Theoretical Interpretations of Large Uncertain and Hyper Language Models: Advancing Natural Uncertain and Hyper Language Processing

Takaaki Fujita [1] *
[1] Independent Researcher, Shinjuku, Shinjuku-ku, Tokyo, Japan.

Abstract

This paper explores the integration of uncertainty frameworks such as Fuzzy, Neutrosophic, and Plithogenic sets into Large Language Models (LLMs) and Natural Language Processing (NLP). We propose novel models, including Large Uncertain Language Models and Natural Uncertain Language Processing, to enhance linguistic representations and processing capabilities. Furthermore, we extend the theoretical foundation of LLMs and NLP by incorporating Hyperstructures and Superhyperstructures, enabling higher-order generalizations and hierarchical modeling. These advancements provide new perspectives for addressing uncertainty and complexity in language understanding and processing. While the paper focuses on theoretical generalizations, practical validation through computational experiments remains an important direction for future work.

Keywords: Hyperstructure, Large Language Models, Natural Language Processing, nth Power set
MSC 2010 classifications: 03E72 - Fuzzy set theory; 68T50 - Natural language processing; 68Q55 - Semantics (including automata theory)

1 Short Introduction

1.1 Large Language Models and Natural Language Processing

Artificial Intelligence (AI) [148, 194, 319], Natural Language Processing (NLP) [25, 50], Machine Learning [316], Graph Neural Networks [13, 157, 304], and Data Analysis [28, 214] have become integral components of modern life. While numerous concepts are being actively researched, this paper focuses on Large Language Models (LLMs) and Natural Language Processing.

Large Language Models (LLMs) are advanced AI systems trained on extensive text datasets to understand, generate, and process human language [23, 51, 133, 208, 314]. For instance, multimodal LLMs, capable of processing not only text but also other modalities such as images and audio, have become increasingly indispensable in practical applications [70, 220, 322, 330, 332, 343].

Furthermore, various specialized LLMs have been developed to meet specific objectives. These include:

- *Instruction-Tuned LLMs*: Models fine-tuned to effectively follow user instructions, improving interaction and response generation [21, 69, 118, 122, 160].
- *Domain-Specific LLMs*: Models tailored for specialized fields such as finance, healthcare, or legal applications [145, 174, 210].
- *Lightweight LLMs*: Optimized for resource-constrained environments, enabling deployment on mobile devices and edge computing platforms [72, 140, 248].
- *Conversational LLMs*: Designed for generating natural and contextually appropriate dialogue, focusing on improving coherence and interactivity in conversations [188, 243, 306, 349].

These variations of LLMs represent active areas of research, addressing diverse challenges and expanding their applicability across a wide range of domains.

Natural Language Processing, on the other hand, involves the computational analysis, interpretation, and generation of human language. It plays a crucial role in communication, decision-making, and knowledge discovery [25, 50, 53, 175, 185, 187]. It is often studied alongside Large Language Models.

- *Multimodal Natural Language Processing*: Expands NLP to process and integrate data from multiple modalities, such as text, images, and audio [73, 168, 221].
- *Context-Aware Natural Language Processing*: Enhances traditional NLP by incorporating contextual information to improve accuracy and relevance [134].

Furthermore, the concept of *HyperLanguage* extends traditional language structures [29, 30, 80]. HyperLanguage finds application in various domains, such as automata theory, where it enhances the representation and processing of complex linguistic constructs [29, 99].

1.2 Uncertain Sets

Set theory, a foundational branch of mathematics, provides the framework for studying collections of objects, known as "sets" [66, 144, 303, 305]. Over time, the classical concept of sets has been extended to address the complexities of real-world uncertainty. Prominent among these extensions are Fuzzy Sets [59, 292, 334, 336–339, 350], Vague Sets [5, 43, 48, 128, 346], Soft Sets [8, 10, 88, 182, 196, 328], Hypersoft Sets [272, 273], Rough Sets [212–217, 219], Hyperfuzzy Sets [87, 106, 154, 286], and Neutrosophic Sets [35, 74, 203, 257, 258, 282, 308].

Each of these frameworks addresses unique aspects of uncertainty. For instance, *Fuzzy Sets* assign a membership degree between 0 and 1 to each element, enabling the modeling of partial belonging [334]. In contrast, *Neutrosophic Sets* extend this concept by introducing three degrees: truth, indeterminacy, and falsehood, providing a comprehensive approach to handling ambiguity in complex systems [257, 258]. Among these, *Plithogenic Sets* stand out as a versatile generalization that incorporates multiple levels of uncertainty and contradiction. These sets have gained attention for their adaptability to complex systems [2, 109, 233, 251, 261, 262, 280, 283, 289].

Research on uncertain sets, including the aforementioned frameworks, has expanded significantly [99, 155, 257, 258, 262]. These studies have led to advancements in uncertain graphs, such as Fuzzy Graphs and Neutrosophic Graphs, which have been applied to a variety of problems [82, 86–88, 92, 96–98, 235].

Moreover, these concepts have seen practical applications in fields such as Neural Networks [9, 123, 166, 176, 180, 229, 293, 294] and decision-making processes [4, 6, 46, 113, 151, 223, 224], showcasing their relevance and versatility in addressing uncertainty across diverse domains.

1.3 Hyperstructure and Superhyperstructure

This subsection explains the concepts of Hyperstructure and Superhyperstructure. Hyperstructures and Superhyperstructures represent hierarchical structures. A *Hyperstructure* is a mathematical construct that extends the concept of power sets to various mathematical frameworks. Building upon this foundation, a *Superhyperstructure* incorporates the notion of n-th power sets, providing a hierarchical and iterative generalization of hyperstructures. Superhyperstructures can be interpreted as repeated applications of hyperstructural principles, enabling deeper levels of abstraction and complexity [278, 279].

For instance, in graph theory, a *Hypergraph* is a hyperstructure, while a *Superhypergraph* is a superhyperstructure, offering a higher level of abstraction and complexity. A *Hypergraph* is defined as a generalization of a traditional graph (cf. [68]) where edges, called hyperedges, can connect more than two vertices [24, 32, 110–112]. Hyperstructures have been extensively studied in the field of AI, including concepts like Hypergraph Neural Networks [49, 79, 101, 125, 132, 135, 296, 312]. In contrast, a *SuperHyperGraph* represents a more generalized class of graphs that incorporates superedges and supervertices. This extension builds upon fundamental graph-theoretic concepts, including traditional graphs and hypergraphs, to achieve greater levels of abstraction and flexibility (cf. [84, 84, 87, 90, 91, 93, 94, 105, 119, 120, 231, 263, 264, 267, 271, 276, 276, 278]). Similarly, SuperHypergraph Neural Networks have also been actively explored in recent studies [85].

Beyond graph theory, *superhyperstructures* have been extensively studied across various mathematical disciplines, including:

- *Topology [163]*: The study of hypertopologies [67, 177, 179, 200] and superhypertopologies [274, 275, 284] has revealed novel insights into topological structures.

- *Functions*: Hyperfunctions [103, 139, 197] and superhyperfunctions [270, 276] extend the functional analysis framework.
- *Soft Sets [182, 196]*: Advances include hypersoft sets [1, 97, 124, 141, 225, 242, 245, 269] and superhypersoft sets [272, 285], broadening the scope of soft set theory.
- *Fuzzy Sets [235, 334]*: Hyperfuzzy sets [106, 154, 286] and superhyperfuzzy sets [87] provide a higher-order generalization of classical fuzzy sets.
- *Group Theory [170]*: The extension of hypergroups [159] to superhypergroups [159] enriches algebraic structures in group theory.
- *Neutrosophic Sets [257]*: Hyperneutrosophic sets [62, 87] and superhyperneutrosophic sets [87] address higher-order uncertainties.
- *Algebra Theory [20, 56, 161]*: Developments in hyperalgebras [57, 137, 227, 290] and superhyperalgebras [142, 143, 255, 265, 284] expand the understanding of algebraic systems.
- *Automata Theory* [31, 131]: The concepts of *Hyperautomata* [29, 241] and *Superhyperautomata* [99] incorporate hyperstructures into the framework of automata theory, extending its theoretical boundaries.
- *Ring Theory [288, 320]*: *Hyperring Theory* [14, 58, 153, 164] and *Superhyperring Theory* [277] explore hyperstructures within algebraic systems, enriching the study of commutative and non-commutative rings.
- *Rough Set Theory* [214, 217, 218]: The notions of *Hyperrough Sets* [87, 253] and *Superhyperrough Sets* [87] generalize classical rough set theory by leveraging hyperstructural principles.

Given this broad scope of applications, research into Hyperstructures and SuperHyperstructures is crucial for advancing mathematical theory and its interdisciplinary applications.

1.4 Our Contribution in This Paper

This subsection provides a concise explanation of our contributions in this paper.

We theoretically propose models for Large Uncertain Language Models and Natural Uncertain Language Processing by incorporating the concepts of Fuzzy, Neutrosophic, and Plithogenic frameworks into the domains of Large Language Models and Natural Language Processing. Additionally, we introduce models that integrate the concepts of Hyperstructure and Superhyperstructure into Large Language Models and Natural Language Processing, expanding their theoretical foundation.

It is important to note that this discussion primarily focuses on theoretical generalizations. The practical feasibility and robustness of these methods in real-world applications require further computational experiments and validation.

1.5 The Structure of the Paper

The format of this paper is described below.

2 Preliminaries and Definitions

In this section, we provide the necessary preliminaries and definitions. As it is not feasible to explain or define every concept within this paper, readers are advised to refer to the respective notations and basic definitions in the relevant literature as needed.

2.1 Basic Set Theory

This subsection introduces foundational concepts of set theory. For a detailed discussion, readers are encouraged to consult standard references [126, 144, 150].

Definition 2.1 (Set)**.** [144] A *set* is a well-defined collection of distinct objects, referred to as *elements*. For any object x, it can be unambiguously determined whether x belongs to the set. If A is a set and x is an element of A, this relationship is expressed as $x \in A$. Sets are often denoted using curly brackets. For example, $A = \{1, 2, 3\}$ represents a set containing the elements 1, 2, and 3.

