Making VR Environments Fully Resonate with Interacting Humans: Core Constructs

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

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

Katsuko T. Nakahira
Nagaoka University of Technology
Nagaoka, Niigata, Japan
Email: katsuko@vos.nagaokaut.ac.jp

***Abstract*— Much of human life is shaped by interacting with artificial objects in the surrounding environment. When these artificial objects behave not according to natural laws but according to algorithms created by artifact designers, the environment can be considered a virtual environment. Our previous research has conceptualized the human–object interaction process as comprising perceptual, cognitive, and motor processes operating synchronously with the environment's changes, alongside memory processes operating asynchronously with the environment's changes, further linked by resonance between them. More specifically, P-resonance triggered by the basic senses comprised of rhythmic, spatial, and number senses connects perceptual process with the Perceptual-Multi-Dimensional Memory Frame, and C-resonance connects the Multi-Dimensional Memory Frame with cognitive process via the structure of Goals, Operators, Methods, and Selection rules (GOMS) established in the Multi-Dimensional Memory Frame. This paper presents the characteristics objects in a virtual environment should possess to enable smooth human–object interactions.**

Keywords- Resonance; Basic Senses; Meme; GOMS; MHP/RT; Symbol Grounding.

I. Introduction

Humans acquire information about the external world through their five senses, activate Perceptual, Cognitive, and Motor (PCM) processes, and select and execute appropriate actions based on the circumstances at that moment. These processes are controlled by the brain. The brain forms a complex mechanism composed of an immense number of cells. However, when viewed functionally as an information processing device that appropriately converts input information to the brain and outputs it to the outside world, it can be understood as a remarkably simple cognitive architecture model.

Card, Moran, and Newell [1] devised Model Human Processor (MHP), a cognitive architecture model that simulates human perception, cognition, and motor processes. By analogy to von Neumann-type computers as information processing devices, they conceived the brain as comprising the components "memory," "processor," and "program," which defines how memory and processor are coupled. Memory is characterized by its storage capacity, decay time, and data type, while a processor is characterized by its cycle time, which is the processing time per unit task. Memory consists of visual and auditory image stores, working memory, and long-term memory, each serving distinct functions. The following situation was envisioned as the simulation target domain for operating the information processing device specified above. This situation involves solving problems within a clearly defined problem space based on the principle of rationality to achieve the specified goal. The knowledge required in this situation is organized as Goals, Operators, Methods, and Selection rules (GOMS).

By encoding perceived information within the perceptual process and representing it as symbols, it becomes possible to *think* within the cognitive process. In the thinking process, memory is utilized, and symbols are successively transformed. Memory itself is also expressed through symbols. Part of the result of thinking provides an expression of a sequence of actions executable in the motor process. Based on the idea that human intelligence can be captured by the thinking process of manipulating symbols, Newell proposed the Physical Symbol System as a theory of human intelligence [2]. This approach provided the foundation for Soar, a cognitive architecture that views thinking as problem-solving process involving manipulation of a problem space [3][4].

However, the real-world environment with which humans interact changes its state moment by moment based on the mechanisms inherent within the environment itself. Mechanisms can be either linear or nonlinear. In the latter case, it is fundamentally impossible to predict the temporal evolution of the state. When confronted with such real-world environments, it is difficult to fully capture the appropriate actions required in response to the environment's state as a rational problem-solving process within a predefined, well-defined problem space. We have developed a cognitive architecture model—the Model Human Processor with Realtime Constraints (MHP/RT)—that can simulate not only behaviors modeled by navigation within well-defined problem spaces targeted by MHP and Soar, but also adaptive and flexible behaviors [5][6]. MHP/RT integrates models for understanding individual behavioral ecology and models for understanding collective behavioral ecology, incorporating the latest research findings from neuroscience and cognitive science into the conceptual framework of MHP.

MHP/RT consists of the following three elements: Perception, Cognition, and Motor. These form the "Perception-Cognition-

