# A Fuzzy Reinforcement Function for the Intelligent Agent to process Vague Goals

Ho-Sub Seo*,
Information Communication R&D Center
DACOM
DACOM Bldg., 706-1, Yeoksam-dong, Kangnam-ku
Seoul 135-610, Korea
garam910@chollian.net

So-Joeng Youn, Kyung-Whan Oh
Computer Science and Engineering Dept.
Sogang University
Shinsoo-dong 1, Mapo-ku,
Seoul 121-742, Korea
sjyoun@ailab.sogang.ac.kr, kwoh@ccs.sogang.ac.kr

**Abstract**

The intelligent agent is one of the most interesting fields of Artificial Intelligence study. Generally, very many kinds of the intelligent agent receive the user's goal and they try to solve it with their expert knowledge. The user's goals or requests can be represented with the human language, and they contain the uncertainties of the human knowledge. While the intelligent agent must represent these vague goals and understand the user's desires or intentions, there have not been enough researches done for the intelligent agents to express the user's goals.

In this paper, we propose a new method to represent the vague goals as well as the uncertain environments. We suggest a fuzzy goal and a fuzzy state representation. We extend the traditional reinforcement learning to the fuzzy reinforcement learning with defining the fuzzy reinforcement function by using the fuzzy goal and the fuzzy state. We, also propose a new FuzzyQ-Learning algorithm. The experiment results show the better performance of the learning, and the reasonability of the fuzzy reinforcement learning.

## 1. Introduction

As the computing environments have been changed to the Internet and to the distributed environments, the agent-based technology has become a new paradigm of Artificial Intelligence and Computer Science[1]. An intelligent agent is a software that performs the user's task instead of the user, and it tries to solve problems through the interaction with the given environments.

The intelligent agent must find optimal solutions from given user's goals and the environments that the agent explores. If the given goals and the environments comprise of any uncertainties, the agents cannot perform the proper actions. These uncertainties are related with the perception of the uncertain environments, the action of other agents, and the knowledge of the intelligent agents. There were many researches attempted to solve these kinds of problems, and mostly by using the probabilistic and stochastic methods [2][3][4][5]. But, based on the previous researches, it is impossible to handle user's goals represented with linguistic values that comprise of uncertainties in human language and knowledge. Let's say, if a user gives a vague goal represented with the linguistic values, such as *"large"*, *"most"*, *"small"*, the agent must understand the user's desire. After understanding it, an agent must perceive the uncertain environments and initiate proper actions to solve the given problem or to get the optimal solutions.

We propose a new approach the intelligent agents to use a fuzzy goal and a fuzzy state to represent the user's vague goal and the uncertain environments. For example, suppose that the given user's goal is *"Find the information about the car with the proper price, by searching the enough number of web sites, and by spending sufficient time to search"*. In this case, the agent must understand the vague meaning of the *"proper"*, *"enough"*, *and "sufficient"*. The agent represents its own goal by a vector of the appropriate fuzzy sets – it is a fuzzy goal. The agent represents the uncertain environments in which it moves to find the optimal solutions with the vector of the fuzzy sets – it is a fuzzy state. Therefore, the intelligent agent can understand the user's vague goal and can perceive necessary information in the uncertain environments by using the fuzzy goals and the fuzzy states.

In this paper, we extend the traditional reinforcement learning by using the fuzzy goal and the fuzzy state representations. We propose a new fuzzy reinforcement function and a FuzzyQ-Learning algorithm. We use the fuzzy reinforcement function to modify the Q-Learning that is one of the well-known reinforcement learning algorithms. In our FuzzyQ-Learning algorithm, we use a new FuzzyQ-Function instead of the Q-Function and a FuzzyQ-Table.

We tested a validity of the proposed learning model through the experiments of the modified grid world problem that is a traditional problem domain in the reinforcement learning. A new FuzzyQ-Learning agent uses the FuzzyQ-Table and the FuzzyQ-Function to learn the optimal policy in the uncertain environments and with the user's vague goals.

