Unraveling the Mysteries of Motivation

Randall C. O'Reilly

Departments of Psychology and Computer, Science Center for Neuroscience, University of California, Davis 1544 Newton Ct, Davis, CA 95816

Abstract

Motivation plays a central role in human behavior and cognition, and yet is not well-captured by widely-used artificial intelligence (AI) and computational modeling frameworks. This paper 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 neuroscience research and theorizing on these questions, a wealth of insights for constraining the development of computational models of motivation can be found.

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. But 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 widely-used artificial intelligence (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.

oreilly@ucdavis.edu.

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

- 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, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. But defining with computational precision exactly what a goal is, and how it actually functions within 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're 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 will not be 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, multi-factorial model of*

goal states (Figure 1). Specifically, we can think of an overall goal state as composed of 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 [20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 17]:

- *Internal representations of biologically and affectively salient outcomes*, e.g., the *unconditioned stimulus (US)* of classical conditioning, including food, water, social (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, 37, 38]. For example, scoring a goal in a game (using a particular strategy) is an action-oriented, plan-level "goal" for achieving the outcome-goal of *winning*. The 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 OFC) against the relative costs and effort associated with the specific plan being considered to obtain that outcome (in 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 maximize overall utility.

Computationally, this process of converging on an overall goal state can be described as a *parallel constraint satisfaction* process [44, 45, 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 step-wise updates to converge on an overall state that (at least locally) maximizes the overall goodness-of-fit between all 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 make sense of empirical data showing that while *individual* neurons throughout all of these areas exhibit tuning for *each* of the above factors, more comprehensive *population-code* data in monkeys shows clearer differences in overall coding across these areas, in ways that correspond nicely with the above functional account [30, 49]. The separability of these areas also aligns with the separate loops of interconnectivity between these frontal areas and corresponding basal ganglia areas [20, information as part of the overall ventral pathway, while ACC encodes action-based affective information as part of the dorsal pathway [51].

Finally, as noted above, 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, 53, 54, 55].

In summary, from the above discussion, the existing systems neuroscience literature can provide a specific and concrete basis for developing 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-Enaged vs. Goal-Selection States

Another major set of questions for computational models concern the temporal dynamics of when these goal states get 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 can't 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 of 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.

Indeed, 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 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* vs. *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 this same strategy, with exercise equipment, point cards, and especially social media apps tapping into this 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. But 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. feeling that it wasn't 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* vs. 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 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, then 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 will likely require built-in motivational drive states associated with all of the above 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 likely 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, converge to require goal-related processing to be fundamentally sequential in nature, and this sequentiality in turn [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. As such, 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 R's": *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. In 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. imposing sequential processing on an otherwise fundamentally parallel distributed neural computer is the basal ganglia [46, 32], as featured in the ACT-R production system model [64, 65, 66]. The basal ganglia has opposing Go vs. 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]), then 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 actually 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 and Outstanding Questions

In summary, there are considerable constraints and relevant data for developing computational models of distributed goal representations that capture the differential value functions associated with goal-selection vs. the 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 opportunity.

Many important further questions arise directly from the ideas discussed here (see the Outstanding Questions box), including major issues regarding the relevant time scales 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 enaged pursuit. In any case, understanding the basic properties of goal dynamics and

learning at the simplest level would likely provide important insights into these bigger questions. There is also recent work within the machine learning / AI community that is adapting traditional RL algorithms to include multiple goals and related motivational issues [71, 72].

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 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

References

[1]. Rescorla RA, Wagner AR, A theory of Pavlovian conditioning: Variation in the effectiveness of reinforcement and non-reinforcement, in: Black AH, Prokasy WF (Eds.), Classical Conditioning II: Theory and Research, Appleton-Century-Crofts, New York, 1972, pp. 64–99.

