

Fuzzy-probabilistic logic for common sense

YKY (King-Yin Yan)

Abstract. P(Z) logic offers a new way to reason about vagueness (ie fuzziness), that treats fuzziness as degrees, distinct from probabilities. One then applies probability distributions over fuzziness. This approach is different from both classical fuzzy logic [26] and possibility theory [1]. P(Z) logic is specially designed for common-sense reasoning.

1 Main idea

Vagueness is pervasive in common-sense reasoning. A **calculus of degrees** allows common-sense statements to be rendered into formal logic and be reasoned about computationally, thus fulfilling a need in logic-based AGI (artificial general intelligence) systems [5]. It is widely believed that a general-purpose uncertainty calculus should somehow combine the ideas in probability theory and those in fuzzy logic; The current approach is not groundbreaking, but rather an intuitive, simple and practical solution based on well-established probability theory over continuous degrees. We acknowledge that it has not been empirically tested.

A more detailed exposition of our AGI theory is [25]. The logic described here deals with propositions, assigning truth values consisting of 2 components: Z = fuzzy, P = probabilistic. We will also refer to B = binary logic, which is subsumed by P(Z).

1.1 Prior research

The theory presented here is a synthesis of old ideas that have been expressed in the literature on fuzzy logic [26] [14] [15] and vagueness, in particular the so-called degree theory [20] [2] [8] [6] [21] [22].

1.2 What is fuzziness?

The central idea is to treat vagueness as degrees, which are distinct from probabilities. A fuzzy value, $z \in [0, 1]$ represents a degree of something. The degree z itself is not really “fuzzy” and is unnatural for use in common sense reasoning. For example, it would be ridiculous for an exact fuzzy value of 0.7 to imply that Mary’s height is *exactly* 1.7m and

not 1.70000001m. To better capture the essence of fuzziness, it makes intuitive sense to distribute probabilities over degrees.¹

Once we conceptualize vagueness in this manner, the calculus for reasoning with vagueness follows rather routinely.

Any physical quantity (eg age, height) in $[0, \infty)$ can be translated into a degree $\in [0, 1]$ via sigmoidal transformations such as:

$$Z(x) = e^{-\ln 2 \cdot (x/\xi)^2} \tag{1}$$

where ξ is the *point of neutrality* for that specific quantity, ie the point at which $z = 0.5$ (for example Fig 1 is the case for *young*). Note that the ratio (x/ξ) makes z *dimensionless*.

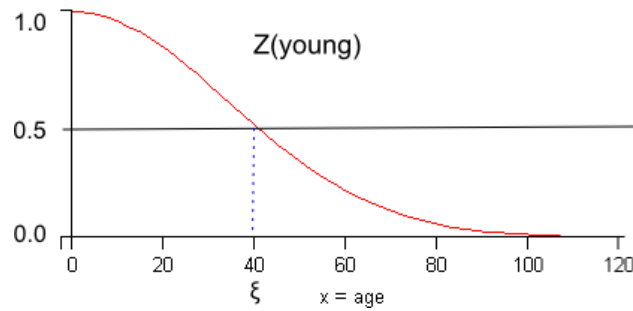


Fig. 1. neutral point

Reference classes: A tall building and a tall person should not be compared by the same scale. For each reference class there would be a characteristic ξ , but it is the job of the logic to decide which ξ to use, whereas this paper focuses on the calculus of propositional truth-values.

The common-sense interpretation of numerical Z values is shown in Table 1.

2 P: probabilistic logic

A variety of probabilistic logics have been proposed, but the most popular approach seems to be “Bayesian logic” (eg [4] [7]).²

¹ Note that P(Z) is different from Fuzzy Random Variables (FRV) [11]. P(Z) values are probability distributions over $[0,1]$, ie, simple random variables; whereas FRVs are probability distributions over fuzzy sets which are themselves membership functions from some arbitrary domains to $[0,1]$. In other words, FRVs explicitly manage transformations such as Eqn (1), whereas our approach simply distributes probabilities over $[0,1]$ and thus is easier to use.

