Artifact Intelligence: Yet Another Approach for Intelligent Robots

Hideaki Takeda
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku
101-8430 Tokyo, Japan
E-mail takeda@nii.ac.jp

Kazunori Terada
Gifu University
Yanagido 1-1, Gifu, Gifu 501-1193, Japan
E-mail terada@info.gifu-u.ac.jp

Tatsuyuki Kawamura
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, 630-0101 Nara, Japan
E-mail tatsu-k@aist-nara.ac.jp

Abstract

In this paper, we propose a new concept for intelligence called Artifact Intelligence that can be another approach to realize intelligent robots. Artifact intelligence means intelligence for artifacts that fits its embodiment, i.e., structures and functions of artifacts. Artifact intelligence differs from natural intelligence in terms of intentionality and automated objects in autonomy. In order to realize artifact intelligence, we investigate two types of affordance with the prototype robots, i.e., active affordance with the autonomous mobile chair and emergent affordance with "AgentBox".

1 Introduction

The recent years is an amazing period for robotics because really applicable robots for home and office appeared. The impact of robots like humanoid robots as ASIMO and SDR-3, and animal-like robot as AIBO seem so successful that people expect that they will live with such robots sooner or later. These robots are surely something that people image as robots different from the existing industrial "robots". But a big problem lies on such life-like robots. Although mechanism and control technologies for such robots are well developed, intelligence for them is still uncertain. Most of researchers would think that the current ASIMO should be dangerous if it would be introduced at home because of lack of flexibility in behavior. The real reason is that there is a gap between their embodiment and intelligence.

The main problem is how to build intelligence that fits embodiment. Creatures have evolved their intelligence according to the evolution of their bodies. Creatures with different complexity of bodies have different complexity of intelligence. Insects never have intelligence like human being (see path A in Figure 1). If we follow the evolution of creatures when developing robots, we should start from robots like low-complex creatures to those like more complex ones (path B in Figure 2). It can be said as "artificial life roots of artificial intelligence" [1]. In this context, behavior-based robots by Brooks [2] are very suggestive. Such behavior-based approach is strongly influenced by biology [1]. Some people blamed his robots as "cockroach intelligence" but their intelligence is enough to their embodiment. It may be the right way but it is probably very difficult because we should repeat the evolution process of creatures. The other approach is "traditional artificial intelligence" approach. Artificial intelligence has mainly concerned intelligence itself apart from embodiment until recently. Research in Artificial Intelligence gives us some of knowledge on high-level complexity of intelligence but it is not well integrated with embodiment (Path C in Figure 1).

In our opinion, both two approaches are still too complex to tackle with the current technologies. For the former, even insects are too complex in their embodiment to analyze and model. For the latter, we do not need such a high-level reasoning.

We here propose a new concept for intelligence called *Artifact Intelligence* that can be another approach to realize intelligent robots. Artifact intelligence means intelligence for artifacts that fits their embodiment, i.e., structures and functions of artifacts. Intelligence should be lower in complexity if the artifact is lower in complexity of its structure and functions (path D in Figure 1). The following two questions are essential to understand artifact intelligence.

- Is it different from intelligence in creatures?

Yes, apparently. It is clear to think intelligence as a design problem. It is impossible to specify aims of creatures, or what we can say at most is that the aim is *survival instinct*, which is too vague as specification. On the other hand, artifacts have clearly-described aims necessarily because they are reasons to be created. They are usually called as *functions* in design research [1]. Artifact intelligence is intelligence that maximizes functionality of the artifacts. It is much easier task to investigate than such vaguely tasks like instincts of creatures.

- Is it different from the existing automated artifacts?

Yes, it is different in that artifact intelligence should be autonomous for their functions. It can be seen as extension of intelligent objects. The approach is in a way similar to the traditional approach for automation, i.e., adding more active functions to artifacts. But in that approach, every

[1] They are clearly specified but the realization of functions as artifacts is another difficult tasks in design [3].

