Category-Based Intrinsic Motivation

Lisa Meeden
Rachel Lee
Ryan Walker
Swarthmore College, USA

James Marshall
Sarah Lawrence College, USA

9th International Conference on Epigenetic Robotics
Venice, Italy
November 12-14, 2009

Research goals

- Design a robot control architecture that implements an ongoing, autonomous developmental learning process

- Test it on a physical robot

- Essential components
  - Categorization
  - Prediction
  - Intrinsic motivation

- “**Categorization** is of such fundamental importance for cognition and intelligent behavior that a natural organism incapable of forming categories does not have much chance of survival”

  —Lungarella et al., Developmental robotics: A survey, *Connection Science*, 2003

- “The ability to make **predictions** is part of the core initial knowledge on top of which human cognition is built”

  —Spelke, Core knowledge, *American Psychologist*, 2000

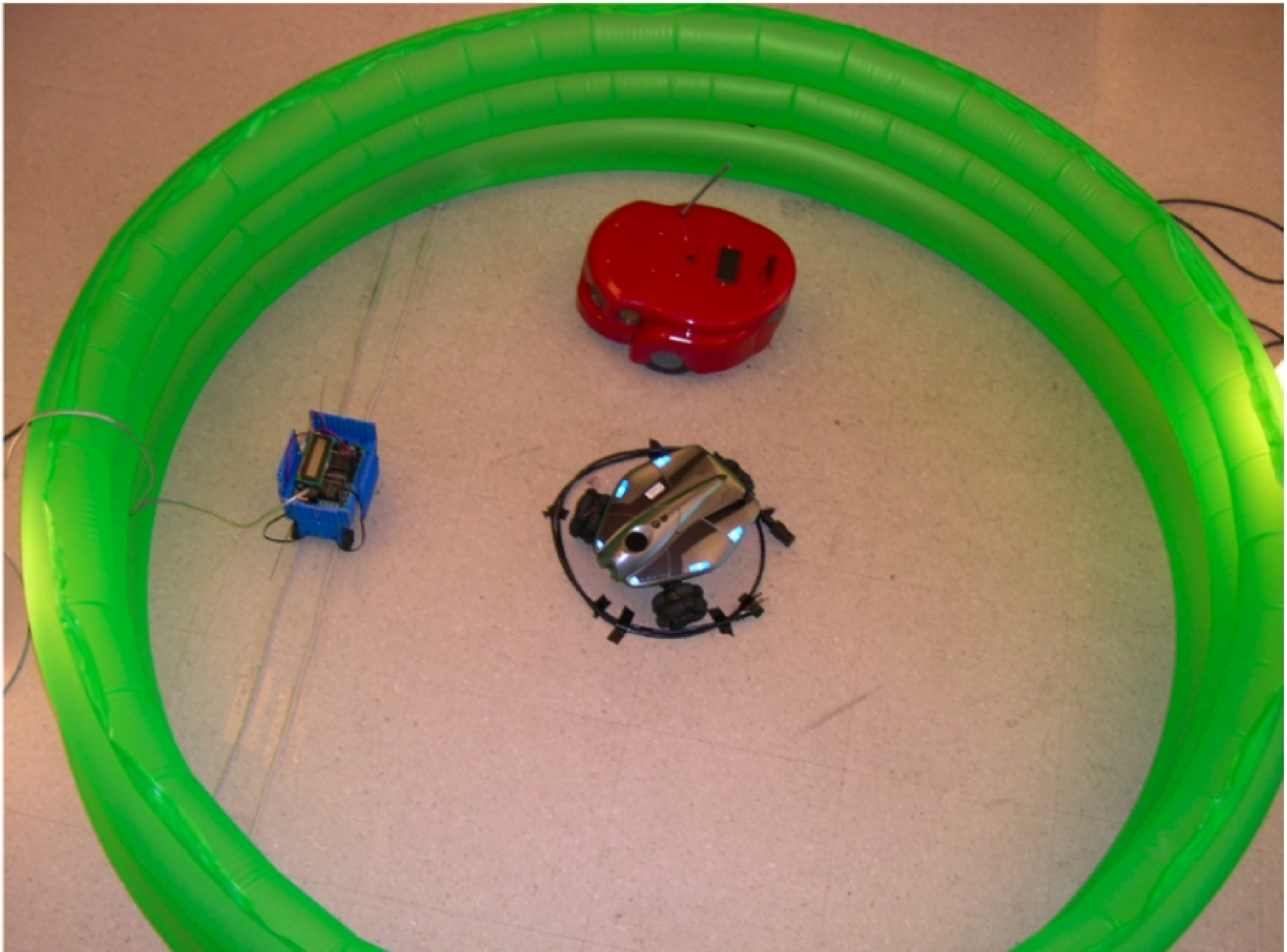
- “**Intrinsic motivation** is the inherent tendency to seek out novelty and challenges, to extend and exercise one's capacities, to explore, and to learn”

  —Ryan and Deci, *American Psychologist*, 2000

Implementation

- Categorization
  - Growing Neural Gas (Fritzke, 1995)
- Prediction
  - Artificial neural networks
- Intrinsic motivation
  - Intelligent Adaptive Curiosity (Oudeyer *et al.*, 2007)
- Physical robot
  - Rovio

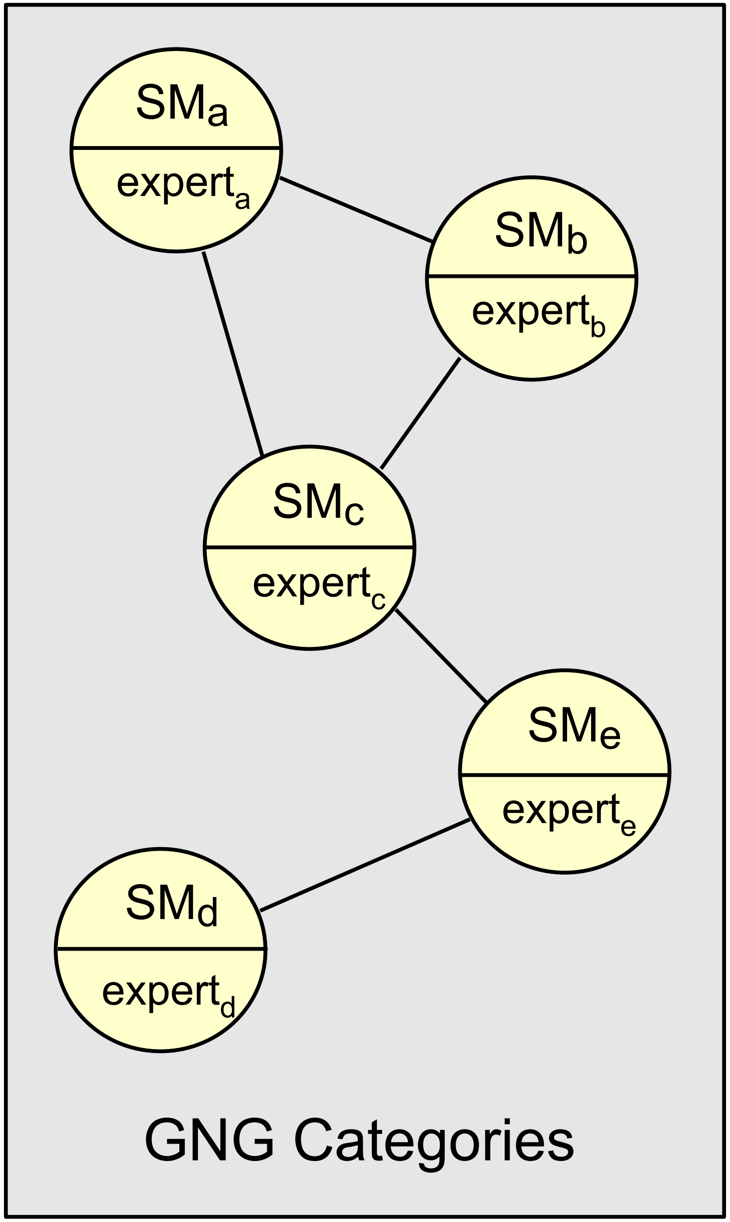
Environment



Overview of experiment

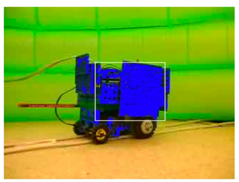
- **Robot observes** its surroundings visually and decides on its own what actions to perform
- **Categorizes** its sensory input based on similiarity to previous inputs
- **Predicts** what it will see on the next time step as a result of performing a particular action
- **Learns** by comparing its prediction to what it actually observes
- **Intrinsically motivated** to choose actions that maximize learning progress

Perception-action loop

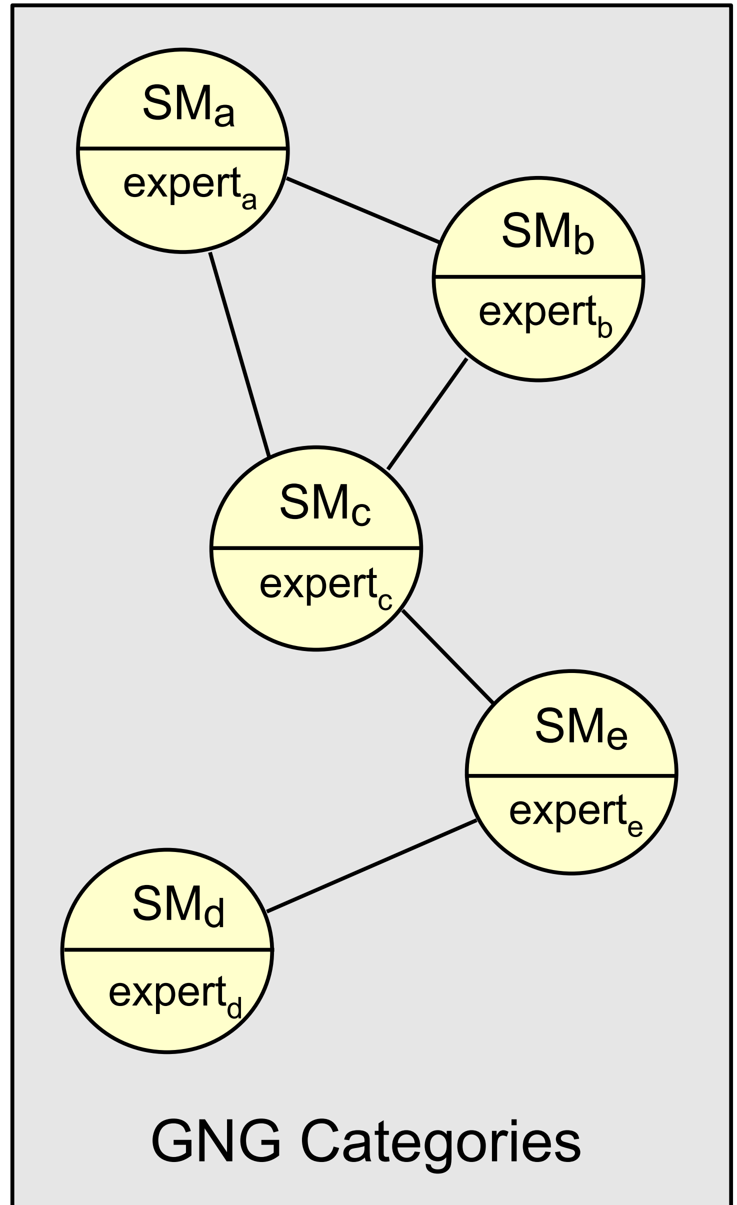


Perceive current sensory input S(t)

S(t)

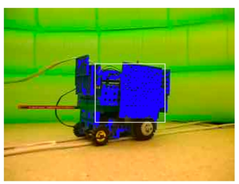


Camera image



Consider possible motor actions

S(t)