 ISBN: 978-1-68558-381-1

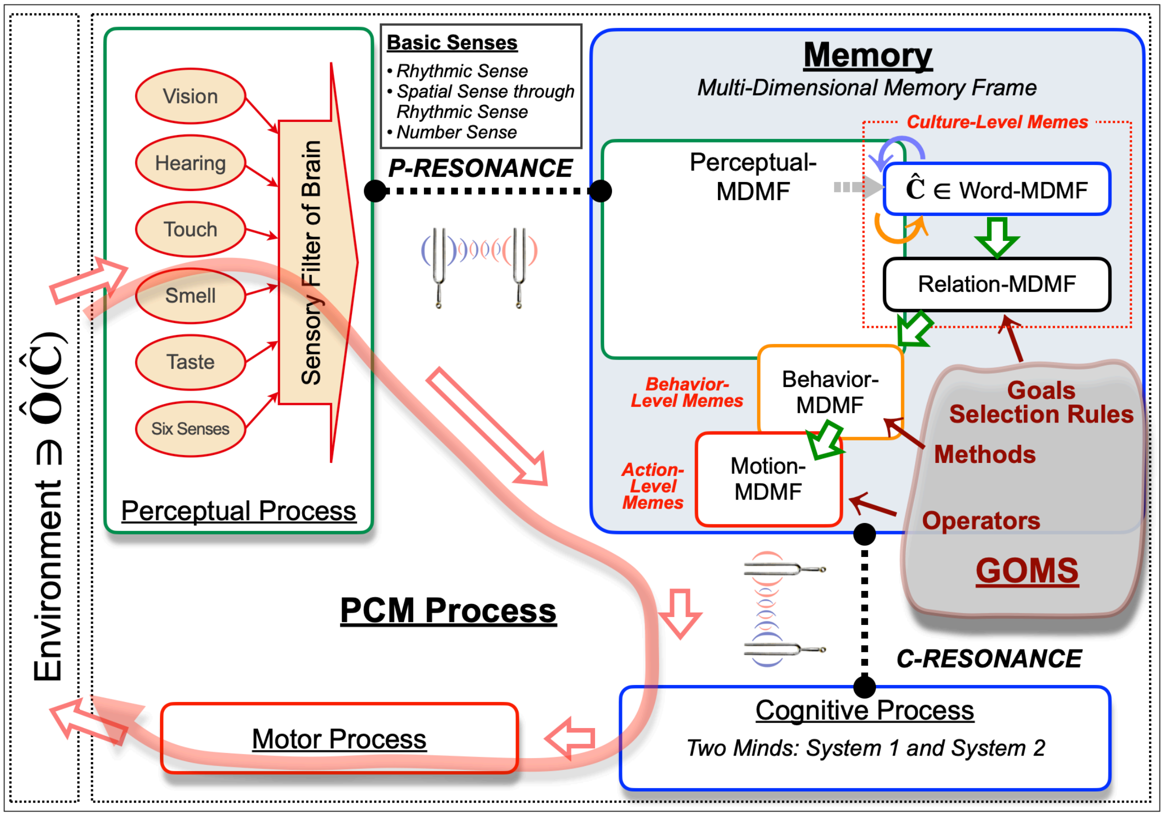

Figure 1. PCM processes operating synchronously with the environment and Multi-Dimensional Memory Frame operating asynchronously with the environment (created by synthesizing [7, Figure 2] and [8, Figure 6(b)] and making modifications).

Motor (PCM) Process," which selects and executes actions that are likely to yield satisfactory results without causing the human system, as a living organism, to fail. This process *synchronizes* human interactions with the real-world environment along the time axis with the state transitions of that environment; the "Memory Process," which operates *asynchronously* with the real environment, accumulating the results of PCM process execution and functioning to be utilized during action selection and execution; and a "resonance" mechanism connecting the environment-synchronized PCM process and the environment-asynchronous memory process. The operation of the PCM process represents individual behavioral ecology, while the contents of memory represent collective ecology—an accumulation of interactions with artifacts within a culture that is inherited across generations.

The key to understanding humans acting within their environment is as follows.

(1) PCM processes operating in various modes depending on the task,
(2) The structure and content of memory acquired while PCM processes are active and utilized by them,
(3) P-resonance linking perceptual processes and memory,
(4) C-resonance linking memory and cognitive processes,

Understanding is achieved by holistically relating these elements and grasping their overall dynamics.

Behavioral events occurring in actual dynamic environments manifest in diverse forms. This paper identifies factors to consider when analyzing and understanding them, structuring insights gained during our research process according to the four points listed above. From the perspective of ensuring *smooth* interaction with the environment, it outlines the necessary requirements for achieving it.

This remainder of this paper is organized as follows. Section II explains the PCM process, memory process, and the P-resonance and C-resonance that connect them. Section III presents considerations for designing smooth interactions as conditions that induce resonance. Section IV summarizes this research.

II. PCM, Memory Processes and Resonance

This section provides an overview of the PCM process and memory process within MHP/RT, a cognitive architecture capable of simulating everyday behavioral selections [5][6][9] as the foundation.

A. Interaction between Human and Objects in the Environment

Interaction between humans and the object $\hat{\mathrm{O}} \in \mathbf{Environment}$ in the environment is realized through PCM processes operated by humans, memory processes utilized by PCM processes and updated as a result of PCM process execution, and processes occurring within $\hat{\mathrm{O}}$. Figure 1 shows its overall outline.

Ô changes state over time. These changes may arise either from human actions affecting Ô or from mechanisms inherent within Ô itself. In either case, for humans to smoothly continue their interaction with Ô, it is necessary to synchronize their own PCM process with changes in the state of Ô.

Meanwhile, memory defines the content executed by the PCM process. Memory accumulates the results of the PCM process's execution. Furthermore, memory is provided during the execution of the PCM process. The processes of accumulation and provision are executed asynchronously with respect to the environment's temporal development.

PCM processes operating synchronously with the environment and memory processes operating asynchronously with the environment collaborate to realize human behavior within the environment—what is executed and how—by linking them through resonance. Resonance includes P-resonance occurring between perceptual processes and memory, and C-resonance occurring between cognitive processes and memory. These enable the PCM process to utilize memory and generate behavior within the environment.

B. PCM Process Operating Synchronously with the Environment

1) PCM Process: When interacting with objects in the environment, humans receive physical and chemical stimuli through sensory nerves located at the interface with the environment and take in environmental information in the body. Further, it generates bodily movements that are suitable for the current environment. The stable and sustainable relationship between the environment and the self is established through continuous coordination between the activity of the self and the resultant changes in the environment, which should affect the self's next action.