Definition 2.2 (Subset)**.** [144] Let A and B be sets. The set A is said to be a *subset* of B, written $A \subseteq B$, if every element of A is also an element of B. Formally:

$$A \subseteq B \iff \forall x\,(x \in A \implies x \in B).$$

If $A \subseteq B$ and $A \neq B$, then A is called a *proper subset* of B, denoted $A \subset B$.

2.2 Uncertain Sets

In this subsection, we focus on *Uncertain Sets*, exploring concepts such as fuzzy sets, Neutrosophic sets, and plithogenic sets.

Fuzzy sets offer a powerful framework for managing uncertainty by assigning degrees of membership to elements [334, 340, 341]. Building upon this foundation, several extensions have been proposed, including bipolar fuzzy sets [7,45,114], intuitionistic fuzzy sets [16–19], hesitant fuzzy sets [76,77,152,192,298,299,323], picture fuzzy sets [54, 162, 252, 313], spherical fuzzy sets [15, 115, 116, 181, 189], pythagorean fuzzy sets [102, 184, 240, 347], and vague sets [43, 48, 128].

Neutrosophic sets extend the concept of fuzzy sets by incorporating the notion of indeterminacy, enabling representation of states that are neither entirely true nor entirely false. This framework has been extensively explored in various fields [257–259]. Furthermore, related concepts include Bipolar Neutrosophic Sets [3, 64, 65, 195, 302], Complex Neutrosophic Sets [11, 12], Single-Valued Neutrosophic Sets [44, 146, 158, 244, 309], Interval-Valued Neutrosophic Sets [329, 333, 344, 345], and Neutrosophic Offsets [83, 256, 260, 266, 268].

Plithogenic sets further enhance these frameworks by accommodating a greater degree of complexity and multi-dimensional uncertainty. They provide a flexible and robust approach for modeling intricate scenarios, making them a versatile tool for uncertainty management [98, 242, 262, 283].

The relevant definitions, theorems, and related concepts are presented below.

Definition 2.3 (Fuzzy set)**.** [334–336,339] A *fuzzy set* τ in a non-empty universe Y is a mapping $\tau : Y \to [0,1]$. A *fuzzy relation* on Y is a fuzzy subset δ in $Y \times Y$. If τ is a fuzzy set in Y and δ is a fuzzy relation on Y, then δ is called a *fuzzy relation on* τ if

$$\delta(y,z) \leq \min\{\tau(y), \tau(z)\} \quad \text{for all } y, z \in Y.$$

Example 2.4 (Fuzzy set)**.** Consider a non-empty universe $Y = \{\text{Cold}, \text{Moderate}, \text{Hot}\}$, which represents temperature levels. A fuzzy set τ maps each element in Y to a degree of membership in the interval $[0,1]$. For example:

$$\tau(\text{Cold}) = 0.8, \quad \tau(\text{Moderate}) = 0.5, \quad \tau(\text{Hot}) = 0.2.$$

This means the degree of ”coldness” is 0.8, ”moderateness” is 0.5, and ”hotness” is 0.2 for the given context (e.g., a day with a temperature of 15°C). This approach accommodates the vagueness of linguistic terms like ”cold” or ”hot.”

A fuzzy relation δ on Y could represent the perceived similarity between temperature levels:

$$\delta(\text{Cold}, \text{Moderate}) = 0.6, \quad \delta(\text{Cold}, \text{Hot}) = 0.2, \quad \delta(\text{Moderate}, \text{Hot}) = 0.7.$$

The fuzzy relation satisfies:

$$\delta(y,z) \leq \min\{\tau(y), \tau(z)\} \quad \text{for all } y, z \in Y.$$

Definition 2.5. [257] Let X be a given set. A Neutrosophic Set A on X is characterized by three membership functions:

$$T_A : X \to [0,1], \quad I_A : X \to [0,1], \quad F_A : X \to [0,1],$$

where for each $x \in X$, the values $T_A(x)$, $I_A(x)$, and $F_A(x)$ represent the degree of truth, indeterminacy, and falsity, respectively. These values satisfy the following condition:

$$0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3.$$

Example 2.6. Let $X = \{\text{Product A}, \text{Product B}, \text{Product C}\}$, representing products in an online store. A Neutrosophic Set A assigns to each product a degree of truth $T_A(x)$, indeterminacy $I_A(x)$, and falsity $F_A(x)$. For example:

$$T_A(\text{Product A}) = 0.7, \quad I_A(\text{Product A}) = 0.2, \quad F_A(\text{Product A}) = 0.1.$$

Here, $T_A(\text{Product A}) = 0.7$ indicates 70% positive reviews, $I_A(\text{Product A}) = 0.2$ reflects 20% uncertain or mixed feedback, and $F_A(\text{Product A}) = 0.1$ signifies 10% negative reviews. These values satisfy:

$$0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3 \quad \text{for all } x \in X.$$

Proposition 2.7. *(cf. [258, 281]) A Neutrosophic Set generalizes a Fuzzy Set.*

Proof. This follows directly from the definition. □

Definition 2.8. [261, 262] Let S be a universal set, and $P \subseteq S$. A *Plithogenic Set PS* is defined as:

$$PS = (P, v, Pv, pdf, pCF)$$

where:

- v is an attribute.
- Pv is the range of possible values for the attribute v.
- $pdf : P \times Pv \to [0,1]^s$ is the *Degree of Appurtenance Function (DAF)*.
- $pCF : Pv \times Pv \to [0,1]^t$ is the *Degree of Contradiction Function (DCF)*.

These functions satisfy the following axioms for all $a, b \in Pv$:

1. *Reflexivity of Contradiction Function*:

$$pCF(a, a) = 0$$

2. *Symmetry of Contradiction Function*:

$$pCF(a, b) = pCF(b, a)$$

Example 2.9. (cf. [83, 96]) The following examples of Plithogenic sets are provided.

- When $s = t = 1$, PS is called a *Plithogenic Fuzzy Set*.
- When $s = 2, t = 1$, PS is called a *Plithogenic Intuitionistic Fuzzy Set*.
- When $s = 3, t = 1$, PS is called a *Plithogenic Neutrosophic Set*.
- When $s = 4, t = 1$, PS is called a *Plithogenic quadripartitioned Neutrosophic Set* (cf. [138, 228, 249]).
- When $s = 5, t = 1$, PS is called a *Plithogenic pentapartitioned Neutrosophic Set* (cf. [26, 55, 183]).
- When $s = 6, t = 1$, PS is called a *Plithogenic hexapartitioned Neutrosophic Set* (cf. [211]).
- When $s = 7, t = 1$, PS is called a *Plithogenic heptapartitioned Neutrosophic Set* (cf. [34, 198]).
- When $s = 8, t = 1$, PS is called a *Plithogenic octapartitioned Neutrosophic Set*.
- When $s = 9, t = 1$, PS is called a *Plithogenic nonapartitioned Neutrosophic Set*.

Libraries and programming frameworks for Fuzzy Sets and Neutrosophic Sets are discussed in various studies. For Fuzzy Sets, references such as [81, 149, 167, 169, 191, 250, 287, 351, 352] provide valuable insights. For Neutrosophic Sets, relevant research can be found in [36, 205, 254], among others. While these references are merely examples, they can be consulted as needed.

2.3 Hyperstructure and Superhyperstructure

This subsection provides an explanation of Hyperstructure and Superhyperstructure. A *Hyperstructure* refers to a mathematical concept characterized by the structure of a power set. The term *Superhyperstructure* denotes the structure defined by the n-th power set [278, 279]. These concepts enable the representation of various hierarchical structures. The definition of the n-th power set is given below.

Definition 2.10 (Base Set)**.** A *base set* is a foundational set S from which more complex structures, such as powersets or hyperstructures, are derived. Formally, a base set is defined as:

$$S = \{x \mid x \text{ is an element of the universe of discourse}\}.$$

All elements in derived structures, such as $\mathcal{P}(S)$ or $\mathcal{P}_n(S)$, originate from the elements of the base set S.

Definition 2.11 (Powerset)**.** [85, 234] The *powerset* of a set S, denoted $\mathcal{P}(S)$, is the set of all subsets of S, including the empty set and S itself. Formally,

$$\mathcal{P}(S) = \{A \mid A \subseteq S\}.$$

Definition 2.12 (n-th powerset)**.** (cf. [85, 255, 278]) The n-th powerset of H, denoted $P_n(H)$, is defined recursively as:

$$P_1(H) = P(H), \quad P_{n+1}(H) = P(P_n(H)) \quad \text{for } n \geq 1.$$

Similarly, the n-th non-empty powerset of H, denoted $P_n^*(H)$, is defined as:

$$P_1^*(H) = P^*(H), \quad P_{n+1}^*(H) = P^*(P_n^*(H)).$$

Proposition 2.13. *An n-th powerset generalizes a power set.*

Proof. This is evident. □

If we were to mathematically define Hyperstructure and Superhyperstructure, the definitions would be as follows.

Definition 2.14 (Hyperstructure)**.** A *Hyperstructure* is a mathematical structure that incorporates elements from the powerset of a base set. Formally, a hyperstructure $\mathcal{H}$ is defined as:

$$\mathcal{H} = (\mathcal{P}(S), \circ),$$

where S is a base set, $\mathcal{P}(S)$ is its powerset, and $\circ$ is an operation defined on subsets of S.

Proposition 2.15. *A Hyperstructure possesses the structure of a Power set.*

Proof. This follows directly from the definition. □

Definition 2.16 (n-Superhyperstructure)**.** (cf. [255, 278]) An *n-Superhyperstructure* generalizes a hyperstructure by iteratively applying the powerset operation n-times. It is formally defined as:

$$\mathcal{SH}_n = (\mathcal{P}_n(S), \circ),$$

where S is the base set, $\mathcal{P}_n(S)$ is the n-th powerset of S, and $\circ$ is a general operation defined on $\mathcal{P}_n(S)$.

