Riegler, A. (2005) Decision-making and anticipations: Why Buridani's ass is an unlikely creature. In: Smit, I., Wallach, W. & Lasker, G. E. (eds.) Cognitive, emotive and ethical aspects of decision making in humans and in AI. Volume IV. Windsor, Canada: The International Institute for Advanced Studies in Systems Research and Cybernetics, pp. 35–40.

Decision making and anticipation: Why Buridani's ass is an unlikely creature

Alexander Riegler

Center Leo Apostel for Interdisciplinary Research, Vrije Universiteit Brussel
Krijgskundestraat 33, B-1160 Brussels, Belgium
ariegler@vub.ac.be

Abstract

Asinus Buridani starved because it could not decide whether to feed from the left or right haystack. Naturally, decisions need to be triggered. Where does this bias come from? Usually conscious free will is held responsible despite warnings such as Huxley's *Helpless Spectator Theory*, which degrades consciousness to mere surveillance unable to do anything, and Freud's claim that human consciousness is not the 'master in its own house'. Similarly, in Libet's empirical results free will appears to be at the mercy of the limbic unconsciousness. Prinz framed this remarkable result as "We don't do what we want, but we want what we do." The consequences are obvious. What is referred to as a 'decision maker' is actually constructed at a level that obviously eludes conscious access. I propose an algorithmic account for decision making in humans and artificial cognitive agents that not only take the empirical results into consideration but also link decision making to the concept of anticipation. My claim is that decisions are the result of internal canalizations that arise from the dynamical hierarchical interlocking of structural elements in the cognitive system. This inevitably forces a particular path of how to react to certain contexts. I present psychological, ethological, and evolutionary evidence that support my account.

Keywords: Consciousness, anticipation, canalization, constructivism, volition.

Introduction

If a hungry donkey is placed between two haystacks it will not hesitate to start feeding on either of them. But what if the two stacks are of equal size, shape, and color, and if the donkey is placed exactly between them, thus excluding any 'outer' causes that could trigger a decision? It may nevertheless choose the left haystack because its senses are not symmetrical, or because it has developed the habit of starting to feed on the left. Any such drives or motivations may lead to decisions. But what if such 'inner' causes could be eliminated as well thus creating an abstract creature, an *asinus Buridani*, standing between two abstract haystacks? The only way to arrive at a decision requires arbitrariness. But since abstract creatures and perfectly symmetrical situations do not exist we can conclude that there must always be inner or outer causes that trigger decisions from among various choices.

While this scholastic thought experiment anticipates the answer to the question of how decisions are taken it does not tell us *who* or *what* is responsible for them. It is taken for granted that the human consciousness takes decisions based on careful considerations of the situational context. In economics this idea has been condensed into 'rational choice theory', which regards rationality as the only criterion a decision maker has to strive for. From a psychological perspective this theory was soon criticized for its inadequacy because it does not take the limitations of human cognition into account. Similarly, in artificial intelligence researchers discovered that successful rational reasoning in microworlds cannot be extrapolated for use in far more complex environments. That is, the computational effort of computing all rationally possible consequences of a propositionally represented context becomes intractable as soon as the size of the system in question is no longer trivial (cf. the 'frame problem', Dennett 1984). The proclaimed solution is 'bounded rationality' (e.g., Simon 1982), which builds on 'fast and frugal heuristics' (Gigerenzer 2000) rather than thorough rationality.

However, supporters of both the rational school and the evolutionary adaptive heuristics camp adhere to the metaphor "of the brain as an information processing system" (Dolan 2002). I claim that this assumption is flawed and present arguments backed by evidence from biology, ethology, and psychology that press for a different account that rejects the information-processing paradigm and, consequently, changes the perspective on decision making.

Decision making

As widely documented in the scientific literature, cognition is rarely rational. If it were the cognitive organism could never make up its mind as there are an infinite number of rationally possible consequences that emanate from a given situational context. Any rational attempt to take a decision requires the arbitrary interruption of the reasoning process before the agent runs out of time. On these grounds, Gigerenzer (2000) argued that decisions in a complex environment must necessarily be adaptive, fast, and frugal if they are to ensure survival. In artificial intelligence, a similar paradigm has emerged that focuses on behavior-based robotics. It does not rely upon *a priori* mathematical analysis of a given situation but rather on a *hinc-et-nunc* strategy that takes *system-internal* drives into account rather than a sophisticated representation defined in terms of the programmer's semantic world. The focus of attention has moved from defining a cognitive system as information *processor* to describing it as information *producer*. Consider a fly crawling over a painting of Rembrandt. It does in no way *process* the visual information presented in the painting as from its perspective there is no painting whatsoever. Only the human observer may wonder which information *filters* the fly applies in order to *ignore* the rich informational input. The fundamental difference arises from the fact that human scientists and engineers (lacking the first-order perspective of observed systems) necessarily concentrate on the perceivable *output* of systems. Cognitive systems, however, take actions in order to control and change their perceptive and proprioceptive *input*, e.g., they avoid the perception of an obstacle or drink to quench their thirst. In this perspective, decision making is based upon *internal* states rather than external states of affair. For an observer, evolutionarily successful organisms happen to make the proper decisions that respond to challenges of the environment. But from the perspective of the organism decisions are only the consequence of their internal cognitive dynamics. Consider the behavior of an incubating goose that decides to use its bill to roll back the egg that has fallen out of its nest. Interestingly, it pays no heed to the fact that an ethologist may have taken away the egg in the meantime (Lorenz and Tinbergen 1939). However, the animal does not filter out environmental changes. Rather, its cognition checks environmental states only at certain, apparently evolutionarily important moments (which do not include the existence of ethologists).

