A Study on the Interpretability of a Fuzzy System to Control an Inverted Pendulum

Bahadir Zeren*, Muhammet Deveci*, Simon Coupland†, Robert John*, Ender Özcan*
ASAP, School of Computer Science, University of Nottingham, Nottingham, UK*
Institute of Artificial Intelligence, De Montfort University, Leicester, UK†
Email: {bahadir.zeren, muhammet.deveci, robert.john, ender.ozcan}@nottingham.ac.uk*, simonc@dmu.ac.uk†

Abstract—Fuzzy systems mimic human reasoning and provide solutions to problems under uncertainty via ‘computing with words’. This particular strength of fuzzy systems is often discarded in some real world applications where the fuzzy sets are designed for control problems or created through training using historical data. This study explores the interpretability of fuzzy systems by generating ‘meaningful’ fuzzy sets using a dictionary constructed by humans and fuzzy transfer learning. The inverted pendulum control problem is used as a case study. The empirical results show that interpretability of a fuzzy system is achievable even for this problem at the expense of a ‘slightly’ reduced performance.

Index Terms—Interpretability, Fuzzy Transfer Learning

I. INTRODUCTION AND MOTIVATION

The motivation behind this research is that the fuzzy logic community often make claims that because the user/expert can interrogate both the fuzzy rules and fuzzy sets and are therefore in some way ‘meaningful’ or ‘interpretable’. For the fuzzy sets by interpretable we mean they have meaning. If a set ‘low’ is used, for example, we would expect it to appear towards the left hand side of the domain. If there are no fuzzy sets below low then one would expect the set to be either a shoulder set or (if triangular) the right hand side of the triangle. The supposed interpretability of fuzzy systems is one of the unique selling points over artificial neural networks that are essentially black box. However, in most real world applications, this is not true:

- Where the parameters of the system are trained using historical data (through ANFIS, Artificial Neural Networks for example), the final trained fuzzy sets are not usually sensible.
- In control applications the sets (whatever shape) have no meaning but are there to provide a mapping from inputs to outputs with the control engineer only interested in performance of the control system.

Zadeh introduced the notion that fuzzy logic can deliver Computing with Words [1] and many researchers have used this terminology in their work, however, we posit that they are not ‘doing’ Computing with Words, but jumping on the bandwagon. This is initial work and we decided to see if we can investigate this through a control application and selected the inverted pendulum control problem.

The inverted pendulum problem on a cart is a classic control system problem and one of the most important problem in dynamics [2], [3]. It is an under actuated, high order, multi-variable, and unstable non-linear system, which is widely used as a benchmark problem in control theory. Therefore, it is an ideal model for testing various theories in the field of control engineering [4]–[7]. A flexible and intuitive way of expressing indefinite responses in system of many different stable controllers can be provided by fuzzy logic controller [7].

Some researchers on the inverted pendulum have been studied in recent years based on various methods including the state-space controller [8]–[10], the proportional integral derivative controller [11]–[13], the artificial neural network control theory [14]–[17], the sliding-mode control theory [15], [18]–[20], linear quadratic regular theory [13], [21], [22] and fuzzy logic controller [7], [23]–[26].

Fuzzy logic control that incorporates ambiguous human logic into computer programs [26] is based on fuzzy sets and fuzzy inference [27]. The fuzzy sets proposed by Zadeh [28] as a mathematical tool for useful modeling various types of uncertainty that include vague, ambiguous and imprecise data [29]. Fuzzy sets have been successfully utilised in many real-world applications, such as, controller design [30], [31], signal processing [32], decision making [33], medicine [34], risk analysis [35], electronic control system [36], [37], process planning [38], and so on.

The use of fuzzy systems for control problems is not new. The stabilisation of an inverted-pendulum is a complex problem requiring a nonlinear model [39]. The use of a fuzzy logic controller for a variant of the problem to stabilise an inverted pendulum on a cart is discussed in [4], [40], [41].

