Evolution of Self-Interested Agents

An Experimental Study

Naoki Yamada

Chiaki Sakama

Wakayama University, Japan

MIWAI 2013, Krabi, Thailand

Cooperation vs. Selfishness

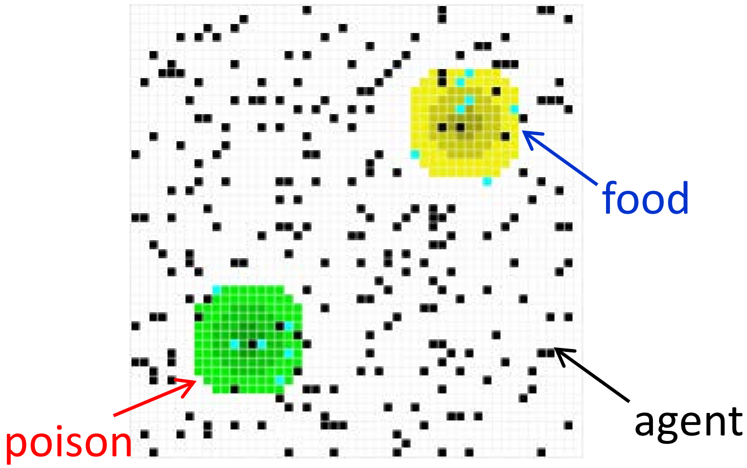
- Animals communicate in many ways to share information.
  (food location, predator attacks, etc)
- Animals compete with one another for limited resources.
  (food, space, mate, etc)
- "Any altruistic system is inherently unstable, because it is open to abuse by selfish individuals, ready to exploit it."
  -- Richard Dawkins: "*The Selfish Gene*" (1976)

Question:

How self-interested behavior evolves in resource-restrictive environments?

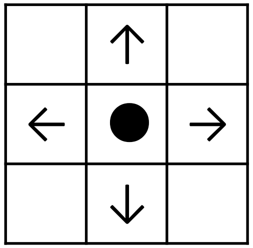
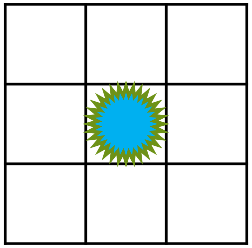
Outline of the Study

- We implement a multiagent system in which agents initially behave cooperatively by sharing information of food.
- After generations of a genetic algorithm, we observe the evolution of self-interested agents who do not share food information.
- We analyze experimental results and see the adaptation of agents in a resource-restrictive environment.



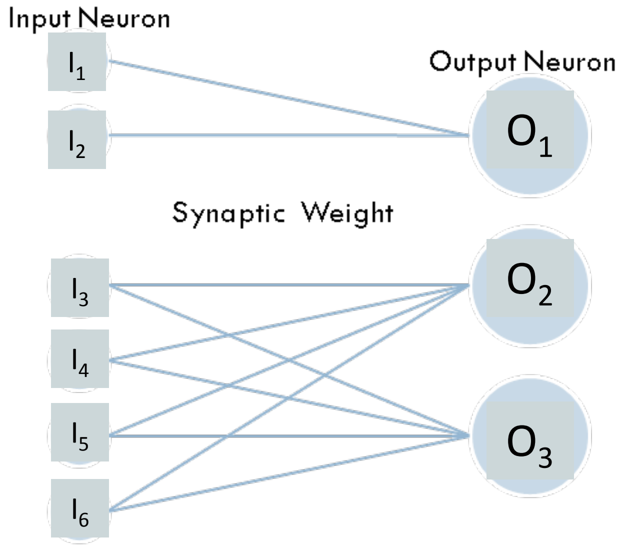
Environment

- We set the environment as a two-dimensional grid of 50×50 cells where 250 agents are living.
- Each agent stays at one cell and two different agents cannot stay at a cell at the same time.
- The environment contains 132 cells of food and 132 cells of poison.
- The spacial constraints on the food allow a maximum of 132 agents to be fed simultaneously.
- The maximum amounts of food or poison in each cell are initially given (range from 1 to 3).