Figure 1—created by combining [7, Figure 2] and [8, Figure 6(b)] and making modifications—illustrates the process by which environmental information is incorporated into the body via sensory nerves, undergoes processing within the brain, and then acts upon the external world via motor nerves, based on MHP/RT [5][6]. This process involves memory, modeled as Multi-Dimensional Memory Frame shown in the upper right of Figure 1, and the PCM processes indicated by the red arrows in Figure 1. The Multi-Dimensional Memory Frame comprises the Perceptual-, Behavior-, Motor-, Relation-, and Word-Multi-Dimensional Memory Frame. The Perceptual-Multi-Dimensional Memory Frame overlaps with the Behavior-, Relation-, and Word-Multi-Dimensional Memory Frame. This allows activity to propagate from Perceptual- to Motor-Multi-Dimensional Memory Frame.

This section explains the relationship between the PCM process and the memory process. Perceptual information taken in from the environment through sensory organs *resonates* with information in the Multi-Dimensional Memory Frame, which is called P-Resonance [10]. In Figure 1, this process is indicated by the symbol •—•. Resonance occurs first in the Perceptual-Multi-Dimensional Memory Frame and activates the memory network. After that, the activation spreads to the memory networks that overlap with the Perceptual-Multi-Dimensional Memory Frame, and finally to the Motor-Multi-Dimensional Memory Frame. In cognitive processing based on the Two Minds [0], conscious processing (System 2) and unconscious processing (System 1) operate in an interrelated manner [8][10]. System 2 utilizes the Word- and Relation-Multi-Dimensional Memory Frame via C-Resonance, while System 1 draws on the Behavior- and Motor-Multi-Dimensional Memory Frame via the same mechanism. Motor sequences are then expressed according to the Motor-Multi-Dimensional Memory Frame. The memories involved in the production of actions are updated to reflect the traces of their use process and influence the future action selection process.

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.*

Figure 2. Four operation modes of MHP/RT and their relationship with the four bands in the time scale of Newell's human action [3, Figure 3-3].

2) Four Operation Modes: Humans interact with the external environment and select appropriate actions to achieve behavioral goals through a cycle of PCM processes. In MHP/RT, the action selection process is controlled by System 1 and System 2. These systems cooperate to link perception and movement, and the degree of cooperation depends on the state of the external environment with which the MHP/RT interacts. Figure 2 shows the Four Operation Modes characterized by the relationship between System 1 and System 2. There are synchronous and asynchronous modes.

The interaction is carried out in one of the four operation modes. Interactions executed in Mode 1 or Mode 2 are sound in the sense that System 2 monitors whether actions executed by System 1 in a timely manner have deviated from the desired trajectory. If signs of undesirable events occurring are detected, feedback control via System 2 activates, enabling a return to normal operation. Meanwhile, Mode 3 and Mode 4 contain unmonitored feedforward System 1 processes. This indicates that instability may be introduced into the interactions.

In Mode 2, System 2 frequently intervenes the PCM processes conducted by System 1. More precisely, the pace of interaction is controlled by System 2. The role of System 1 is to carry out the necessary PCM processes, to advance the main System 2–environment interactions. As interactions progress, whenever the situation changes, the content of actions is determined starting from System 2 and executed by System 1. This does not correspond to the situation described in this paper where smooth interaction is achieved.

In contrast, in Mode 1, System 2's intervention is weak. The external environment supports the automatic processes performed unconsciously by System 1, enabling smooth interaction. This is the smooth interaction addressed in this paper. In this study, we propose that this is achieved through the occurrence of P-resonance and C-resonance. Our previous paper [13] provides guidelines for designing artifacts that realize interactions for Mode 1 and Mode 2. The guideline for supporting Mode 1 interaction is as follows.

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.*

This study details the considerations for implementing the items in this guideline.

C. Memory Processes Operating Asynchronously with the Environment

When the PCM process is running, the contents of Perceptual-Multi-Dimensional Memory Frame are updated in response to the perceptual process, those of Word-, Relation-, and Behavior-Multi-Dimensional Memory Frame are updated in response to the cognitive process, and those of Motor-Multi-Dimensional Memory Frame are updated in response to the motor process. In the memory shown in the upper right of Figure 1, focusing on memories that serve as traces of PCM process operation, Multi-Dimensional Memory Frame is classified into Perceptual-, Word-, Relation-, Behavior-, and Motor-Multi-Dimensional Memory Frame. This represents a structured approach focused on *memory updates* associated with the execution of the PCM process.

Imitation—doing what one observes—forms the basis of behavior. Therefore, the results of imitative behavior are structured and represented within the Multi-Dimensional Memory Frame, and imitative behavior arises through the propagation of activity within the Multi-Dimensional Memory Frame. Imitation occurs between individuals, and imitative behaviors exist that are passed down across generations. Therefore, the Multi-Dimensional Memory Frame utilized by the PCM process and updated through its execution can be organized from the perspective of *memes* inherited across generations [14].

Word is considered the archetype of meme [15]. Words, i.e., symbols, exist within the individual's Word-Multi-Dimensional Memory Frame as language. Through the use of language, individual languages or cultural languages are organized within the Multi-Dimensional Memory Frame. Individual language (behavioral-level memes) is the language used in person-to-person communication, encompassing not only direct usage but also potential metaphorical usage, as well as thesauri—lists of words grouped by synonyms or related concepts. Cultural language (cultural-level memes) is language used within a cultural context where a proper understanding of the established common sense within a specific community is essential for successful communication. These languages develop during the early developmental stage from birth to age three, associated with objects in the environment (motor-level memes) encoded within neural networks. These circulate among people and persist across generations [16].

The relationship between the three levels of memes mentioned above and Multi-Dimensional Memory Frame is as follows:

- C-memes represent culture stored in the Relation- and Word-Multi-Dimensional Memory Frame.
- B-memes represent behaviors in the environment stored in the Behavior-Multi-Dimensional Memory Frame.
- A-memes represent bodily actions stored in the Motor-Multi-Dimensional Memory Frame.

