

Bridging Data and Meaning: A Review of Semantic Information Theory with Fuzzy Linguistic and Possibilistic Techniques

Dr. Jayesh V. Karanjgaonkar

Assistant Professor (Mathematics)

Swami Atmanand Government Eng. Med. Model College Atari, Raipur (C.G.)

Abstract –

Semantic Information Theory (SIT) extends classical information theory by focusing on the meaningful content of communicated data rather than mere symbol transmission. This review paper bridges the foundational theories of semantic information—such as those by Bar-Hillel, Carnap, MacKay, and Floridi—with modern advances utilising fuzzy linguistic and possibilistic techniques. Classical approaches provide quantitative and logical frameworks but often struggle to handle vagueness, context, and meaning in natural language and human communication. Fuzzy set theory and possibility distributions, pioneered by Zadeh and further developed by contemporary researchers, offer robust mathematical tools to model semantic content under uncertainty and imprecision. These modern methods encode meaning as restrictions and graded membership functions, enabling more nuanced measurement and interpretation of semantic information that better aligns with human linguistic and cognitive processes. Integrating these perspectives, the paper highlights the theoretical evolution from syntactic to semantic measures, discusses practical benefits in AI and hybrid systems, and underscores ongoing challenges in formalising a comprehensive, operational semantic information framework. This synthesis provides a cohesive understanding of how data and meaning converge in semantic information theory, fostering progress in knowledge representation, decision making, and intelligent communication systems.

Keywords - Semantic Information Theory, Fuzzy Logic, Possibility Theory, Information Measurement, Zadeh's Restriction Principle, Information and Meaning, Knowledge Representation, Uncertainty Modelling.

I . Introduction -

Semantic Information Theory (SIT) marks a significant advancement beyond Claude Shannon's pioneering work in information theory, which primarily quantified “how much” information is transmitted through a communication channel without concern for its meaning. SIT addresses a more fundamental and practical question: “what” information means, focusing on the semantic content conveyed between sender and receiver. The importance of SIT arises from the recognition that effective communication transcends reliable symbol transmission to encompass understanding, interpretation, and meaningful knowledge exchange, particularly in advanced domains such as artificial intelligence, knowledge management systems, and linguistic computation (Mingers, 1996).

Weaver's classic tripartite classification of the communication problem sets an important framework for the study of information:

- **Technical problem:** The accurate and efficient transmission of symbols through a channel, the focus of Shannon's classical information theory.
- **Semantic problem:** The faithful conveyance of the intended meaning behind the transmitted symbols.
- **Pragmatic problem:** The impact of the received message on the behaviour or response of the receiver (Karanjgaonkar, 2023).

While Shannon's entropy successfully resolved the technical problem by providing a robust quantitative measure of information reliability, it deliberately left the semantic and pragmatic problems unaddressed. The semantic dimension is especially crucial today as the complexity and volume of data rise, and the need for machines to process and understand natural language and ambiguous inputs becomes imperative. SIT hence emerges as a vital theoretical and practical field concerned with modelling, measuring, and processing the meaning and context of information (Karanjgaonkar, 2023).

With the increase in interdisciplinary applications, ranging from natural language processing and expert systems to decision theory and cognitive science, SIT seeks to provide rigorous frameworks that incorporate uncertainty, vagueness, and meaning beyond raw data transmission. This evolution bridges the gap between data and knowledge, enabling more intelligent and human-centred information systems (Mingers, 1996; Karanjgaonkar, 2023).

II. Literature Review –

Semantic Information Theory (SIT) has evolved from classical probabilistic and logical frameworks towards more encompassing approaches that integrate fuzzy logic and possibility theory. This section reviews major theories, their mathematical foundations, strengths, and limitations.

Classical Theories of Semantic Information

Bar-Hillel and Carnap (1952) were pioneers in formulating semantic content mathematically. Their Classical Semantic Theory (CST) models the semantic information content of a statement A based on how many possible states of the world it excludes. The logical probability $m(A)$ measures the likelihood of A, and semantic content and information are defined as:

$$\text{Content}(A) = 1 - m(A) \quad \dots\dots\dots(1)$$

$$\text{Information}(A) = -\log_2 m(A) \quad \dots\dots\dots(2)$$

A tautology ($m(A) = 1$) excludes no states and thus has zero content, whereas a contradiction excludes all states, having maximum content. However, this theory suffers from the Bar-Hillel-Carnap paradox (BCP) where contradictions possess infinite information, which is counterintuitive (Karanjgaonkar, 2023).

