
Proceedings Paper:
jimenez-Rodiguez, A. and Prescott, A.J. orcid.org/0000-0003-4927-5390 (2023) Motivational modulation of consummatory behaviour and learning in a robot model of spatial navigation. In: Meder, F., Hunt, A., Margheri, L., Mura, A. and Mazzolai, B., (eds.) BIomimetic and Biohybrid Systems: 12th International Conference, Living Machines 2023, Genoa, Italy, July 10–13, 2023, Proceedings, Part II. Living Machines 2023, 10-13 Jul 2023, Genoa, Italy. Lecture Notes in Computer Science (LNAI 14158). Springer , Berlin , pp. 240-253. ISBN 9783031395031

https://doi.org/10.1007/978-3-031-39504-8_17

Motivational modulation of consummatory behaviour and learning in a robot model of spatial navigation *

Alejandro Jimenez-Rodriguez[1,3][0000−0001−7172−1794] and Tony J. Prescott[2,3][0000−0003−4927−5390]

[1] Sheffield Hallam University, UK
[2] University of Sheffield, Sheffield, UK
[3] Sheffield Robotics, Sheffield, UK email{a.jimenez-rodriguez, t.j.prescott}@sheffield.ac.uk

Abstract. We present a biomimetic model of motivated behaviour based on the network architecture of the mammalian hypothalamus and its interaction with brain systems involved in reward, memory and decision-making. Specifically, a novel model of the hypothalamus, viewed as a layered structure, is integrated with a previously-developed model of the hippocampal-striatal network controlling a simulated robot in a navigation task. Hypothalamic modulation of model dopamine signals allows the robot to learn the location of reward while regulating simulated food intake. When 'satiated' the robot explores, when 'hungry' it moves towards the learned food source. We discuss the potential uses and future challenges of such models in the development of autonomous robots.

Keywords: Hypothalamus · Motivation · Reinforcement learning · Navigation

1 Introduction

Behaviour in animals is orchestrated to satisfy different needs arising from the homeostatic constraints that make life possible [23]. A hungry animal, for example, needs to forage for food while remembering the location of reliable food sources that can be exploited in the future. On the other hand, a satiated animal is free to explore and engage in other behaviours, typically in order to the satisfy other needs (e.g. mating, nesting, etc.). Robots lack natural needs and therefore have no intrinsic motivational grounding for their behaviour [14]. The absence of any genuine needs has been argued to be a fundamental and immutable difference between robots and animals [4]. On the other hand, one could argue that robots can have needs in the sense that they require certain resources and inputs (power and maintenance, for example) to function properly. The instantiation of goals, provided by humans, in robot control systems can also play a similar role to natural needs in co-ordinating behaviour.

* Supported by Horizon 2020, Horizon Europe, and UK Research and Innovation.

Many animal physiological needs, for instance, to maintain blood oxygen and sugar levels, hydration, body temperature and so on, are regulated by the central and autonomic nervous systems. A layered architecture of brain systems is involved in this regulation [8, 25, 26, 29], where the hypothalamus, a forebrain structure situated near the midline of the brain, plays a critical role. In this work, we propose a biomimetic model of the hypothalamus and its interaction with brain systems involved in decision-making (basal ganglia) and navigation (hippocampus). We show that this model can simulate alternation between consummatory behaviour and learning through the regulation of simulated physiological variables. Our long-term goal is to better understand the neural basis of motivated behaviour in animals, including humans. We consider that such an understanding could help illuminate these critical differences between evolved organisms and intelligent artefacts such as robots.

2 The role of the hypothalamus in motivated behaviour

As noted, in mammalian brains, the hypothalamus plays a critical role in the orchestration of motivated behaviour. In particular, it acts as an integrator for bodily signals that provide an up-to-date picture of homeostatic needs, and delivers outputs based on these signals to a wide network of brain regions that drive different aspects of behaviour [5,25]. The hypothalamus is a heteregeneous structure composed of multiple subregions, of particular importance, to the current work, are the dorsal and lateral areas. The dorsal hypothalamus (DH) is in contact with physiological milieu and generates signals that correlate with lower and higher levels of a variety of substances within the body. Those signals are used to drive a motivational state representation in the lateral hypothalamus (LH) that also integrates signals from other cortical and subcortical structures [3, 10, 24].

