Applied Artificial Intelligence, 21:155–183

ISSN: 0883-9514 print/1087-6545 online
DOI: 10.1080/08839510601170713

# A BIOLOGICALLY INSPIRED METHOD FOR CONCEPTUAL IMITATION USING REINFORCEMENT LEARNING

**Hossein Mobahi, Majid Nili Ahmadabadi, and Babak Nadjar Araabi** □ *AI and Robotics Lab, Control and Intelligent Processing Center of Excellence, Department of Electrical and Computer Engineering, University of Tehran and School of Cognitive Sciences, IPM, Tehran, Iran*

□ *Learning by imitation can provide agents with a natural and effective means for transferring knowledge when brain-to-brain connection is infeasible. Natural mechanisms for imitation demonstrate strong abstraction and conceptualization capabilities, however, computational models that have been proposed for imitative learning barely address these fundamental features.*

*Inspired by functions of human brain constituents and exploiting ideas enthused by mirror neurons and the multi-store model of memory, we propose a new model for learning by imitation capable of developing relational concepts. In our model, memory gradually organizes sensory data into concepts through reinforcement learning and consolidation, while mirror neurons maintain an extendible repertoire of familiar actions connected to corresponding concepts. We also discuss the relation between modeling behavior of concept-oriented agents in terms of mathematical functions and relevant biological evidence of mirror neurons. Eventually, we evaluate our method in a phoneme acquisition experiment through real interaction with humans.*

Humans and social animals acquire some parts of their knowledge through interaction with other members of their societies. Such social learning processes may have different forms. When an agent (human, animal, or even machine) tries to learn a task by observing how others do it, the process is called "imitation." Although there is still debate on the exact definition of imitation (Byrne and Whiten 1988; Tomasello 1990), most researchers agree that it is different from mimicking. While mimicking is merely involved in recording and reproducing observed actions, imitation needs some sort of abstraction and understanding of observations (Arbib 2000; Breazeal and Scassellati 2000).

Learning by imitation is appealing to the engineering community, particularly to those in AI and robotics, due to its distinctive benefits. The most

Address correspondence to Majid Nili Ahmadabadi, AI and Robotics Lab, Control and Intelligent Processing Center of Excellence, Department of Electrical and Computer Engineering, University of Tehran, 14114 Tehran, Iran. E-mail: mnili@ut.ac.ir

notable benefit is that it provides agents with some natural and effective means for transferring knowledge from one to another, when brain-to-brain connection is not possible. This is particularly important for heterogeneous agents that have different brain structures. For instance, in human robot interaction (HRI) applications where the interaction must be natural, imitative learning can be an effective solution. Several works have been carried out to apply imitative learning in HRI applications. For instance, Schaal et al. (1997) trained a robot arm to handle a pendulum-balancing task by observing a human. Kuniyoshi et al. (1994) also proposed a method for learning a stacking blocks problem that is performed by a human.

A part of robotics research looks at learning by imitation from a purely engineering viewpoint. This is the classic approach to imitative learning which was shaped in 1980 when manipulators had to be programmed manually. This was a tedious and time-consuming task. AI, particularly symbolic reasoning, was a common choice at that time to convert a measured human's arm trajectory to a set of "IF-THEN" rules (Lozano-Pérez 1982; Dufay and Latombe 1984). In later works, although this approach was facilitated to some extent, the core method has remained the same. For instance, instead of manually providing the agent with trajectory data, the robot could autonomously extract it using computer vision techniques (Kuniyoshi et al. 1994; Tung and Kak 1995).

The main problem with the classic approach is its domain-limitedness. The solutions are very task specific and can be utilized in highly controlled environments only.

Inspired by imitative learning in natural species, another line of research has been formed to overcome the drawbacks of pure engineering approaches. This approach hopes to achieve the robustness and adaptability existing in biological intelligence. Fortunately neuroscientists and cognitive scientists have been seriously studying learning by imitation to compare similarities and differences in human and animal cognition. Although this goal is different from what engineers aim at, it has provided them with useful information for proposing biologically inspired models of imitative learning. However, there is still a long way to achieve robust and adaptive models like the natural ones.

Existing biologically inspired models are still complex and task-specific. For instance, all of the models introduced by Arbib et al. are tailored for learning how to grasp (Fagg and Arbib 1998; Oztop and Arbib 2002; Oztop et al. 2004). Recently, Jenkins et al. (2004) have addressed the limitations of the models that have been proposed so far. One of these limitations is that the learned behavior is different from what a human teacher desires. They have proposed some reasons, such as limited amount of training set, to cover the whole work space or the over fitting problem where a robot loses its generalization.

We add the inability of imitator to learn abstract concepts to these reasons. If concepts are abstract, then the agent cannot achieve the behavior that the teacher desires. This issue has not been studied much. In fact, the existing models have chosen evaluation testbeds that can be understood by their perceptual characteristics, where some distance metric can be defined for measuring similarities in perceptual space. For instance, Jenkins and Mataric (2003) used spatial isomap (which is based on Euclidian distance among points) to categorize recorded joint angles of a demonstrator's movements. Categorization of perceptual data has a major role in their learning model. In contrast, in this work, our concern is problems that cannot be understood merely from perceptual information.

Abstract concepts may be scattered irregularly in the agent's perceptual space. However, they are grouped together according to some principles, like their functional similarity. This is what the definitions of "learning by imitation" emphasize, but which the existing models lack. In the next section, we will see that neural mechanisms that are believed to be responsible for imitative learning have also an abstract representation of actions. Maybe that is why they can handle a wider range of problems. Obviously, similarity metrics are not always available, especially in abstract problems.

In this article we introduce a new and biologically inspired model for learning by imitation.[1] Unlike previous works, we consider conception/abstraction as an essential element in our method. Abstraction and conception by imitation requires some similarity metrics in the perceptual and the conceptual spaces. Hand design of such similarity metrics is a data-dependent and very difficult task, if not impossible. Therefore, we use reinforcement or emotional learning to help the imitator to learn these metrics through conceptualization and abstraction of the stimuli implicitly and automatically.

## MIRROR NEURONS

A group of researchers studying neuron activations in monkeys could discover an interesting region in their brain named F5 (Rizzolatti and Gentilucci 1988). The behavior of these neurons is such that they possess both perceptual and motor characteristics and respond to both types of stimuli. The most important finding about these neurons is that the mirror neurons which are able to recognize an action are also able to produce it. In other words, mirror neurons have a common representation of doing and observing an action. That is why these neurons are believed to be an important neural mechanism of imitation. Although mirror neurons were first discovered in monkeys, there is some evidence that the Broca area in the human brain is a homologue of mirror neurons and has similar characteristics (Fadiga et al. 1995).

The fact that representation in mirror neurons is action-based has an important byproduct, which is abstraction. For instance, some recent discoveries in the F5 area have identified audiovisual mirror neurons (Keysers et al. 2003). These neurons respond not only when a monkey observes someone breaking a peanut, but also when it only hears the sound of breaking a peanut. This indicates that mirror neurons encode actions not observations. As the two discussed stimuli belong to completely different perceptual spaces, this indicates that mirror neurons possess abstraction capability.

Another interesting feature of mirror neurons is that there seems to be a one-to-one correspondence between perceived concepts and available actions. This idea is supported by experiments that show mirror neurons have different congruency. Some neurons have a broad congruence; for instance, they respond to any type of grasp action. On the other hand, some have a limited congruence; for example, they respond to very specific types of grasp. In fact, although some actions are a subset of others, there is still a separate representation for each action and no combination of basic actions seems to occur in mirror mechanism.

Several models have been proposed to simulate the behavior of mirror neurons. For instance, Arbib et al. (Fagg and Arbib 1998; Oztop and Arbib 2002; Oztop et al. 2004) have introduced models FARS, MNS1, and IGLM. However, all these models are focused on grasp learning. They have hard-coded a mechanism in these models so that their representation becomes invariant to visual changes. If these models were applied to problems other than grasp, this hard-coding must be repeated again and again. This is in contrast with the goal of the biological approach of learning by imitation.

## CONCEPTS

As abstraction and concept-oriented imitation are the motivation and the core of our work, it is useful to first review some concept definitions and taxonomies. A concept is an internal representation of the world in an agent's mind. It can be a set of objects or events that are similar with respect to a principle (Zentall et al. 2002). Utilizing concepts in the mind of an intelligent agent has a number of advantages:

*Generalization*: By experiencing different instances of the same concept and analyzing the similarities and differences, a model for these instances can be constructed. This model is a general description of the concept and is able to satisfy its diverse instances, even the novel ones.

*Communication*: If there is a concept for each observation, during communication, agents can refer to the corresponding concept instead of describing all details of the observation. Note that, in this case, the communicating agents should have a similar interpretation of the concept.

*Cognitive Economy*: Thanks to the compact concepts in an agent's knowledge structure, there is no need to repeat details when thinking. This speeds up and simplifies cognitive processing.

Concept acquisition in natural environments must be able to cope with some constraints (Davidsson 1994). First, concepts should be learned gradually because we do not encounter all instances of a concept at one point in time. Moreover, they should be learned in parallel, i.e., the type or the order of incoming instances is arbitrary. At last, learning must be accomplished relatively fast in the sense that we are able to learn a fairly useful concept representation just by encountering instances of the category on a few occasions.