Proposition 2.17. *A n-Superhyperstructure possesses the structure of a n-th powerset.*

Proof. This follows directly from the definition. □

Proposition 2.18. *An n-Superhyperstructure generalizes a Hyperstructure.*

Proof. By definition, a Hyperstructure $\mathcal{H} = (\mathcal{P}(S), \circ)$ is based on the powerset $\mathcal{P}(S)$ of a base set S, with an operation $\circ$ defined on $\mathcal{P}(S)$.

An n-Superhyperstructure $\mathcal{SH}_n = (\mathcal{P}_n(S), \circ)$ extends this concept by applying the powerset operation n-times, where $\mathcal{P}_n(S) = \mathcal{P}(\mathcal{P}_{n-1}(S))$, and $\mathcal{P}_1(S) = \mathcal{P}(S)$.

For $n = 1$, we have $\mathcal{P}_1(S) = \mathcal{P}(S)$, and hence $\mathcal{SH}_1 = (\mathcal{P}(S), \circ)$, which is equivalent to a Hyperstructure. For $n > 1$, the n-th powerset introduces additional levels of hierarchical structure beyond the base powerset, thereby generalizing the original Hyperstructure.

Thus, an n-Superhyperstructure reduces to a Hyperstructure for $n = 1$, and for $n > 1$, it represents a broader generalization. □

3 Theoretical Considerations of Uncertain Language

In this section, we explore concepts related to languages that incorporate uncertainty, such as Fuzzy Languages, and discuss their theoretical underpinnings.

3.1 Uncertain Natural Language Processing (NLP)

We introduce and mathematically define several frameworks under the umbrella of Uncertain Natural Language Processing (NLP), including:

- *Natural Fuzzy Language Processing*
  A system that handles linguistic uncertainty using fuzzy languages, assigning degrees of membership to words or phrases.
- *Natural Neutrosophic Language Processing*
  A framework incorporating truth, indeterminacy, and falsity degrees for nuanced processing of ambiguous linguistic data.

- *Natural Plithogenic Language Processing*
  A method combining plithogenic languages to manage linguistic data with multiple attributes and higher-order uncertainty.

These frameworks aim to extend the theoretical foundation of NLP to accommodate uncertainty and vagueness inherent in natural language. It is important to note that this discussion focuses on theoretical generalizations. Practical feasibility and robustness of these methods for real-world applications require further computational experiments and validation.

3.1.1 Classic Natural Language Processing (NLP)

Natural Language refers to human languages(cf. [202, 222, 317]) used for communication, encompassing spoken, written, or signed forms, which evolve naturally over time. Natural Language Processing (NLP) involves enabling computers to understand, interpret, and generate human language for purposes of communication and analysis [25,50]. NLP has been extensively studied in various contexts and applications [25,33,50,52,53,75, 108,175,185,187,199,321,331].

The definitions and examples are provided below. Readers interested in learning more about Natural Language Processing are encouraged to consult the relevant survey introductions as needed [71,156,186,209,318].

Definition 3.1 (Formal Language)**.** [95,121,130,232,238] A *formal language* $\mathcal{L}$ is defined as a set of strings (or sequences) formed from a finite alphabet Σ, subject to specific syntactic rules. Formally:

$$\mathcal{L} \subseteq \Sigma^*,$$

where Σ^* is the set of all finite strings over the alphabet Σ. The strings in $\mathcal{L}$ are called *well-formed formulas (WFFs)*.

A formal language $\mathcal{L}$ is typically accompanied by:

- A set of *symbols* (or *alphabet*) Σ, which may include logical connectives (e.g., $\wedge$, $\vee$, $\neg$), quantifiers (e.g., $\forall$, $\exists$), variables, and parentheses.
- A set of *formation rules* that determine which strings in Σ^* are well-formed.

Example 3.2 (Formal Language)**.** Consider the formal language $\mathcal{L}$ over the alphabet $\Sigma = \{a, b\}$, defined as:

$$\mathcal{L} = \{w \in \Sigma^* \mid w \text{ contains an equal number of } a\text{'s and } b\text{'s}\}.$$

This language includes all strings formed from a and b such that the number of occurrences of a in the string equals the number of occurrences of b. Some examples of well-formed strings (words) in $\mathcal{L}$ are:

$$\varepsilon, ab, ba, aabb, abab, bbaa, \ldots$$

where ε represents the empty string.

Formation Rules:

- The empty string ε is in $\mathcal{L}$.
- If $w \in \mathcal{L}$, then $awb \in \mathcal{L}$ and $bwa \in \mathcal{L}$.
- No other strings are in $\mathcal{L}$.

The language $\mathcal{L}$ is an example of a formal language that can be described by a context-free grammar. It ensures that all strings adhere to the rule of equal numbers of a's and b's, representing a well-defined syntactic structure over the alphabet Σ.

Definition 3.3 (Word)**.** (cf. [121, 232]) Let Σ be a finite set of symbols, referred to as an *alphabet*. A *word* over Σ is defined as a finite sequence of symbols from Σ. Formally, a word w is an element of Σ^*, where:

$$\Sigma^* = \bigcup_{n=0}^{\infty} \Sigma^n,$$

and Σ^n denotes the set of all sequences of length n formed from Σ, including the empty sequence ε when $n = 0$.

For a word $w \in \Sigma^*$, the length of w, denoted $|w|$, is the number of symbols in w. If $w = \varepsilon$, then $|w| = 0$. For example:

- If $\Sigma = \{a, b\}$, then $w = aba \in \Sigma^*$ is a word of length $|w| = 3$.
- The empty word $\varepsilon \in \Sigma^*$ is the unique word with $|w| = 0$.

Definition 3.4 (Natural Language)**.** (cf. [25, 50, 185]) A *natural language* is a system of communication composed of words, phrases, and rules, developed naturally among humans for expressing thoughts, emotions, and information. Unlike formal languages, natural languages are characterized by ambiguity, irregularity, and context-dependence, and are primarily governed by implicit grammar rather than strict syntactic rules. Examples include English, Japanese, and Arabic.

Definition 3.5 (Probability Model)**.** (cf. [104, 236, 237]) A *probability model* is a tuple $(\Omega, \mathcal{F}, P)$, where:

- Ω is the sample space,
- $\mathcal{F}$ is a σ-algebra of subsets of Ω,
- $P : \mathcal{F} \to [0, 1]$ is a probability measure satisfying:

  $$P(\Omega) = 1 \quad \text{and} \quad P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i),$$

  for any countable collection of disjoint events $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$.

Definition 3.6 (Natural Language Processing (NLP))**.** (cf. [25, 50, 185]) Let Σ be a finite alphabet representing the vocabulary of a natural language, and let Σ^* denote the set of all finite sequences (words) over Σ. A *language* $\mathcal{L}$ is a subset $\mathcal{L} \subseteq \Sigma^*$.

An NLP system is a tuple:

$$\mathcal{N} = (\Sigma, \mathcal{L}, \mathcal{P}, \mathcal{M}, \mathcal{T}),$$

where:

1. Σ: A finite alphabet of symbols.
2. $\mathcal{L} \subseteq \Sigma^*$: The language, defined by some grammar $\mathcal{G}$.
3. $\mathcal{P} : \mathcal{L} \to [0, 1]$: A probability model [237] assigning probabilities to each $w \in \mathcal{L}$:

   $$\mathcal{P}(w) = P(w \mid \theta),$$

   where θ represents model parameters.
4. $\mathcal{M} : \mathcal{L} \to \mathcal{O}$: A mapping function that transforms each $w \in \mathcal{L}$ into a structured output $o \in \mathcal{O}$ (e.g., a parse tree, a translation).
5. $\mathcal{T} : \mathcal{L} \times \mathcal{L} \to \mathbb{R}$: A similarity measure between pairs of words or sentences.

Libraries and programming frameworks for natural language processing have been explored in various studies, such as [22, 171, 193, 222, 307, 310, 325, 326]. Please refer to these works as needed.

3.1.2 Natural Fuzzy Language Processing

We explore *Natural Fuzzy Language Processing*, which combines the principles of Natural Language Processing and Fuzzy Language [60, 61, 129, 246, 324]. Definitions, theorems, and related concepts are provided below.

Definition 3.7 (Fuzzy Language)**.** [60, 61, 226] Let Σ be a finite alphabet, and let Σ^* and Σ^ω denote the sets of all finite and infinite words over Σ, respectively. A *fuzzy language* is defined as follows:

1. *Fuzzy Set*: A fuzzy set A on a set X is characterized by a membership function $f_A : X \to [0, 1]$, where $f_A(x)$ represents the degree of membership of $x \in X$ in the set A.
2. *Fuzzy Language over Finite Words*: A fuzzy language over Σ^* is defined as a mapping $r : \Sigma^* \to [0, 1]$, where $r(w)$ denotes the degree of membership of the word $w \in \Sigma^*$ in the language.
3. *Fuzzy Language over Infinite Words*: Similarly, a fuzzy language over Σ^ω is a mapping $r : \Sigma^\omega \to [0, 1]$, where $r(w)$ represents the degree of membership of the infinite word $w \in \Sigma^\omega$.
4. *Support of a Fuzzy Language*: The support of a fuzzy language r is the set of words with non-zero membership:
$$\text{supp}(r) = \{w \in \Sigma^* \mid r(w) > 0\}.$$

This framework models uncertainty and partial belonging by assigning a degree of membership to each word in the language.

Example 3.8 (Fuzzy Language Example 1: Room temperature comfort)**.** Consider a linguistic context describing room temperature comfort. In everyday conversation, words like “warm,” “cool,” or “pleasant” do not strictly classify temperature as entirely comfortable or uncomfortable. Instead, they express a degree of comfort. For instance, let $\Sigma = \{\text{"cold"}, \text{"cool"}, \text{"warm"}, \text{"hot"}\}$ be a set of words used to describe temperature. A fuzzy language $r : \Sigma \to [0, 1]$ might assign:

$$r(\text{"cool"}) = 0.3, \quad r(\text{"warm"}) = 0.8,$$

indicating that “cool” has a low degree of membership to the ideal comfort zone (perhaps slightly chilly) while “warm” strongly belongs to the comfort category. The exact membership degrees depend on context or personal preference, reflecting how people naturally express nuances rather than absolute truths.