Another alternative is to use interval fuzzy values, but their inference mechanisms tend to be more complicated (see eg [9]), and the fuzzy intervals tend to diverge to $[0,1]$ very quickly during inference.

² This view is explained in the “AIMA” textbook [19], 2nd ed, in §14.6. Several other ways to lift probabilistic reasoning to first-order settings are briefly described in the textbook [12] §9.3.7.

Table 1.

z	interpretation
1.0	definitely or extremely
0.9	very
0.7-0.8	moderately
0.6	slightly
0.5	neutral
0.4	slightly not
0.3-0.2	moderately not
0.1	very not
0.0	definitely not

In an AGI, when we describe some probabilistic relations, we get a *network* of probabilistic conditionals. For instance, we may say:

- *If a burglary occurs without an earthquake, the alarm will sound with 0.94 chance*
- *If the alarm sounds, Mary will call with 0.9 chance*
- etc etc..

then voila, we already have a Bayesian network (BN). Such BNs can be generated on-the-fly upon each user query. This technique is known as KBMC (knowledge-based model construction) [24] and has become standard for lifting propositional probabilistic logics to first-order. Once we have the BN we can use **belief propagation** to find the truth value of the propositional variable we want.

2.1 Conditionals

At the heart of logical reasoning is the implication operator, often called the “arrow”. In Bayesian networks, nodes represent random variables and links represent **probabilistic conditionals** of the form $P(x|y)$. Probabilistic conditionals correspond to **implications** ($x \leftarrow y$) in classical logic³. P(Z) logic is the extension of Bayesian logic where random variables can take continuous values in [0,1].

Referring to the commutative diagram (Fig 2), which shows the relation between the 4 types of truth values (B, P(B), Z, and P(Z)) and their logical operators. When going from B to P, we replace classical implication ($y \rightarrow x$) with the Bayesian network link ($y \rightarrow_P x$). When going from P(B) to P(Z), we will also give up the classical connectives ($x \vee y$ and $x \wedge y$). The top-right corner requires us to invent a new set of fuzzy-probabilistic operators, a technical issue we won’t go into here.

They include, but not exhaustively: [18]’s Markov Logic Networks which is based on Markov random fields instead of Bayesian networks; and Loopy Logic which is based on [16]’s belief propagation algorithm; [3]’s Probabilistic Relational Models; and [10]’s Bayesian Logic Programs. Relatedly, [17] developed a Stochastic Lambda Calculus, and [13] provides a way to perform probabilistic reasoning within classical higher-order logic.

³ This correspondence is not exact. Indeed, the relation between implications and conditionals is controversial and has resulted in variants of probabilistic logics. Our view is that classical implication should be replaced by probabilistic conditionals.

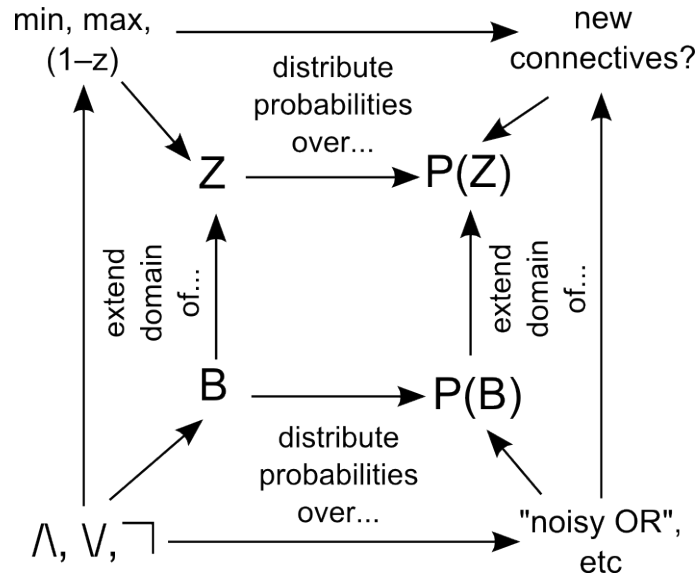


Fig. 2. The inner square shows the 4 types of truth values, the outer square their logical operations

3 P(Z): probability distributions over fuzziness

The logical operators in *pure Z* logic include $\wedge = \min$, $\vee = \max$, and $\neg = 1 - z$. It is a direct consequence of negation as $1 - z$ that 0.5 is neutral and 0.0 represents the *opposite* of a concept.