Approach and avoid signals generated by the LH are used to modulate the downstream dopaminergic neurons in the ventral tegmental area (VTA) [17, 24] which project to major areas of the brain, and prominently to the nucleus accumbens (NAcc), which is an input structure for action selection systems in the basal ganglia [9, 21]. VTA dopamine acts as both a motivational proxy and as a saliency signal that drives learning [13, 16, 21]. Motivational functions of dopamine include stimulation of behavioural activation, effort and appetite [1] and modulation of action selection in the basal ganglia [19], Phasic increases and decreases in dopamine, locked to environmental stimuli, allow for the modulation of reinforcement learning in regions including the striatum and the hippocampus [22]. Hence, in the current model simulated dopamine signals act as both a learning signal and to activate different motivational systems. GABAergic signals from the VTA have been found to act along with dopamine to allow associative learning in the NAcc [7]. In this work we use a similar chollinergic signal to modulate the plasticity of the hippocampal-striatal network.

In previous work, we presented a robotic model of the hippocampal-striatal network that allows the robot to learn the location of a reward [28]; we have also separately introduced a layered dynamical model of motivation [11]. The current

model combines simulation of some of the neuronal networks involved in motivation, representing the different regions discussed in the previous paragraphs. This model is integrated with the hippocampal network in a layered architecture [18, 29] and tested using a simulation of the animal-like robot platform, MiRo-e [15].

In the following sections we introduce the main components of the architecture: the motivational model, the cholinergic network and the navigational model. Then we present some simulation results.

3 The motivational model

As described in the introduction, our motivational model is composed of three layers as shown in figure 1. In each layer we have abstracted a network motif that performs specific computations relevant for the encoding motivational state of the agent and the corresponding modulation of learning and behaviour. Each of the neurons in the model is a rate neuron:

$$\tau \frac{dx_k}{dt} = -x_k + f(\sum w_{jk} x_j + \theta_k), \tag{1}$$

where, x_k is the firing rate of the k_{th} neuron, w_{jk} the synaptic weights and θ_i a bias term. The activation function is sigmoidal $f(x) = 1/(1 + \exp(-\beta x))$. Table 1 shows the different neuron populations in the model.

Table 1: Abbreviations used for the different neuron populations in the model

Variable	Description
x_s	Sensitive neurons in the Hammel network
x_i	Insensitive neurons in the Hammel network
x_h	Hunger neurons
x_s	Satiety neurons
x_{app}	Approach pathway neurons
x_{av}	Avoid pathway neurons
x_{acc}	Nucleus accumbens neurons
x_{vta_gaba}	VTA GABA neurons
x_{da}	Dopaminergic neurons
x_{msn}	Medium Spiny Neurons in the striatum
x_{cin}	Collinergic interneurons

In the upper layer or ventricular part, representing the dorsal hypothalamus, we propose a version of the network motif known as the *Hammel model* [6] in order to encode "hunger" and "satiety" states. This mechanism effectively defines a set-point based upon physiological signals, which are modeled as first order ODEs:

$$\tau \frac{du}{dt} = \alpha(1 - u) - ur(t), \tag{2}$$

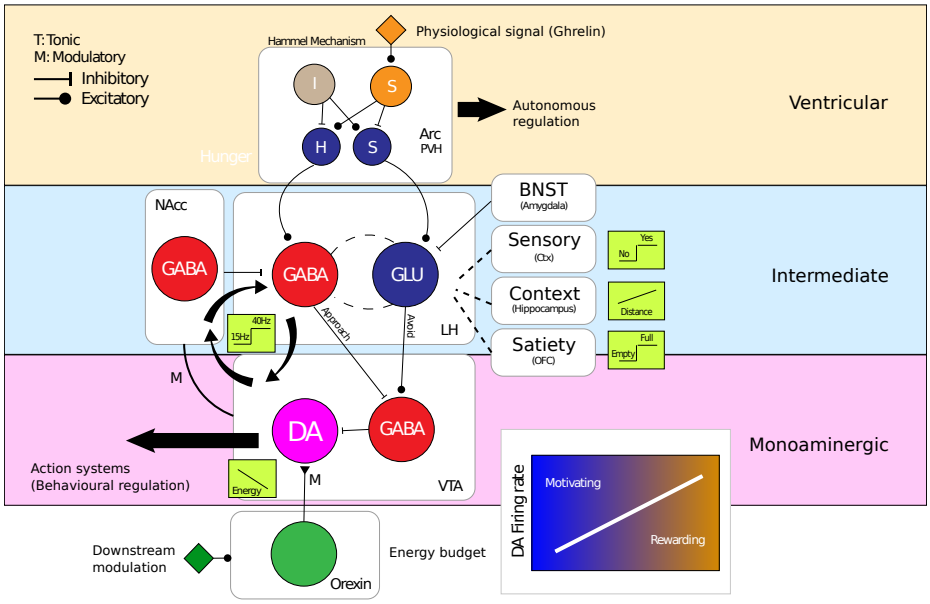

Fig. 1: Motivational circuit inspired by the hypothalamus architecture. The ventricular part, modelled on the dorsal hypothalamus, is in contact with the physiological millieu, and drives hunger (H) and satiety (S) signals that activate approach and avoid channels in the Lateral Hypothalamus (LH). These signals are reinforced by the nucleus accumbens. The approach pathway drives behavioural activation by disinhibiting VTA dopamine neurons (DA), shown here in the monoaminergic layer, by removing tonic inhibition from GABA cells. Inset: From the accumbens perspective, different dopamine regimes drive exploration/explotation trade-offs. See text for further explanation.