There has been a growing number of studies on the control of inverted pendulum by using fuzzy logic control. Ma et al. [42] addressed the analysis and design of the fuzzy controller and the fuzzy observer on the basis of the Takagi-Sugeno (T-S) fuzzy model. The main contribution of the study is to improve the separation: fuzzy controller and fuzzy observer can be designed independently. Wang et al. [43] attempted a design methodology for stabilisation of non-linear system for fuzzy control utilising Takagi-Sugeno fuzzy model. Magana and Holzapfel [44] proposed an experimental setup of a fuzzy logic controller for an inverted pendulum problem using vision feedback. Ochoa et al. [45] investigated the effect of performance tuning the absolute error value of the fuzzy system to control the position of a pendulum. They used seven membership
functions for the position and three membership functions for the speed. El-Bardini and Al-Nagar [46] studied a type-2 fuzzy logic controller for inverted pendulum and then compared it to a type-1 controller. The results showed that the performance of the type-2 controller was significantly improved upon the other one. Jang [47] examined temporal back propagation based self-learning fuzzy controllers. Lee and Takagi [48] built an automatic fuzzy system design method that used a genetic algorithm for the classical inverted pendulum problem. Su et al. [49] presented event-triggered fuzzy control design for the inverted pendulum system. They showed the advantages and effectiveness of the proposed design using three simulation experiments.

In this study, an ‘interpretable’ fuzzy controller for the inverted pendulum problem is proposed. The next section describes the inverted pendulum problem. Sections III and IV present how meaning is assigned to the fuzzy sets based on the concept of ‘computing with dictionary of words’. The results are provided in Section V. Finally, Section VI provides the concluding remarks.

II. THE INVERTED PENDULUM PROBLEM

The inverted pendulum is a typical nonlinear control problem that is of importance in a range of areas, such as robotics, rockets and industrial processes [49]. The schematic diagram of inverted pendulum system is shown in Fig. 1, which consists of a cart, straight line, driving unit and pendulum [4], [7], [41], [50]. The cart can move freely to the right or left. It is assumed that there is no friction in the system [41]. It is clear that the pendulum in the upright position has an unstable balance leading to a non-linear control problem for its stabilisation. The notation for the pendulum system is given in Table I.

![Fig. 1: The schematic diagram of inverted pendulum system on a cart.](image)

The closed-loop fuzzy control system for stabilising the inverted pendulum on a cart is also shown in Fig. 1.

The non-linear dynamic equations of the inverted pendulum system are based on the physical laws [49]. Eqs. 1 and 2 are the differential equations used to simulate system response to controllers for the inverted pendulum.

\[
\ddot{x} = \frac{pg \times \cos \Theta \times \sin \Theta + p \times l \times \dot{\Theta}^2 \times \sin \Theta + F}{M + p \times (1 + \sin \Theta^2)}
\]

**TABLE I: The notations and descriptions of the inverted pendulum system.**

| Notation | Description |
|----------|-------------|
| M        | Mass of the cart |
| p        | Mass of the pendulum |
| l        | Length of the pendulum |
| g        | Acceleration of gravity |
| F        | Force on the cart |
| \(\Theta\) | Angle from the upright position |
| \(\dot{\Theta}\) | Angular velocity of the pendulum |

![Fig. 2: The inverted pendulum closed-loop fuzzy control system.](image)

III. THE DICTIONARY

In this paper we are interested in the interpretability of a typical engineering focused fuzzy system. To do this, we shall investigate whether fuzzy sets in this application can be replaced with fuzzy sets obtained from surveys in a general context. In doing this, we are bringing together two previously separate strands of research computing with words dictionaries [51]–[53] and fuzzy transfer learning [54]. In this section we consider the construction of a dictionary of words from interval data which can be used in to replaced the more synthetic sets in the inverted pendulum application.

In previous work [51], [52] a large scale web based survey was undertaken to obtain interval data about where people feel a set of word should placed on a somewhat arbitrary scale of \([0, 10]\). In this paper we shall use this data to construct a type-1 fuzzy set with a Gaussian membership function for each word in the survey data. The steps used to obtain the parameters for the membership functions were:

1) Clean the data using the preprocessing step given in [51].
2) Construct a histogram (we used a bin count of 201) for each word based on the cleaned interval data.
3) Calculate the mean (c) and standard deviation (\(\sigma\)) of this histogram.