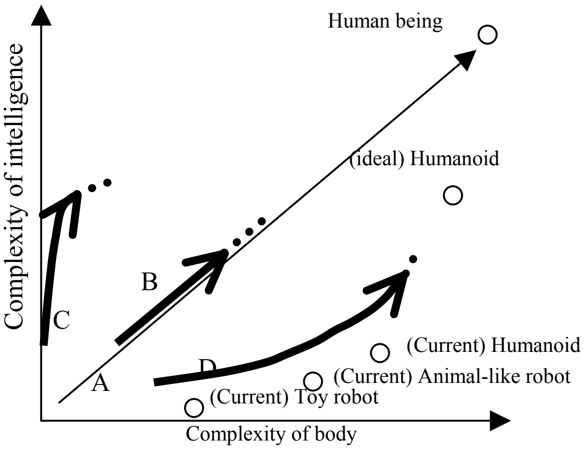


Figure 1. The approaches to intelligent robots.

behavior should be designed. We need autonomy to maintain functions of artifacts according to their structures and environments to realize intelligence for artifacts.

The benefits to pursue artifact intelligence are threefold. The first one is as foundation of intelligence for robots. As we have shown with Figure 1, robots are more appropriate as extension of intelligent artifacts rather than reflection of creatures. Accumulating various levels of artifact intelligence will contribute to intelligent robots. The second is to artificial intelligence research because thinking intelligence for various levels of complexity in bodies will contribute to nature of intelligence. The third is to design methodology for artifacts because artifact intelligence may provide new ways to design complex artifacts by easing designers' tasks.

Then the problem is how to realize artifact intelligence. We show two different but closely related approaches for it, i.e., active affordance and emergent affordance.

2 Affordance for Artifacts

The basic problem for Artifact Intelligence is to establish intentional/physical relationship between humans and artifacts. Intentional relationship means that humans use artifacts in order to archive some tasks and artifacts provide some functions to satisfy such requests. Physical relationship means that such interaction between human and artifacts should be mostly done physically. Neither intentional nor physical interaction alone is not sufficient. For example, remind how difficult it is to define "chair" physically. What artifacts can be "chair" is dependent on what is our intention on "chair".

There are two aspects for both humans and artifacts according to two aspects of interaction. Physical interaction associates directly to embodiment both of humans and of artifacts, while intentional interaction to intention of humans and role of artifacts. Figure 2(a) shows the basic elements and relations for intentional/physical relationship among humans and artifacts.

When artifacts would become autonomous, integration of intelligence and embodiment should be needed because the former is needed to solve intentional interaction and the latter to solve physical interaction.

