of thought when trying to make sense of new events or information, it is not easy to do this. The path of least resistance for our brains – the neurological path with the most chemical receptors at the synapses – is to fit the emerging reality into a pre-existing mental pathway or "frame" (Lakoff, 2004).

If thought is physical, the concepts we think with must be physically instantiated in the synapses and neural circuitry of our brains. Once established, neural circuits do not change quickly (Lakoff, 2004; Levine, 2005). Natural selection, the process that designed our brain, takes a long time to design a circuit of any complexity. The time it takes to build circuits that are suited to a given environment is so slow it is hard to even imagine – being more akin to a stone being sculpted by wind-blown sand. Even relatively simple changes may take thousands of years.

When a word or phrase is repeated many times over a prolonged period, the neural circuits that compute its meaning are activated repeatedly in the brain. As the neurons in those circuits fire, the synapses connecting the neurons in the circuits get stronger. In this way, the circuits may eventually become permanent. Learning a word physically changes your brain, and that word's meaning becomes physically instantiated in your brain. This suggests that premeditated agents of change are unlikely to be able to change other people's minds unless they can find a way of altering the hard-wiring of their brains.

2. Frames and Mental Models

2.1 Frames

"Frames" are mental models of limited scope, with a systematic internal organisation. For example, our simple frame for "war" includes semantic roles: *the nations at war, their leaders, their armies with soldiers and commanders, weapons, attacks, and battlefields*. The war frame includes specific, local knowledge. In Australia, for instance, the Australian Defence Force operates under the direction of the Minister of Defence; war's purpose is to protect the country; the war is over and won when the other army surrenders. All these words are defined with respect to frames.

According to Lakoff, frames shape our thoughts and ideas. Deeper frames structure our long-standing morals and values – known as our *worldview*. Deep framing is the conceptual infrastructure of the mind. Surface frames are more modest, being associated with particular words, phrases and modes of communication. Language can be used to reframe a situation. For example, the reframing of the Iraq War as a "front in the war on terror" was a surface reframing.

2.2 Metaphorical thought

Most human thought uses conceptual metaphors, but we are mostly unaware of the metaphors we think with and live by (Kövecses, 2002; Lakoff, 2006). Metaphorical thought is used constantly, and we act on our metaphors. Take the phrase "nerves of steel," for example. People who are calm and collected in a crisis are admired greatly. Many of us try to emulate them, keeping calm if possible. Thinking in frames and metaphors like this is normal and gives rise to inferences that do not fit the laws of logic or deductive rationality as mathematicians and economists have formulated them. Since metaphors and frames can vary from person to person, not all forms of reasoning are universal.

The frames and metaphors in your brain define your "common sense". However, some of your commonsense ideas may not fit well with the observed reality in the world. Suppose a fact is inconsistent with the frames and metaphors in your brain that define common sense. Then the frame or metaphor will usually stay, and the fact will be ignored. For facts to make sense they must fit existing frames and metaphors in the brain. Furthermore, if you are unaware of your own deep frames and metaphors, then you are unaware of the basis for your moral and political decisions. Problems may arise if you base your own decisions on an inherently imperfect understanding of the situations in which you participate (Kövecses, 2002). Since facts matter, proper framing – both deep and surface – is needed to communicate the truth about our economic, social and political realities.

3. Two Competing Worldviews

Webster's Dictionary defines a worldview as "a comprehensive, personal philosophy or conception of the world and of human life." It is generally agreed that an individual's worldview is mostly founded upon four underlying concepts: economics, politics, religion and science. Because deeper frames and metaphors structure our long-standing morals and values, they shape an individual's perception of these four concepts and our personal understanding of how society is best served by this perception. Thus, worldwide, there could be almost as many worldviews competing for a place in our brains as there are human beings competing for a place on the planet.