Agents

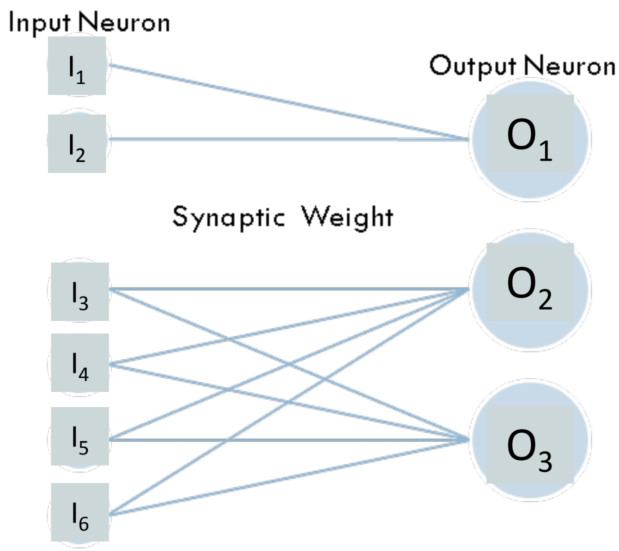
- Agents act synchronously in discrete time steps.
- A generation consists of 200 time steps.
- Each agent can move to neighbor cells (horizontally or vertically adjacent cells), send a signal, and obtain food or poison at every step.
- If an agent stays at a cell where food or poison is located, then it is counted as one at a step. The numbers of food or poison which an agent obtains in one generation are counted.
- Each agent can send a signal when they find food or poison.
- An agent can recognize signals on the 360° field and decides a direction to move at the next step based on the amount and direction of signals.

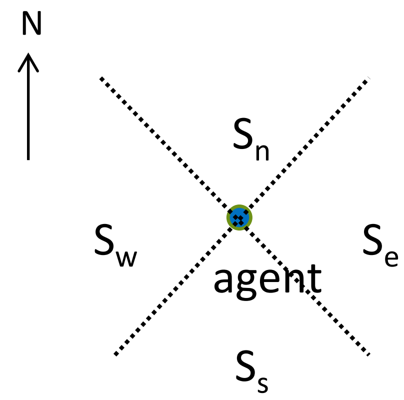


Neural Network

- Each agent has a simple neural network which consists of 6 input neurons I_1, I_2, I_3, I_4, I_5, I_6 and 3 output neurons O_1, O_2, O_3 .
- Input neurons and output neurons are connected through 10 synaptic weights SW_{ij} ($1 \leqq i \leqq 6$; $1 \leqq j \leqq 3$) which represent the strength of connections.

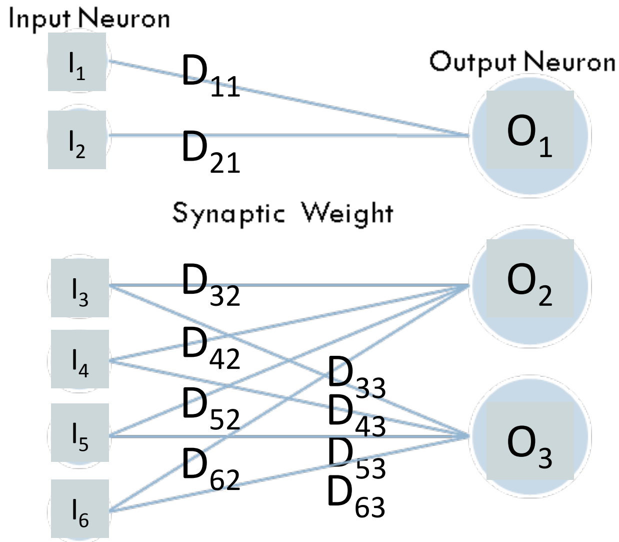
Values of Input Neuron





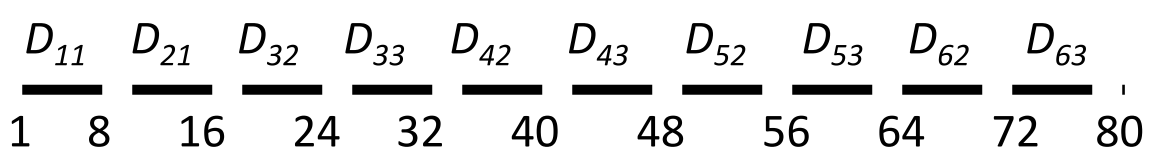
- $v(I_1)=1$ if there is food on the cell where an agent is staying; otherwise, $v(I_1)=0$.
- $v(I_2)=1$ if there is poison on the cell where an agent is staying; otherwise, $v(I_2)=0$.
- $v(I_3)= S_w/S$, $v(I_4)= S_e/S$, $v(I_5)= S_n/S$, $v(I_6)=S_s/S$, where S_w, S_e, S_n and S_s respectively represent the amount of signals from the four sections (west, east, north and south) of 90° each and $S = S_w + S_e + S_n + S_s$.

