Integrating induced knowledge in an expert fuzzy-based system for intelligent motion analysis on ground robots

José M. Alonso[1], Luis Magdalena[1], Serge Guillaume[2], Miguel A. Sotelo[3], Luis M. Bergasa[3], Manuel Ocaña[3]
[1] ETSI Telecomunicación, UPM, {jmam,llayos}@mat.upm.es, Ciudad Universitaria, 28040 Madrid, Spain
[2] Cemagref Montpellier, serge.guillaume@montpellier.cemagref.fr, BP 5095, 34033 Montpellier Cedex, France
[3] Universidad de Alcalá, {michael,bergasa,mocana}@depeca.uah.es, Department of Electronics.
Escuela Politécnica Superior. Campus Universitario s/n. Alcalá de Henares, Spain

Abstract

This work describes the knowledge extraction process for building a fuzzy system using expert and induced knowledge, applied to the detection of motion problems in ground robots. The expert knowledge was used for describing the robot behaviour in order to identify the variables that should be used with the aim of detecting a collision of the vehicle against an undetected obstacle, as well as proposing a suitable recovery action. Data collected in real trials were used for inducing knowledge so as to complete and validate the expert knowledge. Both kind of knowledge were integrated and used in a cooperative way in the final fuzzy-based system, which is interpretable and accurate at the same time.

Keywords: integration, expert and induced knowledge, linguistic knowledge base, interpretability, accuracy, diagnosis, ground robots.

1 Introduction

Interpretability is one of the most powerful features of the Fuzzy Inference Systems (FIS). The fuzzy logic formalism is well known for its linguistic concept modeling ability. The fuzzy rule expression is close to expert natural language. On the other hand, as they are universal approximators, FIS can be used for knowledge induction processes. The problem under consideration in this paper requires the integration of both, expert and induced knowledge, since none of these sources of information seems to offer a complete view of the problem. The cooperation between expert knowledge and induced knowledge let us achieve more accurate systems, but their integration must be done carefully for keeping the interpretability. Ensuring a good trade-off between accuracy and interpretability is one of the most difficult things in the fuzzy modeling [1].

The cooperation framework was proposed in [2] and it is implemented in KBCT[1], an open source software for generating or refining fuzzy knowledge bases. It is a dynamic user-friendly tool with the aim of reducing the effort of knowledge extraction process in generation or refinement of fuzzy-based linguistic knowledge bases. It was designed and developed in the framework of the European research project ADVOCATE II[2] [3]. The purpose of this architecture is to enhance the level of reliability and efficiency of autonomous robotic systems. In order to do that, it adds intelligence into control software of the system integrating different artificial intelligence techniques (Neuro-Symbolic Systems, Bayesian Belief Networks, and Fuzzy Logic). As a result, as a part of this project, a fuzzy logic based intelligent module was developed in order to solve real diagnosis or recovery specific problems working with uncertainty.

The present paper describes one of the application problems of the ADVOCATE II project: diagnosis of collision with undetected obstacles of a ground robot, with the aim to provide recovery actions. The diagnosis module is made up of a fuzzy system that integrates expert knowledge and data sample information, in an interpretable and accurate knowledge base.

[1] http://www.mat.upm.es/projects/advocate/kbct.htm
[2] http://advocate2.e-motive.com

The paper is structured as follows. Section 2 offers a perspective of the problem under consideration. Then section 3 describes the expert knowledge. The induced knowledge is considered in section 4. And the simplification of the knowledge base is explained in section 5. Finally, section 6 offers some conclusions.

2 Problem analysis

This problem was presented in the last EUSFLAT conference [4]. Vehicle dynamics and system behaviour upon obstacle collision were described in detail in [5]. Also, a preliminary knowledge base for the same problem was described in [6]. In short, the goal is to build an interpretable fuzzy knowledge base, using expert and induced knowledge in a collaborative way, in order to detect motion problems due to non visible obstacles using the sensorial capabilities on board the robot. Figure 1 shows the robot and the obstacle used in the trials.

Figure 1: The robot and the obstacle.

The following sections will describe how to build a fuzzy knowledge base for this problem. The overall knowledge extraction process was proposed in [2, 7]. To sum up, the process consists of three different steps: defining a common universe for each of the variables according to both expert knowledge and data distribution (section 3), then inducing rules from data (section 4), and finally integrating the induced rules into the expert knowledge base (section 5).

