Degrees that are not probabilities

Kenneth Pernyér, Reflective Labs
Stockholm, Sweden
kenneth.pernyer@reflective.se · reflective.se · ORCID 0009-0006-4543-1156

Organizations reason in vague predicates. A customer is *large*, an exposure is *material*, a supplier is *strategic*, and none of these has a threshold anyone can defend, which is why the thresholds that end up in software are invented in meetings and then never revisited. This paper describes the fuzzy inference capability of our analytics extension: linguistic variables, membership functions, expert-authored rules, and three inference families — Mamdani with defuzzification, Sugeno, and Tsukamoto — running as Suggestors inside the governed loop of our substrate. Two things make it a paper rather than a library note. The first is a distinction the implementation enforces in the type system: a membership degree and a materiality degree are both numbers in $[0, 1]$ and are not the same quantity, and neither is a probability. Confusing them is the same category error this series has now met three times, and the remedy is the same — make the compiler refuse. The second is a constraint that is genuinely mathematical rather than stylistic: Tsukamoto inference requires monotone consequent membership functions, because the firing strength must be *inverted* to a crisp value, and our implementation returns an error rather than an approximation for the triangular, trapezoidal and Gaussian shapes where the inverse does not exist.

1. Introduction

Ask a credit committee what makes an exposure material and you will get a number, an apology for the number, and three exceptions. The number is not wrong exactly. It is a crisp boundary drawn across a phenomenon that has no crisp boundary, and the exceptions are what the boundary cost.

This is the situation fuzzy set theory was invented for (Zadeh, 1965). Rather than forcing a predicate to be true or false, a fuzzy set assigns a *degree of membership* in $[0, 1]$, and a calculus over those degrees lets a system reason with *large*, *material* and *strategic* without first pretending they are thresholds. The engineering tradition that followed — linguistic rules, an inference procedure, a defuzzification step (Mamdani and Assilian, 1975, Takagi and Sugeno, 1985) — produced controllers that were both effective and, importantly for us, *readable*: the rules are sentences a domain expert wrote and can still read afterwards.

We use it for the same reason. Our substrate promotes facts with provenance (Pernyér, 2025a), and the provenance of a fuzzy conclusion is the set of rules that fired and how strongly. That is an explanation in the sense an auditor means, not in the sense a feature-attribution plot means.

2. The pieces

A linguistic variable is a named quantity — exposure, tenure, concentration — together with fuzzy sets over its domain. Each set carries a membership function $\mu : \mathbb{R} \to [0, 1]$ from the usual shapes: triangular, trapezoidal, Gaussian, and monotone sigmoids.

A rule is an implication over these sets, with an optional weight:

$$\text{IF } x_1 \text{ is } A_1 \text{ AND } x_2 \text{ is } A_2 \text{ THEN } y \text{ is } B.$$

Inference proceeds in three steps. *Fuzzification* evaluates each antecedent term, giving membership degrees. *Aggregation* combines them into a firing strength α_r per rule — minimum for conjunction, maximum for disjunction, in the standard reading. *Defuzzification* turns the aggregate consequent back into a number the rest of the system can use.

We support four defuzzification methods — centroid, bisector, mean of maxima, and height — over a discretized domain with a configurable number of steps. The choice matters more than it is usually given credit for: centroid and mean of maxima can disagree substantially on multi-modal aggregates, which is precisely the case where the rule base is telling you the answer is genuinely ambiguous.

3. Two degrees, and neither is a probability

The implementation carries two distinct newtypes over $[0, 1]$, both clamped at construction:

> **Definition (Membership and materiality).** A **membership degree** $\mu_A(x) \in [0, 1]$ is the extent to which x belongs to fuzzy set A. A **materiality degree** is the extent to which a value bears on a decision — its salience or evidentiary weight. They are separate types and there is no conversion between them.

Why bother? Because they answer different questions and are routinely averaged together by systems that

store both as `f64`. "This exposure is 0.8 *large*" and "this exposure is 0.8 *decision-relevant*" are different claims: a small exposure to a sanctioned counterparty has low membership in *large* and high materiality, and a system that multiplies the two produces a number meaning nothing. Separating them at the type level makes the nonsense unwriteable rather than merely discouraged — the same move the analytics paper describes for promotion authority (Pernyér, 2025b), applied to arithmetic.

