Trends in Cognitive Sciences

CellPress
REVIEWS

Opinion

Unraveling the Mysteries of Motivation

Randall C. O'Reilly[1,*]

Motivation plays a central role in human behavior and cognition but is not well captured by widely used artificial intelligence (AI) and computational modeling frameworks. This Opinion article addresses two central questions regarding the nature of motivation: what are the nature and dynamics of the internal goals that drive our motivational system and how can this system be sufficiently flexible to support our ability to rapidly adapt to novel situations, tasks, etc.? In reviewing existing systems and neuroscience research and theorizing on these questions, a wealth of insights to constrain the development of computational models of motivation can be found.

Highlights

Motivation can be operationalized in terms of goals, which are distributed and multifactorial, including biologically and affectively salient outcomes, sensory-motor plans, and integrated utility representations. Different frontostriatal loops are specialized for each of these factors, involving the OFC, dlPFC, and ACC, respectively.

We appear to have different value functions when considering potential goals to pursue, compared with when we are engaged in pursuing an active goal: the goal-selection phase is relatively conservative because the goal-engaged state is so dominated by pursuit of the active goal. This can explain procrastination and other phenomena.

The coordination and binding across these distributed goal representations during planning and decision making requires serial processing, which also has the benefit of enabling flexible reuse of processing systems across time, resulting in serial Turing-machine-like flexibility.

Motivation for Motivation

Motivation lies at the heart of all cognition and behavior. We only think what we are motivated to think and do what we are motivated to do. At some level, this is an empty tautology. However, if we could construct an artificial cognitive system that exhibits human-like motivation, it would likely revolutionize our understanding of ourselves, by helping to unravel the details of how our motivational system works and how we can make it work better (e.g., most of the major mental disorders can be characterized as disorders of motivation). However, the most widely used AI/computational modeling approaches in this area are based on the reinforcement learning (RL) framework, which is rooted in the classic behaviorist tradition that treats the organism as a relatively passive recipient of rewards and punishments [1,2]. These models typically have only one overriding goal (to accrue as much reward as possible) and although recent 'deep' versions of these models have demonstrated spectacular success in game playing [3], these models remain narrowly focused on maximizing one objective and do little to help us understand the flexibility, dynamics, and overall richness of human motivational life.

By contrast, we are manifestly under the spell of our own internal motivations and goals, with many dramatic examples of the sheer power of these systems. Anyone who has been a parent can appreciate the transition that emerges around the age of 2 years, when your child suddenly starts to strongly want things and really not want other things, to the point of throwing epic tantrums that clearly exceed any rational metabolic cost/benefit analysis. The current tragic epidemic of suicide is another stark indicator that our motivational systems are strong enough to override even our basic survival instincts. Likewise, the rational economic picture of human decision-making runs counter to the large numbers of voters who appear to disregard their own economic, health, and other considerations in favor of more strongly felt social affiliations and other powerful motivational forces. These have not been traditional targets of computational modeling, but perhaps they should become the main focus, to probe the deepest, darkest secrets of our motivational systems.

This Opinion article highlights central challenges and progress in developing computational models of human motivation, focusing on a biologically based approach involving interactions among different brain systems to understand how both the normal and disordered system

[1]Department of Psychology and Computer Science Center for Neuroscience, University of California, Davis, 1544 Newton Ct, Davis, CA 95816, USA

*Correspondence: oreilly@ucdavis.edu (R.C. O'Reilly).

© 2020 Elsevier Ltd. All rights reserved.

functions. Two central questions, the answers to which would significantly advance our computational models, are as follows.

- Can we specify with computationally useful levels of precision the nature and dynamics of the internal goals that drive our motivational system?
- How can the motivational system be sufficiently flexible and open ended to support the huge range of human motivated cognition and behavior, and our ability to rapidly adapt to novel situations, tasks, etc.?

If good answers to these questions can be found, perhaps we will then understand the underlying basis of human flexibility, adaptability, and overall robustness, which remains unmatched in any artificial system, even with all the recent advances in AI.

The Nature and Dynamics of Goal States

It is broadly accepted that we can usefully characterize our motivational systems in terms of internal goals, which serve as the core construct for understanding how motivated behavior is organized toward achieving specific desired outcomes [4–17]. However, defining with computational precision exactly what a goal is, and how it functions in the biological networks of the brain, remains a central challenge. An important point of departure is to at least assert that active goal states ultimately correspond to states of neural activity in the brain, which are architecturally centrally positioned to drive behavior in ways that other neural activity states are not. This central, influential status of goal-related activation states implies, under most mechanistic attempts to define properties of conscious awareness [18,19], that typically they should be something we are aware of. However, that does not mean that we are necessarily aware of the various forces driving the activation of active goal states in the first place [14] – this is a major but separable topic that is not pursued further here.

A major challenge here is that the term goal can refer to many different things. Perhaps this can instead be turned into an opportunity, by adopting a distributed, multifactorial model of goal states (Figure 1). Specifically, we can think of an overall goal state as comprising multiple interacting yet distinguishable components, with good support for at least the following three major components based on a substantial body of systems neuroscience research and theory [17,20–31].