where $r(t)$ is the reward obtained from the environment, α, an accumulation rate, and τ a time constant that determines the timescale. Physiological signals accumulate indicating increasing levels of thirst, hunger, etc. The four neurons in the mechanism evolve according to the dynamics:

$$\begin{aligned}
\tau_s \frac{dx_s}{dt} &= -x_s + f(w_u u), \\
\frac{dx_i}{dt} &= 0, \\
\tau \frac{dx_h}{dt} &= -x_h + f(w_{s\to h} x_s + w_{i\to h} x_i), \\
\tau \frac{dx_{st}}{dt} &= -x_{st} + f(w_{s\to st} x_s + w_{i\to st} x_i).
\end{aligned} \tag{3}$$

Here, the sensitive neurons x_s need to overcome the constant firing rate of the insensitive ones x_i to elicit hunger x_h state; otherwise the output x_{st} indicate a

satiety state. We assume $\tau = 1$ for the rest of the model, except for τ_s which is adjusted to the particular choice of physiological signal.

The intermediate layer is composed of the approach and avoid channels of the LH. The approach channel is GABAergic [10] and the avoid one is glutamatergic. In the current model they do not interact.

$$\begin{aligned} \frac{dx_{app}}{dt} &= -x_{app} + f(w_{acc\to app}x_{acc} + w_{h\to app}x_h), \\ \frac{dx_{av}}{dt} &= -x_{av} + f(w_{st\to av}x_{st}), \end{aligned} \tag{4}$$

Note that this channel is modulated by higher order areas in the brain, however, we focus on the connection with the NAcc whose dynamics we model with the equation:

$$\frac{dx_{Acc}}{dt} = -x_{Acc} + f(wx_{Acc} + \theta_{da}). \tag{5}$$

Here, the dopamine input, acts as a bifurcation parameter. From figure 2 we can see that the accumbens undergoes a cusp bifurcation and it is therefore bistable. As dopamine increases, the population transitions from a high frequency firing regime to a low frequency one. Finally, in the monoaminergic layer, the

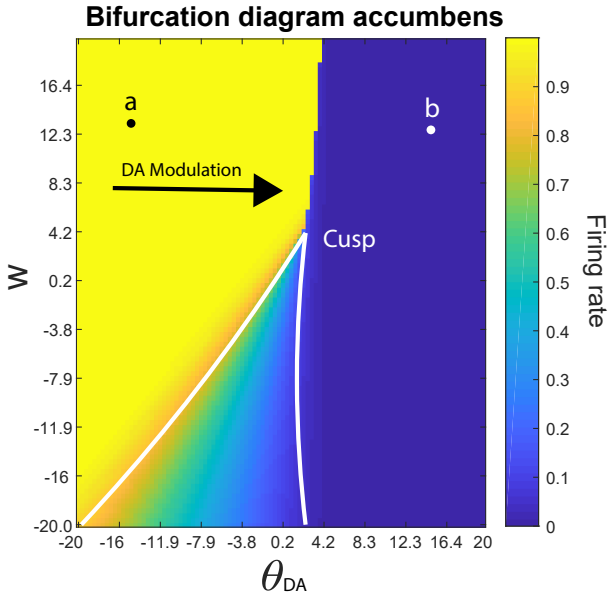

Fig. 2: Bifurcation diagram of the nucleus accumbens showing a cusp catastrophe. The dopamine signal drives the firing rate to a low frequency firing in an abrupt change. This allows for sharp transition between motivated behaviour and exploration.

VTA is composed of two populations. The GABAergic VTA neurons inhibit the dopaminergic neurons and are excited by the the avoid pathway while inhibited by the approach pathway. Thus, behavioural activation works by disinhibition of dopamine neurons. The full model of the VTA is:

$$\begin{aligned} \frac{dx_{vta\_gaba}}{dt} &= -x_{vta\_gaba} + f(w_{app\to vta\_gaba}x_{app} + w_{av\to vta\_gaba}x_{av}) \\ \frac{dx_{da}}{dt} &= -x_{da} + f(w_{vta\_gaba\to da}x_{vta\_gaba} + \theta_{orexin}). \end{aligned} \tag{6}$$

Note that we have added an *orexin* related modulation term to modulate the excitability of the dopaminergic neurons and therefore the timescales used by acetylcholine in the next section. As shown in figure 1, the approach and avoid pathways in the intermediate part are subject to modulation from different brain areas [5].