$P(Z)$ logic is the combination of Z with P . Each $P(Z)$ value is a probability distribution over $[0,1]$ (for example Fig 3). All distributions are restricted to be beta distributions, represented by their means and variances (μ, σ^2) . The beta distribution is chosen because it is the most well-studied continuous distribution over $[0, 1]$, and can be represented by just 2 parameters. With this choice we cannot represent multi-modal distributions, but it seems to be an adequate “first approximation”.

The beta distribution is very versatile and can represent common sense concepts with “fuzzy”, “unary”, and “binary” characters (Fig 4). For example, common sense says that a person is either dead or alive, but upon closer examination we may discover a continuous spectrum (Fig 5).

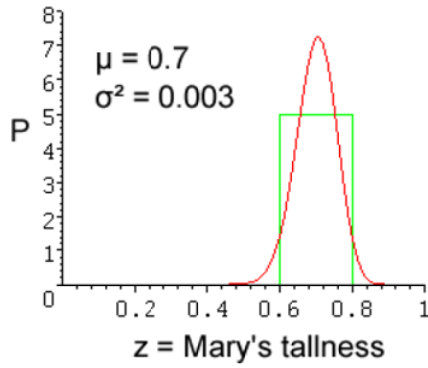


Fig. 3. an example $P(Z)$ distribution

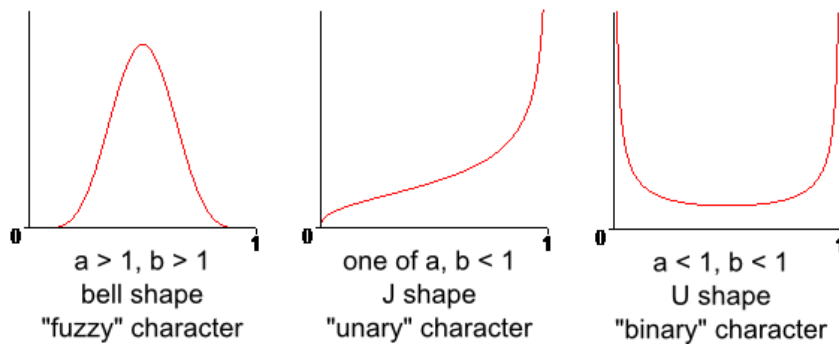


Fig. 4. shapes of beta distributions

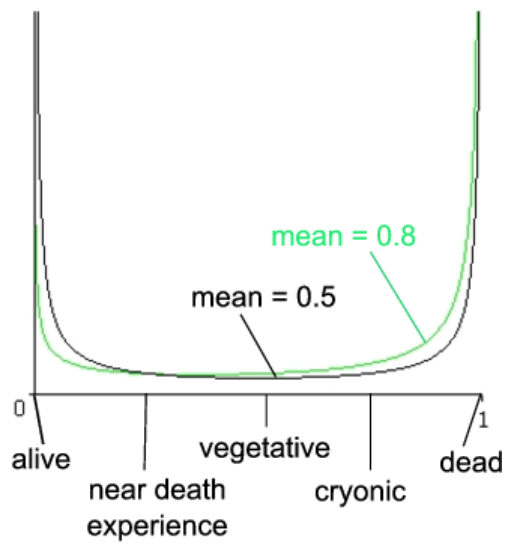


Fig. 5. a concept with binary character

“Java girl” paradox: For example, a boy may judge a girl’s desirability based on traits like “good looks”, “personality”, “intelligence”:

$$\text{desirable} \leftarrow \text{trait}_1 \wedge \text{trait}_2 \wedge \text{trait}_3 \wedge \text{etc...}$$

but imagine a girl who scores 0.9 in the top traits but is also 0.9 “good at Java”; she would be even more desirable, but \wedge in fuzzy logic would still give a value of 0.9, counter-intuitively. A solution could be based on taking into account the variance of the $P(Z)$ distributions, as shown in Fig 6, where we know with more specificity and higher overall certainty that girl A is more desirable.

