An Agent and an Environment: A View on "Having Bodies"

— A Case Study on Behavior Learning for Vision-Based Mobile Robot —

Minoru Asada
Mech. Eng. for Computer-Controlled Machinery
Osaka University, Suita, Osaka 565, Japan

asada@robotics.ccm.eng.osaka-u.ac.jp
http://www-robotics.ccm.eng.osaka-u.ac.jp/Welcome-eg.html

Abstract

This paper discusses how so-called "intelligence" can be emerged as a cognitive process, that is, how an agent can develop its internal representation according to the complexity of the interactions with its environment through its capabilities of sensing and acting. The complexity might be increased by the existence of other active agents, and the development can be possible depending on how the agent can find a new axis in the internal representation in trying to accomplish a given task in the environment including other agents. As an example of such a development, I show a case of a vision-based mobile robot of which task is to shoot a ball into a goal with/without a goal keeper along with preliminary experiments by real robots.

1 Introduction

The ultimate goal of my research is to design the fundamental internal structure inside physical entities having their bodies (robots) which can emerge complex behaviors through the interactions with their environments. The consequences of behaviors can be regarded as "intelligent" or "emergent" from a viewpoint of the observer [1]. This means that to design so-called "intelligence" or " emergence" explicitly and directly is a very hard problem. In order to emerge the intelligent behaviors, bodies are indispensable. I consider the followings in mentioning "having bodies":

1. Sensing capability to sense something from the environment where diverse natures are involved.
2. Acting capability to emerge actions able to influence the environment.
3. Sensing and acting are tightly coupled and not separable.
4. In order to achieve the goal (the final goal is to survive?), the sensor and actuator spaces should be abstracted under the resource bounded conditions (memory, processing power, controller etc.).
5. The abstraction depends on both the fundamental embodiments inside the agents and the experiences (interactions with their environments). The consequences of the abstraction are the agent-based subjective representation of the environment, and its evaluation can be done by the consequences of behaviors.

Design principles are

- the design of the internal structure of the agent which has a body able to interact with its environment, and
- the policy how to provide the agent with tasks, situations, and environments so as to develop the internal structure.

In the following, first I show a method for so-called "segmentation" problem which is one of the most fundamental problems in AI and Robotics. Then, based on it I discuss the complexity of the environments where the robot cam emerge intelligent behaviors, especially the relationship between the complexity and the existence of other agents.

2 Reinforcement Learning and a Soccer Robot

For the readers' understanding of the fundamental problem, "segmentation," I show our first achievement of "soccer robot" [1] [2]. By applying the vision-based re-

[1]Someone claimed that it was not the soccer robot but just a shooting robot when I published the first paper. There were

inforcement learning method, the robot learned to shoot a ball into a goal. The state space consists of the sizes and the positions of both the ball and the goal, and the orientation of the goal in the image plane (see Fig.1).

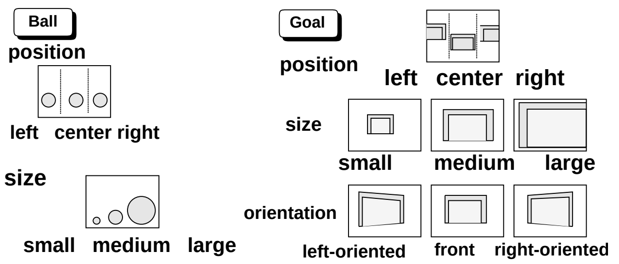


Figure 1: **State space consisting of a ball and a goal**

The action space consists of combinations of two motor commands each of which corresponds to one of forward, stop, and backward motions. Totally, we have nine kinds of actions. We assume that the robot does not know the physical meaning of these actions. The main problem we faced with is that the constructed states and actions do not correspond to each other. Traditional notions of state in the existing applications of the reinforcement learning schemes fit nicely into deterministic state transition models (e.g. one action is forward, backward, left, or right, and the states are encoded by the locations of the agent). However, it seems difficult to apply such deterministic state transition models to real robot tasks. In real world everything changes asynchronously [3].

Figure 2 indicates this problem, the area representing the state "the goal is far" is large, therefore the robot frequently returns to this state if the action is forward. This is highly undesirable because the variations in the state transitions is very large, consequently the learning does not converge correctly.