Human cognition works in similar ways. As the psychological literature documents, human problem-solving is dominated by the *set-effect* (Duncker 1935) and *mechanizations of thoughts* (Luchins 1942) that make subjects repeatedly choose a once successful strategy irrespective of whether another, simpler strategy might be better suited for new problems. Human perception, too, is determined by internal cognitive dynamics that occasionally seek to verify certain anticipations about future input states, as shown in the sequential order of tactile object recognition (Sacks 1995).

All this gives rise to the assumption that cognition is implemented in the form of schemata that accept information as it becomes available at sensory surfaces and that are changed by that information. They direct movements and exploratory activities that make more information available, by which they are further modified (Neisser 1975). This psychological account was formalized as *constructivist artificial life model* (CALM, Riegler 1994) where dovetailed schemata possess *perception-anticipating* slots ('checkpoints') that query certain environmental states in their condition parts. While in traditional production systems conditions are only used to trigger an appropriate action sequence, in CALM they are also used to decide whether a sequence should continue or terminate. In other words, since conditions can be embedded within action sequences, they act as *decision points* for determining whether the current schema is still 'on the right track'. Furthermore, since the ontogenetic expansion of schemata is a sequential process dependencies among schemata will occur in the sense that more recently added schemata become dependent on the existence and configuration of older schemata to which the new ones refer. This results in structural and processual hierarchies, leading to canalizations in the execution of schemata (Riegler 2001). Therefore, decisions are not taken in response to an environmental challenge but are a consequence of internal cognitive dynamics: Asinus Buridani's decision is based on internal cognitive activities rather than on the presence of haystacks. It will not fall prey to starvation because it is highly unlikely that – even it were a human being – its cognitive schemata would first request contemplation about equal distance to and equal size of the haystacks before starting to eat them.

Decision makers

Obviously, the information-producing paradigm leaves us with a contradiction. On the one hand, the stress on the *internal* perspective corroborates the common-sense view that taking decisions is a matter of the conscious free will rather than a matter of determination from the outside. On the other hand, however, the implementation of cognition as dynamical pattern emerging from the increasingly canalized interaction among schemata creates the impression of a rule-governed and therefore constrained cognitive apparatus.

Within the information-processing paradigm, neurophysiology is ready to account for human cognition in the latter sense. Damasio's (1994) *somatic marker hypothesis* (SMH) describes the ventromedial prefrontal cortex as capable of sketching diverse scenarios that may occur as a consequence of possible actions. The unconsciously working limbic system, especially the amygdala, assigns emotional significance to environmental stimuli in order to decide which of the projected actions to take. (In the information-producing perspective, the SMH is equivalent to the decision tree where nodes are conditions querying environmental states and edges are action sequences performed between such checkpoints). Consequently, Damasio distinguishes between emotion (the unconscious mechanism indexing the occurrence of significant events) and feeling (the phenomenological experience of emotional decisions). Ultimately, this means to relieve consciousness of its active role and to consider it a mere observer.

Such a perspective is by no means new. Already in the late 1890s, T. H. Huxley's *Helpless Spectator Theory* claimed that consciousness is not able to modify the working mechanism of the body. In analogy to canalizations, consciousness is in the situation of a train that is bound to go where the tracks are leading to. Jaynes (1986) arrived at the conclusion that about 3000 years ago the ancient Mycenaeans, as described in Homer's *Iliad*, did not (yet) possess consciousness. Whenever a decision had to be made they seemed to hear auditory hallucinations – 'voices of gods' – that took decisions for them. Jaynes claimed that their brain was divided into two disassociated hemispheres with different working modes resulting in a *bicameral mind*: "In his everyday life [bicameral man] was a creature of habit, but when some problem arose that needed a new decision or a more complicated solution than habit could provide, that decision stress was sufficient to instigate an auditory hallucination." This suggests that decision making happened at an unconscious level.