3 Expert knowledge

The expert is invited to describe the behaviour of the main influential variables. In a first step, expert knowledge extraction process was kept at a "high" abstraction level. The expert defined linguistic variables, the range of them according to data distribution, and the number of linguistic terms needed for each given variable without defining the corresponding fuzzy sets. Table 1 shows the main parameters of input variables defined by the expert. UnderShoot Width and Depth, and Decrease Of Battery Voltage have 7 labels because at the beginning the expert didn't know how to model their behaviour. Range Sonar and its derivative, Derivative Of Range Sonar, have 2 labels because they provide information concerning robot is moving or not with respect to its environment. Finally, Measured and Commanded Linear Velocities have 6 labels because their changes are in intervals of 50.

Table 1: Input Variables.

Variable	Range	Labels	Linguistic terms	Units
UnderShoot Width	[1, 5]	7	null, very small, small, medium, medium large, large, very large	ms
UnderShoot Depth	[10, 70]	7	null, very small, small, medium, medium large, large, very large	%
Decrease Of Battery Voltage	[0, 0.8]	7	zero, very small, small, medium, medium large, large, very large	v
Range Sonar	[20, 3000]	2	zero, NOT(zero)	mm
Derivative Of Range Sonar	[0, 1500]	2	zero, NOT(zero)	
Commanded Linear Velocity	[0, 250]	6	zero, very low, low, medium, high, very high	mm/s
Measured Linear Velocity	[0, 250]	6	zero, very low, low, medium, high, very high	mm/s

For each input variable, a strong fuzzy partition [8] with a number of fuzzy sets equal to the number of linguistic terms given by the expert was built. This type of partitioning ensures each fuzzy set can be attached a linguistic label and satisfy semantic constraints [9] on membership functions in order to respect semantic integrity within the partitions.

The expert also defined the set of diagnoses to be provided:

- False Alarm: it means that there is no real collision.
- Vehicle Drags Obstacle: the vehicle has collided against an obstacle not heavy enough to block vehicle movement. Thus, after a transient interval the vehicle controller (adaptive) regains the commanded velocity and keeps on moving by dragging the obstacle on its way.
- Vehicle Stalled: in this case the obstacle is heavy enough to impede the vehicle from moving. The vehicle is stopped as a result of the collision, it gets trapped by the obstacle.

According to previous information, one categorical output with three linguistic terms, one for each possible diagnosis, was defined.

Finally, once defined the characteristics of all variables, a common universe for each of them is achieved. The final semantic agreement is given by the expert and as a result the fuzzy set centers correspond to possible prototypes of the corresponding labels. Then, the expert explained the global behaviour of the system through two expert linguistic rules. These rules are of form "If condition Then conclusion" where both, the premise and the conclusion, use the linguistic terms previously defined.

1. **IF** *UnderShoot Width* is *small OR medium* **AND** *UnderShoot Depth* is *very small OR small OR medium OR medium large OR large* **AND** *Decrease Of Battery Voltage* is *small* **THEN** *Vehicle Drags Obstacle*

2. **IF** *UnderShoot Depth* is *very large* **AND** *Range Sonar* is *NOT(zero)* **AND** *Derivative Of Range Sonar* is *zero* **AND** *Commanded Linear Velocity* is *NOT(zero)* **AND** *Measured Linear Velocity* is *zero* **THEN** *Vehicle Stalled*

4 Induced knowledge

In order to complete and validate the expert knowledge, some real experiments were performed so as to collect data concerning variables defined by the expert. The original data set contains 167 instances, and it has been randomly divided into 2 different subsets (each one with 84 and 83 examples), maintaining the ratio of each diagnosis in both data sets. The number of cases for each possible diagnosis are collected in next table.

Table 2: Data Set.

Data Set	Dragging	Stalled	False Alarm
Learning	45	17	22
Test	44	17	22

Figure 2 shows data collected in one of the tests. According to this figure, the key for detecting a collision is the peak in the measured linear velocity. The width and depth of such peaks are modeled using two input variables, UnderShoot Width and UnderShoot Depth.