[2]. Sutton RS, Barto A, Toward a modern theory of adaptive networks: Expectation and prediction., Psychological Review 88 (2) (1981) 135–170. URL http://www.ncbi.nlm.nih.gov/pubmed/7291377 [PubMed: 7291377]

[3]. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D, Mastering the game of Go with deep neural networks and tree search, Nature 529 (7587) (2016) 484. doi:10.1038/nature16961. [PubMed: 26819042]

[4]. Lewin K, Vorsatz, Wille und Bedürfnis, Psychologische Forschung 7 (1) (1926) 330–385. doi:10.1007/BF02424365.

[5]. Tolman E, Cognitive maps in rats and men., Psychological Review 55 (4) (1948) 189–208. URL http://www.ncbi.nlm.nih.gov/pubmed/18870876 [PubMed: 18870876]

[6]. Miller GA, Galanter E, Pribram KH, Plans and the Structure of Behavior, Holt, New York, 1960.

[7]. Powers WT, Behavior: The Control of Perception, Hawthorne, 1973.

[8]. Klinger E, Consequences of commitment to and disengagement from incentives, Psychological Review 82 (1975) 1–25.

[9]. Schank RC, Abelson RP, Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures, Erlbaum Associates, Hillsdale, N. J., 1977.

[10]. Carver CS, Scheier MF, Control theory: A useful conceptual framework for personality-social, clinical, and health psychology., Psychological Bulletin 92 (1) (1982) 111–135. URL http://www.ncbi.nlm.nih.gov/pubmed/7134324 [PubMed: 7134324]

[11]. Wilensky R, Planning and understanding: A computational approach to human

[12]. Kuhl J, Volitional Aspects of Achievement Motivation and Learned Helplessness: Toward a Comprehensive Theory of Action Control1 1I wish to thank Dr. Charles S. Carver and Dr. Franz E. Weinert for helpful comments on a previous draft of this article, in: Maher BA, Maher WB (Eds.), Progress in Experimental Personality Research, Vol. 13 of Normal Personality Processes, Elsevier, 1984, pp. 99–171. doi:10.1016/B978-0-12-541413-5.50007-3.

[13]. Heckhausen H, Gollwitzer PM, Thought contents and cognitive functioning in motivational versus volitional states of mind, Motivation and Emotion 11 (2) (1987) 101–120. doi:10.1007/BF00992338.

[14]. Bargh JA, Goal and Intent: Goal-Directed Thought and Behavior Are Often Unintentional, Psychological Inquiry 1 (3) (1990) 248–251. doi:10.1207/s15327965pli010314.

[15]. Gollwitzer PM, Goal achievement: The role of intentions, European Review of Social Psychology 4 (1993) 141–185.

[16]. Balleine BW, Dickinson A, Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates., Neuropharmacology 37 (1998) 407–419. URL http://www.ncbi.nlm.nih.gov/pubmed/9704982 [PubMed: 9704982]

[17]. Pezzulo G, Rigoli F, Friston KJ, Hierarchical Active Inference: A Theory of Motivated Control, Trends in Cognitive Sciences 22 (4) (2018) 294–306. doi:10.1016/j.tics.2018.01.009. [PubMed: 29475638]

[18]. Lamme VAF, Towards a true neural stance on consciousness, Trends in Cognitive Sciences 10 (11) (2006) 494–501. doi:10.1016/j.tics.2006.09.001. [PubMed: 16997611]

[19]. Seth AK, Dienes Z, Cleeremans A, Overgaard M, Pessoa L, Measuring consciousness: Relating behavioural and neurophysiological approaches, Trends in Cognitive Sciences 12 (8) (2008) 314–321. doi:10.1016/j.tics.2008.04.008. [PubMed: 18606562]

[20]. Alexander G, DeLong M, Strick P, Parallel organization of functionally segregated circuits linking basal ganglia and cortex., Annual Review of Neuroscience 9 (1986) 357–381. URL http://www.ncbi.nlm.nih.gov/pubmed/3085570

[21]. Ongür D, Price JL, The organization of networks within the orbital and medial prefrontal URL http://www.ncbi.nlm.nih.gov/pubmed/10731217

