# A Simple Goal Seeking Navigation Method for a Mobile Robot using Human Sense, Fuzzy Logic and Reinforcement Learning

Hamid Boubertakh, Mohamed Tadjine, Pierre-Yves Glorennec, Salim Labiod

*Abstract* — **This paper proposes a new fuzzy logic-based navigation method for a mobile robot moving in an unknown environment. This method allows the robot obstacles avoidance and goal seeking without being stuck in local minima. A simple Fuzzy controller is constructed based on the human sense and a fuzzy reinforcement learning algorithm is used to fine tune the fuzzy rule base parameters. The advantages of the proposed method are its simplicity, its easy implementation for industrial applications, and the robot joins its objective despite the environment complexity. Some simulation results of the proposed method and a comparison with previous works are provided.**

*Key-words* — **Fuzzy logic, Reinforcement learning, Mobile robot navigation, Obstacle avoidance.**

## I. INTRODUCTION

Path planning is an important topic in robotics, mainly for its practical applications, and can be classified on two types; the global path planning and the local path planning. The first one is done in off-line manner, and since the path is designed, the user can decide about the control method to use to follow the path by the robot. Many methods are developed in the literature; A* algorithm [1], potential field method [2, 3], cell decomposition method [3] and recently, the ant colony methods [4, 5]. The second one also named obstacle avoidance or navigation is done in on-line reactive manner; it imposes that the robot must be equipped by sensors to have a vision of its neighbourhood, and after, takes the adequate action to achieve its a priori defined task. We can find in the literature a large number of publications dedicated to this purpose where several methods are used; the force field [6], the reinforcement learning [7], and the one witch has known a great success in the last years is the fuzzy logic [1, 8-13]. The success of fuzzy logic in mobile robots navigation is due to its capability to represent the human reasoning and therefore the robot doesn't need an exact vision of its environment. Recently many research woks in vehicles navigation using fuzzy logic are published.

In [8, 9], authors used a fuzzy rule base composed of 243 rules. The use of such a big rule base is unjustified and questions the applicability of the method in real time. In [10-13], the authors proposed simples fuzzy rule bases. But these methods have a common drawback; the robot can be stuck in a local minimum in complex environments. The problem of local minima was treated in [1]; the authors proposed for the goal seeking the fusion of two elementary behaviours; the convex obstacles avoidance behaviour is achieved by 25+25 fuzzy rules, and the wall following behaviour is achieved by 15 fuzzy rules. Many of the above works use the reinforcement learning for parameters adjustment of the fuzzy navigator. In [9] a reinforcement learning algorithm assisted by a supervised learning algorithm is used, at the beginning, the supervised learning is performed by a gradient descent-based algorithm witch is used to tune the premise parameters, and after, the reinforcement learning is used for fine tuning of both premise and the conclusion parameters of the fuzzy rule base. This learning process is very complicated and may lead to a non interpretable rule base. In [10, 12], authors used the fuzzy reinforcement Q-learning algorithm. The main advantage of this algorithm is its ease for real time implementation.

In the present work, we propose a new simple fuzzy logic-based navigation method for a mobile robot. We use a single fuzzy rule base composed of 8 rules inspired from the human reasoning where both obstacles avoidance and goal seeking behaviours are merged. To achieve the fine tune of the navigator conclusion parameters of the rule base, the reinforcement Fuzzy Q-Learning algorithm is used. The proposed navigation method is compared to previous works.

This paper is organized as follows. Section II describes the considered robot architecture, in section III, the proposed fuzzy navigator is presented with the reinforcement fuzzy Q-learning algorithm. Section IV presents the simulation results and finally, section V, concludes the paper.