Values of Synaptic Weights

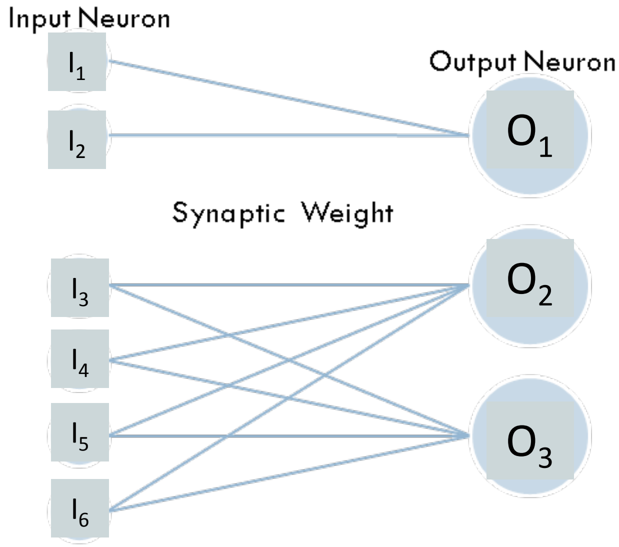


- Each agent has a sequence of 80-bit genes as a binary digit.
- The values of synaptic weights are calculated by: $SW_{ij} = \frac{D_{ij} \times 2}{255} - 1$

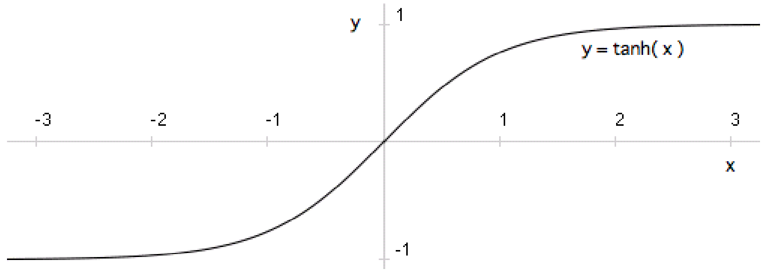
  where D_{ij} represents the decimal number corresponding to 8-bit genes.



- A synaptic weight takes a value of $-1 \leqq SW_{ij} \leqq 1$ (i.e. -1 if D_{ij}=0; 1 if D_{ij}=255)



Values of Output Neurons



- The values of output neurons are computed using the values of input neurons and synaptic weights as follows:

$$O_1 = \tanh\left(\sum_{k=1}^{2}(v(I_k) \times SW_{k1})\right) \qquad O_n = \tanh\left(\sum_{k=3}^{6}(v(I_k) \times SW_{kn})\right) \quad (n = 2,3)$$

- O_j is expressed using the hyperbolic function and takes a value between -1 and 1.

- The values of output neurons are used for deciding action of an agent.

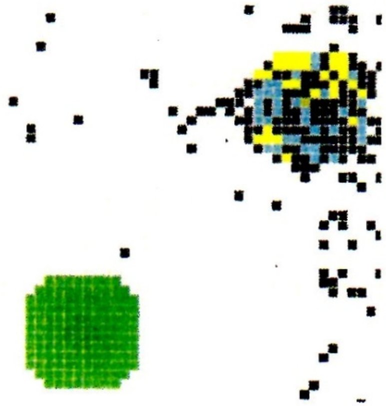
Action Rules

- An agent can take two different actions at each step: sending a signal or moving to neighbor cells.
- An agent sends a signal if the value of the output neuron is $O_1>0$. This may happen when an agent finds food ($I_1>0$) or poison ($I_2>0$), but whether $O_1>0$ or not depends on the values of the synaptic weights.
- A move of an agent is decided by the values of output neurons O_2 and O_3. A move of an agent is expressed by (dx, dy) where dx (resp. dy) represents a movement in the x-axis (resp. y-axis) direction and is defined as follows:

$$dx = \begin{cases} -1 & (O_2 < -\frac{1}{3}) \\ 0 & (-\frac{1}{3} \le O_2 \le \frac{1}{3}) \\ 1 & (\frac{1}{3} < O_2) \end{cases} \qquad dy = \begin{cases} 1 & (O_3 < -\frac{1}{3}) \\ 0 & (-\frac{1}{3} \le O_3 \le \frac{1}{3}) \\ -1 & (\frac{1}{3} < O_3) \end{cases}$$