- Internal representations of biologically and affectively salient outcomes [e.g., the unconditioned stimulus (US) of classical conditioning, including food, water, social connection, etc.]. There is considerable evidence that the ventral-medial prefrontal cortex (vmPFC), specifically including the orbitofrontal cortex (OFC), plays a central role in actively maintaining and tracking these US-like states (which are richer and more stimulus specific than pure 'value'), and OFC damage impairs the ability to adapt behavior to rapid changes in these US outcomes [16]. As in classical conditioning theory, a critical feature of these US-like representations is that they provide a biologically based *a priori* grounding to the entire motivational system: goals are not purely arbitrary and fanciful, but rather are ultimately driven by core biologically based needs. There is an extensive literature on drives going back to Hull [34] and Maslow's hierarchy of needs [35], which relate to this essential component of motivated behavior, and modern neuroscience research suggests that many different specialized subcortical areas, including the amygdala and hypothalamus for example, are the source of neural projections that converge into the OFC to anchor our highest-level cortical goal states [21].
- Sensory-motor plans that specify how behavior and cognition should be directed to achieve the desired outcomes. Extensive research indicates that networks in the dorsolateral prefrontal cortex (dlPFC) support such representations and strongly influence other brain areas in support of such plans [36–38]. For example, scoring a goal in a game (using a particular strategy) is an action-oriented, plan-level 'goal' to achieve the outcome goal of winning. The

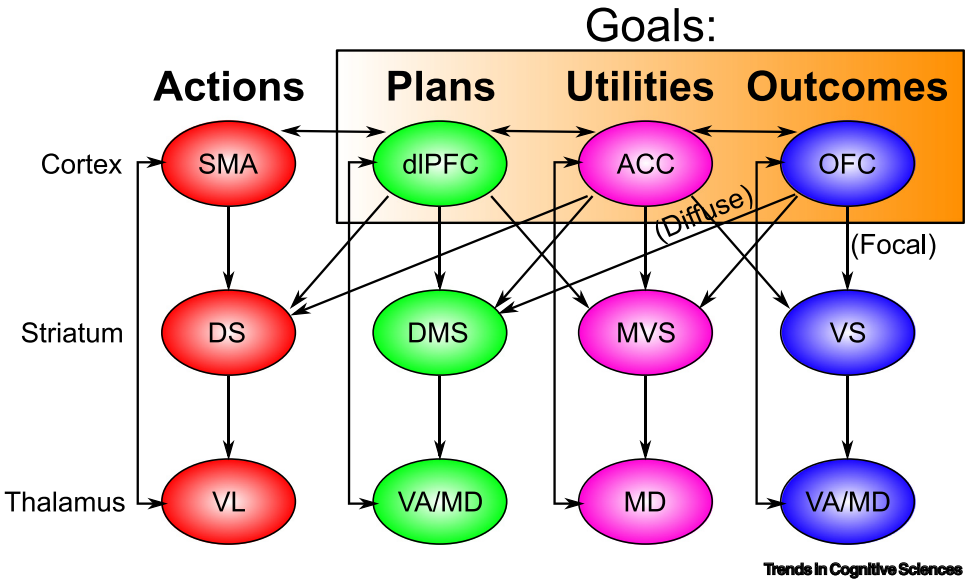

Figure 1. Three Core Elements of Distributed Goal States. Schematic for how four different anatomically defined loops through the frontal cortex and basal ganglia [20] correspond to a three-factor distributed, componential goal state, which collectively drives concrete action selection in the fourth, motor-oriented loop [supplementary motor area (SMA)]. The orbitofrontal cortex (OFC) maintains and tracks biologically salient outcomes, the dorsolateral prefrontal cortex (dlPFC) selects, maintains, and guides sensory-motor plans, and the anterior cingulate cortex (ACC) integrates these factors, along with potential effort and other costs, in terms of overall utility. The consistent frontostriatal loops across these areas support flexible, dynamically gated active maintenance and updating of these distributed goal states (e.g., [32,33]). Abbreviations: DMS, dorsomedial striatum; DS, dorsal striatum; MD, medial dorsal thalamus; MVS, medial ventral striatum; VA, ventral anterior thalamus; VL, ventral lateral thalamus; VS, ventral striatum.

dlPFC also interacts extensively with the parietal lobe, which represents the sensory aspects of the action plan (e.g., the visualization of the ball going into the net).

- Integrated utility combines information from the above components to determine the net balance between the value of the potential outcome (in the OFC) and the relative costs and effort associated with the specific plan being considered to obtain that outcome (in the dlPFC). There is extensive evidence that the anterior cingulate cortex (ACC) encodes this integrated utility information, including uncertainty and conflict signals associated with potential action plans [39,40].