## II. ROBOT MODEL

We use a cylindrical omnidirectional mobile robot model with a radius of 20 cm [9]. The robot is equipped with 24 ultrasonic sensors evenly distributed in a ring as depicted in

Manuscript received September,18, 2008.
H Boubertakh is with LAMEL, Faculté des Sciences de l'Ingénieur, University of Jijel, BP. 98, Ouled Aissa, 18000, Jijel, Algeria
LCP, Département de Génie Electrique, Ecole Nationale Polytechnique, 10 av. Hassen Badi, BP. 182, El Harrach, Alger, Algeria.
INSA de Rennes, France, 20 av. des Buttes de Coëmes-CS 14315-35043 Rennes Cedex, France (email : boubert_hamid@yahoo.com)
Mohamed Tadjine is with Département de Génie Electrique, Ecole Nationale Polytechnique (email : tadjine@yahoo.fr)
Pierre-Yves Glorennec is with INSA de Rennes, France (email:glorenne@irisa.fr)
Salim Labiod is with LAMEL, Faculté des Sciences de l'Ingénieur, University of Jijel

DOI:10.2298/JAC0801023B

Fig.1. Each sensor, $s_i$ for $i$=1,…,24, covers an angular view of 15° and gives the distance to the obstacle $l_i$ in its field of view. T o r educe th e number o f i nputs for the na vigator, sensors in the front of the robot are arranged into tree sensor groups; t he l eft g roup *SL* consists of th e 3 neighbouring sensors $s_i$ (i=1,…,3), th e f ace gr oup *SF* consists of th e 6 neighbouring s ensors $s_i$ (i=4,…,9), a nd th e r ight gr oup *SR* consists of the 3 n eighbouring s ensors $s_i$ (i=10,…,12). T he distances measured by t he t ree g roups *SL*, *SF* and *SR* denoted r espectively by *dl*, *df* and *dr* are ex pressed a s follows:

$$dr = R + \min\left(l_i \; / \; i = 1,2,3\right)$$
$$df = R + \min\left(l_i \; / \; i = 1,..,9\right)$$
$$dl = R + \min\left(l_i \; / \; i = 10,11,12\right)$$

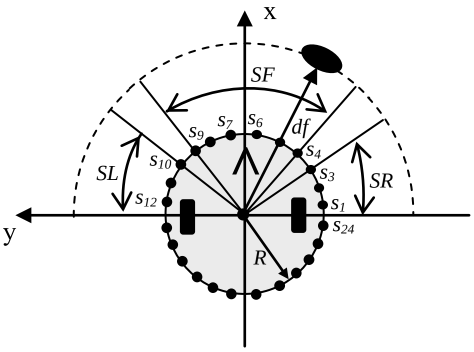


Fig. 1. Mobile robot and the sensor arrangement

We use two coordinate systems; the world coordinate system *XOY* and the mobile robot coordinate system *xoy* where *o* is in the center of the robot and the *x* axis goes in the middle between the two sensors $s_6$ and $s_7$ (see Fig. 2). The robot actions ar e t he change of t he he ading angle $\Delta\Phi$ and the linear velocity *v* of the robot. For a goal seeking behaviour, the robot knows the position of its goal and $\theta$ defined as the angle between the orientation axis and the line connecting the center of the robot to the goal.

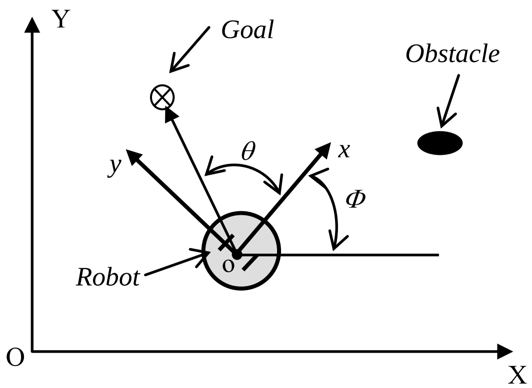


Fig. 2. The system coordinates

## III. FUZZY NAVIGATOR

The f uzzy na vigator inputs are d istances i n t he t hree directions; right *dr*, face *df* and left *dl*, and the outputs are the speed *v* of the robot centre and the change of the steering angle $\Delta\Phi$.