## Abstract Levels

Observations are categorized to concepts with respect to some principles that depend on physical and/or functional characteristics of the items. From this perspective, Zentall et al. have categorized concepts to three levels of abstraction (Zentall et al. 2002) (see Figure 1):

*Perceptual*: These concepts are formed solely by measuring similarity of instances in perceptual space. Such data can be categorized by simple clustering algorithms in an unsupervised fashion.

*Relational*: In this type of concept, although perceptual similarity still contributes to categorization, it is not sufficient to form the correct concepts. External information must link perceptual categorizes and form the right concept. This is achieved by classical conditioning.

*Associative*: In learning these concepts, the stimuli within classes bear no obvious physical similarity to one another, but cohere because of shared functional properties.

## Same/Different Judgment

As we move from perceptual toward associative concepts, more complex cognitive capabilities are required. Fortunately relational concepts

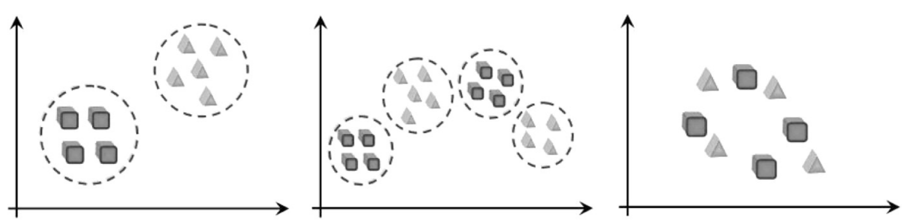

**FIGURE 1** Three types of concept: perceptual (left), relational (middle), and associative (right) in a two-dimensional feature plane.

are in the middle level of abstraction and therefore easier for learning, compared with associative concepts. However, the perceptual capability of an agent is no longer sufficient (but necessary) to form target concepts. Therefore, complementary information must be given to the agent by the teacher. This information can unify scattered clusters that belong to the *same* concept. Now the main question is how this information can be transferred from the teacher to the learning agent.

A solution to learning relational concepts is *same/different* judgment. In this technique, the learning agent is exposed to two stimuli. It must decide whether they belong to the same or to different concepts. A reward or punishment signal is issued by the teacher depending on correctness or incorrectness of the learner's decision. So the learning agent gradually develops abstract concepts to increase its reward. Eventually the agent will be able to correctly classify novel stimuli of the learned concepts. It has been shown that pigeons, parrots, rhesus monkeys, baboons, and chimpanzees are capable of learning abstract concepts by the same/different method (Cook et al. 2003).

The success of relational concept learning by same/different judgment is valuable for the AI and robotics community for a number of reasons. First, relational concepts are the first step toward abstract concepts, yet they are easy to learn due to the contribution of perceptual information. Second, concept learning can be a solution to the symbol grounding problem (Harnad 1990), i.e., relating symbols to their physical characteristics, which is one of the most fundamental problems in AI. Third, since the technique is based on reinforcement learning, the training interface can be highly simplified, e.g., making use of the teacher's emotional state as the reinforcement signal.

## Concept Representation

There are three general theories of concepts, namely, exemplar, prototype, and rule theories.

*Exemplar*: Merely instances of a concept are memorized. A new stimulus is classified according to its similarity to all of the known instances of the various candidate concepts. The specification of contents (exemplars) is not a global summary but is instead a collection of piecemeal information. A perceived stimulus can be formally represented by its values in the perceptual space. The similarity of the stimulus to a memory exemplar corresponds to their proximity in the perceptual space.

*Prototype*: It might seem inefficient or wasteful to remember every instance of a category. Perhaps some sort of summary could be abstracted during learning, and then the individual cases could be safely jettisoned. The summary, also called a prototype, should be representative of the various instances of the category, e.g., average or idealized caricature of instances. Because a prototype has a value on every dimension of the stimuli, it can be formally represented like an exemplar, although a prototype need not correspond to any actually experienced instance.

*Rule*: A rule-based model uses either a strict match/mismatch process or a boundary representation. An example of a rule-based model is one that uses featural rules that specify strict necessary and sufficient conditions that define category membership. Rules are computationally attractive as concept representations because they can be uniformly applied to all stimuli, regardless of the instances actually experienced. Moreover, they can describe feature combinations that are not tied to the specific featural realizations. However, many natural categories are very difficult to specify in terms of content rules.

Our model is based on exemplars and flexible prototypes. In fact, as the number of prototypes per category increases, there can eventually be one prototype per instance, and such models become equivalent to exemplar models (Nosofsky 1984). Therefore, these theories are the extremes of a range of prototypeness. By flexible prototype we mean the one whose prototypeness degree can be adjusted.

In the last two decades we have observed a remarkable progress in the development of computational models for concept learning (Kruschke 1992; Love and Medin 1998). These models do not address how the developed concepts can actually influence an agent's world (through its actions). Conceptual imitation is our suggested solution for linking concepts to actions.

## MEMORY

The importance of memory in our life is obvious. However, what is in our focus about memory is its contribution to concept learning. In fact, psychologists have shown that the contents of long-term memory bias concept learning. For instance, the research of Merriman et al. (1997) on three-month-old infants supports this hypothesis. However, proposed models of concept learning have not seriously studied the role of memory in concept acquisition (Kruschke 1992; Love and Medin 1998). In fact, memory provides our model with acquisition of abstract concepts through interaction with a teacher agent.

A number of computational models for memory have been proposed such as levels of processing (Craik and Lockhart, 1972), transfer-appropriate processing, parallel-distributed processing (Rumelhart and McClelland 1986), and information processing (Atkinson and Schiffrin 1968). Historically, the latter model has been the most successful, influential, and comprehensive one. Therefore, we adopt this model in our work.

Information processing model suggests that the input information must pass through three different stages to be stored. These stages are three types of memory, namely, sensory memory, short-term memory (STM), and long-term memory (LTM). For this reason, this model is also called multi-store memory model.

Sensory memory is a register to keep sensory data. This information is transferred to STM immediately. STM has low capacity and it cannot hold as much information as LTM. Moreover, STM contents have a short life; they will be wiped out if not transferred to LTM. STM is generally used for holding information temporally. It acts as a scratchpad for the mind. Long-term memory has unlimited capacity and nothing decays or gets forgotten there. Information is transferred from STM to LTM by the hippocampus, a region in the human brain, through the process of consolidation.

LTM itself consists of different types of memory, such as episodic memory and semantic memory. Episodic memory stores time/space dependent events, while semantic memory is independent of these factors. Semantic memory stores concepts and semantic associations between them, forming a semantic network. Since this article is concerned with concept acquisition and has nothing to do with episodic memory, in the following, we will use the terms LTM and semantic memory interchangeably.

The hippocampus can be viewed as a control system between WM (Working Memory) and LTM and is responsible for consolidation. Rehearsal contributes to the consolidation process. There are two types of rehearsal, namely, maintenance and elaborative. Maintenance rehearsal means just repeating things over and over. Although this can help to keep information in WM for a short time, no consolidation occurs. In contrast, elaborative rehearsal means relating information in WM to existing knowledge in LTM and it is the main cause of consolidation. It was believed that the hippocampus is only involved in consolidation of episodic memory. However, recent studies indicate that the hippocampus also has an important role in learning and consolidation in semantic memory (Manns et al. 2003).

Scientific evidence indicates that events are better memorized in emotionally arousing situations (Burke et al. 1992; Christianson 1992). In addition, some researchers like Gold and McGaugh (1975) believe that the time that it takes the consolidation process completes is a biological advantage because it provides an opportunity to evaluate the emotional importance of the event. The emotional evaluation of stimuli is up to a part of brain called the amygdala (LeDoux 1996). Once the amygdala decodes the stimulus in terms of its emotional importance, the result propagates to the whole body by generating stress hormones. Although emotionally arousing stimuli are better learned and consolidated, excessive amounts of stress hormones have a negative effect on the hippocampus (LeDoux 1996). Therefore, emotion can have constructive and destructive effects on memory and consolidation. We will make use of this fact in our model to control consolidation by reinforcement learning.

## PROPOSED MODEL

### Goals

As discussed earlier, abstraction is what distinguishes imitation from mimicking. Moreover, the behavior of mirror neurons supports abstract representation of the brain's imitation mechanism. Nevertheless, none of the biologically inspired models in the literature addresses abstraction in sufficient depth. Therefore, achieving a model for conceptual imitation is our major goal. Another goal that we pursue in this work is that the proposed model can be applied to real-world problems. In this line, our model takes situatedness into account and allows the imitator to autonomously acquire its knowledge by interacting with teacher agents. Moreover, making use of reinforcement learning, which is connected to emotion in our work, provides a natural means for training robots by humans. The last purpose of our research is developing a simple and lightweight algorithm. It is important to note that although our work is inspired by biological observations, it is not biologically plausible. In fact, it works based on some rough ideas about the functionality of a few brain modules.