Hintikka (1968) refined this by defining the information content $I(h)$ of a statement h as the reduction in the uncertainty of the receiver:

$$I(h) = -\log_2 p_h = -\log_2 \left(\frac{w_h}{2^K} \right) \dots\dots\dots(3)$$

where p_h is the logical probability, w_h the width of statement h , and K the number of constituent statements. This refines the information measure but was found limited in practical semantic analysis (Karanjgaonkar, 2023).

Dretske (1981) provided a philosophical framework linking semantic information to the receiver's knowledge k and conditional probabilities:

$$r \text{ carries information about } s \text{ is } F \Leftrightarrow P(s \in F \mid r, k) = 1 \text{ and } P(s \in F \mid k) < 1$$

Jumarie's Relative Information Theory (1990) considers communication as a two-stage process: syntactic entropy $H(a)$ and semantic entropy $H(a \mid l)$

$$H_{\text{total}} = H(a) + H(a \mid l) \dots\dots\dots(4)$$

where a is the source string of symbols and l is the meaning assigned by the observer. This subjectively weights semantic information but remains limited to formal domains.

Floridi's Strong Semantic Theory (2004) introduces the concept of semantic deviation v and computes informativeness as:

$$\text{Info}(o_i) = 1 - v^2 \dots\dots\dots(5)$$

for an information instance o_i , with $v \in [-1, 1]$ representing the degree of discrepancy from a reference situation w . Floridi quantifies semantic information by integrating this measure over possible states, but practical calculation challenges exist (Karanjgaonkar, 2023).

Fuzzy Logic and Possibilistic Approaches

To overcome the limitations of sharp logical probabilities, Zadeh's fuzzy set theory (Zadeh, 1965) introduces membership functions $\mu_F(x)$ that grades the degree to which an element x belongs to set F

$$\mu_F: X \rightarrow [0, 1] \dots\dots\dots(6)$$

and the notion of possibility distributions for a variable X :

$$\text{Poss}(X = u) = \mu_F(u) \dots\dots\dots(7)$$

Representing linguistic vagueness as fuzzy constraints, Zadeh conceptualises natural language statements as elastic restrictions:

$$X \text{ is } F \dots\dots\dots(8)$$

where F is a fuzzy set defining the possible values of X with graded membership (Karanjgaonkar, 2023).

Specificity—a measure analogous to entropy—is used to quantify how focused a fuzzy set is:

$$\text{Specificity}(A) = 1 - \frac{\sum_{i=1}^n (a_i - a_{i+1}) \log i}{\log n} \dots\dots\dots(9)$$

where a_i are ordered membership grades and n the number of elements (Yager, 1996). Dubois and Prade (1997) demonstrated that possibility and probability measures can be connected via transformations, enabling graded semantic uncertainty modelling while retaining meaningful mathematical structure (Karanjgaonkar, 2023)

Classical semantic information measures provide valuable formal insight but fail to handle real-world linguistic vagueness and context dependence. Fuzzy set theory and possibility distributions, grounded in Zadeh's approaches, bring flexible, graded, and context-sensitive modelling methods supporting advanced semantic information measurement, crucial for AI and hybrid systems.

The Dempster-Shafer Theory of Evidence (DST)

The Dempster-Shafer Theory of Evidence (DST), also known as evidence theory or belief functions theory, provides a powerful and flexible mathematical framework for reasoning under uncertainty when information is incomplete, imprecise, or conflicting. Unlike traditional probability theory, which assigns probabilities to mutually exclusive events, DST allows belief to be assigned to sets or subsets of possibilities, more naturally capturing epistemic uncertainty and ignorance (Shafer, 1976).

At its core, DST uses a mass function $m: 2^X \rightarrow [0,1]$, defined over the power set 2^X of the frame of discernment X , which assigns belief masses to subsets of X , called focal elements, representing the evidence supporting those subsets:

$$\sum_{A \subseteq X} m(A) = 1, \quad m(\emptyset) = 0 \quad \dots\dots\dots(10)$$

The belief function $\text{Bel}(A)$ is the total belief exactly supporting subset A and all its subsets:

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B) \quad \dots\dots\dots(11)$$

while the plausibility function $\text{Pl}(A)$ considers all subsets that intersect A , capturing how plausible A is given the evidence:

$$\text{Pl}(A) = 1 - \text{Bel}(A^c) = \sum_{B \cap A \neq \emptyset} m(B) \quad \dots\dots\dots(12)$$

These two functions provide lower (belief) and upper (plausibility) bounds on the probability of A , effectively sandwiching uncertainty between what is supported and what is not contradicted by evidence (Shafer, 1976).