Another everyday example might involve affordability. Words like “cheap,” “affordable,” or “expensive” do not neatly classify items into binary categories. Suppose $\Sigma = \{\text{"cheap"}, \text{"affordable"}, \text{"costly"}\}$. A fuzzy language could assign:

$$r(\text{"affordable"}) = 0.6,$$

implying that “affordable” partially belongs to a “reasonably priced” category without asserting a strict boundary.

Example 3.9 (Fuzzy Language Example 2: Climate Comfort)**.** Let $\Sigma = \{\text{"humid"}, \text{"dry"}\}$ represent words describing weather conditions. Define a fuzzy language $r : \Sigma \to [0, 1]$ that measures how well these words fit the notion of a “comfortable climate”:

$$r(\text{"humid"}) = 0.5, \quad r(\text{"dry"}) = 0.4.$$

In this scenario, “humid” is somewhat comfortable (though not ideal), while “dry” is slightly less comfortable, reflecting subjective human perceptions rather than a binary classification.

Example 3.10 (Fuzzy Language Example 3: Food Quality)**.** Consider $\Sigma = \{\text{"ripe"}, \text{"stale"}\}$ referring to food freshness. Define a fuzzy language $r : \Sigma \to [0, 1]$ representing how well each word matches “good to eat”:

$$r(\text{"ripe"}) = 0.9, \quad r(\text{"stale"}) = 0.2.$$

“Ripe” strongly belongs to the “edible and appealing” category, while “stale” barely meets that criterion, illustrating a gradient of acceptability rather than a strict boundary.

Theorem 3.11. *Every formal language $\mathcal{L} \subseteq \Sigma^*$ can be represented as a fuzzy language.*

Proof. Let $\mathcal{L} \subseteq \Sigma^*$ be a formal language. By definition, every string $w \in \Sigma^*$ either belongs to $\mathcal{L}$ ($w \in \mathcal{L}$) or does not ($w \notin \mathcal{L}$).

To represent $\mathcal{L}$ as a fuzzy language, we define a fuzzy membership function $r : \Sigma^* \to [0, 1]$ as follows:

$$r(w) = \begin{cases} 1 & \text{if } w \in \mathcal{L}, \\ 0 & \text{if } w \notin \mathcal{L}. \end{cases}$$

In this case:

- $r(w)$ assigns a membership degree of 1 to strings in $\mathcal{L}$, indicating full membership.
- $r(w)$ assigns a membership degree of 0 to strings not in $\mathcal{L}$, indicating no membership.

Since a fuzzy language allows membership values in the interval $[0, 1]$, this construction is valid. The fuzzy language $r(w)$ corresponds exactly to the formal language $\mathcal{L}$.

Moreover, the support of the fuzzy language, defined as:

$$\text{supp}(r) = \{w \in \Sigma^* \mid r(w) \neq 0\},$$

is equivalent to $\mathcal{L}$, because $r(w) = 1$ for all $w \in \mathcal{L}$ and $r(w) = 0$ for all $w \notin \mathcal{L}$.

Thus, the fuzzy language $r(w)$ faithfully represents the formal language $\mathcal{L}$. □

Corollary 3.12. *Fuzzy languages are a generalization of formal languages, as they can accommodate partial membership values (*$0 < r(w) < 1$*) in addition to the binary membership of formal languages.*

Proof. This is evident. □

Definition 3.13 (Natural Fuzzy Language)**.** A *Natural Fuzzy Language* is a formal framework for representing and processing natural language with inherent uncertainty using fuzzy sets. Formally, let Σ be a finite alphabet representing the vocabulary. A Natural Fuzzy Language $\mathcal{L}_F$ is defined as:

$$\mathcal{L}_F = (\Sigma, \mathcal{M}_F, \mathcal{P}_F, \mathcal{S}_F),$$

where:

- Σ: A finite alphabet representing the set of words.
- $\mathcal{M}_F : \Sigma^* \to [0, 1]$: A fuzzy membership function assigning to each word $w \in \Sigma^*$ a degree of membership $\mathcal{M}_F(w)$. This value captures how strongly the word w belongs to the language, modeling linguistic vagueness (e.g., the word "warm" might have a membership degree $\mathcal{M}_F(\text{"warm"}) = 0.8$ in a language describing comfortable temperatures).
- $\mathcal{P}_F : \Sigma^* \times \Sigma^* \to [0, 1]$: A fuzzy relation measuring semantic proximity or contextual similarity between two words $u, v \in \Sigma^*$. For example, $\mathcal{P}_F(\text{"warm"}, \text{"cozy"}) = 0.7$ may indicate that "warm" and "cozy" are semantically related with a moderately high degree.
- $\mathcal{S}_F$: A set of syntactic or semantic rules represented as fuzzy constraints, guiding the construction and interpretation of sentences. For instance, a fuzzy syntactic rule might state that certain phrases are "somewhat acceptable" with a degree of 0.5, reflecting partial grammaticality or contextual fit.

Example 3.14 (Natural Fuzzy Language)**.** Consider a vocabulary $\Sigma = \{\text{"cold"}, \text{"cool"}, \text{"warm"}, \text{"hot"}\}$ describing temperature-related terms. A Natural Fuzzy Language $\mathcal{L}_F$ could assign:

$$\mathcal{M}_F(\text{"cool"}) = 0.3, \quad \mathcal{M}_F(\text{"warm"}) = 0.8,$$

indicating that "cool" is only somewhat representative of comfortable temperatures, while "warm" strongly fits the notion of comfort. Additionally,

$$\mathcal{P}_F(\text{"warm"}, \text{"hot"}) = 0.6,$$

suggesting that "warm" and "hot" are moderately similar in meaning. Fuzzy syntactic rules might allow for partially acceptable sentence formations, reflecting the inherent gradation in natural language structures.

Example 3.15 (Natural Fuzzy Language in a Japanese Linguistic Context)**.** Japanese is known as a language with a high degree of ambiguity in meaning (cf. [190, 301]). Consider a Natural Fuzzy Language $\mathcal{L}_F$ derived from Japanese vocabulary describing nuances in weather conditions. Let:

$$\Sigma = \{\text{"samui (cold)"}, \text{"suzushii (cool)"}, \text{"atatakai (warm)"}, \text{"atsui (hot)"}\}.$$

In everyday Japanese, these words convey nuanced perceptions of temperature, with interpretations often dependent on personal feelings and context. Define a fuzzy membership function $\mathcal{M}_F : \Sigma^* \to [0, 1]$:

$$\mathcal{M}_F(\text{"suzushii"}) = 0.6, \quad \mathcal{M}_F(\text{"atatakai"}) = 0.8.$$

Here, "suzushii (cool)" might represent a moderately pleasant coolness (0.6), neither too cold nor too warm. "Atatakai (warm)" indicates a higher degree of comfort (0.8), suggesting a more clearly positive and comfortable temperature. In contrast, "samui (cold)" might have a lower membership degree (e.g., 0.3) if we consider the fuzzy language to represent "comfortable living conditions."

A fuzzy relation $\mathcal{P}_F : \Sigma^* \times \Sigma^* \to [0, 1]$ could capture semantic proximity:

$$\mathcal{P}_F(\text{"atatakai"}, \text{"atsui"}) = 0.5.$$

Although "atsui (hot)" implies a higher temperature, it shares some semantic ground with "atatakai (warm)" as both indicate warmth, albeit at different comfort levels.

Fuzzy syntactic rules could assign partial acceptability to certain phrases depending on context. For instance, describing a day as "sukoshi atatakai (slightly warm)" might have a membership of 0.7 in a fuzzy grammar representing "pleasant weather descriptions."

Definition 3.16 (Natural Fuzzy Language Processing (NFLP))**.** Let Σ be a finite alphabet, and let $\mathcal{L} \subseteq \Sigma^*$ be a language defined by some grammar $\mathcal{G}$. A fuzzy language $r : \Sigma^* \to [0, 1]$ assigns to each word $w \in \Sigma^*$ a degree of membership $r(w) \in [0, 1]$.

A *Natural Fuzzy Language Processing (NFLP)* system is a tuple:

$$\mathcal{N}^F = (\Sigma, \mathcal{L}^F, \mathcal{P}^F, \mathcal{M}^F, \mathcal{T}^F),$$

where:

1. Σ: A finite alphabet of symbols.
2. $\mathcal{L}^F \subseteq \Sigma^*$: A language over which a fuzzy membership function is defined.
3. $\mathcal{P}^F : \mathcal{L}^F \to [0, 1]$: A fuzzy membership model assigning to each $w \in \mathcal{L}^F$ a value $\mathcal{P}^F(w) \in [0, 1]$.
4. $\mathcal{M}^F : \mathcal{L}^F \to O$: A mapping function that transforms each $w \in \mathcal{L}^F$ into a structured output $o \in O$.
5. $\mathcal{T}^F : \mathcal{L}^F \times \mathcal{L}^F \to \mathbb{R}$: A similarity measure for comparing pairs of words under fuzzy membership considerations.

Theorem 3.17. *Natural Fuzzy Language Processing (NFLP) generalizes Natural Language Processing (NLP).*

Proof. Let $\mathcal{N} = (\Sigma, \mathcal{L}, \mathcal{P}, \mathcal{M}, \mathcal{T})$ be an NLP system as defined above. In an NLP system, $\mathcal{L} \subseteq \Sigma^*$ is a formal language, and $\mathcal{P}(w)$ assigns probabilities to words $w \in \mathcal{L}$.

For NFLP, let $\mathcal{N}^F = (\Sigma, \mathcal{L}^F, \mathcal{P}^F, \mathcal{M}^F, \mathcal{T}^F)$, where $\mathcal{L}^F \subseteq \Sigma^*$ and $\mathcal{P}^F(w)$ assigns a fuzzy membership degree to words $w \in \mathcal{L}^F$.

To show generalization:

- If $\mathcal{P}^F(w) \in \{0, 1\}$, NFLP reduces to a deterministic NLP system where $\mathcal{P}(w) = 1$ for $w \in \mathcal{L}$ and $\mathcal{P}(w) = 0$ otherwise.