This paper is organized as follows. Section 2 covers a new representation of the fuzzy goal and the fuzzy state. In section 3 we propose a new fuzzy reinforcement function with the proposed representation and extend the traditional Q-Learning algorithm. We do this by using a new fuzzy reinforcement function to a new FuzzyQ-Learning algorithm. In section 4, we make experiments to compare our new FuzzyQ-Learning algorithm with the Q-Learning. And the section presents the reasonability of the proposed FuzzyQ-Learning through the experiment in the modified grid world problem. We make conclusions in section 5.

## 2. Fuzzy Goal and Fuzzy State

Recently, the intelligent agent has become one of the important issues in Artificial Intelligence, and it is the ultimate goal of Artificial Intelligence to make the "thinking" computers. A lot of engineers and scientists have researched about the various aspects of properties of the intelligent agent and its abilities. Therefore, many researchers expect that the intelligent agent will think and act like human to accomplish the given goal in the near future. Generally, a user gives his or her own goal to the intelligent agent, and it is supposed to perform some actions by itself and adapt to its environments to achieve the given goal. First of all, the intelligent must understand the user's goals and purposes to achieve the appropriate result. If the user gives the goal to an agent with the linguistic representation of the human languages, the agent must understand it before choosing any kind of an action. In this case, the agent needs some methods to clarify the user's vague goals. In addition to it, the intelligent agent has to perceive and explore the real world like the Internet to solve the given problem, which has various uncertain information The intelligent agent is confronted with the various kinds of uncertainties along the way, so the agent has to understand not only the user's vague goal, but also the uncertain environments.

In this paper, we propose a new method to represent vague goals and uncertain environments. First of all, we make the intelligent agent to understand a user's vague goal. We make an expression to represent an intelligent agent's vague goal with a sequence of fuzzy set. We define a fuzzy goal as a vector of *n* fuzzy sets from the user's goal, which is represented with a linguistic value. If the *n* fuzzy sets are $\tilde{G}_1, \tilde{G}_2, \ldots, \tilde{G}_n$ from the n linguistic expressions for the user, we define a fuzzy goal $\breve{G}$ of the intelligent agent with the equation (1)

$$\breve{G} = \left(\tilde{G}_1, \tilde{G}_2, \ldots \tilde{G}_n\right) \qquad (1)$$

For example in section 1, if a user gives an Internet search agent a goal – *"Find the information about the car with the proper price, by searching the enough number of web sites, and by spending sufficient time to search"*. In this case, the agent must understand the vague meaning of the *"proper"*, *"enough"*, *"sufficient"*. In this example, the linguistic value of the goal, such as *"proper"*, *"enough"*, *"sufficient"*, can be changed into proper fuzzy sets, MIP(**MI**ddle **P**rice), MAN(**MA**ny **N**umber), and MUT(**MU**ch **T**ime). Then, the fuzzy goal of intelligent agent is

$$\breve{G} = \left(MIP, MAN, MUT\right) \qquad (2)$$

Each fuzzy set corresponds to the linguistic value of the user's vague goal. Similarly, we define a fuzzy state as a vector of fuzzy sets. The fuzzy set in the fuzzy state is obtained from the perception of the uncertain environments. Therefore, if we get the *n* fuzzy sets from the environments - $\tilde{S}_1, \tilde{S}_2, \ldots, \tilde{S}_n$, the fuzzy state from the real world environments is

$$\breve{S} = \left(\tilde{S}_1, \tilde{S}_2, \ldots, \tilde{S}_n\right) \qquad (3)$$

Each fuzzy set in the fuzzy state corresponds to each fuzzy set in the fuzzy goal in order. If the intelligent agent has proper fuzzy sets as its expert knowledge for the given problem domain - goal and environments, it can express the uncertainties from user's vague goal and the uncertain real world environments with this fuzzification methods. Then, the intelligent agent can understand the vague goal and can perceive the uncertain environments.