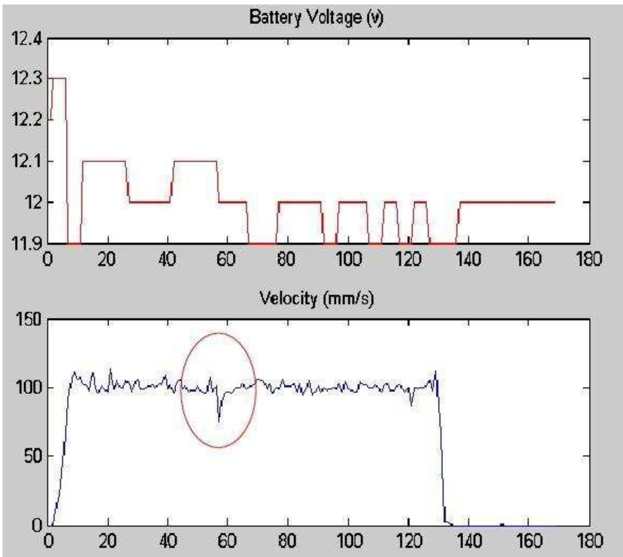

Figure 2: Vehicle Battery Voltage and Linear Velocity during a collision-and-drag case.

4.1 Knowledge Base Quality

The criteria for evaluating knowledge base quality are explained in [10]. It is measured according to the data set through considering next indexes for the output variable:

- **Coverage**: Percentage of examples from data that fires at least one rule with a degree higher than Δ.

- **No classified cases**: Number of cases from data set that don't fire at least one rule with a degree higher than Δ.

- **Error cases**: Number of covered cases from data set that produces error, i.e. observed and inferred values are different, in inference.

- **Ambiguity cases**: Number of covered cases from data set that produces ambiguity, i.e. difference between two output diagnoses is smaller than an established threshold, in inference.

These indexes convey complementary information. A good knowledge base should maximize the coverage and minimize the error indexes. Note that error and ambiguity cases are measured in relation to the covered examples.

4.2 Rule induction

The knowledge base quality indexes are shown in table 3. They are not good as there are a lot of misclassified cases. The coverage over learning and test data sets are equal to 39 % and 30 %, respectively. Most cases of stalling are correctly classified, but only a little cases of dragging are well detected. Obviously, all cases of false alarm are misclassified because the expert doesn't know how to describe these cases.

Table 3: Expert rules quality.

Data Set	Error Cases	Ambiguity Cases	No Classified Cases	Diagnosis
Learning	2	0	28	Dragging
Test	0	0	35	Dragging
Learning	0	0	1	Stalled
Test	0	0	0	Stalled
Learning	1	0	21	False alarm
Test	0	0	22	False alarm

In order to improve the knowledge base quality, some rules can be induced from data. KBCT let us choose between three methods (*Wang and Mendel* (WM) [11], *Fast Prototyping Algorithm* (FPA) [12], and *Fuzzy Decision Trees* (FDT) [13]) which make rule induction with previously defined partitioning. Let us underline that our contribution don't rely on such algorithm development but in their use within an expert data cooperation framework. Please refer to the cited literature for a complete description.

Table 4: Induced rules quality over learning data set.

Method	Induced Rules	Error Cases	Ambiguity Cases	No Classified Cases	Coverage (%)
WM	70	1	0	0	100
expert + WM	72	1	2	0	100
FPA	68	4	0	39	53
expert + FPA	70	6	0	22	72
FDT	43	1	0	0	100
expert + FDT	45	1	1	0	100
FDT + P	26	1	0	0	100
expert + FDT + P	28	1	1	0	100

Table 5: Induced rules quality over test data set.

Method	Induced Rules	Error Cases	Ambiguity Cases	No Classified Cases	Coverage (%)
WM	70	0	0	38	48
expert + WM	72	0	0	27	62
FPA	68	0	0	50	36
expert + FPA	70	0	0	33	55
FDT	43	2	0	0	100
expert + FDT	45	2	0	0	100
FDT + P	26	2	0	0	100
expert + FDT + P	28	2	0	0	100

Tables 4 and 5 give the main results. Four cases were studied: induced rules with WM, induced rules with FPA, induced rules with FDT, and induced rules with pruned FDT (FDT + P). When induced rules give a low coverage, the expert rules can help to increase it. However, if coverage is equal to 100 %, expert rules could produce ambiguity in some cases. The last case was selected due to, as it can be seen in these tables, it produces the best quality, the more interpretable and accurate knowledge base.

5 Integration Process

Induced rules with FDT + P were integrated into the expert knowledge base. As a result the rule base consists of 28 rules, 2 expert rules and 26 induced ones.

During this last step, the fundamental properties of the rule base have to be guaranteed: consistency, lack of redundancy and interpretability. Both kinds of rules use the same linguistic labels thanks to the previously defined common universe. Therefore rule comparison is made at linguistic level only.

