Draft: Please do not cite or quote without permission

Thinking slow about thinking fast

Stuart E. Dreyfus

Department of Industrial Engineering and Operations Research, Etcheverry Hall, University of California at Berkeley, Berkeley, CA 94720,USA

Corresponding author: Dreyfus, S.E. (dreyfus@ieor.berkeley.edu)

In his book, *Thinking, Fast and Slow*, Daniel Kahneman attributes experientially-learned real-world coping skills to an "associative machine" acting on declarative memories of facts and events. While this attribution is probably correct for the unfamiliar types of situations that are the subject of his famous experiments conducted with Amos Tversky, we argue that experientially-learned real-world coping skills are based, instead, on a procedural memory system, currently under intense behavioral neuroscientific investigation, that is surprisingly overlooked in Kahneman's book.

,

Kahneman's book, *Thinking, Fast and Slow*

Daniel Kahneman's recently published, best-selling book, *Thinking, Fast and Slow,* [1] adopts a dual process explanation of human cognition. For purposes of exposition he accepts the previously published designations System 1 and System 2 for these two processes.[2] According to him, the fast System 1 operates automatically and quickly, with little or no effort and no sense of voluntary control. The core of System 1, he adds, is an associative ability that operates on the declarative memory of facts and recallable personal experiences. He includes in System 1 the experience-based fast execution of coping skills. The slow System 2 is meditative. It allocates attention to effortful mental activities such as reasoning and complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration. The book, and therefore this article, ignores the many and diverse cognitive details of System 2.

The bulk of Kahneman's heuristics and biases (H and B) claims are based on the highly revealing, and justly celebrated, experiments that he conducted with Amos Tversky, together with the wealth of work that these experiments spawned. However, these experiments almost ubiquitously involve situations in which the subject has had little or no opportunity for experiential learning. Having thought slowly about his digression into the domain of experientially-learned skillful coping, we have concluded that this is at best incomplete. The coping skills that Kahneman explicitly mentions include firefighting decision making, the choice of good moves by a chess master playing rapidly, driving decisions on an open road, and reading and understanding nuances of social situations. These skills are indeed executed rapidly, automatically, and with little or no effort, but are they produced by the associative System 1?

Kahneman's account of skill

Kahneman seems to assume that declarative memory *can* indeed explain learned skillful coping. He undertakes on page 11 an explanation of fast chess play. He refers in an endnote to a 1992 paper by Herbert Simon [3] in which Simon describes a master's fast chess play: "A large amount of the chess master's expertise lies in his or her intuitive (recognition) capabilities, based, in turn, on *large amounts of stored and indexed knowledge* derived from training and experience." (our italics). In [4] Kahneman writes approvingly, referring to [5], "They estimated that chess masters acquired a repertoire of 50,000 to 100,000 immediately recognizable patterns, and that this repertoire enables them to identify a good move…" (Box 1) Clearly, he envisions an associative link between an input chess position and a remembered situation that provides a good move, and sees this as a template for explaining learned coping skills. The Simon-Kahneman speculation about the production of skill also requires an articulation in memory of cues concerning what is significant about a situation. Even given an articulation of this, their associative system would somehow also have to answer the question "similar with respect to what?" to access the most similar memory.

One cannot deny that some human cognitive decision-making behavior, particularly in unfamiliar types of situations, can be attributed to an *associative machine*, to use Kahneman's term, acting on declarative memories of facts and events. However, we shall show below that there is much neuroscientific experimental evidence concerning a completely different route to skilled coping, based on a sort of memory that is surprisingly omitted in Kahneman's new book.

Procedural Memory

By the end of the period from 1969-–1984 of the very productive collaboration of Kahneman and Tversky, much had been learned about the existence of what is now called procedural memory [7,8]. After considerable experience, this procedural memory system produces *know how* [9,10].