## 3. Fuzzy Reinforcement Function and FuzzyQ-Learning

The reinforcement learning is one of the supervised learning algorithms that have been used for the intelligent agents in the dynamic environments in order to satisfy the autonomy and the adaptability. The reinforcement learning is based on Markov Decision Process and Dynamic Programming. In the reinforcement learning algorithm, each agent acts on a given environment and receives the immediate reinforcement value. The agent chooses the proper action set to get the maximum value of expected sum of the reinforcement value, and we call the proper action set is the optimal policy. The process of getting the action set is the reinforcement learning[6].

The Q-Learning is one of the famous algorithms of reinforcement learning, which uses The Q-Function to evaluate the agent's state. If the agent learns with the one-step Q-Learning, the agent can perform on-line learning in dynamic environments, because it learns based on one action state value whereas other algorithms calculate current sate value according to rewards of all actions. Suppose that the agent chooses the action $a$ in state $S$, and moves to another state $S'$, so it obtains the reinforcement value $r(s,a)$ like the diagram in Figure 1.

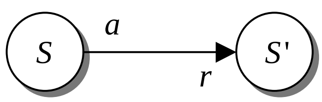


Figure 1. Action State Transition Diagram

Then, the Q-Function can be defined as following.

$$Q(s,a) \leftarrow r(s,a) + \gamma \max_{a'} Q(s',a') \qquad (4)$$

In the equation (4), $a'$ means the action in the state $S'$, and $\gamma$ is the discount factor between 0 and 1. If the agent chooses the action to make the higher Q-Function value, it can learn the optimal policy. The agent that learns by using the Q-Function has the Q-Table that is made up of the state, action, and Q-Function value. It updates the Q-Table at every learning step to make the Q-function value much higher[6].

We extend the traditional reinforcement learning by using the fuzzy goal and the fuzzy state representations. We show a new fuzzy reinforcement function and FuzzyQ-Learning algorithm with the fuzzy goal and the fuzzy state. We use the fuzzy reinforcement function to modify the Q-Learning. In our FuzzyQ-Learning algorithm, we use a new FuzzyQ-Function instead of the Q-Function. The FuzzyQ-Function is produced with the new fuzzy reinforcement function. The fuzzy reinforcement learning agent has a FuzzyQ-Table, and it updates the FuzzyQ-Table at each step of actions to find the optimal solutions.

In Figure 2, we extend the general action state transition to the fuzzy state transition first. Suppose that the intelligent agent that has a fuzzy goal $\breve{G}$ choose an action $a$ in a fuzzy state $\breve{S}$, and moves to a fuzzy state $\breve{S}'$, in Figure 2. Then, the agent get a fuzzy reinforcement value $FR(\breve{S},a)$, also. We define this process of an action is “Fuzzy State Transition”.

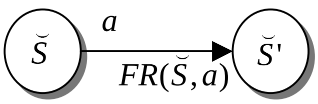

Figure 2. Fuzzy State Transition

At this time, let the i^th fuzzy set in the fuzzy state $\breve{S}$, $\breve{S}'$ and the fuzzy goal $\breve{G}$ are $\tilde{S}_i$, $\tilde{S}'_i$, and $\tilde{G}_i$. The ith fuzzy reinforcement value $fr_i$ can be expressed in Equation 5 and Figure 3.

$$fr_i = \max\{\mu_{\tilde{S}_i} \wedge \mu_{\tilde{G}'_i}\} = \max\{\min\{\mu_{\tilde{S}_i}, \mu_{\tilde{G}'_i}\}\} \qquad (5)$$

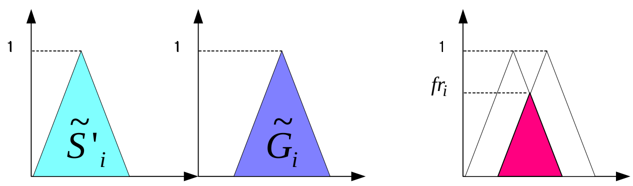

Figure 3. Calculation of the fuzzy reinforcement value by using the fuzzy goal and the fuzzy state