D. P-Resonance

This section explains the mechanism by which P-resonance occurs between the perceptual process and Perceptual-Multi-Dimensional Memory Frame, based on our previous work [10].

1) Characteristics of the Perceptual-Multi-Dimensional Memory Frame: Humans act within an environment that changes moment by moment. The actions generated are observed as a continuous sequence of events containing environmental information in time and space.

Memory is formed by sampling this continuous sequence. Therefore, at the instance when memory is formed, information regarding absolute time and spatial coordinates is lost. The memory represented by the Multi-Dimensional Memory Frame, which has no position and time data, represents relationships between objects. Therefore, the contents of the memory can be reused by performing a topological transformation to match the current time and spatial scales. Here, by binding spatiotemporal information about actions to memories that contain only relational information between events, these memories become executable in the real world.

The accuracy of current perceptual information can be enhanced based on past memories and associated time and space values. This is because the Perceptual-Multi-Dimensional Memory Frame is simply a memory of object relationships; the Behavior-Multi-Dimensional Memory Frame remembers object relationships associated with the time of the entity's actions in the procedures of physical behavior; the Relation-Multi-Dimensional Memory Frame stores object relationships associated with procedures and temporal concepts recognized by System 2. There is an overlap between Behavior- and Relation-Multi-Dimensional Memory Frame, and through this overlap, the information in Perceptual-Multi-Dimensional Memory Frame is enhanced.

2) Basic Senses: To generate appropriate actions synchronized with the environment in ever-changing situations, it is necessary to bind spatiotemporal information to Multi-Dimensional Memory Frame and render Motor-Multi-Dimensional Memory Frame executable. As a foundational concept to solve this binding problem, the concept of basic senses was proposed, which was identified as the actual manifestation of P-resonance. [10]. P-Resonance occurs between the external stimulus and the Perceptual-Multi-Dimensional Memory Frame by rhythmic and spatial senses. Subsequently, cognitive objects are generated by utilizing the number sense.

a) Rhythmic Sense: The changes brought about by actual human action are micro changes on cyclic activity. These minute changes alter the relative situation between the actors themselves and the environment that involves others. However, from the three-dimensional understanding of human perception, these changes are perceived as continuous changes along the time axis. Conversely, if we look at the organs active in the human body, they have evolved and developed under circadian rhythms. Consequently, periodically active organs such as the heart have been formed to provide unique rhythms.

Turning to the environmental side, changes with various reproducible rhythms occur under the cyclic activity of the earth. Thus, to adapt to changes in the environment, a “rhythmic sense of basic perception” should be formed in the connection circuit of the circulatory network formed by various procedural memories. This is called rhythmic sense that enables flexible binding of memory and perceptual information on the time axis in P-Resonance, with hearing as the core and perception in general.

Here, through P-resonance based on rhythmic sense, the region that matches the temporal cyclic pattern of the memory held within the Perceptual-Multi-Dimensional Memory Frame is activated.

b) Spatial Sense: Bodily activity includes movement that involves changing the position of one’s own body part in the three-dimensional space. Recognition of the current situation of the three-dimensional space is necessary for constructing executable bodily activities from the information stored in the Motor-Multi-Dimensional Memory Frame, that is free from absolute positions. The unique dimensions associated with movement are distance and time, which are required to make the move. The time is associated with body’s internal rhythms, which define the scale for measuring distance. Thus, the information concerning distance between objects in the external environment is conceived through the rhythm-based scale. This is the second basic sense, the “spatial sense” defined through the rhythmic sense.

Furthermore, the scale can change overtime because human behavior changes its orientation and range of circulation as it grows. Consequently, the cyclic trajectory thus formed evolves into a complicated web reflecting the range of variations of movement; humans expand their activity bandwidth. Inevitably, “spatial sense of basic perception”, which serves as the basis for spatial cognition, should be formed in the connection circuit of the circulatory network formed by various procedural memories, which works in P-Resonance.

Here, procedural memory is formed by overlapping Perceptual-Multi-Dimensional Memory Frame with Behavior- and Relation-Multi-Dimensional Memory Frame, so P-resonance causes activation in that region.

c) Number Sense: For humans to select appropriate actions in a timely manner in an ever-changing environment, information related to quantitative comparisons such as larger or smaller for size, more or fewer for the number of objects, farther or closer for distance, and longer or shorter for duration is indispensable. When this information is combined with the reward response that reflects the appropriateness of the choice, a basic sense of quantitative discrimination is formed. This is the third basic sense, the “number sense” [17].

Perceptual information activates Perceptual-, Relation-, and Behavior-Multi-Dimensional Memory Frame through P-resonance mediated by rhythmic and spatial senses. This process takes place in the midst of synchronization between the environment and human activity, which is weak synchronization, synchronization within the width of the time or activity bandwidth [18]. Number sense integrates them into cognitive objects that can be consciously manipulated.

The cognitive objects generated by the number sense are relational networks composed of the following items.

1) The Perceptual-Multi-Dimensional Memory Frame activated by the rhythmic sense,
2) Existing in its overlapping domains, and associated with Relation-, Behavior-Multi-Dimensional Memory Frame, and Perceptual-Multi-Dimensional Memory Frame activated by spatial sense,
3) Symbols existing in the overlapping region of 1 and 2 within the Word-Multi-Dimensional Memory Frame.

