Guidelines for Designing Interactions Between Autonomous Artificial Systems and Human Beings to Achieve Sustainable Development Goals

Muneo Kitajima
Nagaoka University of Technology
Nagaoka, Japan
Email: mkitajima@kjs.nagaokaut.ac.jp

Makoto Toyota
T-Method
Chiba, Japan
Email: pubmtoyota@mac.com

Jérôme Dinet
Université de Lorraine, CNRS, INRIA, Loria
Nancy, France
Email: jerome.dinet@univ-lorraine.fr

***Abstract*—Human beings live in an environment that consists of various artifacts, such as physical or virtual tools, information systems, and social systems. With IT advancement, the wider the network of artifacts, the more autonomous they become. However, the ultimate goal of developing these artifacts is to achieve the Sustainable Development Goals (SDGs) through the exploration of the design space for realizing a sustainable society. The artifacts that human beings interact with apply this mechanism for utilizing the artifacts, by selecting the subsequent actions for a given situation. This mechanism includes Perceptual, Cognitive, and Motor (PCM) processes and the memory process. The cognitive process is characterized by the bounded rationality and by the satisficing principle proposed by Simon, and Two Minds of unconscious and conscious processes proposed by Kahneman. The state-of-the art cognitive architecture, Model Human Processor with Realtime Constraints (MHP/RT), developed by Kitajima and Toyota, defines these processes as autonomous systems and proposes a resonance mechanism between the PCM and memory processes. The purpose of this study is to propose guidelines to conduct strategical explorations in design space. Based on the simulation of human–artifact interaction processes through the MHP/RT cognitive architecture, the guidelines are grouped into three levels: goal, mode, and process levels. Moreover, hints are provided for applying the proposed guidelines to narrow down the design space.**

***Keywords*—Design guidelines, Human–artifact interaction; Autonomous systems; Cognitive architecture; MHP/RT.**

I. Introduction

This paper is based on the previous work originally presented in COGNITIVE 2022 [1]. It extends the concepts described in Section IV by providing a new Section IV-A.

When viewed as an individual, *each human being is composed of autonomous systems* that control perception, cognition, and movement in synchronization with changes in the environment, in addition to a memory autonomous system that works to link perception and movements [2][3]. The environment in which humans interact and live is composed of various artifacts. With the progress of networking technology, a large number of artifacts have become related to each other, overcoming the physical constraints of time and space. In this case, the central management method of the set of artifacts and the environment design to achieve their goals is not effective. It would rather be effective *to design the environment as a set of autonomously operating artifacts equipped with Parallel Distributed Processing (PDP)*, which can be referred to as the Artificial PDP (A-PDP) system, and to design them so that they function as a whole and achieve their goals.

There exist interfaces between the above-mentioned autonomous systems, which have to be properly designed for well-being. The interfaces and internal algorithms defining their behaviors must support activities conducted by human beings; they attempt to achieve their happiness goals by utilizing artifacts. A research question that arises is – *how can such interfaces and internal algorithms be designed for the two autonomous systems?* Our daily life is based on interactions with a wide variety of artifacts. The purpose of interactions, for human-beings, is to achieve well-being through activities in domains such as health (e.g., bio-monitoring), mobility (e.g., driving an electric vehicle), education (e.g., learning on Massive Open Online Course (MOOC)), and entertainment (e.g., playing e-sports). The artifacts support human activities through the interface at each moment of interaction. There are multiple autonomous systems on both sides of the interface with complex relationships. The purpose of this study is to propose a set of guidelines that should be applied when designing the interfaces of autonomous artifacts, for supporting activities carried out by autonomous human beings. Traditionally, designers adopt a top-down or a bottom-up approach. It is advised to combine both approaches while designing; the proposed guidelines are useful for this task. The top-down design would be implemented without deviating from the objective by following the guidelines; the bottom-up design would be facilitated by observing user behavior to ensure that the design is in line with the guidelines.

The remainder of this paper is organized as follows. Section II outlines the human-artifact interaction to define the specific perspective for considering the complex situation of interaction, i.e., both sides are autonomous systems. Section III briefly reviews the Model Human Processor with Realtime Constraints (MHP/RT), developed by Kitajima and Toyota [2][3] and defines a framework for developing the guidelines. A set of guidelines are described. In general, guidelines are intended to provide direction for designing artifacts, not to indicate how to proceed with the design according to the specific guidelines provided. Therefore, designers cannot immediately use them in their design activities. To address this issue, Section IV introduces some hints that will be useful in designing interactions according to the guidelines given in Section III.

II. Human–Artifact Interaction

There are human beings on this side of the interface and artifacts on the other side. From the viewpoint of a user that perceives the information provided on an interface to select the next action for accomplishing his/her goal, a complete understanding of the detailed processing of an artifact to generate information on the interface, e.g., the knowledge of implementing the internal algorithms, is unnecessary; similarly, a detailed understanding of the internal processing of an input to an artifact is unnecessary for them to continue the interaction cycle of execution and evaluation. Although the internal processes are not known to the user, s/he has to comprehend the mechanism at the interface level in order to proceed, i.e., "bridging the gulf between execution and evaluation [4, Figure 3.2]." This also applies to the artifacts. For designers to specify the interfaces of the I/O for the systems by developing internal algorithms to support human activities, there is no need for them to have a complete understanding of human reactions to the output of the artifacts and of human expectations attached to the input to the artifacts.

As Simon pointed out [5], an interface is characterized by an artificial system between two environments – inner and outer, i.e., human beings and artifacts, respectively. These environments lie in the province of "natural science" where the systems of artifacts and human beings are the focus of research, but the interface linking them is the realm of "artificial science." Therefore, the research question that this study addresses is in the realm of artificial science. The two sides, i.e., the behaviors of human beings and artifacts, are governed by their own principles, and they have to interact with each other by simultaneously considering the behaviors of the either side at *the appropriate approximation levels* in hope of a successful development. Their articulation could be formalized as guidelines, which is the form of an answer to the research question that this study addresses.

The interface between the two systems can be conceived from a variety of perspectives or dimensions. One of them is the dimension that focuses on the Perceptual, Cognitive, and Motor (PCM) processes and the manner in which memory is acquired, used, and developed in the use of artifacts. This study specifically focuses on the ongoing PCM processes and the manner in which they use the memory in the human–artifact interaction process. Our previous study [6] focused on the acquisition and development process, and proposed guidelines for designing artifacts, which could cause the evolution of artifacts. In the process of evolution, the *techniques* used in the development of artifacts are received and absorbed by users as *skills* by applying the PCM and memory processes, which is simulated by MHP/RT. The techniques could turn into skills if the conditions derived by the simulations based on MHP/RT are satisfied. When this spiral evolution occurs, the socio-cultural ecology, wherein the artifacts are embedded, evolves to exhibit a splicing evolution. The focus of this paper is not on the evolution that occurs at the interfaces but on the ongoing events.

Another dimension, which this study effectively focuses on, is the structure of human goals, which can be used by human beings to organize their behaviors. Our previous paper [7] proposed an effective method for achieving Sustainable Development Goals (SDGs) through the behavior of individual human beings, by applying the knowledge of cognitive science; the idea is to connect the daily activities of human beings when trying to accomplish task goals through real world problem-solving [8], i.e., activities in the COGNITIVE Band of Newell's time scale of human action [9, page 122, Fig. 3-3], through any of the SDGs, that concerns social ecology and resides in the SOCIAL Band by finding the non-linear mappings between the goals in different bands. The interfacing situation this study deals with is analogous to the one above. Each individual human being conducts activities to accomplish his or her behavioral goal. This activity is non-linearly mapped onto the autonomous artifacts, which have the goal of any of the SDGs, where the gulfs of execution and evaluation have to be bridged.