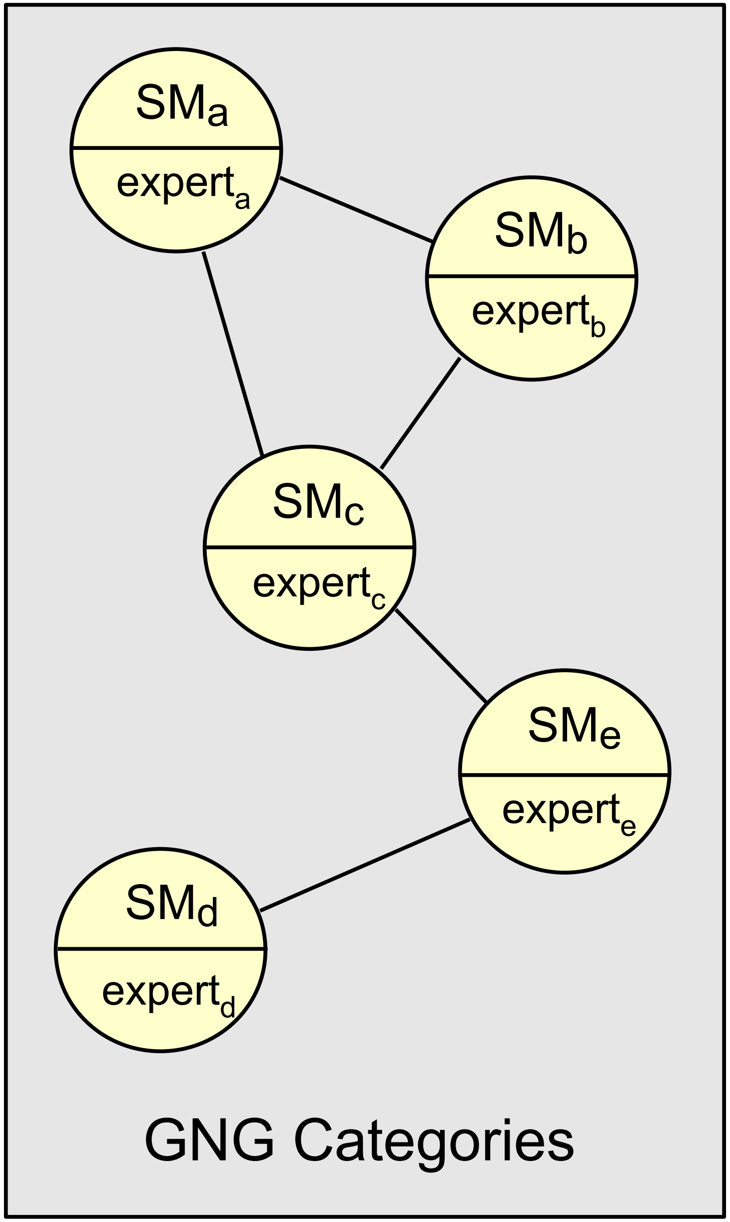
Camera image

$M_1(t)$

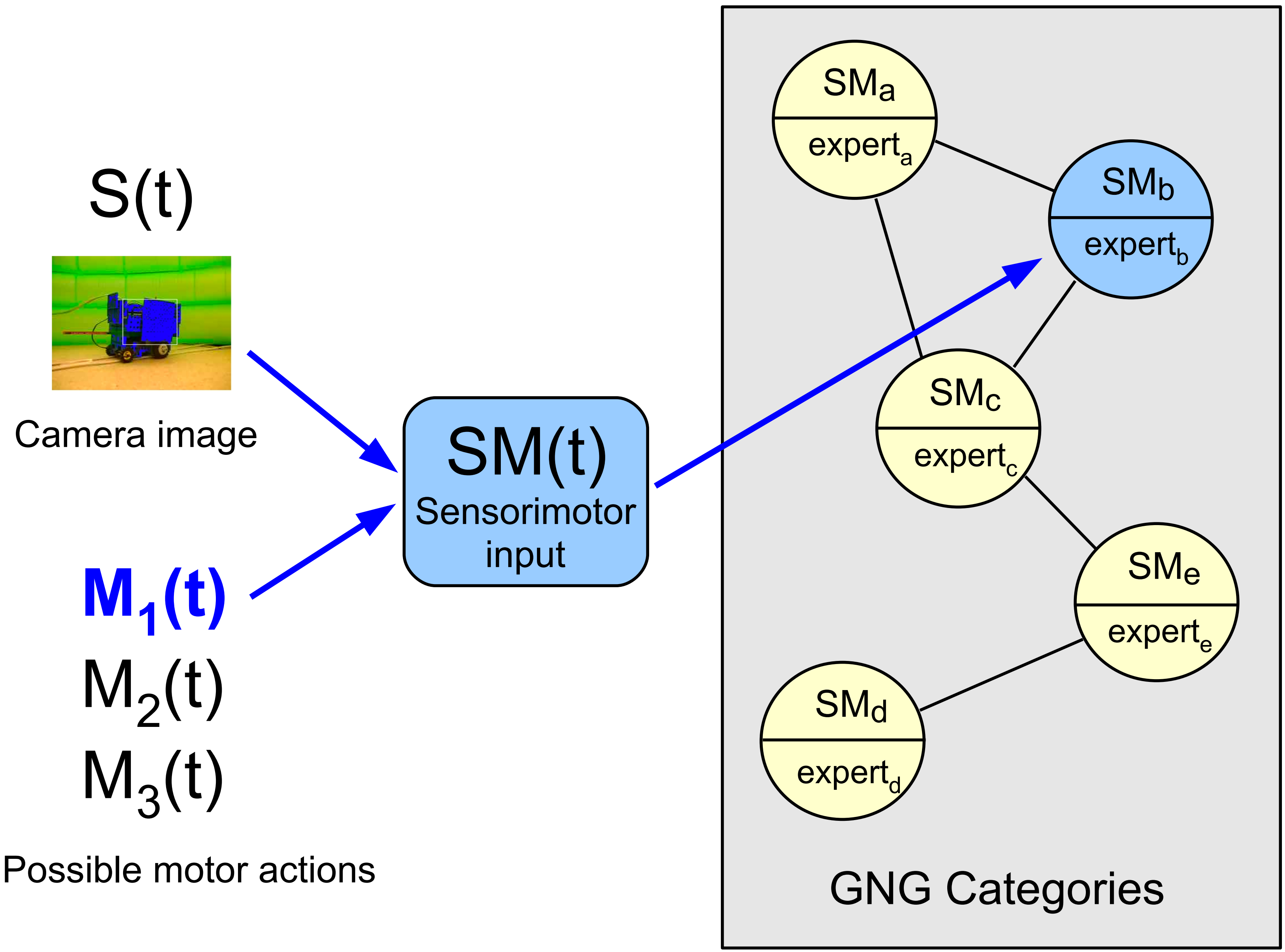
$M_2(t)$

$M_3(t)$

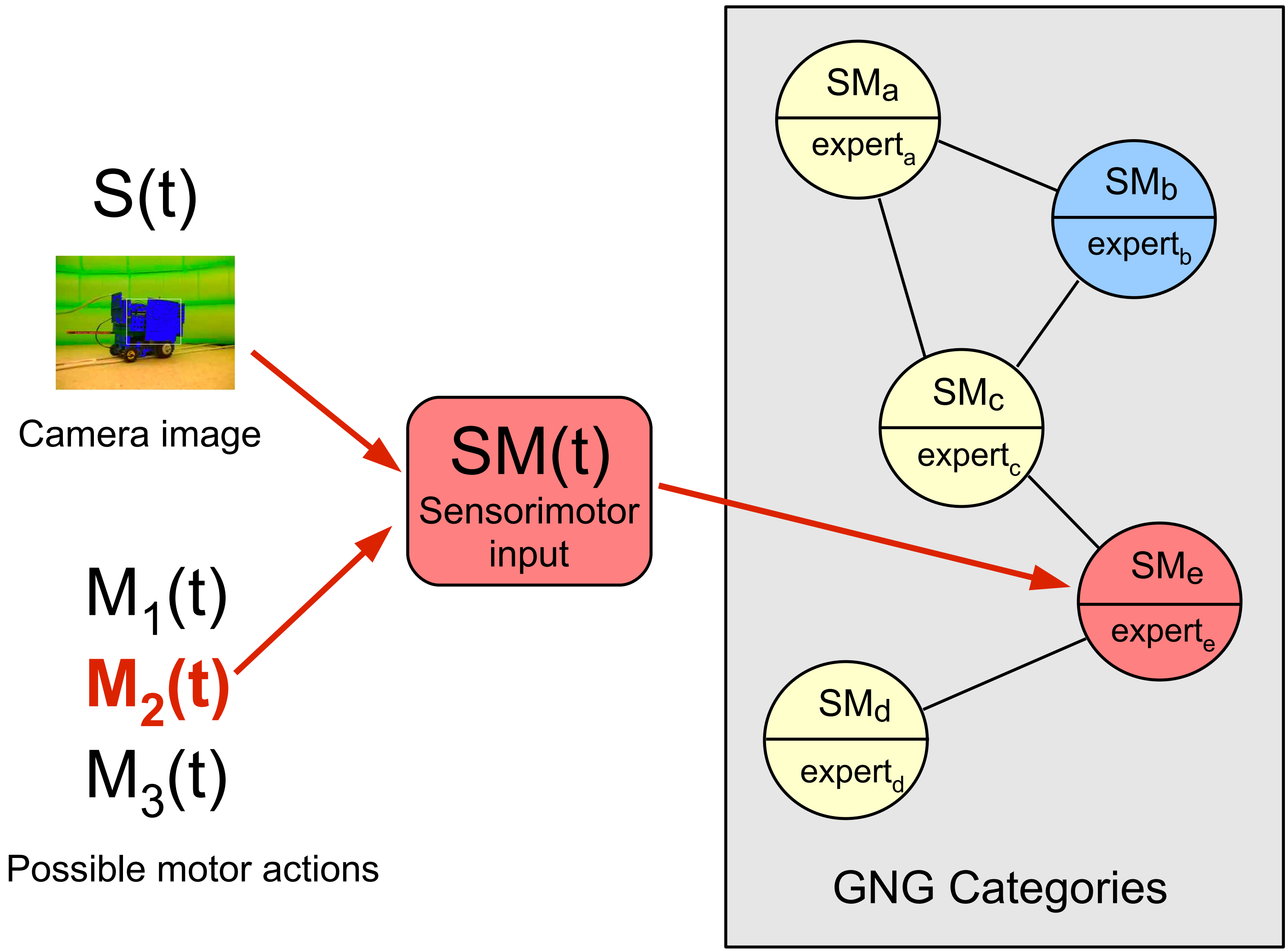
Possible motor actions



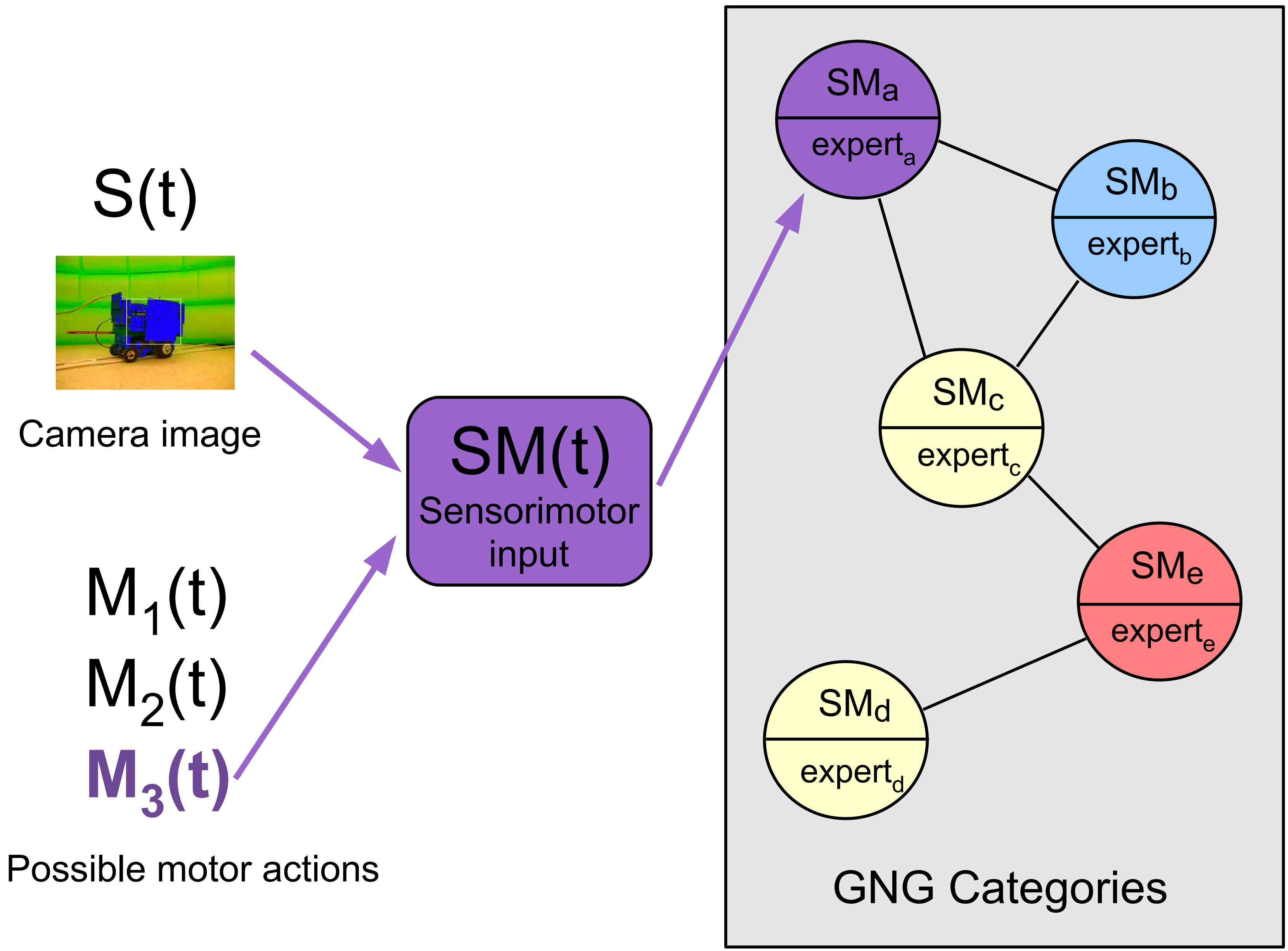
Find best matching categories