### One-to-One Correspondence in Mirror Neurons

Before introducing our model in detail, we once again return to the one-to-one correspondence hypothesis that mirror neurons seem to possess. In this section we will see the mathematical benefit of such one-to-one representation. In fact, if we assume that all modules of an arbitrary concept-oriented agent can be described by mathematical functions—not necessarily analytical ones, even a lookup table—we will prove that there always exists an agent with a different internal structure, particularly with one-to-one concept to action mapping, which is behaviorally equivalent to the original agent.

Such one-to-one correspondence provides the agent with a repertoire of known actions. By known actions we mean that the agent knows the abstract meaning of each concept in a way similar to what exists in mirror neurons through a link to its corresponding action and knowing how to perform the action in terms of motor parameters.

In order to analyze a concept-oriented agent mathematically, we formalize its behavior with respect to some basic but general building blocks. Although the model that we will introduce later has more details and it is easier to interpret, yet this mathematical representation fits well to the general structure that we analyze here.

As one of the goals was applicability of the model to real-world problems, particularly HRI, we assume that the teacher does not have access to the agent's brain and its behavior is the only thing that he can observe.[2] In fact, the teacher evaluates the agent's actions and issues reinforcement signals accordingly. Based on this assumption, we denote the ideal behavior of the learning agent by a mapping from the sensory data to the motor commands. By ideal behavior we mean the one which maximizes incoming reinforcement signal.

In our model, the agent itself has to conceptualize events in its mind. That means linking regions of continuous input and output spaces through a discrete concept set. So there are two functions involved, sensory-concept ($f$) and concept-motor ($g$) mappings; see Figure 2. Formally, we have:

$$\begin{aligned} &f : \Re^m \to \mathbf{N}_p \\ &g : \mathbf{N}_p \to \Re^n \\ &\mathbf{N}_p = \{1, 2, 3, \ldots, p\}, \end{aligned} \tag{1}$$

where $\Re^m$ and $\Re^n$ are sensor and motor spaces, respectively, and $p$ is the number of concepts. Abstract concepts can be used by a symbol manipulator for complex cognitive tasks like language learning. Currently, we do not use symbols for this purpose, but the role of symbol manipulator is shown in Figure 2 for extended works.

In one hand, the behavioral structure must be able to reconstruct the ideal behavior $gof(x) = h(x)$, i.e., maximizing the expected reward. On the other hand, it must minimize the number of concepts to keep them as general as possible (extremely, each sample can be assigned to a single concept). These two constraints act in opposite directions, because if $p$ is minimized too much, the structure becomes too restricted to reconstruct the ideal behavior, and if high expected reward ($E(R)$) is desired, $p$ must

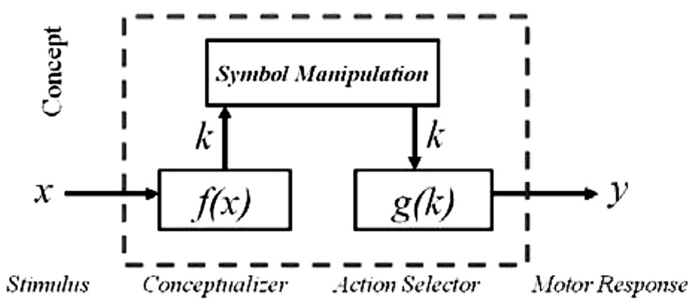


**FIGURE 2** The general model of a concept-oriented agent.

remain big enough. These constraints and the discrete nature of concepts make this problem a non-linear multi-objective optimization task

$$ObjectiveFunctions\begin{cases} Maximize\ (E(R)) \\ Minimize\ (p) \end{cases} \tag{2}$$

Concepts are linked to continuous regions of input and output spaces. We represent each region by a prototype; the nearest prototype in the perceptual space catches a given stimulus. Then the prototype is translated into a concept. Similarly, concepts are mapped to motor prototypes. From the prototype perspective, functions $f$ and $g$ can be decomposed as shown in (3).

$$\begin{aligned}
&f_1 : \Re^m \to \mathbf{N}_q;\ f_1 = \underset{i=1,\dots,q}{ArgMin}\|x - x_i\| \\
&f_2 : \mathbf{N}_q \to \mathbf{N}_p \\
&f = f_2(f_1(x)) = f_2\, o\, f_1(x) \\
&g_1 : \mathbf{N}_p \to \mathbf{N}_r;\ r \le p \\
&g_2 : \mathbf{N}_r \to \{y_1, y_2, \dots, y_r\} \subset \Re^n \\
&g(k) = g_2(g_1(k)) = g_2\, o\, g_1(k);\ k \in \{1, 2, \dots, p\},
\end{aligned} \tag{3}$$

where $f_1$ maps stimulus to a perceptual prototype and $f_2$ maps the prototype to corresponding concept. Note that due to the abstract nature of concepts, $f_2$ is possibly many-to-one. Similarly, $g_1$ maps each concept to a corresponding action (motor prototype) and $g_2$ realizes the action by physical motor parameters in motor space. Obviously, $q$ and $r$ are the number of sensor and motor prototypes and $p$ is the number of concepts. Finding the prototype vectors and concept/prototype mappings are up to the learning algorithm. Note that $p \le q$ and $r \le p$ are necessary conditions for $f$ and $g$ to be functions.

Nevertheless, due to the discrete nature of concept and prototype sets, theoretically any many-to-many mapping of them can be converted to many-to-one (function) by extending the set of indices to its power set. Hence, we will let the function constraint remain for the sake of simplicity.

So far four functions have been introduced for behavior construction, namely, $f_1$, $f_2$, $g_1$, $g_2$. Here we claim that for any arbitrary structure $h$ obtained by combining these functions, there always exists an equivalent structure $h'$ whose $g_1{}'$ is one-to-one. If $g_1$ is not one-to-one, then there exists at least two values with the same map.

$$g_1(n_1) = g_1(n_2) = k; \quad n_1 \ne n_2 \tag{4}$$

However, $n_1$ and $n_2$ are themselves obtained from $f_2(\cdot)$

$$n_1 = f_2(m_1); \quad n_2 = f_2(m_2) \tag{5}$$

Now we define our $f_2{}'$ and $g_1{}'$ functions as

$$\begin{aligned} g_1'(n_1) &= k \\ f_2'(m_1) &= f_2'(m_2) = n_1 \end{aligned} \tag{6}$$

It is easy to check that

$$g_1(f_2(m)) = g_1'(f_2'(m));\ \ m \in \{m_1, m_2\} \tag{7}$$

Now let's consider $g_2{}' = g_2$ and $f_1{}' = f_1$; so while $h' = h$, $g_1{}'$ is one-to-one. Although $h$ and $h'$ are behaviorally the same, they are different in the number of concepts. As $g'_1$ is one-to-one so $p' = r'$ and for non one-to-one cases like $g_1$ $p > r$. But $r = r'$ because $g'_1 = g_2$. Combining these results gives $p > p'$. Since the one-to-one instance takes fewer number of concepts, it is preferred according to (2).

In fact, there could be a dual case, where $f_2$ was one-to-one and $g_1{}'$ was many-to-one. However, relational concepts require different regions to be mapped to the same concept. Therefore, for efficiently representing the relational concepts (without power sets), this case is avoided.

Returning to the one-to-one $g_1$, which can be now denoted by $g_1 : \mathbf{N}_p \to \mathbf{N}_p$, its task is now limited to a simple permutation. This is immaterial and can be eliminated. The model is therefore simplified to

$$\begin{aligned} &h = g_2\ o\ f_2\ o\ f_1 \\ &f_1 : \Re^m \to \mathbf{N}_q \\ &f_2 : \mathbf{N}_q \to \mathbf{N}_p \\ &g_2 : \mathbf{N}_p \to \{y_1, y_2, \ldots, y_p\} \end{aligned} \tag{8}$$

## The Model

Previously, we reviewed evidence of conceptual representation in mirror neurons and then we discussed that concepts are stored in semantic memory. So there may be a relation between these two areas of brain. Moreover, we could not find any biological evidence that contradicts with this hypothesis. For instance, Keysers et al. (2003), who work on auditory mirror neurons, emphasize that there is no evidence for existence of direct connection between the F5 area and the auditory cortex, according to their correspondence with Matelli. They state that auditory information may reach F5 neurons along complex cortico-cortical routes (Romanski et al. 1999) or even involve corticosubcortical loops (Fries 1984). So it is not exactly clear that what these connections are, but we guess that relation of semantic memory to mirror neurons can be a possibility.