A key innovation is Dempster's rule of combination, which provides a method for fusing independent pieces of evidence from multiple sources to update belief succinctly while managing conflicts. If two independent mass functions m_1 and m_2 are defined over the same frame, their combined mass m is:

$$m(A) = \frac{1}{1-K} \sum_{B \cap C = A} m_1(B)m_2(C), \quad K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C) \quad \dots\dots\dots(13)$$

where K measures the degree of conflict between sources (Sentz & Ferson, 2002).

DST is especially valuable in real-world situations where data sources provide partial, uncertain, or even conflicting information. It explicitly models ignorance and incomplete knowledge, allowing systems to make cautious yet informed decisions, updating beliefs as new evidence arrives. Applied across sensor

fusion, decision support, and AI reasoning, DST affords robustness beyond classical probabilistic models by supporting a flexible spectrum between full belief and complete ignorance (Taroun & Yang, 2019).

In summary, DST offers a mathematically rigorous, yet operationally practical, framework to represent, combine, and reason with multifaceted evidence in semantic information contexts, thus strongly advocating its inclusion in any comprehensive general theory of information and uncertainty.

III. Modern Methods: Fuzzy Linguistic and Possibilistic Measures

Fuzzy Linguistic Approach

The fuzzy linguistic approach recognises that much of human knowledge and communication is imprecise or graded, not binary or crisp. Fuzzy set theory, formalised by Zadeh in 1965, introduced the possibility of modelling vague concepts by assigning each element in a universe a degree of belonging, or membership, between 0 and 1 (Zadeh, 1965; Karanjgaonkar, 2023).

[0,1]

In this framework, linguistic terms—such as "tall," "very hot," or "quite expensive"—are directly mapped to fuzzy sets that capture their vagueness in mathematical form. Modern information modelling capitalises on this property by constructing membership functions (vertical, horizontal, similarity, or empirically based) for such linguistic variables, so that human statements can be processed and compared quantitatively. This approach supports the representation of meaning and context in information analysis, making it vital in expert systems, AI, and knowledge management.

Zadeh's Fuzzy Restriction Principle

Zadeh's restriction principle asserts that many natural language statements can be interpreted as elastic constraints ("restrictions") on variable values (Zadeh, 1978; Karanjgaonkar, 2023). For example, "The temperature is high" restricts the temperature variable to values within a fuzzy subset of possible values labelled 'high.' These restrictions are mathematically described by fuzzy sets—specifically, the statement "X is F" indicates a restriction on X such that its value belongs to fuzzy set F. A possibility distribution is then induced, where the possibility of $X = u$ is given by the membership $\mu_F(u)$.

$$\text{Poss}(X = u) = \mu_F(u)$$

This paradigm, deeply embedded in the methodology called Computing with Words, underpins natural language processing, AI, and reasoning under uncertainty.

Possibility Distributions and Restriction-Based Methods

Possibility theory, an outgrowth of fuzzy set theory, represents information and uncertainty through possibility distributions rather than probabilities (Karanjgaonkar, 2023). In possibility theory, the possibility measure of a set A of outcomes is defined as the highest possibility assigned to its elements, formally:

$$\text{Poss}_F(A) = \sup_{u \in A} r_F(u) \dots \dots \dots (14)$$

Here $r_F(u)$ is the possibility degree associated with the value u in fuzzy set F. The associated necessity measure is given by:

$$\text{Nec}_F(A) = 1 - \text{Poss}_F(A^c) \dots \dots \dots (15)$$

Possibility distributions are especially useful in modelling epistemic uncertainty (ignorance, partial knowledge), with practical fitting methods including similarity-based assignments, measurement scales, and aggregation of empirical data. They are also central to modelling and analysing information, scaling from fully specific (singleton/sure) to completely ignorant (all values possible).

The process of "instantiation and precisiation" converts natural language restrictions into formal possibility distributions—the minimum specificity distribution that conforms to observed or described data. This enables robust integration of semantic restriction into information systems (Karanjgaonkar, 2023).

Benefits and Advances

- Context-sensitive representation: Fuzzy and possibilistic measures model both the qualitative and quantitative variability in linguistic data, supporting refined approaches to knowledge representation and reasoning.
- Bridging uncertainty and meaning: They provide a natural way to reconcile the imprecision found in natural language with quantitative analysis, supporting hybrid reasoning in AI, decision-making systems, and knowledge processing.
- Practical advantages: Restriction-based possibilistic approaches offer common-sense applicability, faster approximate reasoning, and robustness against incomplete or noisy data, which are limitations of crisp (probabilistic) methods.
- Expanding the semantic scope: These methods have led to advances in the analysis of specificity (how much information is focused), similarity between distributions, and systematic conversion between possibility and probability representations, granting wide scope for applications in hybrid and adaptive systems.