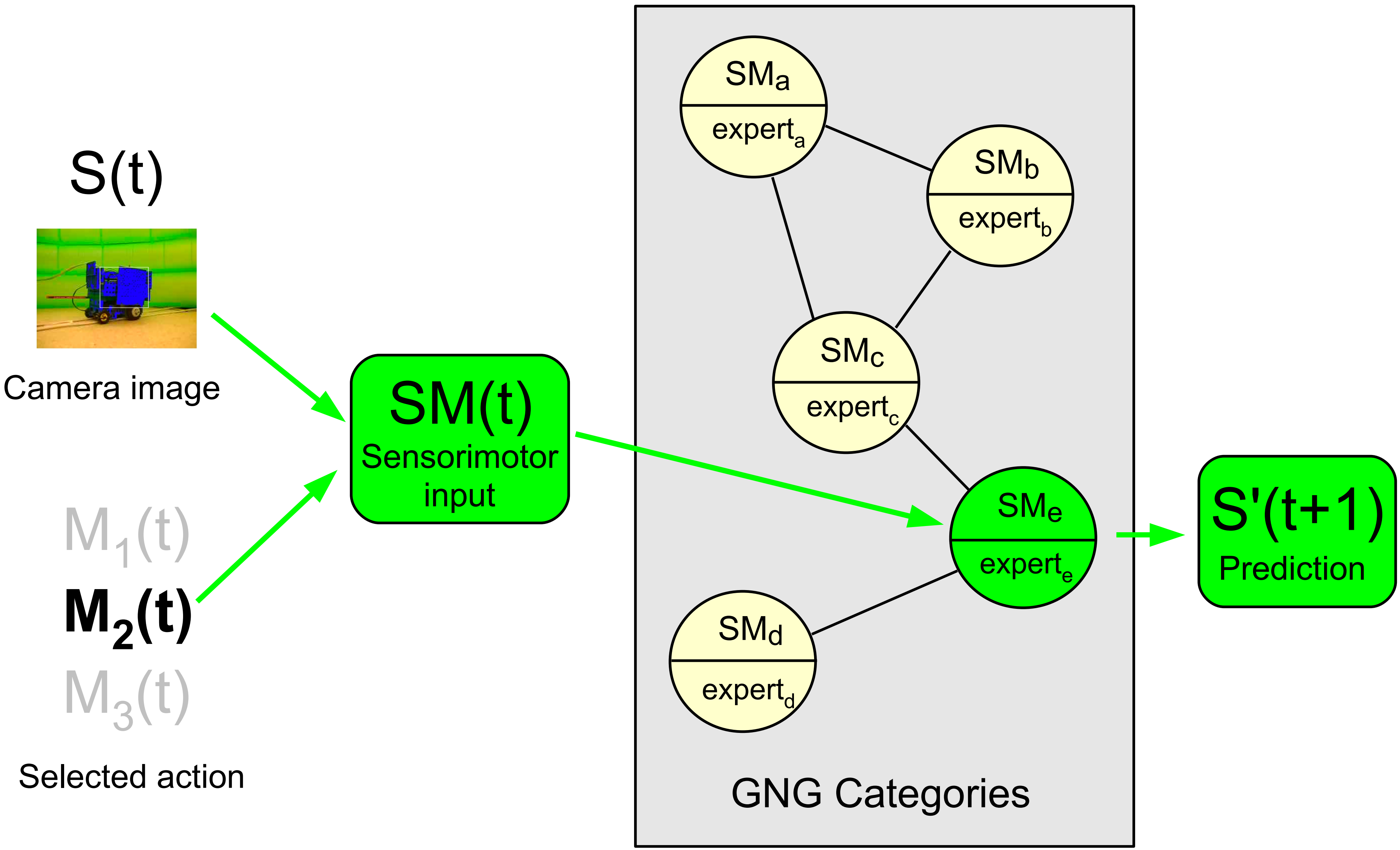
Find best matching categories



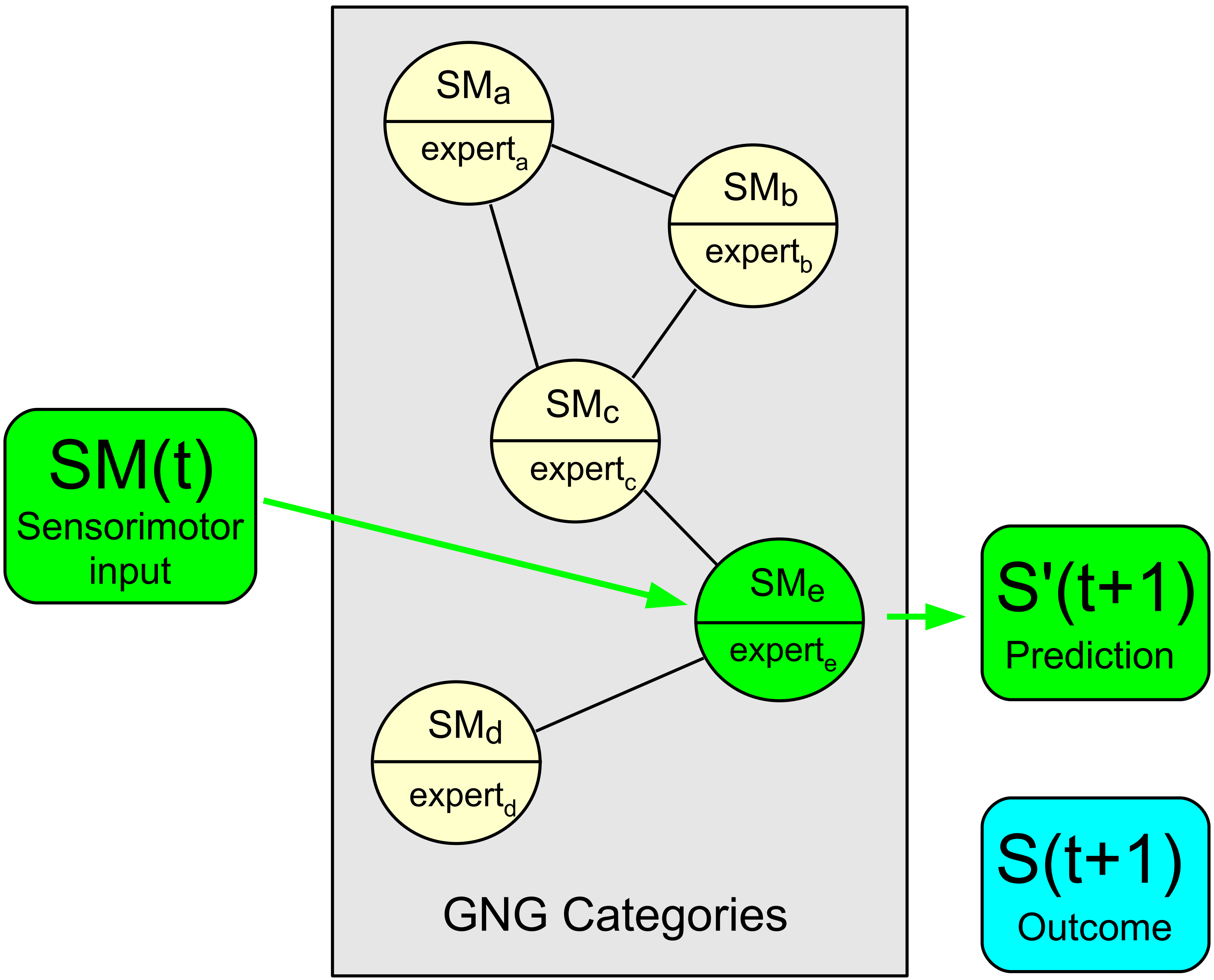
Find best matching categories



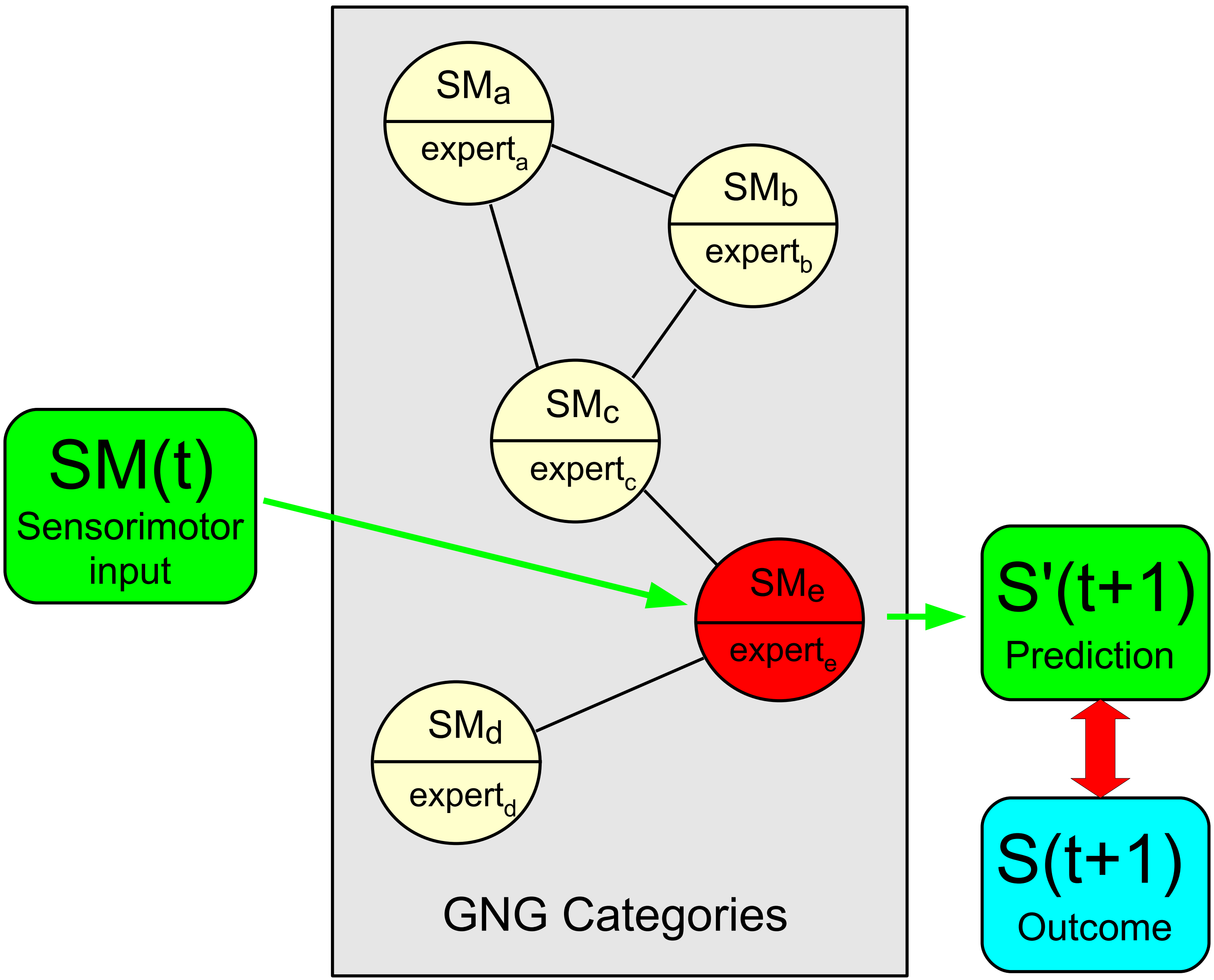
Select expert with maximal learning progress