To avoid this problem, we reconstruct the action space as follows. Each action is regarded as an action primitive. The robot continues to take one action primitive until the current state changes. This sequence of the action primitives is called an action.

Figure 3 shows some kinds of behaviors obtained by our method. The difference in character of robot players due to the discounting factor γ is shown in (a) and (b) in which the robot started from the same position.

two reasons for us to call it soccer robot. First, we were worrying about that people might associate gun-shooting when they heard of "shooting robot." Second, we stressed ourselves to build up a team of soccer playing robots by declaring "soccer robot." However, we have published few papers entitled something "soccer robot."

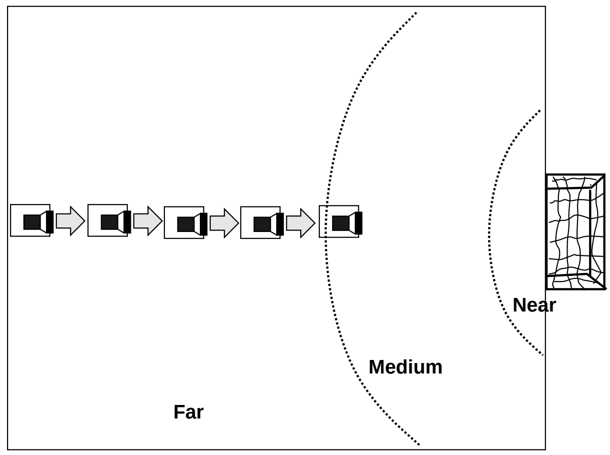
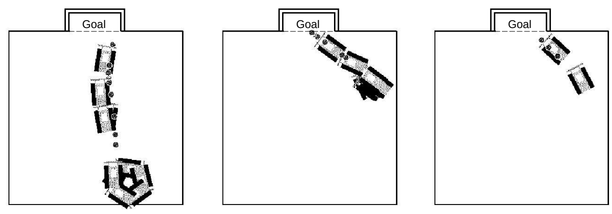


Figure 2: **A state-action deviation problem**



(a) finding, dribbling, and shooting
(b) shooting ($\gamma = 0.999$)
(c) shooting ($\gamma = 0.6$)

Figure 3: **Some kinds of behaviors obtained by our method**

In the former, the robot takes many steps in order to ensure the success of shooting because of a small discount, while in the latter the robot tries to shoot a ball immediately because of a large discount. In (c), the robot started at a position from where it could not view a ball and a goal, then found the ball by turning, dribbled it towards the goal, and finally shot the ball into the goal. This is just a result of learning. We did not decompose the whole task into these three tasks.

Reinforcement learning is generally regarded as one of the unsupervised learning methods. However, it needs the well defined state and action spaces so that the learning can converge, which makes it difficult to be regarded as unsupervised. That is, the state space tells the robot how to segment the world. The more essential problem than "state-action deviation" seems that the robot should construct the state and action spaces for itself.

3 State space segmentation

The state space construction problem is divided into two subproblems:

1. how to find the feature vector which includes the information necessary and sufficient for the robot to achieve the goal, and

2. how to segment the selected feature vector into states for the learning.

In the case of soccer robot [2], the segmentation problem might be much easier because the ball (red) and the goal (blue) regions in the image have been already extracted. For the first subproblem, we applied the principal component analysis for the fundamental region features such as area, centroid, and moments from many sequences of image data, and as a result, the first five principal components correspond to ball size, ball position, goal size, goal position, and goal orientation [4].

Basic ideas to cope with the second subproblem, how to segment the selected feature vector, are that we define a state as a cluster of of input vectors from which the robot can reach the goal state or the state already obtained by a sequence of one kind action primitive regardless of its length, and that this sequence is defined as one action.

The initial state space consisting of the goal state and the other is iteratively separated into several states. Fig 4 and Table 1 show the results. In the figure, the final state space is projected into two dimensional space in terms of the ball size and the goal size (when their positions are frontal and the orientation of the goal is horizontal) where the gird lines indicate state segmentation designed by the programmer that are quite different in shape and size from the obtained states. The success rate has been improved although the number of states is drastically decreased, and therefore the search time is very short.

Table 1: **Comparison with existing methods**

	Number of States	Search Time	Success Rate (%)
Previous work[2]	243	500M*	77.4
Proposed method	33	41M	83.3
cf. Fixed action length	107	222M	71.5

* indicates Q-learning time.