III. Interaction Levels and Guidelines

Figure 1 shows the top-level view of the human–artifact interaction. The artifacts placed at the center should exist as entities for achieving any of the SDGs by providing appropriate support for the individual human beings who try to achieve any of the seventeen happiness goals. This section begins by introducing MHP/RT [2][3] in Section III-A focusing on the levels of interactions with artifacts. It follows Section III-B and Section III-C with suggestions for enabling conditions that artifacts have to satisfy to help human beings achieve a smooth coordination between System 1 and System 2. Section III-D presents the relationships between the happiness goals of human beings and the SDGs that the artifacts are expected to achieve. The top-level constraint for developing guidelines is that any artifact that complies with the guidelines has to provide a stable human-artifact interaction; unstable interactions should result in unpredictable results, which do not come with the SDGs.

A. MHP/RT and Interaction Levels

Kitajima and Toyota [2][3] constructed a comprehensive theory of action selection and memory, MHP/RT, that provides a basis for constructing any model to understand the daily behavior of human beings. MHP/RT is an extension of the Model Human Processor (MHP) proposed by Card, Moran, and Newell [10], which can simulate routine goal-directed behaviors. MHP/RT extends the MHP by the following assumptions to consider the fact that the processes involved in action selection are a dynamic interaction that evolves in the irreversible time dimension:

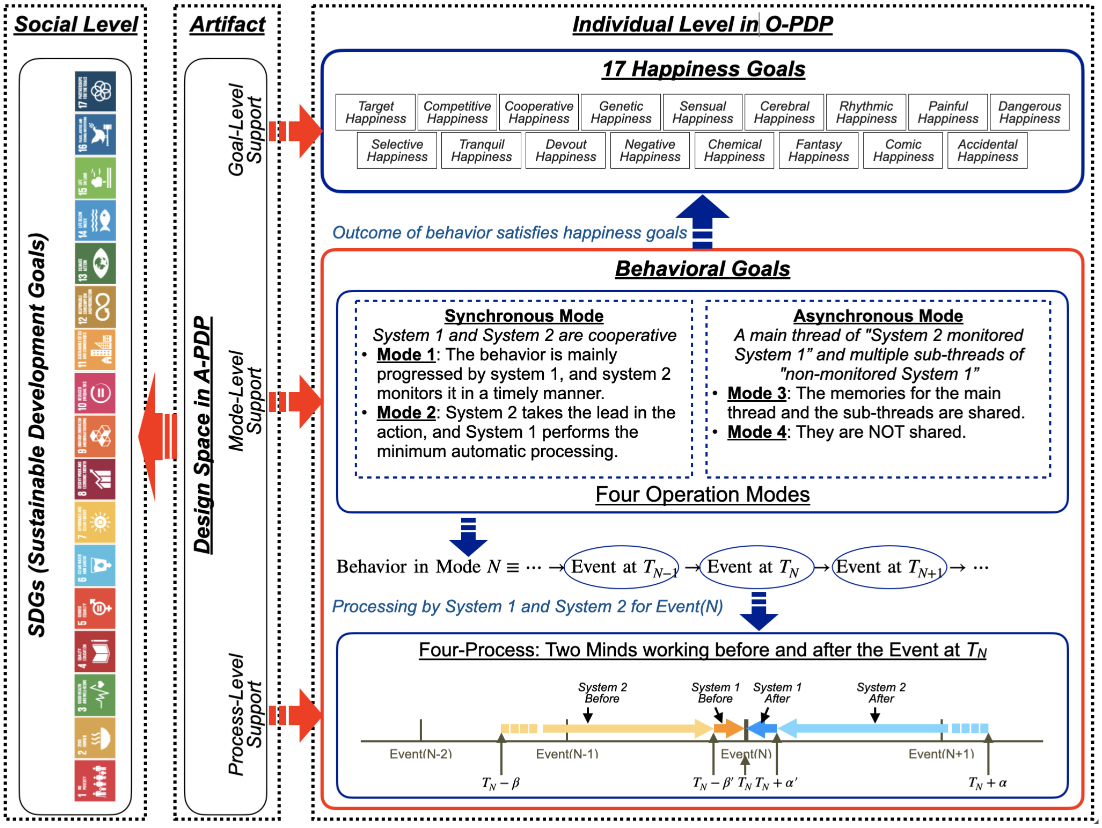
Social Level
SDGs (Sustainable Development Goals)
Artifact
Design Space in A-PDP
Goal-Level Support
Mode-Level Support
Process-Level Support
Individual Level in O-PDP
17 Happiness Goals
Target Happiness
Competitive Happiness
Cooperative Happiness
Genetic Happiness
Sensual Happiness
Cerebral Happiness
Rhythmic Happiness
Painful Happiness
Dangerous Happiness
Selective Happiness
Tranquil Happiness
Devout Happiness
Negative Happiness
Chemical Happiness
Fantasy Happiness
Comic Happiness
Accidental Happiness
Outcome of behavior satisfies happiness goals
Behavioral Goals
Synchronous Mode
System 1 and System 2 are cooperative
• Mode 1: The behavior is mainly progressed by system 1, and system 2 monitors it in a timely manner.
• Mode 2: System 2 takes the lead in the action, and System 1 performs the minimum automatic processing.
Asynchronous Mode
A main thread of "System 2 monitored System 1" and multiple sub-threads of "non-monitored System 1"
• Mode 3: The memories for the main thread and the sub-threads are shared.
• Mode 4: They are NOT shared.
Four Operation Modes
Behavior in Mode $N \equiv \cdots \rightarrow$ Event at $T_{N-1} \rightarrow$ Event at $T_N \rightarrow$ Event at $T_{N+1} \rightarrow \cdots$
Processing by System 1 and System 2 for Event(N)
Four-Process: Two Minds working before and after the Event at T_N
System 2 Before
System 1 Before
System 1 After
System 2 After
Event(N-2)
Event(N-1)
Event(N)
Event(N+1)
$T_N - \beta$
$T_N - \beta'$
T_N
$T_N + \alpha'$
$T_N + \alpha$

Fig. 1. Top-level view of human–artifact interaction. (adapted from [1])

1) The fundamental processing mechanism of the brain is PDP [11], which leads to a collection of the autonomous systems having specific functions for generating an organized human behavior. It consists of autonomous systems for perception, cognition, motor movements, and memory, which collectively form the Organic PDP (O-PDP) system in the development of MHP/RT.
2) Human behavior emerges as a result of the cooperation of the dual processes of System 1, i.e., fast unconscious processes for intuitive reaction with feedforward control, which connect perception with motor movements, and System 2, i.e., slow conscious processes for deliberate reasoning with feedback control. System 1 and System 2 are referred to as Two Minds [12].
3) Human behavior is organized under 17 happiness goals [13].

Human beings use artifacts to accomplish certain behavioral goals for realizing the desired state of affairs. The human–artifacts interaction is a cycle of PCM processes. The MHP/RT simulates the PCM processes as follows. The cognitive process is to select the next actions that are appropriate for accomplishing the behavioral goals, given the comprehension results of the perceived information. System 1 directly connects to the motor process, whereas System 2 can only indirectly affect the motor process via System 1. The MHP/RT assumes a resonance mechanism for connecting the PCM processes and memory, where the records of the results of the PCM processes are accumulated in a layered and partially overlapped structure of multidimensional memory frames. The cognitive process is carried out by coordinating System 1 and System 2 appropriately to accomplish the behavioral goals. System 1 and System 2 interact simultaneously with the multi-dimensional memory frames to select an appropriate action and carry it out in a timely manner in the ever-changing environment. The former is the issue of coordination, while the latter is that of synchronization. Section III-B and Section III-C will address these issues.

B. *Mode Level: Coordination of Two Minds According to the Goals*

Individual beings interact with artifacts to accomplish their behavioral goals by selecting appropriate actions, by running a cycle of PCM processes. The MHP/RT assumes that the action selection processes are controlled by System 1 and System 2. System 1 and System 2 cooperate to connect perception with motion, and the degree of cooperation varies depending on the external environmental conditions, i.e., the state of the artifact that the MHP/RT is interacting with.