All of these brain areas (and other related neighboring areas, such as the anterior insula and broader vmPFC) interact through extensive interconnectivity in the process of converging on an overall goal state. This goal state in turn drives subsequent behavior in accord with the engaged action plan, with extra attentiveness to progress toward, and opportunities consistent with, the expected outcome. The ACC state may play a critical role in energizing behavior in proportion to the net overall utility [41,42]. A critical advantage of incorporating these multiple factors together under the goal construct is that they all should mutually interact in the goal-selection process: the specific sensory-motor plans necessary to achieve a given affectively salient outcome must be appropriate to the specifics that particular outcome (indicating the need for more than purely abstract value representations), and the integrated utility obviously depends on both of these components, but also feeds back to potentially modify plans and outcome goals to maximize overall utility.

Computationally, this process of converging on an overall goal state can be described as a parallel constraint satisfaction process [44–46], where many individual factors, encoded directly

via rates of neural firing communicated over learned synaptic weights, mutually influence each other over a number of iterative stepwise updates to converge on an overall state that (at least locally) maximizes the overall goodness of fit between all of these factors. The properties of this system may explain how unconscious factors can come to influence overt behavior in the course of solving the reduction problem of choosing one plan among many alternatives [47,48]. We have implemented this process in two related models [43,46] as illustrated in Figure 2, both of which demonstrate the importance of this interactive, constraint-satisfaction process of goal selection.

This notion of tightly interacting yet functionally separable brain areas can help to make sense of empirical data showing that while individual neurons throughout all of these areas exhibit tuning for each of the aforementioned factors, more comprehensive population-code data in monkeys show clearer differences in overall coding across these areas, in ways that correspond nicely with the aforementioned functional account [30,49]. The separability of these areas also aligns