[22]. Saddoris MP, Gallagher M, Schoenbaum G, Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex., Neuron 46 (2) (2005) 321–331. URL http://www.ncbi.nlm.nih.gov/pubmed/15848809 [PubMed: 15848809]

[23]. Frank MJ, Claus ED, Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal., Psychological Review 113 (2) (2006) 300–326. URL http://www.ncbi.nlm.nih.gov/pubmed/16637763 [PubMed: 16637763]

[24]. Rushworth MFS, Behrens TEJ, Rudebeck PH, Walton ME, Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour., Trends in Cognitive Sciences 11 (4) (2007) 168–176. URL http://www.ncbi.nlm.nih.gov/pubmed/17337237 [PubMed: 17337237]

[25]. Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK, A new perspective on the role of the orbitofrontal cortex in adaptive behaviour., Nature Reviews Neuroscience 10 (12) (2009) 885–892. URL http://www.ncbi.nlm.nih.gov/pubmed/19904278 [PubMed: 19904278]

[26]. Kouneiher F, Charron S, Koechlin E, Motivation and cognitive control in the human prefrontal cortex., Nature neuroscience 12 (9) (2009) 659–669. URL http://www.ncbi.nlm.nih.gov/pubmed/19503087

[27]. Kennerley SW, Behrens TEJ, Wallis JD, Double dissociation of value computations in orbitofrontal and anterior cingulate neurons, Nature Neuroscience 14 (12) (2011) 1581–1589. doi:10.1038/nn.2961. [PubMed: 22037498]

[28]. Pauli WM, Hazy TE, O'Reilly RC, Expectancy, ambiguity, and behavioral flexibility: Separable and complementary roles of the orbital frontal cortex and amygdala in processing reward expectancies., Journal of Cognitive Neuroscience 24 (2) (2012) 351–366. URL http://www.ncbi.nlm.nih.gov/pubmed/22004047 [PubMed: 22004047]

[29]. Rudebeck PH, Murray EA, The orbitofrontal oracle: Cortical mechanisms for the prediction and evaluation of specific behavioral outcomes, Neuron 84 (6) (2014) 1143–1156. doi:10.1016/j.neuron.2014.10.049. [PubMed: 25521376]

[30]. Rich EL, Wallis JD, Decoding subjective decisions from orbitofrontal cortex, Nature Neuroscience 19 (7) (2016) 973–980. doi:10.1038/nn.4320. [PubMed: 27273768]

[31]. Hunt LT, Malalasekera WMN, de Berker AO, Miranda B, Farmer SF, prefrontal cortex T. E. J., bioRxiv (2017) 171173 doi:10.1101/171173 .

[32]. O'Reilly R, Biologically based computational models of high-level cognition., Science 314 (5796) (2006) 91–94. URL http://www.ncbi.nlm.nih.gov/pubmed/17023651 [PubMed: 17023651]

[33]. O'Reilly RC, Frank MJ, Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia., Neural Computation 18 (2) (2006) 283–328. URL http://www.ncbi.nlm.nih.gov/pubmed/16378516 [PubMed: 16378516]

[34]. Hull CL, Principles of Behavior, Appleton, 1943.

[35]. Maslow AH, A Theory of Human Motivation, Psychological Review 50 (1943) 370–396.

[36]. Miller EK, Cohen JD, An integrative theory of prefrontal cortex function., Annual Review of Neuroscience 24 (2001) 167–202. URL http://www.ncbi.nlm.nih.gov/pubmed/11283309

[37]. Desimone R, Neural Mechanisms for Visual Memory and Their Role in Attention., Proceedings of the National Academy of Sciences 93 (24) (1996) 13494–13499. URL http://www.pnas.org/content/93/24/13494

[38]. O'Reilly RC, Braver TS, Cohen JD, A Biologically Based Computational Model of Working Memory, in: Miyake A, Shah P (Eds.), Models of Working Memory: Mechanisms of Active Maintenance and Executive Control., Cambridge University Press, New York, 1999, pp. 375–411.