IV. Synthesis: Recent Research and Applications

Recent research in semantic information theory highlights the need for an integrated, general theory of information and uncertainty—one that unites syntactic, semantic, and pragmatic aspects and supports real-world applications where data, vagueness, and meaning must be processed together.

Toward a General Theory: The Trinity Model

Modern theoretical advances (Zhong, 2017) emphasize a “trinity” model for information:

- Syntactic information (structure),
- Semantic information (meaning),
- Pragmatic information (utility, action),

suggesting that only the combination of these facets yields truly useful information for intelligent systems and human cognition.

This “information ecosystem” (Zhong, 2017) can be modelled mathematically as the transformation of raw object information S into syntactic (X), semantic (Y), and pragmatic (Z) forms:

$$Y = f(X, Z) \quad y = f(x, z), \text{ where } x \in X, y \in Y, z \in Z$$

Uncertainty: Fusing Probability and Possibility

Classical information theory uses Shannon entropy:

$$H = -\sum_i p_i \log p_i \quad \dots\dots\dots(16)$$

but this only captures average uncertainty in discrete settings. In many real-world scenarios (sensor fusion, decision-making with vague data, intelligent agents), information is incomplete, vague, or context-dependent.

Possibility theory generalizes this, employing possibility distributions $\Pi(x)$ and related measures for both “epistemic” (knowledge-based) and “physical” (feasibility-based) uncertainty:

$$\text{Poss}(A) = \sup_{x \in A} \Pi(x)$$

$$\text{Nec}(A) = 1 - \text{Poss}(A^c)$$

Practical systems often require fusion of probability- and possibility-based information (e.g., multimodal sensory data in robotics, semantic command interpretation in hybrid human–machine systems) for robust reasoning and action.

Scenario 1: Hybrid Human–Machine Systems

In collaborative AI (e.g., smart environments, robotics, cognitive assistants), machines must interpret both precise sensor data (probabilistic) and human commands (possibilistic, fuzzy). Semantic information analysis here follows:

- **Step 1:** Data stream from sensors $X = \{x_1, x_2, \dots, x_n\}$ with frequencies $F = \{f_1, \dots, f_n\}$
- **Step 2:** Construction of both probability and possibility distributions:

$$P(x_i) = \frac{f_i}{N}, \pi_m(x_i) = \frac{f_i}{f_m} \quad \dots\dots\dots(17)$$

- **Step 3:** Selection of actions/decisions is then based on the fusion of these measures, evaluated via proximity, specificity and similarity indices (Karanjgaonkar & Jha, 2018)

$$\text{Proximity Value(PV)} = \sum_{i=1}^n |\pi_m(x_i) - P(x_i)| \quad \dots\dots\dots(18)$$

This approach enables robust AI decision-making the presence of vagueness and partial knowledge.

Scenario 2: Semantic Interpretation in Multimodal Data (Cognitive Systems)

Bhatt & Kersting (2017) highlight applications in cognitive vision, human–robot interaction, and activity/situation interpretation from multimodal sources (text, video, sensor streams):

- Integrated frameworks use declarative abstraction and neural-symbolic models to assign semantics (meanings, intentions) to observed behaviour in uncertain, dynamic environments.
- Deep semantic and possibilistic reasoning guides the assignment of meaning and prediction of actions, crucial in assistive technologies for healthcare, smart homes, and autonomous vehicles.

Scenario 3: Decision Support and Resource Planning

Possibilistic semantic information models are applied in production and service planning:

- Representation of vague demand as possibility distributions,
- Aggregation of expert input using fuzzy restrictions,
- Optimisation using scenario analysis and similarity-specificity measures to select feasible plans under uncertainty.

For example, in aggregate service planning, each demand scenario is modelled as a possibility distribution, and the optimal plan is chosen by maximising similarity and specificity (Karanjgaonkar & Jha, 2018).

$$\text{Similarity}(\pi_1, \pi_2) = 1 - \frac{1}{n-1} \sum_{i=1}^{n-1} |\pi_1(x_i) - \pi_2(x_i)| \dots\dots\dots(19)$$

Advocacy for a Unified Theory

Across these domains, researchers (Zhong, 2017; Jayesh, 2023; Bhatt & Kersting, 2017) repeatedly find that no single probabilistic or possibilistic framework alone suffices. The fusion of syntactic, semantic, and pragmatic information—coupled with representations for both stochastic and possibilistic uncertainty—is essential for robust, adaptable decision support, knowledge systems, and intelligent agents.