Evolution

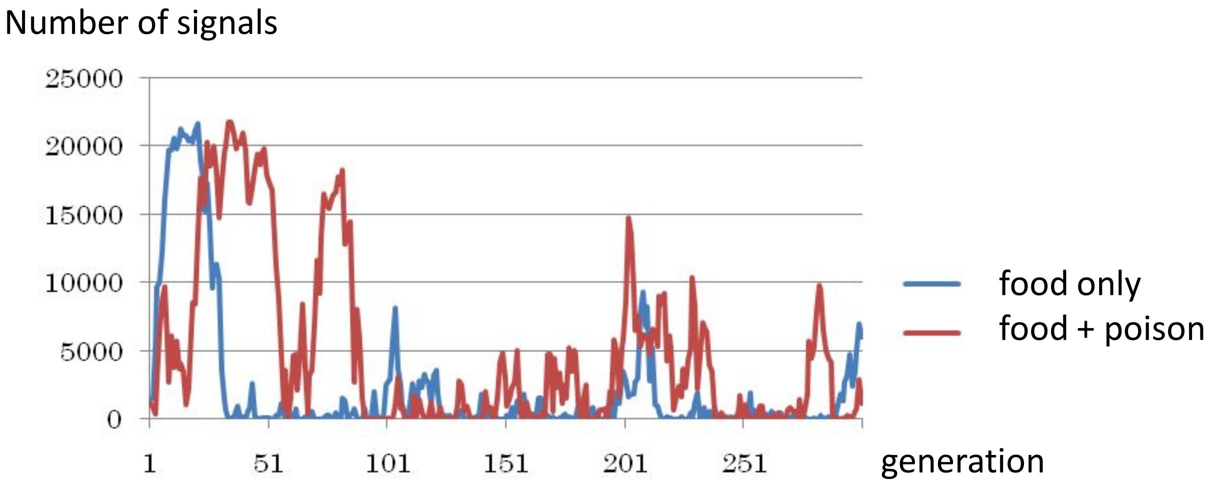
- The movement of an agent depends on the values of its synaptic weights which are calculated by a sequence of 80-bit genes of the agent.
- The initial genes are randomly generated, and then evolve under a fitness condition using the genetic algorithm.
- The fitness of a gene is computed by $f = F - P$ where F and P respectively represent the numbers of food and poison which an agent obtained in one generation.
- In each generation, agents having higher fitnesses are selected, and their genes are modified (by crossover and mutation) to form a new generation.



Experiments

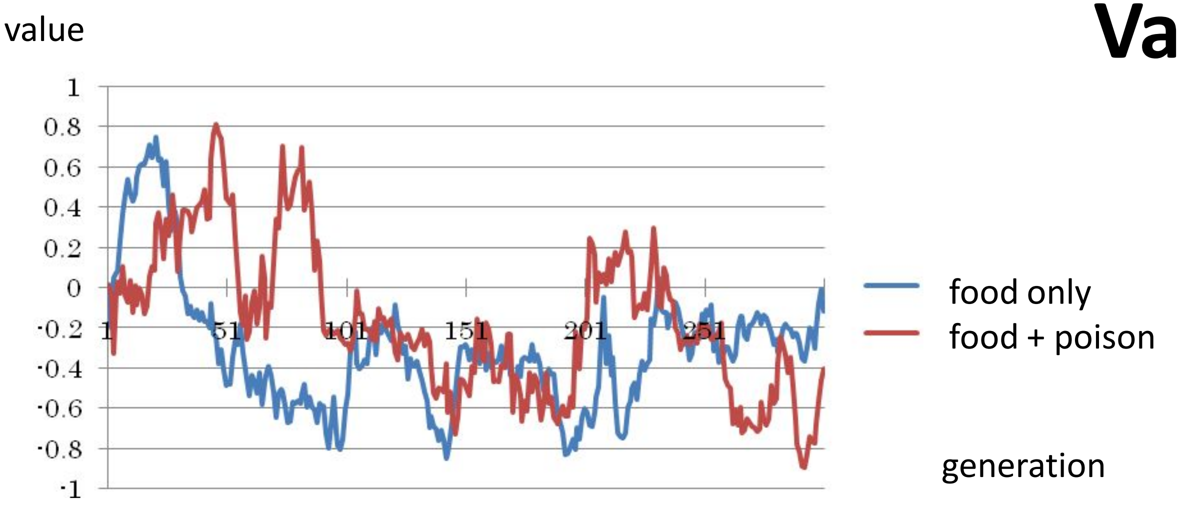
- Experiments are performed in three different situations: (i) the field contains food only; (ii) the field contains poison only; and (iii) the field contains both food and poison.
- Food is located in the north-east corner of the field and poison is located in the south-west corner of the field.
- The initial location of 250 agents is decided randomly and the evolution of agents is observed in 300 generations.
- With these settings, we observe the following changes over generations: (i) the number of signals in the field, and (ii) the average values of the synaptic weights among all agents.