In economics and social science, however, fewer worldviews are discussed. Two competing worldviews are portrayed (under several guises): *Cartesians* and *Stochasts*; *Sheep* and *Explorers*; *Conservatives* and *Liberals*. These worldviews correspond to two extremes in terms of risk-taking behaviour or creativity. The *Cartesians* (or *Sheep* or *Conservatives*) are mostly risk-averse, preferring to follow well-established principles and tactics rather than test untried ones. Their views and strategies are influenced by traditional practices and a more restricted set of fundamental beliefs. In the context of fishing strategies, for example, Allen and McGlade (1986) found that *Stochasts* search randomly for better fishing sites using their own intuition. They are risk-takers seeking higher returns commensurate with the higher risks they take. Conversely, *Cartesians* lie at the other extreme, being conservative skippers who are unwilling to take any risks and who go only to the fishing zones promising the best known returns.

Does this behavioural dichotomy extend beyond the bounds of fisheries and other natural resources? In any mobile population, for example, we find some risk-takers and some who are risk-averse. The number of vehicles turning up on a specific road each day is unpredictable. Of interest are the adaptive strategies of drivers who find themselves exposed regularly to traffic jams. Downs (1962) identified two behavioural classes of driver: those with a low propensity to change their mode or route strategy, called *Sheep*, and those with a propensity to change, called *Explorers*. *Explorers* search for better options to save time. They are quick learners and can hold several heuristics in mind simultaneously. *Sheep* are more conservative and prone to following the same option. Empirical work in North America has confirmed the presence of *Sheep* and *Explorer* behaviour in real traffic (Conquest, 1993).

The parallels between *Cartesian* or *Stochast* fishing strategies and *Sheep* or *Explorer* driving strategies appear striking. Yet they are less surprising when seen as general traits of human behaviour. In the world of technology, risk-averters and risk-takers appear under different guises: *Imitators* and *Innovators*. If we acknowledge the coexistence of imitative and innovative traits within a population, both must be treated as co-evolutionary variables dependent on the unfolding of events (Batten, 2000; 2007).

In American politics, Lakoff has identified two parallel worldviews: the *Conservatives* and the *Liberals* (also known as the *Progressives*). The "stricter father" family is the model for radical conservative behaviour, whereas the "nurturing parent" is the model for progressive behaviour. The stricter father model assumes that the world is a dangerous place, and it always will be, because there is evil out there. The world is always difficult because it is competitive. There will always be winners and losers. There is an absolute right and an absolute wrong. Children are born bad, in the sense that they just want to do what feels good, not what is right. Therefore, they have to be made good. What is needed in this kind of world is a strong, strict father who acts as a moral authority to protect and support the family and teach his children right from wrong in this dangerous world (Lakoff, 2004, p.7).

By way of contrast, the nurturing parent model assumes that the world, despite its dangers and difficulties, is basically good, can be made better, and that both parents share the responsibility to do that. Accordingly, children are born good and parents can make them better. The parents' job is to nurture the children and raise their children to be nurturers. Nurturing has two aspects: empathy and responsibility. These two aspects may be recognised as progressive political values, and from them more caring and innovative policies follow (Lakoff, 2004, p.40). Many people retain active versions of both the *Conservative* and the *Liberal* models in different parts of their brain, and use them in different parts of their lives. Fearful situations can trigger the strict father model, activating that model. For example, there are strict fathers in the classroom who still have progressive politics. Union employees tend to be strict towards their employers but nurture their union members. Military leaders act as strict fathers with young cadets but nurture older colleagues and friends.

4. Censors in the Brain

Simon Levine, a psychologist at the University of Texas at Austin, agrees that an individual agent (or collective) can select what types of behaviour it wishes to promote or what types of behaviour it wishes to forbid. The parenting model we experience as children has a strong influence on what we classify as good or bad – which Levine calls *angels* and *demons* after the painting by Escher (Figure 1) in which the spaces between white angels look like black devils. The brain's neural censors of angels and devils are its decisions about what classes of behaviour to enhance or suppress. The tendency to create such classifications and to act, at least partially on them, seems to be universal and hard-wired.