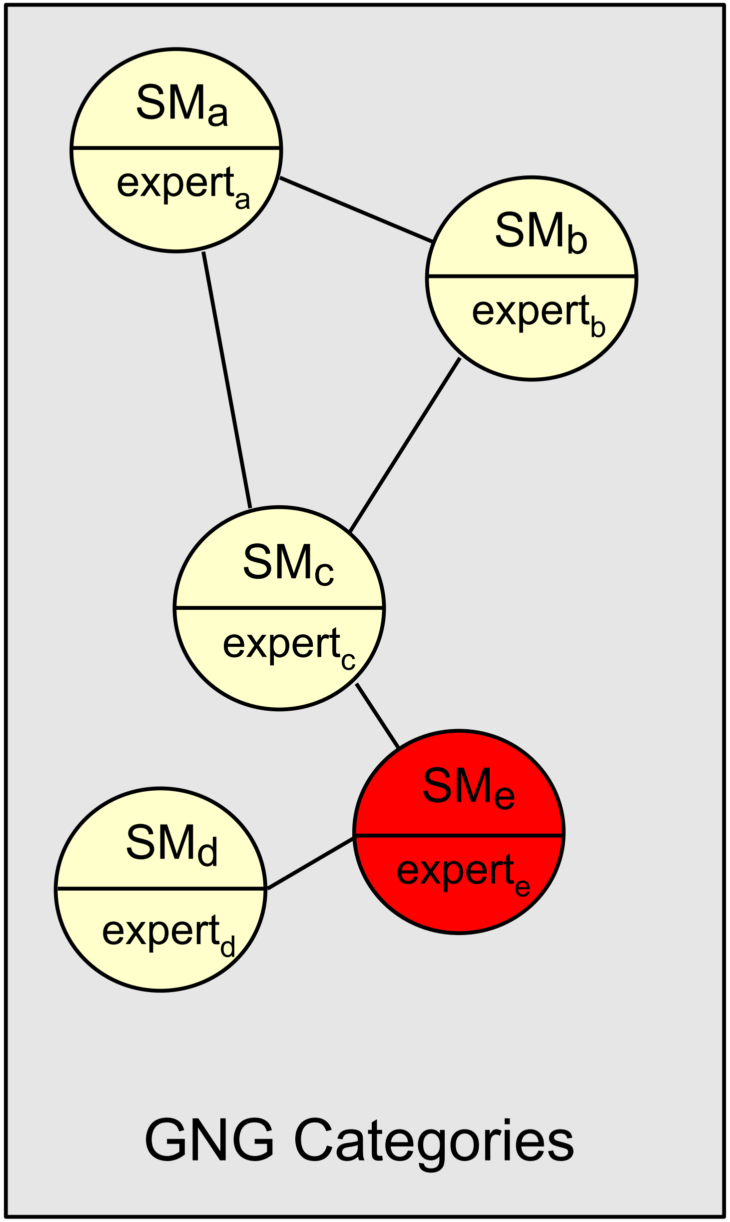
Perform selected action and observe outcome



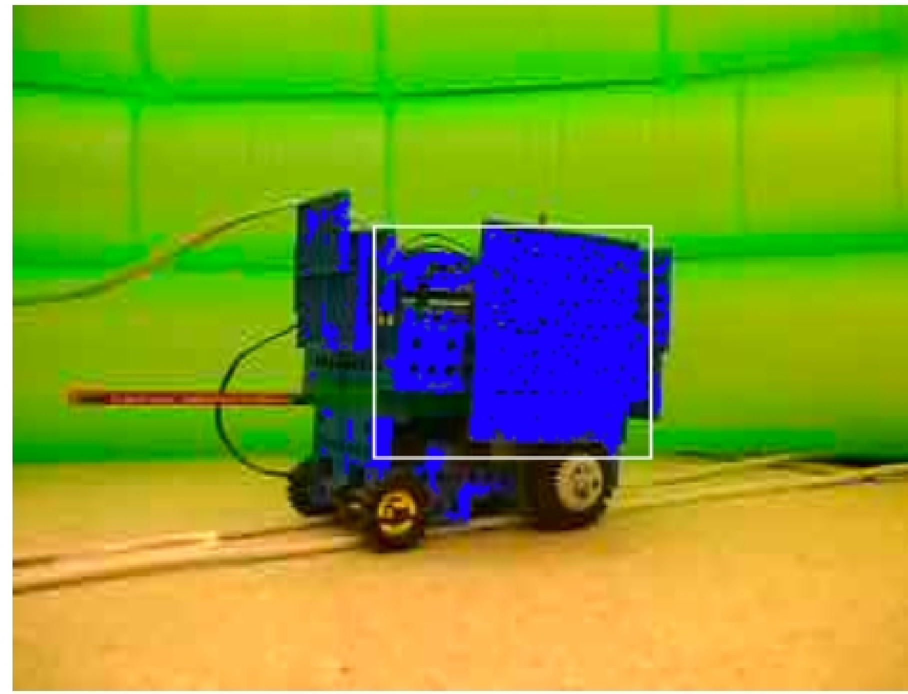
Update expert based on prediction error



Adjust GNG categories

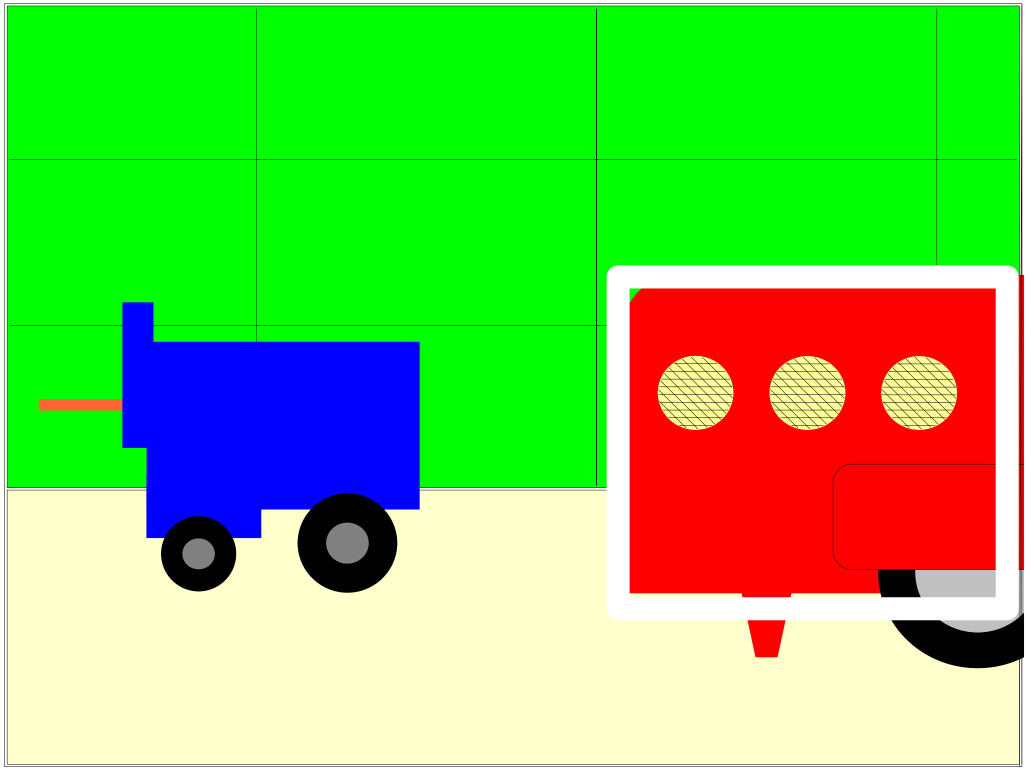


Perceptions



- Camera images
- Red, green, and/or blue can be detected
- Robot **chooses** which color to focus on
- Sensory vector S(t) = (*red*, *green*, *blue*, *area*, *position*)
- Example: (0, 1, 1, 0.12, 0.5)

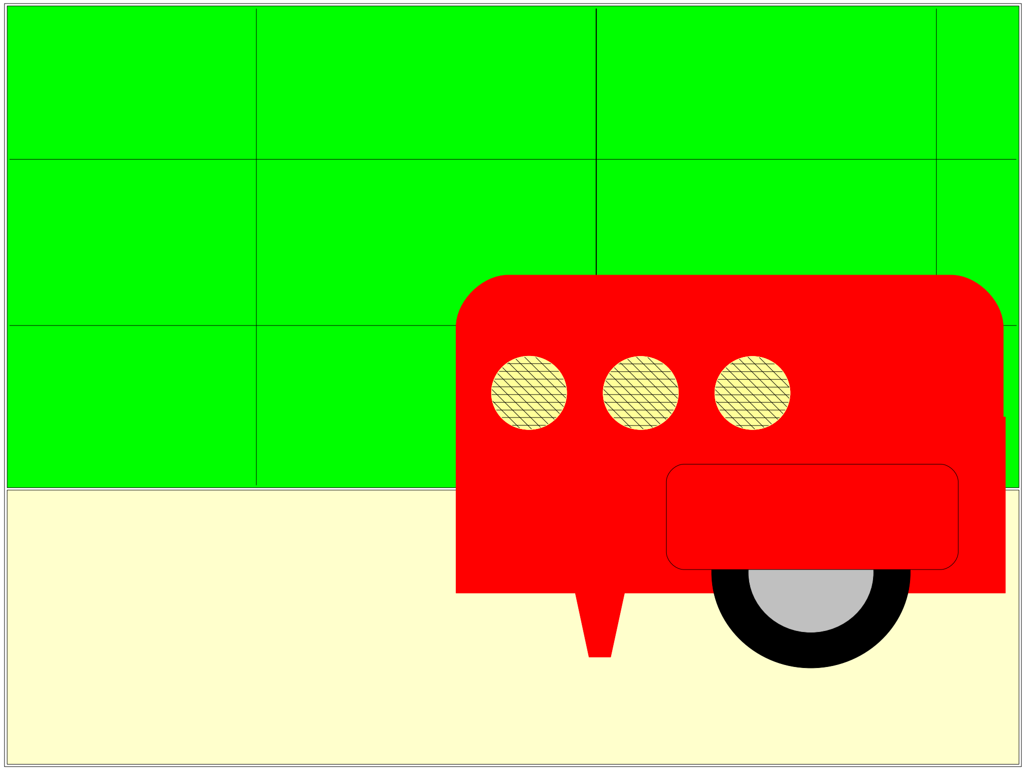
Example: Robot chooses to look at red



S(t) = (*red*, *green*, *blue*, *area*, *position*)

= (1, 1, 1, 0.23, 1.0)

Example: Robot chooses to look at blue



S(t) = (*red*, *green*, *blue*, *area*, *position*)

= (1, 1, 0, 0, 0)

Motor actions

- Which color to focus on
- How much to rotate
- Motor vector M(t) = (*colorFocus*, *rotation*)
  - *colorFocus* $\in$ [0...1] $\Rightarrow$ [Red...Green...Blue]
  - *rotation* $\in$ [0...1] $\Rightarrow$ [Left...Right]

