［論　文］

Revisiting Architectural Requirements for Visual Processes

視覚的プロセスのアーキテクチャラル環境要件の再評価

LYNCH, Gavin *

Abstract

Investigation into the function of vision in animals, including humans, is summarised from the standpoint of Artificial Intelligence and philosophy and combined with theoretical psychology, showing that biological evolution is quietly busy solving various issues that arise in nature. Although experts in the field require specially controlled environments to investigate underlying mechanisms, most of the examples given in this paper require only careful notice to everyday visual architectures. Computer models and neural theories in current use cannot shed enough light on the mechanisms behind such visual architectures, so progress is needed in explaining the requirements for new systems waiting to be built. However, defining such requirements can prove to be a difficult task. Competences that are familiar to us are first investigated and, as we begin to understand the need to have more precise research, such as how visual competences are held by particular species, how the environment affects performance and how individuals in a group differ, we begin to get an insight into which requirements are more important to include in visual process models. The end of this paper will debate the need for a new type of experimental machinery. The work of behavioural neuroscience researchers is reviewed in this paper.

Keywords：Visual Processes, Philosophical Theory, Architectural Proposals

1 Introduction: Ideas of mental architectures

1.1 Vision viewed from a different angle

This paper focuses on areas to be explained, and not on providing a model (a solution) for use in the study of vision. Contemporary experts reject Kant's proposal in 1781 that knowledge of mathematics should neither be viewed empirically nor synthetically. However, Sloman (1962) tried to defend Kant's views in the light of multi-disciplinary knowledge of fields including education, neuroscience, linguistics, psychology, etc.

Attempts are made to show how visual ability in some animals, humans and androids is governed by the ability to reason about their environment, including not only the presence but the absence of processes. Therefore, we need to pay attention to what is not happening, a step ignored by most researchers. Visual researchers such as Kaneff (1970) often dealt with image structure, Marr (1982) thought of vision as providing physical facts, Berthoz (2000) and others think of vision as a behaviour control tool and Gibson and Pick (2000) view vision as environmental potential ability input.

1.2 Complete systems are important

Vision is dependent on interactions between the sensory systems, action-control systems and central control areas. Each part has been changed by evolution to become what they are today, and continue to

---

＊ リンチ・ギャビン
北陸学院大学短期大学部　コミュニティ文化学科
OS・ネットワーク概論

change as we develop as an individual. Therefore, the visual architecture can differ due to differences in species, and also due to differing individual development within a species. One example of individual difference is that produced by culture – the interpretation of gestures. Another may be a variation of the reading of written symbols.

Understanding of the complete system is important. After all, one may sense a facial expression such as happiness but, if the underlying, central mechanisms are not equipped to make use of the incoming information, the sensory system will have little meaning.

Instead of asking what happens under certain conditions, psychologists would be better off asking “how” things happen under a holistic design-based approach. We could then use that knowledge gained to produce lists of requirements. These lists would be used to investigate which functions are best incorporated into artificial neural systems, whether for further experimentation and analysis or for commercial projects/products.

We also need to understand what happens in the absence of certain aspects of a design. We already have input for this from studies of cases of brain damage and sensory reduction, among others. Understanding of this may help us to treat future problems of a similar sort, underlining the importance of building and studying incomplete systems, yet focusing on the whole.

Sloman (2000) tells us that we need to understand niche space (the space of possible sets of requirements) and *design space* (the space of possible designs). Furthermore, we need to understand the interactions and relationships between the two spaces within one ecosystem, and how changes in one space (niche or design) may affect other niche or design spaces.

Analysis of sets of functions should give us a list of possibilities for a system (visual or mental). A logical topography of states could be subdivided, in a commonsensical way, into logical geographies, as laid out in Ryle, 1949. This lets us see that classification from a common-sense point of view has a place within a scientific explanation.

1.3 General languages of representation

Current researchers often are committed to particular modelling forms, as they already have the tools for creating such designs. However, a more generalised semantic has been recognised in biological systems. Researchers’ commitments can limit them in the scenarios they can set up and results they can produce.

Very young children and many animals can react to processes using an internal and general language (i.e., not a learned language from their society). This helps them to reason about changes they see happening and, possibly, about what they notice not to be happening. Various scales of abstraction can be represented without having to use external, learned structures.

2 Interactive systems (wholes) and scaled up systems (parts)

2.1 Joining parts to create a “scaled out” system

A complete system, whether it is an animal or a robot, needs components which can interact with other, different components. This is called scaling out and is in contrast to the often focused on scaled up system which is one that can handle larger and deeper inputs. Human systems are often scaled out components, and it is such interaction that distinguishes us from machines. Most systems artificially created by researchers work well on their own in experiments, but do not work well with each other. They can scale up, but cannot efficiently scale out. This tells us that some systems, while they work as well or even better than their biological counterparts, are not useful as part of a whole system. For example, a computer system designed to teach a foreign language may not be effective when asked to learn a new language itself, while its biological peer may activate another component to complete the task.