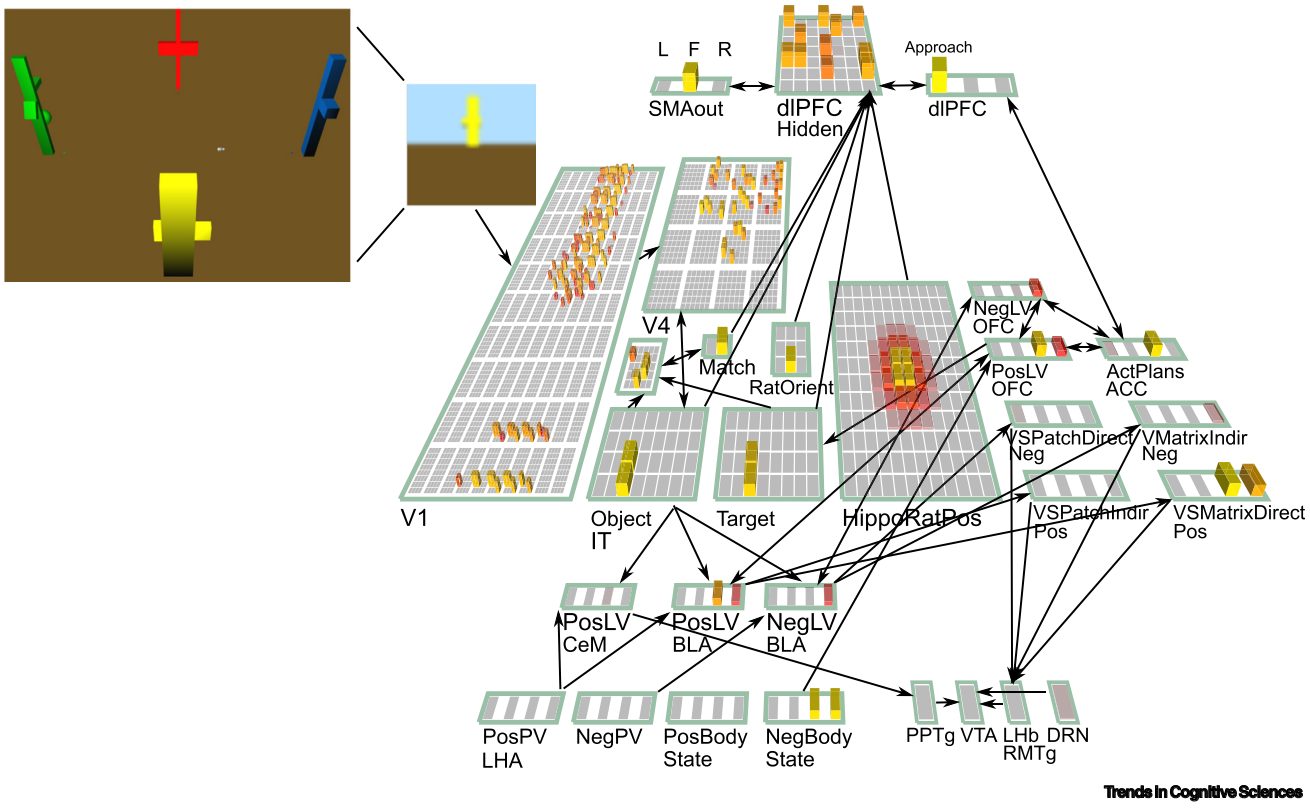

Figure 2. Computational Model of Constraint-Satisfaction Goal Selection. Computational model demonstrating how multiple separate goal-state components interact to produce coherent overall goal selection [43]. A simulated rat (named 'Emery') forages in a plus-maze-like environment with four different possible primary value (PV) locations (red, meat; blue, water; yellow, sugar; green, veggies). In addition to these positive PV outcomes, each location can also have negative outcomes (e.g., veggies can be bitter, meat can be rotten). For each run, Emery experiences two states of negative body state deprivation (e.g., thirst and need for vitamins from veggies), which drive goal selection. Emery processes a bitmap first-person camera view into the environment, to extract an invariant Object IT representation. The anterior cingulate cortex (ACC) ActPlans units encode the four different potential targets to approach, coordinating the orbitofrontal cortex (OFC) positive learned value (PosLV) representations that learn the specific target features associated with different PV outcomes, with the specific dorsolateral prefrontal cortex (dlPFC) motor plan required (Approach Target in this case). The negative learned value (NegLV) OFC representation encodes negative outcomes associated with different targets and helps to constrain the target-selection process. All of these prefrontal cortex (PFC) areas interact through bidirectional connections to settle on the target that best satisfies all of these factors (constraint satisfaction). At that point, those PFC states are gated in [via simulated ventral striatum (VS) gating] and maintained during the goal-engaged period. These top-down signals bias the dopamine-driving pathways in simulated amygdala (CeM, BLA) to produce phasic dopamine signals in response to increments of progress toward the target state.

with the separate loops of interconnectivity between these frontal areas and corresponding basal ganglia areas [20,50] (Figure 1) and with the overall idea that the OFC encodes stimulus-driven motivationally salient information as part of the overall ventral pathway while the ACC encodes action-based affective information as part of the dorsal pathway [51].

Finally, as noted earlier, these high-level, cortical systems interact extensively with the many different subcortical affective/motivational areas including the amygdala, hypothalamus, lateral habenula, and ventral striatum (nucleus accumbens), which in turn modulate the firing of deep midbrain neuromodulatory systems including the ventral tegmental area (VTA) dopaminergic system and the serotonergic system in the dorsal raphe nucleus. These neuromodulators in turn shape learning throughout the affective/motivational system, focusing learning on affectively significant, unexpected outcomes. Computational models of these systems at multiple levels of analysis show how these interacting systems can produce the signature phenomena of classical and instrumental conditioning, thereby explaining how perceived rewards and punishments can modulate behavior [52–55].

In summary, from the earlier discussion, the existing systems neuroscience literature can provide a specific and concrete basis for the development of computational models of how componential, distributed goal states could work together to guide behavior in motivationally appropriate ways. The detailed way in which any given specific goal state might come into activation may be complex, chaotic, and difficult to explain [14], but hopefully we can at least articulate the broader principles and organization of the neural systems involved, and the computational models can provide a critical bridging link between these neural systems and overall adaptive behavior.

Temporal Dynamics of Goal-Engaged versus Goal-Selection States

Another major set of questions for computational models concern the temporal dynamics of when these goal states become activated, how long they are maintained, and how the system knows when a goal has been completed and can therefore move on to another. Interestingly, our everyday subjective vocabulary characterizes these goal dynamics well, providing at least some potential insight into how the neural system functions. We experience satisfaction and pleasure when our goals are achieved, frustration and anger when they are impeded, and disappointment, sadness, and rage when we are forced to finally give them up. Boredom dominates when we cannot find anything interesting to engage in, and the aversiveness of this state suggests the overall importance of having actively engaged goals [56]. The broad scope and likely universal nature of these states associated with goal-driven processing suggests that our brains have biologically grounded, primary motivational/affective states associated with keeping ourselves in a productive, goal-engaged mode of functioning.

One of the most effective components of the standard treatment for major depressive disorder is the concerted reestablishment of basic goal-engaged behaviors, known as behavioral activation [57]. Self-actualization, or the achievement of major life goals, confers a broad sense of satisfaction, and soldiers and parents report high levels of life satisfaction even while experiencing higher levels of adversity and challenge on a moment-to-moment basis [58]. The importance of these goal-oriented states for our everyday lives again points to the centrality of motivation for understanding human cognition, and a key computational modeling challenge is to construct the necessary metacognitive monitoring and motivationally significant grounding to enable a model to capture the corresponding goal dynamics. We really need our models to experience something like frustration when we impede their ability to make progress on goal states that were activated through their own constraint-satisfaction process.

Further insight into these dynamics comes from considering the puzzle of procrastination: if goal achievement is so rewarding, why do we procrastinate so much? This and other interesting phenomenology can be explained by considering the different forces in play at two different points in the overall goal dynamics [8,12,13,43,59]: when you are currently engaged in pursuing an active goal (goal engaged) versus when you are selecting the next goal state to engage in (goal selection). These are labeled postdecisional or volitional versus predecisional or motivational according to the Rubicon model [13]. In the goal-engaged state, the brain develops a form of tunnel vision and can become obsessed with completing the goal, sometimes to the point of neglecting other important needs. Video games in particular have perfected the engagement of this state, by providing incremental positive rewards and indicators of progress toward the goal, which seem particularly important for sustaining the goal state. In recent years, many activities have become 'gamified' by adopting the same strategy, with exercise equipment, point cards, and especially social media apps tapping into the same drive to 'keep surfing that wave of progress toward your goal'. Giving up on our goals is aversive, so even if we no longer value the outcome that much, it is hard to stop. How many times have you finally managed to disengage from an addictive game or binge-watching a TV series, only to realize with disgust how much time has been wasted on a seemingly meaningless activity?