1) Four Operation Modes: The conduction of the cooperation can be understood by observing the interaction processes for a certain amount of time, to identify the feature of the interaction in terms of the mode. The processes carried out

TABLE I
FOUR OPERATION MODES OF MHP/RT. (ADAPTED FROM [1])

Synchronous Modes	
- Mode 1:	Unconscious mechanism driven mode *A single set of perceptual stimuli initiates feedforward processes at the BIOLOGICAL and COGNITIVE bands to act with occasional feedback from an upper band, i.e., COGNITIVE, RATIONAL, or SOCIAL.*
- Mode 2:	Conscious mechanism driven mode *A single set of perceptual stimuli initiates a feedback process at the COGNITIVE band, and upon completion of the conscious action selection, the unconscious automatic feedforward process is activated at the BIOLOGICAL and COGNITIVE bands for action.*
Asynchronous Modes	
- Mode 3:	In-phase autonomous activity mode *A set of perceptual stimuli initiates feedforward processes at the BIOLOGICAL and COGNITIVE bands with one and another intertwined occasional feedback process from an upper band, i.e., COGNITIVE, RATIONAL, or SOCIAL.*
- Mode 4:	Heterophasic autonomous activity mode *Multiple threads of perceptual stimuli initiate respective feedforward processes at the BIOLOGICAL and COGNITIVE bands, some with no feedback and others with feedback from the upper bands, i.e., COGNITIVE, RATIONAL, or SOCIAL.*

by System 1 and System 2 are independent for some time durations but are totally dependent on each other in other domains. This provides a macroscopic view of the manner in which the human–artifact interaction is organized.

Four qualitatively different modes are identified [14]. System 1 is a fast feedforward control process with the characteristic time range of <150 ms to connect the perceptual process with the motor process, which makes it possible to behave synchronously with the ever-changing environment. There could be multiple System 1 processes that correspond to active perceptual–motor controls. However, System 2 is a slow feedback control process, which takes a significantly longer time. The time range can be months or years as long as feedback from the past event could affect the ongoing processing. System 2 is a serial process. It can process only one thing at a time; the process could be monitoring one of the active threads of System 1 to check for possible deviations of the results of System 1 from the expected course of actions.

Table I lists four modes, each of which is characterized by the relationships between System 1 and System 2. Modes 1 and 2 are characterized by a single major System 1 process monitored by System 2. The differences between them is the degree of intervention of System 2 for checking the output of System 1. In Mode 1, the occasional feedback from System 2 is sufficient to conduct the behavior. In Mode 2, a frequent monitoring is necessary to organize the behavior appropriately in the environment. Mode 3 corresponds to the situation wherein a single set of perceptual stimuli initiates System 1 processes with one and another intertwined occasional feedback processes by System 2. Mode 4 corresponds to the situation where multiple threads of perceptual stimuli initiate the corresponding System 1 processes, some with no feedback and others with feedback from System 2.

2) Guidelines for Supporting Mode Level Interactions: The human–artifact interaction is carried out in one of the four operation modes of MHP/RT. For the viewpoint of a sound human–artifact interaction, the artifacts should support the interactions that are carried out in Mode 1 and Mode 2. Mode 3 and Mode 4 include unmonitored feedforward System 1 processes, which might cause an instability in the human–artifact interaction. The safety of the human–artifact interaction is realized by allowing the artifact to intervene through System 2, causing the human being to restore to the normal interaction. In other words, the resilience of the human–artifact interaction is realized by maintaining the interaction in Mode 1 and Mode 2. Achieving resilience is a necessary condition for the sustainability of the artifacts to achieve the SDGs.

a) Supporting Mode 2 Interaction: In Mode 2, System 2 frequently intervenes the PCM processes conducted by System 1. More precisely, the pace of interaction with the artifact is controlled by System 2. The role of System 1 is to carry out the necessary PCM processes, to advance the main System 2–artifact interactions. Because System 2 operates on language, the appropriate input from the artifact by means of language is of critical importance. Because the processes of System 1 are carried out in the context defined by those of System 2, the appropriate interactions from the artifact for supporting the processes of System 1 have to be provided considering the context.

Guideline [A]

1. *Converse with System 2.*
2. *Intervene System 1 for facilitating the main conversation with System 2.*

b) Supporting Mode 1 Interaction: In Mode 1, where the intervention of System 2 is weak, language is not an appropriate medium for communication. The interaction from the artifact has to support the unconsciously carried out automatic processes by System 1. However, in Mode 1, the timely examination of the progress is critical for a smooth interaction. The triggers for initiating the examinations carried out by System 2 could be provided internally or externally, i.e., from the artifact. There could be a situation where the examination by System 2 has not been carried out when necessary. In this situation, the intervention from the artifact is necessary for maintaining Mode 1 interaction.

Guideline [B]

1. *For a normal Mode 1 interaction, provide information to both System 1 and System 2, so that System 1-led processes can run smoothly.*
2. *For an intensive Mode 1 interaction, e.g., video games and e-sports, focus on System 1 support.*

c) Supporting Transition Between Mode 1 and Mode 2: When the interaction running in Mode 1 breaks down, it becomes impossible to continue. In this case, the accomplishment of the goal via the interaction being advanced is either given up or a remedial action is taken to return from the failed state to the original normal state and resume to the execution in Mode 1. Mode 2 addresses the recovery process.

Guideline [C]

1. *On the detection of the intensive behavior of System 2 during Mode 1 support, switch from Mode 1 support to Mode 2 support.*

C. Process Level: Synchronization of Two Minds with the Environment

The mode-level support described in Section III-B is defined for the interactions that span the extended time along the time dimension. Therefore, its basis is a macroscopic bird's-eye view of the interactions. However, process-level support is defined for each event that occurs along the time dimension. Its basis is a microscopic view for the interaction at the level of each PCM process. The MHP/RT defines four processing modes by considering the manner in which System 1 and System 2 concern the event occurring at time T.

1) Four Processing Modes: Conscious/Unconscious Processes Before/After an Event: Experiences associated with the activities of an individual are characterized by a series of events, each of which is recognized by a person consciously. As shown in Figure 1, the behavior is defined as a time series of events, "$\cdots \rightarrow$ [Event at T_{N-1}] $\rightarrow$ [Event at T_N] $\rightarrow$ [Event at T_{N+1}] $\cdots$." focus on *a particular event* that occurs at the absolute time T_N. For the event to occur at T_N, the MHP/RT assumes that there should have existed the conscious processes of System 2 and unconscious processes of System 1 before T_N. For the executed event at T_N, the MHP/RT assumes that there should exist unconscious System 1 processes and conscious System 2 processes, concerning the event after T_N. The behavior of the MHP/RT appears as though it works in one of four processing modes [2][15] at a time before and after the event at T_N. They are shown at the bottom of Figure 1.

Two of the four processing modes concern the processes carried out *before* the event:

- System 2 Before Mode: In the time range of $T_N - \beta \leq t < T_N - \beta'$, where $\beta' \sim 500$ms and β ranges a few seconds to hours or even to months, the MHP/RT uses a part of the memory for System 2 to *consciously* prepare for future events.
- System 1 Before Mode: In the time range of $T_N - \beta' \leq t < T_N$, the MHP/RT *unconsciously* coordinates motor activities to the interacting environment. This mode uses the part of the memory for System 1.

The other two modes concern the processes carried out *after* the event:

- System 1 After Mode: In the time range of $T_N < t \leq T_N + \alpha'$, where $\alpha' \sim 500$ms, the MHP/RT *unconsciously* tunes the connections between the sensory inputs and motor outputs for a better performance for the same event in the future. This mode updates the connections within the part of the memory for System 1.
- System 2 After Mode: In the time range of $T_N + \alpha' < t \leq T_N + \alpha$, the MHP/RT *consciously* recognizes an event in the past and then modifies the memory concerning the event, where α ranges a few seconds to minutes or even to hours. This mode modifies the connections of the part of the memory for System 2.