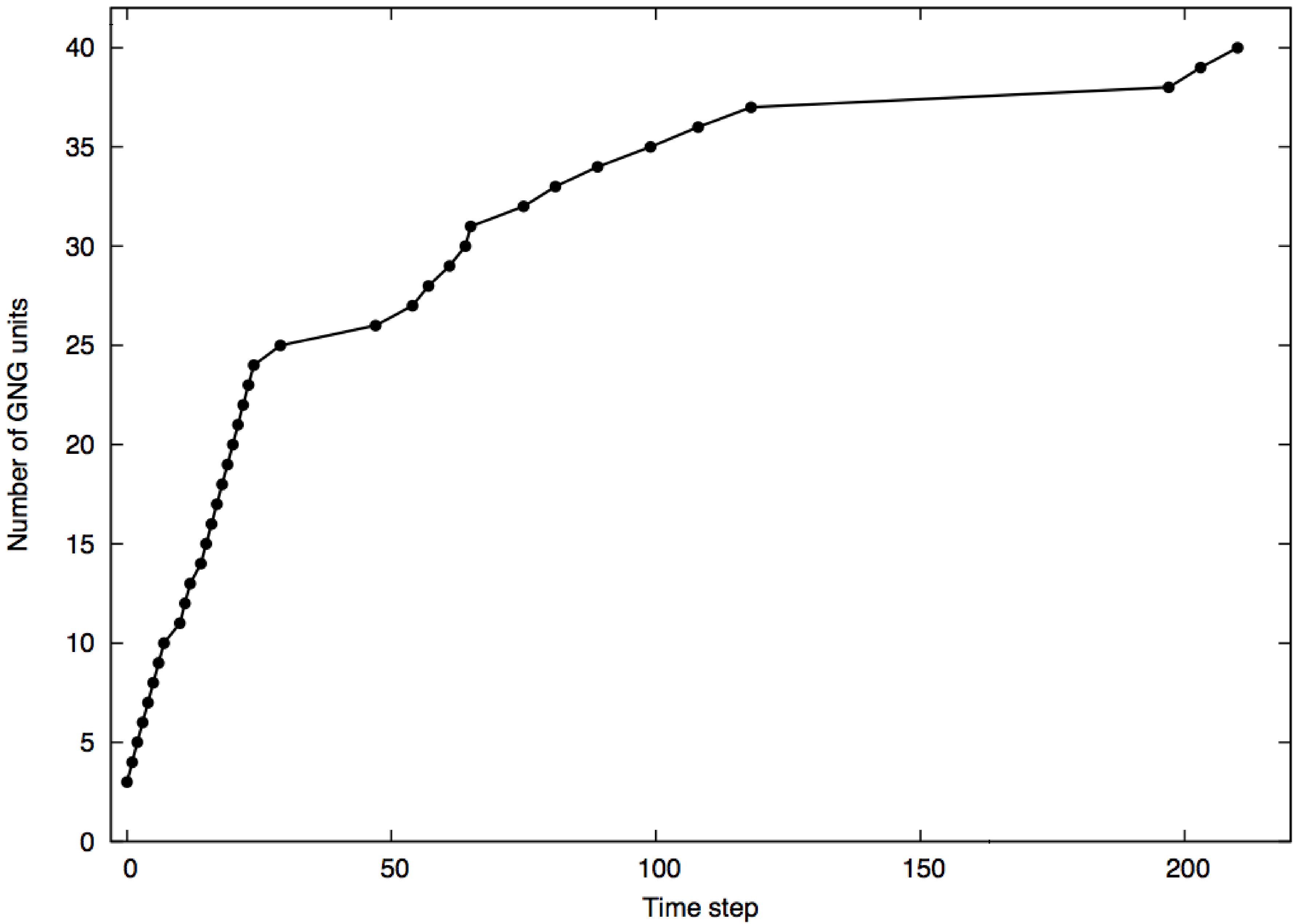
- Example: (0.8, 0.2) = focus on blue, turn left

- Sensorimotor vectors SM(t) are **7-dimensional**

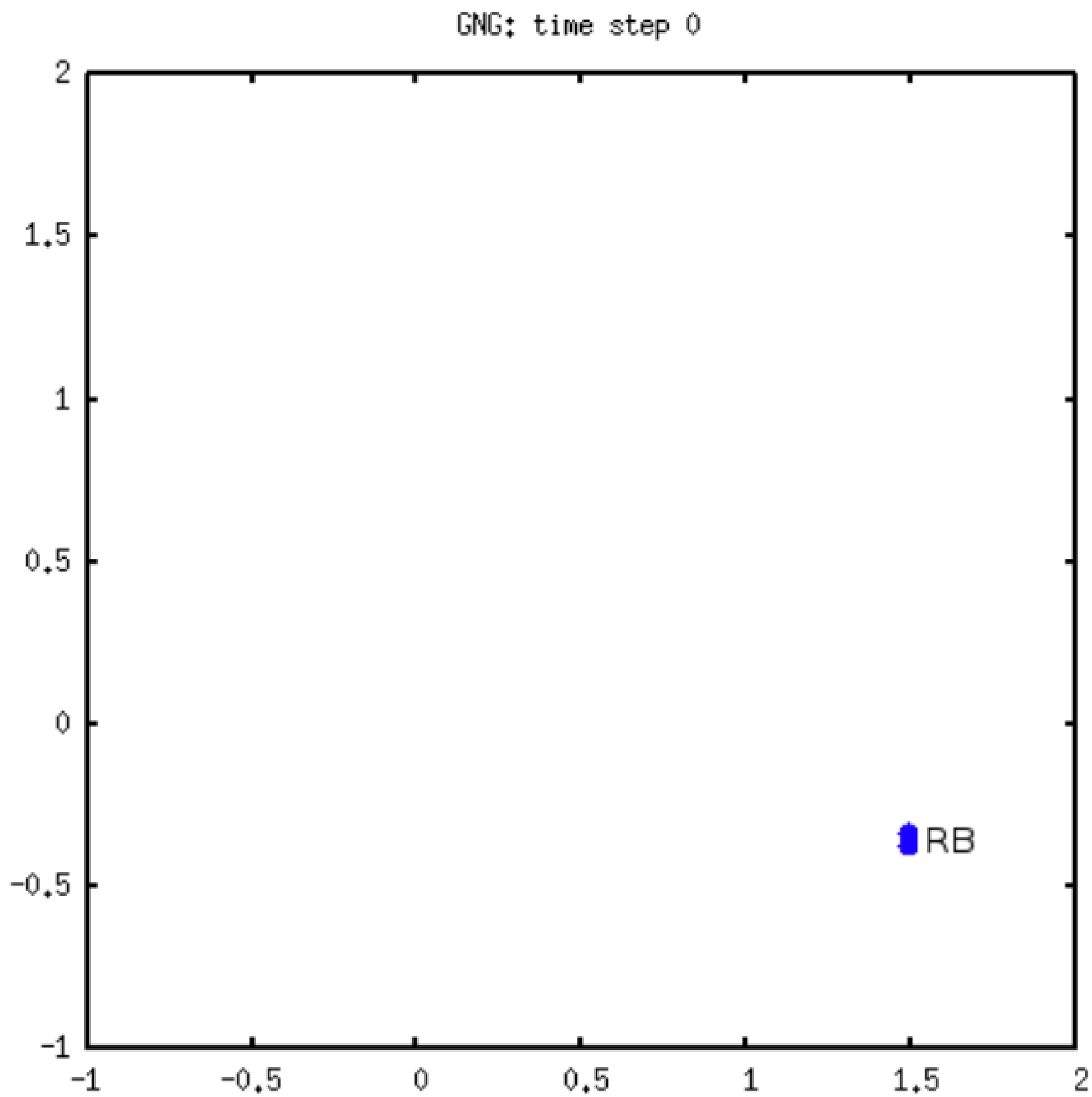
Learning opportunities

- **Green walls** offer a constant background, which is easy to predict
- **Red static robot** is also predictable, but is only visible from a few positions
- **Blue moving robot** is smaller and harder to predict