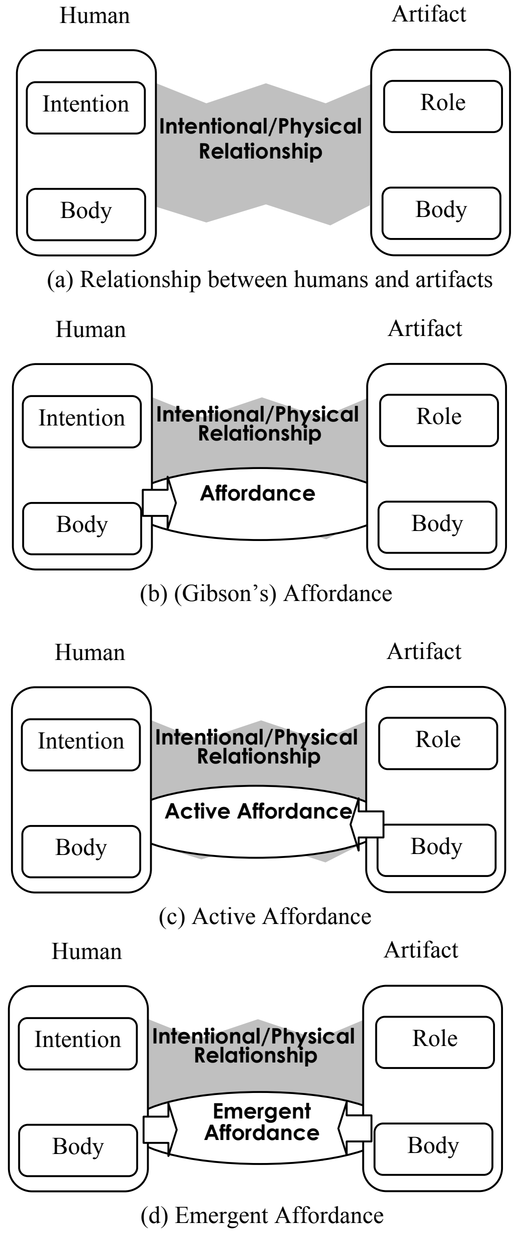


Figure 2. Realization of relationship between humans and artifacts

Although it is a very complicated problem, the concept of affordance [4] provides a good insight for this problem. Affordance refers to the possibilities for the action that available in the environment or the object, and which are is revealed by interaction between the human and the environment (see Figure 2(b)).

We extend the concept of affordance to fulfil requirements for Artifact Intelligence. The first extension is to introduce the approach from artifacts to humans. Gibson's affordance is passive for artefacts because it is uncertain until a particular interaction initiated by humans is done. We propose the concept as active affordance in which artifacts themselves participate to establish affordance (see Figure 2(c)). Active affordance is realized with actions by artifacts that can understand human intentions to them.

The next extension is to combine two approaches, i.e., one from humans to artifacts and one from artifacts to human in order to find roles of artifacts cooperatively (see

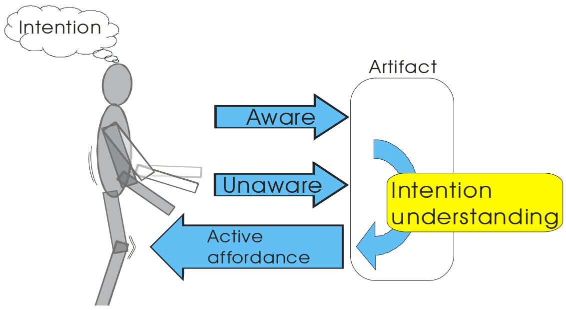


Figure 3. Communication between human and artifact

Figure 2(d)). Ability of affordance is probably innate, but affordance occurs heuristically. Human can afford new environments with establishing new relationship to the environments. Intelligent artifacts should have such ability, i.e., they should find their roles with new environments. We call this ability as emergent affordance.

In the following sections, we investigate possibilities of artifact intelligence according to the above discussion with implementation of prototype systems. We explain active affordance and active artifacts as implementation in Section 3, and emergent affordance and Agent Box concept as implementation.

3 Active Affordance

Active affordance is realized by artifact itself. The concept of affordance was introduced by the psychologist J.J.Gibson [4] As mentioned, affordance refers to the possibilities for the action available in the environment or the object, and it is revealed by interaction between the human and the environment. For example, the affordance of a chair is that it allows to a human to sit on the chair, but is not manifested until the human generates the action of sitting down. While affordance is considered to be realized by the user's action[5], active affordance is realized by the artifact's action.

The key issues of active affordance are as follows.

1) The communication between human and artifact. The artifact should comprehend the user's intention so that it can afford its functions appropriately and timely.
2) Embodied interaction between human and artifact. The artifact should have control strategies for realizing affordance that strongly depend on the embodiment relation.

3.1 Communication between Human and Artifacts

Figure 3 shows our model of human-artifact communication and the concept of active affordance we propose. The communication can be divided into the following parts.

(1) Conveyance of intention from the human to the artifact which includes
 (1-a) aware communications and
 (1-b) unaware communications, and
(2) realization of functions from the artifact to the human.

While human-human communication is bi-directional, human-artifact communication is mono-directional. The purpose of human-artifact communication

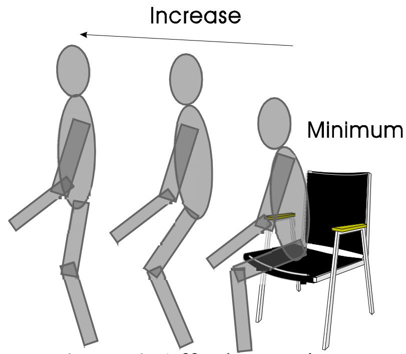


Figure 4. Affordance Distance.

is to convey human intention form human to artifact so that the artifact affords its function appropriately and timely. A human has an intention when she/he is going to carry out some task. The intention is conveyed to the artifact through the communications channel. There are two modes for the communications channel; 1) aware communications, the means of which include natural language, sign language, and gesture and 2) unaware communication channel, which refers to nonverbal behavior.