[39]. Alexander WH, Brown JW, Medial prefrontal cortex as an action-outcome predictor., Nature Neuroscience 14 (10) (2011) 1338–1344. URL http://www.ncbi.nlm.nih.gov/pubmed/21926982 [PubMed: 21926982]

[40]. Shenhav A, Botvinick MM, Cohen JD, The expected value of control: An integrative theory of anterior cingulate cortex function., Neuron 79 (2) (2013) 217–240. URL http://www.ncbi.nlm.nih.gov/pubmed/23889930 [PubMed: 23889930]

[41]. Stuss DT, Alexander MP, Is there a dysexecutive syndrome?, Philosophical transactions of the Royal Society of London 362 (1481) (2007) 901–915. URL http://www.ncbi.nlm.nih.gov/pubmed/17412679 [PubMed: 17412679]

[42]. Holroyd CB, Yeung N, Motivation of extended behaviors by anterior cingulate cortex., Trends in cognitive sciences 16 (2) (2 2012). URL http://www.ncbi.nlm.nih.gov/pubmed/22226543

[43]. O'Reilly RC, Hazy TE, Mollick J, Mackie P, Herd S, Goal-driven cognition in the brain: A computational framework, arXiv:1404.7591 [q-bio] (Apr. 2014). arXiv:1404.7591. URL http://arxiv.org/abs/1404.7591

[44]. Hopfield JJ, Neurons with graded response have collective computational properties like those America 81 (1984) 3088–3092. URL http://www.ncbi.nlm.nih.gov/pubmed/6587342

[45]. Ackley DH, Hinton GE, Sejnowski TJ, A learning algorithm for Boltzmann machines, Cognitive Science 9 (1) (1985) 147–169.

[46]. Herd S, Krueger K, Nair A, Mollick J, O'Reilly R, Neural Mechanisms of Human Decision-Making, Cognitive Affective and Behavioral Neuroscience (2019). URL Preprint:https://arxiv.org/abs/1912.07660

[47]. Bargh JA, What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior, European Journal of Social Psychology 36 (2) (2006) 147–168. doi:10.1002/ejsp.336. [PubMed: 19844598]

[48]. Huang JY, Bargh JA, The Selfish Goal: Autonomously operating motivational structures as the proximate cause of human judgment and behavior, Behavioral and Brain Sciences 37 (2) (2014) 121–135. doi:10.1017/S0140525X13000290. [PubMed: 24775120]

[49]. Hunt LT, Malalasekera WMN, de Berker AO, Miranda B, Farmer SF, Behrens TEJ, Kennerley SW, Triple dissociation of attention and decision computations across prefrontal cortex, Nature Neuroscience 21 (10) (2018) 1471–1481. doi:10.1038/s41593-018-0239-5. [PubMed: 30258238]

[50]. Haber SN, Chapter 24 - Integrative Networks Across Basal Ganglia Circuits, in: Steiner H, Tseng KY (Eds.), Handbook of Basal Ganglia Structure and Function, Vol. 20 of Handbook of Behavioral Neuroscience, Elsevier, 2010, pp. 409–427. doi:10.1016/B978-0-12-374767-9.00024-X.

[51]. O'Reilly RC, The What and How of prefrontal cortical organization, Trends in Neurosciences 33 (8) (2010) 355–361. doi:10.1016/j.tins.2010.05.002. [PubMed: 20573407]

[52]. Hazy TE, Frank MJ, O'Reilly RC, Neural mechanisms of acquired phasic dopamine responses in learning., Neuroscience and Biobehavioral Reviews 34 (5) (2010) 701–720. URL http://www.ncbi.nlm.nih.gov/pubmed/19944716 [PubMed: 19944716]

[53]. Mollick JA, Hazy TE, Krueger KA, Nair A, Mackie P, Herd SA, O'Reilly RC, A systems-neuroscience model of phasic dopamine (In Press).

