Hierarchical Functional Concept Formation using Reinforcement Learning

Hamide Vosoughpour Majid Nili Ahmadabadi Maryam S. Mirian Babak Nadjar Araabi

Abstract

Reinforcement learning methods are slow and inefficient in large and continuous perceptual spaces. Discretizing such perceptual spaces generates a large number of states which slows down learning process. This motivates us that inspiring from concept learning in humans and cognitive science propose a framework for learning the concepts of the agent's functionality space in an unsupervised fashion. These concepts are extracted through mapping of perceptual space's samples to the agent's functionality space formed based on the agent's action-values. The proposed approach is realized in an online categorization problem and the results confirm that applying proposed learning method achieves a good generalization and high efficiency in terms of average reward and speed of convergence in the spaces with relational and associative concepts. In addition, this approach may help heterogeneous agents to share their abstract knowledge.

Keywords: *Concept Learning, Knowledge Abstraction, Concept Abstract Hierarchy, Reinforcement Learning*

1. Introduction

Reinforcement learning methods are among the best choices for interactive intelligent systems' learning because of their unsupervised nature. In these methods, agents learn action-values in each state by receiving a reinforcement signal from the environment. Another characteristic of Reinforcement Learning methods is the ability of online learning which is required for interacting with dynamic environments. It means that the agent can improve its learning while interacting with the environment and there is no need for a certain offline learning phase. These kind of methods are mainly inspired by the way human infants learn interacting with the environment.

But, there is one more important property in human learning which helps them to learn faster when dealing with a new environment. Researches show that, this property is the humans' ability to extract abstract concepts from the environment during the learning process [5]. These abstract concepts when formed in the mind can be used in unvisited environments to make a faster and more efficient learning. For example, a person who has learned riding a bicycle can learn faster how to drive a car than a person with no pre-knowledge about this, while obviously there may be no similarity in sensory spaces of these devices.

In fact, this is because of the common abstract concepts which exist in dealing with all types of vehicles; such as: breaking, accelerating and steering. These abstract concepts are not formed in the perceptual space. In fact, it is the similarity in functionality of different objects, situations or even stimuli which makes them members of the same concept in the human mind. In other words, although the perceptual space may be continuous and complex in real environments, often there are few functional concepts which help the agent to generalize those learned concepts to new stimuli. Fig. 1 shows the schematic view of stimuli prototypes formed in perceptual space and extracted functional concepts in agent's functionality space.

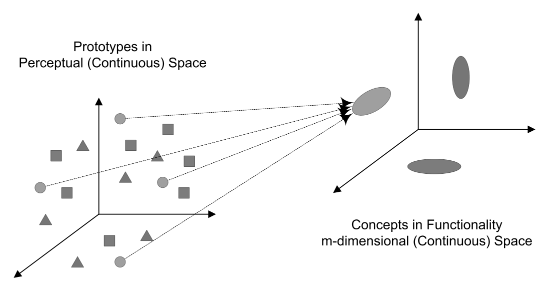

Figure 1. The schematic view of prototypes and concepts in functionality space.

In this paper, we propose a learning framework for unsupervised formation of abstract concepts and their hierarchy in the agents functionality space. By using extracted concepts and abstract concept hierarchy formed based on those concepts, and generalizing them to new observations, the agent can speed up the convergence of action-values and consequently accelerate learning in a continuous and complex perceptual space.

The paper is organized as follows: In section 2 the definitions and models of concepts from artificial intelligence and cognitive science perspectives are reviewed, then our point of view is mentioned. In section 3, our proposed approach for learning functional concepts, extracting abstract hierarchy, and using them in decision making is explained in details. Then, in section 4, some experiments designed for evaluation of methods as well as simulation results and some analysis are presented.

H. Vosoughpour and M. S. Mirian are with University of Tehran, Department of Electrical and Computer Engineering, Control and Intelligent Processing Center of Excellence, AI and Robotics Lab. (email: {vosoughpour, mmirian}@ut.ac.ir)

M. Nili Ahmadabadi and B. N. Araabi are with University of Tehran, Department of Electrical and Computer Engineering, Control and Intelligent Processing Center of Excellence, AI and Robotics Lab and School of Cognitive Sciences, IPM. (email: {mnili,araabi}@ut.ac.ir)

Finally, last section proposes the conclusion and future works of this research.