This suggests the followings:

1. The state space designed by the programmer is not always appropriate for the robot to accomplish a given task. Robot should construct the state space from its experiences (interaction with its environment) for itself, and this problem can be considered as "segmentation problem." How to see the world depends on the agent (capabilities of sensing and action) and its environment.

2. Because of the inter-dependence between state and action spaces[2] which resembles to the well known "chicken and egg" problem, the minimum constraint should be introduced. Here, we adopt "action primitive" as a minimum constraint, and parameterize its length as one action until states change. This can be regarded as a process of parallel construction of the state and action spaces.

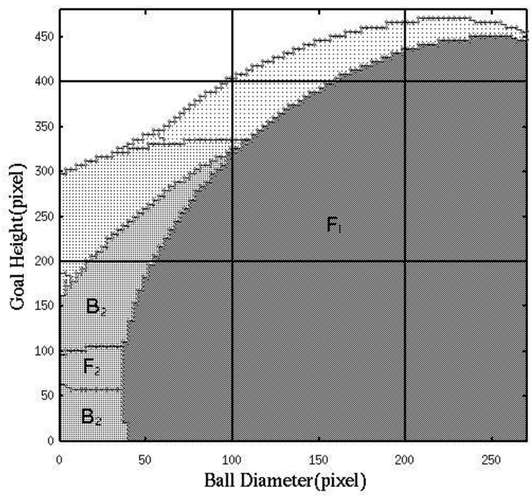


Figure 4: **2-D projection of the result of state space construction**

The latter can be regarded as temporal abstraction of the action space under the constraint of "action primitive." However, abstraction is mainly based on perception since the number of degrees of freedom is too few (just two) to abstract the action space.

We, human beings, can easily grasp an object by our hand directly when it is very close to ourselves, by reaching out our arm towards it when it is further, and by standing up and approaching to it when it is much further. Although the number of physical degrees of freedom we have is many and therefore it is actually a very difficult problem for the robot to control all of them, we

[2]Generally, the design of the state space in which necessary and sufficient information to accomplish a given task is included depends on the capability of agent actions. On the other hand, the design of the action space also depends on the capability of perception.

seem to easily perform the task. When grasping, one degree of freedom (open-close) seems sufficient regardless of the physical number of degrees of freedom of our hand. How can we acquire such abstraction of the action space? I have claimed that abstraction (segmentation) of the sensory information such as vision depends on the action capability of the agent. On the other hand, abstraction (segmentation) of the action space also depends on the sensing capability of the agent. It seems important to decide what is the minimum action capability necessary to abstract the both.

The essence of "having a body" is that the agent abstracts a variety of sensing modalities and movements with many degrees of freedom in order to achieve the goal (the ultimate goal seems to survive?) under the constraint of resource bounds as a physical entity. The consequences of behaviors after such abstraction can be seen "intelligent." In the following, I try to explain in case of vision that the complexity of the environment, especially the existence of other agent will cause to level up the abstraction.

4 Emergence, Intelligence, and Existence of Other Agents

Since each species of animals can be regarded to have its own intelligence, difference of intelligence seems to depend on the agent (capabilities in sensing, acting, and cognition) and its environment. If agents have the same bodies, differences or levels in intelligence can occur in the complexity of interactions with their environments. In case of our soccer playing robot with vision, the complexity of interactions may change due to other agents in the field such as agents of common side, opponents, judges and so on. In the following, I attempt at showing my view about the levels of complexity of interactions, especially from a viewpoint of the existence of other agents. To simplify the discussion, the following assumptions are adopted: 1) the agents has a vision sensor which can observe the consequences of the agent's actions, 2) the agent has n-DOFs, each of which has its own servo loop and/or the result of motor command execution can be measured, and 3) the environment almost consists of stationary objects, and other few objects can be in it as discussed below.

1. Self definition (boundary of the body): The area in which an agent capable of action can directly correlate between motor commands and visual information.

2. Static environment: By direct correlation between motor commands the agent sent and the visual information observed during the motor command executions, the agent can discriminate the static environment from others.