Aware communications

If an artifact has a modality for aware communications, a human is able to use the channel for aware communications by utilizing the methods for human-human communication. However, the communications protocol should be predefined and the system needs a large database of vocabulary. In our implementation described in Section 3.3, we use gesture as one means for aware communication. We thus compare the user's motion with a set of predetermined gesture pattern.

Unaware communications

Some psychological researchers have concluded that more than 65 percent of the information exchanged during a face-to-face interaction between humans is expressed nonverbal [6].

The unaware communications channel is important for human-human communication. Cassell pointed out that speech and nonverbal behavior join together to convey the communicative intent of a speaker[7]. The unaware communications channel should also be important in human-artifact communication because people sometimes fail to find affordance with artifacts even the affordance is intended for the artifacts. The reasons for this failure are as follows.

- The function of the artifact is unknown.
- A user does not know how to use the artifact (the user is, however, conscious of his intention).
- A user is not conscious of his intention.

The agent can utilize user's unconscious motion for intention comprehension. The unconscious motion is always generated when a human manipulates some artifact, and it varies according to the artifact's physical properties and functions. The artifact might take advantage of the peculiarities of the various forms of motion to detect the user's intention. We can use following heuristics to realize the unaware communication.

- Physical contact always occurs in object manipulation, and indicates a critical state.

Figure 5. Autonomous Mobile Chair.

- The distance between the surfaces of user's body and the artifact is reduced by the action of reaching.

This reduction of distance indicates that the user intends to manipulate the given artifact.

3.2 How to afford the functions

Once the intention has been 'understood' by the artifact, it should produce behavior that is appropriate in response to the intention. In this section, we describe the method used to generate such behavior.

Affordance Distance

In the following sub-section, we describe an architecture, which allows an artifact to realize its function autonomously. Firstly, we consider a method of describing the functional relation between a human and an artifact based on the concept of affordance.

As was mentioned by Gibson, a physical relation and, in particular, relation of surfaces between the human's body and the environment is important for the human's behavior in physical world. Such a relation is relative to the human's body as a standard. In order to describe the relation between the surfaces of a human's body and the surfaces of the object (artifact), we introduce the concept of *affordance distance*.

Affordance distance has a value and defined in the following way.

- Affordance distance will be as its minimum value at the end of an action sequence, i.e., a tactile state.
- Affordance distance increases as the agent has longer distance from the tactile state.
- Affordance distance is defined between a point on artifact's body and a point on the human's body.

For example, in the case of the chair drawn in Figure 4, the affordance distance will be minimum value when the human is sitting, and will increase as the human goes away the chair. Affordance distance is not Euclidean but corresponds to the cost of the action, which is required of the artifact when it goes to contact with human. That is because the optimal action path that minimizes the distance between two points is not necessarily the shortest distance. The optimal action path depends on the locomotive ability of the artifact and the relative angle between the two surfaces.

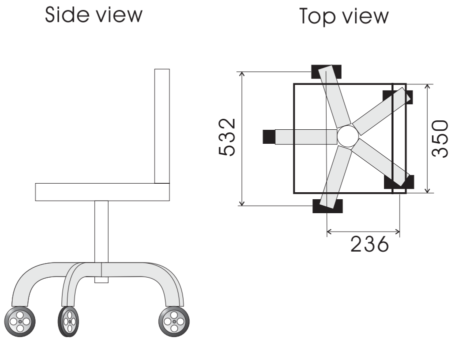


Figure 6. The model of the autonomous mobile chair

Calculation and minimizing of the affordance distance

From our viewpoint, the purpose of an artifact is to search for and move to a state where the affordance distance is at its minimum. Then we need a method for calculating the affordance distance and the method for controlling the artifact. We employ a utility function to express the affordance distance. The utility function is a widely used in a research in autonomous agents[8]. The utility function is capable of representing the distance to the goal state, considering the locomotive ability of agent. Details of calculation are found in [9].