It is precisely this obsessive, locked-in nature of the goal-engaged state that requires the goal-selection process to be cautious and careful. In effect, your brain knows that whatever goal you end up selecting to engage in, you will run the risk of overcommitting, so it works extra hard to ensure that only the best goals are selected. However, what is best for your brain's reward pathways may not be what you rationally consider to be the best. Your brain evolved in times of scarcity and challenge and thus a quick, safe, highly rewarding outcome is generally preferred over a longer-term, more uncertain gamble [60] – this is the same dynamic for why we prefer sugary and fatty foods instead of the longer-term benefits of healthy vegetables etc. This is also why you tend to gravitate toward simpler, more satisfying tasks when faced with the unwelcome prospect of writing that overdue paper, paying those bills, doing your taxes, or replying to those more difficult emails that have been festering in your inbox. Your brain is just being 'rational' in choosing the most immediately rewarding goals for you. To overcome this biological bias, you must somehow align your sense of value with those things that you actually need to do, and that is where deadlines come in: they force the issue and finally get that difficult task over threshold.

Inevitably, once you start on those challenging tasks, you experience the ever-surprising feeling that it was not so bad after all (only to forget it again in time for the next such difficult task). This discrepancy clearly demonstrates that we have fundamentally different value functions operating during goal selection versus the goal-engaged state, and computational models will need to incorporate these state-dependent differences in value. Heckhausen, Gollwitzer, and colleagues have published several studies demonstrating the differences between these states, providing further potential targets for computational models [13].

Interestingly, the opposite of procrastination, precrastination, has also been found, where people prefer to get a more difficult task out of the way first [61]. This reversal of the usual pattern may be attributable to the desire to reduce the dread or weight of more difficult tasks, which is another countervailing motivational force that helps to mitigate against procrastination. The exact balance of these forces in any given situation may have to do with the perceived inevitability of the different tasks: if it is clearly not possible to put off the more difficult task, it makes more sense to get it out of the way first, as that will lighten your mental burden.

These kinds of pervasive mental phenomena that are strongly tied to the dynamics of goal activation and completion provide concrete targets for computational modeling and are further

evidence for the central importance of goal states in shaping our mental lives. Computational models are likely to require built-in motivational drive states associated with all of the aforementioned dynamics (satisfaction, disappointment, etc.) to keep the system productively engaged in a sequence of goal-selection and -completion cycles. The breakdown of various elements in this overall dynamic could lead to the emergence of a depressive state, which produces a self-reinforcing feedback loop of disappointment and sadness associated with the lack of goal engagement (and associated feelings of lack of overall self-efficacy and self-worth).

The prevalence of depression, and the diversity of etiologies that all converge on the same general dysfunctional behavior, suggests that the balance of this goal-driven machinery may be relatively delicate overall. Although it sounds a bit creepy in an uncanny-valley sense, there is no obvious theoretical reason why computational models could not exhibit all of these kinds of dynamics and thus provide significant insights into the complex emergent dynamics of these motivational systems. The existence of such models would be likely to undermine the widely held belief that it is precisely this capacity for complex emotional and motivational dynamics that makes us uniquely human.

Flexibility, Novelty, and the Importance of Sequential Processing

Let us now turn to the second major question posed at the outset: how can the motivational system be sufficiently flexible and adaptive to account for these signature properties of human behavior and cognition? The central hypothesis considered here is that the distributed, multifactorial nature of goal states, along with basic physical constraints, converges to require goal-related processing to be fundamentally sequential in nature, and this sequentiality in turn affords important benefits for flexibility and the ability to deal effectively with novel situations [46].

Physically, we can generally only do one thing at a time (walking and chewing gum being the exception rather than the rule), and this constraint appears to extend into the cognitive realm as well [62]; apparent multitasking is usually accomplished by rapid switching between tasks. Thus, goal selection represents the fulcrum by which otherwise parallel processing systems in the brain must produce a serial sequence of individual actions, carefully selected to optimize this extremely limited serial resource. The parallel constraint-satisfaction process that drives goal selection must therefore properly integrate the many different factors relevant to the current situation (internal body state, external cues and constraints, etc.) to come up with a plan that satisfies as many of these constraints as possible within a given plan of action.

The transition to a serial mode of processing, while slower overall, also affords many benefits, which can be summarized by 'the three Rs': reduce, reuse, and recycle. Serial processing reduces binding errors that would otherwise arise from considering multiple options at the same time, across multiple distributed brain areas. For example, if the ACC activates an 'effortful' representation in response to multiple different dlPFC plans being considered in parallel, how does the system know which plan is being so evaluated? This is a widely recognized problem for parallel distributed systems [63], which is avoided by only considering one potential plan at a time, so that the associated representations across the rest of the brain areas can be assumed to reflect the evaluation of that one plan. This is a key feature of our existing models of this parallel distributed constraint-satisfaction process [43,46] (Figure 2). However, even when the focus is on one particular goal, other goals (e.g., longer-term ones) can still inject relevant constraints into the process and thus influence and integrate across time scales, etc.