   Theoretically, discrimination between "self body" and "static environment" is a hard problem because the definition of "static" is relative and depends on the selection of the base coordinate system which also depends on the context of the given task. From the third assumption, many visible parts in the image can be regarded as parts of the static environment. The following is an example based on this assumption.

   Nakamura and Asada [5] proposed "motion sketch" as an internal representation of vision-based mobile robot in behavior learning by optical flow. In their early stage, the action space is categorized by the direct correlation between motor commands and optical flows against the static environment. Conversely, the static environment can be discriminated as areas which can be directly correlated with motor commands.

   Without the third assumption, the discrimination has less meaning. Suppose a space robot of which base coordinate system can be set at arbitrary location in the space.

   Hosoda and Asada [6] proposed an adaptive visual servoing method which performs an on-line estimation of image Jacobian by tracking a visual target and a feedforward control of the robot arm to accomplish a given task (trajectory tracking) without any a priori knowledge about the structure of the robot arm (kinematics) or camera parameters. This means that the parts which has a direct correlation with motor commands such as the self body or the static environment can be found in a sense of that the robot can estimate the image Jacobian on it. Without the third assumption, further discrimination seems difficult from only visual information because any part of the body can be a part of the static environment (suppose a space robot, again). If the static environment has an important role for the robot to accomplish a given task, another sensation such as acceleration would be helpful to support the perception of the ground plane.

3. Passive agents: as a result of actions of the self or other agents, passive agents are moving or stop. In Asada95b, the ball is a passive agent. In our work of autonomous sensor space segmentation [4], the ball and the goal (a part of the static environment) are included, therefore the complexity of the environment can be regarded higher than a task of goal

achieving in a static environment.

Takahashi et al. [7] proposed a method of incremental (on-line) sensor space separation by which a real robot could learn to shoot a ball into a goal in one and half hours. In their method, a ball is modeled as a static environment until the robot reach it, then the ball is modeled as a part of self body because its shape and size are constant. Therefore, linear combination model can cope with it, however, the current method cannot cope with other active agents because of nonlinearity of their actions.

4. Active (other) agents: active other agents do not have a simple and straightforward relationship with the self motions. In the early stage, they are treated as noise or disturbance because of not having direct visual correlation with the self motor commands. Later, they can be found as having more complicated and higher correlation (coordination, competition, and others). The complexity is drastically increased.

As discussed above, the complexity of the interaction with environment seems to develop the internal structure inside the robot, and as a result, robot may emerge a variety of behaviors. Unfortunately, we have not proposed an exact unified design theory to realize such an internal structure. As one of the technical issues for the realization, a method for understanding other agents' behaviors is necessary in order to make clear the complicated relationship between other agents' behaviors and self ones. We have been trying to discriminate other agent's behavior struggling with the limitation of partial observation due to visual sensing (the visual angle of the robot is less than 60 degs, therefore more than five sixthes of surrounding information is lost) [8].

5 Development and Guidance

Dynamical systems approach seems very attractive as a model of development that essentially involves dynamic changes [9]. The existence of strange attractors and their changes resemble that of affordance and its organization [10]. However, the design policy for the generation of strange attractors and their control does not seem evident. Therefore, the approach seems unsatisfactory from a viewpoint of robotics research.

We have argued about the design of the internal structure capable of development. The another issue is something related to guidance. There are two kinds of guidances. One is an explicit guidance, that is, teaching what to do to accomplish a given task at each step, what's wrong, what's correct, and so on. The other is an implicit guidance, no direct teaching, but providing environments or situations so as to make the development of the robot smooth and easy.

We call such a policy "Learning from Easy Missions" (in short, LEM) to make robots learn smoothly and efficiently. In [11], we have set up situations in the order of easiness by assuming the continuity of the state space. Theoretically, the learning time which is usually in the exponential order of the size of the state space can be reduced into the linear order.

Another example of LEM is the study on how the learning agent can improve its performance by the behavior of other agents [12]. The task is to shoot a ball into a goal avoiding a goal keeper (an opponent). Intuitively, we can see the learning agent cannot learn at all if the opponent has the optimal policy to block the learner because of no success. Therefore, we started with a stationary opponent (stationary obstacle), and then increase its velocity until the maximum one of the agent. Figure 5 shows the success rates in terms of number of trials [3] with and without LEM, where solid and broken curves indicate both cases, respectively. With LEM, the agent started learning with a stationary opponent, and then with one of half speed (from the first arrow), and finally with one of the maximum speed (from the second arrow). While, without LEM, the agent starts from an opponent with the maximum speed, therefore the success rate is low, and has not converged to the level with LEM. This figure tells that LEM seems essential for the learning from other competitive agents.