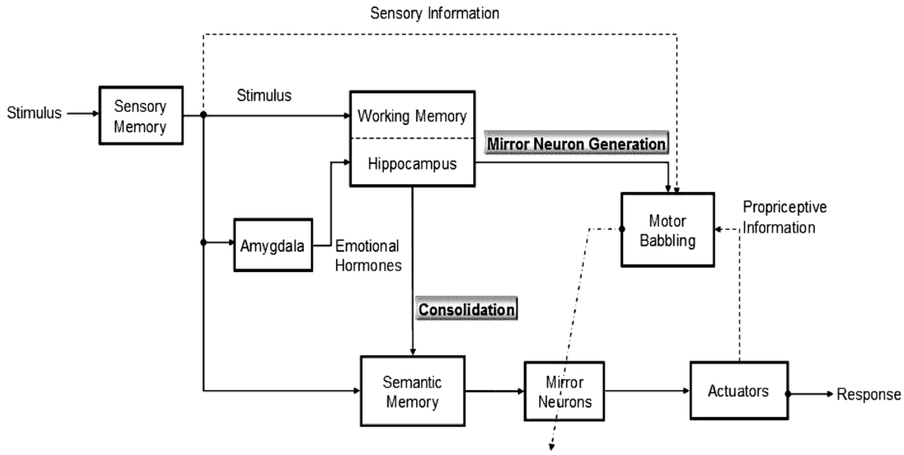


**FIGURE 3** Proposed model (interactive and non-interactive phases are indicated by solid and dashed arrows, respectively).

We put together all the ideas obtained so far to develop a new model for conceptual imitative learning and schematically show them in Figure 3.

This model is composed of two functional modes, namely, interactive and non-interactive. During the interactive mode, the agent interacts with the teacher in order to learn stimulus-concept relations. The behavior of the agent is modified by reinforcement signals that the teacher issues. Initially the agent transfers sensed stimulus to working memory. In parallel, emotion is extracted from stimulus to direct rehearsal and consolidation. Consolidated concepts are kept in semantic memory and sent to the mirror neuron system. The latter system is composed of a repertoire of actions and their relations to concepts. Ultimately, the mirror neuron system sends the matched action to the motor system.

In the non-interactive mode, the agent detaches its contact with the teacher and merely attempts to find how to imitate the observed action. This phase is triggered when none of the actions in the agent's repertoire matches the observed action. So it learns how to produce that action itself by motor babbling. Now we will review each component of our model separately. Interaction of these components will be described in the next subsections in the context of learning algorithm.

***Sensory Memory***

As the rate of stimulus generation may be higher than the time that the agent requires to sample, process, and respond to the stimulus, it is required that the perceived stimulus be captured and kept in a register until the end of the processing cycle. So, sensory memory keeps a duplication of the perceived stimulus.

***Amygdala***

This unit extracts emotional cues from stimulus and influences the hippocampus using-stress hormones. This hormone is modeled by a scalar value where its sign indicates destructiveness or constructiveness of the emotion and its magnitude indicates the strength of the hormone. So it acts like a reinforcement signal. Emotional cues can be extracted from stimulus using simple features like facial expression for visual stimuli and intonation and loudness of the voice for auditory stimuli.

***Working Memory***

Similar to its biological counterpart, the contents of this memory are represented perceptually using exemplars. In our model, working memory temporarily stores incoming stimulus to be later transferred to semantic memory, if needed. While a stimulus is in working memory, it can be compared with available concepts in semantic memory to be possibly associated with one of them (elaborative rehearsal).

***Semantic Memory***

The representation in semantic memory is based on organized information. Since our model relies on relational concepts, semantic memory in our model consists of a number of prototypes in perceptual space that are grouped to form different concepts. In other words, each concept is a set of prototypes in semantic memory that have the same meaning with respect to a criterion that the teacher has in mind.

The motivation of using prototypes instead of exemplars was the long duration of information in long-term memory. While information in working memory is continually transferred to semantic memory, the information in semantic memory is not moved to anywhere else. On the other hand, since in a real environment the flow of new stimuli may be nonstop, semantic memory can be exploded by an unbounded number of exemplars. However, only a few prototypes are sufficient to represent a large number of (even infinite) stimuli.

***Hippocampus***

This component looks like a control mechanism for working memory. Similar to what *denate gyrus* does in the brain for the hippocampus, it can recruit new units to be used by working memory. This provides the agent with the ability of incremental learning. It is particularly valuable for interactive learning applications where the number of concepts is not known a priori. Another task of the hippocampus is controlling the consolidation process. This happens when the stimulus is likely to belong to an exemplar of working memory, not a prototype in semantic memory.

***Mirror Neurons***

According to our hypothesis, mirror neurons get their conceptual input from semantic memory. This highly simplifies the work of mirror neuron system. According to the one-to-one evidence discussed previously, mirror neurons merely form a repertoire of known actions. By putting these actions into a sequence, a variety of complex actions can be constructed. Mirror neurons learn with the help of the motor babbling module to realize actions by physical motor responses.

***Babbling***

Babbling is a fundamental and basic mechanism in imitation (Meltzoff and Moore 1997). It means issuing random motor commands and inspecting their effects in order to discover self's motor-percept model, also called forward model. When babbling becomes goal-directed, it provides a basic way to imitative learning. Although goal-directed babbling starts from the issuance of random motor commands, the learning agent refines these commands gradually as the learning proceeds to achieve a target effect. In our model, babbling extends the repertoire of familiar actions to be used by the mirror neuron system.

Besides the mentioned components, there are also two processes in our model that are explained in the following.

***Consolidation***

As discussed previously, elaborative rehearsal is a necessary condition for consolidation. Another condition for consolidation is that the environment and the teacher provide consolidation opportunity for the agent through emotional factors and repeating chance. In fact, each time a stimulus similar to an exemplar of working memory is experienced, the hippocampus relates it to one of the concepts. The agent performs the corresponding action and receives reinforcement. The reinforcement may support or destroy that relation. If all relations are destroyed, it means no concept could represent the exemplar, so a new concept is created for that exemplar. The set of the new concept is initialized by one prototype whose position in perceptual space is the same as the mentioned exemplar.

***Mirror Neuron Generation***

Since mirror neurons establish a connection between concepts and actions, once a new concept is formed in semantic memory, its corresponding action must be determined for the mirror neuron system. At this time, the hippocampus, which is able to interpret emotions in a more abstract fashion than the amygdala, switches to the non-interactive mode. In this mode, babbling tries to learn how to act in order to reconstruct the observation, which belongs to the new concept. Once babbling completes, the system switches back to the interactive mode.

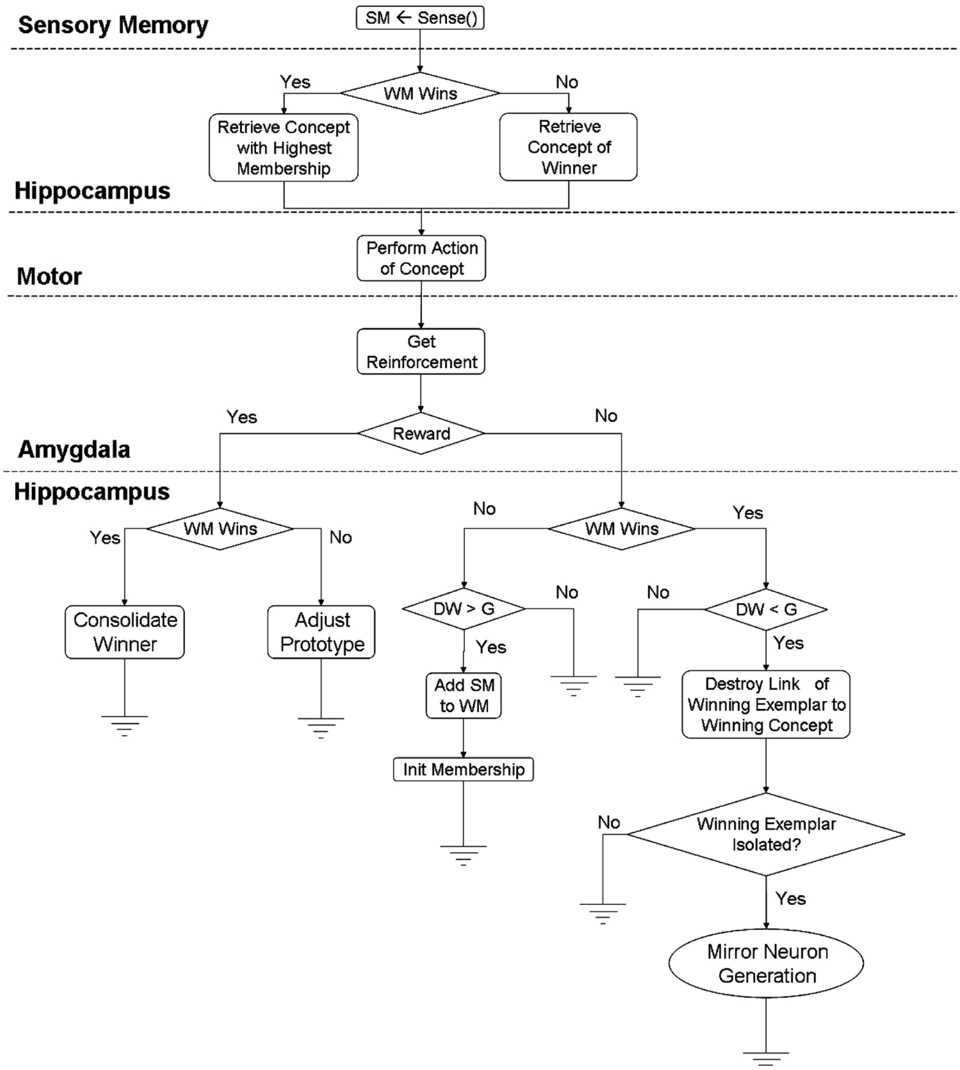


**FIGURE 4** Flowchart of the proposed method.

## Learning Algorithm

The learning algorithm is an iterative procedure. Each learning cycle begins when a new stimulus arrives. The algorithm is illustrated in Figure 4 and its pseudo-code is shown in Figure 5. In the pseudo-code a number of variables can be seen which are going to be introduced here. Before going into the details of the pseudo-code, first we introduce its variables.