In addition, serial processing enables the same representations to be reused over time, greatly facilitating the ability to transfer knowledge from one situation to another and supporting the ability

to make reasonable decisions in novel situations. For example, a single common distance representation could be used in many different sequential goal-selection contexts to help compute expected time and effort costs, whereas a parallel system would require multiple such representations to avoid interference.

Finally, serial processing allows prior states to recycle or reverberate over time, contextualizing the evaluation of subsequent processing with some of the main conclusions derived from earlier processing steps. By contrast, purely parallel systems require dedicated, separate substrates for each type of processing being in done in parallel, and it is often challenging or even impossible to ensure that the proper information and constraints are communicated across each of these parallel channels. This is the fundamental reason why even relatively simple machines can be universal computational devices (i.e., Turing machines) but parallel computers must be carefully configured to achieve specialized computational functions (and many problems are simply not computable in parallel due to mutual interdependencies). Thus, to achieve this same kind of universal, flexible computational ability, the brain must likewise rely on fundamentally serial processing.

Multiple approaches have converged on the idea that the neural engine responsible for imposing sequential processing on an otherwise fundamentally parallel distributed neural computer is the basal ganglia [32,46], as featured in the ACT-R production system model [64–66]. The basal ganglia has opposing Go versus NoGo pathways [67,68] that compete to decide whether to engage a proposed goal state that is currently activated and being evaluated in parallel across the distributed cortical representations [46]. If the basal ganglia registers a NoGo (based on its history of dopamine-modulated learning [55,69]), the process is iterated again with a new plan that emerges from the ashes of the previous one, and so on until something gets the basal ganglia's Go approval, subject to relevant time and other constraints, etc.

Thus, the basic limitations against doing multiple things at the same time can give rise to a much greater level of flexibility enabled by serial processing. The frontal cortex and basal ganglia may have initially evolved to support basic motor action selection, but it is notable that these very structures are among the most enlarged in humans, and our unique symbolic cognitive abilities may represent the further development of these serial processing mechanisms to produce a powerful integration of both massively parallel and flexible serial processing [70].

Concluding Remarks

In summary, there are considerable constraints on and relevant data for the development of computational models of distributed goal representations that capture the differential value functions associated with the goal-selection versus goal-engaged state. Furthermore, if these models operate with significant serial dynamics, they can achieve greater flexibility, at the cost of slower overall functioning. Managing the complexities of these serial dynamics is notoriously difficult in complex recurrent networks, so this represents a major challenge as well as an opportunity.

Many important further questions arise directly from the ideas discussed here (see Outstanding Questions), including major issues regarding the relevant timescales over which goals operate and how the longer-term goals interact with the more immediate, active goal states that drive online behavior. Hierarchical models are appealing [17], but a more heterarchical framework may be more flexible, with longer-term goals providing various forms of context and constraint that guide the ongoing dynamics of goal selection and goal-engaged pursuit. In any case, understanding the basic properties of goal dynamics and learning at the simplest level would be likely to provide important insights into these bigger questions. There is also recent work in the machine

Outstanding Questions

What other components or factors are represented in distributed goal states and which brain areas are critical for such representations?

What kinds of monitoring or metacognitive signals are required to implement computational models of motivational states like satisfaction, frustration, boredom, etc.? These involve tracking progress toward the achievement of desired outcomes – how is this progress tracking encoded in the brain and how does it connect to central affective brain systems involved in dopamine regulation, etc.?

The most important specific version of the previous question is how we determine when a goal has been accomplished? Is it just the first sight or first taste of a desired outcome or do you have to consume the whole thing? Likewise, how do secondary reinforcers like money, which can be very abstract these days, drive primary reward pathways? These issues have critical implications for all aspects of learning, as they determine when a phasic dopamine signal occurs.

How are different timescales of goals encoded? Are the intuitive inner and outer loops of subgoals and goals represented by distinct neural systems or is everything just interleaved within a common distributed system that unfolds over time?

How are dlPFC plans represented such that they can bias the unfolding of a sequence of actions over time to accomplish specific goals? In effect, they function like a computer program – how close is this analogy and are there neural equivalents of core elements such as loops, conditionals, subroutines, variable binding, etc.?

To what extent can existing reinforcement-learning computational frameworks be extended to incorporate the richer motivational systems implicated in this Opinion article?

How can novel goal states be learned? Many models assume a preexisting, limited vocabulary of possible goals, but human goals are open ended and endlessly creative. The ultimate challenge here is to understand how

learning/AI community that is adapting traditional RL algorithms to include multiple goals and related motivational issues [71,72].

'arbitrary' patterns of neural activity turn into something that can motivate and guide behavior toward the achievement of specific, desired outcomes.

Thus, there are many exciting open challenges in understanding the nature of human motivation, and this untapped frontier has great potential for basic and applied scientific benefits. Fundamentally, we all seek control and self-determinism most strongly, and understanding how we achieve at least the illusion of it (and strive to maintain this illusion through various dissonance reduction mechanisms) can provide deep insights into why we can at once be so manifestly irrational and yet much more adaptive, robust, and flexible than any existing artificial system.