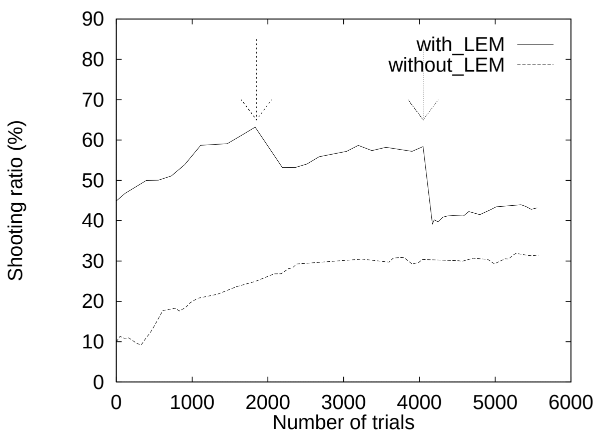


Figure 5: **Efficient learning by LEM**

[3] one trial ends when the agent succeeds in shooting or crosses over the field line

6 Concluding Remarks

Along with examples of soccer robots, I have claimed the importance of the design of the internal structure which reflects the complexity of the interactions with the agent's environment, and of the scheme of providing the agent with tasks, situations, and environments that encourage the agent to develop the internal structure.

Although the task and the environment seem simple and limited, the design of the soccer robots includes a variety of the fundamental and important issues as a standard problem in AI and robotics [13]. I expect that more agents in the field cause much higher interactions among them, which emerges a variety of behaviors.

I appreciate fruitful discussions with Dr. Koh Hosoda and other members in my laboratory.

References

[1] D. McFarland and T. Bösser. *Intelligent Behavior in Animals and Robots.* MIT Press, 1993.

[2] M. Asada, S. Noda, S. Tawaratsumida, and K. Hosoda. Vision-Based Reinforcement Learning for Purposive Behavior Acquisition. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pp. 146–153, 1995.

[3] M. Mataric. "Reward Functions for Accelerated Learning". In *Proc. of Conf. on Machine Learning-1994*, pp. 181–189, 1994.

[4] M. Asada, S. Noda, and K. Hosoda. Action-Based Sensor Space Categorization for Robot Learning. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1996 (IROS96)*, pp. ???–???, 1996.

[5] T. Nakamura and M. Asada. Motion Sketch: Acquisition of Visual Motion Guided Behaviors. In *Proc. of IJCAI-95*, pp. 126–132, 1995.

[6] K. Hosoda and M. Asada. Versatile Visual Servoing without Knowledge of True Jacobian. In *Proc. of IROS'94*, pp. 186–193, 1994.

[7] Y. Takahashi, M. Asada, and K. Hosoda. Reasonable Performance in Less Learning Time by Real Robot Based on Incremental State Space Segmentation. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1996 (IROS96)*, 1996 (to appear).

[8] E. Uchibe, M. Asada, and K. Hosoda. Agnet Classification Using Subsapce Method and Reinforcement Learning in the Environment Including Other Agnets. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, 1997 (submitted).

[9] E. Thelen and L. B. Smith. *A Dynamic Systems Approach to the Development of Coginition and Action.* The MIT Press, 1994.

[10] J. J. Gibson. *The Ecological Approach to Perception.* Houghton-Mifflin, Boston, 1979.

[11] M. Asada, S. Noda, S. Tawaratumida, and K. Hosoda. Purposive Behavior Acquisition for a Real Robot by Vision-Based Reinforcement Learning. *Machine Learning*, Vol. 23, pp. 279–303, 1996.

[12] M. Asada, E. Uchibe, and K. Hosoda. Agents That Learn from Other Competitive Agents. In *Proc. of Machine Learning Conferen Workshop on Agents That Learn from Other Agents*, pp. 1–7, 1995.

[13] H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda, and E. Osawa. "RoboCup: The Robot World Cup Initiative". In *Proc. of IJCAI-95 Workshop on Entertainment and AI/A-life*, 1995.