$\mathbf{x} := \text{Sense}()$

$i := \underset{\mathbf{x}_k \in W}{ArgMin} \left\| \mathbf{x} - \mathbf{x}_k \right\|$

$j := \underset{\mathbf{x}_k \in L}{ArgMin} \left\| \mathbf{x} - \mathbf{x}_k \right\|$

$\textit{if } \left\| \mathbf{x} - \mathbf{x}_i \right\| < \left\| \mathbf{x} - \mathbf{x}_j \right\|$

$\qquad k := \underset{p \in C}{ArgMax}\ \mathbf{W}_{ip},\ \mathbf{y} := \mathbf{y}_k$

$\quad \textit{else} \quad k := c_j,\ \mathbf{y} := \mathbf{y}_k$

$\text{Perform}(\mathbf{y})$

$R := \text{Get\_Reinforcement}()$

$\textit{if } (R > 0, \left\| \mathbf{x} - \mathbf{x}_i \right\| < \left\| \mathbf{x} - \mathbf{x}_j \right\|)$

$\qquad p := \text{New\_L}(),$

$\qquad \mathbf{x}_p := \mathbf{x}_i,\ c_p := k,$

$\qquad \text{Delete\_W}(i)$

$\textit{if } (R > 0, \left\| \mathbf{x} - \mathbf{x}_i \right\| \geq \left\| \mathbf{x} - \mathbf{x}_j \right\|)$

$\qquad \Delta \mathbf{x}_j := \eta (\mathbf{x} - \mathbf{x}_j)$

$\textit{if } (R < 0, \left\| \mathbf{x} - \mathbf{x}_i \right\| \geq \left\| \mathbf{x} - \mathbf{x}_j \right\|, \left\| \mathbf{x} - \mathbf{x}_i \right\| > G)$

$\qquad p := \text{New\_W}(),\ \boldsymbol{\mu}_p := \mathbf{y},\ \mathbf{x}_p := \mathbf{x},$

$\qquad \underset{q \in C}{\mathbf{W}_{pq}} := \exp(-\left\| \mathbf{y}_q - \boldsymbol{\mu}_p \right\|^2)$

$\textit{if } (R < 0, \left\| \mathbf{x} - \mathbf{x}_i \right\| < \left\| \mathbf{x} - \mathbf{x}_j \right\|, \left\| \mathbf{x} - \mathbf{x}_i \right\| < G)$

$\qquad \mathbf{W}_{ik} := 0$

$\qquad \textit{if } (\forall \mathbf{y}_k \in M, \mathbf{W}_{ik} = 0)$

$\qquad\qquad \text{Find } \mathbf{y}^* \text{ such that } \text{Perform}(\mathbf{y}^*) = \mathbf{x}_i$

$\qquad\qquad q := \text{New\_C}()$

$\qquad\qquad \mathbf{y}_q := \mathbf{y}^*$

$\qquad\qquad p := \text{New\_L}()$

$\qquad\qquad \mathbf{x}_p := \mathbf{x},\ c_p := q$

$\qquad\qquad \text{Delete\_W}(i)$

$\qquad\qquad \forall \mathbf{x}_p \in W : \mathbf{W}_{pq} = \exp(-\left\| \mathbf{y}_q - \boldsymbol{\mu}_p \right\|^2)$

**FIGURE 5** Pseudo-code for the proposed algorithm.

- $W = \bigcup \mathbf{x}_i; \mathbf{x}_i \in \Re^m$: This is a set of exemplars in the working memory of the agent that it has experienced. Each exemplar is in fact a vector in perceptual space.
- Long-term memory maintains consolidated prototypes and their semantic interconnections. These two components are denoted by sets $L$ and $C$,

respectively. Elements of $L$ are vectors in perceptual space and elements of $C$ are simple functions each of which maps a prototype index to a concept index. For instance, $c_p$ returns the number of the concept associated with prototype $p$.

$$L = \bigcup \mathbf{x}_i; \ \ \mathbf{x}_i \in \Re^m$$
$$c : \mathbf{N}_{|L|} \rightarrow \mathbf{N}_{|C|}$$

- $\mathbf{x} \in \Re^m$ is a variable which simulates sensory memory and keeps the perceived stimulus until the end of the cycle.
- $\mathbf{y} : \mathbf{N}_{|C|} \rightarrow \Re^n$ is a function that maps a concept to its motor response. For instance, $\mathrm{y_k}$ is a vector in motor space which returns motor response to concept $k$.
- **W** stands for the matrix which determines the internal state of the hippocampus in terms of rehearsal and consolidation. Rows of this matrix correspond to the exemplar indices of working memory ($W$) and its columns correspond to concept indices. The elements of **W** describe the membership likelihood of an exemplar in working memory to a concept in semantic memory.
- $\mu : \mathbf{N}_{|W|} \rightarrow \Re^n$ is a function which maps the index of a given prototype to the motor response of the concept with likely correspondence. This mapping is created once a new entry in working memory is created. This function is later used for computation of membership likelihood in **W** in Gaussian-like fashion as $\exp(-\|y_q - \mu_p\|^2)$.
- $G$: This is a data-dependent constant that determines the degree of generalization. Small values for $G$ are suitable for compact prototypes, while its large values are appropriate for bulky prototypes. Currently, $G$ must be set manually in a trial and error fashion.

To ease explaining the algorithm, assume that it is being executed for some cycles and both memories have some contents. Moreover, we refer to each line in the pseudo-code using # sign.

In #1, the stimulus is stored in SM (sensory memory). Both working memory (WM) and long-term memory (LTM) have access to SM and compete to capture the stimulus. In each memory, the most similar unit (#2 for exemplars in WM and #3 for prototypes in LTM) is chosen and the winning memory is the one that has the most similar unit with the stimuli. We used Euclidian distance in perceptual space as dissimilarity measure.

Next the concept related to the stimulus and the corresponding action is retrieved in #4, #5, and #6. If the winning unit belongs to LTM (#6), there exists a solid relation between the winner and its corresponding concept, so the concept can be recalled definitely. However, when the winner belongs to WM (#5), there is no solid connection between the memory units and the concepts. Rather, each unit is connected to all concepts with different membership degrees. In this case, the concept with the highest membership is chosen as the matching concept of the stimulus. Formally, there is a membership matrix **W** with WM units as its rows and concepts as its columns and each element indicates membership of unit in concept.

In either case, once the concept is retrieved, mirror neurons convert it to its related action and motor response (#5 and #6). The agent performs the action (#7) and logically the teacher reacts (#8). Taking advantage of the teacher's reinforcement, the agent examines whether the action performed by the teacher and the one performed by itself belong to the same or different concepts. This is where same/different judgment gets involved. Depending on type of the reinforcement and the winning memory, four cases happen:

- *Reward and WM*: In this case (#9), the correct relation between WM unit and corresponding concept is found and therefore consolidation constraints are satisfied. Hence, this unit is removed from WM (#12) and is added to LTM (#10) and connected to the concept that caused the reward (#11).
- *Reward and LTM*: This situation (#13) means that the LTM unit has correctly attracted the stimulus. In this stage, the agent adjusts the winning prototype a bit toward the stimulus so that the prototype can represent any instance of it (#14).
- *Punishment and LTM*: In order to enter this branch, another constraint is imposed which requires that the distance of the winning unit to the stimulus (denoted by *DW*) is below a threshold *G*. This case is formulated entirely in #15. *G* is a data dependent constant; while compact clusters are best represented by small values of G, wide clusters need larger values of it. Recall that, this situation occurs when the LTM unit has wrongly attracted the stimulus. Since the contents of LTM are all correct (otherwise they would not get reward to enter to LTM), there is no need to change anything there. Instead, the stimulus may belong to another old concept or to a new one, but there is no prototype to attract and relate the stimulus to the correct concept. In both cases, a new unit is created in WM to represent the stimulus (#16) and determine its concept in further rehearsals. Moreover, since a new prototype is created, its membership to all concepts must be computed in #17 by adding a new row to membership matrix **W** and initializing the columns of the row. Memberships are estimated using distance in motor space.
- *Punishment and WM*: If during rehearsal, WM makes a wrong decision and the distance of stimulus from the winning WM unit falls below threshold *G* (#19), the negative reinforcement weakens consolidation in that

channel (unit to concept) in #20. Therefore, next time that this unit attracts a stimulus, another concept will be rehearsed. If all channels are destroyed (#21), it indicates that the unit belongs to a new concept. In this case, first a new action is learned through babbling such that it can result in a similar effect (#22). Next a new concept is created in LTM (#23) and the winner is transferred from WM to LTM (#25, #26, and #27). In other words, from one side the new concept is related to the transferred prototype (#26) and, from the other side, the concept is connected to the action resulted by babbling (#24). Eventually since a new concept is created, a new column is added to membership matrix **W** and the rows (prototypes) of that column (new concept) are initialized using motor estimation in motor space (#28).