Acknowledgments

R.C.O'R. is supported by ONR grants N00014-18-1-2116, N00014-14-1-0670/N00014-16-1-2128, and N00014-18-C-2067.

Disclaimer Statement

R.C.O'R. is Chief Scientist at eCortex, Inc., which may derive indirect benefit from the work presented here.

References

1. Rescorla, R.A. and Wagner, A.R. (1972) A theory of Pavlovian conditioning: variation in the effectiveness of reinforcement and non-reinforcement. In *Classical Conditioning II: Theory and Research* (Black, A.H. and Prokasy, W.F., eds), pp. 64–99, Appleton-Century-Crofts
2. Sutton, R.S. and Barto, A. (1981) Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88, 135–170
3. Silver, D. *et al.* (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 484
4. Lewin, K. (1926) Vorsatz, wille und bedürfnis. *Psychol. Forsch.* 7, 330–385 (in German)
5. Tolman, E. (1948) Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208
6. Miller, G.A. *et al.* (1960) *Plans and the Structure of Behavior*, Holt
7. Powers, W.T. (1963) *Behavior: The Control of Perception*, Hawthorne
8. Klinger, E. (1975) Consequences of commitment to and disengagement from incentives. *Psychol. Rev.* 82, 1–25
9. Schank, R.C. and Abelson, R.P. (1977) *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*, Erlbaum Associates
10. Carver, C.S. and Scheier, M.F. (1982) Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. *Psychol. Bull.* 92, 111–135
11. Wilensky, R. (1983) *Planning and Understanding: A Computational Approach to Human Reasoning*. https://www.osti.gov/biblio/5673187
12. Kuhl, J. (1984) Volitional aspects of achievement motivation and learned helplessness: toward a comprehensive theory of action control. In *Progress in Experimental Personality Research. Normal Personality Processes* (Vol. 13) (Maher, B.A. and Maher, W.B., eds), pp. 99–171, Elsevier
13. Heckhausen, H. and Gollwitzer, P.M. (1987) Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motiv. Emot.* 11, 101–120
14. Bargh, J.A. (1990) Goal and intent: goal-directed thought and behavior are often unintentional. *Psychol. Inq.* 1, 248–251
15. Gollwitzer, P.M. (1993) Goal achievement: the role of intentions. *Eur. Rev. Soc. Psychol.* 4, 141–185
16. Balleine, B.W. and Dickinson, A. (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419
17. Pezzulo, G. *et al.* (2018) Hierarchical active inference: a theory of motivated control. *Trends Cogn. Sci.* 22, 294–306
18. Lamme, V.A.F. (2006) Towards a true neural stance on consciousness. *Trends Cogn. Sci.* 10, 494–501
19. Seth, A.K. *et al.* (2008) Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn. Sci.* 12, 314–321
20. Alexander, G. *et al.* (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381
21. Ongür, D. and Price, J.L. (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* 10, 206–219
22. Saddoris, M.P. *et al.* (2005) Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. *Neuron* 46, 321–331
23. Frank, M.J. and Claus, E.D. (2006) Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol. Rev.* 113, 300–326
24. Rushworth, M.F.S. *et al.* (2007) Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn. Sci.* 11, 168–176
25. Schoenbaum, G. *et al.* (2009) A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat. Rev. Neurosci.* 10, 885–892
26. Kouneiher, F. *et al.* (2009) Motivation and cognitive control in the human prefrontal cortex. *Nat. Neurosci.* 12, 659–669
27. Kennerley, S.W. *et al.* (2011) Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.* 14, 1581–1589
28. Pauli, W.M. *et al.* (2012) Expectancy, ambiguity, and behavioral flexibility: separable and complementary roles of the orbital frontal cortex and amygdala in processing reward expectancies. *J. Cogn. Neurosci.* 24, 351–366
29. Rudebeck, P.H. and Murray, E.A. (2014) The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. *Neuron* 84, 1143–1156
30. Rich, E.L. and Wallis, J.D. (2016) Decoding subjective decisions from orbitofrontal cortex. *Nat. Neurosci.* 19, 973–980
31. Hunt, L.T. *et al.* (2017) Triple dissociation of attention and decision computations across prefrontal cortex. *bioRxiv* Published online August 1, 2017. https://doi.org/10.1101/171173
32. O'Reilly, R. (2006) Biologically based computational models of high-level cognition. *Science* 314, 91–94
33. O'Reilly, R.C. and Frank, M.J. (2006) Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput.* 18, 283–328
34. Hull, C.L. (1943) *Principles of Behavior*, Appleton
35. Maslow, A.H. (1943) A theory of human motivation. *Psychol. Rev.* 50, 370–396
36. Miller, E.K. and Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202
37. Desimone, R. (1996) Neural mechanisms for visual memory and their role in attention. *Proc. Natl. Acad. Sci. U. S. A.* 93, 13494–13499
38. O'Reilly, R.C. *et al.* (1999) A biologically based computational model of working memory. In *Models of Working Memory:*