Number of Signals around Food



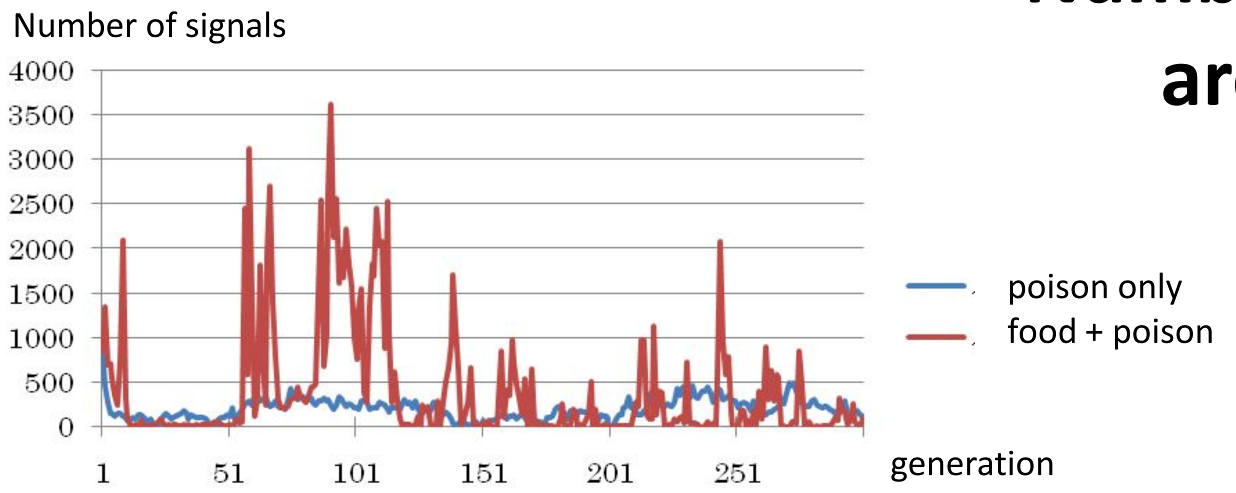
- As agents increases around the food, the number of signals increases accordingly. In several generations, it results in the crowd around food.
- Agents who actively send signals around the food cannot obtain food as before, which results in decreasing fitness values of those agents. In contrast, agents which do not actively send signals would have relatively high fitness values.
- Then the probability of selecting agents who actively send signals around the food reduces, which results in the decrease of signaling agents in the next generation.
- When the field contains both food and poison, the values sharply oscillate compared with the case of food only. This is because agents around the poison also send signals, which eliminates the effect of decrease of signals in the field.

Value of SW_{11}



- The average synaptic weight SW_{11} increases at first, while it decreases after the peak around the 20th generation.
- After the 50th generation, SW_{11} mostly takes negative values which indicates that most agents do not send signals around the food.
- However, agents who send signals around the food do not die out because reduction of signals has the effect of solving overpopulation around the food.
- When the field contains both food and poison, the average synaptic weight SW_{11} also oscillates between positive and negative values.