4 Cholinergic signal

To modulate learning, we generate a transient signal λ_{ach} by combining the GABAergic and dopaminergic outputs from the VTA in two steps circuit similar to [7]:

$$\begin{aligned} \frac{dx_{msn}}{dt} &= -x_{msn} + f(w_{vta\_gaba\to msn}x_{vta\_gaba} + \theta_{msn}) \\ \frac{dx_{cin}}{dt} &= -x_{cin} + f(w_{da\to cin}x_{da} + w_{msn\to cin}x_{msn} + \theta_{cin}) \end{aligned}, \tag{7}$$

with $\lambda_{ach} = \Theta(x_{cin} - b)$. Note that the GABA input disinhibits the cholinergic output which is being excited by dopamine; the parameters are chosen in such a way that the CIN output tracks the start and end of the consummatory behaviour (see figure 3). On the other hand, the teaching signal is a binary value that determines when learning should occur (see next section).

5 The navigation model

The hippocampal-striatal model has been previously published in [28], here we show only an overview of the important parts of its integration with the motivational model. The main parts of the model are shown in figure 4. For the model we assume the MiRo-e is located in a simulated open arena and its coordinates are known at each step.

5.1 Hippocampal network

The hippocampal network is composed of 100 place cells connected in a grid to their immediate neighbours (figure 4). Their firing rate is given by:

$$x_j = \begin{cases} 0 & \text{if } x_j^{'} < 0 \\ 100 & \text{if } x_j^{'} > 100 \\ x_j^{'} & \text{otherwise.} \end{cases} \tag{8}$$

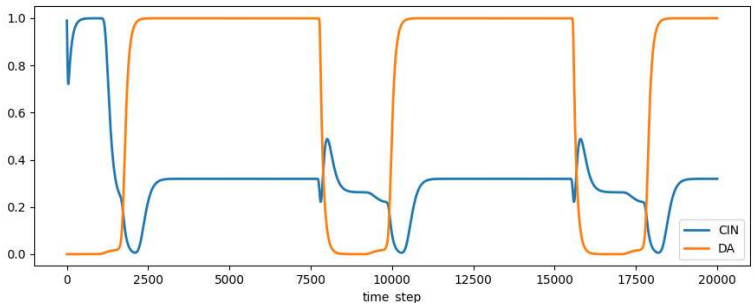

Fig. 3: The activity of the Cholinergic interneurons tracks the start and end of behavioural activation. Shown is a fraction of the simulation DA and CIN time series to illustrate this behaviour.

with the rate defined as rectified function of its current:

$$x_j^{'} = \alpha \left(I_j - \epsilon \right)$$

The dynamics of this current are given by:

$$\tau_I \frac{d}{dt} I_j = -I_j + \psi_j I_j^{syn} + I_j^{place}, \tag{9}$$

where, I_j^{syn} is the synaptic current from neighboring place cells, ψ_j is the intrinsic plasticity (see paper [28] for details) and I_j^{place} is the input due to a place field given by:

$$I_j^{place} = I_{max}^{p} \exp \left[- \frac{(x_{MiRo}^c - x_j^c)^2 + (y_{MiRo}^c - y_j^c)^2}{2d^2} \right], \tag{10}$$

where the position of the robot is assumed to be known. The synaptic input is modulated by the cholinergic output of the motivational model (see previous section). Additionally it possesses depression and facilitation terms D_k and F_k to ensure the correct dynamics during learning:

$$I_j^{syn} = \lambda_{ach} \sum_{k=1}^{8} w_{jk}^{place} x_k D_k F_k, \tag{11}$$

see [28] for details on the depression and facilitation dynamics.

5.2 Action cells

There are 72 action cells, each coding for 1/360 degrees of orientation. Firing rates are drawn from a normal distribution