2.2 Mechanisms versus whole designs

Brooks (1991) recommends us to build whole systems instead of the mechanisms being currently built. Neural networks are one mechanism often used in AI but we need to move beyond the parts and find out what the whole can do in order to study what requirements can be satisfied by new designs. For example, we may be able to find out what parts of the brain are activated while completing or understanding a task using brain imaging, but maybe who should instead focus on what the whole brain is exactly doing and what it is not doing, along with what messages are being passed and not passed.

We may be focused on creating null hypotheses while we should instead, as put forward by Lakatos (1980), separate research into purely empirical science based on laws, and the science of possibilities.

2.3 Vision and differences from mathematical reasoning.

Returning to Kant, he tells us that new truths can be discovered that expand our knowledge, which are not analytical and not empirical. We can visualise what happens when a change is made, even if we do not actually make such a change. In a mathematical formula, for instance, we can see that a different result is produced without actually making the change. Such reasoning can be visual (in the case of mentally drawing a line perpendicular to another line in order to find its midpoint) or can be a result of perception of temporal and spatial relationships (such as mapping or relating a series of objects). However, both types use related competences which are expected to work together, i.e., to scale out.

2.4 Growth of competences.

Young children cannot see all relationships, nor think about their existence. For example, a young child playing with a shape puzzle may be able to pick the pieces up and be able to place them in approximately the correct area, but may not be able to make them fit. The boundaries of the pieces and the holes may not match exactly, but the young child will try to press the piece in anyway. The young child needs to develop their awareness of objects and their boundaries using at least the following three aspects: (1) representation of objects; (2) acceptance of new ideas; and (3) manipulation of representations. Suavy and Suavy (1974) tell us that these skills are then built upon to understand what structures and processes can help us and which can hinder us in our tasks.

3 Realisation of processes and possibilities

3.1 Understanding changes and lack of changes

The ability to reason visually needs (1) the ability to see parts and their relationship in a complex system, (2) the skill to be able to decide what changes can be made, including constraints and possibilities, and (3) the awareness to be able to see what the consequences of such changes would be.

The ability to see both constraints and possibilities concurrently are important in building visual models as the ability to perceive what might happen if a constraint were to be removed is important in truly understanding the surrounding environment. For example, seeing that a car would crash if a competent driver (a constraint to a crash) were not present is important in creating a robotic system that might otherwise decide that the driver is redundant and remove him from, or fail to (re)place him in a system.

Warneken and Tomasello in 2006 showed with video footage how some primates and also human infants can react to such constraints on behalf of others using their generalised language abilities.

3.2 Evolution of independent sensorimotor signals

Visual skills help us when grasping objects using independent tools, such as hands or teeth. When the 3-D world in taken into consideration, the number of possible patterns is huge, yet we are able to process such options quickly due to our holistic biological design. Creating a design for artificial 3-D grasping needs not only sensorimotor signals, or not even such signals, but instead a language and design using spatial processes and structures capable of producing results in 3-D. The design will need to have reusable, generic mappings between processes, structures and motor signal patterns.

4 General conceptions of vision

4.1 Generative representations and limitations

Animals are adept at recognising what can impede their progress, or the progress of others. This allows the process of creation to begin, as evidenced by Weir, Chapell and Kacelnik in 2002 on their observations of a crow which used a tool (a bent hook) to get around a problem of not being able to reach some food.

AI vision researchers can omit to focus on what animals perceive in their world, the realisation that something does not exist but that it could exist. They focus on the acquirement of information from the environment and therefore miss the chance to have their machine create or manipulate the environment the way biological systems can do. They end up creating an artificial system that does not representation a biological system. In fairness to them, however, very young infants do not display this ability to manipulate their environment (Gibson and Pick, 2000) so it can be claimed that AI does represent the innate state of humans in that case.

4.2 Combining processes in complex systems

It must be remembered that processes occur in the environment of time and space, and that their parts are affected in different ways by such an environment. They can be combined in the following ways (Sloman 2009):

- a sequence of processes forming a complex process;
- parallel processes;
- chronological overlap, with the second and third starting before the first (think of a soccer match where the striker has already moved into position anticipating another player receiving the ball and passing it to him, with the striker being marked by another player);
- spatial overlap (think of 3-D work, such as carpentry);
- mutual modification (one process can slow another, which speeds up a related process, in a feedback, or reverse feedback loop);
- one process kick-starts another, for example a computer bootstrapping.