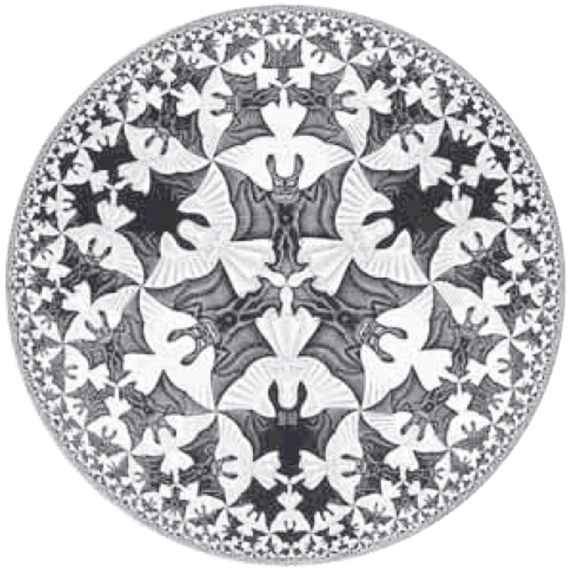

Figure 1: "Circle Limit" by M.C. Escher
(Source: Levine, 2005)

Many religious scholars agree that differences between faiths (Christianity, Islam, Judaism, Hinduism, etc.) are less profound than differences between the *Liberal* and *Conservative* wings of each faith (Beyer, 1994). As explained above, *Liberals* and *Conservatives* differ greatly in their devils. Ultimately, both the *Conservative* and *Liberal* impulses are rooted in our evolutionary programming. The *Conservative* impulse is closely tied to the characteristic *fight-or-flight* responses that humans, like other animals, have developed to cope with danger or stress (see Eisler and Levine, 2002). Families arise from the part of evolution that makes humans and other animals bond with others of their species and tend their offspring (Taylor, 2002). This part of evolution generates the compassion and cooperation that drives the *Liberal* impulse (Loye, 2002).

Eisler and Levine (2002) describe possible brain mechanisms for implementing behaviors that tend to fit these two competing models. Specifically, they discuss neural pathways for *fight-or-flight* responses and for what Taylor (2002) terms *tend-and-befriend* responses. Levine suggests that a dynamical systems perspective, centered on neural networks, can be useful for ordering our thinking about how internal psychic and neurodynamic processes lead us to behave in one way rather than another. He proposes that our relatively stable personality types (or character types) may correspond to distinct attractors in some dynamical system that arises by virtue of interactions between various parts of the brain. Adopting this view, the mental process of "changing one's mind" may correspond to the physical process switching between attractors in our brain. Attractor neural networks represent an attempt to develop more realistic and dynamic models of the brain's neural circuitry (Amit, 1989). Aiming to put more flesh on this abstract picture, Levine discusses a number of specific brain regions, their apparent psychological functions and the interactions between them, and on this basis proposes a speculative and provocative schematic diagram for the brain's overall decision-making apparatus (Levine, 2005).

Additionally, one may conceive of the brain as a synergetic system, which implies that it operates close to instabilities and achieves its preferred states by self-organization (Haken, 2006). Its activity is then governed by collective variables known as *order parameters* that enslave the individual elements, i.e. the neurons. Haken uses a detailed model of a neural network that, he believes, captures the essential features of real neurons. Among his conclusions are: (1) that all logical processes in the brain are fundamentally nonlinear; (2) that the brain is neither an attractor network nor can its function be described by chaos in the sense of chaos theory; (3) that the brain state moves from one quasi-attractor to another quasi-attractor, with attention being the essential variable that keeps quasi-attractors open and closes off others.

