Tonic dopamine integration in a bio-inspired robot controller

Cristóbal Nettle[1], María-Jose Escobar[1] and Arthur Leblois[2].
[1] Advanced Center for Electrical and Electronic Engineering, Valparaíso, Chile
[2] CNRS, Laboratoire de neurophysique et physiologie, UMR 8119, Université Paris Descartes, Paris, France.

Introduction

From neuroscience

- Dopamine (DA) has been related to the exploration-exploitation trade-off in basal-ganglia related decision-making processes.
- Striatal dopamine receptors are reduced in patients with anxiety disorders.

About artificial intelligent agents

- Exploration-exploitation trade-off is fundamental.
- The integration of emotions might be a great benefit in:
  - Learning processes.
  - Interactions with humans and the environment.

Proposal

- Integration of tonic DA effects in a biologically based decision-making mechanism.
- Integrate the model in a robot controller to:
  - Set its exploration-exploitation trade-off.
  - Indirectly integrate an emotion related effect.

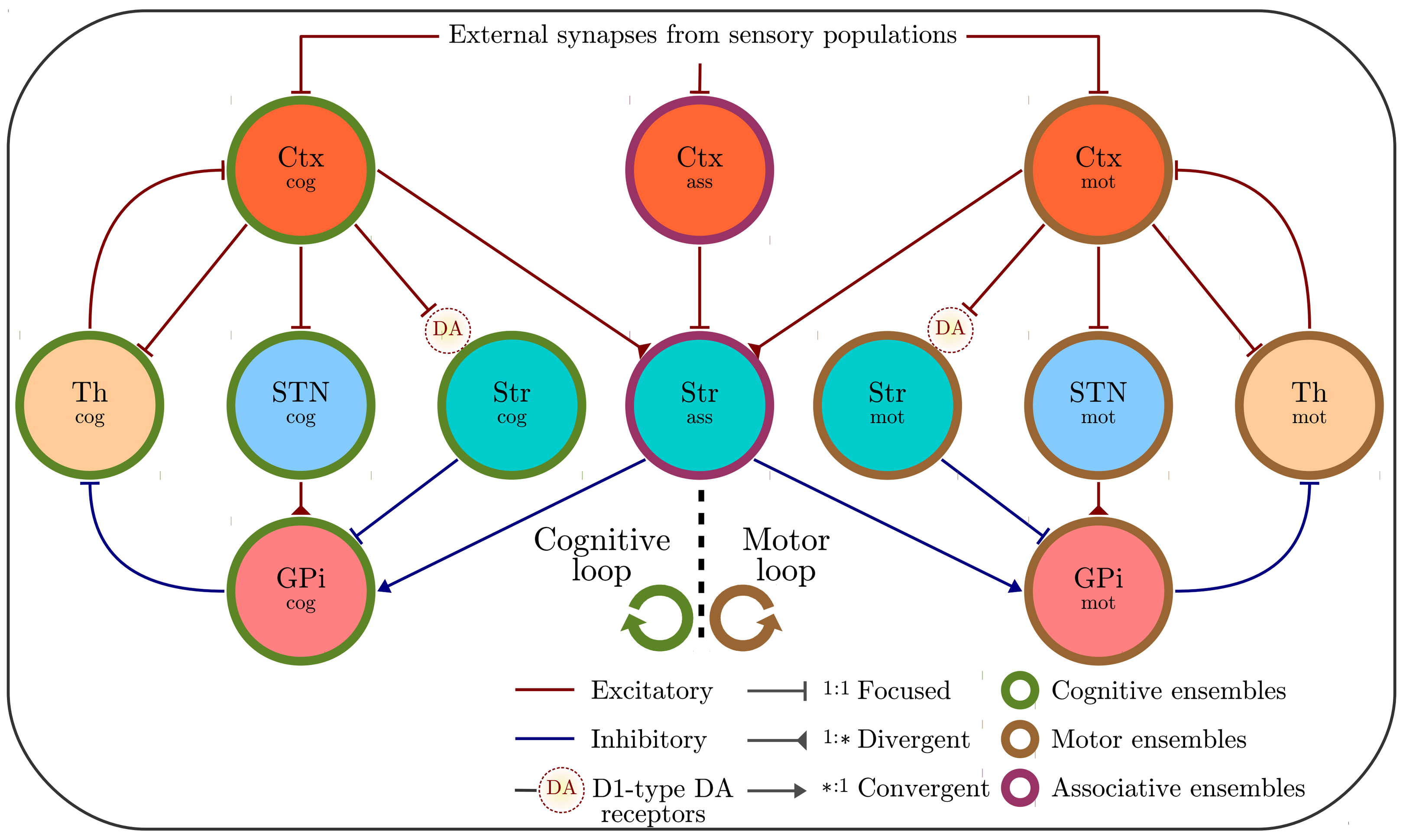
Cortico-Basal Ganglia (CBG) model

- **Basal ganglia loops model [1]**
- Multiple parallel loops implementation:
  - Two symmetrical loops → two selections.
  - A loop is composed by two competing pathways:
    - Direct pathway: focussed positive feedback through the striatum.
    - Hyper-direct pathway: spread negative feedback through the subthalamic nucleus.
  - There is a crosstalk between loops.
- Selection process follows a winner-take-all behavior.
- Reward based learning in corticostriatal synapses.
  ⇒ Learning leads to the selection of the better option.

- **Tonic DA integration [2]**

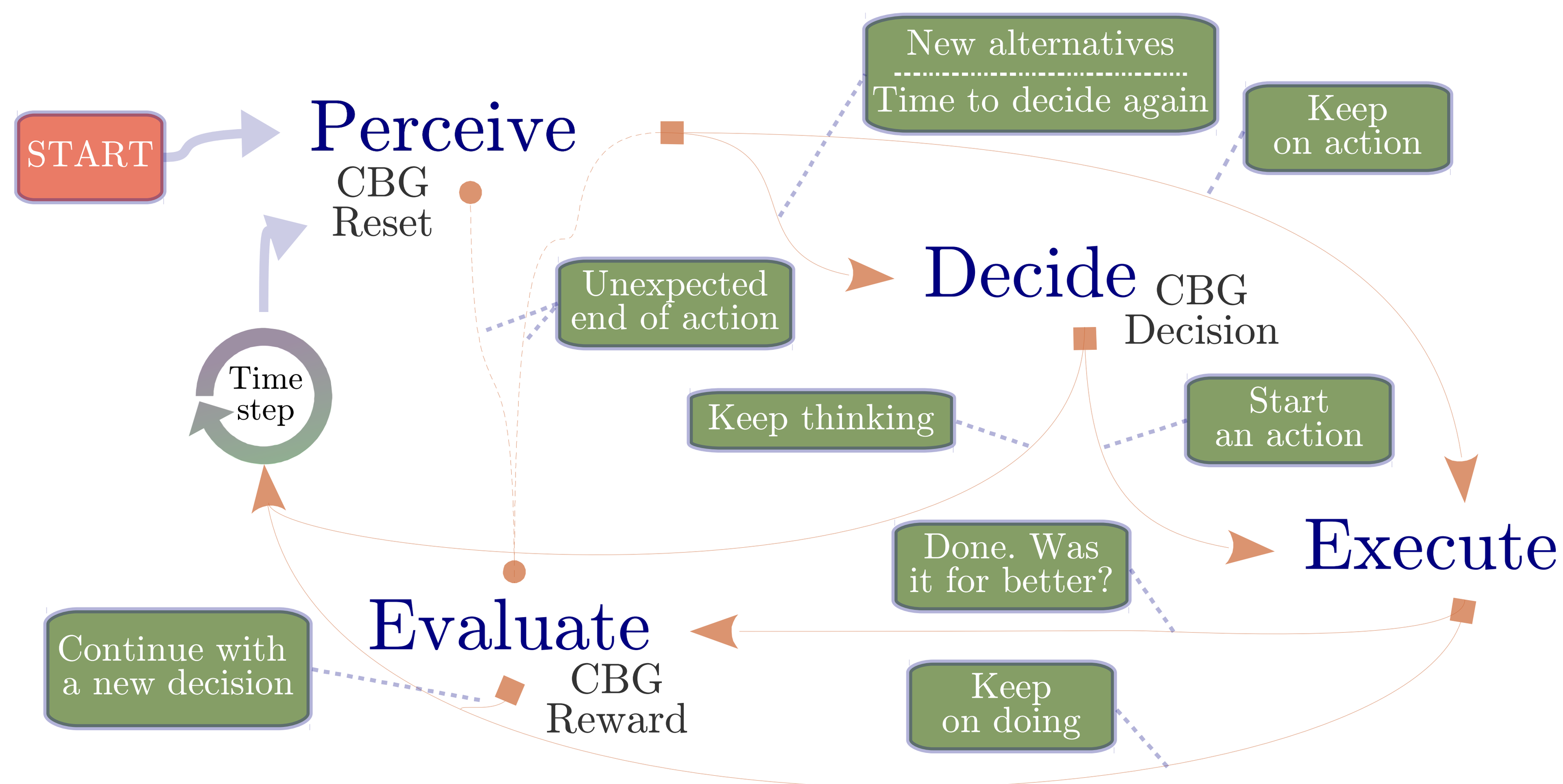
Tonic dopamine D1-type has a strong effect over the exploration-exploitation trade-off.

