Homeostasis after injury: how intertwined inference and control underpin post-injury pain and behaviour

Pranav Mahajan
Nuffield Department
of Clinical Neurosciences
University of Oxford
pranav.mahajan@ndcn.ox.ac.uk

Peter Dayan
Max Planck Institute
for Biological Cybernetics
University of Tübingen
dayan@tue.mpg.de

Ben Seymour
Nuffield Department
of Clinical Neurosciences
University of Oxford
ben.seymour@ndcn.ox.ac.uk

Abstract

Several arguments suggest that the brain might have a dedicated representation of the state of injury. This would provide an internal control system to modulate behaviour given the changed homeostatic priorities associated with injury, including pain, anxiety and mood changes appropriate to the need for heightened protection and recuperation during healing. Here, we propose a computational architecture for how this might be constructed, treating the injury as a partially observable Markov decision process (POMDP), and proposing a Bayesian decision-theoretic solution that combines inference with optimal control. We show how this offers an explanation of two core paradoxical observations: behaviours such as rubbing an injured area (conventionally viewed under the lens of gate control theory), and high propensity of transition to pathological chronic pain states. Overall, this provides a quantitative framework for mapping injury homeostasis to neural substrates, with potential for identifying novel chronic pain targets. Full paper: https://www.biorxiv.org/content/10.1101/2025.02.04.636410v2

Keywords: Bayesian Decision Theory; POMDP; Computational Pain Neuroscience; Fear-Avoidance model

Introduction

Injuries often lead to a period of vulnerability, reduced functionality, and characteristic behavioural changes considered adaptive for safe recovery (Wall, 1979; Williams, 2019), accompanied and potentially mediated by pain. Seymour et al. (2023) offered a new perspective on tonic pain, noting the brain's inevitable uncertainties about injury and recovery. They proposed that the brain continually integrates multisensory inputs to infer an uncertain injury state, a representation tied to action choice and generating internal signals interpreted as pain. This framework extends Bayesian models of pain perception (Seymour et al., 2013; Büchel et al., 2014) to include control under uncertainty, and critically suggests that protective behaviours, by restricting access to informative signals about recovery, might lead to persistent injury beliefs and chronic pain.

We formalize these ideas as a partially observable Markov decision process (POMDP; Kaelbling et al. (1998)), providing the first concrete computational realisation of Seymour et al. (2023) within a reinforcement learning context. Our work addresses several gaps: it (i) unifies Bayesian inference and control approaches in pain research (Seymour & Mancini, 2020), (ii) mathematically formalises how information restriction about injury resolution may drive chronic pain states (Seymour et al., 2023), and (iii) makes explicit how the value of information influences pain-related behaviour (Seymour, 2019). We focus on the belief about the injury state, which underlies tonic pain and is informed by integrating multiple, potentially unreliable, information sources over time (Debanne, 2004; Höfle et al., 2010; Mancini et al., 2022). Our simulations, using a minimal POMDP model detailed in the full paper, explore how the costs of information gathering (e.g., using an injured part to assess recovery) can, counter-intuitively, lead to pathological chronic pain states, even without ongoing peripheral nociception (Fitzcharles et al., 2021). We proceed by studying the properties of its optimal policy, tying the abstract model to observed pain phenomena.

Theory sketch

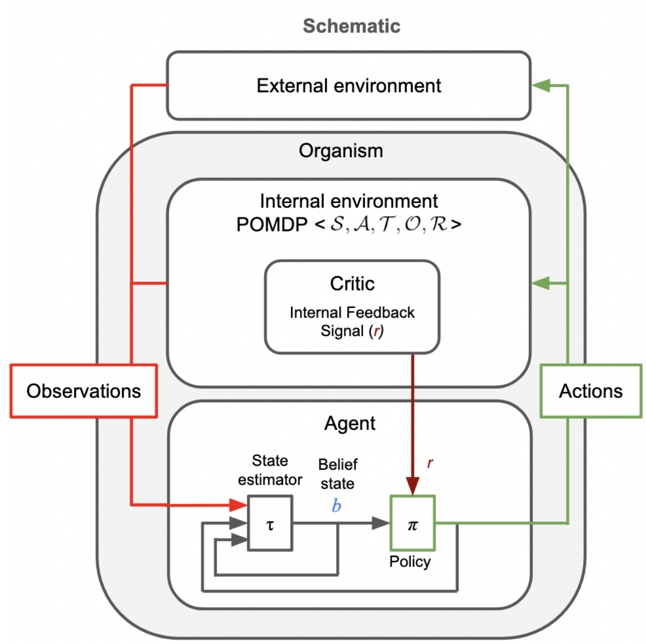

Figure 1: (A) Schematic of the injury POMDP, with an internal environment generating observations and conferring utilities and an internal agent inferring a belief state (Kaelbling et al., 1998) to choose optimal actions.