5. Direct and Systemic Causality

Lakoff goes even further by suggesting that radical *Conservatives* rely on *direct causality*, whereas *Liberals* rely on *systemic causality* (Lakoff, 2006, chapter 7). Although such distinctions are rarely hard and fast, there is evidence of *Conservatives* arguing on the basis of direct causation (i.e. by a single individual) and *Liberals* arguing on the basis of systemic causation when discussing various complex systems. Systemic causation is a causal relation involving at least one complex system. As well as the human brain, other complex systems are weather systems, a stock market, a power grid, an ecosystem, an epidemic, a health care system, a culture and an electorate. The behaviour of these systems cannot be pinned down to single events or even to single classes of events.

Beginning with Aristotle, the concept of causality has enjoyed a controversial history in philosophy and science and remains controversial to this day. This paper does not intend to delve into the formal or epistemological nature of causality for its own sake, but to take a pragmatic approach to causality as the production and propagation of neural effects in the brain, with the aim of supporting the understanding of real world human behaviour. For a discussion of circular causality in the brain, see Freeman (1999).

4.1 Direct causality

For several centuries, the dominant frame in science has been reductionism or direct causality. One agent exerts a single force on something and it moves or changes as a result. You flip a light switch and the light comes on. By definition, direct causality excludes any intermediate causes or the actions of multiple agents. Being the simplest kind of causality, it may be relatively easy for the brain's neural circuitry to accept and reinforce over time.

Models based on simple causality and linear causal chains are attractive because they seem tractable, unlike the confusing complexity of real systems such as a brain or the weather. Much of our cultural conditioning is predicated on a naïve view of linear causal chains – examples include: finding 'the single cause' of an effect, or 'the person' to be held responsible for something, or 'the perfect cure' for an illness. This focus on singular or direct causes makes it difficult to intervene effectively in complex systems and engineer desirable outcomes, without attendant undesirable ones (so-called 'side-effects' or unintended consequences).

4.2 Systemic causality

What makes a system (like our brain) complex is the network of interdependencies between the elements of the system. This means that the consequences of any event or property unfold through many interacting pathways. Also, if we want to know how a particular thought or property came about, we find that many cross-linked pathways contributed to it. In such a system as the brain, therefore, we cannot expect direct causality (one cause → one effect), or linear causal chains (each effect inexorably causing the next, like a Rube Goldberg machine), to hold in general.

It is possible to make a complex system *appear* simple by restricting the scope of our attention to a particular subsystem or pathway. If the scope is widened to include other pathways, or if unexpected side-effects that have propagated through those pathways are linked back and suddenly manifested within the restricted scope, then we are quickly reminded that the causal chain was just one of many pathways through the network. Such systems are not characterised by linear causal chains, but by networks of causal relationships through which consequences propagate and interact. For example, the firing of a particular neuron in our brain is influenced by the input signals received from many other neurons, and its own firing in turn influences the probability of firing for many neurons further downstream. Looking at the entire network of interconnected neurons we see firing patterns which are associated with complex higher-level functions. Such networks of interactions between contributing factors can exhibit emergent behaviours which are not readily attributable or comprehensible. This simple observation has some significant consequences for our ability to understand and influence complex systems.

Because systemic causation is abstract and poorly understood, it may be difficult for another unprepared human brain to accept. Thus arguments that rely upon it (e.g. those made by *Liberals*) are less likely to change the minds or brains of others (e.g. *Conservatives*), unless they are introduced and reinforced in particular ways.

6. Competing Memes

Competition between the *Liberal* and *Conservative* worldviews is reminiscent of competing *memes*, the notion proposed by evolutionary biologist, Richard Dawkins. Dawkins coined the term "memes" to describe how human ideas, objects and artifacts can be treated as agents evolving separately from their human hosts, owing their existence as entities to contingent facts about human brains and their interactions (Dawkins, 1989). He assumed that the dynamics of memes are governed by the principles of Universal Darwinism (Dennett, 1995). Our contention in a recent paper (Batten and Bradbury, 2009) is that a science of co-evolving ideas, habits, actions and artefacts – in fact, all elements of human culture – could be built on Dawkins' notion of the meme using the analytical tools of complex systems (see Anderson, 1988; Epstein and Axtell, 1996; Holland, 1998). As well as multi-agent simulation, this toolkit includes several other nonlinear dynamic modeling tools, such as Bayesian belief networks and neural network models.