$$y_i \sim \mathcal{N} \left(\tilde{y}_i, \sigma^2 \right), \tag{12}$$

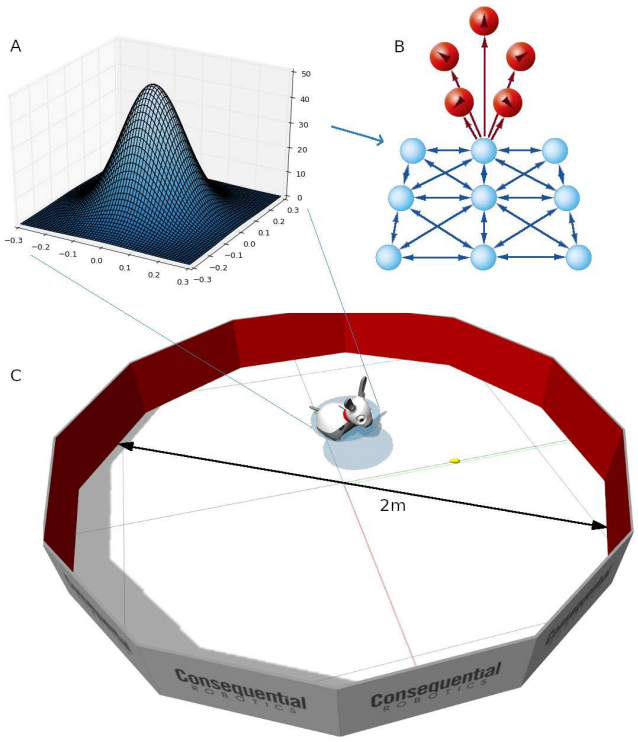

Fig. 4: Hippocampal striatal network. A. Each place field is a Gaussian centered at a given preferred position. B. The network is composed of the hippocampal network that has a feed-forward projection to action cells. C. The simulations are performed in a circular arena with the robotic platform MiRo-e.

where the mean is computed according to

$$\tilde{y}_i = \frac{x_{da}}{1 + \exp\left[-c_1 \sum_{j=1}^{100} w_{ij}^{PC\text{-}AC} x_j - c_2\right]}, \tag{13}$$

and where x_{da} is the output of the dopaminergic cells in the VTA. Note that firing rate is modulated by the place cells. The target direction of the robot is computed by tallying the preferred directions of all cells weighted by their firing rates:

$$\theta_{target} = \arctan\left(\frac{\sum_i y_i \sin\theta_i}{\sum_i y_i \cos\theta_i}\right). \tag{14}$$

The magnitude of the activation in that direction is computed as the magnitude of the population vector.

$$m_{target} = \sqrt{\left(\sum_i y_i \sin\theta_i\right)^2 + \left(\sum_i y_i \cos\theta_i\right)^2}, \tag{15}$$

this is used in action selection (exploration vs. consummatory behaviour).

5.3 PC → plasticity

The synapses from the place cells are plastic and are updated using a policy gradient RL methods (see full derivation in [28]). The weights are updated according to:

$$\frac{dw_{ij}^{PC\text{-}AC}}{dt} = \frac{\eta}{\sigma^2} R e_{ij}, \tag{16}$$

where $R = \lambda_{ach}$ in this case and e_{ij} is the eligibility trace that evolves according to

$$\frac{de_{ij}}{dt} = -\frac{e_{ij}}{\tau_e} + (y_i - \tilde{y}_i)(1 - \tilde{y}_i)\tilde{y}_i x_j. \tag{17}$$

6 Results

We performed a continuous simulation during 300000 steps where the simulated MiRo-e robot could behave freely in an open arena using Gazebo 11 and the MiRo MDK (v201904). In each step of the simulation, the motivational model is integrated first and then its output is fed to the navigational model whose output is the presence or absence of reward. Once a Acetylcholine signal is detected, behaviour is stopped for as long as it lasts; during this time, the weights are updated using the policy gradient method.

In figure 5 we show the dynamics of the main variables of the motivational model during the whole session. We chose the hormone *Ghrelin* to act as the physiological proxy of hunger. The robot manages to keep the hormone levels from saturating and therefore is able to survive. Note that the location of the reward is not known before hand. If learning is turned off (figure 6), the robot only randomly satisfies its needs when passing by the location of the rewards.

It is useful to investigate the trajectories from the start to the end of behavioural activation, as indicated by the cholinergic signal (figure 3). When learning is activated, the trajectories clearly converge to the reward position during the consummatory behaviour and away from it when the robot is satiated (figure 7). The transparency of the trajectories has been made proportional to the index of the consummatory event. On the other hand, when learning is deactivated, trajectories remain random (figure 8)

7 Conclusions and future work

We have proposed a model of motivated consummatory behaviour that integrates hypothalamic and hippocampal-striatal networks showing that this model is able to modulate learning whilst satisfying certain homeostatic constraints. The simulated MiRo-e robot, with this control architecture, is able to find and remember the reward location, satisfying its needs continuously, while performing exploratory behaviour whenever the energy budget allows.