Once the utility function is obtained, it can be used to estimate affordance distance with any start and goal situations.

3.3 Experimental results

To show the validity of our proposed method, we performed experiments using a computer simulation and a real robot. In this section we describe our experimental system and results.

Autonomous mobile chair

We have built an autonomous mobile chair as an example of the active artifact. The purpose of the autonomous mobile chair is to reach its back reclining on human's back. We remodeled some parts of an aluminum chair to allow it to move around (See Figure 5).

The chair has five legs, radiating in every-spaced directions, and each leg has a caster, which freely rotates in the horizontal plane. In our system, we replaced two of the casters with powered wheels, each of which is fixed to the leg. The autonomous mobile chair is equipped motion-capture system made by Ascension Technology, which enables measurement of the position and orientation of the chair's body. The motion-capture system employs pulsed-DC magnetic-field transmission and sensing technology to measure the positions and orientations of miniaturized sensors that are attached to the measuring equipment. The autonomous mobile chair is controlled by a Linux PC to which it is connected via RS232C cable. A subject in this experiment also has to carry a motion sensor so that the autonomous mobile chair is able to determine the reaching point.

Modeling of the autonomous mobile chair and state space

In order to perform computer simulation, the environment and the autonomous mobile chair ware modeled in the following way. The environment is a floor

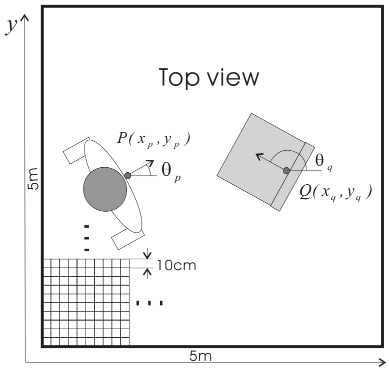
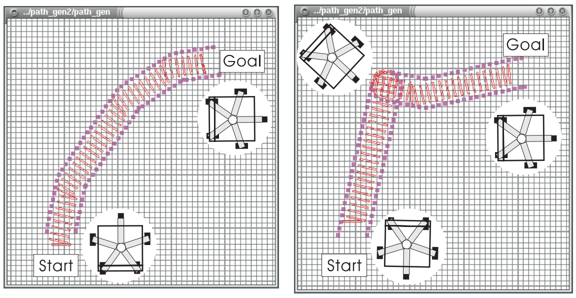


Figure 7. The model of the environment and state space



(a) Start: $(10,10,-\pi/2)$ (b) Start: $(10,10,\pi/2)$

Figure 8. Generated paths

of 5 m square. The size of the chair and the arrangement of its wheels are as shown in figure 6.

The state space is constructed in the following way (see figure 7). To simplify the simulation, we assume that the dimension of two for both the autonomous mobile chair and the environment, that is, height is disregarded and the artifact is able to only move in the 2D plane. As a result, the dimension of the state space is three; (x, y, θ). The floor is divided into a 50×50 grid. The angle of the normal vector of the surface is discretized into 16 steps. As a result, the discrete state space has $50\times50\times16$ states. We also assume that we are able to send six action commands to the autonomous mobile chair. The commands are executed by specifying the speeds of the motors. The commands are *A1(-V, V)*, *A2(0, V)*, *A3(V, V)*, *A4(V, 0)*, *A5(V, -V)*, and *A6(-V, -V)*. In this experiment, *V* is *0.3m/sec*. We define the action unit as the segment of the artifact's motion that precedes the observation of a change of state.

Calculation of utility value

Firstly, we calculated the transition probability model. Each of the six action commands is executed from 100 uniform points in a grid. This operation is performed for each of the 16 angles. Next, the utility function is calculated by using transition probability model. The goal point, i.e., the point of contact, is set as *(40, 40, 0)*. The reward given for reaching the goal point has a value of 1.