Scenario applications—ranging from autonomous decision systems to semantic reasoning in hybrid environments—demonstrate the operational need for a unified general theory of information and uncertainty, as both foundational and context-driven models must be flexibly integrated to tackle real-world complexity.

Conclusion

Semantic Information Theory has matured into an indispensable interdisciplinary field linking mathematics, artificial intelligence, linguistics, cognitive science, and complex systems research, yet it remains under active development. Classic challenges—the gap between raw data and meaningful content, the insufficiency of purely syntactic measures, and the absence of a widely accepted general theory of information—are being addressed from multiple complementary directions. Fuzzy logic and possibility theory offer rigorous tools to represent graded membership, vagueness, and context, allowing natural language descriptions, expert judgments, and incomplete observations to be encoded as fuzzy sets and possibility distributions instead of forced into crisp probabilistic models. Zadeh's restriction-based paradigm further reframes information as generalized constraints of the form “X is R,” and, when combined with measures such as specificity, similarity, and proximity, enables systematic analysis and fusion of heterogeneous sources of semantic content in hybrid human-machine environments. Parallel developments in semantic generalizations of Shannon theory, ecological models of information-knowledge-intelligence, and evidence-based approaches such as Dempster-Shafer theory collectively point toward a unified account of information that can accommodate syntax, meaning, utility, and multiple kinds of uncertainty within a single conceptual framework. There remains, however, a clear need for deeper formalization, empirical validation on large-scale real-world systems, and tighter integration with modern machine learning and cognitive architectures, especially as semantic communication systems,

autonomous agents, and socio-technical infrastructures become more complex, adaptive, and context dependent.

References

1. Mingers, J.C., 1996. An Evaluation of Theories of Information with regard to the Semantic and Pragmatic Aspects of Information Systems. *Systems Practice*, 9(3), pp. 235-257.
2. Karanjgaonkar, J., 2023. On Study of Semantic Information Theory - A Possibilistic Approach. Thesis, Pt. Ravishankar University, Raipur.
3. Bar-Hillel, Y. and Carnap, R., 1952. An Outline of a Theory of Semantic Information. *Studies in Logical Theory*.
4. Hintikka, J., 1968. Logical Foundations of Information Content. *Philosophica*.
5. Dretske, F., 1981. Knowledge and the Flow of Information. MIT Press.
6. Jumarie, G., 1990. Relative Information Theory. *Information Sciences*.
7. Zadeh, L.A., 1965. Fuzzy Sets. *Information and Control*, 8(3), pp.338-353.
8. Yager, R.R., 1996. Measures of Specificity and Possibility Distributions. *Fuzzy Sets and Systems*, 77, pp.271-278.
9. Dubois, D. and Prade, H., 1997. The Three Semantics of Fuzzy Sets. *Fuzzy Sets and Systems*, 90(2), pp.141-150.
10. Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press.
11. Sentz, K. and Ferson, S., 2002. Combination of Evidence in Dempster-Shafer Theory of Evidence. Sandia National Laboratories, Albuquerque, NM, Report SAND2002-0835.
12. Taroun, A. and Yang, J.B., 2019. Dempster-Shafer theory of evidence: potential usage for construction project risk assessment. *Journal of Engineering, Design and Technology*, 17(2), pp.370-388. <https://doi.org/10.1108/JEDT-04-2018-0050>
13. Zadeh, L.A., 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1), pp.3-28.
14. Zhong, Y., 2017. A Theory of Semantic Information. Proceedings of the IS4SI 2017 Summit – Digitalisation for a Sustainable Society, Gothenburg, Sweden, 12–16 June 2017. MDPI Proceedings, 1(4), p.129. <https://doi.org/10.3390/IS4SI-2017-04000>
15. Bhatt, M. and Kersting, K., 2017. Semantic Interpretation of Multi-Modal Human Behaviour Data. *KI-Künstliche Intelligenz*, 31(4), pp.317-320. <https://doi.org/10.1007/s13218-017-0511-y>
16. Karanjgaonkar, J. and Jha, P., 2018. Possibilistic Analysis of Uncertainty and Vagueness in Semantic Communication. *International Journal of Recent Advances in Multidisciplinary Topics*, 5(3), pp.309-312.