Neuroscience has established that procedural memory is produced by a system of brain areas centered on the basal ganglia, including cortico-striatal loops that are subject to influences from neuromodulators such as dopamine, and also connections with limbic areas such as the amygdala. Experimental evidence strongly suggests that subcortical striatal areas are at least the initial home bases of *know how*. The limbic system provides reward signals needed, as we shall explain later, for the experiential reinforcement learning of skill. There are experimental reasons to believe that certain connected prefrontal regions are involved in providing to the striatal areas appropriate saliencing, sometimes called top-down modulation [11.12], of the incoming sensory stimuli. This saliencing creates a foreground-background distinction that is necessary for an organism to act in accordance with what is variously called, depending upon the context and researcher, a set, a goal, a task, an affordance (in the vocabulary introduced by the psychologist J. J. Gibson), a sense of the situation, a perspective (our preferred all-inclusive name [13] that is therefore used in this article), and without doubt by other names. All action is taken under a perspective so appropriate perspective and how it changes as stimuli change must be learned prior to, or simultaneously with, learning skilled action. Skilled coping is a hierarchical process.

Let us now delve into how the procedural process actually works. It is *not* by using Simon's stored and indexed knowledge.

Machine learning of skill

The story begins in the 1980s with the introduction of a machine-learning procedure called temporal difference reinforcement learning (TDRL) [14]. Under the assumption that the experimental learning of a skill requires feedback of a reward signal measuring the quality of performance, this is a family of algorithms that, when implemented on a computer, allows the computer to learn procedural skills including skills requiring a *sequence* of coordinated reward-producing actions, even actions that tend to maximize *expected total reward in stochastic environments*. TDRL can be accomplished by several means. We are concerned with a version using artificial neural networks. Each real or simulated act provided by an actor network during one step in a sequential process is accompanied by an observed reward signal and by outputs of an auxiliary artificial neural network called a critic. The critic evaluates both the expected total outcome quality of the current situation and of the remaining sequential behavior after the current step. This information is combined to calculate what is called a TD error. This error, as the computer proceeds to experiment with ways of accomplishing a sequential task, is used to modify the artificial synapses in the network that determines its choice of act and also those that determine the output of the critic. The TD error acts as a surrogate reinforcement signal in the sense of Thorndike [15], favoring actions that lead to better than currently expected total outcome quality. Thus the learned actor has the capability of directly mapping inputs from salienced situations into appropriate actions. We advisedly use *mapping* rather than "*associating* outputs with inputs" to clearly distinguish the actor network's contribution from Kahneman's *associative* machine explanation. The procedure described above is called model-free actor-critic learning. The totality of experiences input into the computer system is encapsulated in the synapses of the actor and critic nets. *No memory of the computer's individual actions and outcomes during experimental learning need be stored in a declarative memory!*

Furthermore, this machine-learning approach can optimize decision sequences in stochastic domains without seeking to explicitly learn the *probabilities* of state transitions or the *rewards* associated with them. The existence of this sort of model-free learning, if implemented in a brain, seems to answer Kahneman's query on page 13: "Why is it so difficult for us to think statistically?" The answer is that such a system can experientially optimize its behavior in stochastic situations *without thinking statistically* in the way Kahneman asks of his experimental subjects.

Behavioral neuroscientists have since the 1990s hypothesized that something resembling TDRL algorithms that work for machine learning *are* actually executed in living animals by their procedural striatal brain system.[16-21] The direct-mapping actor behavior, when learned, is an example of what is termed *habit* in the neuroscientific literature (Box 2). Using both fMRI imaging and electrodes embedded in various areas of laboratory animal brains while the animals learn and then habitize procedural skills such as maze running, a great deal of supporting experimental evidence has been amassed. This past three decades, marked first by the discovery of the procedural memory system, and then by the inroads that behavioral neuroscience has begun to make in discerning its operating principles, is what Kahneman's book surprisingly overlooks.

System 0 implications

Having now established the importance of at least three, not two, systems, it is time for slow thinking about rearranging this cognitive furniture. For reasons of simplicity and symmetry, what we have thus far designated the striatal system will now be called *System 0*.

We describe System 0 as the very fast, very effortless, inexplicable (except to say, "it's due to my experience-produced synapses"), experientially-trained method of skillfully and automatically coping with the everyday world, when there is feedback on the quality of performance. This is the proper system for explaining Kahneman's fast coping skills and, in fact, almost all learned and then habitized coping skills.