According to the equation (5) and Figure 3, as the fuzzy state gets closer to the fuzzy goal at the next step, the fuzzy reinforcement agent gets the much bigger reinforcement value. If the next fuzzy state $\breve{S}'$ is the fuzzy goal $\breve{G}$, the agent takes the maximum reinforcement. We define the fuzzy reinforcement function $FR(\breve{S},a)$ from the equation (5) and Figure 3.

$$FR(\breve{S},a) = r \cdot \min_{fr_i^m}\{fr_1^m, fr_2^m, \ldots, fr_n^m\} \qquad (6)$$

In the equation (6), $r$ means the maximum fuzzy reinforcement value ($\because 0 \le \min_{fr_i^m}\{fr_1^m, fr_2^m, \ldots fr_n^m\} \le 1$), and $m$ is the positive integer value ($2 \le m$) for getting the highest reinforcement value when the agent move from the adjecent fuzzy state to the fuzzy goal. We extend the Q-Learning to the FuzzyQ-Learning using the fuzzy reinforcement function in the equation (6). We define the FuzzyQ-Function as follow.

$$FuzzyQ(\breve{S},a) \leftarrow FR(\breve{S},a) + \gamma \max_{a'} FuzzyQ(\breve{S}',a') \quad (7)$$

In this equation, $\gamma$ is the discount factor like the Q-Function. The fuzzy reinforcement has the FuzzyQ-Table, which is composed of the fuzzy state, agent’s action, and FuzzyQ-Function value. In Each step of the learning, the agent updates this FuzzyQ-Table, and performs the fuzzy reinforcement learning, FuzzyQ-Learning. A new algorithm of the FuzzyQ-Learning is as shown in Table 1.

Table 1. FuzzyQ-Learning Algorithm

1. For each fuzzy state $\breve{S}$ and action $a$, initialize the table entry $FuzzyQ(\breve{S},a)$ to zero
2. Observe the current fuzzy state $\breve{S}$
3. Do forever:
   - Select an action $a$ and execute it.
   - Receive immediate fuzzy reward $FR(\breve{S},a)$
   - Observe the new fuzzy state $\breve{S}'$
   - Update the table entry for $FuzzyQ(\breve{S},a)$ as the equation (7)
   - Move to the new fuzzy state $\breve{S}'$

## 4. Experiments

In this paper, we tested the validity of the fuzzy reinforcement function in the equation (6) and FuzzyQ-Learning algorithm in the Table 1. The experiment domain is the grid world problem that is one of the popular problem domains in the traditional reinforcement learning. We modified the grid world problem to in order to make about our proposed algorithm experiments. The grid world problem is a very simple problem domain. In the grid world, there are a start position and a goal position. The reinforcement agent starts from the start position, and tries to find the goal position as fast as possible, to compare the various kinds of the reinforcement learning algorithms.

In the experiment, we used the next equation to update the FuzzyQ-Table. $\alpha$ means the learning rate to control the learning speed of an agent. Moreover, we used the Boltzmann exploration strategy as the strategy to choose the action in the experiments. The Boltzmann exploration strategy has been generally used in the reinforcement learning algorithms.

$$FuzzyQ(\breve{S},a) \leftarrow FR(\breve{S},a) + \alpha\{FR(\breve{S},a) + \gamma \max_{a'} FuzzyQ(\breve{S}',a') - FuzzyQ(\breve{S},a)\} \tag{8}$$

We made two experiments. The first one was a comparison between the traditional Q-Learning algorithm and the new FuzzyQ-Learning algorithm. The second experiment was performed to show the reasonability of the FuzzyQ-Learning algorithm in the modified gird world problem domain – the fuzzy goal of the intelligent agent.

1) *Experiment 1*: Comparison with the Q-Learning.