The philosophies of Huxley and Jaynes received empirical backing from Libet's (1985; Nørretranders 1998) experiments focusing on the *readiness potential* (RP), a specific electroencephalographic pattern that accompanies volitional acts. The RP starts about 0.5–1.5s before the act. Libet found out that the conscious experience of making a decision occurs about 0.2s before the action takes place. This means the RP, i.e., the sign that the unconsciously working part of the brain has started to prepare the 'volitional' act, commences at least 300ms *before* consciousness decides to act. In other words, "we don't do what we want, but we want what we do" (Prinz 1996). Our feeling of being in conscious control over our decisions derives from the fact that the conscious experience of events is projected backwards in time.

Since the readiness potential appears even if the test subject decides to cancel an already chosen action, Libet still grants an active role to consciousness in the sense that it can veto decisions made by the unconscious. Other neurophysiologists such as Roth (2001), however, negate even this possibility. They argue that the limbic system responsible for unconscious evaluations of emotional events is superior to any consciously working instance in the brain. In their view, the ultimate decision maker is not the conscious mind (Riegler 2003).

Conclusion

In the dominating information-*processing* paradigm the prefrontal cortex generates scenarios that may occur as a consequence of possible actions based on current perceptual stimuli. In order to enable fast decisions the amygdala not only attributes weights to these alternatives but is also involved in early object perceptual processing needed for the formulation of the scenarios (Dolan 2002). In this paradigm, unconscious emotions play a supreme role to counterbalance the inability of rational mechanisms to cope with the information flood in complex environments. Following the insights of Libet, consciousness is demoted to a mere observer.

The information-*producing* paradigm emphasizes the primacy of the internal cognitive dynamics over influences from the outside. Decisions are checkpoints embedded in action sequences. Both conditions and actions form schemata that populate the cognitive apparatus. Their dovetailing is continuously changing in function of phylogenetic and ontogenetic learning leading to a hierarchical organization and, ultimately, to canalizations that force certain paths of schema execution. Canalizations make sure that the vast majority of possible decisions remain outside the reach of cognition. In this paradigm, emotions do not play a superior role because there is no need to cope with the computational costs of filtering and evaluating the flood of perceptual stimuli. Cognition, then, becomes (again) the faculty concerned with what can rationally be known (Glasersfeld 1995).

References

Damasio, A. R. (1994); Descartes' Error; Grosset/Putnam

Dennett, D. C. (1984); Cognitive Wheels: The Frame Problem of AI; Minds, Machines, and Evolution (ed. C. Hookway); Cambridge University Press (pp. 129–151)

Dolan, R. J. (2002); Emotion, Cognition, and behavior; Science Vol. 298 (pp. 1191–1194)

Duncker, K. (1935); Zur Psychologie des produktiven Denkens; Springer

Gigerenzer, G. (2000); Adaptive Thinking; Oxford University Press

Glasersfeld, E. von (1995); Radical Constructivism; Falmer Press

Jaynes, J. (1986); Consciousness and the Voices of the Mind. Canadian Psychology, Vol. 27, No. 2 (pp. 128–148)

Libet, B. (1985); Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. Behavioral and Brain Sciences, Vol. 8 (pp. 529–566)

Lorenz, K. Z. and Tinbergen, N. (1939) Taxis und Instinkthandlung in der Eirollbewegung der Graugans. Zeitschrift für Tierpsychologie Vol. 2 (pp. 1–29)

Luchins, A. S. (1942); Mechanization in Problem Solving; Psychological Monographs Vol. 54

Neisser, U. (1975); Cognition and Reality; Freeman

Nørretranders, T. (1998); The User Illusion: Cutting Consciousness Down to Size; Viking

Prinz, W. (1996); Freiheit oder Wissenschaft?; Freiheit des Entscheidens und Handelns. (eds. M. von Cranach and K. Foppa); Asanger (pp. 86–103)

Riegler, A. (1994); Constructivist Artificial Life; PhD Thesis; Vienna University of Technology

Riegler, A. (2001); The Cognitive Ratchet: The Ratchet Effect as a Fundamental Principle in Evolution and Cognition; Cybernetics and Systems, Vol. 32 (pp. 411–427)

Riegler, A. (2003); Whose Anticipations? Anticipatory Behavior in Adaptive Learning Systems (eds. M. Butz, O. Sigaud and P. Gerard); Springer-Verlag (pp. 11–22)

Roth, G. (2001); Die neurobiologischen Grundlagen von Geist und Bewusstsein. In: Neurowissenschaft und Philosophie. (eds. M. Pauen and G. Roth); Fink/UTB (pp. 155–209)

Sacks, O. (1995); An Anthropologist on Mars; Knopf

Simon, H. A. (1982); Models of Bounded Rationality; MIT Press