These steps are depicted in Figure 3.

This process produced a dictionary of 32 words, each modelled by a fuzzy set with a Gaussian membership function. Table II lists these 32 words with the parameters and Figure 4 depicts these fuzzy sets.
TABLE II: Parameters of the Fuzzy Set which make up the Dictionary.

| Word               | c    | σ    | Word               | c    | σ    |
|--------------------|------|------|--------------------|------|------|
| A bit              | 1.00 | 2.32 | Modest amount      | 4.18 | 1.82 |
| A lot              | 8.30 | 1.96 | None to very little| 0.08 | 2.19 |
| A smidgen          | 0.15 | 2.06 | Quite a bit        | 4.18 | 1.81 |
| Considerable amount| 7.23 | 1.98 | Sizeable           | 6.00 | 1.96 |
| Extreme amount     | 9.95 | 2.06 | Small              | 1.70 | 1.80 |
| Fair amount        | 5.50 | 1.63 | Some               | 3.31 | 1.94 |
| Good amount        | 0.47 | 1.74 | Some to moderate   | 4.70 | 1.30 |
| High amount        | 9.00 | 2.20 | Somewhat small     | 1.78 | 1.83 |
| Huge amount        | 9.85 | 2.24 | Substantial amount | 7.63 | 2.01 |
| Humongous amount   | 9.95 | 2.39 | Teeny-weeny        | 0.15 | 3.08 |
| Large              | 8.35 | 1.50 | Tiny               | 0.15 | 2.01 |
| Little             | 0.85 | 2.31 | Very high amount   | 9.90 | 1.74 |
| Low amount         | 1.20 | 1.93 | Very large         | 9.95 | 2.43 |
| Maximum amount     | 9.90 | 0.91 | Very little        | 0.15 | 1.80 |
| Medium             | 5.03 | 1.10 | Very sizeable      | 8.13 | 2.29 |
| Moderate amount    | 5.13 | 1.38 | Very small         | 0.20 | 1.66 |

IV. HOW THE FUZZY SETS WERE COMPARED

In order to identify which fuzzy sets from the dictionary best match the sets from the inverted pendulum application we first transfer the dictionary to each of the three domains in the inverted pendulum system. We then identify which sets best match the exist sets.

A. Transferring the Dictionary

Fuzzy transfer learning [54] suggests how fuzzy sets from one context can be transferred to another. In the simplest terms this mean translating and scaling the parameters of membership function from one domain to another. In this paper, we only use Gaussian membership functions. As such the transfer of our dictionary to any other domain becomes a problem of translating and scaling values of \( c \) and scaling values of \( \sigma \). Translating the centre of our Gaussian sets from one domain to another is trivial. We calculate the distance between the centre of each domain and move all values of \( c \) by that distance. The domain of the dictionary is \([0, 10]\) which has a centre of 5. The first input domain in the inverted pendulum application is \([-40, 40]\) which has a centre of 0. As such we must translate all values of \( c \) in the dictionary fuzzy sets by \(-5\). Next the values of \( c \) and \( \sigma \) must be scaled. To do this we calculate a scaling factor as the width of the new domain over the width of the old domain. Again looking as the first input domain of the inverted pendulum application \((-40, 40)\) we see it has width of 80 whilst the dictionary \((0, 10)\) has a width of 10. This gives 80/10 so we have a scaling factor of 8. Every value of \( c \) and \( \sigma \) must now be multiplied by this scaling factor. Consider the example set from the dictionary \( c = 5.03 \) and \( \sigma = 1.10 \). In order to transfer these parameters to the domain \([-40, 40]\) we first translate the value of \( c \) by \(-5\) giving 0.03. We then scale both \( c \) and \( \sigma \) by 8 giving a values of \( c = 0.24 \) and \( \sigma = 8.80 \). Applying the same procedure to Maximum amount \( (c = 9.90, \sigma = 0.91) \) yields values of \( c = 39.20 \) and \( \sigma = 7.82 \). The calculation of the new values of \( c \) and \( \sigma \) is given formally in equations 3 and 4.