- If $\mathcal{P}^F(w)$ takes values in $[0,1]$, NFLP allows partial membership for w, capturing uncertainty or vagueness, which NLP cannot.

Since NLP is a special case of NFLP when $\mathcal{P}^F(w) \in \{0,1\}$, NFLP generalizes NLP. □

Theorem 3.18. *Natural Fuzzy Language Processing (NFLP) possesses the structure of a fuzzy language.*

Proof. By definition, NFLP operates over $\mathcal{L}^F \subseteq \Sigma^*$ with a fuzzy membership function $\mathcal{P}^F : \mathcal{L}^F \to [0,1]$. This aligns with the definition of a fuzzy language, where $r : \Sigma^* \to [0,1]$ assigns membership degrees to words.

In NFLP:

- Σ^* is the set of all finite sequences over the alphabet Σ.
- $\mathcal{P}^F$ is equivalent to the membership function r in a fuzzy language, as it maps each word $w \in \Sigma^*$ to $[0,1]$.
- The support of $\mathcal{P}^F$, defined as $\text{supp}(\mathcal{P}^F) = \{w \in \Sigma^* \mid \mathcal{P}^F(w) \neq 0\}$, corresponds to the set of words with non-zero membership.

Thus, NFLP inherits the structure of a fuzzy language. □

3.1.3 Natural Neutrosophic Language Processing

Natural Neutrosophic Language Processing is a concept that combines the principles of Neutrosophic Language and Natural Language Processing. Definitions and related theorems are provided below.

Definition 3.19 (Neutrosophic Language)**.** Let Σ be a finite alphabet. A *Neutrosophic Language* over Σ^* is a function:

$$N : \Sigma^* \to [0,1]^3,$$

where for each word $w \in \Sigma^*$, $N(w) = (T(w), I(w), F(w))$ with $T(w), I(w), F(w) \in [0,1]$ and

$$0 \leq T(w) + I(w) + F(w) \leq 3.$$

Here:

- $T(w)$ represents the *truth-membership degree* of w.
- $I(w)$ represents the *indeterminacy-membership degree* of w.
- $F(w)$ represents the *falsity-membership degree* of w.

A Neutrosophic Language generalizes the notion of membership beyond the single membership function of a fuzzy language by explicitly incorporating degrees of truth, indeterminacy, and falsity.

Example 3.20 (Neutrosophic Language Example 1: the word "balanced")**.** In a neutrosophic language $N : \Sigma^* \to [0,1]^3$, each word is assigned three values $(T(w), I(w), F(w))$ representing truth, indeterminacy, and falsity. Consider using words to describe political opinions on a new policy. Let the word "balanced" describe the policy's approach. Different individuals may view this policy variably:

$$N(\text{"balanced"}) = (T(\text{"balanced"}), I(\text{"balanced"}), F(\text{"balanced"})) = (0.5, 0.4, 0.1).$$

This could mean:

- T("balanced") = 0.5 : A moderate portion of people find this description appropriate.
- I("balanced") = 0.4 : There is notable uncertainty or disagreement about whether "balanced" is the right term.
- F("balanced") = 0.1 : A small fraction strongly disagrees that the policy is balanced.

Similarly, consider a weather description like "fair." Some days are clearly sunny or rainy, but "fair" might carry uncertainty:

$$N(\text{"fair"}) = (0.7, 0.2, 0.1),$$

implying the weather is mostly agreeable ($T = 0.7$), somewhat uncertain or ill-defined ($I = 0.2$), and only rarely considered an incorrect descriptor ($F = 0.1$).

These examples illustrate how fuzzy and neutrosophic languages model the gradations and uncertainties present in everyday natural language usage.

Example 3.21 (Neutrosophic Language Example 2: Product Descriptions)**.** Consider

$$\Sigma^* = \{\text{"reliable"}, \text{"controversial"}\}$$

describing products. Define a Neutrosophic Language $N : \Sigma^* \to [0, 1]^3$:

$$N(\text{"reliable"}) = (0.6, 0.3, 0.1), \quad N(\text{"controversial"}) = (0.3, 0.4, 0.3).$$

"Reliable" has a moderate truth value, some uncertainty, and a low falsity, reflecting mostly positive but not unanimous opinions. "Controversial" has lower truth, higher uncertainty, and increased falsity, representing mixed and polarized views.

Example 3.22 (Neutrosophic Language Example 3: Information Accuracy)**.** Let $\Sigma^* = \{\text{"accurate"}, \text{"misleading"}\}$ represent terms describing information reliability. Define $N : \Sigma^* \to [0, 1]^3$:

$$N(\text{"accurate"}) = (0.8, 0.1, 0.1), \quad N(\text{"misleading"}) = (0.2, 0.5, 0.3).$$

"Accurate" predominantly conveys correctness with minimal uncertainty or falsity. "Misleading" shows considerable uncertainty (0.5) and a non-negligible falsity score (0.3), indicating that not everyone views this word as fitting the truth, and there is notable disagreement about its appropriateness.

Theorem 3.23 (Neutrosophic Language generalizes Fuzzy Language)**.** *Every fuzzy language is a special case of a neutrosophic language.*

Proof. A Fuzzy Language $F : \Sigma^* \to [0, 1]$ assigns to each word w a single membership value $F(w) \in [0, 1]$.

Consider a Neutrosophic Language $N : \Sigma^* \to [0, 1]^3$ with $N(w) = (T(w), I(w), F(w))$. If we restrict ourselves to the case where:

$$I(w) = 0 \quad \text{and} \quad F(w) = 0,$$

then $T(w)$ alone determines the membership, and we have:

$$N(w) = (T(w), 0, 0).$$

If we set $T(w) = F(w)$ from the fuzzy language, the neutrosophic language reduces exactly to the given fuzzy language. Hence, fuzzy languages are included as a special case of neutrosophic languages. □

Definition 3.24 (Natural Neutrosophic Language)**.** A *Natural Neutrosophic Language* incorporates the notion of indeterminacy into the modeling of natural language, extending beyond the single membership function of a fuzzy language. Let Σ be a finite alphabet. A Natural Neutrosophic Language $\mathcal{L}_N$ is defined as:

$$\mathcal{L}_N = (\Sigma, \mathcal{M}_N, \mathcal{P}_N, \mathcal{S}_N),$$

where:

- Σ: A finite alphabet representing the set of words.
- $\mathcal{M}_N : \Sigma^* \to [0,1]^3$: A neutrosophic membership function assigning to each word $w \in \Sigma^*$ a triplet $(T(w), I(w), F(w))$, where $T(w)$ is the truth-membership degree, $I(w)$ the indeterminacy-membership degree, and $F(w)$ the falsity-membership degree. These values satisfy:

$$0 \leq T(w) + I(w) + F(w) \leq 3.$$

  This triple encodes more nuanced linguistic uncertainty. For example, a word might be considered partly true, partly indeterminate, and partly false in a given linguistic context.
- $\mathcal{P}_N : \Sigma^* \times \Sigma^* \to [0,1]^3$: A neutrosophic relation that assigns to each pair (u, v) a triplet representing truth, indeterminacy, and falsity of their semantic similarity or contextual relation.
- $\mathcal{S}_N$: A set of syntactic or semantic rules represented as neutrosophic constraints. These rules can express degrees of truth, uncertainty, and contradiction in sentence formation and interpretation.

Example 3.25 (Natural Neutrosophic Language)**.** Consider a vocabulary $\Sigma = \{\text{"balanced"}, \text{"fair"}, \text{"complex"}\}$. In a Natural Neutrosophic Language $\mathcal{L}_N$, we might have:

$$\mathcal{M}_N(\text{"balanced"}) = (T(\text{"balanced"}), I(\text{"balanced"}), F(\text{"balanced"})) = (0.7, 0.2, 0.1),$$

indicating that "balanced" is considered 70% true, 20% indeterminate, and 10% false within a certain context (e.g., describing a policy that is mostly fair but not universally agreed upon).

A neutrosophic relation could be defined as:

$$\mathcal{P}_N(\text{"balanced"}, \text{"complex"}) = (0.5, 0.3, 0.2),$$

suggesting that "balanced" and "complex" have a moderate truth-related similarity (0.5), a noticeable indeterminacy (0.3), and a small element of falsity (0.2) in their relationship. Neutrosophic syntactic rules might allow certain sentences to be formed that express ambiguous or partially contradictory meanings, reflecting the nuanced and often uncertain nature of human language.

Example 3.26 (Natural Neutrosophic Language in a Japanese Linguistic Context)**.** Consider a Natural Neutrosophic Language $\mathcal{L}_N$ with a vocabulary

$\Sigma = \{$"teinei (polite)" [291], "bimyou (subtle or questionable)" [78, 107], "tekitou (appropriate, sometimes careless)" [300]$\}$

. These words represent various subjective qualities in communication or behavior.

Define a neutrosophic membership function $\mathcal{M}_N : \Sigma^* \to [0,1]^3$. For instance:

$$\mathcal{M}_N(\text{"bimyou"}) = (T(\text{"bimyou"}), I(\text{"bimyou"}), F(\text{"bimyou"})) = (0.4, 0.4, 0.2).$$

The word "bimyou (subtle or questionable)" in Japanese often conveys subtlety, uncertainty, or something that is "not clearly good or bad." Here:

- $T(\text{"bimyou"}) = 0.4$: There is some sense of truth or correctness in calling something "bimyou."
- $I(\text{"bimyou"}) = 0.4$: A high indeterminacy reflects the ambiguity and difficulty in categorizing "bimyou" definitively.
- $F(\text{"bimyou"}) = 0.2$: There is a small falsity component, recognizing that some may find "bimyou" to be clearly one way or another.

In contrast, consider "teinei (polite)." We might have:

$$\mathcal{M}_N(\text{"teinei"}) = (0.7, 0.1, 0.2),$$

indicating a general consensus that "teinei" is positively true (0.7), with low uncertainty (0.1) and a small degree of falsity (0.2), accounting for contexts where someone might consider an action "not truly polite."