## EXPERIMENTAL RESULTS

To evaluate our method, we carried out two experiments in phoneme acquisition domain. We chose phoneme acquisition as our testbed for a number of reasons. First, our model learns by babbling and babbling is best known in phoneme acquisition by children. Second, our results indicate that vowels are relational concepts because of forming disjoint clusters in perceptual space. Third, the number of available vowels and the clusters they form are very limited and therefore reinforcement learning is applicable. Fourth, using loudness as an easy-to-extract emotional cue, the human user can interact with the agent in a natural and simple way.

In order to use a compact representation (not necessarily optimal) such that the acoustic signal can be described by an appropriate feature vector in perceptual space, a dimensionality reduction must be performed. However, extracting appropriate features from acoustic signal is itself a difficult problem in speech processing and recognition domain. Since machine learning is the main concern of this paper, we tried to avoid involvement in complicated speech processing problems. Therefore, we focused on the vowel learning problem because formants[3] are easy to compute and effective features for vowel representation. We also used an articulatory model of speech synthesis. An articulatory model is comprised of a set of simulated muscles whose contraction is controllable. By adjusting muscle parameters, one can produce various voices. We implemented our algorithm in C language under a Linux operating system.

### Experiment I

In this experiment, we used the Peterson and Barney data set, which is a well-known benchmark in vowel learning (Peterson and Barney 1952). This database contains formants of 10 American English language vowels spoken by 76 people (33 men, 28 women, and 15 children). In this experiment, each vowel is uttered twice, so there are totally 1520 samples in this data set. High overlap among samples of this data set has made it a difficult learning problem. In this data set there are even 21 pairs whose first and second formants are the same. Since such ambiguous data are not distinguishable, we removed one instance of each pair. So the number of samples in the data set was reduced to 1499.

Since the proposed algorithm requires mutual interaction between the imitator and the teacher, in this experiment, we defined virtual teachers from the data set. In fact, in each stage, one teacher (speaker) is chosen randomly and utters one of the 10 vowels arbitrarily, i.e., the agent gains access to its first two formants. Next, the imitator tries to learn how to utter something similar to the heard sound using babbling. Then, the teacher—the software in this experiment—finds the closest sample to the agent's uttered sound in the whole dataset and compares its label with the label of the original vowel uttered by the teacher. If these labels are the same, the teacher issues a positive reinforcement; otherwise, a negative reinforcement is given to the agent.

In this experiment, instead of using a realistic articulatory model and extracting formants from real acoustic signal, we used a simple affine transform to simulate forward model. The synthesis system is composed of two muscle parameters, namely, $m_1$ and $m_2$. These parameters are then converted to self formants (*F1,F2*) using equation (9). Coefficients $\theta_1$ to $\theta_6$ are constant and they define the physical characteristics of the forward model. Care was taken to avoid choosing singular coefficients. Babbling process tries to minimize the difference between self and the heard formants. Since an analytical formulation of the forward model is available, babbling can be realized using gradient decent optimization.

$$\begin{aligned} F_1 &= \theta_1 m_1 + \theta_2 m_2 + \theta_3 \\ F_2 &= \theta_4 m_1 + \theta_5 m_2 + \theta_6 \end{aligned} \tag{9}$$

Returning to the learning algorithm, the learning rate $\eta$ for adapting prototypes was set to 0.1 in the experiment. This experiment was repeated twice, once with $G = 100$ and another time with $G = 10$. Results indicate that the agent could accurately develop 10 concepts, each of which corresponds to one of the vowels. Figure 6 shows the average reinforcement computed from execution of the program 10 times. Due to the discrete nature of reinforcements in our model $(1, -1)$, the plot was smoothed so that the underlying behavior becomes clear. The number of iterations required to achieve an acceptable level of reinforcement was 6000 steps. This may seem large for an algorithm that is supposed to be used in interactive applications.

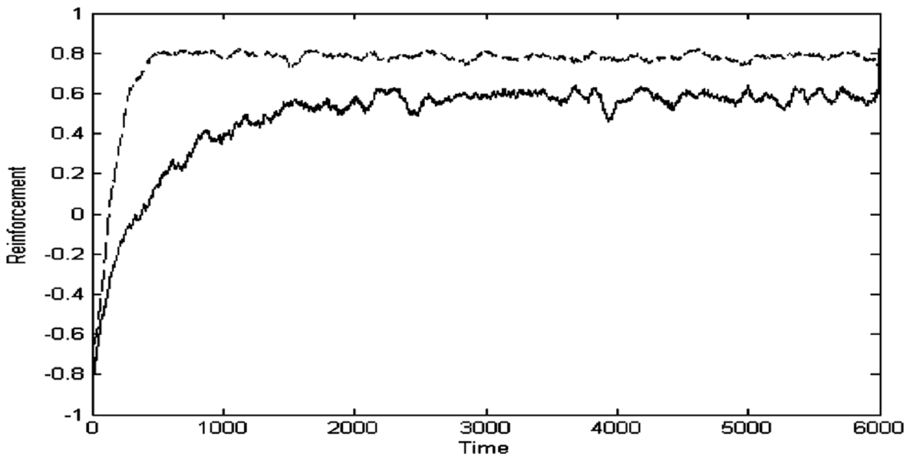


**FIGURE 6** Reinforcement over time. Solid and dashed plots correspond to G = 100 and G = 10, respectively.

The main reason for this large number of steps is the large number of speakers and vowels. In fact, if we divide 6000 by the number of speakers (76) and vowels (10), we will see that the average number of steps required to learn each vowel uttered by each teacher is 7, which is not so high.

Average and standard deviation of the number of remained units in WM and LTM and the number of concepts as well as the accuracy of the agent's performance are shown in Table 1.

There are two important issues in this table. First, the agent has learned exactly 10 concepts. Second, the number of prototypes in LTM—denoted by μLTM—is much larger than the number of concepts (10) and it indicates that vowels could not be represented by perceptual categorization, but relational concepts. In fact what the algorithm learns is to find these clusters as well as the semantic relations among them. Therefore the agent has developed the same abstract concepts that the teacher had in mind.

Another interesting piece of data in the table is that the number of prototypes in LTM is very close to the number of samples in the training set (1499). Maybe this is due to the relatively small number of utterances per vowel for each teacher (2), comparing with the number of speakers (76). Whatever the reason is, it is obvious that the appropriate representation for this data set is exemplar method (low values of *G*). This fact is

**TABLE 1** Statistics Obtained After 10 Executions in Experiment I

| G | μA% | σA% | μWM | σWM | μLTM | σLTM | μC | σC |
|---|---|---|---|---|---|---|---|---|
| 10 | 92.15 | 4.86 | 72.6 | 16.61 | 1246.1 | 110.57 | 10 | 0 |
| 100 | 79.25 | 4.91 | 27.7 | 3.56 | 949.8 | 58.48 | 10 | 0 |

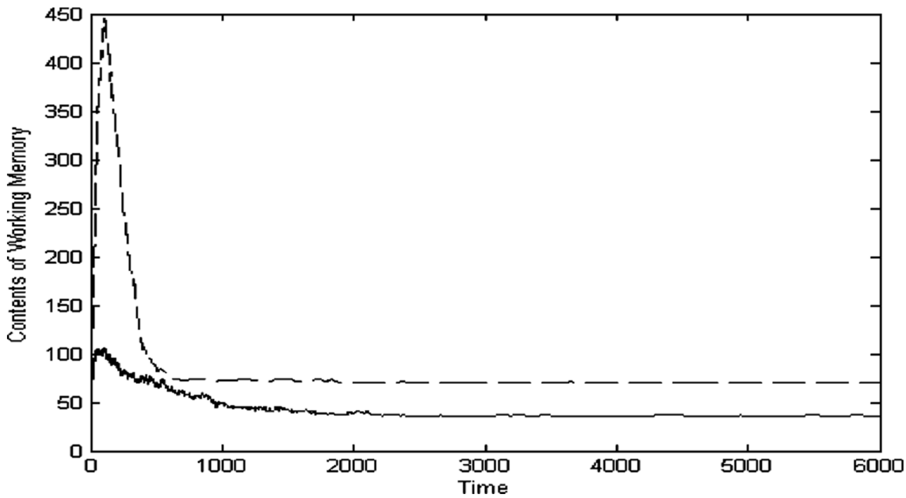


**FIGURE 7** Load of working memory. Solid and dashed plots correspond to G = 100 and G = 10, respectively.

also reflected in two plots of Figure 6. The accuracy of the results is denoted by *A* in Table 1. It is computed when the algorithm has passed all 6000 steps. At that time the whole data set is presented to the agent and its correct classifications are counted. Then accuracy is defined as the percentage of correct classifications to the whole number of samples.