System 1, the system of the associative machine, is fast, effortless, and inexplicable (except to say, "it's due to my experience-produced associations and who knows or cares how synapses cause *that*"), and it becomes an operative system when relatively fast decision making is needed in a situation such as those in H and B experiments where one has never had vast experience with feedback evaluating performance. But it is certainly not the explanation of most ongoing everyday experientially-learned coping.

The various biases that are possible when System 1 is appropriately employed have been thoroughly studied by Kahneman and his colleagues. We speculate, given our belief in the importance of perspective (Box 3), that some of these biases occur because the experimental subject, unlike in everyday life, is suddenly placed in a situation without having available a contextual perspective determining what information is salient. Hence perspective, and therefore the decision, can be manipulated by the experimenter.

The fully habitized System 0 procedural brain responds to situations *really* fast and effortlessly compared to the associative machine. It doesn't think, in the conventional use of the word, it simply *knows how*; it breaks the thought barrier. It is the brain of our everyday skillful coping while we are *not* experiencing ourselves as decision making. It can do many things simultaneously. The driver of a manual shift car can attentively carry on a conversation while navigating to work along a normal path, while accelerating or decelerating as required, while shifting gears when appropriate, and while making the complicated manual motions to do so. System 0 is far from infallible. In bringing salience to its situation, it can overlook a glass that it then knocks over at a dinner party. More importantly, having learned by pleasing its reward system, it can produce addictive or sociopathic behavior if its reward system has been hijacked. Except for an assist from the associative machine on an important but novel decision such as a marriage proposal, System 0 plays the predominant role in determining *who we are* [22].

The procedural System 0 is not evolutionarily designed for coping with changes in the everyday world that imply that acting in accordance with experiential learning is inappropriate. That requires System 1 or 2. System 0 *can,* however*,* observe when its critic's evaluation of performance is not correct and realize that new learning of perspective and/or action is needed. System 0's TDRL apparatus alone also is not designed for detecting when experiential learning is becoming appropriate, but that sufficient experience to trust the learning has not yet occurred and a concurrent process more resembling thinking should be employed. A related brain area might do this adjudication [23].

Conclusion

We opine, therefore, that the book, *Thinking, Fast and Slow,* brilliantly illuminates associative-machine-provided fast choices in situations with which the subject lacks experience. At the same time, the book fails to inform the reader about the source of really fast experience-based habitized coping skill and of the fact that TDRL has demonstrated that there is no need for thinking statistically when coping experientially with stochastic real-world situations.

Acknowledgements

The author thanks Peter Dayan for indispensable advice in the preparation of this paper. He also acknowledges the generous financial support provided by Statoil ASA, through Project Academy StatoilHydro.

References

1 Kahneman D (2011). *Thinking, Fast and Slow* (New York:Macmillan)
2 Stanovich KE and West RF (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences,* 23, 645–726.
3 Simon HA (1992) What is an "explanation" of behavior? *Psychological Science*, 3,150-161
4. Kahneman, D and Klein, G (2009) Conditions for intuitive expertise: A failure to disagree, Am. Psychol. 64 (6) 515-526
5 Chase WG and Simon SA. (1973). The mind's eye in chess. In W G. Chase (Ed.). *Visual information processing* (pp. 215-281). (San Diego,CA: Academic Press).
6 Wan X et al, (2011) The neural basis of intuitive best next-move generation in board game experts, *Science,* 331 (2011) 341–346.
7 Squire LR (2004) Memory systems of the brain: a brief history and current perspective. *Neurobiol Learn Mem* 82:171–177.
8 Squire LR (2009) Memory and Brain Systems: 1969-2009, *J. Neuro.,* 29(41),12711–12716
9 Cohen NJ and Squire LR (1980) Preserved learning and retention of pattern analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science* 210:207–210
10 Cook, SDN and Brown, JS. (1999). Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing. *Organ. Sci.* 10(4) 381–400.
11 Miller EK and Cohen JD (2001), An Integrative Theory of Prefrontal Cortex Function, *Annu. Rev. Neurosci.* 24, 167
12 Gazzaley A and Nobre A (2012) Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.*, 16 (2): 129-135,
13 Dreyfus HL and Dreyfus SE. (1988) *Mind Over Machine*: *The Power of Human Intuition and Expertise in the Era of the Computer,* paperback edition pg. 28 (New York:The Free Press)
14 Sutton, RS and Barto, AG (1998) *Reinforcement Learning* (Cambridge, MA: MIT Press)
15 Thorndike, EL (1911). *Animal Intelligence; Experimental Studies* (New York:Macmillan).
16 Barto, AG (1995). Adaptive critics and the basal ganglia. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 215-232). (Cambridge, MA: MIT Press).
17 Schultz W et al (1997). A neural substrate of prediction and reward. *Science*, **275**, 1593-1599