We propose that the brain treats injury in terms of a partially observable Markov Decision Problem (POMDP; Fig. 1) (Drake, 1962; Astrom et al., 1965; Sondik, 1971; Kaelbling et al., 1998). For the purpose of our simulations, we construct a simplified injury POMDP describing a patient contemplating a demanding activity whilst uncertain about their injury. The true state characterizes the injury; we consider just two possible states $s_t \in \{0, 1\}$ for healthy or injured. However, crucially, the brain lacks full information about the state; interoception being incomplete and noisy so the agent has a belief state $b_t \in [0, 1]$, a probabilistic distribution over states (Kaelbling et al., 1998). We radically simplify actions to three to four: a_{act} (physically demanding, resource-collecting); $a_{\text{r\&r}}$ (recovery); and a_{que} (assessing injury), sometimes including a null action a_{nul}.

Utilities are central to determining the optimal policy and are a contentious aspect of the POMDP. In standard RL, utilities are externally provided, unlike in nature. This has inspired work on homeostatic reinforcement learning (Keramati & Gutkin, 2014) and intrinsic rewards (Chentanez et al., 2004; Barto & Simsek, 2005; Singh et al., 2009; Dayan, 2022). Here, as a simplification, we assume an intrinsic reinforcement function $r(s, a)$, defining immediate affective consequences of action a in true state s. For instance, this function might be large and negative/positive for activity/rest actions when injured; small and negative for investigating while injured; and negative for resting while uninjured (as proxies for long-run effects, further detailed in the full paper). Since the agent only knows its belief b_t about s, the expected utility is internally constructed. Utilities and observation models described in more detail in the full paper.

For concreteness, as a substantial simplification, we associate the belief state b_t with tonic pain – higher belief in being injured corresponds to greater pain. This pain becomes chronic if the agent fails to act or gather evidence to revise its belief. We link the expected negative reinforcement from a_{que} to phasic pain caused by injury investigation during the episode. This expectation averages over the belief state, providing a mechanism for precisely tuning the feedback (Seymour, 2019) in case of correct inferences, but is also susceptible to incorrect or underinformed inferences.

Normative consequences

Why do we investigate injury despite it being painful? Our model provides a normative explanation for investigating an injury despite it being painful, a behaviour difficult to explain through simple inferential or control-theoretic approaches alone. In this instance, a_{que} when injured incurs a cost (e.g., $r(s = 1, a_{\text{que}}) = -4$) but provides more accurate information about the internal state than other actions. For this demonstration, we omit a_{nul} for simplicity, though its inclusion does not alter the core finding (details in full paper).

When starting from an uncertain belief state ($b_0 = 0.5$), the agent chooses a_{que} at the cost of some phasic pain due to the value of information (Fig. 2). With each sample, the agent updates its belief until committing to a_{act} or $a_{\text{r\&r}}$ (Fig. 2, red and blue arrows). If not injured ($s = 0$), beliefs update (red arrows) leading to a_{act}; if injured ($s = 1$), beliefs update (blue arrows) leading to $a_{\text{r\&r}}$. This demonstrates agents choosing costly investigative actions to accrue evidence for optimal decisions. Such actions extend beyond ”rubbing the injured area” to include non-contact explorations like moving a painful joint.

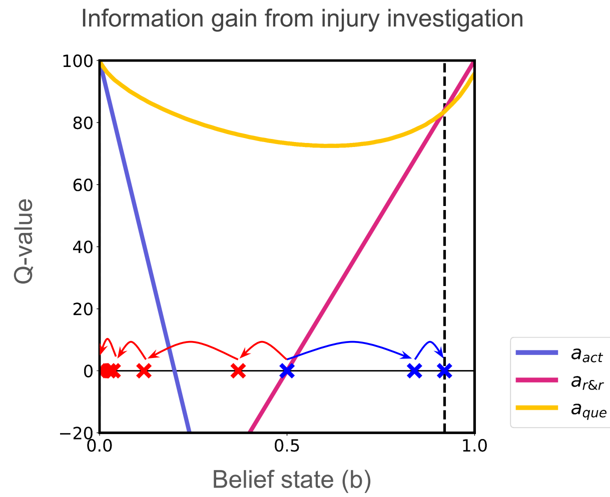

Figure 2: Action values after value iteration. Red arrows show belief updates under true state $s = 0$ (not injured), whereas blue arrows show belief updates under true state $s = 1$ (injured). The agent takes multiple a_{que} actions, reaching belief thresholds where choosing a_{act} or $a_{\text{r\&r}}$ is more valuable than a_{que}, thus terminating the episode.

For further results on the trade-off between phasic pain and information gain, and results on dysfunctional consequences due to information-restriction and aberrant priors, please refer to the full paper.