2. Concepts

There are different definitions and models for concept in the literature of artificial intelligence and cognitive science. Among them, we based our research on the definition which is presented by Zentall in [10], in which concepts are categorized into tree classes of abstraction:

Perceptual concepts: in this abstraction level, the similarity of observations in perceptual space is sufficient to map them to one concept

Relational concepts: in this abstraction level, some kinds of external knowledge is needed to form exact concepts in perceptual space, although the similarity of observations still contributes to map them to one concept.

Associative concepts: in this abstraction level, the similarity between observations in functionality space maps them to one concept and against the relational concepts; the similarity of observations in perceptual space has no contribution to map them to one concept.

According to Zentall's definition [10], the concept is the agents representation of its environment. So, the most important points of concept learning methods are how to form that internal representation, and how to use it in decision making.

While different perspectives of concept learning are mainly focused on learning in perceptual space [3, 4], Mobahi *et al.* [6] maps perceptual concepts into best agent's response to a stimulus. They express relational concepts in perceptual space by mirror neurons model [8, 7, 2] and the way of forming them. In this model, abstraction is done only in one level and this causes a bounded learning speed and generalization.

Abstraction and generalization are two important properties of intelligent systems improving their speed and quality of learning. Ability of generalization means being able to estimate the best response based on just a few experiments of a stimulus. Abstraction is to extract high level concepts with no details of observations. We can use abstract concepts, to estimate optimum agent's response to new or less observed stimuli.

In reinforcement learning methods, as the most usual online and interactive learning methods, representation of continuous space as a series of base kernels (such as RBF) is called generalization [9]. The studies show that either when online modification of kernels—which is a difficult process and critically affects convergence of learning—is not possible or if concepts in perceptual space and neighborhood of associative concepts is very scattered (and irregular), concept formation in perceptual space without any prejudgment about their shapes, is preferred to other methods [6, 1]. Therefore, we use clustering to descritize continuous space.

3. Learning method

In the learning method, as [6], prototypes are formed by clustering observations in perceptual space. By prototyping, the continuous perceptual space is mapped to some discrete prototypes. The radius of clusters is assumed to be fixed (G) for all clusters.

As mentioned before, against other related works for concept learning in continuous sensory space [6, 1] which define concepts as the best response to the stimulus, we define concepts in the entire functionality space (Section 3.1). The proposed method to learn and form these concepts are explained in this section.

3.1. Associative concepts in functionality space

According to the definition of associative concepts, similarity in action-value vectors maps the prototypes to the same concept. In other words, the concept F is composed of all prototypes in perceptual space with almost equal action-value vectors, defined below:

$$F = \{s| \left\| Q_s - Q_F \right\| \leq \varepsilon, \varepsilon \geq 0\}, \tag{1}$$

in which:

$$Q_s = \langle Q(s, a_1), \ldots, Q(s, a_m) \rangle, \tag{2}$$

and Q_F stands for the functionality vector of concept F in which the value of each action in the concept is specified. These concepts are defined in the whole space of agent's functionality and form the zero-level concepts in hierarchy.

It is noticeable that the problem we focused on in this paper is learning how to interactively categorize stimuli in an n-dimensional continuous perceptual space. Obviously, this can be modeled as a single step episode RL mechanism. In other words, the agent is situated in different perceptual states and receives a reinforcement signal after taking one of its m predefined actions. Therefore, its next state is independent of both its current and previous states.

The values of vector Q_s is updated using the last interaction with the environment while observing s, according to the update Q-learning formula [9] which is modified for the single step episode problem definition:

$$Q(s, a) = Q_{\text{old}}(s, a) + \alpha[r - Q_{\text{old}}(s, a)], \tag{3}$$

in which α is the learning rate and r is the received reinforcement signal from the environment. The vast diversity of the agent's interaction with the environment by each action causes different uncertainties in different dimensions of vector Q_s.