Reaching the goal

To show the validity of the calculated utility function, we carried out two cases of reaching experiments in computer simulation. One of the cases, i.e., case (a), is that where the state of starting point is $(10,10,-\pi/2)$ while in the other case, i.e., case (b), the state is $(10,10,\pi/2)$. Figure 8 shows the paths generated in the two cases. In case (a), the autonomous mobile chair changes direction gradually as it approaches the goal. In case (b), the autonomous mobile chair reverses once, changes direction, and then proceeds to the goal. In both cases, the autonomous mobile chair generates an appropriate path and reaches its goal.

Figure 9. Intention understanding and behaviour generation

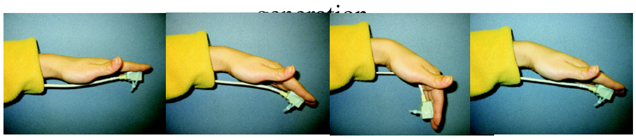


Figure 10. Beckoning gesture

Comprehension of intention and generation of behavior

We conducted experiment for comprehension of intention and behavior generation. In this experiment we use aware communication channel, i.e., gesture. Figure 9 shows snapshots of our experiment in which a subject calls a chair with beckoning gesture (see figure 10) and the chair moves to the subject. Gesture recognition was performed by utilizing simple rules described as follows. If the following two conditions are satisfied, beckoning gesture is detected.

- Hand height is between 110cm to 150cm.
- Angular velocity of finger is greater than 7rad/sec.

The gesture is captured using motion capture system.

4 AgentBox

AgentBox is a hardware that can discover its role through interaction with human [10]. Although conventional artifacts are given their roles by designers, e.g., a "chair" as an artifact to sit down, future intelligent artifacts will find their roles that are possible with their bodies and also needed by humans. We propose *emergent affordance* as a novel interaction methodology between humans and artifacts to find their roles.

4.1 Emergent Affordance of AgentBox

A mechanism of emergent affordance is shown in Figure 11. In the methodology of emergent affordance, an agent can decide its role through actions to the artifact. An artifact driven by emergent affordance is just like an anorganic substance before a user decides a certain intension for the artifact. When the user supposes a specific function of an artifact and does the suitable action

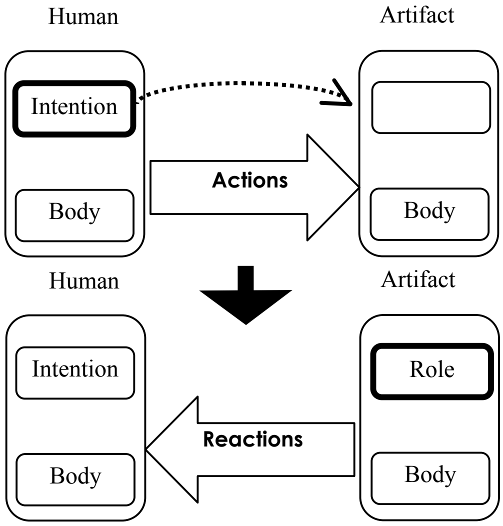


Figure 11. The process of emergent affordance

to the artifact, the artifact hypothesizes its role from the user's action, changes a mode of itself and then reacts to the user.

4.2 A Mixed Real Society

Interaction between human and artifacts is not restricted to one to one, but should be many-to-many. We propose that a *mixed real society* among humans and agents. For realizing such a socialized world, we add a communication function to the design of AgentBox for transporting other agents. Conventional artifacts are hardly involved with each other without semantic relations. Figure 12 illustrates the mixed real society that consists of various human and various agents which live together in the real world and interact to each other. We can divide the interaction of the society into the following three elements: 1) Person interacts with an agent, 2) One agent works together with other agents in a cooperative task, and 3) Human-Human interaction can be operated via the agents.

We introduce the following simple examples for explaining the idea of the mixed real society.