The crucial idea about memes – what makes them uniquely Darwinian – is that if a meme can get itself copied it will. Copying and imitation come naturally to human beings. In fact, what makes humans different from other animals is our ability to imitate (Blackmore 1999). For example, an idea that gets passed on can be passed on again, and again, and thus take on a life of its own. And that is what makes the meme a replicator and gives it replicative power. It has been said that a chicken is just an egg's way of making another egg. Perhaps a scholar is just a library's way of making another library (Dennett, 1995). A meme's continued existence depends on physical embodiment in one medium or another. Like genes, memes are potentially immortal. But, also like genes, they depend for their existence on a continuous chain of physical vehicles. Books, buildings and music are relatively permanent, as are inscriptions on monuments. But unless all of these are under the protection of human conservators, they tend to dissolve over time.

If memes, as well as genes, build humans and their culture to further their own interests, then even bigger questions loom. Where is free will in this system description? Modern genetics has undermined the belief that humans are born with the freedom to shape their individual destinies. Scientists now recognize that genes shape our minds as well as our bodies. If memes have a hand in shaping our minds as well, then who is really in charge – ourselves or our memes? Can the human brain possibly survive as the ruling vehicle in the face of such a complex mix of memetic influences operating at vastly different speeds?

There are no simple answers to these questions. You may be appalled by the idea of your brain being "a sort of dungheap in which the larvae of other people's ideas renew themselves, before sending out copies of themselves in an informational diaspora" (Dennett 1995). It does seem to rob your mind of its importance as both author and critic. Most of us would like to think of ourselves as godlike creators of ideas, manipulating and controlling them and the minds of others as our whim dictates. But, as we saw earlier with our competing worldviews, even with the most creative *Liberal* minds, this is seldom, if ever, the reality. As Mozart observed of his own "brainchildren":

"*When I feel well and in a good humour, or when I am taking a drive or walking after a good meal, or in the night when I cannot sleep, thoughts crowd into my mind as easily as you would wish. Whence and how do they come? I do not know and I have nothing to do with it. Those which please me I keep in my head and hum them; at least others have told me that I do so*" (Dennett, 1995).

Memes can exploit our limited capacity to deal collectively with complex problems, undermining our efforts to grapple with the complexity of the situation. Simple memes may propagate better for the mechanical reason that something simple can be copied with greater fidelity than something complicated. Whether good or bad, simple memes propagate more effectively than complicated ones. They are copied faithfully by politicians and populations alike. One attractively packaged meme can become omnipotent, dominating any local region of political meme-space. Our own complexity-based work suggests that there is room for only one powerful meme in any local region of meme-space — and such memes will usually be simple. Moreover, simple memes are often the constructs of *Conservatives*, just like simple or direct causality. Thus the *Conservative* message spreads more easily and efficiently than the *Liberal* one.

7. Concluding Remarks

The competing frames of direct and systemic causality can provide insights into controversies involving science and religion. Bob Rosen, who wrote several books about neural networks in the 1930s and started writing about complexity in the 1950s, showed that reductionist thinking resulted in the need for an external, intelligent designer. By way of contrast, complexity theory shows that organisms are different from machines and have no need of an external causal agent (Rosen, 1997). Truly complex systems involve systemic causation and cannot be explained using direct causation unless a deity or other supernatural agent is invoked. Complexity science is posing a threat to the dominant worldview of direct causality, just as progressive thinking is often regarded suspiciously in politics.