Discussion

We present a theoretical framework for understanding the computational logic of a dedicated homeostatic state for injury (Seymour et al., 2023), where internal states are partially observable and require inference. This perspective normatively explains behaviours such as probing an injury despite immediate phasic pain to acquire information, and identifies ”fault lines” (e.g., information restriction, aberrant priors detailed in the full paper) that can lead to suboptimal behaviour and chronic pain. Our framework complements the Fear-Avoidance model (Vlaeyen & Linton, 2000; Vlaeyen et al., 2016); by formalising information restriction dynamics, it provides a computational, inference-based account of how avoidance can sustain maladaptive beliefs about injury resolution. This approach, potentially relevant to understanding clinical conditions like sciatica and phenomena such as boom-bust cycles (further discussed in the full paper), offers a mathematical foundation for making theories of this complex condition explicit, which is of translational importance.

Acknowledgments

PM would like to thank anonymous CCN 2025 reviewers, Michael Browning, Rafal Bogacz, Suyi Zhang, Charlie Yan, Shuangyi Tong, Danielle Hewitt for their feedback on presentations or earlier drafts of the manuscript. The work was funded by Wellcome Trust (214251/Z/18/Z, 203139/Z/16/Z and 203139/A/16/Z), IITP (MSIT 2019-0-01371) and JSPS (22H04998). This research was also partly supported by the NIHR Oxford Health Biomedical Research Centre (NIHR203316). PD is funded by the Max Planck Society and the Humboldt Foundation. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

Astrom, K. J., et al. (1965). Optimal control of markov processes with incomplete state information. *Journal of mathematical analysis and applications*, *10*(1), 174–205.

Barto, A. G., & Simsek, O. (2005). Intrinsic motivation for reinforcement learning systems. In *Proceedings of the thirteenth yale workshop on adaptive and learning systems* (pp. 113–118).

Büchel, C., Geuter, S., Sprenger, C., & Eippert, F. (2014). Placebo analgesia: a predictive coding perspective. *Neuron*, *81*(6), 1223–1239.

Chentanez, N., Barto, A., & Singh, S. (2004). Intrinsically motivated reinforcement learning. *Advances in neural information processing systems*, *17*.

Dayan, P. (2022). "liking" as an early and editable draft of long-run affective value. *PLoS Biology*, *20*(1), e3001476.

Debanne, D. (2004). Information processing in the axon. *Nature Reviews Neuroscience*, *5*(4), 304–316.

Drake, A. W. (1962). *Observation of a markov process through a noisy channel*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Fitzcharles, M.-A., Cohen, S. P., Clauw, D. J., Littlejohn, G., Usui, C., & Häuser, W. (2021). Nociplastic pain: towards an understanding of prevalent pain conditions. *The Lancet*, *397*(10289), 2098–2110.

Höfle, M., Hauck, M., Engel, A. K., & Senkowski, D. (2010). Pain processing in multisensory environments. *e-Neuroforum*, *1*, 23–28.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial intelligence*, *101*(1-2), 99–134.

Keramati, M., & Gutkin, B. (2014). Homeostatic reinforcement learning for integrating reward collection and physiological stability. *Elife*, *3*, e04811.

Mancini, F., Zhang, S., & Seymour, B. (2022). Computational and neural mechanisms of statistical pain learning. *Nature Communications*, *13*(1), 6613.

Seymour, B. (2019). Pain: a precision signal for reinforcement learning and control. *Neuron*, *101*(6), 1029–1041.

Seymour, B., Crook, R. J., & Chen, Z. S. (2023). Post-injury pain and behaviour: a control theory perspective. *Nature Reviews Neuroscience*, 1–15.

Seymour, B., Dolan, R., et al. (2013). Emotion, motivation, and pain. *Textbook of Pain*, 248–255.

Seymour, B., & Mancini, F. (2020). Hierarchical models of pain: Inference, information-seeking, and adaptive control. *NeuroImage*, *222*, 117212.

Singh, S., Lewis, R. L., & Barto, A. G. (2009). Where do rewards come from. In *Proceedings of the annual conference of the cognitive science society* (pp. 2601–2606).

Sondik, E. J. (1971). *The optimal control of partially observable markov processes*. Stanford University.

Vlaeyen, J. W., Crombez, G., & Linton, S. J. (2016). The fear-avoidance model of pain. *Pain*, *157*(8), 1588–1589.

Vlaeyen, J. W., & Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*, *85*(3), 317–332.

Wall, P. D. (1979). Three phases of evil: the relation of injury to pain. *Brain and Mind*, *69*, 293.

Williams, A. C. d. C. (2019). Persistence of pain in humans and other mammals. *Philosophical Transactions of the Royal Society B*, *374*(1785), 20190276.