The number sense activates these regions, making the Multi-Dimensional Memory Frame available for cognitive processes via C-resonance.

E. C-Resonance

This section explains the mechanism by which C-resonance arises between cognitive processes and Word-, Relation-, Behavior-, and Motor-Multi-Dimensional Memory Frame, based on our previous work [8].

1) Characteristics of the C-Resonance: C-resonance occurs between cognitive objects represented as activation patterns in Perceptual-, Behavior-, Relation-, and Word-Multi-Dimensional Memory Frame and System 1 and System 2, driving the cognitive process. As a result, activation propagates to Motor-Multi-Dimensional Memory Frame, mapping to the brain’s representation of motor movement. This is converted into actionable information through spatiotemporal interpolation during the motor process, and movement occurs via motor nerves. The body plan (skeleton) underlies the entire movement as an interpolation program base, functioning as a default value.

The Multi-Dimensional Memory Frame contains three hierarchically structured memes—C-, B-, and A-memes—as shown in the upper right of Figure 1. These memes correspond to each memory in the Multi-Dimensional Memory Frame. Specifically, C-memes corresponds to Word- and Relation-Multi-Dimensional Memory Frame, B-memes corresponds to Behavior-Multi-Dimensional Memory Frame, and A-memes corresponds to Motor-Multi-Dimensional Memory Frame. C- and B-memes are linked to each other by sharing the Perceptual-Multi-Dimensional Memory Frame. B- and A-memes are related through the overlap between the existence of the former in Behavior-Multi-Dimensional Memory Frame and the existence of the latter in Motor-Multi-Dimensional Memory Frame. These associations form a structure that binds them together. This ensures the reality derived from perceptual information linked to the real world.

2) Implementation of C-Resonance by GOMS: C-resonance operates under time constraints where the PCM process must synchronize with environmental changes while selecting and executing actions, linking the Multi-Dimensional Memory Frame with cognitive processes. In our previous work [8] proposing C-resonance, we introduced GOMS, proposed by Card, Moran, and Newell [1], as a mechanism to directly link the Word-, Relation-, Behavior-, and Motor-Multi-Dimensional Memory Frame without going through Perceptual-Multi-Dimensional Memory Frame.

Using memes in an interpreter-like manner via Perceptual-Multi-Dimensional Memory Frame ensures corporeality, but it requires monitoring the situation step by step while advancing the PCM process, making it inefficient. This situation corresponds to MHP/RT running in Mode 2 under System 2 control. Meanwhile, the use of Multi-Dimensional Memory Frame in the form compiled by GOMS corresponds to situations where MHP/RT is running in Mode 1 under System 1 control. This allows us to consider GOMS as implementing C-resonance.

3) Binding Memes via GOMS: In GOMS, behavioral goals form a robust hierarchical structure. The goal structure mediates the organization of behavior. Achieving a goal, G, requires achieving the subgoals underneath it, G's. This structure does not hold the time as its primary parameter. The order between G's is important. The time elapsed for executing G' is associated with the operators located at the bottom layer, which connect to the motor process of PCM that implements the contents of Motor-Multi-Dimensional Memory Frame—i.e., the operators—in the real world.

One may reasonably assume that GOMS is used to structure A-, B-, and C-memes which do not contain absolute spatiotemporal information as a method of realizing behavior generation without breaking down, while keeping in sync with the real world where the situation changes from moment to moment. GOMS should correspond to the phenomenon of A-, B-, and C-memes binding without the Perceptual-Multi-Dimensional Memory Frame when encountering certain situations, indicating the entity of the phenomenon of C-resonance. This may correspond to a shortcut that may be formed within the Multi-Dimensional Memory Frame. Here, GOMS is positioned as a shortcut that guarantees symbol grounding without going through perception.

4) Symbol Grounding via GOMS: The embodiment of one's conceptualized ideas in the real world as envisioned—symbol grounding—is crucial for facilitating smooth interaction.

Figure 1 top right shows GOMS substituting the mapping from the Word- to Motor-Multi-Dimensional Memory Frame, where symbol grounding is guaranteed. The self propagates activations through Perceptual-Multi-Dimensional Memory Frame to Relation- and Behavior-Multi-Dimensional Memory Frame, ultimately activating Motor-Multi-Dimensional Memory Frame, for the concept $\hat{\mathbf{C}} \in$ C-memes that the self intends to realize in the real world. The goal expressed here is a kind of concept. Then, through the motor process, the object $\hat{\mathbf{O}}(\hat{\mathbf{C}})$ embodies itself in the real world. What is embodied here is the state in which the goal has been achieved, existing within the environment as a perceivable object. This object is perceived by the self, and the Perceptual-Multi-Dimensional Memory Frame is activated via P-resonance. The activation propagates to Word-Multi-Dimensional Memory Frame, and when the object is recognized as $\hat{\mathbf{C}}$, the relationship, $\hat{\mathbf{C}} \equiv \hat{\mathbf{O}}(\hat{\mathbf{C}})$, holds true, meaning the envisioned concept has been successfully realized in the real world. In other words, it means that the envisioned goal has been realized in the environment. The establishment of this relationship also signifies the resolution of the symbol grounding problem [19], which has been a challenge in the field of artificial intelligence [7].