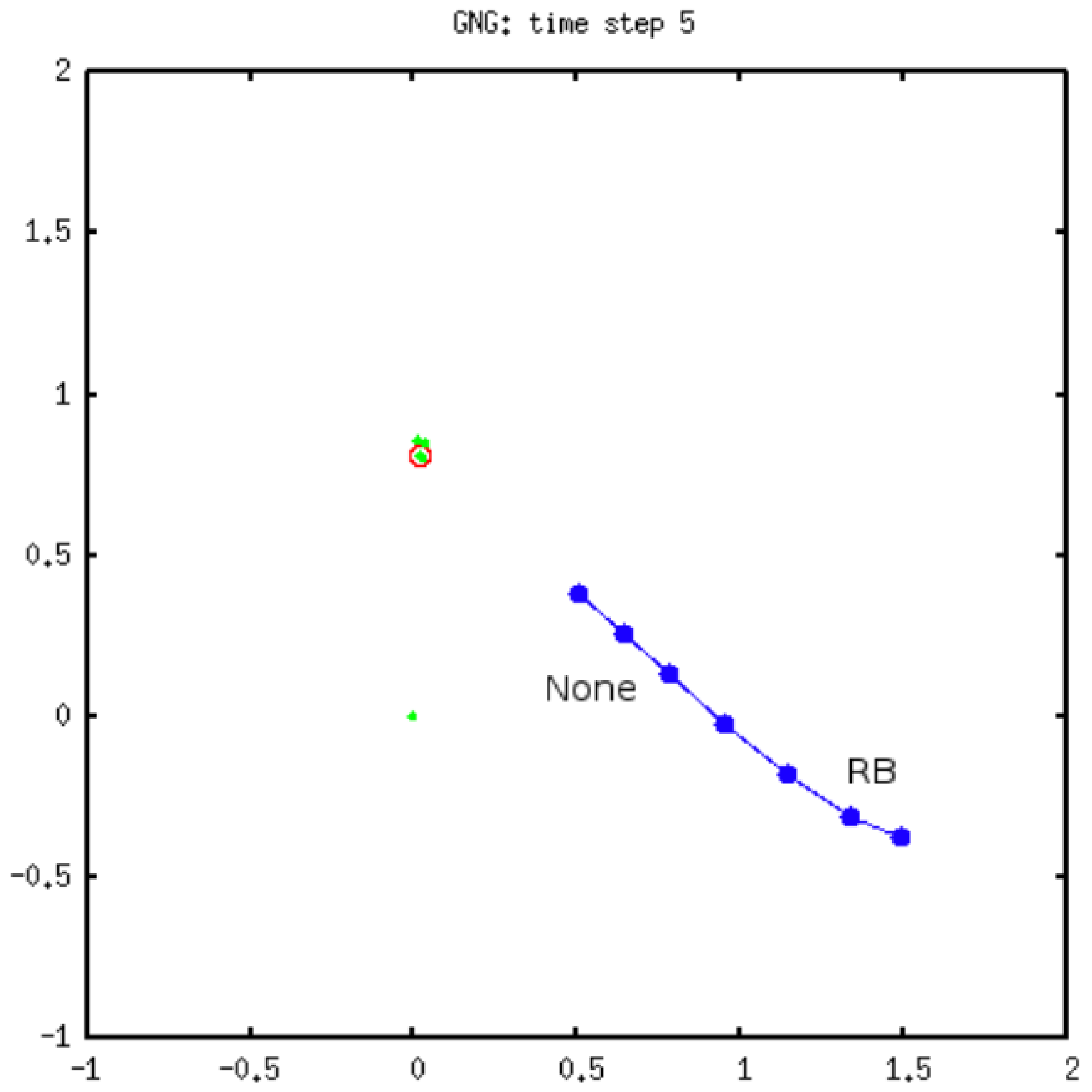
GNG growth over time



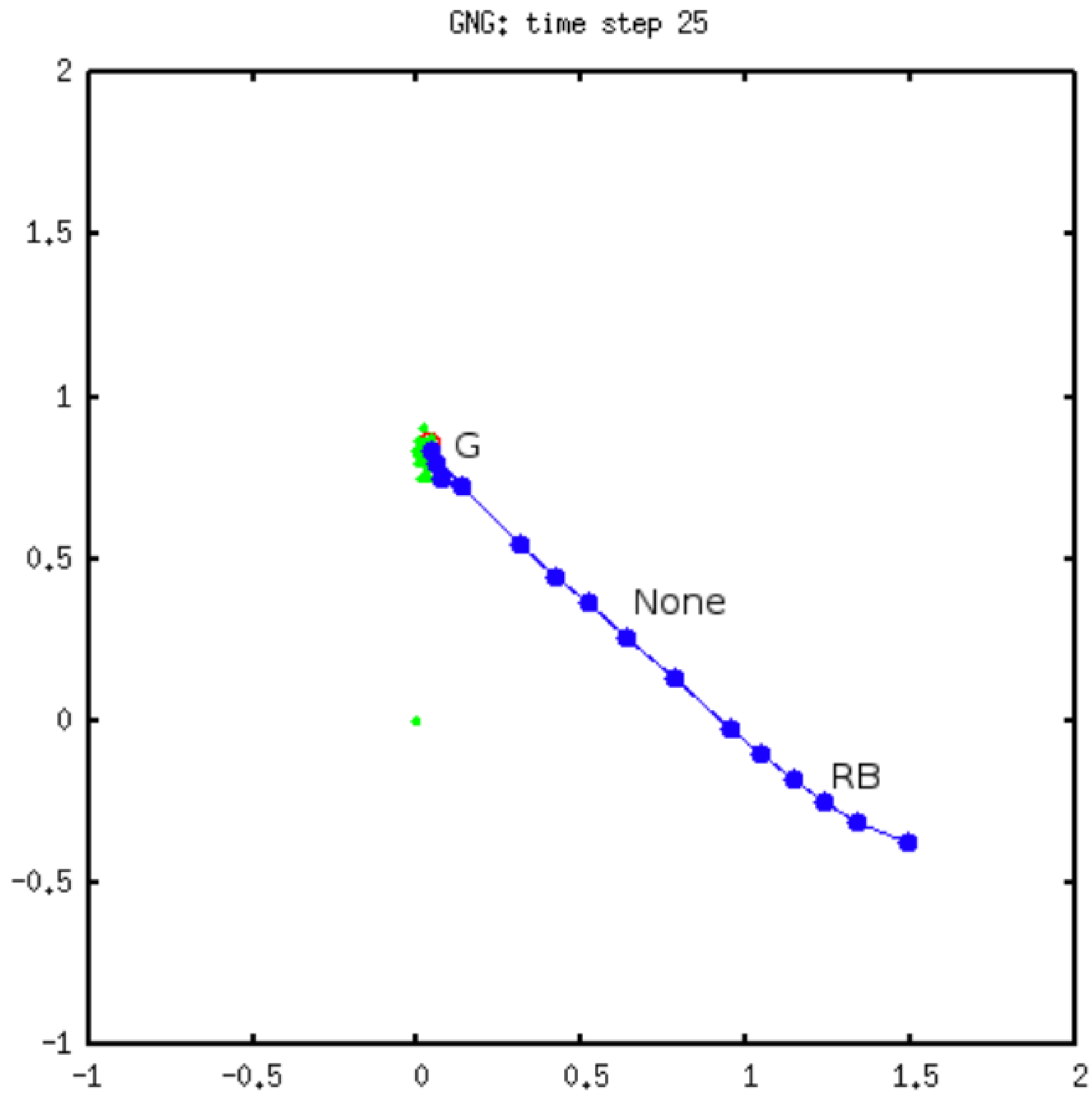
Categories formed over time



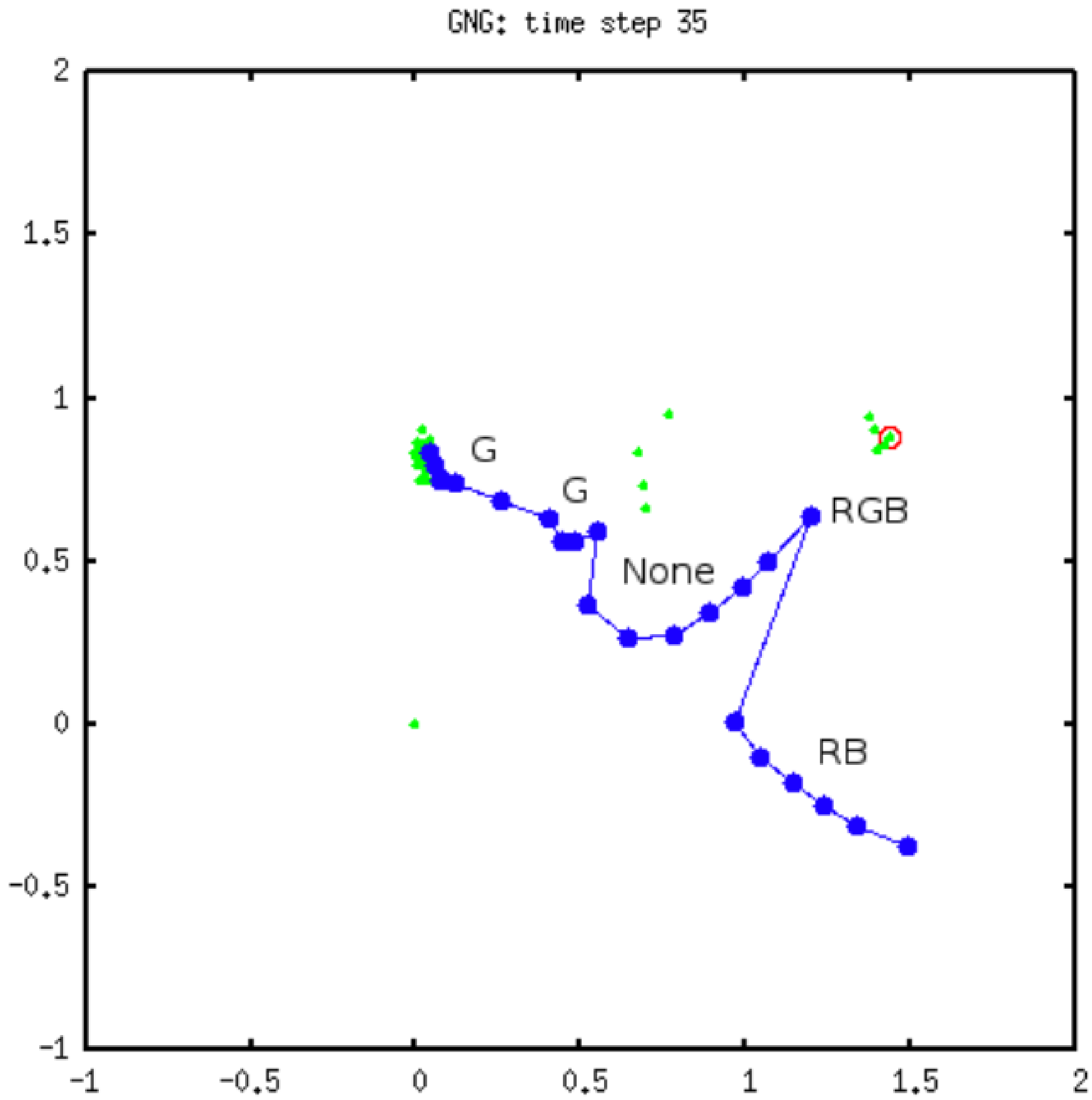
Categories formed over time



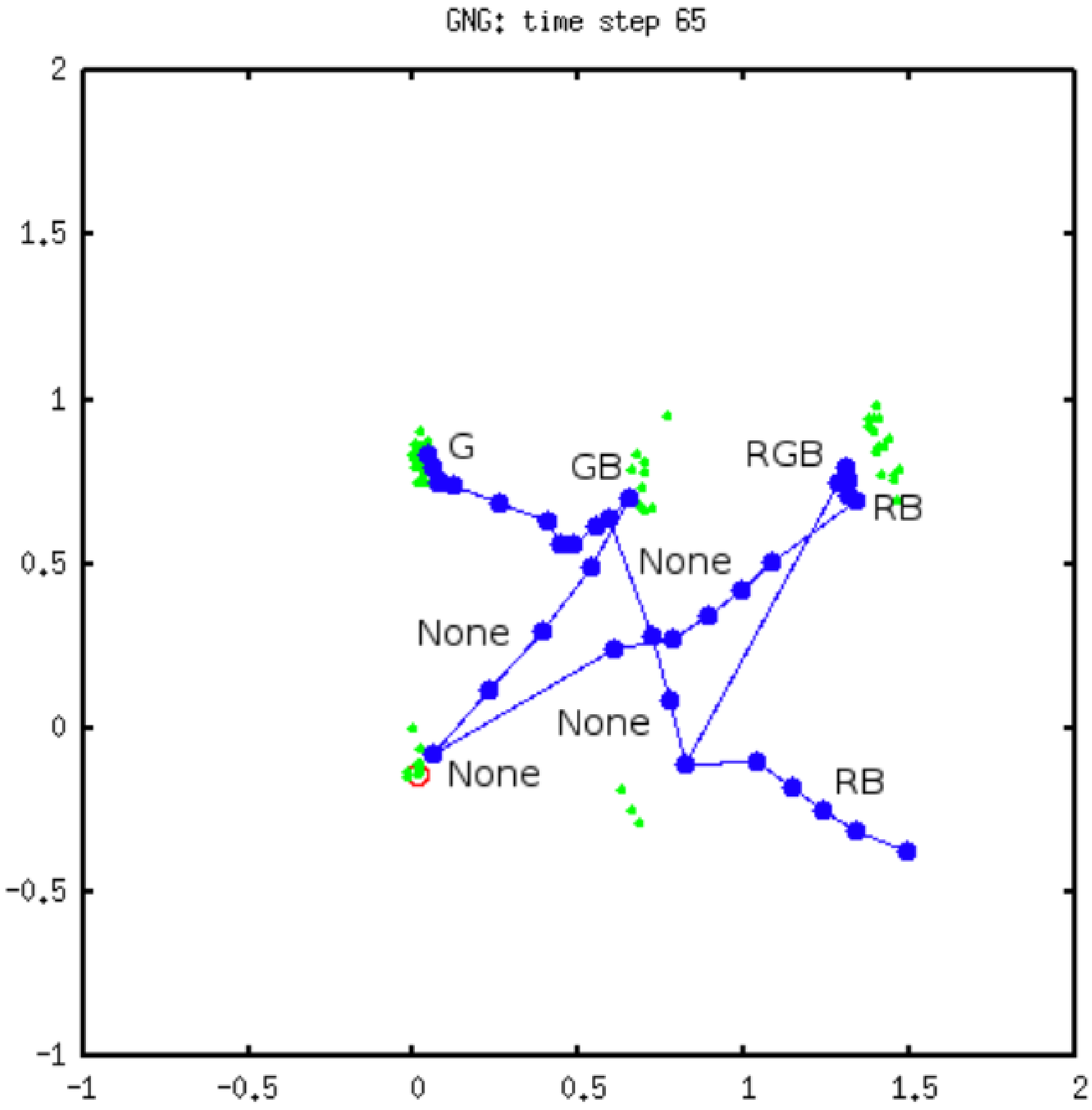
Categories formed over time



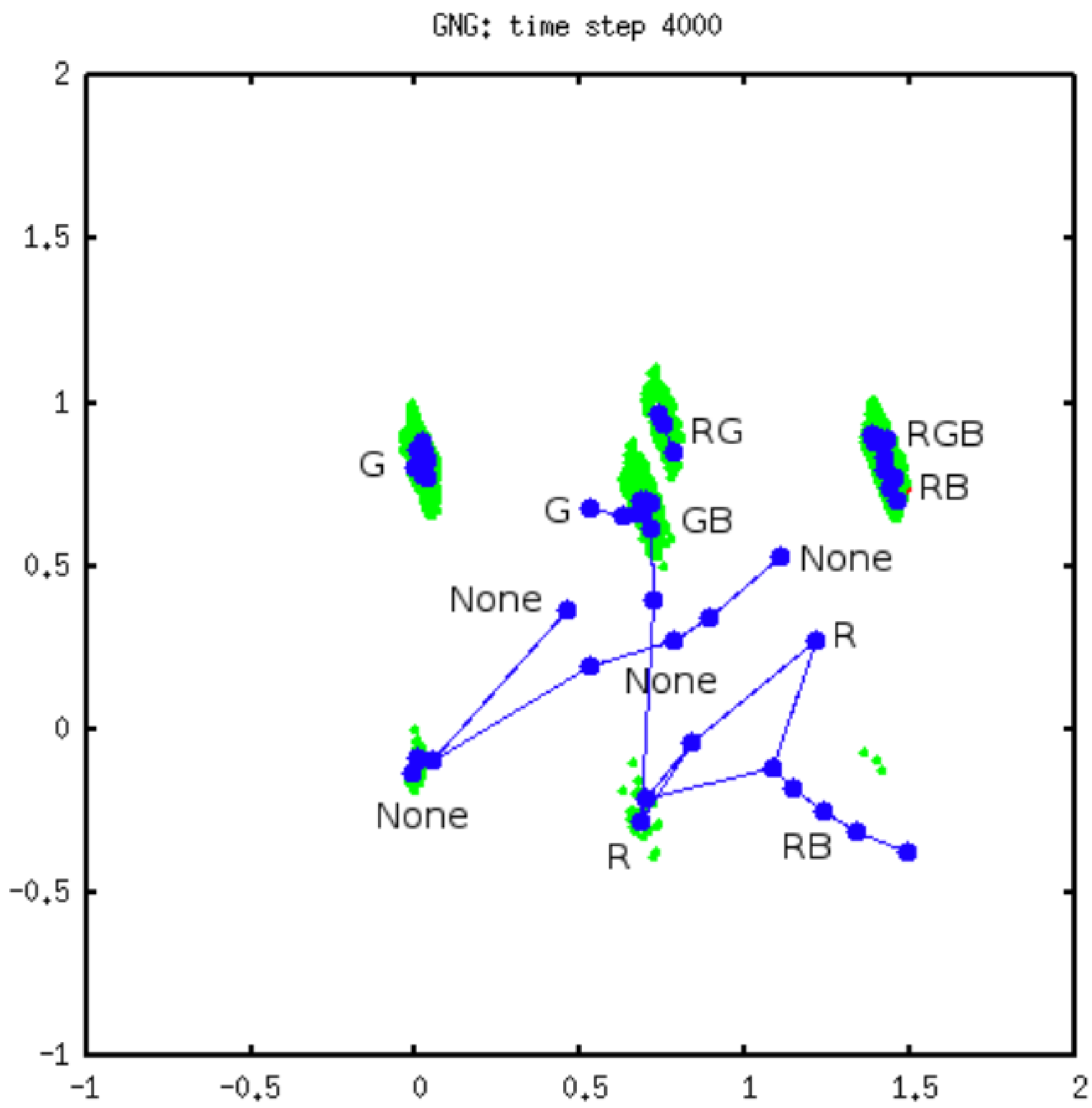
Categories formed over time



Categories formed over time



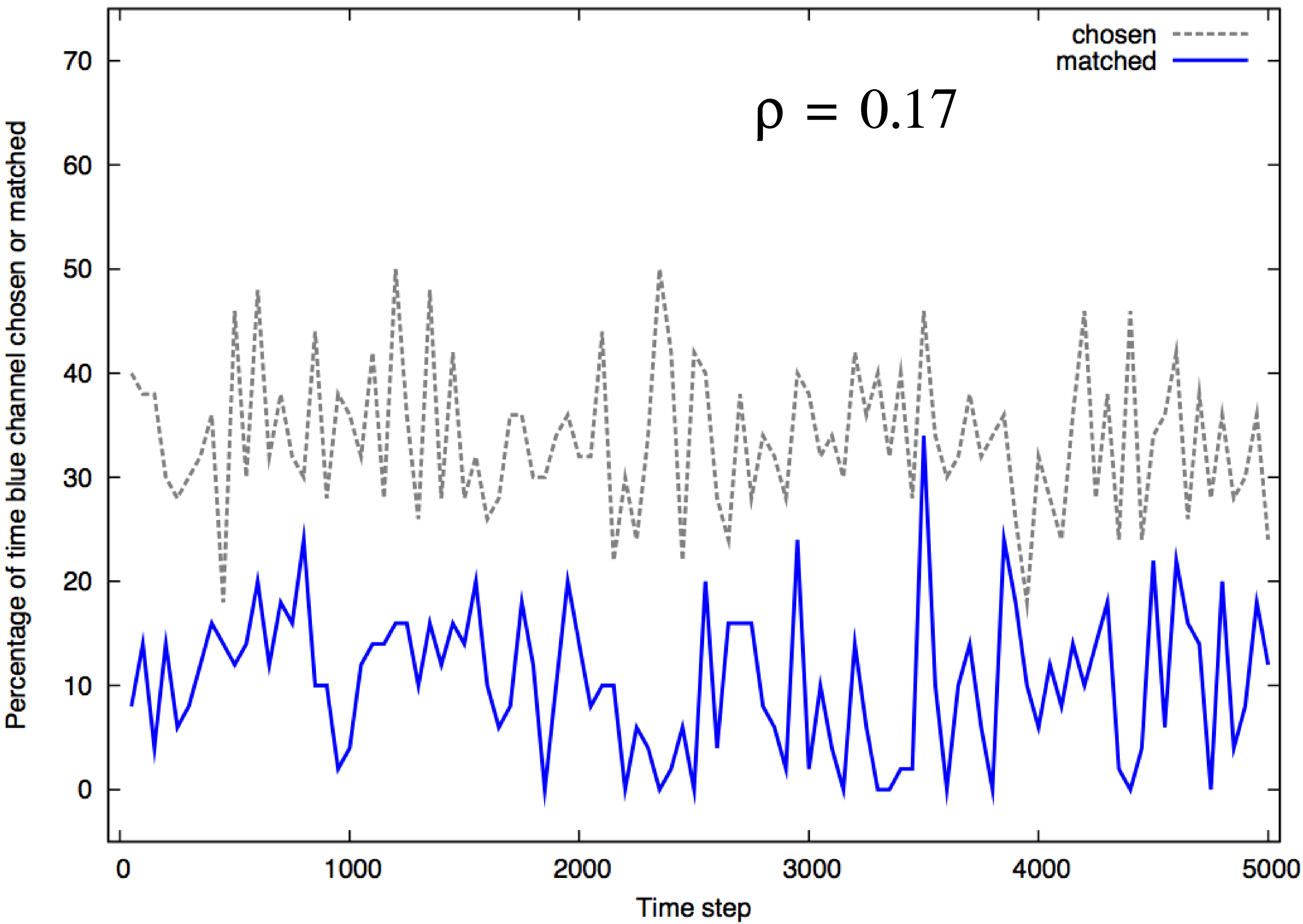
Categories formed over time



Results: Random controller

- Choosing actions **at random** causes the robot to focus on blue about 33% of the time, regardless of whether blue is actually present in the image
