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”

• “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”
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 similarity 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

GNG Categories

SM_a
expert_a

SM_b
expert_b

SM_c
expert_c

SM_d
expert_d

SM_e
expert_e
Perceive current sensory input $S(t)$

$S(t)$
Consider possible motor actions

$S(t)$

Camera image

$M_1(t)$
$M_2(t)$
$M_3(t)$

Possible motor actions

GNG Categories

$S_{M_a}$  
expert$_a$

$S_{M_b}$  
expert$_b$

$S_{M_c}$  
expert$_c$

$S_{M_d}$  
expert$_d$

$S_{M_e}$  
expert$_e$
Find best matching categories

$S(t)$
Camera image

$M_1(t)$
$M_2(t)$
$M_3(t)$
Possible motor actions

$SM(t)$
Sensorimotor input

$SM_a$
$SM_b$
$SM_c$
$SM_d$
$SM_e$
GNG Categories

expert_a
expert_b
expert_c
expert_d
expert_e
Find best matching categories

$S(t)$

Camera image

$M_1(t)$

$M_2(t)$

$M_3(t)$

Possible motor actions

$\text{SM}(t)$

Sensorimotor input

$\text{SM}_a$

expert$_a$

$\text{SM}_b$

expert$_b$

$\text{SM}_c$

expert$_c$

$\text{SM}_d$

expert$_d$

$\text{SM}_e$

expert$_e$

GNG Categories
Find best matching categories

$S(t)$

Camera image

$M_1(t)$
$M_2(t)$
$M_3(t)$

Possible motor actions

SM(t)
Sensorimotor input

GNG Categories

$SM_a$
expert$_a$

$SM_b$
expert$_b$

$SM_c$
expert$_c$

$SM_e$
expert$_e$

$SM_d$
expert$_d$
Select expert with maximal learning progress

\[ S(t) \]

Camera image

\[ SM(t) \]

Sensorimotor input

\[ S'_t(t+1) \]

Prediction

GNG Categories

- \( SM_a \) [expert_a]
- \( SM_b \) [expert_b]
- \( SM_c \) [expert_c]
- \( SM_d \) [expert_d]
- \( SM_e \) [expert_e]

Selected action

- \( M_1(t) \)
- \( M_2(t) \)
- \( M_3(t) \)
Perform selected action and observe outcome

GNG Categories

SM(t)
Sensorimotor input

SMa
expert_a

SMb
expert_b

SMc
expert_c

SMe
expert_e

S'(t+1)
Prediction

S(t+1)
Outcome
Update expert based on prediction error

SM(t)
Sensorimotor input

SM(t+1)
Prediction

S(t+1)
Outcome

GNG Categories

SMa
expert
Sm
expert
Sc
expert
Sd
expert
e
expert

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

\[ S(t) = (\text{red}, \text{green}, \text{blue}, \text{area}, \text{position}) \]

\[ = (1, 1, 1, 0.23, 1.0) \]
Example: Robot chooses to look at blue

\[ S(t) = (\text{red}, \text{green}, \text{blue}, \text{area}, \text{position}) \]

\[ = (1, 1, 0, 0, 0) \]
Motor actions

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

- Example: $(0.8, 0.2) = \text{focus on blue, turn left}$

- Sensorimotor vectors $\text{SM}(t)$ 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

![Graph showing GNG growth over time](image-url)
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

\[ \rho = 0.17 \]
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

\[ \rho = 0.57 \]
Intrinsically motivated controller

\[ \rho = 0.42 \]
Intrinsically motivated controller

\[ \rho = 0.59 \]
Intrinsically motivated controller

\[ r = 0.65 \]
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

Phase 1

Phase 2

Phase 3

Percentage of time red or blue channel matched

Time step

0 1000 2000 3000 4000 5000
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 $SM(t)$