A neutrosophic relation $\mathcal{P}_N : \Sigma^* \times \Sigma^* \to [0,1]^3$ could represent how words relate:

$$\mathcal{P}_N(\text{"teinei"}, \text{"tekitou"}) = (0.3, 0.5, 0.2),$$

suggesting that "teinei (polite)" and "tekitou (appropriate but sometimes careless)" share some conceptual ground (0.3 truth), a large area of uncertainty (0.5 indeterminacy), and a small falsity component (0.2).

Neutrosophic syntactic or semantic rules could allow sentences to reflect partial truth, uncertainty, and contradiction. For example, describing someone's behavior as "teinei da ga bimyou (polite but questionable)" might yield a neutrosophic membership indicating partial agreement, substantial uncertainty, and some degree of falsity regarding the nature of the politeness. This highlights how the Japanese linguistic context can emphasize nuanced, context-dependent meanings effectively captured by neutrosophic language modeling.

Definition 3.27 (Natural Neutrosophic Language Processing (NNLP))**.** A *Natural Neutrosophic Language Processing (NNLP)* system is a tuple:

$$\mathcal{N}^N = (\Sigma, \mathcal{L}^N, \mathcal{P}^N, \mathcal{M}^N, \mathcal{T}^N),$$

where:

1. Σ: A finite alphabet.
2. $\mathcal{L}^N \subseteq \Sigma^*$: A language with neutrosophic membership.
3. $\mathcal{P}^N : \mathcal{L}^N \to [0,1]^3$: A neutrosophic membership function giving $(T(w), I(w), F(w))$ for each w.
4. $\mathcal{M}^N : \mathcal{L}^N \to O$: A mapping function from words to structured outputs.
5. $\mathcal{T}^N : \mathcal{L}^N \times \mathcal{L}^N \to \mathbb{R}$: A similarity measure under neutrosophic membership.

Theorem 3.28. *Every Natural Fuzzy Language Processing system is a special case of a Natural Neutrosophic Language Processing system.*

Proof. A fuzzy membership assigns $\mathcal{P}^F(w) \in [0,1]$. A neutrosophic membership assigns $(T(w), I(w), F(w)) \in [0,1]^3$.

By setting $I(w) = 0$ and $F(w) = 0$, we have:

$$\mathcal{P}^N(w) = (T(w), 0, 0).$$

If we identify $T(w) = \mathcal{P}^F(w)$, the neutrosophic model reduces to the fuzzy model. Thus, NNLP generalizes NFLP. □

Theorem 3.29. *Natural Neutrosophic Language Processing (NNLP) inherently possesses the structure of a Neutrosophic Language.*

Proof. By definition, an NNLP system $\mathcal{N}^N = (\Sigma, \mathcal{L}^N, \mathcal{P}^N, \mathcal{M}^N, \mathcal{T}^N)$ includes:

- A finite alphabet Σ.
- A language $\mathcal{L}^N \subseteq \Sigma^*$ with neutrosophic membership.
- A membership function $\mathcal{P}^N : \mathcal{L}^N \to [0,1]^3$, assigning to each word $w \in \mathcal{L}^N$ a triplet $\mathcal{P}^N(w) = (T(w), I(w), F(w))$, where:
$$0 \leq T(w) + I(w) + F(w) \leq 3.$$
This aligns exactly with the definition of a Neutrosophic Language, where $T(w)$, $I(w)$, and $F(w)$ represent the degrees of truth, indeterminacy, and falsity, respectively.

Moreover, the functions $\mathcal{M}^N$ and $\mathcal{T}^N$ provide additional structure to NNLP, supporting tasks such as mapping words to structured outputs and computing similarity under neutrosophic membership. While these components extend NNLP's applicability, they do not alter the foundational neutrosophic membership structure.

Thus, $\mathcal{L}^N$ equipped with $\mathcal{P}^N$ satisfies all the conditions of a Neutrosophic Language:

$$N(w) = (T(w), I(w), F(w)) \quad \text{for all } w \in \mathcal{L}^N.$$

Therefore, an NNLP system $\mathcal{N}^N$ possesses the structure of a Neutrosophic Language. □

3.1.4 Natural Plithogenic Language Processing

Natural Plithogenic Language Processing is a concept that combines the principles of Plithogenic Language and Natural Language Processing. Relevant definitions and theorems are presented below. As briefly mentioned in the introduction, the Plithogenic concept is particularly advantageous due to its flexibility in defining the number of parameters related to uncertainty. This flexibility makes it a promising framework for various applications, and the authors believe it will inspire extensive research in the future.

Definition 3.30 (Plithogenic Language)**.** Consider a *Plithogenic Set* $PS = (P, v, Pv, pdf, pCF)$ as defined in [261, 262], where:

- P is a subset of a universal set S.
- v is an attribute.
- Pv is the range of possible values for attribute v.
- $pdf : P \times Pv \to [0, 1]^s$ is the Degree of Appurtenance Function (DAF).
- $pCF : Pv \times Pv \to [0, 1]^t$ is the Degree of Contradiction Function (DCF), satisfying $pCF(a, a) = 0$ and $pCF(a, b) = pCF(b, a)$ for all $a, b \in Pv$.

A *Plithogenic Language* over Σ^* (with parameters s, t) is a function:

$$PL : \Sigma^* \to [0, 1]^s,$$

such that the membership vector assigned to each word $w \in \Sigma^*$ is determined by the plithogenic structure (via pdf and influenced by pCF). The values of s and t define the dimensionality of the membership and contradiction degrees.

Example 3.31 (Examples of Plithogenic Languages)**.** (cf. [83, 96])

The following examples illustrate Plithogenic languages, categorized based on the parameters s and t of the associated Plithogenic set:

- When $s = t = 1$, the language corresponds to a *Plithogenic Fuzzy Language*.
- When $s = 2, t = 1$, the language corresponds to a *Plithogenic Intuitionistic Fuzzy Language*.
- When $s = 3, t = 1$, the language corresponds to a *Plithogenic Neutrosophic Language*.
- When $s = 4, t = 1$, the language corresponds to a *Plithogenic Quadripartitioned Neutrosophic Language* (cf. [138, 228, 249]).
- When $s = 5, t = 1$, the language corresponds to a *Plithogenic Pentapartitioned Neutrosophic Language* (cf. [26, 55, 183]).
- When $s = 6, t = 1$, the language corresponds to a *Plithogenic Hexapartitioned Neutrosophic Language* (cf. [211]).
- When $s = 7, t = 1$, the language corresponds to a *Plithogenic Heptapartitioned Neutrosophic Language* (cf. [34, 198]).
- When $s = 8, t = 1$, the language corresponds to a *Plithogenic Octapartitioned Neutrosophic Language*.
- When $s = 9, t = 1$, the language corresponds to a *Plithogenic Nonapartitioned Neutrosophic Language*.

Theorem 3.32 (Plithogenic Language generalizes Neutrosophic and Fuzzy Languages)**.** *Consider a Plithogenic Language* $PL : \Sigma^* \to [0, 1]^s$ *with a contradiction function* $pCF : Pv \times Pv \to [0, 1]^t$.

1. *For $s = 3$ and $t = 1$, a plithogenic language can represent a neutrosophic language, as the triple $(T(w), I(w), F(w))$ of neutrosophic membership degrees fits into the plithogenic framework by interpreting the three-dimensional membership (and a single-dimension contradiction) appropriately.*
2. *For $s = 1$ and $t = 1$, a plithogenic language reduces to a fuzzy language scenario. In this case, we have essentially one membership dimension and a single contradiction dimension that can be fixed, resulting in a structure identical to a fuzzy language.*

Proof. When $s = 3, t = 1$, choose the plithogenic structure so that the three-dimensional membership vector $[0, 1]^3$ corresponds to $(T(w), I(w), F(w))$ of a neutrosophic language. The single contradiction dimension $t = 1$ can represent additional uncertainty, but can also be fixed if needed. Thus, neutrosophic languages are embedded within the plithogenic framework.

When $s = 1, t = 1$, the plithogenic language reduces to a single membership dimension and one contradiction dimension. By setting the contradiction dimension appropriately and ignoring it or treating it as a constant, we get a single membership value per word, which corresponds exactly to the definition of a fuzzy language. Thus, fuzzy languages are obtained as a special case of plithogenic languages.

Hence, plithogenic languages generalize both neutrosophic languages (when $s = 3, t = 1$) and fuzzy languages (when $s = 1, t = 1$). □

Definition 3.33 (Natural Plithogenic Language)**.** A *Natural Plithogenic Language* integrates the principles of plithogenic sets into the modeling of natural language, capturing complex, multi-attribute uncertainty and contradictions within linguistic expressions. Formally, let Σ be a finite alphabet representing the vocabulary. A Natural Plithogenic Language $\mathcal{L}_{PL}$ is defined as:

$$\mathcal{L}_{PL} = (\Sigma, \mathcal{M}_{PL}, \mathcal{P}_{PL}, \mathcal{S}_{PL}, pdf, pCF),$$

where:

- Σ: A finite alphabet representing the set of words.
- $pdf : P \times Pv \to [0, 1]^s$: The Degree of Appurtenance Function (DAF) from a Plithogenic Set, assigning multi-dimensional membership degrees to elements (words) with respect to given attributes.
- $pCF : Pv \times Pv \to [0, 1]^t$: The Degree of Contradiction Function (DCF), quantifying contradictions between possible attribute values. This allows the model to handle conflicting attributes inherent in language interpretation.
- $\mathcal{M}_{PL} : \Sigma^* \to [0, 1]^s$: A plithogenic membership function assigning an s-dimensional membership vector to each word $w \in \Sigma^*$. Each dimension represents a particular attribute of uncertainty, such as truth, indeterminacy, falsity, or more complex measures.
- $\mathcal{P}_{PL} : \Sigma^* \times \Sigma^* \to [0, 1]^s$: A plithogenic relation measuring semantic proximity or contextual similarity between words. This relation can incorporate contradictions among attributes, reflecting the nuanced relationships found in natural language.
- $\mathcal{S}_{PL}$: A set of syntactic or semantic rules formulated as plithogenic constraints. These rules manage how words combine to form phrases and sentences, capturing complex patterns of agreement, contradiction, and multi-faceted meaning.