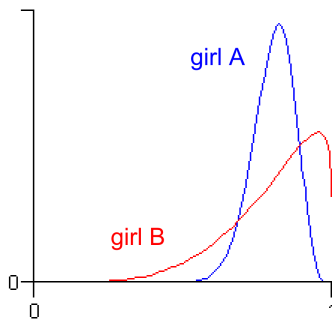


Fig. 6. Same mean, different variance

4 Fuzzy modifiers

Fuzzy modifiers are introduced to handle natural-language hedges such as “moderately”, “very”, “extremely”. They can be any smooth function $F : [0, 1] \rightarrow [0, 1]$, such as those in Figure 7.

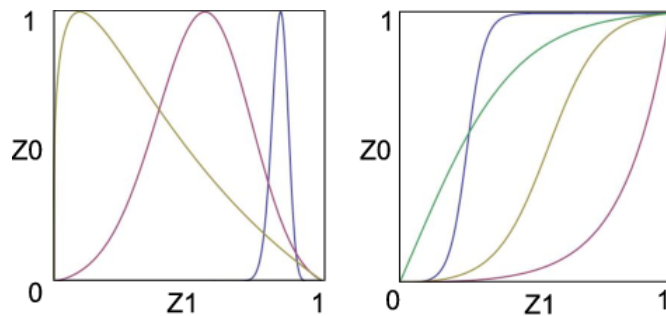


Fig. 7. fuzzy modifiers

5 Inference

All formulas in P(Z) logic has this general syntax:

$$H \leftarrow B_1, B_2, B_3, \dots$$

where H is the head and B_i is the body (which may be empty). Some P(Z) logic operators, with numerical parameters, will determine the CPT (conditional probability table) of $P(H|B_i)$.

As an example, the inference rule for $z_0 \leftarrow z_1 \vee z_2$ is:

$$\mu_0 \approx \max(\mu_1, \mu_2)$$

$$v_0 \approx \begin{cases} v_1 & : \mu_1 > \mu_2 \\ v_2 & : \mu_2 > \mu_1 \end{cases}$$

and the rule for \wedge is similar. The rule for negation ($z_0 \leftarrow \neg z_1$) is simply:

$$\mu_0 = 1 - \mu_1, \quad v_0 = v_1$$

These rules are obtained by considering the result z_0 , a random variable, as a function of other random variables (the operands). Some of the formulas are obtained by simple infinitesimal analysis, others by empirical computations.

Fig 8 shows the behavior of $\max(f_1, f_2)$ against f_1 and f_2 . The distribution of f_0 seems to “eat up” the probability masses of f_1 and f_2 , whichever comes from the right first, until it is “full” (reaches 1). Therefore the mean of f_0 is not much different from the max of the means of f_1 and f_2 , resulting in the approximate rule above. Note that there is in general a slight *right-shift* for max, this $\Delta\mu$ may be included in a more accurate version...