2) Guidelines for Supporting Process Level Interactions: The human–artifact interaction needs to be synchronized for the cyclic PCM processes to run smoothly in any mode, i.e., Mode 1 through 4 defined in Section III-B, the interaction is in. The synchronization between the artifact and user is discussed in [16] in the case of a multi-modal interaction using the concepts of four processing modes. "Synchronization" and its derived concept of "weak synchronization" are defined as follows [17]:

> $\cdots$ a system and a user is synchronized if every system event at T_{sys} occurs as a user event at T_{user} with some amount of time allowance of Δ, $|T_{\text{user}} - T_{\text{sys}}| < \Delta$, where the actual values of Δ depend on the nature of interactions.
>
> $\cdots$
>
> However, a person's activity related with an event has to be considered from the four processing modes, which ranges relatively long time before and after the actual time the event happens. Therefore, "synchronization" has to be considered alternatively as the phenomena a person's activities during the time range of $[T - \beta,\ T + \alpha]$, which are linked with the specific recognizable system event at time T through a sequence of processes carried out in either of the four processing modes: all the processes have some link with the system event at T. When this is satisfied, the event is considered synchronized with a person's activities, which is called *weak synchronization* [16].

The human–artifact interaction has to provide a smooth flow of the four processing modes. It can break when a person has to adjust his/her activity while s/he is in the *System 1*

TABLE II
HAPPINESS GOALS [13] AND THEIR RELATION TO SOCIAL LAYERS. (ADAPTED FROM [1])

				Social Layers		
Category	No.	Name of Happiness	Types	Individual layer	Community layer	Social-system layer
I	8	Painful Happiness	The Masochist	+++		
	11	Tranquil Happiness	The Mediator	+++		
	14	Chemical Happiness	The Drug-taker	+++		
	15	Fantasy Happiness	The Day-dreamer	+++		
II	7	Rhythmic Happiness	The Dancer	+++	+++	
	16	Comic Happiness	The Laugher	+++	+++	
	4	Genetic Happiness	The Relative	+++	+++	
	5	Sensual Happiness	The Hedonist	+++	+++	
III	10	Selective Happiness	The Hysteric	+++	++	
	13	Negative Happiness	The Sufferer	+++	++	
IV	9	Dangerous Happiness	The Risk-taker	+++	++	+
	6	Cerebral Happiness	The Intellectual	+++	+++	++
V	1	Target Happiness	The Achiever	+++	+++	+++
	17	Accidental Happiness	The Fortunate	+++	+++	+++
VI	12	Devout Happiness	The Believer		+++	++
	2	Competitive Happiness	The Winner		+++	+++
	3	Cooperative Happiness	The Helper		+++	+++

+'s denote the degree of relevance of each goal to each layer, i.e., Individual, Community, and Social system, respectively. +++: most relevant, ++: moderately relevant, and +: weakly relevant.

Before Mode in such a way that his/her movement should be in synchrony with the current environment. This is the situation that the interaction has to avoid. This is because when this happens, the condition for weak synchronization is not satisfied. To remedy this, s/he has to make extra efforts to re-establish a weak synchronization by adjusting his/her movement. This leads to the following guidelines.

Guideline [D]

1. *Provide appropriate language-level support for System 2 while the user is in* System 2 Before Mode*,* $T_N - \beta \leq t < T_N - \beta'$*.*
2. *Provide appropriate perceptual- and motor-level support for System 1 while the user is in* System 1 Before Mode*,* $T_N - \beta' \leq t < T_N$*.*

D. Goal Level

The mode-level support described in Section III-B and the process-level support described in Section III-C concern direct interactions with the environment, to accomplish behavioral goals in problem-solving activities, e.g., real-world problem solving [8], or routine goal-oriented skilled activities by applying well-organized knowledge of Goals, Operators, Methods, and Selection rules (GOMS) [10]. As shown in Figure 1, the MHP/RT assumes that the behavioral goals are subordinate to happiness goals; the accomplishment of the behavioral goals are likely to be accompanied by the unconscious feeling of happiness, i.e., achieving a certain happiness goal.

1) Happiness Goals and their Relationship with the Behavioral Goals: Morris [13] listed 17 happiness goals. The left portion of Table II presents them, including goals such as "the inherent happiness that comes with the love of a child," "the competitive happiness of triumphing over your opponents," "the sensual happiness of the hedonist," and so on. Each happiness goal is associated with a type, e.g., the people "the achiever" should have "target happiness," "the winner" should have "competitive happiness," and "the drug-user" should have "chemical happiness."

Kitajima et al. [18] proposed the maximum satisfaction architecture (MSA). MSA assumes that the human brain pursues one of the 17 happiness goals defined by Morris [13] at every moment and switches to another happiness goal when appropriate by evaluating the current circumstances. Each of the happiness goals is associated with one or multiple layers of society. The right portion of Table II presents tentative assignments of the degree of relevance of each happiness goal to each social layer. The middle portion of Figure 1 suggests that any activities for achieving specific behavioral goals would be conducted by individual persons in the pursuit of any of the 17 happiness goals in the social layers presented in the right portion of Table II. Happiness goals define the value structure of the person when he or she makes decisions by running the PCM and memory processes under specific circumstances, while selecting his or her next actions. As such,

	Society of Linear Systems	*Society of Autonomous Systems*
Dimensionality	All systems are 4 dimensional systems	Each system has its own cognitive dimensions
Communication Method	Synchronization	Cooperative autonomous synchronization
Coordination Method	Reconstruction of the entire systems	Autonomous coordination by each system
Number of Systems	Small number of systems acceptable	Large number of systems can participate
Soft Link between Humans	Large burden	Small burden

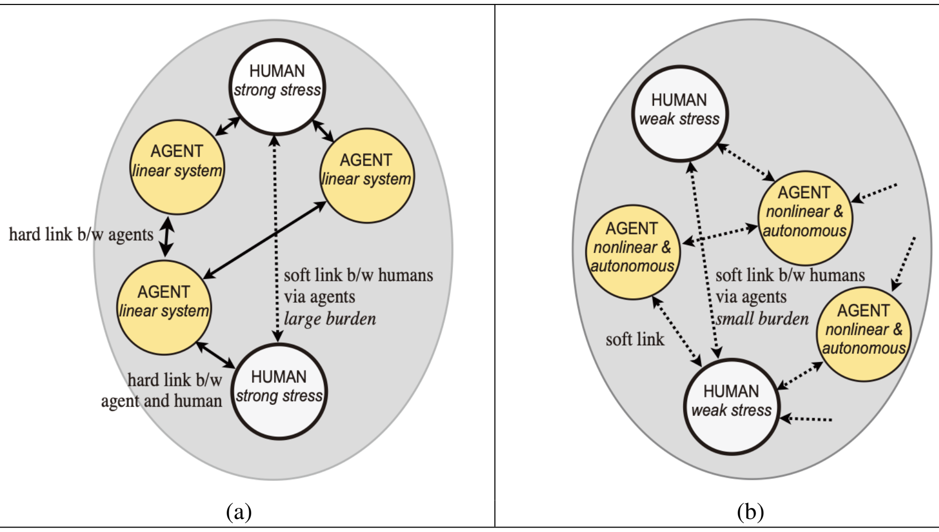

Fig. 2. Society of linear systems and society of autonomous systems.

it is vital to assume the correct happiness goal when supporting the next action selection process of a person, to accomplish the behavioral goals.