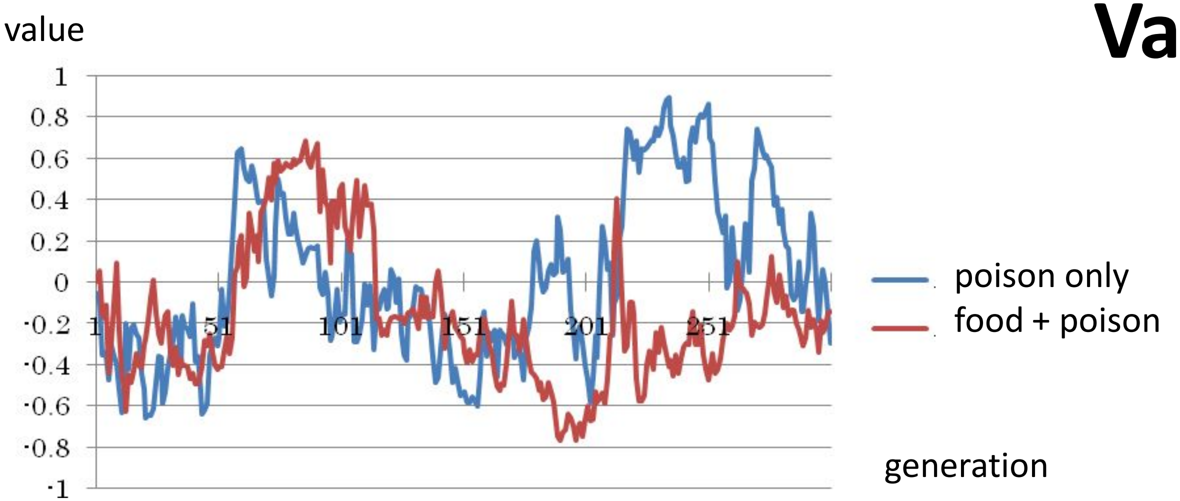
Number of Signals around Poison



- When the field contains poison only, the number of signals is relatively small through generations . The reason is that in this case agents who obtained poison have low fitness values by the equation, which results in the evolution that agents are directed away from signals.

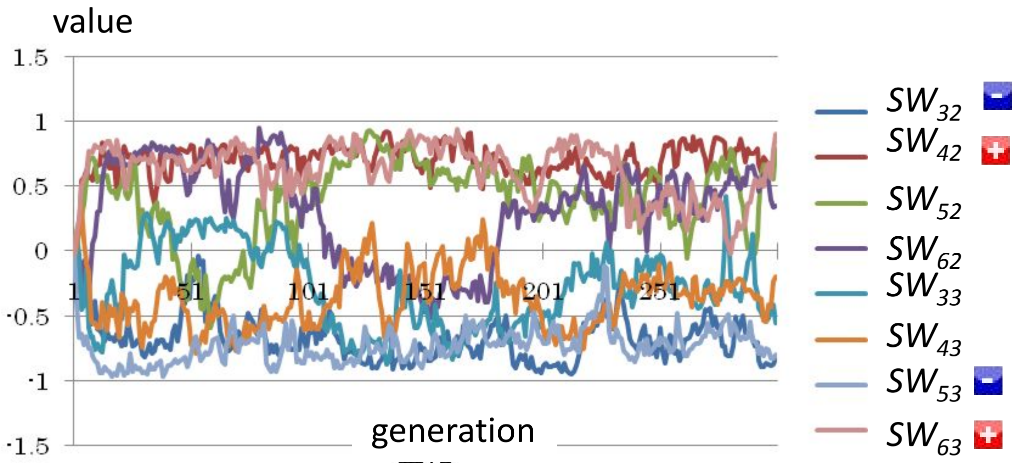
- When the field contains both food and poison, the number of signals randomly oscillates. Such a chaotic behavior is due to the mixture of signals from food and poison.

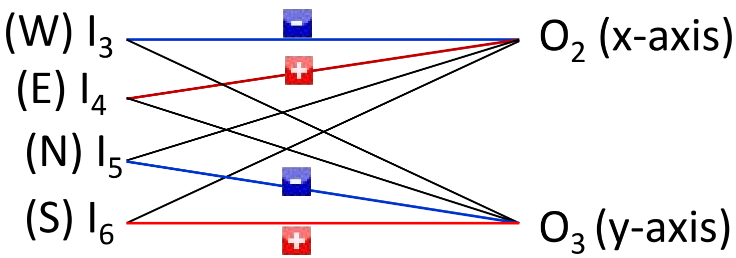
Value of SW_{21}



- The average synaptic weight SW_{21} oscillates between positive and negative values.
- This means that in case of poison stopping signals does not imply any advantage for an agent, which results in no particular evolution of the synaptic weight SW_{21} .