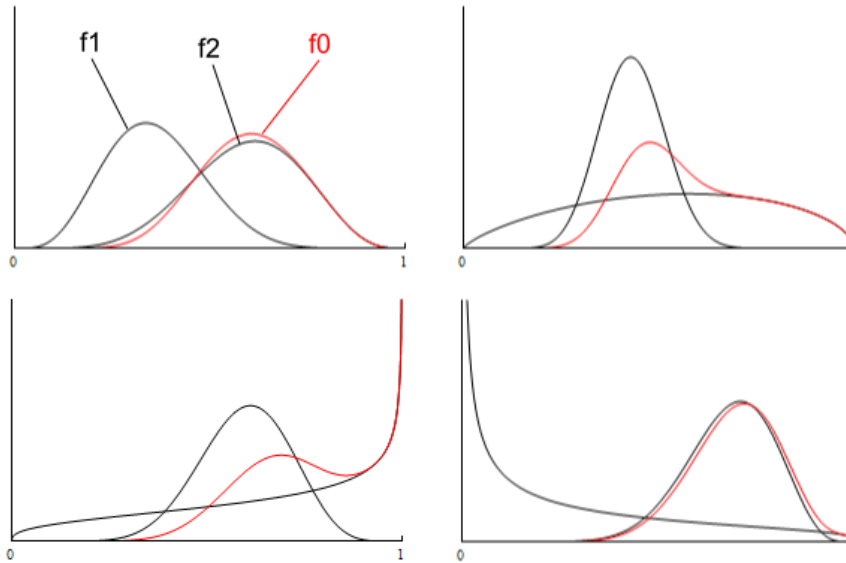


Fig. 8. example PDFs of $z_0 := z_1 \vee z_2$

Acknowledgments

Abram Demski and Anna Nachesa advised me on the probabilistic aspects. I am also indebted to Pei Wang [23] and Ben Goertzel [5] for their seminal ideas.

References

1. Dubois and Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.
2. Edgington. Validity, uncertainty and vagueness. *Analysis*, 52:193–204, 1992.
3. Getoor, Friedman, Koller, and Pfeffer. Learning probabilistic relational models. *Journal of relational data mining*, pages 307–335, 2001.
4. Getoor and Taskar, editors. *Introduction to statistical relational learning*. MIT Press, 2007.
5. Goertzel and Pennachin. *Artificial general intelligence*. Springer, 2007.
6. Graff and Williamson, editors. *Vagueness*. Ashgate, Dartmouth, 2002.
7. Heckerman, Meek, and Koller. *Statistical relational learning*, chapter 7: Probabilistic entity-relationship models, PRMs, and plate models. MIT, 2007.
8. Keefe. *Theories of Vagueness*. Cambridge university press, 2000.
9. Kenevan and Neapolitan. *Fuzzy logic for the management of uncertainty*, chapter A model theoretic approach to propositional fuzzy logic using Beth Tableaux, pages 141 – 157. John Wiley & Sons, Inc. New York, NY, USA, 1992.
10. Kersting and de Raedt. Bayesian logic programs. In Cussens & Frisch, editor, *Work-in-progress reports of the 10th international conference on inductive logic programming (ILP 2000)*. 2000.
11. Kruse and Meyer. *Statistics with vague data*. Reidel Publishing Company, 1987.
12. Luger. *Artificial intelligence: structures and strategies for complex problem solving*. Pearson Education, 2009.
13. Ng and Lloyd. Probabilistic reasoning in a classical logic. *Journal of applied logic*, 7-2:218–238, 2009.
14. Nguyen and Walker. *A first course in fuzzy logic (1st, 2nd, and 3rd edition)*. Chapman & Hall /CRC, 1997, 2000, 2006.
15. Novak, Perfilieva, and Mockor. *Mathematical principles of fuzzy logic*. Kluwer Academic Press, 1999.
16. Pearl. *Probabilistic Reasoning in Intelligent Systems - Networks of Plausible Inference*. Morgan Kaufmann, California, 1988.
17. Pless and Luger. Towards general analysis of recursive probabilistic models. *Proceedings of conference of uncertainty in artificial intelligence (San Francisco)*, 2001.
18. Richardson and Domingos. Markov logic networks. *Journal of machine learning*, 62:107–136, 2006.
19. Russell and Norvig. *Artificial Intelligence - a modern approach*. Prentice Hall, 2003.
20. Sainsbury. Degrees of belief and degrees of truth. *Philosophical Papers*, 15:97–106, 1986.
21. Shapiro. *Vagueness in Context*. Oxford university press, 2006.
22. Smith. *Vagueness and degrees of truth*. Oxford, 2008.
23. Wang. *Rigid Flexibility - The Logic of Intelligence*. Springer applied logic series, 2006.
24. Wellman, Breese, and Goldman. From knowledge bases to decision models. *Knowledge Engineering Review*, 7(1):35–53, 1992.
25. Yan. *Genifer - an artificial general intelligence*. <http://launchpad.net/agi-book> (currently electronic version only), 2010.
26. Zadeh. Fuzzy sets. *Information and control*, 8:338–353, 1965.