- The presence or absence of blue is relatively **uncorrelated** with the robot's choice of color channel
- This is to be expected

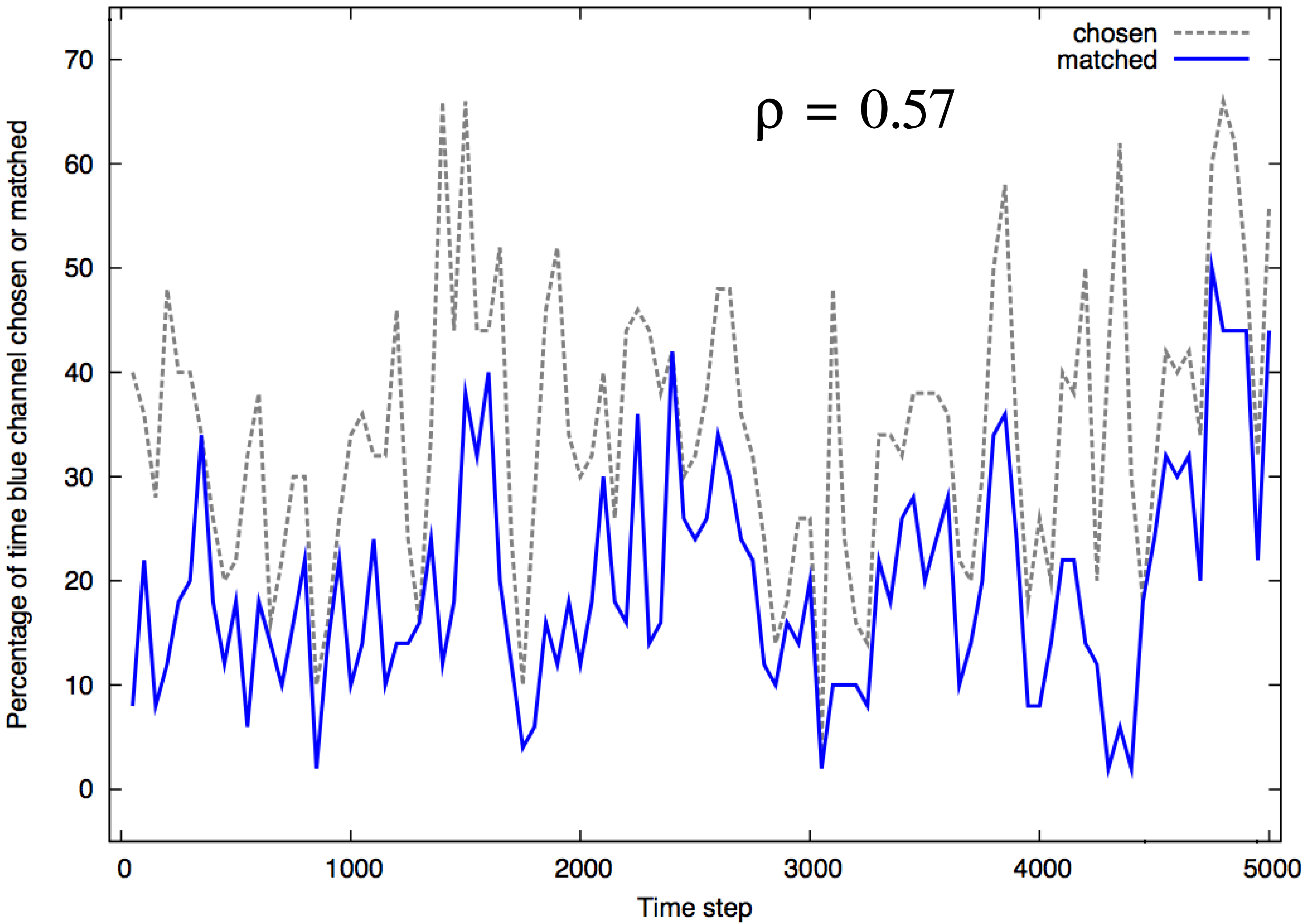
Random controller



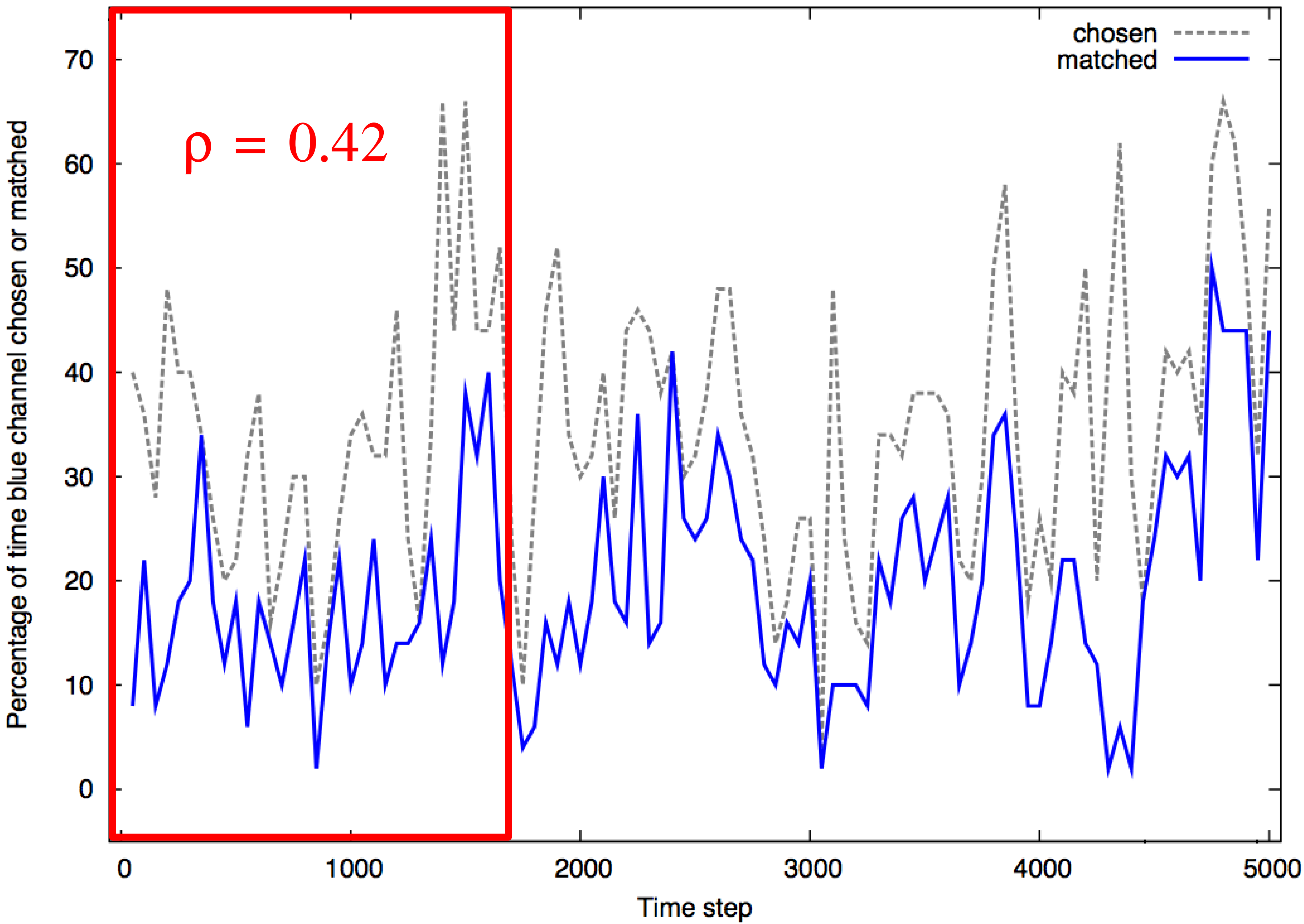
Results: Intrinsically motivated controller

- Choosing **intrinsically-motivated actions** causes the robot to focus on blue much more often when blue is present in the image
- The **correlation coefficient** increases from 0.17 to 0.57
- This correlation becomes **progressively stronger** over time, showing that the Rovio is learning to track the smaller blue robot