### A. Fuzzification

Two fu zzy l abels ar e u sed t o describe e ach of t he t hree distances; Near ( N) a nd Far (F ) as shown in Fig. 3. T he fuzzy labels have the following membership functions:

$$\mu_N(d) = \min\left(\max\left(0, \frac{d - d_s}{d_m - d_s}\right), 1\right)$$

$$\mu_F(d) = \min\left(\max\left(0, \frac{d - d_m}{d_s - d_m}\right), 1\right)$$

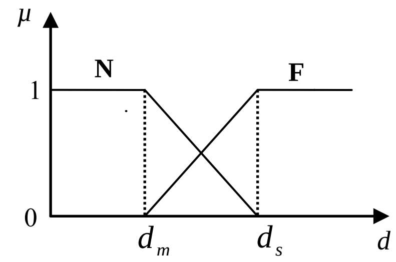


Fig. 3. Inputs membership functions

where $d_m$ is The minimum permitted distance to an obstacle, and $d_s$ is the safety di stance be yond which t he r obot c an move at high speed.

### B. Fuzzy Rules base

The rule base for both obstacles avoidance and goal seeking is c onstructed ba sed o n t he hum an r easoning. I t can b e interpreted as:

- If t he robot is f ar f rom o bstacles i n i ts t hree directions, then the robot steers to the goal and goes with its maximum speed;
- If the goal is not in the front of the robot and there exist obstacles, then the robot follows the nearest obstacle o n i ts r ight o r l eft, according t o the smallest distance to the obstacle.
- If both the goal and obstacles are in the front of the robot, then the robot tries to steer to the goal and follows the nearest obstacle o n i ts r ight o r l eft, according to the smallest distance to the obstacle.

Let us define two parameters *a* and *b* as follows:

$$a = \begin{cases} 0 \; if \;\; g \le d \;\; et \;\; g \le d_M \\ 1 \; else \end{cases} \text{, and } \; b = \begin{cases} 1 \; if \;\; 90° < \theta < 270° \\ 0 \; else \end{cases}$$

where $d_M$ is the maximum distance that can be detected by the sensors.

The two parameters *a* and *b* allow the navigator to select its behaviour m ode; i f *p*=1 (*p*=0) t hen t he right ( left) w all following is selected, and if *q*=1 the goal seeking behaviour is activated else the wall following is activated.

Now, w e group the t hree distances i n a t riplet ( dl,df,dr). Then, the fuzzy rule base is expressed as:

*R1:If (NNN) Then* $v_1$ *is ZR and* $\Delta\Phi_1$ *is a*PB+(1-p)*NB*

*R2:If (NNF) Then* $v_2$ *is ZR and* $\Delta\Phi_2$ *is NB*

*R3:If (NFN) Then* $v_3$ *is* $C*V_{\max}$ *and* $\Delta\Phi_3$ *is ZR*

*R4:If (NFF) Then* $v_4$ *is* $V_{\max}$ *and* $\Delta\Phi_4$ *is a* NS+(1-p)*PS*

*R5:If (FNN) Then* $v_5$ *is ZR and* $\Delta\Phi_5$ *is PB*

*R6:If (FNF) Then* $v_6$ *is ZR and* $\Delta\Phi_6$ *is a*PB+(1-p)*NB*

*R7:If (FFN) Then* $v_7$ *is* $V_{\max}$ *and* $\Delta\Phi_7$ *is a*NS+(1-p)*PS*

*R8:If (FFF) Then* $v_8$ *is* $V_{\max}$ *and* $\Delta\Phi_8$ *is* $q*\theta+(1-q)*(p*NM+(1-p)*PM)$

where $V_{\max}$ is the maximum velocity of the robot, $\Delta\Phi_i$ and $v_i$ are the fuzzy rule conclusions, and $C$ is the speed decrease coefficient. In the simulation part we take $V_{\max}=1$ m/s and C=0.1.
The rules conclusions are singletons expressed linguistically by: PB (*Positive Big*), : PB (*Negative Medium*), PS (*Positive Small*) , ZR ( *Zero*), NS (*Negative Small*), NM ( *Negative Medium*), and NB (*Negative Big*).