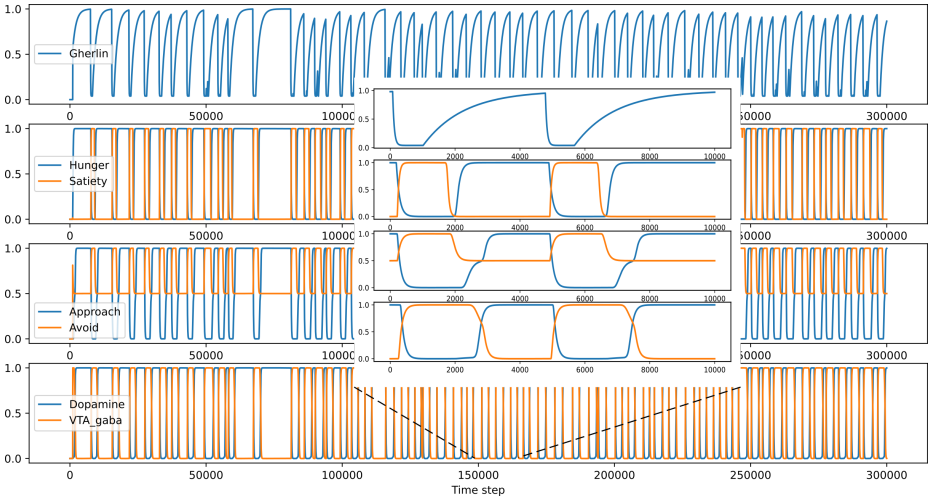

Fig. 5: Time series of the main variables of the motivational model during the whole recording session. Note that the needs are better satisfied over time. From top to bottom. The hormone representing hunger increases over time and decreases once the reward has been consumed. The hunger and satiety signals alternate depending upon the hormone levels. The approach and avoid pathways drive behavioural activation. Dopaminergic and GABAergic signals in the VTA alternate to modulate behaviour and learning.

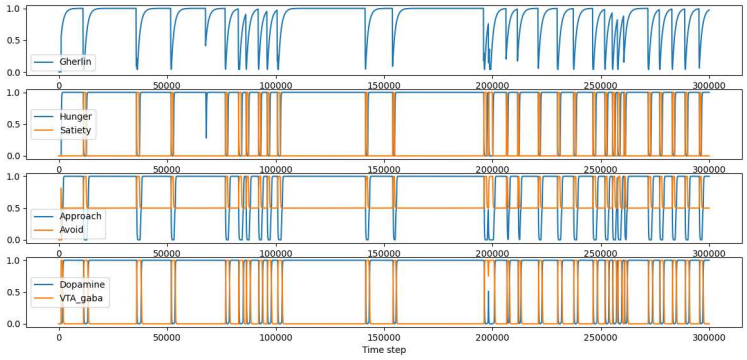

Fig. 6: Time series without learning. The robot satisfies its needs only when it finds the reward by accident.

Although our motivational model simplifies the dynamics of each of the regions involved we consider that it captures important roles of key circuits involving the hypothalamus and VTA in the control of behavioural activation.

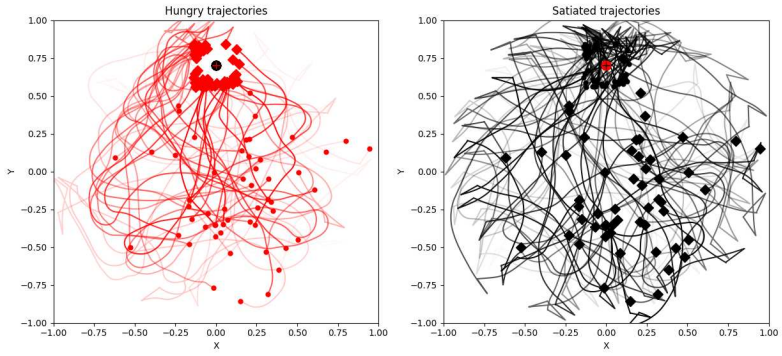

Fig. 7: Trajectories of MiRo during consummatory and exploratory behaviours. The starting point of the trajectory is marked with a circle, while the ending point with a diamond. The place of the reward is marked with a cross. The alpha channel represents time, with opaque trajectories representing later trials. Left. Trajectories of consummatory behaviours when the robot gets hungry, it can be seen that the location of the rewards has been learn and performance improves over time. Right. Once the robot is satiated, it is free to engage of exploration away from the reward until its energy reserves are depleted.

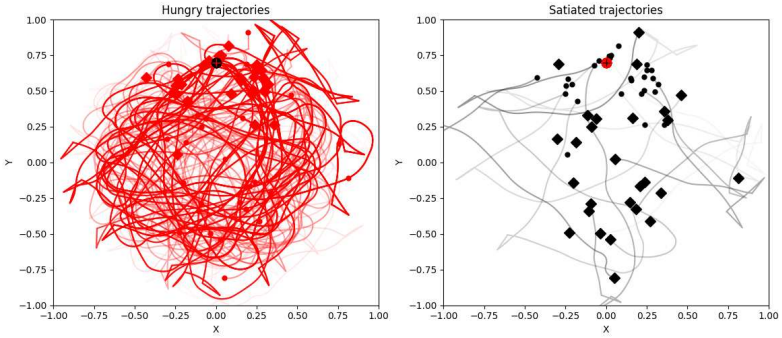

Fig. 8: Trajectories without learning. Same as before. Note that the black trajectories are almost absent, meaning that the robot is hungry most of the time.