III. Resonant Artificial Object Design

Figure 3 illustrates that the following occurs for the concept $\hat{\mathbf{C}} \in$ the Word-Multi-Dimensional Memory Frame—C-memes—corresponding to the state one intends to achieve. First, the cognitive process resonates with GOMS, which includes the corresponding goal, activating System 2 and System 1, thereby generating the object $\hat{\mathbf{O}}(\hat{\mathbf{C}})$ in the environment. Subsequently, this object captures the basic senses, inducing P-resonance and forming cognitive objects in the Multi-Dimensional Memory Frame. If the cognitive objects are linked to $\hat{\mathbf{O}}$, symbol grounding is guaranteed. Below are key design considerations to ensure sufficient C-resonance and P-resonance in artificial objects.

A. Key Considerations for Promoting C-Resonance

Since the true nature of C-resonance is GOMS, it is necessary to represent situations where virtual environments are utilized using GOMS. This is achieved by constructing a GOMS representation of human behavioral ecology within the target virtual environment. GOMS organizes C-, B-, and A-mimes into a hierarchical structure as shown in Figure 3. Therefore, the procedure for constructing GOMS begins with extracting

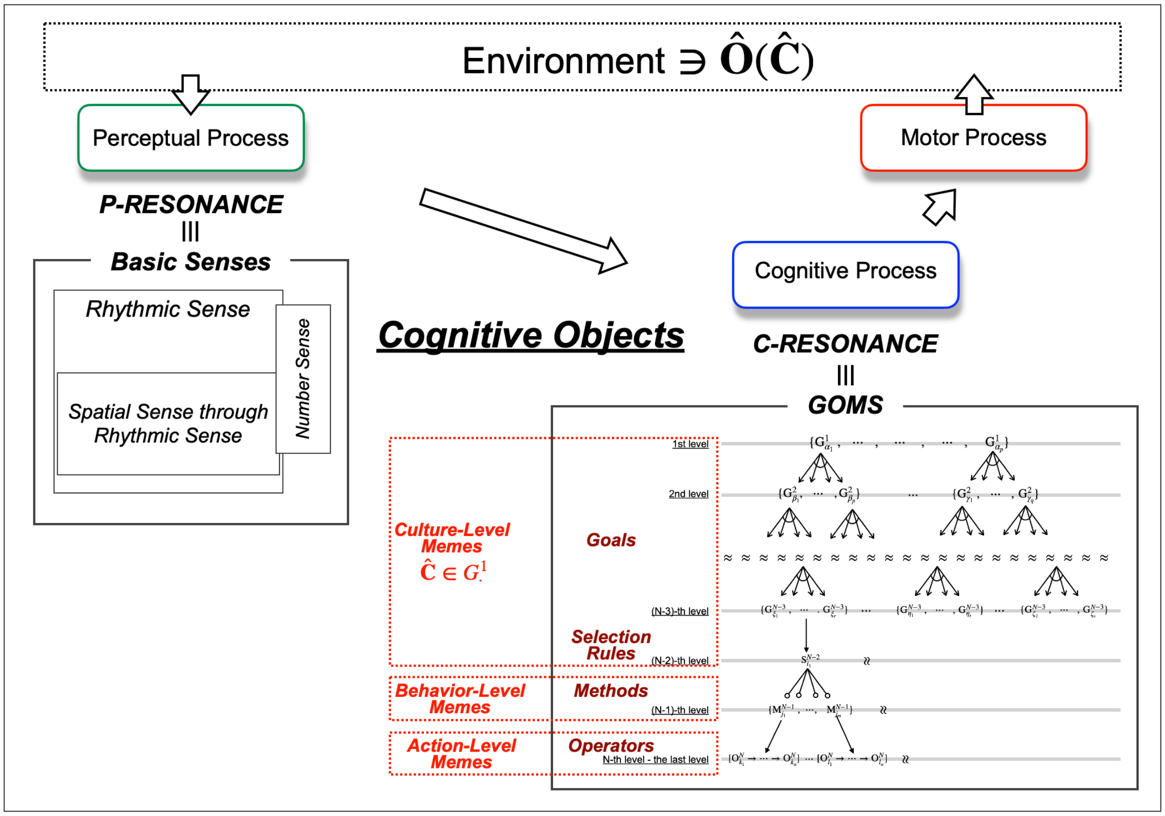

Figure 3. Activation of environment-synchronized PCM processes and environment-asynchronous Multi-Dimensional Memory Frame—P-resonance via basic senses, C-resonance via GOMS, and the cognitive objects connecting them.

memes that are inherited across generations within the target behavioral ecology. Here, a method known as Cognitive Chrono-Ethnography (CCE) [20] proves effective. This involves conducting behavioral simulations using MHP/RT to identify parameters characterizing behavioral patterns, observing human actions exhibiting distinctive combinations of these parameter values, and thereby providing detailed descriptions of behavioral ecology. This method has been implemented in several case studies [21][22].

A-memes are observable. Mutually distinguishable A-memes define the operators "O". B-memes are pattern sequences associated with methods "M", which are patterns reproducibly generated when conditions are met by sequences of A-memes. The conditions define selection rules "S", expressed as "IF condition THEN pattern sequence name" consciously processed by System 2. Upon this, goal structures unfold.

At the top, goals related to happiness and satisfaction—such as winning competitions and gaining intellectual fulfillment—are placed, with the hierarchy unfolding downward. A goal structure is expressed by linking goals represented as concepts in the form of goals and subgoals. Note that a higher-level goal is achieved by accomplishing all the lower-level subgoals connected to it.