Here, because of the geometrical symmetry of the robot, we take: NB=-PB, NM=-PM, NS=-PS. Hence, the number of parameters is reduced and the rule base becomes:

*R1:If (NNN) Then* $v_1$ *is ZR and* $\Delta\Phi_1$ *is (2*p-1)*PB*

*R2:If (NNF) Then* $v_2$ *is ZR and* $\Delta\Phi_2$ *is NB*

*R3:If (NFN) Then* $v_3$ *is* $C*V_{\max}$ *and* $\Delta\Phi_3$ *is ZR*

*R4:If (NFF) Then* $v_4$ *is* $V_{\max}$ *and* $\Delta\Phi_4$ *is (2*p-1)*NS*

*R5:If (FNN) Then* $v_5$ *is ZR and* $\Delta\Phi_5$ *is PB*

*R6:If (FNF) Then* $v_6$ *is ZR and* $\Delta\Phi_6$ *is (2*p-1)*NB*

*R7:If (FFN) Then* $v_7$ *is* $V_{\max}$ *and* $\Delta\Phi_7$ *is (2*p-1)*NS*

*R8:If (FFF) Then* $v_8$ *is* $V_{\max}$ *and* $\Delta\Phi_8$ *is* $q*\theta+(1-q)*(2*p-1)*NM$

### C. Inference

In order to determine control actions; the steering angle $\Delta\Phi$ and the robot linear speed $v$, we use the Sugeno fuzzy inference method [14].

$$\Delta\Phi = \frac{\sum \alpha_i . \Delta\Phi_i}{\sum \alpha_i}$$

$$v = \frac{\sum \alpha_i . v_i}{\sum \alpha_i}$$

where $\alpha_i$ is the truth value of the $i$th rule.

### D. Fuzzy reinforcement Learning

In reinforcement learning, or "learning with a critic", the received signal is a behaviour punishment (positive, negative or neutral). This signal indicates what you have to do without saying how to do it. The agent uses this signal to determine a policy permitting to reach a long-term objective. It exist several reinforcement learning algorithms. Some are based on policy iteration such as *Actor Critic Learning* and others on value iterations as *Q-learning* or *Sarsa*. In the present work, we are interested by the use of the fuzzy version of the Q-learning algorithms named Fuzzy Q-Learning FQL algorihm presented on Tab. 1. The basic theory of the algorithm can be found in [13]. Here, several competing conclusions are associated with each fuzzy rule. To each conclusion is associated a q-value that is incrementally updated. The learning process consists then in determining the best set of rules, the one that will optimize the future reinforcements. The initial rule base is composed therefore of N rules such as:

**If** x is $S_i$ **Then** $y=c[i,1]$ ***with*** $q[i,1]=0$
**or** $y=c[i,2]$ ***with*** $q[i,2]=0$
...
**or** $y=c[i,J]$ ***with*** $q[i,J]=0$

To simplify the use of knowledge, we adopt the method in [15]. We fix a interval for each linguistic conclusion $c(\text{j,j})\in[a(\text{i}), b(\text{i})]$ and the number $J$ of the potential candidate conclusions is evenly distributed on this interval.

Tab. 1. Algorithm of Fuzzy Q-Learning

1. $t$=0, observe the state $x_t$.
2. For each rule, $i$, compute $\alpha_i(x_t)$.
3. For each rule, $i$, choose k[i] with an exploration /exploitation policy
4. Compute the action $A(x_t)$ and its correspondence quality $Q(x_t,A(x_t))$.
5. Apply the action $A(x_t)$. Observe the new state, $x_{t+1}$.
6. Receive the reinforcement, $r_t$.
7. Compute $\alpha_i(x_{t+1})$.
8. Compute a new evaluation of state value.
9. Update parameters $q[i,j]$ using this evaluation
10. t←t+1, go to 3

## IV. SIMULATION RESULTS

The problem consists of equip the robot with the capability of obstacles avoidance and goal seeking without being stuck in local minima and without collusion with obstacles. For this purpose we use the mobile robot model presented in section II, we assume that the effective range of the ultrasonic sensors is 10 cm - 250 cm. We use the proposed fuzzy navigator presented in section III, and to learn the robot, we use the fuzzy Q-Learning algorithm presented in the subsection III-D.