- Affects cognitive and motor corticostriatal synapses.
- Simulates D1-type as a multiplicative factor:
  $I^T(t) = (1 + DA)\ I(t)$



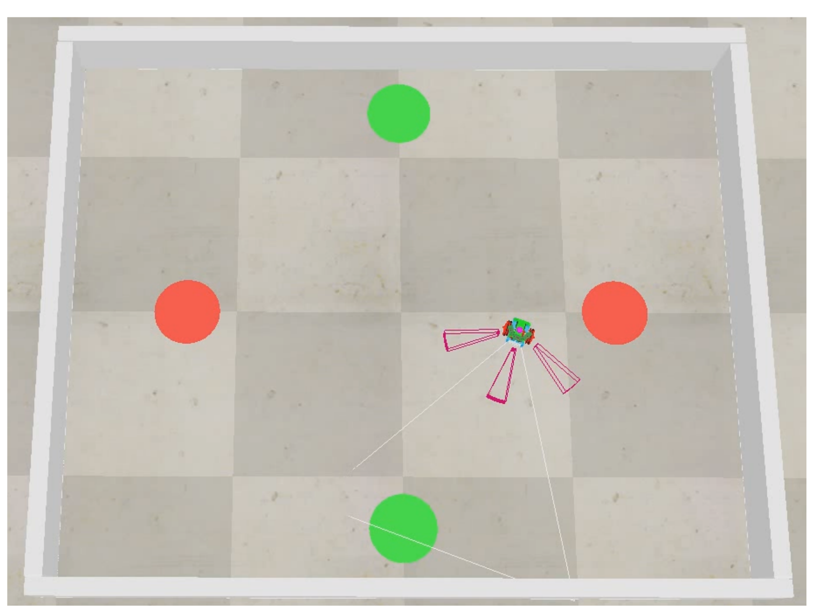
Implementation of the CBG model as a robot controller

- The model is integrated in a robot controller implemented as a finite state machine.
  1. Perceive: detects the environment and then, the robot's possible actions.
  2. Decide: performs a selection using the CBG model.
  3. Execute: controls the robot movements.
  4. Evaluate: performs the reward evaluation and learns.
- The robot learns on-line which option is better for its current internal state.
- The system is tested using the Virtual Robot Experimentation Platform (V-REP) simulator.