There could be associations between the processes of accomplishing behavioral goals and the recognized happiness goals, which could be useful to connect a behavioral goal with a happiness goal. The associations, however, could be vary among individuals. A single observed behavior under a behavioral goal, described in terms of the four operation modes and four processing modes, may have multiple associations with the happiness goals. This is because the condition for feeling happiness is strongly related with individual experiences and the manner in which the reward system functions for that experience [19]. Therefore, the mappings between the behavioral and happiness goals have significant individual and situational differences; a single person could feel different types of happiness when accomplishing a single behavioral goal for different contexts.

2) Guidelines for Supporting Goal Level Interactions: The purpose of designing artifacts has to be linked with any of the SDGs. The design space for artifacts could be explored strategically by associating the targeted SDGs with possible happiness states the user may want to achieve, which is indirectly connected with the behavioral goal of the user [7]. The mode and process level support are truly at the level at which the user could directly interact with. However, the goal-level support is at the level of motivation. The types of happiness goals have discernible aspects of behavioral ecology characterized by individual and contextual differences. Therefore, goal- and contextual-dependent support are needed.

The happiness goals listed in Table II are sorted into six categories according to the degree of relatedness with the social layers, i.e., individual, community, and social-system layers. The categories roughly define the context that the associated behavioral goals are trying to accomplish. The happiness goals in category I could be accomplished individually without any connections with the community or social-system. Those in the category II could be accomplished individually or with the members the individual belongs to. The rest of the categories could be characterized in a similar way. The interface for supporting happiness goals could be designed by category.

Guideline [E]

1. *Provide individually appropriate support for the identified happiness goal that the user might hold when trying to accomplish the behavioral goal.*
2. *Provide contextually appropriate support for the social layer in which the interaction might be conducted.*

IV. Application of the Guidelines: Two Hints

In general, guidelines are policies that indicate the direction to take when designing artifacts, and they do not indicate how the design can be carried out according to the specifically presented guidelines.

The guidelines given in Section III primarily *narrow down the design space in A-PDP* while designing artifacts, as displayed in the center of Figure 1. This can be accomplished

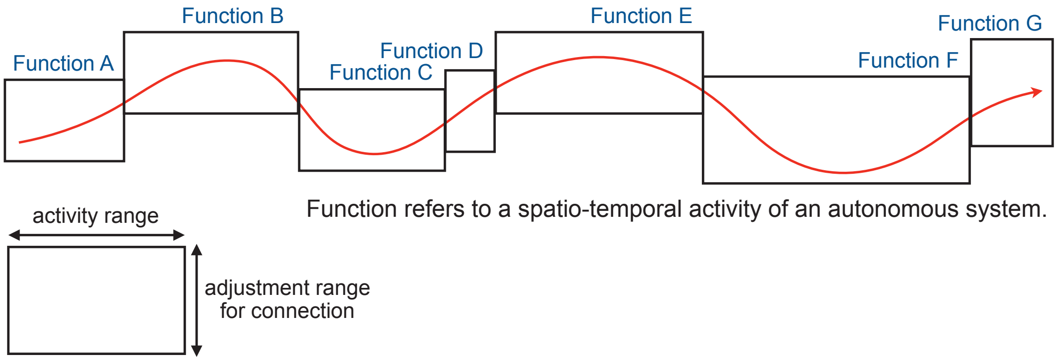
Function A
Function B
Function C
Function D
Function E
Function F
Function G
activity range
adjustment range for connection
Function refers to a spatio-temporal activity of an autonomous system.

Fig. 3. Successive functions are connected within the adjustable band in the spatio-time dimension (adapted from [17, Figure 6]).

through two approaches. The first approach, presented in Section IV-A, involves deriving a specification of the autonomous decentralized A-PDP system introduced in Section I. This system will enable it to function as a whole and achieve its goals, which are any of the SDGs, and then to apply to it the guidelines derived from human operating principles. The second approach, presented in Section IV-B, involves a reverse engineering approach. First, the interaction process with successful artifacts is observed ethnographically via MHP/RT to understand how the guidelines function in the interaction with the artifact in question. This approach helps ensure that artifacts designed as derivatives of successful artifacts are in line with the guidelines.

A. Autonomous Systems Interaction Design (ASID)

Users behave autonomously, and hence it seems natural to design interactive systems as autonomous systems, which are displayed in Figure 1 as A-PDP. This section reviews a framework for designing interactions between the two autonomous systems, A-PDP for artifacts and O-PDP for individuals, proposed briefly by [20]. Design activities are expected to proceed smoothly by following the design guidelines proposed in Section III.

1) Defining Design Space for ASID: Traditional interactive systems transform their input from the environment to output in the environment by using a set of rules. However, these systems are not intelligent enough to respond to an ever-changing environment including users. Therefore, inputs to a system may drift too far to be treated by the set of rules, and the system might respond inappropriately. In these cases, users may have to deal with the output of the system with some efforts that would not be needed if the system is well designed.

The key idea is to treat interactive systems as autonomous systems that interact with users that are other autonomous systems, and designing interactive systems implies designing autonomous systems interaction that establishes natural cooperation among them. By definition, autonomous systems establish appropriate relationships among themselves at any moment by means of autonomous cooperative synchronization. The environment of an autonomous system is defined by the rest of the autonomous systems. As such, the relationships among autonomous systems are symmetric. In other words, there is no asymmetry in the autonomous systems interaction such as “System A transmits data X (output of System A) to System B (input of System B).” The focus of interactive system design shifts from designing the I/O relationship between the systems and its environment to designing autonomous cooperative synchronization among systems, which we call autonomous systems interaction design (ASID).

Living organisms, O-PDP, establish appropriate relationships with their surrounding environment, A-PDP, by means of autonomous cooperative synchronization. This mechanism is flexible and robust enough to achieve timely and automatic coordination with the environment. The mechanism is modeled by MHP/RT, and the design guidelines shown in Section III are derived by considering the conditions for realizing flexible and stable interactions between O-PDP and A-PDP. When determining the specifications of autonomous systems operating on A-PDP, it is possible to narrow the scope by considering the characteristics of the O-PDP that will be interacting with them. For this purpose, the guidelines described in Section III should be applied.

2) Society of Autonomous Systems: The environment human beings interact with also includes interactive systems. This section starts by describing a society of systems that is linear or autonomous, followed by the needs that those systems must satisfy and the proposal of autonomous system interaction that should meet the requirements.

a) Linear Systems: Behaving objects in the environment are defined in four-dimensional space-time coordinates. A human being viewed as a linear system acquires the information of behaving objects, i.e., humans other than oneself and artifacts surrounding oneself, via its sensory organs as two-dimensional data. The four-dimensional data are reduced to two-dimensional data in this process. The axes that make up these two dimensions are the time axis and a one-dimensional axis representing the feature to which attention is directed.