First of all, a consistency analysis [7] of the knowledge base is made in order to detect conflicts at the linguistic level. Afterwards a simplification process is applied with the goal of achieving a more compact knowledge base, with a smaller size so that improving interpretability [14], but without getting worse accuracy of the original knowledge base.

The simplification process is described in detail in [10]. This work describes its application results in the real problem under consideration. Each iteration in the simplification process comprises of two steps:

1. **Data Base Reduction**:
   - Look for variables which are used by none of the rules and propose to remove them.
   - Look for labels which are used by none of the rules and propose to remove them.
   - Look for adjacent labels which are always used together and propose to merge them into a new one.
2. **Rule Base Simplification**:
   - Remove redundant rules:
     - Rules with the same premise and the same conclusion.
     - The input space covered by one rule is included into the one covered by the other, and both rules have the same conclusion.
   - Merge rules.

Table 6: Quality over learning data set.

KB	Error Cases	Ambiguity Cases	No Classified Cases	Coverage (%)
original	1	1	0	100
simplified	1	0	0	100

Table 7: Quality over test data set.

KB	Error Cases	Ambiguity Cases	No Classified Cases	Coverage (%)
original	2	0	0	100
simplified	2	0	0	100

Tables 6, 7 and 8 give the main results. The final knowledge base is more compact, with a smaller number of rules which are incomplete and more general, and a smaller number of labels. Only three error cases, corresponding to confusion between Vehicle Drags Obstacle and False Alarm, were detected over the whole, learning and test, data set. Hence, the simplification process improves interpretability without getting worse accuracy. Definitions of input variables after simplification are shown in table 9.

After removing or merging labels, fuzzy partitions still are strong fuzzy partitions. In order to maintain this structure property, the adjacent fuzzy sets are expanded. This makes the control surface of the fuzzy inference system smoother. Figure 3 shows UnderShoot Depth partition before and after simplification.

The final rule base is made up of thirteen rules, the two expert rules and eleven new induced ones, which are easily interpretable. The simplification process builds more general rules, as expert rules usually are, and it makes the system more robust and more interpretable. According to these rules, two variables (UnderShoot Width and UnderShoot Depth) are crucial

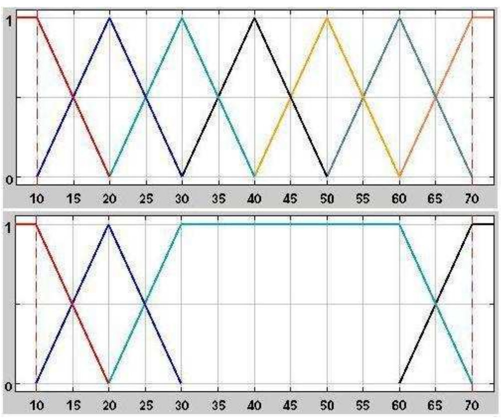

Figure 3: UnderShoot Depth and its simplification.

Table 8: Simplification results.

KB	Rules	Labels
original	28	37
simplified	13	23

Table 9: Simplified Input Variables.

Variable	Range	Labels	Linguistic terms
UnderShoot Width	[1, 5]	6	very small, small, medium, medium large, large, very large
UnderShoot Depth	[10, 70]	4	null, very small, small OR medium OR medium large OR large, very large
Decrease Of Battery Voltage	[0, 0.8]	4	zero, very small, small, medium OR medium large
Range Sonar	[20, 3000]	2	zero, NOT(zero)
Derivative Of Range Sonar	[0, 1500]	2	zero, NOT(zero)
Commanded Linear Velocity	[0, 250]	3	zero, very low, low OR medium OR high OR very high
Measured Linear Velocity	[0, 250]	2	zero, NOT(zero)

for determining whether the vehicle really collided with an obstacle that is being dragged, or on the contrary, whether the undershoot is due to measurement noise and it is a false alarm. This fact confirms, as it can be seen in figure 2, that the key for detecting a collision is the peak in the measured velocity. As an example, here are three induced rules, one for each possible diagnosis.

1. **IF** *UnderShoot Width* is *medium* **AND** *UnderShoot Depth* is *very small* **AND** *Decrease Of Battery Voltage* is *medium OR medium large* **AND** *Commanded Linear Velocity* is *very low* **THEN** *False Alarm*

2. **IF** *UnderShoot Width* is *small OR medium* **AND** *UnderShoot Depth* is *very large* **THEN** *Vehicle Drags Obstacle*

3. **IF** *UnderShoot Width* is *medium large OR large OR very large* **AND** *UnderShoot Depth* is *very large* **THEN** *Vehicle Stalled*

6 Conclusions

The two kinds of knowledge, expert knowledge and data, convey complementary information. Fuzzy logic, and fuzzy inference systems, are likely to offer a common framework. However, the cooperation of expert knowledge and data in system design remains an open problem, especially when the goal is to get a system which is both accurate and interpretable.