18 Suri, RE and Schultz, W (1998). Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Experimental Brain Research*, **121**, 350-354.
19 O'Doherty, JP et al (2003) Temporal difference models and reward-related learning in the human brain. *Neuron*, **38**, 329-337.
20 Haruno, MK et al. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: A functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci*, **24**, 1660-1665.
21 Graybiel AM (2008) Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 31:359 –387.
22 Quartz, SR and Sejnowski, TJ (2002). *Liars, lovers, and heroes: What the new brain science reveals about how we become who we are* (New York: HarperCollins)
23 Beierholm, UR et al (2011) Separate encoding of model-based and model-free valuations in the human brain, *NeuroImage*, 58 (3) 955-962

Box 1. On chess

It is indisputable that a master chess professional has experienced during his or her study a vast number of famous chess positions with masterful moves and positional evaluations. Many of these situations doubtless remain in memory. Chess, because its strict rules constraining admissible decisions and determining their results, allows look-ahead planning unavailable in most real-world situations such as firefighting or driving. In *slow* chess, at certain critical points, this ability will be exercised, but never in the all-move manner of world-class chess programs. This look ahead results in various future positions that require comparative evaluations. It is likely that a master's evaluative process sometimes makes reference to this declarative memory of positions and evaluations to check the valuation of his learned TDRL critic that is explained in the *Machine learning of skill* section.

Move generation in fast chess has been located in the brain's procedural memory system that we describe in our *Procedural memory* section. The authors of [6], however, seem unaware that this brain system executes actor-critic TDRL as discussed in the *Machine learning of skill* section.

Box 2. Habit

The term *habit,* as used in the literature of Instrumental-conditioning studies concerning how animals and humans choose actions appropriate to the affective structure of an environment, should not be confused with a simple stimulus-response association in which an identical stimulus is repeatedly presented and a single response is evoked. Habit, more generally in instrumental conditioning, is used to describe any experientially learned behavior that is only very slowly sensitive to gross manipulations of the experimental environment such as devaluation of reward.

Box 3. Outstanding questions

Habitized behavior exhibited under a fixed perspective is flexibly responsive to changing values of salient stimuli. (Think of a baseball infielder catching a pop-up on a windy day.) The learning and execution of this behavior is well explained by the actor-critic TDRL model. In a laboratory setting, such behavior is only very slowly sensitive to gross manipulations of the experimental environment such as devaluation of reward. In the real world such gross changes rarely occur, but sudden changes in perspective are frequent, resulting in sudden changes in behavior.

This leaves unanswered:

1. How and where does the real-world brain experientially create a representation of a skill-related perspective? (Think of a baseball outfielder identifying a fly ball as catchable by him.)
2. How does the brain experientially learn what stimuli should be treated as salient, given a perspective? (At least the direction and velocity of the batted ball, wind velocity and direction, and, if daytime, the location of the sun potentially matter to a baseball outfielder.)
3. What is the mechanism that allows the experiential learning of when real-world stimuli as seen from a learned perspective become such that a different previously learned perspective becomes appropriate? (Think of the baseball outfielder's brain, as he runs to catch the ball, learning when to suddenly abandon the attempt and shift into a retrieve-the-ball perspective with its set of salient stimuli.)