[54]. Gershman SJ, Blei DM, Niv Y, Context, learning, and extinction., Psychological Review 117 (1) (2010) 197–209. [PubMed: 20063968]

[55]. Frank MJ, When and when not to use your subthalamic nucleus: Lessons from a computational model of the basal ganglia., Modelling Natural Action Selection: Proceedings of an International Workshop (2005) 53–60.

[56]. Wilson TD, Reinhard DA, Westgate EC, Gilbert DT, Ellerbeck N, Hahn C, Brown CL, Shaked A, Just think: The challenges of the disengaged mind, Science 345 (6192) (2014) 75–77. doi:10.1126/science.1250830. [PubMed: 24994650]

[57]. Richards DA, Ekers D, McMillan D, Taylor RS, Byford S, Warren FC, Barrett B, Farrand PA, Gilbody S, Kuyken W, O'Mahen H, Watkins ER, Wright KA, Hollon SD, Reed N, Rhodes S, Fletcher E, Finning K, Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): A randomised, controlled, non-inferiority trial, The Lancet 388 (10047) (2016) 871–880. doi:10.1016/S0140-6736(16)31140-0.

[58]. Nelson SK, Kushlev K, Lyubomirsky S, The pains and pleasures of parenting: When, why, and how is parenthood associated with more or less well-being?, Psychological Bulletin 140 (3) (2014) 846–895. doi:10.1037/a0035444. [PubMed: 24491021]

[59]. Achtziger A, Gollwitzer PM, Motivation and Volition in the Course of Action, in: Heckhausen J, Heckhausen H (Eds.), Motivation and Action, Springer International Publishing, Cham, 2018, pp. 485–527. doi:10.1007/978-3-319-65094-412.

[60]. Green L, Myerson J, A discounting framework for choice with delayed and probabilistic rewards., Psychological bulletin 130 (5) (2004) 769–792. URL http://www.ncbi.nlm.nih.gov/pubmed/15367080 [PubMed: 15367080]

[61]. Rosenbaum DA, Gong L, Potts CA, Pre-Crastination: Hastening Subgoal Completion at the Expense of Extra Physical Effort., Psychological Science 25 (7) (2014) 1487–1496. URL http://www.ncbi.nlm.nih.gov/pubmed/24815613 [PubMed: 24815613]

[62]. Pashler H, Dual-task interference in simple tasks: Data and theory., Psychological bulletin 116 (1994) 220–244. URL http://www.ncbi.nlm.nih.gov/pubmed/7972591 [PubMed: 7972591]

[63]. Treisman A, The Binding Problem., Current Opinion in Neurobiology 6 (1996) 171–178. [PubMed: 8725958]

[64]. Anderson JR, Lebiere C, The Atomic Components of Thought., Erlbaum, Mahwah, NJ, 1998.

[65]. Stocco A, Lebiere C, Anderson J, Conditional Routing of Information to the Cortex: A Model of the Basal Ganglia's Role in Cognitive Coordination., Psychological Review 117 (2010) 541–574. [PubMed: 20438237]

[66]. Jilk D, Lebiere C, O'Reilly RC, Anderson J, SAL: An explicitly pluralistic cognitive architecture, Journal of Experimental & Theoretical Artificial Intelligence 20 (3) (2008) 197–218.