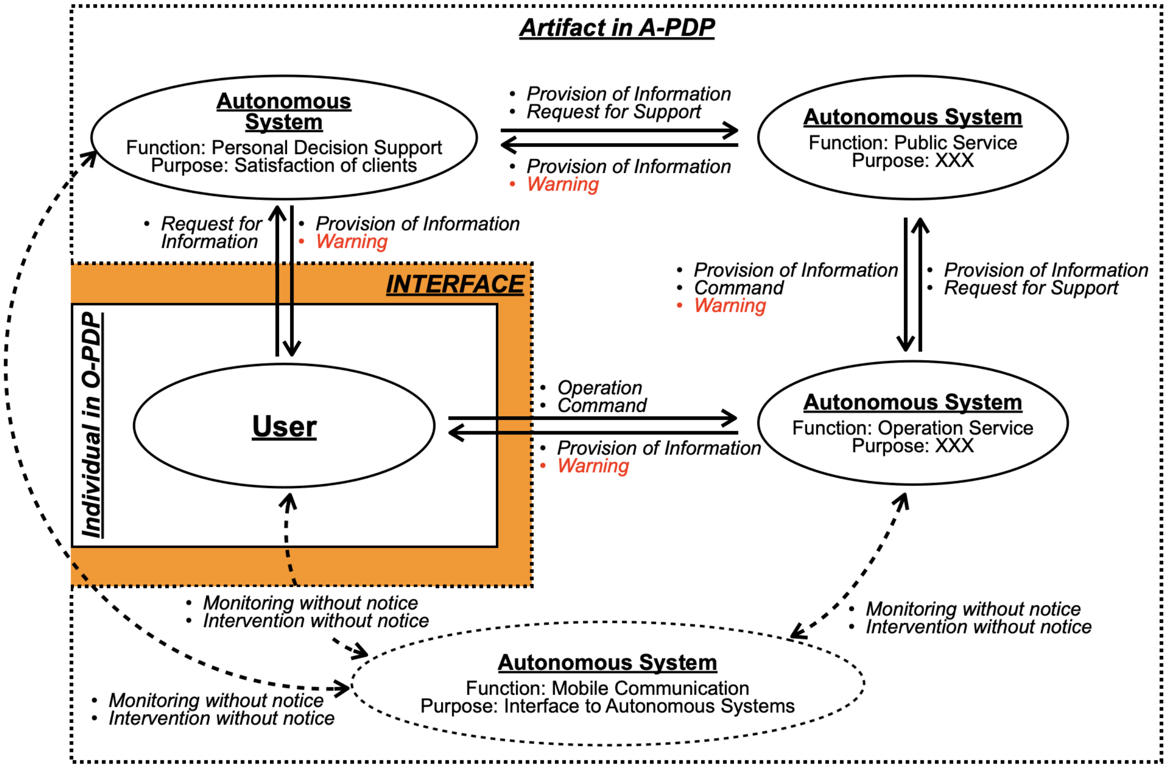
Artifact in A-PDP
Autonomous System
Function: Personal Decision Support
Purpose: Satisfaction of clients
• Provision of Information
• Request for Support
• Provision of Information
• Warning
Autonomous System
Function: Public Service
Purpose: XXX
• Request for Information
• Provision of Information
• Warning
INTERFACE
Individual in O-PDP
User
• Provision of Information
• Command
• Warning
• Provision of Information
• Request for Support
• Operation
• Command
• Provision of Information
• Warning
Autonomous System
Function: Operation Service
Purpose: XXX
• Monitoring without notice
• Intervention without notice
• Monitoring without notice
• Intervention without notice
Autonomous System
Function: Mobile Communication
Purpose: Interface to Autonomous Systems
• Monitoring without notice
• Intervention without notice

Fig. 4. An illustration of ASID.

The input data are used for representing their characteristics by means of static linear functions. The purpose of the linear functions is to predict the future states of objects. When an objective of a behavior is given, the linear system will behave by deriving static solutions through the use of the linear functions that best match the current situation.

Figure 2(a) illustrates a society of linear systems managing various situations by tuning the relationships among the constituent systems. However, there are situations where the current organization of the systems causes a large amount of stress in spite of efforts made to resolve the situations, and they cannot behave properly. In these situations, the systems have to change themselves. However, the change may or may not produce good results. In the worst cases, the change may cause a rapid increase of stress and crash the system.

b) Autonomous Systems: Human beings viewed as autonomous systems represent behaving objects, i.e., humans other than oneself and artifacts surrounding oneself, in the four-dimensional space-time environment via sensory organs. For example, the sense of taste is represented by six-dimensional data, and the sense of sight is represented by four-dimensional data. The input data are processed mainly by System 1 and optionally by System 2, and they are used to define functions that work in the multi-dimensional memory frames and MSA for accomplishing some of the happiness goals under the real-time constraints for establishing *stable synchronization* with the environment. The functions accumulate personal experience continuously in the multi-dimensional memory frames to be used in the distinctive layers of Newell's Biological, Cognitive, Rational, and Social Bands [9]. When an objective of a behavior is given, the autonomous system will behave by deriving effective regions so that the self will behave properly by using the functions. When an autonomous system communicates with another one, it uses the effective region at each moment. This assures less stressful communication among autonomous systems than among linear systems (Figure 2(b)).

The stable synchronization described above is accomplished by means of weak synchronization [17]. Autonomous elements are weakly synchronized with the external world, and the way they actually work indirectly reflects the circularity of the existing environment, i.e., autopoiesis [21], and fluctuations inherent in the environment. This situation is schematically shown by Figure 3. A function, C, is connected with another function, D, using the region of the overlapping edge for maintaining continuity of the activities. Function C can be a series of conscious activities performed in the Rational Band [9] to plan ahead a sequence of actions for controlling the car by consulting the contents of the relation multi-dimensional memory frame. Function C is followed by Function D, which can be an unconscious activity for tuning the planned activities for the particular road conditions by using the bottom layer of the memory structure, i.e., the perceptual, behavior, and motion multi-dimensional memory frames.

c) An Illustration of Autonomous Systems Interaction (ASI): Using the example presented in Figure 4, the following

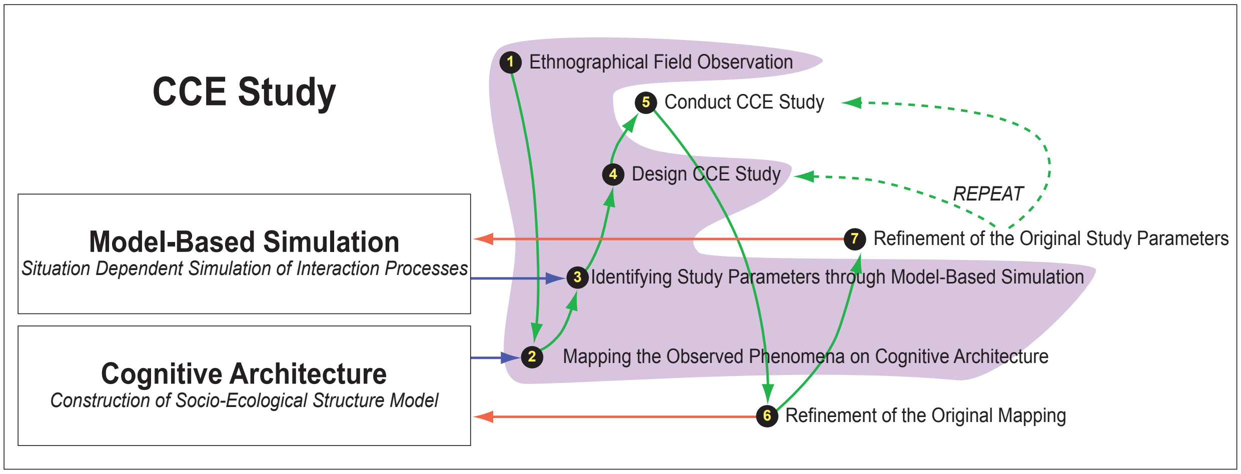
CCE Study
Model-Based Simulation
Situation Dependent Simulation of Interaction Processes
Cognitive Architecture
Construction of Socio-Ecological Structure Model
1 Ethnographical Field Observation
5 Conduct CCE Study
4 Design CCE Study
REPEAT
7 Refinement of the Original Study Parameters
3 Identifying Study Parameters through Model-Based Simulation
2 Mapping the Observed Phenomena on Cognitive Architecture
6 Refinement of the Original Mapping

Fig. 5. The CCE procedure [3, Figure 5.1].

illustrates what the interaction of an autonomous system looks like. In real-world interactions between autonomous systems in the A-PDP, the only thing that is predetermined is the communication protocol. Information flows over a bus to an unspecified number of autonomous systems. Figure 4 displays an extract of a portion of this information. A total of five autonomous systems are shown in Figure 4, four in the A-PDP system and one in the O-PDP system, labeled as “User”. The two systems in the A-PDP, the personal decision support system and operation service system, interact directly with the user and provide information or issue warnings when the user requests information or executes operations or commands. Another system in the A-PDP, the public service system, interacts directly with the two systems mentioned above, provides information, and issues warnings in response to requests for support and the given information from them.