Another important issue in this experiment is the load of WM, as shown in Figure 7. As it can be seen, initially the load of the memory is very high. The reason for this phenomenon could be the behavior of simulated teachers. They choose vowels randomly with a uniform distribution and without attention to the imitator's progress. So before the agent finds the opportunity to complete consolidation of a vowel, another one arrives. That increases the load of WM, however, as the learning proceeds the load decreases because the agent finds sufficient chance for consolidation.

## Experiment II

This experiment was conducted in the real world, where human subjects naturally interact with the agent. The agent receives acoustic input from a microphone and responds using speakers connected to the computer; see Figure 8. Emotional expressions of the human caretaker guide the process of concept learning. Since the focus of this experiment was on machine learning and not speech processing, we utilized a speech analysis/synthesis tool named PRAAT (Boersmo and Weenink 1992). This program could interact with our program in on online manner using interprocess communication techniques.

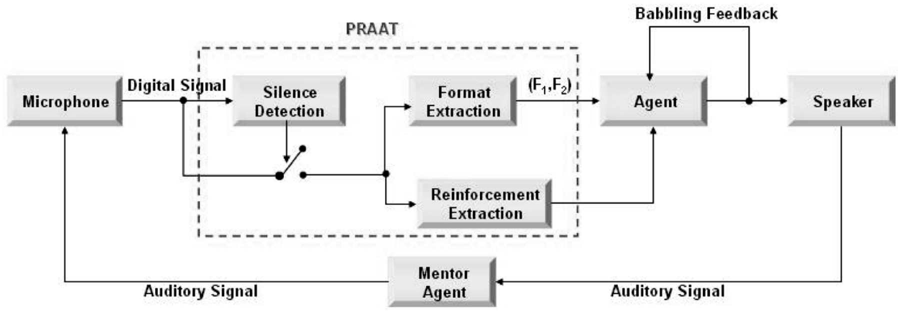


**FIGURE 8** The setup for the second experiment.

Initially, PRAAT segments vowels by silence detection. Thanks to the relatively high power in vowel spectrum, it can be easily segmented by choosing a simple threshold for the power. Next PRAAT provides the agent with the first two formants by extracting them from the segmented vowels. In parallel, reinforcement signal is also extracted from the caretaker's voice using a very simple cue. If the speaker's voice is abnormally loud (to cause fear), the agent interprets it as a negative reinforcement. A similar punishment is used when parents want to train children who are not yet able to understand language. So this protocol is natural enough to humans to train our agent. Continuing interaction with the agent in a normal manner is also a sign of the caretaker's interest and is considered as positive reinforcement. This relationship and the utilized processes in PRAAT are illustrated in Figure 8.

PRAAT is not only an important component in the sensory circuitry of our agent, but also in its motor system. It has a complex articulatory model for speech synthesis composed of 29 muscles. The amount of contraction of each muscle over time can be controlled independently. Obviously, goal-directed motor babbling starts from a random contractions and gradually refines these values to produce the desired sound. However, unlike Experiment I, the analytical model of the articulatory system and consequent gradient is not available here. Therefore, each time, a muscle is chosen randomly and perturbed. If the resulted voice is closer to the goal, it is kept as the best solution found so far; otherwise, it is ignored.

Iterating this scheme over and over yields continuing improvement in the babbling. Babbling is stopped when learning cost drops below a threshold value. The learning cost is a weighted sum of formant error like the one used in Experiment I, plus the inverse of the signal's energy. The latter term forces the agent to escape from silence. Since our experiment was merely concentrated on vowels, only a subset of muscles[4] was chosen for learning. Moreover, a few muscles,[5] which were invariant during synthesis of any vowel, were manually set to appropriate constant values. The rest of the muscles remained inactive.

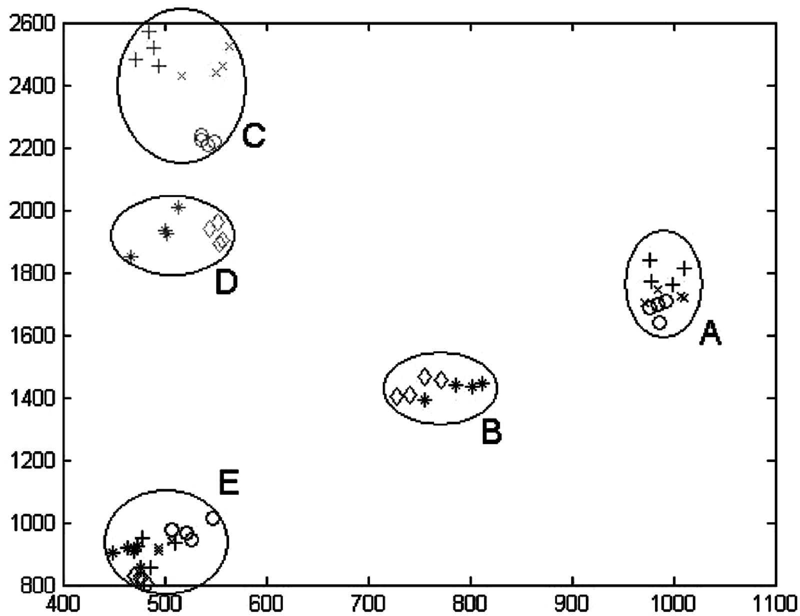


**FIGURE 9** Distribution of three Persian vowels in (F1, F2) plane. Male subjects are represented by diamond and star and female subjects by plus, cross, and circle.

Five human subjects (3 females, 2 males) participated in this experiment. They were told to only use three Persian vowels throughout their interaction with the agent. The distribution of some of these vowels in (F1,F2) plane is visualized in Figure 9. In this figure, each teacher is depicted by a different spot shape. The algorithm reached five prototypes for representing the experimental data set. Prototypes A and B belong to the first vowel, C and D to the second vowel, and E to the last vowel. Obviously, some concepts (vowels) require more than one prototype for proper representation. This means that phoneme acquisition in formant space forms relational concepts. In our study, the speaker's gender is the reason that disjoins prototypes of the same vowel. For instance A and C belong to female samples, while B and D belong to male samples. However, the distribution of the third vowel is relatively compact regardless of the speakers' gender.

The goal of the learning agent is not only how to utter the heard vowels, but also to understand when they belong to the same concept. For instance, it must learn that although samples in regions A and B trigger different prototypes, they go to the same concept.

In this experiment parameter $\eta$ was set to 0.1 and $G$ was set to 150. Once the gained average reward becomes fairly stable, the learning is considered to be in a satisfactory level. It took near to 100 iterations for

**TABLE 2** Information Obtained in the Steady State of Experiment 2

| G | A% | WM | LTM | C |
|---|---|---|---|---|
| 150 | 100 | 0 | 5 | 3 |

the agent to reach such steady state. The agent's brain was further explored to study more details. It is observed that exactly three concepts are formed. In addition, similar to what is shown in Figure 9, five prototypes were created to represent the three concepts and each concept was accurately associated to one of the vowels. In early stages of learning, the amount of contents in the working memory was high, which was gradually transferred to semantic memory through consolidation. In steady state, there was no item left in working memory. The information related to this experiment is shown in Table 2. Since real subjects were involved in this experiment and it was not convenient to repeat the experiment several times, the values in this table are one-time measurements.

The output of the program is compared against natural vowels uttered by one of the female speakers, as shown in Figure 10. The figure is, in fact, a spectrogram where horizontal axis corresponds to time and vertical axis corresponds to frequency. The darker a point is, the higher contribution that frequency has at that time. In PRAAT we had to draw five first formants even though we needed the first two ones only. Formants are denoted by

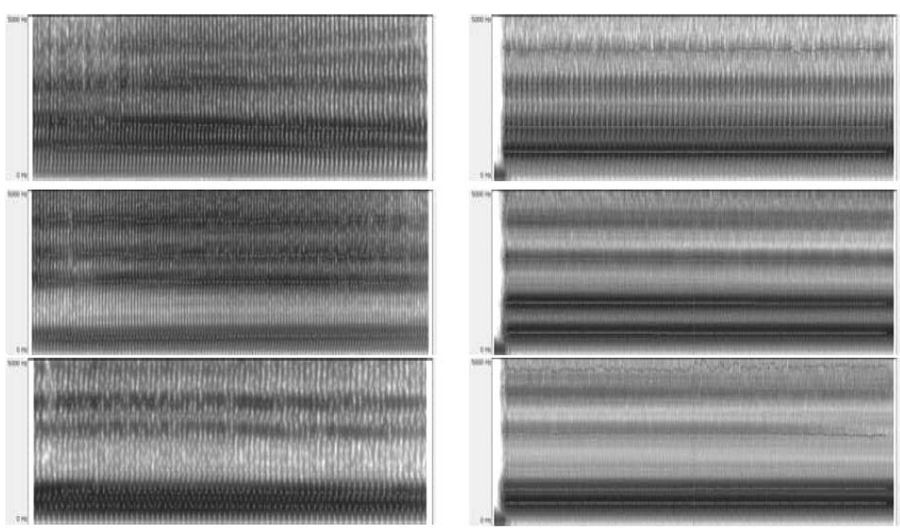

**FIGURE 10** Spectrogram for three Persian vowels. The darker a point is the more contribution its frequency has at that time. Left: a female teacher. Right: reconstructed signals by the agent.

dashed horizontal lines along the dark areas. Obviously the two bottom ones denote the first two formants, which the agent has learned to imitate. It is apparent that the synthesized signals could induct their underlying vowels adequately.