Since the proposed concept learning algorithm is online, the uncertainty of the vectors Q_s in some dimensions may be so high that a concept can not be formed for corresponding actions. In fact, while the learning is not yet converged in a dimension, its corresponding concept will not be formed. In such a situation, the concept is formed in a subspace in which the uncertainty is less than a specific threshold, rather than the whole functionality space.

Therefore, a more applicable definition for functional concepts is proposed: A functional concept $F_{R^i_j}$ (the jth concept in space R^i) as a region in agent's functionality space (R^m) or a subspace of that ($R^i \subset R^m$) is defined as follows:

$$F_{R^i_j} = \{s| \left\| Q_{F_{R^i_j}} - Q_{s_{R^i}} \right\| \leq \varepsilon, \varepsilon \geq 0\}, \tag{4}$$

in which:

$$Q_{s_{R_i}} = \langle Q(s, a_x) | x \in R^i \rangle, \tag{5}$$

and $Q_{F_j^i}$ is the functionality vector of the concept $F_{R_j^i}$ in the subspace R^i.
The subspace R^i is a space in which each dimension represents an index of an action in the concept. For the sake of simplicity, the space is represented as a set of action indices:

$$R^m = \{1, 2, \ldots, m\}, \quad (6)$$
$$R^i = \{n_1, \ldots, n_j\}, n_j \in \{1, \ldots, m\}, 1 \le j \le i. \quad (7)$$

3.2. Zero-level concept extraction

Zero-level concepts are extracted by clustering vectors Q_s in functionality space. As mentioned before, the vector Q_s of each prototype s has uncertainty during the learning phase. The value of the uncertainty of Q_s in dimension i shown by $\sigma_{s,i}^2$ is reduced during learning phase as the number of experiments of action a_i increases or the variance of recent values of $Q_s(a_i)$ decreases. The uncertainty of the value of Q_s in the ith dimension is defined as follows:

$$\sigma_{s,i}^2 = \frac{\bar{\sigma}_{s,i}^2}{KC_{s,i} + 1}, \quad (8)$$

in which, $C_{s,i}$ is the number of experiments of action a_i in prototype s, $\bar{\sigma}_{s,i}^2$ is the variance of recent values of $Q_s(a_i)$ and K is a constant weight.
Therefore, Q_s's are fuzzy points in space R^m, also called *functional points*. Clusters formed by clustering Q_s's in the functionality space are called *functional clusters*. A fuzzy distance measure is needed to calculate the distance between functional points and also to find the center and members of each cluster. The distance of two fuzzy points in m-dimensional space with different fuzzinesses (uncertainties) in ith dimension should fulfill the following conditions:

- Symmetry:
$$d_i(s, s') = d_i(s', s),$$
- Zero distance:
$$d_i(s, s') = 0 \Leftrightarrow Q_s(a_i) = Q_{s'}(a_i),$$
- Effect of centers' distance:
$$|Q_s(a_i) - Q_{s'}(a_i)| < |Q_s(a_i) - Q_{s''}(a_i)|, \sigma_{s',i}^2 = \sigma_{s'',i}^2 \Rightarrow d_i(s, s') < d_i(s, s'')$$
- Effect of uncertainty:
$$|Q_s(a_i) - Q_{s'}(a_i)| = |Q_s(a_i) - Q_{s''}(a_i)|, \sigma_{s',i}^2 < \sigma_{s'',i}^2 \Rightarrow d_i(s, s') > d_i(s, s'')$$

One option for distance function definition satisfying the above conditions is:

$$d_i(s, s') = \frac{|Q_s(a_i) - Q_{s'}(a_i)|}{(\sigma_{s,i}^2 + 1)(\sigma_{s',i}^2 + 1)}, \quad (9)$$

$$d(s, s') = \sqrt{\sum_{i=1}^{m} d_i^2(s, s')}, \quad (10)$$

in which, s and s' are two prototypes in perceptual space and $d(s, s')$ specifies the distance of corresponding functional points of these prototypes in functionality space. While (9) is computed for each individual dimension, (10) is calculated over all dimensions.
As functional points, the functional clusters have also uncertainty in each dimension. The uncertainty vector of the functional clusters is actually an intra-cluster distance criteria, defined for the clusters with fuzzy point members. The uncertainty of functional clusters depends on the scattering of fuzzy point members and their uncertainties. The following equation defines the uncertainty value of the functional cluster C as an average of its members' distance to the cluster center c:

$$\sigma_{C,i} = \frac{\sum_{s \in C} d_i(s, c)}{\|C\|}, \quad (11)$$

in which $d_i(s, c)$ is defined by (9). From each functional cluster, a zero-level functional concept is formed as follows:

$$R^f = \{i \in R^m | \sigma_{C,i} < \beta\}, \quad (12)$$
$$F_{R_j^f} = \langle c_i | i \in R^f \rangle, \quad (13)$$

where $F_{R_j^f}$ is the concept formed in subspace R^f and specifies the value of actions in subspace R^f while has no idea about other actions.

3.3. Abstract concept hierarchy extraction

Assume that $F_{R_j^i}$ and $F_{R_k^p}$ are two functional concepts according to definition (4). The maximal common subspace of R^i and R^p in which functionality vectors of two concepts are similar (R^u) is the space in which $F_{R_j^i}$ and $F_{R_k^p}$ have an overlapping concept. This concept is shown by $F_{R_{k,j}^u}$.

$$R^u = \{x \in R^i \cap R^p | \left| Q_{F_{R_j^i}}(a_x) - Q_{F_{R_k^p}}(a_x) \right| \le \varepsilon\}, \quad (14)$$

$$Q_{F_{R_{k,j}^u}} = \langle Q_{F_{R_j^i}}(a_x), x \in R^u \rangle, \quad (15)$$

$$F_{R_{k,j}^u} = \{s | \left\| Q_{F_{R_{k,j}^u}} - Q_{s_{R^u}} \right\| \le \varepsilon, \varepsilon \ge 0\}. \quad (16)$$

$F_{R_{k,j}^u}$ is a higher level concept in abstract concept hierarchy which is formed by the similarity between $F_{R_j^i}$ and $F_{R_k^p}$ in some action-values.
When we move upward through the hierarchy, functional subspace of concepts becomes smaller and consequently there will be more related zero-level concepts. On the other hand, a zero-level concept may be related to several higher level concepts in different functionality subspaces. Therefore, one many-to-many relation among zero-level concepts and higher level concepts will be formed.
In the proposed approach for concept learning, abstract concepts and their mapping to perceptual space is formed during the online learning process. Fig. 2 shows the pseudo code of the abstract hierarchy extraction process in which concepts in all levels of hierarchy is formed using "is-a" relation among them.

```
REPEAT
    FOR each {a, b} in concepts
        s = maximal subspace in which a is similar to b
        c = new concept in s with action values of a (or b)
        IF c is a new concept THEN
            add c to current concepts
        add relation "a is-a c" to hierarchy
        add relation "b is-a c" to hierarchy
UNTIL no new concept was added
```

Figure 2. Pseudocode of the abstract hierarchy extraction process.

3.4. Decision making based on abstract concept hierarchy

When a stimulus of a prototype in perceptual space is observed, the similarity values among its functional points and formed concepts in functionality space are evaluated. This similarity value is evaluated by the distance between prototype's functional point and the center of functional concept in the functionality space. Since, functional points and functional concepts have the same type and defined in one space, their distance function is defined as in (9). The only difference is that each concept may be formed in a subspace of the whole functionality space. All concepts close to s and concepts above it in the hierarchy, will be candidates to be considered in decision making process about s.

During initial observations of a prototype, when the uncertainties of all actions are high, there are two strategies for action selection. The first strategy is action selection based on action-value vector Q_s and the second strategy is to utilize abstract concept hierarchy, if it is formed, and select an action corresponding to the highest level of candidate concepts hierarchy.

After the initial phase, the action-values of the candidate concepts form the final action-value vector used for the agent's decision making. The value of final action-value vector in each dimension (action) is obtained from the concept with minimum uncertainty at the corresponding action among all candidate concepts.

It is shown in section 4 that the strategy of action selection based on the hierarchy in initial phase of each prototype leads to an improvement in the results.