[Example 1: a desk and a desk lamp]

Suppose a desk agent that hypothesizes expected roles when she/he is writing something on the agent. The desk agent infers its expected role, i.e., as a writing desk, and then generates and performs a task, e.g.. keeping the top flat and stable. Furthermore, the desk agent transports the information of the user's purpose to a desk lamp agent that is located on the desk agent, and then the desk lamp agent decides its role. The desk lamp agent turns the light on at last. Or the lamp agent turns the light off, when the user sleep on the desk agent which is supposed as a pillow by her/him, i.e., its may make the top soft or keep its temperature good for sleep if possible. We consider that other tasks also use this type of interaction. For example, when the user is looking for a lost book around the desk, the desk agent and a bookshelf agent can cope with the task together.

[Example 2: a whimsical AgentBox orchestra]

A user, who is sitting at a chair agent in her/his room, is hitting a desk agent rhythmically. The desk agent then decides its role and does reactions as an interactive

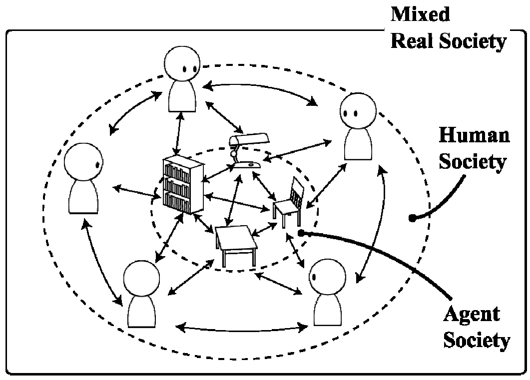


Figure 12. Socialized Agents in the real world

percussion. The desk agent transports the information of user's rhythm to other agents. The chair agent joins with the interaction if the emotional condition of the chair agent is good and has no other task at that moment. Furthermore, some of other agents would synchronize the interaction. By increasing of agents synchronizing the interaction, a whimsical AgentBox orchestra might be organized occasionally. Understandably, other users in the room are able to join to the interaction by hitting a certain agents, and a new collaborative music is generated.

4.2 A-Box : A Prototype Artifact using AgentBox

We have developed a prototype artifact that is augmented by AgentBox. We named the artifact A-Box. Figure 12 illustrates an appearance of A-Box. A-Box is shaped a 50cm cube in which a PC is supended with wire. Its frame is made of aluminum blocks. Its surface is surrounded with acrylic boards. Touch panels which are made by MicroTouch are attached to inside of the surface. A-Box, the current implementation of AgentBox, can accept only the very basic user's actions, i.e., touching, hitting, and stroking by using the touch panel. It has only a signle output method, i.e., making sound with the implanted speaker in the PC of A-Box.

AgentBox Component (ABC) for anthropomorphising of artifacts consists of the following basis modules : 1) Perceptual Module that can sense a input signal of user's touch for interaction to A-Box, 2) Activity Module that controls an activity of A-Box which represents awakening and sleeping, 3) Character Module that sets certain temperamental parameters in order to generate various kinds of reactions in the same reaction, 4) Emotion Module that can raise dynamic behavior of A-Box per the user's interactions, 5) Communication Module that can exchange current states of A-Box selves, 6) Action Module that is a set of primitive reactions for realizing emergent affordance, and 7) State Management Module that gives A-Box a temporal role in a person-artifact interaction by using an adaptive selection of a certain primitive reaction.

Parameters in Emotion and Charater Modoles are designed to realize various interpretations on user inputs. There are enormous possibilites how inputs are interpreted. We should restrict a small set of intepretation for practical reasons but do not want to fix them. So these parameters are used as control of mapping from inputs to a limited set of interpretatoin. Parameters on Emotion Modules varies

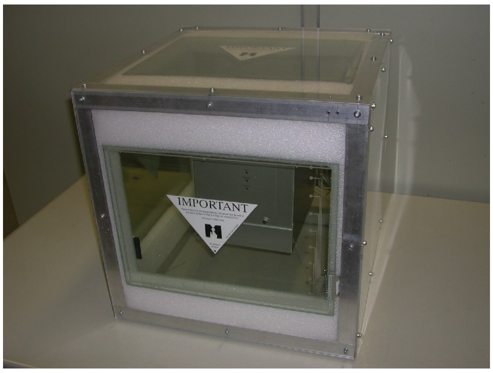


Figure 13. An appearance of A-Box

according to time, while those in Charater Modules are constant in a single artifact.