## CONCLUSION AND FUTURE WORKS

In this article, we proposed concept-oriented imitation as a new and important avenue in machine learning research for the AI and robotics community. Our model was inspired by findings in cognitive science and neuroscience domains, particularly, the representation method in the mirror neuron system, as well as the role of memory in concept learning. Moreover, by mathematical analysis of a simple but general behavioral structure for a concept-oriented agent, we related the use of functions for constructing such a model with the task of the mirror neuron system in a biological system.

Relational concepts, which are the simplest form of abstract concepts according to Zentall's categorization, were chosen as the basis of this work. These concepts are, in fact, disjoint clusters in an agent's perceptual space, which belong to the same concept. The clusters are unified by their semantic and functional properties. The sameness of clusters is learned by reinforcement learning using same/different technique. Moreover, inspired by the role of memory in forming abstract concepts and the process of memory consolidation, the model incorporated some similar characteristics for concept learning. The experimental results on vowel imitative learning showed the effectiveness of the proposed approach.

There are few points that should be considered in future works. For instance, in the proposed model, all concepts are relational and therefore in the same abstraction level. If associative concepts are incorporated into the model, a hierarchy of concepts can be defined, i.e., associative concepts in terms of relational concepts or less abstract associative concepts. Biologically, this is similar to congruency notion in mirror neurons, to some extent. For example, in the phoneme-learning problem, phonemes could be learned in the lowest level and then more abstract concept entities like syllables and words could be constructed.

Automatic adjustment of parameter $G$ is another issue that can be addressed in the future. This is a data-dependent parameter whose manual tuning requires trial and error. Problems with compact clusters are best represented by small values of $G$, while large values of $G$ are appropriate for sparse clusters. A simple idea for the automatic finding of $G$ could be starting from a small $G$ and gradually increasing the value as long as no considerable drop in the average reward results.

Finally, this article was concerned mainly with the algorithmic and machine learning side of imitation. We conducted experiments in the vowel learning domain in order to evaluate the proposed method in a real problem. However, there is still a long way to assess the proposed method in a vast range of problems in the AI and robotics domain.

## REFERENCES

Arbib, M. A. 2000. The mirror system, imitation, and the evolution of language. In: *Imitation in Animals and Artifacts*, eds. Chrystopher Nehaniv and Kerstin Dautenhahn, pp. 229–280, Cambridge, MA: MIT Press.

Atkinson, R. C. and R. M. Shiffrin. 1968. Human memory: A proposed system and its control processes. In: *The Psychology of Learning and Motivation: Advances in Research and Theory*, ed. K. W. Spence, Vol. 2, 89–195. New York: Academic Press.

Breazeal, C. and B. Scassellati. 2000. Challenges in building robots that imitate people. In: *Imitation in Animals and Artifacts*, eds. Chrystopher Nehaniv and Kerstin Dautenhahn, pp. 363–390, Cambridge, MA: MIT Press.

Burke, A., F. Heuer, and D. Reisberg. 1992. Remembering emotional events. *Mem. Cognit.* 20:277–290.

Boersma, P. and D. Weenink. 1992. *Praat: A System for Doing Phonetics by Computer*, available from www.praat.org

Byrne, R. W. and A. Whiten. 1988. *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans.* Oxford: Clarendon Press.

Christianson, S. A. 1992. *Handbook of Emotion and Memory: Current Research and Theory.* Hillsdale, NJ: Erlbaum.

Cook, R. G., D. M. Kelly, and J. S. Katz. 2003. Successive two-item same-different discrimination and concept learning by pigeons. *Behavioral Processes*, 62:125–144.

Craik, F. and R. Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Thinking and Verbal Behavior* 11:671–684.

Davidsson, P. 1994. *Concepts and Autonomous Agents*, Licentiate thesis, Department of Computer Science, Lund University.

Dufay, B. and J. C. Latombe. 1984. An approach to automatic robot programming based on inductive learning. *Int. J. Robot. Res.* 3:3–20.

Fadiga, L., L. Fogassi, G. Pavesi, and G. Rizzolatti. 1995. Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology* 73:2608–2611.

Fagg, A. H. and M. A. Arbib. 1998. Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks* 11:1277–1303.

Fries, W. 1984. Cortical projections to the superior colliculus in the macaque monkey: A retrograde study using horseradish peroxidase. *J Comp Neurol* 230:55–76.

Gold, P. E. and J. L. McGaugh. 1975. A single-trace, two process view of memory storage processes. In: *Short-Term Memory*, eds. D. Deutsch, and J. A. Deutsch, 355–378. New York: Academic Press.

Harnad, S. 1990. The symbol grounding problem, *Physica, D* 12:335–346.

Jenkins, O. C. and M. J. Mataric. 2003. Automated derivation of behavior vocabularies for autonomous humanoid motion. In *Autonomous Agents and Multiagent Systems* (*AAMAS 2003*) 225–232.

Jenkins, O. C., M. Nicolescu, and M. J. Mataric. 2004. Autonomy and supervision for robot skills and tasks learned from demonstration. In: *Proceedings of the AAAI Workshop on Supervisory Control of Learning and Adaptive Systems*, pp. 27–29, California, USA: San Jose.

Keysers, C., E. Kohler, M. A. Umilta, L. Nanetti, L. Fogassi, and V. Gallese. 2003. Audiovisual mirror neurons and action recognition. *Exp BrainRes* 153:628–636.

Kruschke, J. K. 1992. ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99(1):22–44.

Kuniyoshi, Y., M. Inaba, and H. Inoue. 1994. Learning by watching: Extracting reusable task knowledge from visual observation of human performance. *IEEE Transactions on Robotics and Automation* 10(6):799–822.

LeDoux, J. 1996. *The Emotional Brain.* New York: Simon & Schuster.

Love, B. C. and D. L. Medin. 1998. SUSTAIN: A model of human category learning. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)*, pages 639–644, Madison, Wisconsin.

Lozano-Pérez, T. 1982. Task-planning. In *Robot Motion: Planning and Control*, eds. M. Brady, J. M. Hollerbach, T. L. Johnson, T. Lozano-Pe'rez, and M. T. Mason, 473–498. Cambridge, MA: MIT Press.

Manns, J. R., R. O. Hopkins, and L. R. Squire. 2003. Semantic memory and the human hippocampus. *Neuron* 38(1):127–133.

Meltzoff, A. N. and M. K. Moore. 1997. Explaining facial imitation: A theoretical model. *Early Development and Parenting* 6:179–192.

Merriman, J., C. Rovee-Collier, and A. Wilk. 1997. Exemplar spacing and infants' memory for category information. *Infant Behavior and Development* 20:219–232.

Nosofsky, R. M. 1984. Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10:104–114.

Oztop, E. and M. A. Arbib. 2002. Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics* 87(2):116–140.

Oztop, E., N. S. Bradley, and M. A. Arbib. 2004. Infant grasp learning: A computational model. *Exp. Brain Res.* 158:480–503.

Peterson, G. E. and H. L. Barney. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24:175–184.

Rizzolatti, G. and M. Gentilucci. 1988. Motor and visual-motor functions of the premotor cortex. In *Neurobiology of Neocortex*, eds. P. Rakic and W. Singer, 269–284. Chichester: Wiley.

Romanski, L. M., B. Tian, J. Fritz, M. Mishkin, P. S. Goldman-Rakic, and J. P. Rauschecker. 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci* 12:1131–1136.

Rumelhart, D. and J. McClelland. eds. 1986. *Parallel Distributed Processing: Explorations in the Micro Structure of Cognition.* Cambridge, MA: MIT Press.

Schaal, S. 1997. Learning from demonstration. *Advances in Neural Information Processing Systems* 9: 1040–1046.

Tomasello, M. 1990. Cultural transmission in the tool use and communicatory signaling of chimpanzees. In: *Language and Intelligence in Monkeys and Apes: Comparative Developmental Erspectives* 274–311.

Tung, C. P. and A. C. Kak. 1995. Automatic learning of assembly task using a DataGlove system. In *IEEE/RSJ Int. Conf. Intell. Robots Systems*, 1–8.

Zentall, T. R., M. Galizio, and T. S. Critchfield. 2002. Categorization, concept learning, and behavior analysis: An introduction. *Journal of The Experimental Analysis of Behavior* 3(78):237–248.

## ENDNOTES

1. Our approach is inspired by some findings in the domain of neurosciences; however, our model is not necessarily biologically plausible.
2. Note that the teacher could be an internal entity in some cases. By having an internal teacher we mean that the agent can evaluate itself provided that it has a sufficient level of intelligence or emotion.
3. Formants are peaks in frequency spectrum of the auditory signal.
4. Hyoglossus, styloglossus, geniogloussus, orbicularis oris, and masseter.
5. Lungs, levator palatini, interarytenoid, and cricothyroid.