Movement (food only case)



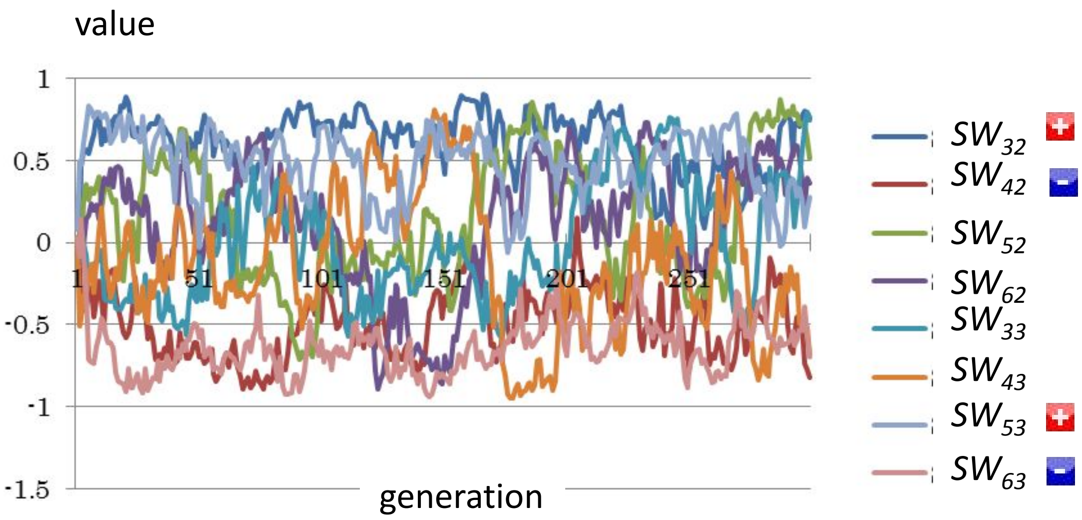


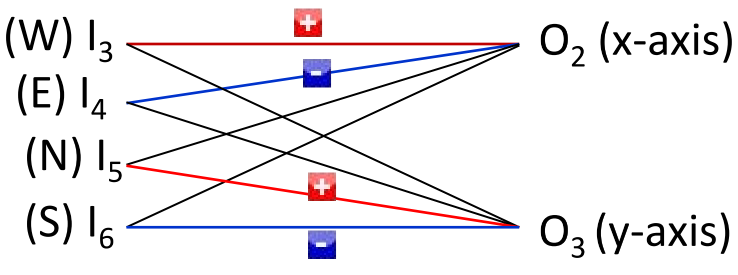
- SW_{42} and SW_{63} take positive values, SW_{32} and SW_{53} take negative values, and others oscillate between positive and negative values.

- SW_{42} depends on the signals from the east and SW_{32} depends on the signals from the west. Both SW_{42} and SW_{32} control movements in the *x*-axis direction.
  SW_{53} depends on the signals from the north and SW_{63} depends on the signals from the south. Both SW_{53} and SW_{63} control movements in the *y*-axis direction.

- Positive SW_{42} contributes to making O_2 positive (hence movement dx=+1) and negative SW_{53} contributes to making O_3 negative (hence movement dy=+1). This leads agents toward the food location in the north-east section of the field.

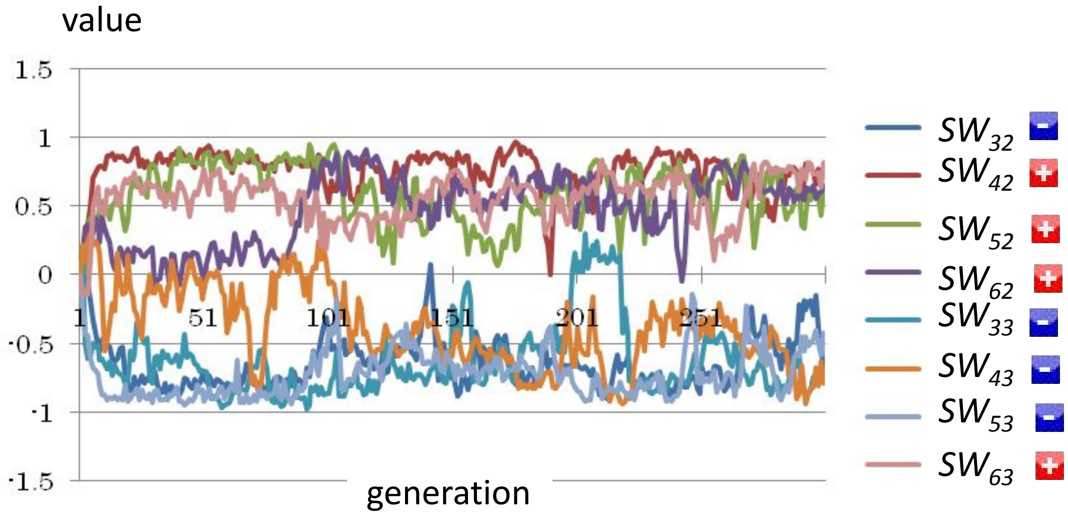
Movement
(poison only case)