GOMS does not consciously evaluate the results of actions because higher-level goals are set by System 2, and System 1 executes actions via feedforward control—execution in Mode 1. To promote C-resonance, it is necessary to accurately grasp in advance the goal structure that humans are attempting to achieve, identify the goal being pursued from the observed sequence of actions, and then appropriately infer the next goal that will likely be set.

B. *Key Considerations for Promoting P-Resonance*

The rhythmic sense, spatial sense, and number sense are applied to generate cognitive objects from the resonated portion of the Perceptual-Multi-Dimensional Memory Frame for further cognitive processing carried out by System 1 and System 2. The accuracy of adaptive behavior can be increased by repeating the behavior while developing the rhythmic, spatial, and number senses in the reuse of the Perceptual-Multi-Dimensional Memory Frame. Objects in the environment $\hat{O}$ serve as the starting point for P-resonance, determining the activation patterns of the Perceptual-Multi-Dimensional Memory Frame through rhythmic, spatial, and number senses. The content of objects that individuals can interact with within their limited time influences the development of basic senses. Therefore, it is necessary to design objects $\hat{O}$ in the environment to elicit responses through basic senses.

1) Conditions that Trigger Rhythm Sense: Humans form the rhythmic sense as fundamental perception in the connection circuits of a cyclic network formed by various procedural memories in order to adapt to environmental changes. A

Copyright (c) IARIA, 2026. ISBN: 978-1-68558-381-1

perceived rhythm is a subjective grouping of events occurring successively along the time axis generated through the rhythmic sense, and its nature is primarily determined by the length of the intervals between events. When events recur periodically according to a certain pattern, the temporal structure of that pattern is recognized as rhythm. The most fundamental element of rhythm is the existence of regularity or repetitive patterns in the intervals between events. Through this regular interval, humans can predict what will happen next and perceive it as a coherent whole. The regularity and periodicity of the intervals between events play a central role in perceiving and defining rhythm. Therefore, the events arising from interaction generating patterns along the time axis become a necessary condition for rhythm perception to be triggered.

2) Conditions that Trigger Spatial Sense: The determination of rhythm signifies that the unit time $\hat{T}$ characterizing events occurring in the environment is established. Event $E(T)$ occurring at time T becomes comparable to an event occurring at time $T + \hat{T}$, and the difference $D(E(T + \hat{T}), E(T))$ between the events becomes discernible. When events occurring in three-dimensional space are the subject, spatial cognition—the cognitive representation of space—is generated based on perceived differences. This spatial cognition is referenced when recognizing and predicting changes in the relative position between the environment and oneself over time.

Objects in the environment can evolve over time within three-dimensional space according to their own programmed dynamics and generate events. However, it is necessary to consider that the receiving human side quantizes events according to a rhythm characterized by the unit time $\hat{T}$.

3) Conditions that Trigger Number Sense: By utilizing spatial sense, the temporal development of the relative relationship between the environment and the self is generated as activation patterns of Perceptual-, Behavior-, and Relation-Multi-Dimensional Memory Frame. To advance the situation toward the immediate behavioral goal, it is necessary to consolidate its activation patterns into a symbol and evaluate them against the behavioral goal defined by System 2. Number sense is used to select information concerning quantitative comparisons of about three items, which is necessary for choosing appropriate actions at the right time in an environment that changes moment by moment. The evaluation can be conducted in a multifaceted manner, but comparisons involving three or fewer items are preferable. The parameters characterizing programmed changes in the external environment must be linked to the evaluation of whether the goal has been achieved.

IV. Conclusion and Future Work

It is desirable for people to strive toward achieving their goals, interact with their surrounding environment, evaluate the results, and gain a sense of accomplishment. When considering a situation where the environment returns pre-programmed responses to human actions, the environment can be regarded as virtual in the sense that it is not the actual environment. As a result of one's own actions, some objects are returned from the virtual environment. If one can identify the content of their own actions within these objects and use it to set the content for the next action, the interaction between the human and the virtual environment can be considered to be progressing smoothly.

Interaction is executed through a PCM process that synchronizes with the ever-changing environment. The execution content is determined by the activation patterns of the Multi-Dimensional Memory Frame, which operate independently of environmental changes. To maintain smooth interaction, it is important to control the content of P-resonance generated by basic senses and the content of C-resonance generated by GOMS. This paper has outlined the matters that should be considered in these controls.

Guidelines are generally expressed in abstract terms, so to apply them to real-world artificial environment design scenarios, implementation methods for the guidelines must accompany them. This study showed that designs aiming to induce P-resonance and C-resonance, and designs enabling the functioning of basic senses and GOMS, are effective as practical implementations of the guidelines [13] proposed in relation to the MHP/RT operating modes (specifically Mode 1) shown in Figure 2. Whether the basic senses function can be examined by providing representations for the perceptual information provided. GOMS can be examined by representing the knowledge required for action selection. Each of these is fundamentally based on static analysis.

In P-resonance, it was pointed out that it is important to consider that the activation of Perceptual-, Relation-, and Behavior-Multi-Dimensional Memory Frame originating from perceptual information occurs in situations where the environment and human activity are weakly synchronized. When a system and a human resonate, the two systems intertwine and synchronize at a specific point in time. Meanwhile, weak synchronization [18][23] is a concept of synchronization applied to human activities occurring within a temporal or activity bandwidth, extending the aforementioned concept of synchronization. Specifically, actions executed at time T involve $System\ 2$ and $System\ 1$ executed beforehand by MHP/RT, and $System\ 1$ and $System\ 2$ executed afterward. These four processing procedures synchronize with the system event at time T. This synchronization method is called weak synchronization. We also propose design guidelines related to weak synchronization [13].