Neither is a probability, and the distinction is older and better documented than our use of it (Dubois and Prade, 1993). A probability is a degree of belief about which of several *crisp* alternatives holds, and it is constrained by additivity: the probabilities of an exhaustive partition sum to one. A membership degree is a degree to which a *vague* predicate applies to a definite object, and it obeys no such constraint. An exposure can be 0.7 *large* and 0.6 *medium* simultaneously; that is not an incoherent probability assignment, it is the ordinary situation of a quantity near a boundary between two overlapping linguistic terms.

The failure this prevents is concrete. Take a fuzzy risk score of 0.8 and a model's output probability of 0.8 and combine them — average, multiply, feed both to a threshold — and the result cannot be interpreted, because there is no operation under which *degree of membership* and *degree of belief* compose. This is the third instance in this series of the same mistake: certificates averaged with probabilities (Pernyér, 2025c), BM25 scores averaged with cosine similarities (Pernyér, 2025d), and now memberships averaged with likelihoods. The pattern is worth naming. **Quantities that share a range do not share a semantics**, and a system that stores them in the same primitive type will eventually combine them.

4. Tsukamoto and the invertibility constraint

The three inference families differ in what they do with the consequent, and one of them imposes a real constraint that our implementation enforces rather than approximates.

Mamdani (Mamdani and Assilian, 1975) clips or scales the consequent fuzzy set by the firing strength, aggregates the clipped sets, and defuzzifies the aggregate. It is the most interpretable and the most expensive, since it needs a discretized domain.

Sugeno (Takagi and Sugeno, 1985) makes the consequent a function of the inputs — typically constant or affine — and takes the firing-strength-weighted average. There is no defuzzification step, and the output is smooth in the inputs, which is why it dominates in control.

Tsukamoto requires each consequent membership function to be **monotone**, and takes the crisp consequent to be the value at which the consequent function attains the firing strength:

$$y_r = \mu_{B_r}^{-1}(\alpha_r), \quad y = \frac{\sum_r \alpha_r y_r}{\sum_r \alpha_r}. \tag{1}$$

Eq. 1 is well-defined only when μ_{B_r} is invertible, which for a membership function means monotone on its support.

> **Proposition (Invertibility).** A membership function admits a unique inverse on $[0, 1]$ if and only if it is strictly monotone on its support. The triangular, trapezoidal and Gaussian shapes are not: each attains its maximum on a set of positive measure or on both sides of a mode, so $\mu^{-1}(\alpha)$ is either a pair of values or an interval.

The implementation exposes `is_monotonic` on the membership function and returns an error from `inverse` for the non-monotone shapes. It does not pick the left branch, or the midpoint, or the mean of the two solutions.

That refusal is the design decision worth defending. Every one of those fallbacks produces a number, and a number that came from silently choosing a branch of a non-invertible function is indistinguishable, downstream, from a number that was computed correctly. In a governed system where the output becomes a proposal with provenance attached, "we approximated an inverse that does not exist" is not a provenance anyone can audit. Refusing means the rule base has to be authored with monotone consequents when Tsukamoto is the chosen family — a constraint on the modeller, stated up front, rather than a silent distortion at inference time.

5. Rules as explanations

A fuzzy conclusion arrives with the rules that produced it, each with its firing strength — the activated rule set. This is the property that makes fuzzy inference attractive in a governed context and it is worth being precise about what it does and does not give.

It gives a *faithful* account: the rules that fired, with weights, are the computation, not a reconstruction of it. There is no gap between the explanation and the mechanism of the kind that post-hoc attribution methods have to bridge, and the argument for preferring intrinsically interpretable models in consequential settings (Rudin, 2019) applies here with unusual force, because the model *is* a set of sentences the organization wrote.

It does not give correctness. A rule base can be faithfully reported and wrong, and the readability of the rules makes them easier to argue about — which is the actual benefit. The organization argues about the rules instead of arguing about the score.

6. Limitations

The rule base is authored, and authoring is the whole cost. Ten variables with five terms each is a combinatorially large rule space that no expert enumerates, so real rule bases are sparse and their coverage is uneven in ways nobody has measured. We do not currently report *how much* of the input space a rule base covers, or warn when an input falls in a region where no rule fires strongly.

Membership functions are chosen by hand. There is a well-developed literature on learning them from data — adaptive neuro-fuzzy systems (Jang, 1993) fit exactly these parameters by gradient descent — and that is deliberately not here: it is the business of the extension that fits models, for the reasons the previous paper gives (Pernyér, 2025b). A fuzzy system with learned parameters is no longer closed-form, and the substrate treats it accordingly.