\[
c' = c + \frac{(d^o_e - d^o_s)}{\sigma^o} \times \frac{d^o_e - d^o_s}{d^o_e - d^o_s}
\]

(3)

where \( c' \) is the value of \( c \) after the transfer, \( d^o_e \) and \( d^o_s \) are the start and end of the new (destination) domain and \( d^o_e \) and \( d^o_s \) are the start and end of original (source) domain.

\[
\sigma' = \sigma \times \frac{d^o_e - d^o_s}{d^o_e - d^o_s}
\]

(4)

where \( \sigma' \) is the value of \( \sigma \) after the transfer.

This process yields a collection fuzzy sets transferred from the dictionary with its somewhat arbitrary domain of \([0, 10]\) to the three domains in the invert pendulum application. When considering the values obtained from this process, it becomes clear whilst the values of \( c \) have transferred across perfectly sensibly, there is something of an issue with the values of \( \sigma \). The values of \( \sigma \) given in the transferred dictionaries are significantly higher than those in the inverted pendulum system. Whilst we leave a discussion of the reasons for this to Section VI we now propose an additional step in the transfer
process which we believe will make the transferred sets more usable. We proposed normalising the width of the transferred fuzzy sets. The method we propose for doing is simple. For each transferred domain calculate the mean $\sigma$ value of the sets from the inverted pendulum system. Scale the transferred fuzzy values of $\sigma$ so that the mean values of $\sigma$ match across the original and transferred sets. This is trivial to do we simply need to obtain a scaling factor for the transferred values of $\sigma$ for each domain which gives the same value for mean $\sigma$. For the dictionary calculated here the respective scale factors for the domains is 0.26, 0.37 and 0.20. Applying these scaling factors clear reduces the width of the membership functions. however we believe it is a sensible step. If we consider the area under a membership function to in some way represent the uncertainty captured by that set it makes sense to have sets being used on the same system which capture the same levels of uncertainty.

B. Selecting Sets from the Dictionary

Our goal is to replace each set in the existing inverted pendulum application with a set from our two respective dictionaries, with and without scaled values of $\sigma$. To do this, we simply compute the similarity of each set in the inverted pendulum system, scale the transferred fuzzy values of $\sigma$ so that the mean values of $\sigma$ match across the original and transferred sets. This is trivial to do we simply need to obtain a scaling factor for the transferred values of $\sigma$ for each domain which gives the same value for mean $\sigma$. For the dictionary calculated here the respective scale factors for the domains is 0.26, 0.37 and 0.20. Applying these scaling factors clear reduces the width of the membership functions. however we believe it is a sensible step. If we consider the area under a membership function to in some way represent the uncertainty captured by that set it makes sense to have sets being used on the same system which capture the same levels of uncertainty.

III gives a complete list of all membership functions used in this paper.

\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - y_i^*)^2} \] (6)

where $n$ is the number of samples (taken at each time interval), $y_i$ is $F$, $\Theta$, or $x$ for the $i^{th}$ sample, and $y_i^* = 0$.

Although they are both interpretable fuzzy sets, the performance of $SN$ is better than $S$ in terms of RMSE for the controller output and system responses of $F$, $\Theta$ and $x$. Moreover, $SN$ employs the least force to control the inverted pendulum generating an RMSE of 0.7601 when compared to $T$, $G$ and $S$.

VI. Conclusion

Interpretability is one of the main advantages of using a fuzzy system. However, that advantage disappears in some applications, for example, while solving control problems using a fuzzy system, since synthetic fuzzy sets are used. In this initial study, we presented an approach to make a fuzzy system more interpretable by making the fuzzy sets more ‘meaningful' based on a dictionary and fuzzy transfer learning. We applied our approach to the well-known problem of inverted pendulum obtaining encouraging results. We will investigate the performance of the proposed approach on other real-world problems and look into the trade-off between interpretability and performance.