The last system in the A-PDP behaves differently from the other systems. This system is the mobile communication system, indicated by the dotted oval in Figure 4. The purpose of this system is to allow the autonomous systems in the A-PDP and O-PDP to synchronize with each other and to facilitate processing. To achieve this, this autonomous system monitors other autonomous systems unnoticed and intervenes with them unnoticed.

As demonstrated in the next section, this autonomous system, which monitors and intervenes unnoticed by other autonomous systems, plays a major role in synchronizing and organizing the interactions among the autonomous systems in A-PDP and O-PDP. However, it is not generally guaranteed that the results of an unnoticed intervention will be meaningful. It depends largely on the circumstances. Therefore, it is necessary to create a feedback system to monitor the impact of the intervention, and if the transition is not in a favorable direction, further interventions to return the system to a normal state are necessary. This would give the autonomous systems in A-PDP and the user in O-PDP the resilience to return to the normal state, thus allowing the entire system to operate in a stable manner.

d) Application of the Guidelines to ASID: The guidelines presented in Section III apply to the design of the portion of the interface displayed in orange in Figure 4. When designing “Provision of Information” and “Warning” displayed in Figure 4, those guidelines that mention “*provide* $\cdots$ *support*”, the Guidelines [D], should be applied. For guideline [D], depending on whether the user is in System 1 Before Mode or System 2 Mode, [D-1] or [D-2] shall apply. It is preferable to determine the user’s state unnoticeably to not affect the natural interaction between the user and the system. For this reason, it is necessary to use the mobile communication system displayed in Figure 4 when applying this guideline.

Guidelines [A], [B], and [C] are to be applied based on the determination of which mode the user is interacting with the system. As this judgment should be made without disturbing the natural interaction between the user and the system, the mobile communication system displayed in Figure 4 should be used. For example, the system intervenes in the interaction between the user and the system without being noticed by the user, naturally prompting the user to request information from the system or providing information from the system to the user that matches the user’s operation mode.

B. CCE

This section introduces a methodology for conducting field experiments to understand human behaviors as a hint for carrying out a strategically principled search in the design space to obtain design specifications that conform to the guidelines this study proposes. The methodology, cognitive chrono-ethnography (CCE) [22], should complement the MHP/RT by providing the real data of human behavior for specific situations that should define constraints on the functioning of PCM and memory processes.

1) CCE for Narrowing Down the Design Space: CCE combines three concepts. “Cognitive” declares that CCE deals with interactions between consciousness and unconsciousness in the PCM cycles. “Chrono (-logy)” is about time ranging from ∼100 ms to days, months, and years, and CCE focuses on these time ranges. “Ethnography” indicates that CCE takes ethnographical observations as the concrete study method because in daily life, the Two Minds of people tend to re-use experientially effective behavioral patterns, which are biases and might have individual and contextual differences.

To conduct a CCE study, study participants (elite monitors) are selected. Each defining study field has values. The study question is “what would certain people do in certain ways in certain circumstances (not average behavior)?” Therefore, elite monitors, certain persons, are selected by consulting the parameter space. In this process, it is necessary that the points in the parameter space, which correspond to the elite monitors, are appropriate for analyzing the structure and dynamics of the study field. The methodology is not for human–artifact interaction but for every aspect of the daily life of human beings. Regarding the relationship between CCE and the design space, CCE focuses on understanding the process of interaction between successful artifacts and users and is intended for existing artifacts. Therefore, it is out of scope to predict the kind of interaction that occurs between the user and a non-existent artifact that no one has discussed. The role of CCE is to enable the design space to be narrowed down by a solid understanding of the success stories of existing artifacts, thereby defining the successful areas of new designs. With that in mind, it will be possible to come up with alternatives to successful artifacts. Whether or not new and innovative artifacts are accepted by users is discussed in another guideline paper published by us [6].

2) CCE Procedure for the Human–Artifact Interaction: Figure 5 shows the seven steps to conduct a CCE study [3, Figure 5.1]. The following describes the CCE steps adapted to human–artifact interaction. Necessary additions appear after the general descriptions for the CCE procedures.

(1) *Ethnographical Field Observation:* Use the basic ethnographical investigation method to clarify the outline of the structure of social ecology that underlies the subject to study.

> The subject of study is to understand the manner in which the existing artifacts in question are used successfully by the current users. The ultimate goal of the artifacts is to achieve any of the SDGs through their use by potential users; the current users may be a part of them. The range of users could be widened by appealing appropriately to the right segments of the users. The users could be characterized as an individual, a member of a community or a social-system. Depending on the social layers the users belong to, the happiness goals that could be achieved could vary. The kinds of SGDs that the artifacts with the current or appropriately enhanced specifications could achieve may be widened or corrected. In this step, it is necessary to clarify the outline of the structure of social ecology in terms of the segment of potential users, the social layer they belong to, and the happiness goals they may achieve by referring to Table II.

(2) *Mapping the Observed Phenomena on Cognitive Architecture:* With reference to the behavioral characteristics of people which have been made clear so far and cognitive architectures, consider what kind of characteristic elements of human behavior are involved in the investigation result in (1).

> This study proposes the use of MHP/RT as the cognitive architecture for this step. As this study proposes, the human–artifact interaction is characterized at three levels, i.e., the mode, the process, and the goal levels. This is based on the MHP/RT cognitive architecture. Because the artifacts in question realize successful interactions with the users, it is assumed that their design should conform to the guidelines in specific ways. In this step, it is necessary to describe the manner in which they conform to the guidelines, i.e., the mode-level support provided, the process-level support, and the goal-level support.

(3) *Identifying Study Parameters through Model-Based Simulation:* Based on the consideration of (1) and (2), construct an initial simple model with the constituent elements of activated memories, i.e., meme, and the characteristic PCM processing to represent the nature of the ecology of the study space.

> In this step, the “what” question answered in (2) is operationalized by turning it into the “how” question. This is answered by constructing an MHP/RT model that could simulate successful users of the artifact in question. The model could run by specifying (a) the likely happiness goals, (b) the possible modes of the assumed interaction, (c) the possible ways of weak synchronization establishment, and (d) the kinds of memes of the simulated user [23][24]. The successful users could be characterized by combining the values assigned to (a) ∼ (d), which constitute the study parameters.

(4) *Design a CCE Study:* Based on the simple ecological model, identify a set of typical behavioral characteristics from a variety of people making up the group to be studied. Then formulate screening criteria of elite monitors who represent a certain combination of the behavioral characteristics, and define ecological survey methods for them.

> This step follows the standard CCE procedure.

(5) *Conduct CCE Study:* Select elite monitors and conduct an ethnographical field observation. Record the monitors' behavior. The elite monitors are expected to behave as they normally do at the study field. Their behavior is recorded in such a way that the collected data is rich enough to consider the results in terms of the parameter space and as un-intrusively as circumstances allow.

> This step follows the standard CCE procedure.

(6) *Refinement of the Original Mapping:* Check the results of (5) against the results of (2) for appropriateness of the mapping. If inappropriate, back to (2) and redo from there.

This step follows the standard CCE procedure.

(7) *Refinement of the Original Study Parameters:* If the result of (5) is unsatisfactory, go back to (4) and re-design and conduct a revised CCE study, otherwise go back to (3) to redo the model-based simulation with a set of refined parameters.

This step follows the standard CCE procedure.

On completion of the CCE cycle, the existing social ecology that characterizes the successful use of the artifact is understood. This understanding is used to widen the range of successful use of the artifact and contribute to determining the direction of strategic development for the maximum utilization of the artifact.

V. Conclusion

The purpose of this study was to contribute to realizing a sustainable society of human beings and artifacts. The focus was on the human–artifact interaction, which occurs at the interface between human beings (the interface is composed of multiple autonomous systems, i.e., PCM and memory systems), and artifacts, which are a collection of autonomous systems. This study used a theory-based approach to derive guidelines for application when designing artifacts that should realize a sustainable society.