[67]. Mink JW, The basal ganglia: Focused selection and inhibition of competing motor programs., Progress in Neurobiology 50 (4) (1996) 381–425. URL http://www.ncbi.nlm.nih.gov/pubmed/9004351 [PubMed: 9004351]

[68]. Collins AGE, Frank MJ, Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive., Psychological Review 121 (3) (2014) 337–366. URL http://www.ncbi.nlm.nih.gov/pubmed/25090423 [PubMed: 25090423]

[69]. Gerfen CR, Surmeier DJ, Modulation of striatal projection systems by dopamine., Annual Review of Neuroscience 34 (2011) 441–466. URL http://www.ncbi.nlm.nih.gov/pubmed/21469956

[70]. O'Reilly RC, Petrov AA, Cohen JD, Lebiere CJ, Herd SA, Kriete T, How Limited Systematicity Emerges: A Computational Cognitive Neuroscience Approach, in: Calvo IP, Symons J (Eds.), The Architecture of Cognition: Rethinking Fodor and Pylyshyn1s Systematicity Challenge, MIT Press, Cambridge, MA, 2014 URL http://books.google.com/books?hl=en\&lr=\&id=uwI_AwAAQBAJ\&oi=fnd\&pg=PA191\&dq=How+Limited+Systematicity+Emerges:+A+Computational+Cognitive+Neuroscience+Approach\&ots=gz2FpFqziJ\&sig=9X3e5EqqrYmG8TC4oJ30dI0BlBI

[71]. Berseth G, Xie C, Cernek P, Van de Panne M, Progressive Reinforcement Learning with Distillation for Multi-Skilled Motion Control, arXiv:1802.04765 [cs, stat] (Feb. 2018). arXiv:1802.04765. URL http://arxiv.org/abs/1802.04765

[72]. Colas C, Fournier P, Sigaud O, Chetouani M, Oudeyer P-Y, CURIOUS: Intrinsically motivated modular multi-goal reinforcement learning, in: Proceedings of the 36th International Conference on Machine Learning, 2019, pp. 1331–1340. URL http://proceedings.mlr.press/v97/colas19a.html

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 to 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.

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 for implementing computational models of motivational states like *satisfaction*, *frustration*, *boredom*, etc? These involve tracking progress toward achieving 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 time scales 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 paper?

- 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 "arbitrary" patterns of neural activity turn into something that can actually motivate and guide behavior toward achieving specific, desired outcomes.

Three Core Elements of Distributed Goal States
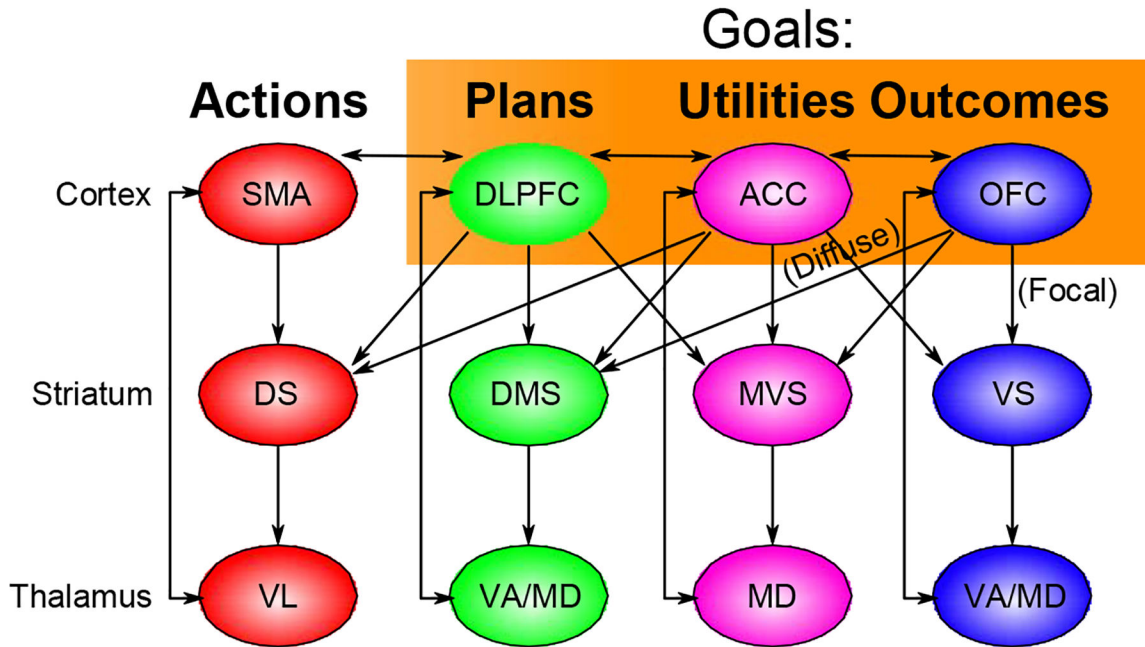
Goals:

Figure 1:

Schematic for how 4 different anatomically-defined loops through the frontal cortex and basal ganglia (Alexander, De-Long, & Strick, 1986) 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). Orbitofrontal cortex (OFC) maintains and tracks biologically-salient outcomes, 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 supports flexible, dynamically-gated active maintenance and updating of these distributed goal states (e.g., [32, 33])
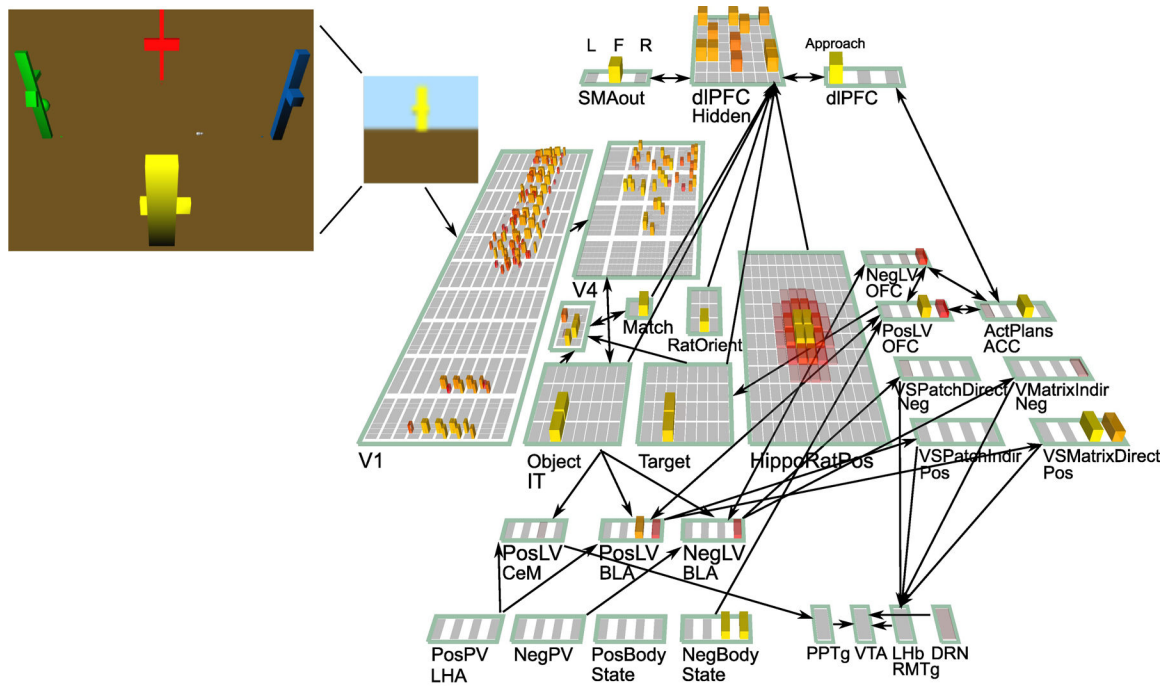
Figure 2:

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, and meat can be rotten. For each run, emery experiences 2 states of negative body state depravation (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 *ACC ActPlans* units encode the 4 different potential *Target*s to approach, coordinating the *OFC PosLV* (positive learned value) representations that learn the specific Target features associated with different PV outcomes, with the specific *dlPFC* motor plan required (approach target in this case). The *NegLV OFC* representation encodes negative outcomes associated with different targets, and helps constrain the target selection process. All of these 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.