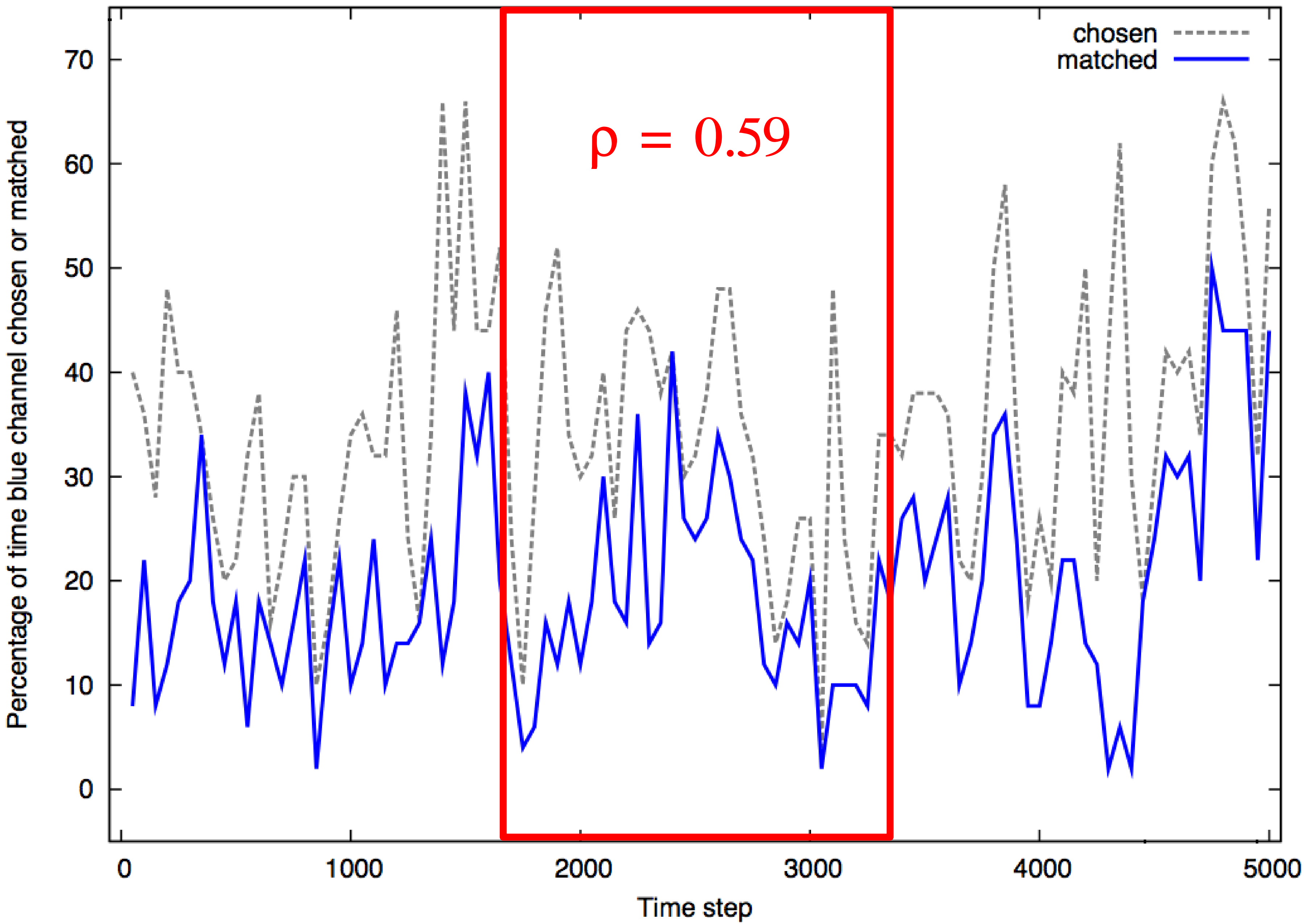
Intrinsically motivated controller



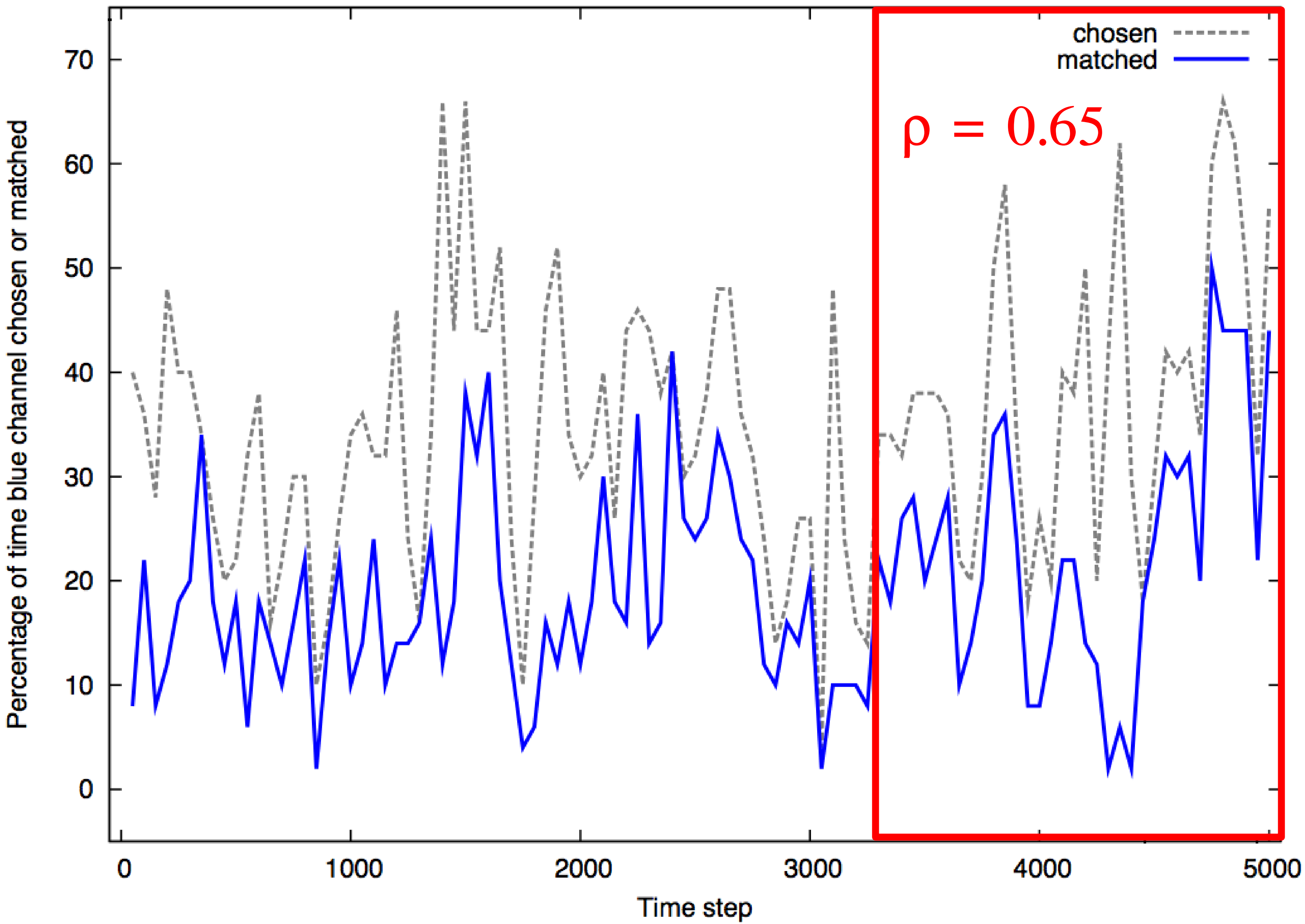
Intrinsically motivated controller



Intrinsically motivated controller



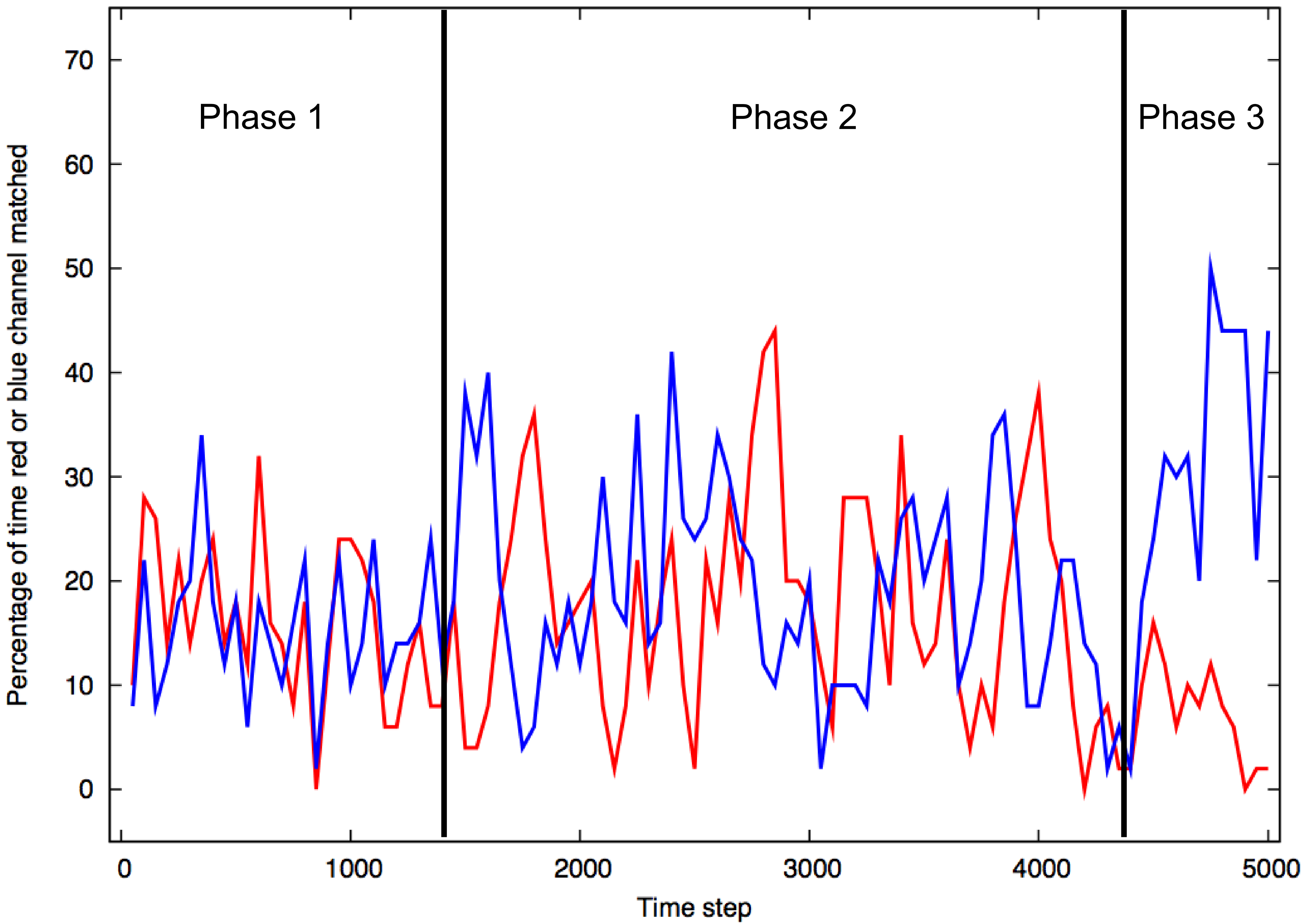
Intrinsically motivated controller



Results: Developmental trajectory

- By comparing the results for the red color channel and the blue color channel, evidence for a **developmental trajectory** can be seen
- In the early stages of learning, the Rovio focuses more closely on the predictable **red** robot
- Later on, the Rovio shifts to tracking the harder-to-predict **blue** robot more closely

Developmental trajectory



Conclusions

- Combining GNG's categorization with IAC's measure of learning progress allows the robot to develop an effective set of categories adapted to its environment
- GNG grows only as much as necessary to capture the significant relationships within the sensorimotor data
- The robot gradually shifts from learning about features of its world that are easy to predict to learning about features that are harder to predict

Results

- The **green** color channel is active on nearly every time step, and is the easiest to predict, yet always has the **lowest** overall focus
- In 10 experiments performed with **intrinsically-motivated** action selection, the robot consistently focused on blue and/or red more often than green
- In 5 experiments performed with **random** action selection, there was no significant difference in focus between red, green, or blue

Perception-Action Loop

- Perceive current sensory input S(t)
- Consider possible motor actions $M_1(t)$, $M_2(t)$, $M_3(t)$, ...
- Find best matching category in memory for each sensorimotor combination $SM_1(t)$, $SM_2(t)$, $SM_3(t)$, ...
- Choose action M(t) associated with category with maximal learning progress, and predict outcome S'(t+1)
- Do M(t) and observe actual outcome S(t+1)
- Use prediction error to update neural network weights
- Adjust GNG categories to better match chosen SM(t)