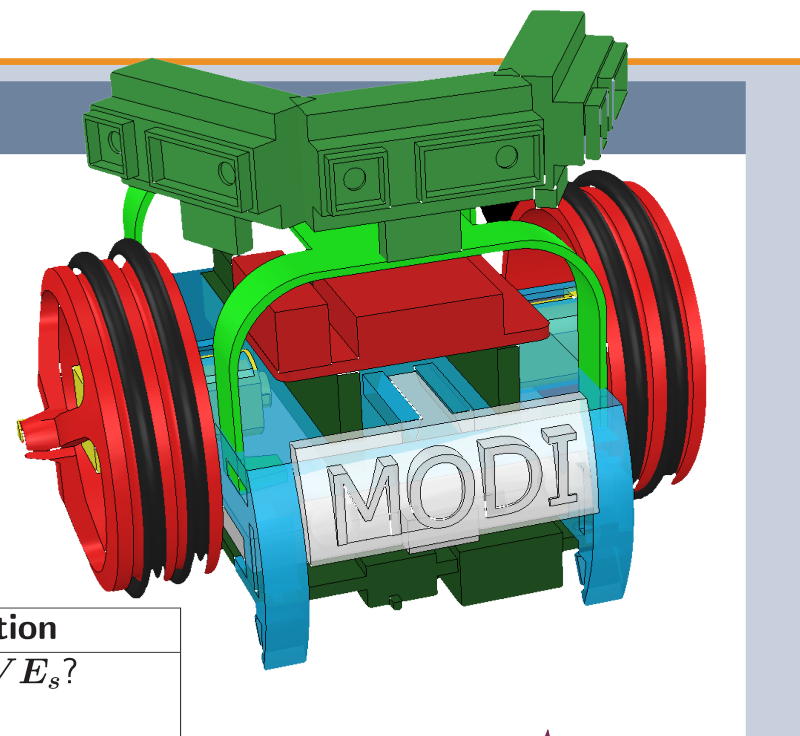
Two-resource survival task

- Minimal scenario for evaluating decision-making mechanisms.
- The agent has two intrinsic energy levels:
  - Potential energy PE: a food like energy level.
  - Vital energy VE: decreases in time leading to death. Its acquisition requires potential energy.
- The robot can reload each energy type being placed on a respective energy source.
  - There are two of energy sources for each type.



Virtual scenario

Agent: the MODI robot



I just want to live for a while!

- MODI (MODular Intelligent) is a compact open-hardware sensorless robot.
- Proximity sensors where attached.
- Virtual sensing of the energy sources is considered.
  - The robot instantaneously knows the position of any energy source inside a range of vision.
- Reward conditions prioritized from top to bottom

Seeking behaviour	Activation	Reward condition
PE seeker	PE ≤ 0.2	Is the robot closer to VE_s?
VE seeker	VE ≤ 0.5	Is the robot closer to PE_s?
Both	Otherwise	Is the robot closer to either VE_s or PE_s?

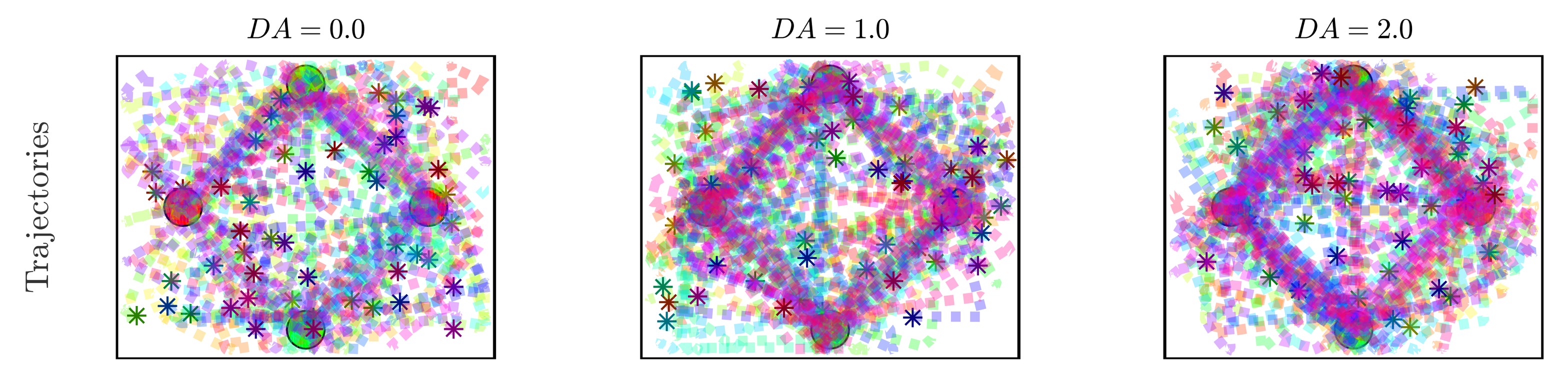
- Agent capabilities:

Motor actions	
forward	Move forward.
turn_left	Turn left on its own axis.
turn_right	Turn right on its own axis.
stop	Stops any movement reducing its VE consumption ratio to the half.
reload	Reload an energy type if the conditions are satisfied.