It is because of two reasons. The first is the agent's tendency to select the action with a lower risk which is actually in the highest level of hierarchy. It seems that, this strategy is often used by humans in the real life because of a reduction in ambiguity achieved when acting based on a higher level concept. For example, when we are not sure that the new object belongs to the concept a or b, it is better to do an action which is rewarding in both. This action is corresponding to a higher level concept located on top of concepts a and b in hierarchy.

The second justification for using the hierarchy is the advantage of finding a proper starting policy in agent's action selection over learning from scratch. In this strategy, even if the selected action is not rewarding, this action selection causes the pruning of candidate concepts' hierarchy for the coming experiments of the prototype.

Fig. 3 shows the flowchart of the learning approach.

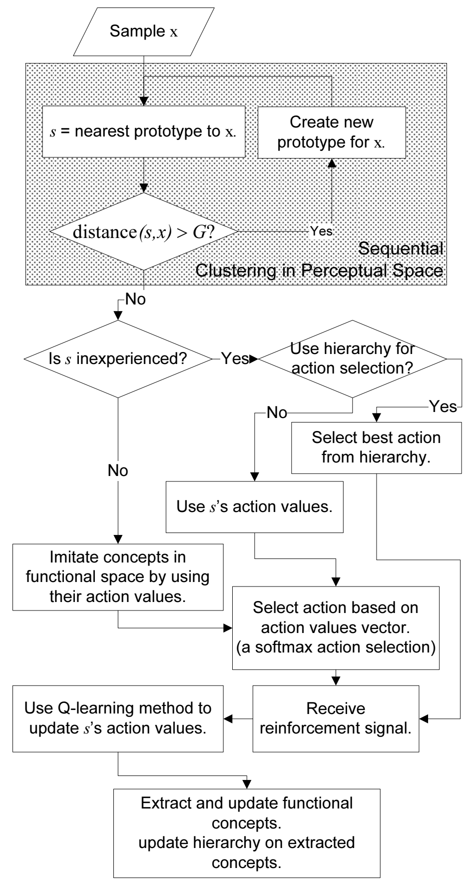

Figure 3. The decision making flowchart of the proposed approach.

4. Experiments and results

Test data sets are generated in an abstract n-dimensional continuous space to evaluate the proposed approach and to compare it with relational model of concept learning proposed in [6].

As mentioned before, each stimulus is an n-dimensional vector representing a point in the agents perceptual space. In data set generation, it is assumed that all stimuli in the perceptual space can be represented by p prototypes. This way, each prototype stands for all its surrounding stimuli and maps to one concept in the agents mind. Stimuli around each prototype are generated using a normal distribution around it with $3\sigma = G$ in order to guarantee that about 90% of stimuli will be placed inside a mass cluster with radius G.

Let $m = 14$, $n = 2$ and $p = 100$ and assume there are 8 different functional concepts in the environment. This means that there are 100 prototypes in perceptual space mapped to 8 functional concepts in agents 14-dimensional functionality space in which the values of all actions in each concept are specified. Note that the number of prototypes (p), and the number of concepts, are assumed unknown by the agent.

The learning process is performed on about 10000 randomly generated episodes in each experiment. Each episode is consisted of: generating a stimulus as explained, presenting the stimulus to the agent, selecting an action by the agent, and giving appropriate reinforcement signal to the agent. This reinforcement signal depends on the selected actions value in the corresponding concept of the given stimulus.

The experiments are iterated over several data sets generated in different perceptual spaces and different arrangements of prototypes in each perceptual space. Then the results are averaged and shown in the charts.

4.1. First experiment

In this experiment, we use different data sets to find those situations in which the proposed functional concept extraction method overcomes relational model for concept extraction proposed in [6]:

1. In this data set, neighbor prototypes in perceptual space, have dissimilar functionalities. This data set models the associative concepts.
2. In this data set, neighbor prototypes in perceptual space, have similar functionalities. This data set models the relational concepts.
3. In this data set, the prototypes in perceptual space randomly map to functionalities. This data set models the real environment which contains relational and associative concepts.