For the first experiment, we made the 8-by-8 grid world environment, which has only one goal that settles the specific position. The fuzzy reinforcement agent tried to find the goal position from a random start position. The agent can move only in four directions – up, down, left or right. We made two fuzzy reinforcement agents that had the fuzzy goal to find the goal. The first one had the fuzzy goal in the equation (9) and its four fuzzy sets are in Figure 4.

$$\breve{G}_{Distance} = (SHD) \tag{9}$$

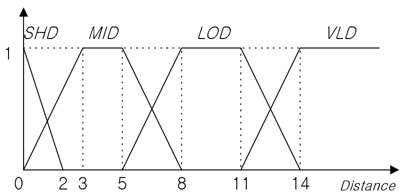


Figure 4. Fuzzy set for the distance between the start position and the goal

The second agent had more information for the fuzzy reinforcement learning. This agent had the fuzzy goal with two fuzzy sets. First fuzzy se represented the distance along to the x-axis; the other one was the distance along to the y-axis. The fuzzy goal of the second agent is as the following equation (10) and its eight fuzzy sets are in Figure 5.

$$\breve{G}_{xy} = (SHD_x, SHD_y) \tag{10}$$

We compared these two fuzzy reinforcement agents with the traditional reinforcement agent, which used the Q-Learning instead of the FuzzyQ-Learning. The result of the experiment is shown in the following graph of Figure 6.

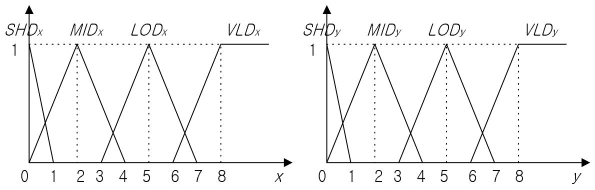


Figure 5. Fuzzy sets for the distance along to the x-axis and the y-axis.

In Figure 6, FQ1 means the first fuzzy reinforcement agent, FQ2 is the second one, and Q means the Q-Learning agent. “Epoch” means the one period of learning, which is the process of agent from the start position to the goal position, and “Step” means the number of the each agent’s transfer.

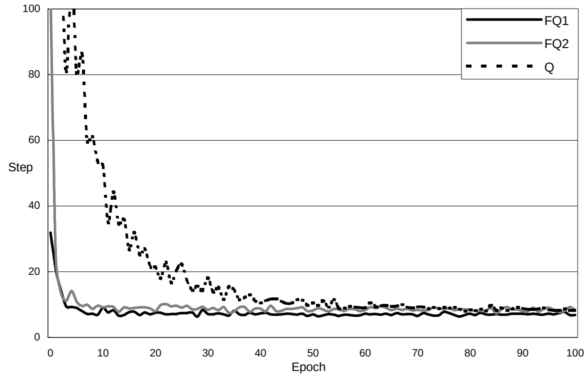


Figure 6. Comparing the Fuzzy Q-Learning and the Q-Learning

According to the graph shown above, FQ1 agent converged fewer than 10 epochs and FQ2 agent did, too. On the other hand, Q agent converged about 50 epochs. The two FuzzyQ-Learning agents were much faster convergence speed than the Q-Learning agent. The reason of the different convergence speed, two fuzzy reinforcement agents had the smaller number of entries in FuzzyQ-Table than those of the Q-Table in the Q-Learning agent. FQ1 agent had 16 entries of the FuzzyQ-Table and FQ2 agent had 64 entries to have to update in the learning process, while the Q agent had 256 entries of the Q-Table. This difference of the table size caused the gap of the learning speed. The FuzzyQ-Learning agent had a much better learning efficiency and was able to find the optimal policy for the vague goal.

2) *Experiment 2*: Fuzzy goal problem

In the second experiment, we made the same size of gird world, which had one reward point in (7,7) coordinates and one hazard point in (0,7). As the agent went closer to the reward point, it got more benefit. In the same way, as it went near to the hazard point, it spent much more costs. The benefit and the cost of the agent were as shown in the equation (11). In the equation (11), $b$, $c$ mean the amount of benefit and cost. $B$, $C$ are the maximum value of benefit and cost. The value of $d_r$, $d_h$ means the distance from the agent to the reward and the distance from the agent to the hazard.

$$b = \begin{cases} B & \text{if } d_r = 0 \\ B/(d_r + 0.5)^2 & \text{if } d_r \neq 0 \end{cases}, \qquad c = \begin{cases} C & \text{if } d_h = 0 \\ C/(d_h + 0.5)^2 & \text{if } d_h \neq 0 \end{cases} \tag{11}$$