In particular, it demonstrates that a push-pull motivational system that operates through parallel approach and avoid channels can allow the orchestration of motivational systems in order to satisfy needs and to generate useful reward prediction errors. We intend to explore potential uses for this model in biomimetic control of autonomous robots for continual learning, and as a platform to investigate motivation-related disorders such as addiction and compulsive eating.

Future work will also include developing a more detailed model of the avoid pathway that involves serotonin and different regions of the nucleus accumbens [27]. This will allow extinction and punishment to be explored alongside reward. The role of acetylcholine in the model should also be refined and, importantly, the roles of nucleus accumbens and amygdala in motivational phenomena such as "wanting" and "liking" [2], the current model can be considered to represent only wanting.

Even though the model comprises just one explicit motivation vs. an explicit drive to explore, we think that the layered architecture presented could harbor multiple motivations as a result of the separation between physiological state (in the ventricular layer) and motivational state (in the intermediate layer). Indeed, this excitatory-inhibitory balanced networks have been shown to support the creation of multiple assemblies and are a good target for modulation of higher areas as shown in figure 1.

Overall, this work is a contribution to the ongoing effort to integrate motivational aspects in robotic cognitive architecture and in machine learning in general [12, 20, 29]. Looking further ahead, in animals, needs and motivation are closely related to emotions and feelings, and are processed in similar areas of the brain. This relationship between the motivational and experiential aspects of being is poorly understood, however, as noted by Craig:

"...our affective feelings derive from the brain networks that generate flexible and adaptable emotional behaviors, which evolution built by expanding upon the ancient homeostatic neural systems that automatically take care of the body. In other words, the affective feelings that you experience are interoceptive reflections of emotional motivations, which are expressed by activity throughout the peripheral and central autonomic systems of your body and your brain and which produce behavior that you "feel" as it happens." [8](p. 12).

We therefore consider that robot modelling that illuminates the neural basis for motivation could ultimately provide a better understanding of the role of the brain in generating affective experience.

Acknowledgements This work was supported by the EU H2020 Programme as part of the FET Flagship Human Brain Project (HBP-SGA3, grant no. 945539), and specifically, through the CATRA (Cognitive Architecture for Therapy Robots and Avatars) project which was supported by the EBRAINS Research Infrastructure Voucher Programme. It was also supported by the UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee for the EIC Pathfinder CAVAA project. We are grateful to Matthew Whelan and Eleni Vasilaki with whom the hippocampal model was developed.

Declaration of Interest TJP is a director and shareholder of Consequential Robotics Ltd which develops the MiRo robot, and Bettering Our Worlds (BOW) ltd which develops robot software. AJR has no competing interests.

Code availability The code is available at `https://github.com/aljiro/motivation_lm`