The improvement rate of average reward gained during the agent's life in our approach and the approach in [6] (relational concept learning) is plotted for three data sets (Fig. 4).

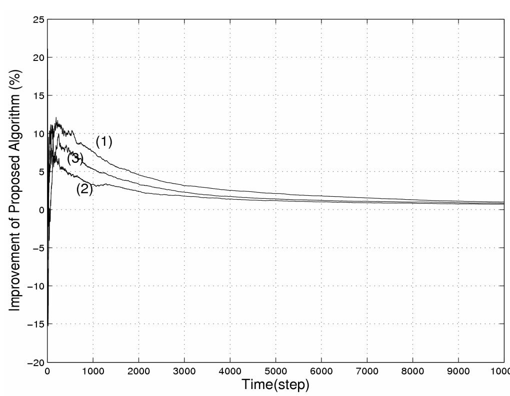

Figure 4. The improvement by the proposed approach compared to relational concept extraction in three data sets of the first experiment.

As it is shown in Fig. 4, the proposed method has shown a reasonable amount of improvement in all data sets. The improvement rate is lower in data set 2 and it may be because similarity in perceptual space helps the relational concept learning method to achieve a better average reward. We claim it is the concept formation in functionality space which makes the ability of generalization in all data sets.

In data set 3 which is a random mapping between sensory and functionality space, the curve is in the middle of two other data sets. It seems that the random mapping data set generation is more close to the model of a real environment containing both relational and associative concepts. So, in the next experiments the data sets are generated randomly in more complex perceptual and functionality spaces.

4.2. Second experiment

In this experiment four methods of learning in same situation of simulation are compared:

- A1. The method of relational concept learning in [6].
- A2. The proposed algorithm for extracting functional concepts and their abstraction hierarchy during the learning with no pre-knowledge in perceptual and functionality spaces.
- A3. The method in which the agent knows the functional concepts and the abstraction hierarchy (generated by method A2) as an abstract pre-knowledge in functionality space. But it has no information about the perceptual space.
- A4. The method is similar to method A3 but the agent's concepts and abstraction hierarchy are the ideal concepts which are manually extracted by the supervisor.

Fig. 5 compares the learning of the agent in these four methods in terms of their average reward during their life.

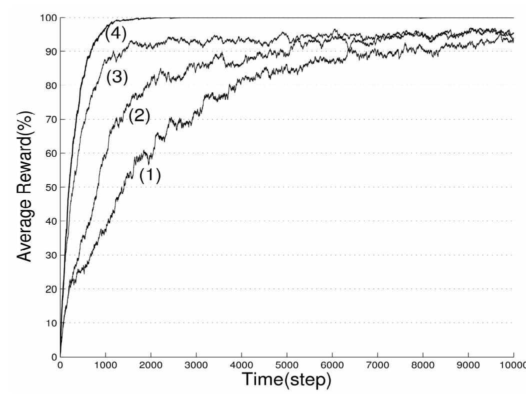

Figure 5. The average rewards of the four methods of the second experiment during the agent's life.

From learning speed perspective, we can compare these four approaches in different situations. In methods A1 and A2, the agent begins to learn the environment with the same pre-knowledge. But the abstraction of concepts in functionality space and extraction of abstract concept hierarchy helps the agent in method A2 to have generalization in confronting new or less observed stimulus.

In methods A3 and A4 the agent's rate of learning increases significantly. It seems that the reason is the agent's abstract pre-knowledge which is formed only in functionality space and is utilized for generalization in a new perceptual space.

Comparing method A3 and method A4 shows the difference in learning's average reward once when we have ideal pre-knowledge and once with the pre-knowledge that is learned by another agent and now is shared for this agent. It is necessary to pay attention to this fact that, the abstract knowledge is only in functionality space while their sensory spaces are completely different.

In table 1 the learning speed of four methods is shown. When the agent arrives to 90% of its maximum average reward, the number of passed episodes is counted and reported in the table.

Table 1. Comparision of the learning speed of methods of the second experiment.