The constraints imposed on the derivation were: 1) the ultimate purpose of artifacts for realizing a sustainable society should be the achievement of any of the SDGs, and 2) human interactions with the artifacts should be theorized by the MHP/RT cognitive architecture. These constraints were related with each other via the concept of resilience of the interaction processes. On the one hand, the stability of the human–artifact interaction at the mode level, i.e., either System 2 dominant or System 1 dominant processes should be carried out stably, was the necessary condition for the accomplishment of behavioral goals using the artifact. On the other hand, the accomplishment of behavioral goals is linked with the feeling of achieving any of the 17 happiness goals, defined at the three social layers. The behavioral goals do not necessarily have a direct connection with the SDGs; rather, the accomplishment of behavioral goals indirectly contributes to the achievement of any of SDGs as by-products [7]. Because both the happiness goals and SDGs focus on social ecology, the mapping between them could be established [7]. This would complete the links from the stable accomplishment of behavioral goals to the achievement of happiness goals and SDGs for realizing a sustainable society.

Generally, guidelines are useless, unless they are practiced.

The second hint is related to a method for applying the derived guidelines based on CCE, which defines the experimentation procedure for complementing the theory of cognitive architecture, MHP/RT. CCE and MHP/RT are the two-wheels of a vehicle to understand the daily behavior of human beings [22]; evidently, the human–artifact interaction is part of it. CCE is used to understand observed behavior. Therefore, it is most useful for extending the existing interaction processes by deliberately extrapolating them by the provision of new interface designs, which should conform to the relevant guidelines. For example, at the process level, weak synchronization has to be realized in the interaction process between the new design and the user. If this interaction is carried out as routine goal-oriented skills in Mode 1, the behavior of the users could be represented as several versions of the GOMS models [10] that are suitable for accomplishing respective behavioral goals. The appropriateness of the new design has to be considered, as to whether it could establish weakly synchronized interaction, given the existing GOMS models.

An artifact is defined as a set of specifications, which are sufficient for engineering to realize a working product. The raison d'être of the artifact would be to contribute to the achievement of any of the SDGs and to make its users feel any of the happiness goals, to realize a sustainable society through the human–artifact interaction. This study proposed a method for bridging these goals as a set of guidelines on the basis of the scientific understanding of human behavior, provided by the cognitive architecture, MHP/RT, the methodology for experimentation, CCE, and the design concept for autonomous systems, ASID, to narrow down the design space.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 20H04290. The author would like to thank Editage (www.editage.com) for English language editing.

References

[1] M. Kitajima, M. Toyota, and J. Dinet, "Guidelines for Designing Interactions Between Autonomous Artificial Systems and Human Beings," in *COGNITIVE 2022 : The Fourteenth International Conference on Advanced Cognitive Technologies and Applications*, 2022, pp. 23–31.

[2] M. Kitajima and M. Toyota, "Decision-making and action selection in Two Minds: An analysis based on Model Human Processor with Realtime Constraints (MHP/RT)," *Biologically Inspired Cognitive Architectures*, vol. 5, pp. 82–93, 2013.

[3] M. Kitajima, *Memory and Action Selection in Human-Machine Interaction*. Wiley-ISTE, 2016.

[4] D. A. Norman, "Cognitive engineering," in *User Centered System Design: New Perspectives on Human-Computer Interaction*. CRC Press, 1986, ch. 3, pp. 31–61.

[5] H. A. Simon, *The Sciences of the Artificial*, 3rd ed. Cambridge, MA: The MIT Press, 1996.

[6] M. Kitajima and M. Toyota, "Guidelines for designing artifacts for the dual-process," in *Procedia Computer Science, BICA 2015. 6th Annual International Conference on Biologically Inspired Cognitive Architectures*, vol. 71, 2015, pp. 62–67.

[7] M. Kitajima, "Cognitive Science Approach to Achieve SDGs," in *COGNITIVE 2020 : The Twelfth International Conference on Advanced Cognitive Technologies and Applications*, 2020, pp. 55–61.

[8] V. Sarathy, "Real World Problem-Solving," *Frontiers in human neuroscience*, vol. 12, p. 261, 2018.

[9] A. Newell, *Unified Theories of Cognition (The William James Lectures, 1987)*. Cambridge, MA: Harvard University Press, 1990.

[10] S. K. Card, T. P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.

[11] J. L. McClelland and D. E. Rumelhart, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition : Psychological and Biological Models*. The MIT Press, 6 1986.

[12] D. Kahneman, "A perspective on judgment and choice," *American Psychologist*, vol. 58, no. 9, pp. 697–720, 2003.

[13] D. Morris, *The nature of happiness*. London: Little Books Ltd., 2006.

[14] M. Kitajima and M. Toyota, “Simulating navigation behaviour based on the architecture model Model Human Processor with Real-Time Constraints (MHP/RT),” *Behaviour & Information Technology*, vol. 31, no. 1, pp. 41–58, 2012.

[15] ——, “Four Processing Modes of *in situ* Human Behavior,” in *Biologically Inspired Cognitive Architectures 2011 - Proceedings of the Second Annual Meeting of the BICA Society*, A. V. Samsonovich and K. R. Jóhannsdóttir, Eds. Amsterdam, The Netherlands: IOS Press, 2011, pp. 194–199.

[16] J. Dinet and M. Kitajima, “Immersive interfaces for engagement and learning: Cognitive implications,” in *Proceedings of the 2018 Virtual Reality International Conference*, ser. VRIC ’18. New York, NY, USA: ACM, 2018, pp. 18/04:1–18/04:8. [Online]. Available: https://doi.org/10.1145/3234253.3234301[retrieved:December, 2022]

[17] M. Kitajima, J. Dinet, and M. Toyota, “Multimodal Interactions Viewed as Dual Process on Multi-Dimensional Memory Frames under Weak Synchronization,” in *COGNITIVE 2019 : The Eleventh International Conference on Advanced Cognitive Technologies and Applications*, 2019, pp. 44–51.

[18] M. Kitajima, H. Shimada, and M. Toyota, “MSA:Maximum Satisfaction Architecture – a basis for designing intelligent autonomous agents on web 2.0,” in *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, D. S. McNamara and J. G. Trafton, Eds. Austin, TX: Cognitive Science Society, 2007, p. 1790.

[19] M. Kitajima and M. Toyota, “Two Minds and Emotion,” in *COGNITIVE 2015 : The Seventh International Conference on Advanced Cognitive Technologies and Applications*, 2015, pp. 8–16.

[20] ——, “Autonomous Systems Interaction Design (ASID) based on NDHB-Model/RT,” in *Proceedings of the 31th Annual Conference of the Cognitive Science Society*. Cognitive Science Society, 2009.

[21] H. Maturana and F. Varela, *Autopoiesis and Cognition (Boston Studies in the Philosophy and History of Science)*, softcover reprint of the original 1st ed. 1980 ed. D. Reidel Publishing Company, 8 1991.

[22] M. Kitajima, “Cognitive Chrono-Ethnography (CCE): A Behavioral Study Methodology Underpinned by the Cognitive Architecture, MHP/RT,” in *Proceedings of the 41st Annual Conference of the Cognitive Science Society*. Cognitive Science Society, 2019, pp. 55–56.

[23] M. Toyota, M. Kitajima, and H. Shimada, “Structured Meme Theory: How is informational inheritance maintained?” in *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, B. C. Love, K. McRae, and V. M. Sloutsky, Eds. Austin, TX: Cognitive Science Society, 2008, p. 2288.

[24] M. Kitajima, M. Toyota, and J. Dinet, “The Role of Resonance in the Development and Propagation of Memes,” in *COGNITIVE 2021 : The Thirteenth International Conference on Advanced Cognitive Technologies and Applications*, 2021, pp. 28–36.