Cognitive alternatives	
Wander	The agent randomly selects and executes a movement between *forward*, *turn_left* and *turn_right*.
Rest	Reduce vital energy consumption: executes the *stop* motor movement.
$Wall_{av}$	Avoid collisions with walls. Depending on where is the wall placed, the agent turns or moves away.
$Reload_{VE}$	Has the goal to increase its VE level. The agent moves closer to a VE_s turning or moving forward, or, if the agent is close enough, reloads.
$Reload_{PE}$	Similarly as $Reload_{VE}$, the agent acts in order to increase its PE level.

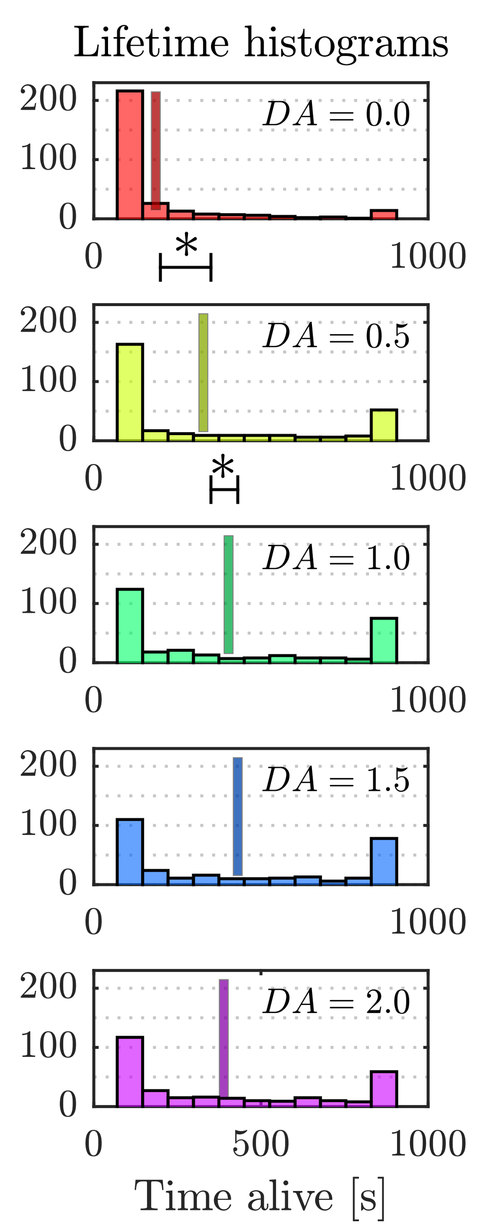
Tonic DA level effects in the behavior of the robotics agent