In order to understand the above, humans (and, as a consequence, human intelligence mapping AI) need methods of perceiving such interactions.

4.3 Processes occurring in space: reasoning

We can learn empirically about interactions in spatial and temporal planes but humans and a number of animals may need to be able to predict consequences of combinations that do not exist but may exist. Such combinations may include walking on an uneven surface while carrying a fragile object. It is better to be able to predict what may happen and to take action to prevent negative circumstances than to adopt a wait-and-see approach. We need to learn how one process can cause a new process to begin or, at least, how one process can provide information about a potential different but related process.

Actions can have consequences, often caused by new situations being created by such actions. A pot may become hot by another process heating it, something that can be detected by visually analysing the environment with background knowledge. Figure 1, shows how a new situation is created as a consequence of shifting position relative to an open door.

Figure 1. The closer you get to the door, the more access you will have about the contents of the room. In contrast, the farther away you move, the less

information you will have. A move to the left or right will change the visual input you receive.

4.4 Reasoning about processes and impedances that produce creativity

Objects work together in holistic systems, sometimes causing problems instead of solutions. It is up to the central controller or the receiver of the input to figure out solutions to such problems.

When processes occur, limitations of each can combine to create large limitations of the whole. In the 3-D world, spatial issues exist as in Figure 2. A chair may be lifted from place to place and a door may offer access to different areas of the house. However, when they come together we may end up failing to be able to access the functions of each process due to certain spatial limitations of the door and the chair. We may find that the chair cannot fit through the door and this failure may cause the subject (a human) to discover that rotating the moveable object a certain number of degrees about its axes can result in the desired goal (that of being able to move the chair through the door) begin achieved. This can all be done visually too, yet we are not able to explain what brain patterns are active in performing such a task. Traditional AI models cannot replicate the sophistication required to do this using their current systems.

Figure 2. A person has the ability to figure out how to move a chair through a door by working out translations and spatial rotations in 3-D about the chair's axes, some processes done in sequence and others in parallel. Current AI designs are limited in their ability to handle such a holistic task.

5 Conclusions

Sloman (1989) proposes that we replace architectures of the "modular" sort with those of the "labyrinthine" type in order to realistically represent the various components in a visual system and the interconnectivity between subsystems of a visual nature and other subsystems such as auditory, action control, etc. This research review confirms that proposal.

Human vision needs to be related to previous evolutionary models, including the visual systems possessed by other animals. Newer (more highly evolved) systems should not replace older (more primitive) ones, but be built upon them to create a layered infrastructure with concurrent levels of abstraction. Each layer may use different semantics and mechanisms, yet all be interconnected in some way and all be registered partly with the optic array.

The environment needs to be taken into account as focusing solely on the internal side of the organism will not lead to a full understanding of the system, in essence, the environment forms part of the organism at that time. The environment has unique meaning for each organism, even if all organisms are situated in the same place. Each creates its own niche. Neisser (1976) points out that the study of minds and brains is meaningless without the study of the environments they were evolved to function in.

The modelling of human vision needs mathematical reasoning about geometric structures. Humans need to reason about what they perceive in order to plan, predict and explain what is going on which, in turn leads to the control of processes which are highly interdependent. It follows that spatial reasoning is needed if truly useful and accurate models of human vision are to be designed.

< References >

Berthoz, A. (2000). *The brain's sense of movement.* London: Harvard University Press.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139-159.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development.* New York: Oxford University Press.

Kaneff, S. (Ed.). (1970). *Picture language machines.* New

York: Academic Press.

Kant, I. (1781). *Critique of pure reason*. London: Macmillan. (Translated (1929) by Norman Kemp Smith)

Lakatos, I. (1980). The methodology of scientific research programmes. In J. Worrall & G. Currie (Eds.), *Philosophical papers, Vol I*. Cambridge University Press.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: Freeman.

Ryle, G. (1949). *The concept of mind*. London: Hutchinson.

Sauvy, J., & Sauvy, S. (1974). *The child's discovery of space: From hopscotch to mazes – an introduction to intuitive topology*. Harmondsworth, UK: Penguin Education. (Translated from the French by Pam Wells)

Sloman, A. (1989). On designing a visual system (towards a Gibsonian computational model of vision). *Journal of Experimental and Theoretical AI*, 1(4), 289-337.

Sloman, A. (2000). Interacting trajectories in design space and niche space: A philosopher speculates about evolution. In M. Schoenauer et al. (Eds.), *Parallel problem solving from nature – ppsn vi* (pp. 3-16). Berlin: Springer-Verlag.