### A. Initialisation

The number of potential candidate conclusions is J=5. And the possibility interval of each fuzzy rule conclusion is given by Tab. 2.

Tab. 2. Fuzzy rule conclusions for $\Delta\Phi$

| | NS | NM | NB | ZR | PS | PB |
|---|---|---|---|---|---|---|
| **MIN** | -40° | -40° | -80° | 0° | 5° | 80° |
| **MAX** | -5° | -5 | -40° | 0° | 40° | 40° |

As a reinforcement signal, we want to punish any conclusion of an activated rule witch can cause collusion with an obstacle:

$$r = \begin{cases} -1 & if\ \min(dl, df, dr) < d_m \\ 0 & otherwise \end{cases}$$

***B. Learning phase***

In the learning phase (Fig. 4), only the wall following behaviour is selected (b=0), the robot is kept in unknown environment and after a sufficient learning time, 600 steps in the simulation, the reinforcement learning algorithm will be stopped.

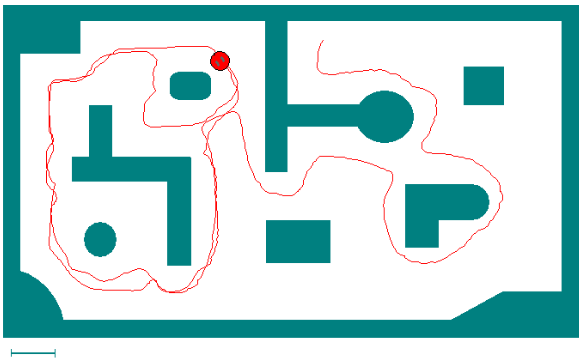


Fig. 4. Learning phase

***D. Goal Seeking***

After the learning stage, the rule conclusions with the best qualities are chosen by the ε-greedy policy [13]. Fig. 5 shows the simulation results of a goal seeking task when the robot starts from different initial positions.

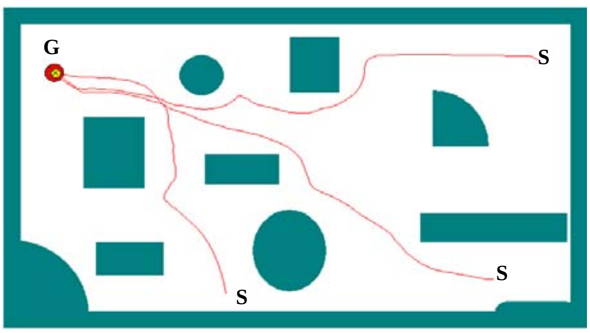


Fig. 5. Goal seeking

***C. Comparison with the related methods***

Fig. 6 shows the simulation results when one of the rule bases of [10, 11, 12] is used; the robot is suck in a local minimum in relatively complex environment. Fig. 7 shows simulation results with the proposed navigation method; the robot can easily escape from the local minima.