Mechanisms of Active Maintenance and Executive Control (Miyake, A. and Shah, P., eds), pp. 375–411, Cambridge University Press
39. Alexander, W.H. and Brown, J.W. (2011) Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14, 1338–1344
40. Shenhav, A. *et al.* (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79, 217–240
41. Stuss, D.T. and Alexander, M.P. (2007) Is there a dysexecutive syndrome? *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 362, 901–915
42. Holroyd, C.B. and Yeung, N. (2012) Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn. Sci.* 16, 122–128
43. O'Reilly, R.C. *et al.* (2014) Goal-driven cognition in the brain: a computational framework. *arXiv* Published online April 30, 2014. http://arxiv.org/abs/1404.7591
44. Hopfield, J.J. (1984) Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. U. S. A.* 81, 3088–3092
45. Ackley, D.H. *et al.* (1985) A learning algorithm for Boltzmann machines. *Cogn. Sci.* 9, 147–169
46. Herd, S. *et al.* (2019) Neural mechanisms of human decision-making. *arXiv* Published online December 16, 2019. https://arxiv.org/abs/1912.07660
47. Bargh, J.A. (2006) What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior. *Eur. J. Soc. Psychol.* 36, 147–168
48. Huang, J.Y. and Bargh, J.A. (2014) The selfish goal: autonomously operating motivational structures as the proximate cause of human judgment and behavior. *Behav. Brain Sci.* 37, 121–135
49. Hunt, L.T. *et al.* (2018) Triple dissociation of attention and decision computations across prefrontal cortex. *Nat. Neurosci.* 21, 1471–1481
50. Haber, S.N. (2010) Integrative networks across basal ganglia circuits. In *Handbook of Basal Ganglia Structure and Function. Handbook of Behavioral Neuroscience* (Vol. 20) (Steiner, H. and Tseng, K.Y., eds), pp. 409–427, Elsevier
51. O'Reilly, R.C. (2010) The what and how of prefrontal cortical organization. *Trends Neurosci.* 33, 355–361
52. Hazy, T.E. *et al.* (2010) Neural mechanisms of acquired phasic dopamine responses in learning. *Neurosci. Biobehav. Rev.* 34, 701–720
53. Mollick, J.A. et al. A systems-neuroscience model of phasic dopamine. *Psychol. Rev.* (in press).
54. Gershman, S.J. *et al.* (2010) Context, learning, and extinction. *Psychol. Rev.* 117, 197–209
55. Frank, M.J. (2005) When and when not to use your subthalamic nucleus: lessons from a computational model of the basal ganglia. In *Modelling Natural Action Selection: Proceedings of an International Workshop* (Seth, A.K. *et al.*, eds), pp. 53–60, AISB
56. Wilson, T.D. *et al.* (2014) Just think: the challenges of the disengaged mind. *Science* 345, 75–77
57. Richards, D.A. *et al.* (2016) Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial. *Lancet* 388, 871–880
58. Nelson, S.K. *et al.* (2014) The pains and pleasures of parenting: when, why, and how is parenthood associated with more or less well-being? *Psychol. Bull.* 140, 846–895
59. Achtziger, A. and Gollwitzer, P.M. (2018) Motivation and volition in the course of action. In *Motivation and Action* (Heckhausen, J. and Heckhausen, H., eds), pp. 485–527, Springer
60. Green, L. and Myerson, J. (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* 130, 769–792
61. Rosenbaum, D.A. *et al.* (2014) Pre-crastination: hastening subgoal completion at the expense of extra physical effort. *Psychol. Sci.* 25, 1487–1496
62. Pashler, H. (1994) Dual-task interference in simple tasks: data and theory. *Psychol. Bull.* 116, 220–244
63. Treisman, A. (1996) The binding problem. *Curr. Opin. Neurobiol.* 6, 171–178
64. Anderson, J.R. and Lebiere, C. (1998) *The Atomic Components of Thought*, Erlbaum
65. Stocco, A. *et al.* (2010) Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. *Psychol. Rev.* 117, 541–574
66. Jilk, D. *et al.* (2008) SAL: an explicitly pluralistic cognitive architecture. *J. Exp. Theor. Artif. Intell.* 20, 197–218
67. Mink, J.W. (1996) The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* 50, 381–425
68. Collins, A.G.E. and Frank, M.J. (2014) Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol. Rev.* 121, 337–366
69. Gerfen, C.R. and Surmeier, D.J. (2011) Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* 34, 441–466
70. O'Reilly, R.C. *et al.* (2014) How limited systematicity emerges: a computational cognitive neuroscience approach. In *The Architecture of Cognition: Rethinking Fodor and Pylyshyn's Systematicity Challenge* (Calvo, I.P. and Symons, J., eds), MIT Press
71. Berseth, G. *et al.* (2018) Progressive reinforcement learning with distillation for multi- skilled motion control. *arXiv* Published online February 13, 2018. http://arxiv.org/abs/1802.04765
72. Colas, C. *et al.* (2019) CURIOUS: intrinsically motivated modular multi-goal reinforcement learning. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 1331–1340, PMLR