By integrating plithogenic concepts, a Natural Plithogenic Language generalizes and extends frameworks such as fuzzy or neutrosophic languages. It represents a comprehensive model that can handle multiple attributes and their contradictions, providing a richer and more flexible representation of the inherent complexity and ambiguity in natural human language.

Example 3.34. Consider a vocabulary $\Sigma = \{\text{"equitable"}, \text{"ambiguous"}, \text{"temporal"}\}$. Suppose we are interested in multiple attributes such as truthfulness, cultural specificity, temporal stability, and potential contradiction among these attributes. The plithogenic membership function $\mathcal{M}_{PL}$ might assign to the word "ambiguous" a vector:

$$\mathcal{M}_{PL}(\text{"ambiguous"}) = (0.5, 0.4, 0.1),$$

where these three components could represent degrees of truth, uncertainty, and another attribute capturing cultural dependency, respectively. The pdf and pCF functions would be defined to handle how attribute values map onto membership degrees and contradictions. For instance, if two words share similar cultural attributes but differ strongly in temporal stability, the plithogenic relation $\mathcal{P}_{PL}(\text{"equitable"}, \text{"temporal"})$ might yield a vector indicating partial similarity in some attributes and high contradiction in others.

Thus, a Natural Plithogenic Language allows for a nuanced, multi-dimensional modeling of words and their interactions, capturing the layered and often contradictory qualities of natural language usage.

Definition 3.35 (Natural Plithogenic Language Processing (NPLP))**.** A *Natural Plithogenic Language Processing (NPLP)* system is a tuple:

$$\mathcal{N}^{PL} = (\Sigma, \mathcal{L}^{PL}, \mathcal{P}^{PL}, \mathcal{M}^{PL}, \mathcal{T}^{PL}),$$

where:

1. Σ: A finite alphabet.
2. $\mathcal{L}^{PL} \subseteq \Sigma^*$: A language with plithogenic membership.
3. $\mathcal{P}^{PL} : \mathcal{L}^{PL} \to [0,1]^s$: A plithogenic membership function.
4. $\mathcal{M}^{PL} : \mathcal{L}^{PL} \to O$: A mapping function from words to structured outputs.
5. $\mathcal{T}^{PL} : \mathcal{L}^{PL} \times \mathcal{L}^{PL} \to \mathbb{R}$: A similarity measure under plithogenic membership.

Theorem 3.36 (NPLP generalizes NNLP and NFLP under specific parameters)**.** *The following properties hold for Natural Plithogenic Language Processing.*

1. *For $s = 3$ and $t = 1$, a Natural Plithogenic Language Processing system reduces to a Natural Neutrosophic Language Processing system.*
2. *For $s = 1$ and $t = 1$, a Natural Plithogenic Language Processing system reduces to a Natural Fuzzy Language Processing system.*

Proof. NPLP to NNLP (s=3, t=1): Choose $s = 3$, giving a three-dimensional membership vector $(T(w), I(w), F(w))$. By suitably defining the plithogenic structure (e.g., selecting pdf and pCF functions), we obtain a triple analogous to neutrosophic membership. With $s = 3, t = 1$, the NPLP framework mirrors NNLP exactly.

NPLP to NFLP (s=1, t=1): For $s = 1$, we have a single membership dimension, akin to a fuzzy membership. The additional $t = 1$ contradiction dimension can be fixed or simplified, leaving a single scalar membership. Thus, the NPLP model collapses to an NFLP model.

Hence, depending on the parameter choices (s, t), NPLP generalizes both NNLP and NFLP. □

3.2 Large Uncertain Language Model

A Large Language Model (LLM) is an AI system trained on extensive text datasets to understand, generate, and process human language [23, 51, 133, 208, 314]. This subsection discusses the concept of the Large Uncertain Language Model.

3.2.1 Classic Large Language Model

The definition of a Classic Large Language Model is provided below [23, 51, 133, 208, 314]. Readers seeking more detailed information are encouraged to refer to introductory notes or surveys as needed [136, 147, 348].

Definition 3.37 (Classic Large Language Model)**.** (cf. [23, 51, 133, 208, 314]) Let Σ be a finite alphabet representing the vocabulary. Let Σ^* denote the set of all finite sequences (strings) over Σ. A *Large Language Model (LLM)* is a probabilistic model designed to predict the likelihood of sequences in Σ^* and perform linguistic tasks. Mathematically, an LLM is defined as:

$$\mathcal{M}_{\text{LLM}} = (\Sigma, \mathcal{P}, \mathcal{T}, \mathcal{G}, \mathcal{O}),$$

where:

1. *Vocabulary*: Σ is the set of tokens (words, characters, or subwords).
2. *Probability Distribution*: $\mathcal{P} : \Sigma^* \to [0, 1]$ is the probability distribution over sequences, defined such that for any $w \in \Sigma^*$,
$$\mathcal{P}(w) = \prod_{t=1}^{|w|} P(w_t \mid w_1, w_2, \dots, w_{t-1}),$$
where w_t is the t-th token of w, and $P(w_t \mid w_1, w_2, \dots, w_{t-1})$ is the conditional probability of w_t given its preceding tokens.
3. *Training Process*: $\mathcal{T}$ represents the training procedure, optimizing the parameters θ of the model to minimize the negative log-likelihood:
$$\mathcal{L}(\theta) = -\frac{1}{N} \sum_{i=1}^{N} \log \mathcal{P}_\theta(w^{(i)}),$$
where $\{w^{(i)}\}_{i=1}^{N} \subset \Sigma^*$ is the training dataset.
4. *Model Architecture*: $\mathcal{G}$ defines the architecture (e.g., Transformer networks), which maps the input tokens to embeddings and computes the conditional probabilities.
5. *Output Space*: $\mathcal{O} \subseteq \Sigma^*$ is the output space of generated sequences, which can include completions, translations, or answers to queries.

Libraries and open source tools for Classic Large Language Models have been studied in works such as [27,47,100,117,165,173,178,204,207,247,295,297,311,315,327,342]. While not exhaustive, these references may be useful as needed.

3.2.2 Large Fuzzy Language Model

A Large Fuzzy Language Model is a definition that integrates the concept of fuzzy language into LLMs. The definitions of the Large Fuzzy Language Model are provided below. It is anticipated that practical research on these models will advance in the future.

Definition 3.38 (Large Fuzzy Language Model (LFLM))**.** Let Σ be a finite alphabet, and Σ^* the set of all finite words over Σ. Consider a fuzzy language $r : \Sigma^* \to [0, 1]$, as defined previously. A *Large Fuzzy Language Model (LFLM)* is defined as a tuple:

$$\mathcal{M}_{\text{LFLM}} = (\Sigma, \mathcal{P}^F, \mathcal{T}^F, \mathcal{G}^F, \mathcal{O}^F),$$

where:

1. Σ: A finite alphabet.
2. $\mathcal{P}^F : \Sigma^* \to [0,1]$ is a fuzzy membership function assigning to each word $w \in \Sigma^*$ a degree of membership $\mathcal{P}^F(w) \in [0,1]$.
3. $\mathcal{T}^F$: A training procedure that adjusts parameters θ to fit observed (word, membership) pairs, aiming to improve consistency with a given fuzzy linguistic environment.
4. $\mathcal{G}^F$: The model architecture (e.g., a neural network) capable of encoding and decoding fuzzy membership contexts.
5. $O^F \subseteq \Sigma^*$: The output space of generated sequences, where each candidate output w is associated with a fuzzy membership degree $\mathcal{P}^F(w)$.

Unlike a standard LLM that uses probabilities, an LFLM uses a fuzzy membership function to represent how well a word fits certain linguistic criteria, without requiring a normalization constraint as in probability distributions.

Theorem 3.39. *The Large Fuzzy Language Model (LFLM) generalizes the Classic Large Language Model (LLM).*

Proof. A Classic Large Language Model $\mathcal{M}_{\text{LLM}} = (\Sigma, \mathcal{P}, \mathcal{T}, \mathcal{G}, O)$ assigns a probability distribution $\mathcal{P} : \Sigma^* \to [0,1]$ to sequences. This probability distribution satisfies:

$$\sum_{w \in \Sigma^*} \mathcal{P}(w) = 1,$$

indicating normalization across all possible sequences.

In contrast, an LFLM $\mathcal{M}_{\text{LFLM}} = (\Sigma, \mathcal{P}^F, \mathcal{T}^F, \mathcal{G}^F, O^F)$ assigns a fuzzy membership function $\mathcal{P}^F : \Sigma^* \to [0,1]$ to sequences. The fuzzy membership function $\mathcal{P}^F(w)$ represents the degree to which the sequence w belongs to a specific linguistic context, without requiring normalization:

$$\sum_{w \in \Sigma^*} \mathcal{P}^F(w) \leq 1.$$

For any normalized probability distribution $\mathcal{P}(w)$, we can define an equivalent fuzzy membership function $\mathcal{P}^F(w) = \mathcal{P}(w)$, where $\mathcal{P}^F$ satisfies the fuzzy membership property. Thus, every LLM can be seen as a special case of an LFLM, where the fuzzy membership degrees are constrained by normalization. Hence, LFLM generalizes LLM. □

Theorem 3.40. *The Large Fuzzy Language Model (LFLM) possesses the structure of a Fuzzy Language.*

Proof. By definition, a fuzzy language is a function $r : \Sigma^* \to [0,1]$, where $r(w)$ represents the degree of membership of a word w in the language. An LFLM defines a fuzzy membership function $\mathcal{P}^F : \Sigma^* \to [0,1]$, where $\mathcal{P}^F(w)$ denotes how well the word w fits certain linguistic criteria. This directly aligns with the definition of a fuzzy language.

Additionally, the support of the fuzzy membership function $\mathcal{P}^F$ is:

$$\text{supp}(\mathcal{P}^F) = \{w \in \Sigma^* \mid \mathcal{P}^F(w) > 0\},$$

which corresponds to the well-formed sequences in the fuzzy linguistic context. Therefore, $\mathcal{M}_{\text{LFLM}}$ inherently satisfies the properties of a fuzzy language. □

3.2.3 Large Neutrosophic Language Model

The Large Neutrosophic Language Model is a concept that integrates the principles of Large Language Models and Neutrosophic Language. Definitions and relevant theorems are provided below.