Just as practice became possible by focusing on basic senses and GOMS—concepts external to MHP/RT—the next challenge is to make artifact design guidelines, derived from weak synchronization (the conditions for P-resonance) associated with MHP/RT's four processing modes, practicable by introducing concepts external to MHP/RT. Here, dynamic analysis focusing on the nature of interactions over time is fundamental. For example, the degree of weak synchronization in situations where latent information about visual objects presented in a VR learning environment is provided auditorily can be examined by measuring visual behavior and pupil responses, and analyzing the relationship between these results and memories of audiovisual experiences [24][25].

Acknowledgment

This work was supported by JSPS KAKENHI (Grant Numbers 19K12246 / 20H04290 / 22K12284 / 23K11334) and the National University Management Reform Promotion Project.

References

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

[2] A. Newell, "Physical Symbol Systems", *Cognitive Science*, vol. 4, no. 2, pp. 135–183, 1980. DOI: 10.1207/s15516709cog0402_2[retrieved:February,2026].

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

[4] J. E. Laird, A. Newell, and P. S. Rosenbloom, "Soar: An architecture for general intelligence", *Artificial Intelligence*, vol. 33, pp. 1–64, 1987.

[5] 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, ISSN: 2212-683X. DOI: 10.1016/j.bica.2013.05.003[retrieved: February,2026].

[6] M. Kitajima, *Memory and Action Selection in Human-Machine Interaction*. Wiley-ISTE, 2016, ISBN: 9781848219274.

[7] M. Kitajima, M. Toyota, and K. T. Nakahira, "Why the Symbol Grounding Problem Matters in Virtual Reality: A Meme-Focused Solution Based on the Model Human Processor with Real-Time Constraints", *International Journal On Advances in Intelligent Systems*, vol. 18, no. 3 & 4, pp. 162–172, 2025.

[8] M. Kitajima, M. Toyota, J. Dinet, and K. T. Nakahira, "Transforming Conscious Goals into Unconscious Actions in Real-world Interactions: Real-world Use of Behavioral Ecological Memes via GOMS", *International Journal On Advances in Intelligent Systems*, vol. 18, no. 3 & 4, pp. 173–186, 2025.

[9] 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. DOI: 10.1080/0144929X.2011.602427[retrieved:February,2026].

[10] M. Kitajima et al., "Basic Senses and Their Implications for Immersive Virtual Reality Design", in *AIVR 2024 : The First International Conference on Artificial Intelligence and Immersive Virtual Reality*, 2024, pp. 31–38.

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

[0] D. Kahneman, *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux, 2011.

[13] M. Kitajima, M. Toyota, and J. Dinet, "Guidelines for Designing Interactions Between Autonomous Artificial Systems and Human Beings to Achieve Sustainable Development Goals", *International Journal on Advances in Intelligent Systems*, vol. 15, pp. 188–200, 2022.

[14] 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*, 2008, p. 2288.

[15] D. C. Dennett, *From Bacteria to Bach and Back: The Evolution of Minds*. W W Norton & Co Inc, Feb. 2018.

[16] M. Kitajima, M. Toyota, and J. Dinet, "How Resonance Works for Development and Propagation of Memes", *International Journal On Advances in Systems and Measurements*, vol. 14, pp. 148–161, 2021.

[17] S. Dehaene, *The Number Sense: How the Mind Creates Mathematics, Revised and Updated Edition (English Edition)*. Oxford University Press, Apr. 29, 2011, p. 604.

[18] J. Dinet and M. Kitajima, "Immersive interfaces for engagement and learning: Cognitive implications", in *Proceedings of the 2015 Virtual Reality International Conference*, ser. VRIC '18, Laval, France: ACM, 2018, 18/04:1–18/04:8. DOI: 10.1145/3234253.3234301[retrieved:February,2026].

[19] S. Harnad, "The symbol grounding problem", *Physica D: Nonlinear Phenomena*, vol. 42, no. 1, pp. 335–346, 1990, ISSN: 0167-2789. DOI: 10.1016/0167-2789(90)90087-6[retrieved: February,2026].

[20] 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.

[21] M. Kitajima, M. Toyota, and J. Dinet, "Art and Brain with Kazuo Takiguchi - Revealing the Meme Structure from the Process of Creating Traditional Crafts -", in *COGNITIVE 2023 : The Fifteenth International Conference on Advanced Cognitive Technologies and Applications*, 2023, pp. 1–10.

[22] K. T. Nakahira, M. Kitajima, and M. Toyota, "Understanding Practice Stages for a Proficient Piano Player to Complete a Piece: Focusing on the Interplay Between Conscious and Unconscious Processes", *International Journal on Advances in Life Sciences*, vol. 16, pp. 164–177, 2024.

[23] 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.

[24] Y. Kurihara, M. Shino, K. Nakahira, and M. Kitajima, "Visual behavior based on information foraging theory toward designing of auditory information", in *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 1: HUCAPP*, INSTICC, SciTePress, 2024, pp. 530–537. DOI: 10.5220/0012474600003660[retrieved:February,2026].

[25] R. Hirabayashi., M. Shino., K. N. T.., and M. Kitajima., "How auditory information presentation timings affect memory when watching omnidirectional movie with audio guide", in *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 2: HUCAPP,*, INSTICC, SciTePress, 2020, pp. 162–169. DOI: 10.5220/0008966201620169[retrieved: February,2026].