This work follows the approach presented in [7] to build fuzzy inference systems through using both, expert and induced knowledge, by focusing in the interpretability. Also it includes the application of a new simplification process of linguistic knowledge bases [10]. The results are encouraging. They show that the followed approach is appropriate. The final knowledge base is accurate and highly interpretable as it was desired.

The final system designed, using this knowledge base, in the framework of ADVOCATE II is able to provide diagnoses, with a very low error rate, as well as recovery actions upon circumstances of collisions with non visible obstacles provoking motion problems. These results are of prime importance for autonomous operation of ground robots in real environments.

Acknowledgments

We would like to acknowledge the partners of the european research project ADVOCATE II. The consortium consists of the following partners: Universidad de Alcalá, Getronics Decan, Universidad Politécnica de Madrid, ATL Elektronics, Ifremer, Hugin Expert A/S, Innova S.p.A and E-motive. The European Commission (IST-2001-34508) supports this work.

References

[1] J. Casillas, O. Cordon, F. Herrera, and L. Magdalena. “Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: an overview”. In [14].

[2] S. Guillaume and L. Magdalena. “Expert guided integration of induced knowledge into a fuzzy knowledge base”. *Soft Computing*. In Press, 2005.

[3] M. A. Sotelo et al. “ADVOCATE II: Advanced on-board diagnosis and control of autonomous systemsII”. *In Computer Aided Systems Theory*. Vol. 2809 of LNCS, pp. 302-313, Springer-Verlag, 2004.

[4] S. Guillaume et al. “Expert and induced knowledge for intelligent motion analysis of ground robots”. In *EUSFLAT’03*, Zittau, Germany, 2003.

[5] Luis M. Bergasa et al. “Knowledge-based intelligent motion analysis of ground robots”. In *ISORA’04*. June 28 - July 2, 2004, Sevilla, Spain.

[6] L. Magdalena, Miguel A. Sotelo, S. Guillaume, Luis M. Bergasa, and José M. Alonso. “Fuzzy-based system for intelligent diagnosis of motion problems on ground robots”. In *IPMU’04*. Information Processing and Management of Uncertainty in Knowledge-Based Systems. July, 4-9, 2004, Perugia, Italy. Proceedings, Vol. 1, pp. 151-158.

[7] José M. Alonso, L. Magdalena, and S. Guillaume. “KBCT: A knowledge extraction and representation tool for fuzzy logic based systems”. *FUZZ-IEEE 2004*. July, 25-29, 2004, Budapest, Hungary. IEEE Internacional Conference on Fuzzy Systems. Proceedings, Vol. 2, pp. 989-994.

[8] E. H. Ruspini. “A new approach to clustering”. *Information and Control*. Vol.15, pp. 22-32. 1969.

[9] J. Valente de Oliveira. “Semantic constraints for membership function optimization”. *IEEE Transactions on Systems, Man and Cybernetics*. Part A, Systems and Humans, Vol. 29, pp. 128-138, January 1999.

[10] José M. Alonso, L. Magdalena and S. Guillaume. “A simplification process of linguistic knowledge bases”. *EUSFLAT 2005*. September, 7-9, 2005, Barcelona, Spain.

[11] L.-X. Wang and J. M. Mendel. “Generating fuzzy rules by learning from examples”. In *IEEE Transactions on Systems, Man and Cybernetics*. Vol. 22 (6), pp. 1414-1427, 1992.

[12] P.-Y. Glorennec, “Algorithmes d′apprentissage pour systèmes d′inférence floue”. *Editions Hermès*. Paris, 1999.

[13] H. Ichihashi, T. Shirai, K. Nagasaka, and T. Miyoshi. “Neuro-fuzzy ID3: A method of inducing fuzzy decision trees with linear programming for maximizing entropy and an algebraic method for incremental learning”. *Fuzzy Sets and Systems*. Vol. 81, pp. 157-167. 1996.

[14] J. Casillas, O. Cordon, F. Herrera, and L. Magdalena, editors. “Interpretability issues in fuzzy modeling”. *Studies in Fuzziness and Soft Computing*, Vol. 128, Springer-Verlag, Heidelberg, 2003.