Then, we gave the agent the fuzzy goal as shown in the equation (12) and the fuzzy sets in Figure 7, it meant, the agent had to learn a policy in order to make the benefit very large and keep the cost very low.

$$\breve{G}_{BandC} = (VL_B, VL_C) \tag{12}$$

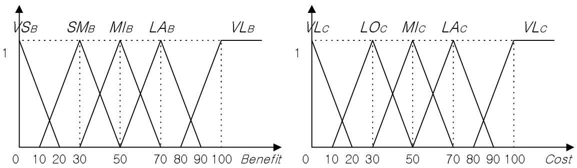


Figure 7. Fuzzy sets for the benefit and the cost of the fuzzy reinforcement agent.

After the experiment, the intelligent agent learned the action set in Figure 8. The agent divided the virtual environment into 14 fuzzy states, and it chose a consistent action in each fuzzy state. The fuzzy reinforcement agent tried to avoid the hazard and went to the reward to get more benefit. But the Q-Learning agent could not learn the optimal policy with any kind of the learning arguments – learning rate, discount factor, and Boltzmann constant.

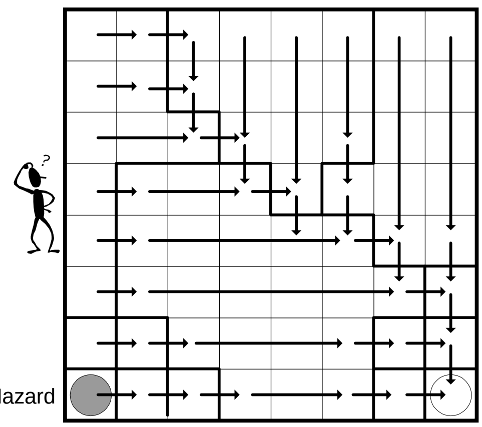


Figure 8. Fuzzy intelligent agent considering the benefit and the cost.

## 5. Conclusion

In this paper, we proposed a new method for the intelligent agent to represent the user's vague goal and the uncertain environments with the fuzzy goal and the fuzzy state. In addition to the new representation, we extended Q-Learning, which is one of the famous traditional reinforcement algorithms, to the FuzzyQ-Learning. For this extension, we proposed the fuzzy reinforcement function with the fuzzy goal and the fuzzy state, and the fuzzy reinforcement agent that we designed could learn the optimal policy from the user's vague goal and uncertain environments. We showed the fuzzy reinforcement agent learned faster than the Q-Learning.

However, some limitations still remain in this method. First of all, we need the exact analysis to get the proper fuzzy sets to the vague goal and the uncertain environments. Moreover, the intelligent agent should be able to explore the dynamic environments like the Internet, so we need to extend our algorithm to adapt to the dynamic and distributed environments. In spite of these limitations, the new method could be used in the various cases of the uncertainties in the intelligent agents.

## Acknowledgement


This work was supported by the Brain Korea 21 Project.


## Reference


[1] Katia P. Sycara, "Multiagents Systems", *AI magazine*, Summer, pp.79-92. 1998

[2] Leslie Pack Kaelbling, Michael L. Littman and Anthony R. Cassandra, "Planning and Acting in Partially Observable Stochastic Domains", *Artificial Intelligence*, Vol. 101, 1998

[3] Ping Xuan, Victor R. Lesser, "Handling Uncertainty in Multi-Agent Commitments", *UMASS Computer Science Technical Report*, 1999-05, Jan 18, 1999.

[4] Sidi O. Soueina, Behorouz H, Far, T. Katube, Z. Koono, "MALL:A Multi-Agent Learning Language for Competitive and Uncertain Environments", *Special Issue on Knowledge-Based Software Engineering*, V.E 81-D N.12, 1998.12

[5] David Poole, "The independent choice logic for model-ling multiple agents under uncertainty", *Artificial Intelligence*, Vol. 94, July 1, 1997

[6] Tom M. Mitchell, "Machine Learning", *WCB/McGraw-Hill*, 1997