As we mentioned, A-Box has the very limited methods for interactions. We consider here how users and artifacts can interact to each other with even such a condition. We have installed the following two primitive reactions into A-Box. These reactions are relaxation chair and interactive percussion. Both of them employed a musical reaction. Relaxation chair uses an existing music, and interactive percussion generates tune dynamically.

[Relaxation chair]

A-Box decides its role to the relaxation chair when the user takes the weight off his/her feet to A-Box during a certain time. A-Box weaves various existing melodies via a condition of emotion value.

[Interactive percussion]

A-Box becomes the interactive percussion when the user hits A-Box rhythmically[2]. The interactive percussion has the following three functions.

1) The user can freely decide a rithm (1/4, 1/8, 1/16) and a tempo. Any decided tempos are, however, recognized within the predifined range of tempo from 61 to 239.
2) A-Box makes a tune with MIDI sound souces to the decided rithm and tempo, and sounds the tune.
3) A-Box can dinamically cope with an out of tempo between the user and A-Box.

User and A-Box can make more deep communication by using a phenomenon of entrainment [11] when he/she plays the interactive percussion.

5. Conclusion

In this paper, we propose a new aspect to investigate intelligence called Artifact Intelligence. It is a research methodology to realize intelligent artifacts from the very primitive objects to complicated robots. We focus on intentional/physical relationship between humans and artifacts, and raise a new concept, *affordance for artifacts*, i.e., active affordance and emergent affordance. We are investigating these ideas with implemented systems. The current research status is still a just starting point, but we believe that our approach will open the new way for intelligence.

Acknowledgement

We thank all members of "Agent Box" project including Takahiro Ueoka, Tomohiro Fukuhara, Masaki Chikama, Kingo Koshiishi, and Masahiro Hamasaki. This paper includes research supported by Synsophy Project at Communications Research Laboratory (CRL). We thank Prof. Toyoaki Nishida at the University of Tokyo and CRL.

References

[1] Steels, L. *The artificial life roots of artificial intelligence*. Artificial Life, 1(1):75--110, 1994.

[2] Brooks, R. A., *New Approaches to Robotics*, *Science* (253), September 1991, pp. 1227–1232

[3] Umeda Y., Takeda H., Tomiyama T., and Yoshikawa Y., *Function, behaviour, and structure*. In J.S. Gero, editor, Applications of Artificial Intelligence in Engineering V, volume 1, pages 177-194. Springer-Verlag, Berlin, 1990.

[4] Gibson, J. J. *The Ecological Approach to Visual Perception*. Houghton Mifflin Company. 1979

[5] Norman, D. A. *The psychology of everyday things"*, Basic Books Inc. 1988

[6] Argyle, M. *Bodily Communication*. Methuen & Co. 1988.

[7] Cassell, J. Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In Cassell, J.; Sullivan, J.; Prevost, S.; and Churchill, E., eds., *Embodied conversational agents*. The MIT Press. chapter 1, 1–27. 2000.

[8] Sutton, R. S., and Brato, A. G. *Reinforcement Learning*. The MIT Press. 1998.

[9] Terada K. and Nishida T., *Active artifacts: for new embodiment relation between human and artifacts*. In *Intelligent Autonomous System 7 (IAS-7)*, pp. 333-340, 2002.

[10] Kawamura T., et *al.*, AgentBox: a Hardware for a Novel Interaction between Human and Artifacts, 16th Annual Conference of JSAI, 2002 (in Japanese).

[11] Watanabe K., Okubo M., *Physiological Analysis of Entrainment in Communication*, *IPSJ Journal*, vol. 39, no. 5, 1998, pp. 1225-1231. (in Japanese)

[2] The other possible interpretation is that the user is irritated. If A-box interprets so, it may make itself "relaxation chair".