Defuzzification method selection is left to the caller, with no guidance beyond this paper. Centroid is the sensible default and is wrong when the aggregate is strongly bimodal, which is the case a committee would most want flagged.

Finally, we have no comparative evaluation. We claim the outputs are explainable and the semantics are sound; we do not claim, and have not measured, that a fuzzy rule base predicts better than a fitted classifier on any of our domains. It very likely does not, and that is not what it is for.

7. Conclusion

Vagueness in organizations is not a defect to be thresholded away, and a formalism that represents it directly turns an argument about where to draw the line into an argument about the shape of a membership function — which is a better argument, because it is about the phenomenon rather than about the software.

Two things are worth carrying out of this paper. Quantities that share a range do not share a semantics: membership, materiality, confidence and probability all live in $[0, 1]$ and none of them compose with the others, so a system that cares should make the compiler say so. And when a computation requires an inverse that does not exist, the right output is an error. A plausible number with an impossible provenance is the worst artifact a governed system can produce, because everything downstream will treat it as ordinary.

> **Code and correspondence.** The work described here is implemented, not sketched. The source behind every claim in this paper — the engine, the extension, the fixtures and the tests that hold the invariants — can be shared on request, including the parts that did not work. Write to kenneth.pernyer@reflective.se. We are equally interested in being told where the argument is wrong.

References

Dubois, D., Prade, H., 1993. Fuzzy sets and probability: Misunderstandings, bridges and gaps. Proceedings of the Second IEEE International Conference on Fuzzy Systems 1059–1068.

Jang, J.-S.R., 1993. ANFIS: Adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics 23, 665–685.

Mamdani, E.H., Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies 7, 1–13.

Pernyér, K., 2025a. Proposal, gate, commitment: deterministic convergence as a substrate for organizational decisions. Reflective Labs technical report.

Pernyér, K., 2025b. Nothing to fit: closed-form analytics as governed Suggestors. Reflective Labs technical report.

Pernyér, K., 2025c. Confidence is a certificate: provable optimization inside a probabilistic loop. Reflective Labs technical report.

Pernyér, K., 2025d. Recall is a proposal: governed memory for multi-agent organizations. Reflective Labs technical report.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence 1, 206–215.

Takagi, T., Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man, and Cybernetics SMC-15, 116–132.

Zadeh, L.A., 1965. Fuzzy sets. Information and Control 8, 338–353.

A. The fuzzy surface, operationally

F1 Types `MembershipDegree` and `MaterialityDegree` are distinct clamped newtypes over $[0, 1]$. Construction clamps; there is no conversion.

F2 Membership functions Triangular, trapezoidal, Gaussian and monotone shapes, each with `evaluate`, `is_monotonic` and `inverse`. `inverse` returns an error for non-monotone shapes.

F3 Rules A rule carries an antecedent expression, a consequent, and a weight in $[0, 1]$. Inputs are validated before inference; an input naming an unknown variable fails rather than defaulting.

F4 Inference Mamdani with a discretized domain and a choice of centroid, bisector, mean-of-maxima or height defuzzification; Sugeno by weighted average; Tsukamoto by Eq. 1, which requires monotone consequents.

F5 Output The activated rule set travels with the result — every rule that fired, with its strength — and is carried into the proposal as provenance.

B. Why the three families differ

Let rules $r = 1...n$ have firing strengths α_r and consequent sets B_r over an output domain Y.

Mamdani. Each consequent is clipped, $\mu'_{B_r}(y) = \min\big(\alpha_r, \mu_{B_r}(y)\big)$, and the clipped sets are aggregated by maximum. The result is a fuzzy set over Y, and defuzzification reduces it — centroid takes the first moment of the discretized aggregate, height takes the maximizing point. The whole aggregate shape is available for inspection, which is why this family is the most interpretable and the most expensive.

Sugeno. Each consequent is a function f_r of the inputs, and the output is $\sum_r \alpha_r f_r(x) / \sum_r \alpha_r$ directly. No set is constructed over Y, so there is nothing to defuzzify and nothing to inspect beyond the weights.

Tsukamoto. Each consequent is a monotone set, and the crisp value is recovered by inverting it at the firing strength per Eq. 1. It sits between the other two: cheaper than Mamdani because no aggregate set is built, more constrained than Sugeno because the consequents must be invertible. The constraint is the price of getting a crisp value out of a set rather than out of a function.