References

1. Beeler, J.A., Frazier, C.R., Zhuang, X.: Putting desire on a budget: dopamine and energy expenditure, reconciling reward and resources. Frontiers in integrative neuroscience **6**, 49 (2012)
2. Berridge, K.C.: Food reward: brain substrates of wanting and liking. Neuroscience & Biobehavioral Reviews **20**(1), 1–25 (1996)
3. Berthoud, H.R., Münzberg, H.: The lateral hypothalamus as integrator of metabolic and environmental needs: from electrical self-stimulation to optogenetics. Physiology & behavior **104**(1), 29–39 (2011)
4. Boden, M.: Robot says: Whatever (2018), `https://aeon.co/essays/the-robots-wont-take-over-because-they-couldnt-care-less`
5. Bonnavion, P., Mickelsen, L.E., Fujita, A., De Lecea, L., Jackson, A.C.: Hubs and spokes of the lateral hypothalamus: cell types, circuits and behaviour. The Journal of physiology **594**(22), 6443–6462 (2016)
6. Boulant, J.A.: Neuronal basis of hammel's model for set-point thermoregulation. Journal of Applied Physiology **100**(4), 1347–1354 (2006)
7. Brown, M.T., Tan, K.R., O'Connor, E.C., Nikonenko, I., Muller, D., Lüscher, C.: Ventral tegmental area gaba projections pause accumbal cholinergic interneurons to enhance associative learning. Nature **492**(7429), 452–456 (2012)
8. Craig, A.D.: How Do You Feel? Princeton University Press (2015)
9. Humphries, M.D., Prescott, T.J.: The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Progress in Neurobiology **90**, 385–417 (2010). `https://doi.org/10.1016/j.pneurobio.2009.11.003`, `http://dx.doi.org/10.1016/j.pneurobio.2009.11.003`
10. Jennings, J.H., Rizzi, G., Stamatakis, A.M., Ung, R.L., Stuber, G.D.: The inhibitory circuit architecture of the lateral hypothalamus orchestrates feeding. Science **341**(6153), 1517–1521 (2013)
11. Jimenez-Rodriguez, A., Prescott, T.J., Schmidt, R., Wilson, S.: A framework for resolving motivational conflict via attractor dynamics. In: Biomimetic and Biohybrid Systems: 9th International Conference, Living Machines 2020, Freiburg, Germany, July 28–30, 2020, Proceedings 9. pp. 192–203. Springer (2020)
12. Keramati, M., Gutkin, B.: A reinforcement learning theory for homeostatic regulation. Advances in neural information processing systems **24** (2011)
13. Kremer, Y., Flakowski, J., Rohner, C., Lüscher, C.: Context-dependent multiplexing by individual vta dopamine neurons. Journal of Neuroscience **40**, 7489–7509 (9 2020). `https://doi.org/10.1523/JNEUROSCI.0502-20.2020`
14. McFarland, D., Bösser, T., Bosser, T.: Intelligent behavior in animals and robots. Mit Press (1993)

15. Mitchinson, B., Prescott, T.J.: Miro: a robot "mammal" with a biomimetic brain-based control system. In: Biomimetic and Biohybrid Systems: 5th International Conference, Living Machines 2016, Edinburgh, UK, July 19-22, 2016. Proceedings 5. pp. 179–191. Springer (2016)
16. Mohebi, A., Pettibone, J.R., Hamid, A.A., Wong, J.M.T., Vinson, L.T., Patriarchi, T., Tian, L., Kennedy, R.T., Berke, J.D.: Dissociable dopamine dynamics for learning and motivation. Nature **570**(7759), 65–70 (2019)
17. Nieh, E.H., Vander Weele, C.M., Matthews, G.A., Presbrey, K.N., Wichmann, R., Leppla, C.A., Izadmehr, E.M., Tye, K.M.: Inhibitory input from the lateral hypothalamus to the ventral tegmental area disinhibits dopamine neurons and promotes behavioral activation. Neuron **90**(6), 1286–1298 (2016)
18. Prescott, T.J., Redgrave, P., Gurney, K.: Layered control architectures in robots and vertebrates. Adaptive Behavior **7**(1), 99–127 (1999)
19. Redgrave, P., Prescott, T.J., Gurney, K.: The basal ganglia: A vertebrate solution to the selection problem? Neuroscience **89**, 1009–1023 (1999). https://doi.org/10.1016/S0306-4522(98)00319-4
20. Rosado, O.G., Amil, A.F., Freire, I.T., Verschure, P.F.: Drive competition underlies effective allostatic orchestration. Frontiers in Robotics and AI **9** (2022)
21. Salamone, J.D., Correa, M.: The mysterious motivational functions of mesolimbic dopamine. Neuron **76**(3), 470–485 (2012)
22. Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., Buzsáki, G.: Cocaine place conditioning strengthens location-specific hippocampal coupling to the nucleus accumbens. Neuron **98**(5), 926–934 (2018)
23. Sterling, P.: Allostasis: a model of predictive regulation. Physiology & behavior **106**(1), 5–15 (2012)
24. Stuber, G.D., Wise, R.A.: Lateral hypothalamic circuits for feeding and reward (1 2016). https://doi.org/10.1038/nn.4220
25. Swanson, L.W.: Cerebral hemisphere regulation of motivated behavior. Brain research **886**(1-2), 113–164 (2000)
26. Swanson, L.W.: Brain Architecture: Understanding the Basic Plan. Oxford University Press (2003)
27. Wert-Carvajal, C., Reneaux, M., Tchumatchenko, T., Clopath, C.: Dopamine and serotonin interplay for valence-based spatial learning. Cell Reports **39**(2), 110645 (2022)
28. Whelan, M.T., Jimenez-Rodriguez, A., Prescott, T.J., Vasilaki, E.: A robotic model of hippocampal reverse replay for reinforcement learning. Bioinspiration & Biomimetics **18**(1), 015007 (2022)
29. Wilson, S.P., Prescott, T.J.: Scaffolding layered control architectures through constraint closure: insights into brain evolution and development. Philosophical Transactions of the Royal Society B **377**(1844), 20200519 (2022)