Trajectories



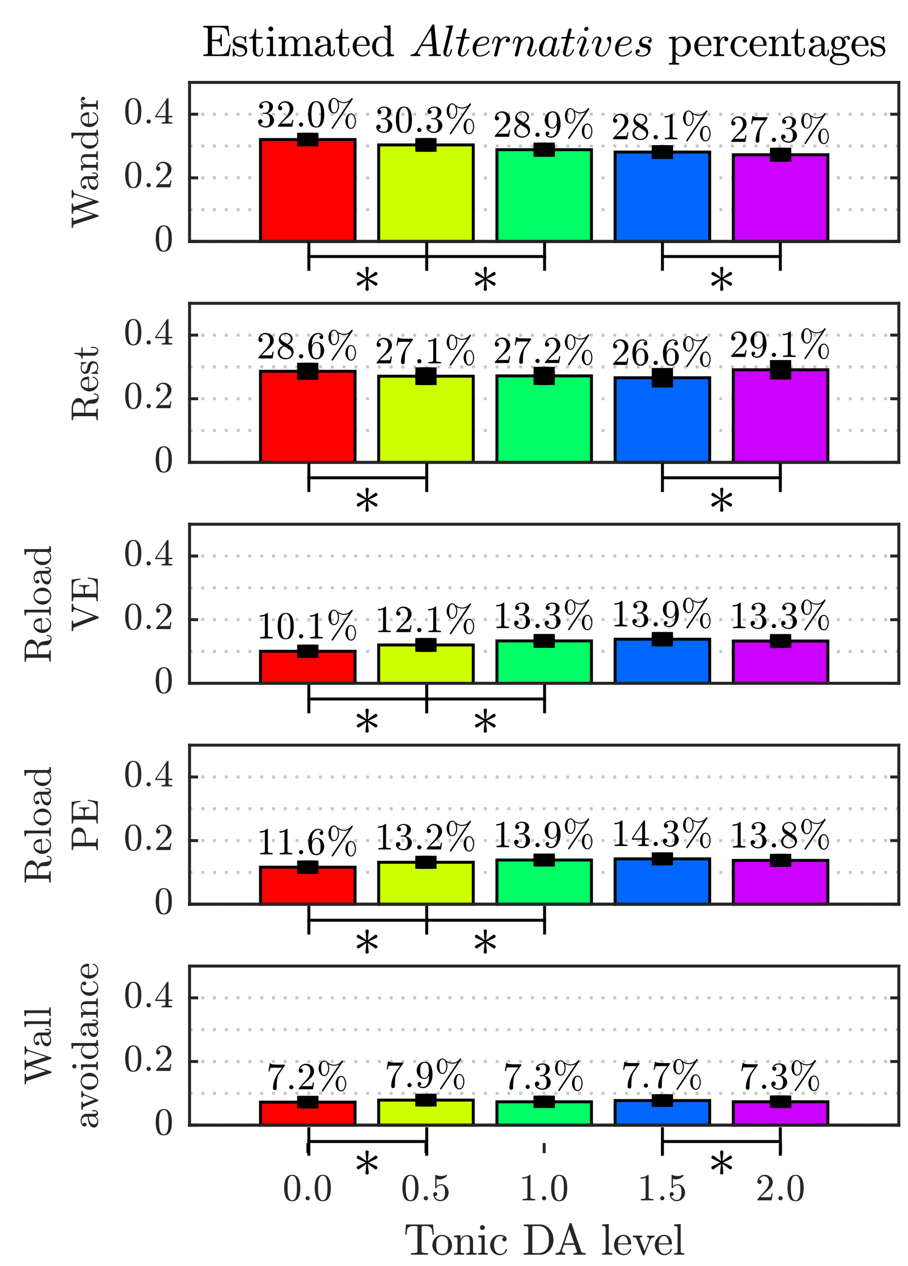
An **increase in the tonic DA level** produces an increase in the trajectories between sources, showing **more exploitative related behaviors**.

Effects in the life expectancy



Vertical lines show the mean value, representing the life expectancy of the robot for each level of tonic DA. **Life expectancy is modulated by the tonic DA level.**

Distribution of the selected actions



Mean time spent performing each alternative. The height of the black rectangles above each bar corresponds to two times the variance. The remaining percentages correspond to the time spent during the *deciding* process. **Higher tonic DA levels turns out into more energy reloads and less exploration of the environment.**

*: statistically significant shifts following the Mann-Whitney test ($p < 0.05$).

Conclusions

- The controller implementation shows the feasibility of using the proposed CBG model as a decision making mechanism in an artificial intelligent agent.
- The system is able to learn on-line its best option, given its current energy levels.
- Tonic DA controls the agent's exploration-exploitation trade-off. Direct related to its tonic DA level, the agent
  - modifies the probability of selecting the better option.
  - modifies its surviving skills.

References

M. Guthrie, A. Leblois, A. Garenne, and T. Boraud, "Interaction between cognitive and motor cortico-basal ganglia loops during decision making: a computational study," *Journal of neurophysiology*, vol. 109, no. 12, pp. 3025–3040, 2013.

M. D. Humphries, M. Khamassi, and K. Gurney, "Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia," *Frontiers in neuroscience*, vol. 6, 2012.

cristobal.nettle@gmail.com