Method	A1	A2	A3	A4
Time to reach 90% of maximum reward	8360	5060	1309	680

4.3. Third experiment

In this experiment, the importance of using abstract concept hierarchy is shown. The average reward of two methods is compared:

B1. The method is similar to method A4 of the second experiment.
B2. Similar to B1, but the agent uses only functional concepts and ignores the abstract concept hierarchy.

These two methods are compared in Fig. 6 in term of average reward of the agent during its life. As it is shown, using abstract concept hierarchy speeds up the learning.

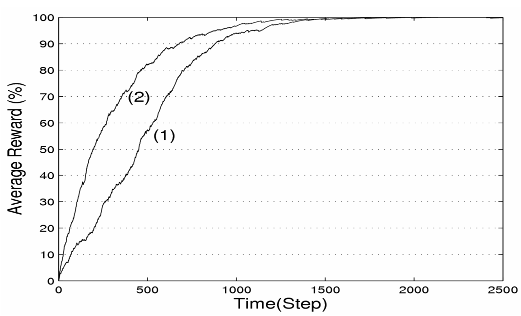

Figure 6. The average rewards of the two methods in the third experiment during the agent's life.

One reason is that, decision making based on higher level concepts in the hierarchy helps the agent to make decisions with less ambiguity especially when the agent is very uncertain about the real concept. Additionally, this strategy may help to reduce ambiguity in real concept recognition, in coming decision makings.

5. Conclusions and future works

In this research we focused on a new space in which concepts have more suitable and compact representations. This space is the agent's functionality space. In this space, each point is an action-value vector and its neighborhood contributes to the same concept too. In fact, extremely distant points in perceptual space which belong to one concept, map to neighboring points in functionality space. We proposed an approach for extraction and formation of concepts in this space in an online and autonomous manner based on a partial similarity between these concepts which forms an abstract concept hierarchy in functionality space.

The results show that the representation of concepts in functionality space helps agent to generalize learnt concepts to new observations and gain a better performance in comparison with usual reinforcement learning methods, specially when the concepts in perceptual space is associative and the similarity in this space does not contribute to have the same concept. Moreover, it is shown in the experiments that using the abstract concept hierarchy, which may be extracted by other agents or produced manually by the supervisor, accelerate the learning process four times.

Consequently, the proposed learning framework may be used and modified as a knowledge abstraction strategy for sharing knowledge in heterogeneous agents in different sensory spaces. One more next step is extraction of other types of relations among concepts; such as: has-a for composite concepts and affects/causes/... for temporal concepts. Finally, we can add attention control in both perceptual space and functionality space to help the agent to narrow down its field of view for more efficient operation.

Acknowledgement

I want to thank Iran Telecom. Research Center (ITRC) for the financial support of this research.

References

[1] S. Amizadeh, M. Nili Ahmadabadi, B. Nadjar Araabi, and R. Siegwart. A bayesian approach to conceptualization using reinforcement learning. In *Proc. of 2007 IEEE/ASME Intl. Conf. on Advanced Intelligent Mechatronics*, Switzerland, September 2007.

[2] M. A. Arbib. *The Mirror System, Imitation, and the Evolution of Language, Imitation in Animals and Artifacts*. MIT Press, 2002.

[3] J. Krichmar and G. Edelman. Machine psychology: Autonomous behavior, perceptual categorization and conditioning in a brain-based device. *Cerebral Cortex*, 12(8):818–830, August 2002.

[4] M. Luiz, A. Oliveira, R. Grupen, D. Wheeler, and A. Fagg. Tracing patterns and attention: humanoid robot cognition. *IEEE Intelligent Systems*, 15(4):70–77, July/August 2000.

[5] T. Matsuzawa. Comparative cognitive development. *Developmental Science*, 10(1):97–103, 2007.

[6] H. Mobahi, M. Nili Ahmadabadi, and B. Nadjar Araabi. A biologically inspired method for conceptual imitation using reinforcement learning. *Applied Artificial Intelligence*, 21(3):155–183, March 2007.

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

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

[9] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

[10] T. R. Zentall, M. Galizio, and T. S. Critchfied. Categorization, concept learning, and behavior analysis: an introduction. *The Experimental Analysis of Behavior*, 3(78):237–248, November 2002.