The results from this study do present some issues when attempting to use fuzzy sets in the form of words elicited from humans in engineering applications.

Engineering application often require (at least output fuzzy sets) to have some exact central datum, in the case of the inverted pendulum a value of zero. Fuzzy sets elicited from
### TABLE III: Membership Functions of all Fuzzy Sets used in this Paper.

| Label  | $a$   | $m$   | $e$  | $c$  | $\sigma$ | Word            | $J_s$ | $c$  | $\sigma$ | Word            | $J_s$ |
|--------|-------|-------|------|------|----------|----------------|------|------|----------|----------------|------|
| NVB    | -40   | -40   | -25  | -40  | 5        | Very Small      | 0.35 | -39.4| 4.61     | None to very little | 0.96 |
| NB     | -40   | -25   | -10  | -25  | 5        | Very little     | 0.46 | -25.8| 3.855    | Somewhat small   | 0.76 |
| N      | -2    | -10   | 0    | -10  | 3.33     | Some to moderate| 0.25 | -13.5| 4.08     | Some             | 0.46 |
| ZO     | -5    | 0     | 5    | 0    | 1.66     | Medium          | 0.19 | 0.2  | 2.31     | Medium          | 0.72 |
| P      | 0     | 10    | 20   | 10   | 3.33     | Moderate amount | 0.25 | 11.73| 3.67     | Good amount      | 0.67 |
| PB     | 10    | 25    | 40   | 25   | 5        | Large           | 0.48 | 25.00| 4.83     | Very sizeable    | 0.97 |
| PVB    | 25    | 40    | 40   | 40   | 5        | Maximum amount  | 0.63 | 39.60| 5.03     | Humongous amount | 0.93 |

Fig. 5: Controller output: (a) Force, system response: (b) $\Theta$ and (c) position ($x$).

The sets at either end of the domain have lower values of $\sigma$ than those in the centre. It is not surprising that humans find it easier to agree about what maximum and minimum are compared to medium. We do need to consider this fact and how we might address it when using such sets in an engineering context.

The sets in the dictionary are placed on what is essentially an arbitrary axis of $[0, 10]$. The level of uncertainty in these sets is unlikely to be suitable for a given engineering application. In this paper we solved this problem by matching the level of uncertainty to an existing application. Future work should consider how to achieve this when no such application exists.

### VII. Acknowledgements

We would like to acknowledge the work of Jerry Mendel and his students in the area of interval fuzzy sets which informed, and gave labels to the dictionaries used in this paper.

### REFERENCES

1. L. A. Zadeh, “Fuzzy logic = computing with words,” *IEEE Transactions on Fuzzy Systems*, vol. 4, pp. 103–111, May 1996.
2. M. I. El-Hawwary, A.-L. Elshafei, H. M. Emara, and H. A. A. Fattah, “Adaptive fuzzy control of the inverted pendulum problem,” *IEEE Transactions on Control Systems Technology*, vol. 14, no. 6, pp. 1135–1144, 2006.
3. E. J. Minnaert, B. Hemmelman, and D. Dolan, *Inverted pendulum design with hardware fuzzy logic controller*. PhD thesis, South Dakota School of Mines and Technology, Rapid City, 2009.
4. J. Yi and N. Yubazaki, “Stabilization fuzzy control of inverted pendulum systems,” *Artificial Intelligence in Engineering*, vol. 14, no. 2, pp. 153–163, 2000.
5. L. Hong-ju and F. Jian, “An inverted pendulum fuzzy controller design and simulation,” in *Computer, Consumer and Control (IS3C), 2014 International Symposium on*, pp. 557–559, IEEE, 2014.
6. W. Chen and N. Theodomile, “Simulation of a triple inverted pendulum based on fuzzy control,” *World Journal of Engineering and Technology*, vol. 61, pp. 31–47, 2017.
7. A. I. Roose, S. Yahya, and H. Al-Rizzo, “Fuzzy-logic control of an inverted pendulum on a cart,” *Computers & Electrical Engineering*, vol. 46, pp. 263–269, 1992.