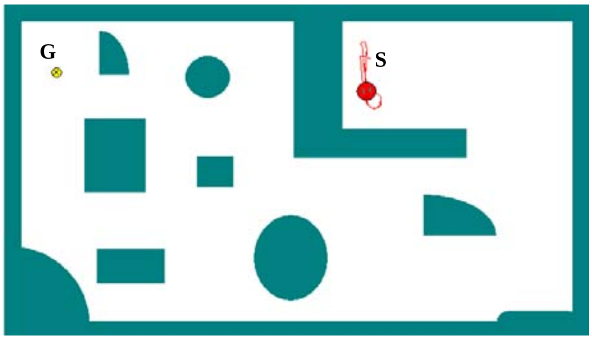


Fig.6. Local minima with some previous proposed fuzzy rule base

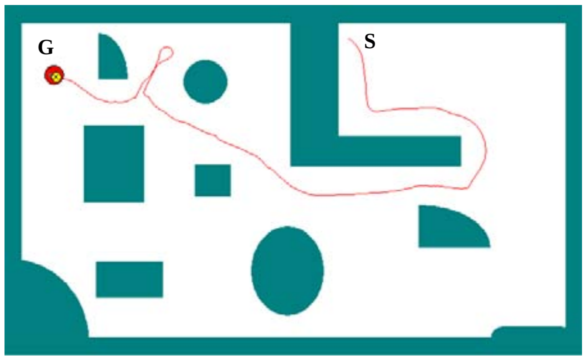


Fig.7. Goal seeking in a complex environment

## V. CONCLUSION

In this paper, we have proposed a solution to the local minima problem in goal seeking navigation for mobile robots in unknown environments. The proposed navigator is based on fuzzy logic. The rule base of the navigator is inspired from human reasoning and the parameters fine tuning is get by the Fuzzy Q-Learning algorithm. This method allows the robot the obstacle avoidance and goal seeking without being stuck in local minima. The merits of the proposed method are its simplicity and its easy implementation for industrial applications. The efficiency of the proposed method is demonstrated trough simulation results.

## REFERENCES

[1] Maaref H., and Barret C.," Sensor Based Navigation of an Autonomous Mobile Robot in an Indoor Environment", Control Engineering Practice, Vol. 8, pp. 757-768, 2000.

[2] Latombe J-C., *Robot Motion Planning*, Kluwer Academic Publishers, 1991.

[3] Koren Y. and Borenstein J.,"Potential Field Methods and Their Inherent Limitations for Mobile Robot Navigation", Proc. of the IEEE International Conference on Robotics and Automation, Sacramento, California, April 1991.

[4] Tan G.-Z., He H., and Sloman A.,"Ant Colony System Algorithm for Real-Time Globally Optimal Path Planning of Mobile Robots", ACTA AUTOMATICA, Vol. 33, No. 3, March 2007.

[5] Liu G., Li T., Peng Y. and Hou X. "The ant algorithm for solving robot path planning problem", Proc. of the IEEE/ICTTA'05, 2005.

[6] Brooks R. A." A robust Layered Control System for a mobile robot", IEEE Trans. On Robotics and Automation, Vol. RA-2, No. 1, pp. 14-23, Mar. 1986.

[7] Suwimonteerabuth D. and Chongstvatana P. " Online robot learning by reward and punichment for a mobile robot, Proc. of the IEEE/RSJ

international conference on Intelligent Robots and Systems, Lausanne, Swizerland, Oct. 2002.

[8] Beom H. B., and Cho H. S.,"A Sensor Based Navigation for a Mobile Robot using Fuzzy Logic and Reinforcement Learning", IEEE Trans. On Systems, Man, and Cybernetics, Vol. 25, No. 3, Mar. 1995.

[9] Ye Cang, and Nelson H.C. Yung, "A Fuzzy Controller with Supervised Learning Assisted Reinforcement Learning Algorithm for Obstacle Ovoidance", IEEE Trans. On Systems, Man, and Cybernetics, Vol. 33, No. 1, Feb. 2003.

[10] Zavlangas P. G. and Tzafestas S. G., "Motion control for Mobile Robot Obstacle Avoidance and Navigation: A Fuzzy Logic-Based Approach", Systems Analysis Modelling Simulation", Vol. 43, No. 12, Dec. 2003, pp. 1625-1637

[11] Dahmani Y. and Benyettou A., "Fuzzy Reinforcement Rectilinear Trajectory Learning", Journal of Applied Science, Vol. 4, No. 3, pp.388-392, 2004.

[12] Kermiche S., Saidi M. L., and Abbassi H. A.," Gradient Descent Adjusting Takagi-Sugeno Controller for a Navigation of Robot Manipulator", Journal of Engineering and Applied Science, Vol. 1, No 1, pp.24-29, 2006.

[13] Glorennec P. Y. and Jouffe L.," A Reinforcement Learning Method For Autonomous Robot", Proceedings of Eufit'96, Fourth European Congress on Intelligent Techniques and Soft Computing, Aachen, Germany, Sept. 1996.

[14] Yager R. R., and Filev D. P., *Essential of fuzzy modelling and control*, John Wily & Sons inc, 1994.

[15] Boubertakh H. and Glorennec P. Y., "Optimization of a Fuzzy PI Controller Using Reinforcement Learning", Proc. of the IEEE/ITTA'06, 2006.