To conclude, this paper has revealed some parallel threads in the ideas of Lakoff, Levine, Haken and others studying the brain's structure and likely behaviour under certain conditions and the nature of systemic causality in human systems. It recognizes that some major challenges must be overcome before more tools from the complex systems science toolkit are applied to the problem of systemic causation and emergent behaviour in the human brain. Neural networks and Bayesian belief models would appear to hold promise, especially when coupled with an understanding of the brain's psychophysiology and Dawkins' notion of selfish memes. In the future, application of such nonlinear tools might show that direct causation is of no use in two key areas of science: evolutionary theory and climatology.

Quotation

"It is worthy of remark that a belief constantly inculcated during the early years of life, while the brain is impressible, appears to acquire almost the nature of an instinct – and the very essence of an instinct is that it is followed independently of reason."

Charles Darwin: in *Descent of Man*.

References

Allen, P.M., McGlade, J.M. (1986): Dynamics of discovery and exploitation: the case of Scotian shelf fisheries. *Canadian Journal of Fisheries and Aquatic Sciences 43*, 1187-1200.

Amit, D.J. (1989): *Modeling Brain Function: The world of attractor neural networks*. Cambridge: Cambridge University Press.

Anderson P., Arrow, K., Pines, D. (1988): *The economy as an evolving complex system*. New York: Addison Wesley.

Arthur, W.B. (1994): Inductive behaviour and bounded rationality. *The American Economic Review 84*, 406-411.

Batten, D.F. (2000): *Discovering Artificial Economics: How agents learn and economies evolve*. New York: Westview Press.

Batten, D.F. (2007): Are some human ecosystems self-defeating? *Environmental Modelling and Software 22*, 649-655.

Batten, D.F., Bradbury, R. (2009): Simple memes and complex cultural dynamics. In K. Haynes, C. Karlsson and R. Stough (Eds.), *Innovation, Dynamic Regions, and Regional Dynamics* (in press) Berlin: Springer.

Beyer, P. (1994): *Religion and Globalization*. New York: Sage.

Blackmore, S. (1999): *The Meme Machine*. Oxford: Oxford University Press.

Conquest, J. (1993): Effect of motorist information on commuter behaviour. *Transport Research B 18*, 101-110.

Dawkins, R. (1989): *The Selfish Gene*. Oxford: Oxford University Press.

Dennett, D.C. (1995): *Darwin's dangerous idea: evolution and the meanings of life*. New York: Simon & Schuster.

Downs, A. (1961): The law of peak-hour expressway congestion. *Traffic Quarterly 16*, 393-409.

Eisler, R., Levine, D. S. (2002): Nature, nurture and caring: we are not prisoners of our genes. *Brain and Mind 3*, 9-52.

Epstein J., Axtell, R. (1996), *Growing artificial societies: Social science from the bottom up*. Washington: Brookings Institution Press.

Freeman, W.J. (1999): Consciousness, intentionality, and causality. *J. of Consciousness Studies 6*, 143-172.

Haken, H. (2006): Synergetics of brain function. *Int. J. of Psychophysiology 60*, 110-124.

Holland J.H. (1998): *Emergence: From chaos to order*. Oxford: Oxford University Press.

Lakoff, G. (2004): *Don't think of an elephant*. White River Junction, VT: Chelsea Green.

Lakoff, G. (2006): *Whose freedom?* New York: Farrar, Strauss and Giroux.

Levine, D.S. (2005): Angels, devils, and censors in the brain. *Complexus 2*, 35-59.

Loye, D. (2002): The moral brain. *Brain and Mind 3*, 133-150.

Kövecses, Z. (2002): *Metaphor: A Practical Introduction*. New York: Oxford University Press.

Rosen, R. (1997): *Life Itself*. New York: Columbia University Press.

Taylor, S.E. (2002): *The Tending Instinct*. New York: Holt.

Author Biography

DAVID BATTEN is an Adjunct Research Fellow at the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Melbourne and Coordinator of Theme 5 (Cellular automata, agent-based modelling and simulation) of the Australian Research Council's Complex Open Systems Research Network (COSNet): (http://www.complexsystems.net.au/).