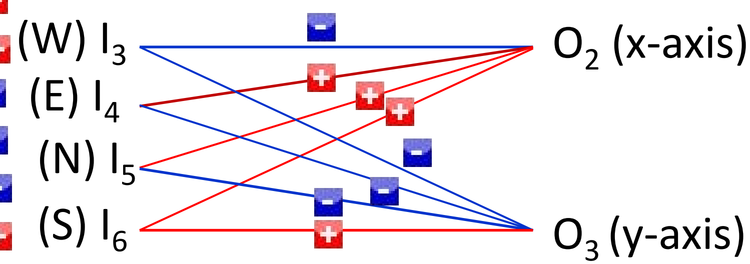




- SW_{42} and SW_{63} take negative values, SW_{32} and SW_{53} take positive values, and others oscillate between positive and negative values.
- The positive-negative patterns are the reverse of the food only case.
- This is explained that in the food-only case agents are evolved to be attracted to food, while in the poison-only case agents are evolved to be kept away from poison.

Movement
(food + poison)





- Synaptic weights, which are oscillated in the food-only or poison-only case, converge on either positive or negative values. SW_{52} and SW_{62} take positive values, SW_{33} and SW_{43} take negative values.
- When signals from the north and the south, the positive weights SW_{52} and SW_{62} make O_2 positive, which will lead an agent to proceed forward the x-axis direction (the east).
- When signals from the east and the west, the negative weights SW_{33} and SW_{43} make O_3 negative, which will lead an agent to proceed forward the y-axis direction (the north).
- In each case, the weights lead agents close to food and apart from poison.

Related Work

S. Mitri, D. Floreano and L. Keller:
"The evolution of information suppression in communicating robots with conflicting interests"
In: *Proc. National Academy of Sciences* 106(37), 2009.

- Robots are placed in an arena containing a food source and a poison source that both emit red. The robots earn points for how much time they spent near food as opposed to poison.

- The robots emit blue light randomly which other robots could perceive. As robots become efficient at finding food and staying nearby, the concentration of blue light near food also increases. Thus, blue light plays an inadvertent cue providing information on the food location.

- Spacial constraints around the food source allow a maximum of 8 robots of 10 to feed simultaneously and result in higher robots density and increased competition and interference near the food.

- After 100 rounds, the robots with the highest scores are selected for the next round. By the 50th generation, robots are selected to decrease the rate of blue light emission. Thus, selection is acting toward suppressing information on the food location.

Comparison

- In [Mitri et al. 2009] unintentional communication develops useful information, which generates self-interested behaviors of robots. By contrast, we set agents to send signals when they find food or poison. We then observe that agents who behave cooperatively at first also turn to become self-interested in evolution.
- In [Mitri et al. 2009] each robot has a neural network which is more complicated than ours. The length of genes encoded in a robot is more than 3 times longer than ours.

	[Mitri et al. 2009]	Our model
Input Neuron (#)	11	6
Outout Neuron (#)	3	3
Synaptic Weight (#)	33	10
length of gene (bit)	264(=8*33)	80(=8*10)

Conclusion

- We experimentally realized a multiagent system to observe the evolution of self-interested agents to survive in a resource-restrictive environment.
- The evolution of self-interested nature of agents from simple action rules would explain a reason for the emergence of selfish behaviors of animals in resource-limited environments in nature.
- In nature animals often send deceptive signals in order to keep others away from food. Further refinement of social models is needed to realize the evolution of agents who may act dishonestly.