Definition 3.41 (Large Neutrosophic Language Model (LNLM))**.** Let Σ be a finite alphabet, and consider a Neutrosophic Language $N : \Sigma^* \to [0, 1]^3$, where for each $w \in \Sigma^*$,

$$N(w) = (T(w), I(w), F(w)), \quad 0 \leq T(w) + I(w) + F(w) \leq 3.$$

A *Large Neutrosophic Language Model (LNLM)* is defined as:

$$\mathcal{M}_{\text{LNLM}} = (\Sigma, \mathcal{P}^N, \mathcal{T}^N, \mathcal{G}^N, \mathcal{O}^N),$$

where:

1. Σ: A finite alphabet.
2. $\mathcal{P}^N : \Sigma^* \to [0, 1]^3$ is the neutrosophic membership function that assigns to each w a triple $(T(w), I(w), F(w))$.
3. $\mathcal{T}^N$: A training procedure to fit parameters θ to observed data with neutrosophic membership annotations.
4. $\mathcal{G}^N$: The model architecture that can handle the tripartite membership representation.
5. $\mathcal{O}^N$: The output space of generated sequences, each associated with a neutrosophic membership triple.

Theorem 3.42 (LNLM generalizes LFLM)**.** *Every Large Fuzzy Language Model is a special case of a Large Neutrosophic Language Model.*

Proof. A Large Fuzzy Language Model uses $\mathcal{P}^F : \Sigma^* \to [0, 1]$. Consider an LNLM with $\mathcal{P}^N : \Sigma^* \to [0, 1]^3$. If we restrict the neutrosophic membership to:

$$I(w) = 0, \quad F(w) = 0,$$

then:

$$\mathcal{P}^N(w) = (T(w), 0, 0).$$

Set $T(w) = \mathcal{P}^F(w)$ from the fuzzy model. Under this restriction, the LNLM reduces exactly to the LFLM. Hence, the LNLM framework generalizes the LFLM. □

3.2.4 Large Plithogenic Language Model

The Large Plithogenic Language Model is a concept that combines the principles of Large Language Models and Plithogenic Language. Relevant theorems and definitions are provided below.

Definition 3.43 (Large Plithogenic Language Model (LPLM))**.** Consider a Plithogenic Language $PL : \Sigma^* \to [0, 1]^s$ defined with parameters s, t, as per the plithogenic set structure. A *Large Plithogenic Language Model (LPLM)* is defined as:

$$\mathcal{M}_{\text{LPLM}} = (\Sigma, \mathcal{P}^{PL}, \mathcal{T}^{PL}, \mathcal{G}^{PL}, \mathcal{O}^{PL}),$$

where:

1. Σ: A finite alphabet.
2. $\mathcal{P}^{PL} : \Sigma^* \to [0, 1]^s$ assigns to each $w \in \Sigma^*$ an s-dimensional membership vector derived from a plithogenic structure, potentially influenced by a contradiction function $pCF : Pv \times Pv \to [0, 1]^t$.
3. $\mathcal{T}^{PL}$: A training procedure to learn parameters θ for the plithogenic framework.

4. $\mathcal{G}^{PL}$: The model architecture accommodating multi-dimensional membership and contradiction information.

5. O^{PL}: The output space of generated sequences, each associated with a multi-dimensional membership vector.

Theorem 3.44 (LPLM generalizes LNLM and LFLM under specific parameters)**.** *1. For $s = 3$ and $t = 1$, a Large Plithogenic Language Model reduces to a Large Neutrosophic Language Model.*

2. For $s = 1$ and $t = 1$, a Large Plithogenic Language Model reduces to a Large Fuzzy Language Model.

Proof. LPLM to LNLM (s=3, t=1): In the plithogenic setting, each word w is assigned a membership vector in $[0, 1]^s$. For $s = 3$, let $(T(w), I(w), F(w))$ represent these three membership dimensions. By appropriately defining the plithogenic framework (e.g., choosing pdf and pCF functions to mimic neutrosophic conditions), we obtain the same triple structure as a neutrosophic language model. The extra contradiction function dimension $t = 1$ can be fixed or integrated to match neutrosophic constraints. Thus, for $s = 3, t = 1$, the LPLM coincides with the LNLM.

LPLM to LFLM (s=1, t=1): If we set $s = 1$, the plithogenic model assigns a single membership value per word, just like a fuzzy language. The additional $t = 1$ contradiction dimension can be fixed or nullified, leaving a single membership value per word. Hence, the LPLM matches the LFLM structure when $s = 1, t = 1$.

Thus, depending on parameter choices for s and t, the LPLM framework can specialize to LNLM or LFLM, demonstrating that LPLM generalizes both LNLM and LFLM. □

4 Theoretical Considerations of n-Superhyperlanguage

In this section, we explore the theoretical considerations of n-superhyperlanguage. It is important to note that this discussion focuses on theoretical generalizations. Practical feasibility and robustness of these methods for real-world applications require further computational experiments and validation.

4.1 n-Superhyperword and n-Superhyperlanguage

In this subsection, we define the notions of a hyperlanguage and an n-superhyperlanguage. Intuitively, a hyperlanguage [29, 30, 80] generalizes the concept of a language by allowing its elements to be sets of words rather than individual words. We then extend this idea hierarchically to n-superhyperlanguages, which are based on iterated power sets of the set of words(cf. [89, 99]).

Definition 4.1 (Hyperword and Hyperlanguage)**.** [29, 30, 80, 89, 230] Let Σ be a finite alphabet, and let Σ^* denote the set of all finite words over Σ.

1. A *hyperword* over Σ is a nonempty subset of Σ^*. In other words, a hyperword is an element of the power set $\mathcal{P}(\Sigma^*)$.

2. A *hyperlanguage* over Σ is a set of hyperwords over Σ. Thus, a hyperlanguage H is a subset of $\mathcal{P}(\Sigma^*)$. Formally:

$$H \subseteq \mathcal{P}(\Sigma^*).$$

A hyperlanguage can therefore be viewed as a *set of sets of words* over Σ.

Example 4.2 (Hyperword and Hyperlanguage)**.** Consider a large collection of written documents describing various topics—e.g., cooking recipes. Let Σ be an alphabet representing characters, and Σ^* represent all possible words that can appear in these recipes (e.g., “salt,” “tomato,” “roast,” “bake”).

A *hyperword* is a nonempty subset of Σ^*. For instance, consider the following subsets of words:

$$H_1 = \{\text{"tomato"}, \text{"onion"}, \text{"garlic"}\}, \quad H_2 = \{\text{"roast"}, \text{"grill"}, \text{"bake"}\}.$$

Each H_i is a hyperword representing a set of related culinary terms. For example, H_1 might represent ingredients commonly used together, and H_2 might represent cooking methods.

A *hyperlanguage* is a set of hyperwords. Suppose we consider:

$$\mathcal{H} = \{H_1, H_2, H_3, \dots\}$$

where each $H_i \subseteq \Sigma^*$ is a set of words grouped by some semantic or thematic criterion. For instance:

$$H_3 = \{\text{"dessert"}, \text{"pastry"}, \text{"sorbet"}\}.$$

In this scenario, $\mathcal{H}$ could be seen as a collection of ingredient sets, method sets, and category sets, each representing a cluster of related words (e.g., ingredients in H_1, cooking techniques in H_2, and dessert types in H_3). Thus, a hyperlanguage $\mathcal{H}$ is essentially a set of sets of words.

Theorem 4.3. *A hyperword generalizes the concept of a word, and a hyperlanguage generalizes the concept of a language.*

Proof. Let Σ be a finite alphabet, and let Σ^* denote the set of all finite words over Σ.

A word $w \in \Sigma^*$ is a single finite sequence of symbols from Σ. In contrast, a hyperword $W \subseteq \Sigma^*$ is a nonempty subset of Σ^*, allowing for collections of words instead of individual sequences.

For example:

- If $\Sigma = \{a, b\}$, a word w could be $w = aba \in \Sigma^*$.
- A hyperword W could be $W = \{aba, abb\}$, representing a set of words rather than a single sequence.

Clearly, every word $w \in \Sigma^*$ can be viewed as a hyperword by identifying it with the singleton set $\{w\}$. Thus, the set of words Σ^* is embedded in the power set $\mathcal{P}(\Sigma^*)$, and hyperwords generalize words by allowing subsets of Σ^* as elements.

A language $L \subseteq \Sigma^*$ is a subset of words from Σ^*. A hyperlanguage $H \subseteq \mathcal{P}(\Sigma^*)$ is a subset of hyperwords, i.e., a set of sets of words.

For example:

- If $\Sigma = \{a, b\}$, a language L could be $L = \{aba, abb\}$.
- A hyperlanguage H could be $H = \{\{aba\}, \{abb, baa\}\}$, where each element of H is a hyperword, i.e., a subset of Σ^*.

Every language $L \subseteq \Sigma^*$ can be viewed as a hyperlanguage by identifying it with the set of singleton hyperwords $\{\{w\} \mid w \in L\}$. Therefore, hyperlanguages generalize languages by allowing sets of hyperwords as elements. □

Definition 4.4 (n-Superhyperword and n-Superhyperlanguage)**.** [89] We now generalize this construction to multiple levels. Define the iterated power sets as follows:

$$\mathcal{P}^0(\Sigma^*) := \Sigma^*, \quad \mathcal{P}^{k+1}(\Sigma^*) := \mathcal{P}(\mathcal{P}^k(\Sigma^*)), \text{ for all } k \geq 0.$$

1. An *n-superhyperword* over Σ is an element of $\mathcal{P}^n(\Sigma^*)$. In particular:

$$\mathcal{P}^1(\Sigma^*) = \mathcal{P}(\Sigma^*) \text{ consists of hyperwords,}$$

$$\mathcal{P}^2(\Sigma^*) = \mathcal{P}(\mathcal{P}(\Sigma^*)) \text{ consists of sets of hyperwords, and so forth.}$$

2. An *n-superhyperlanguage* over Σ is a subset of $\mathcal{P}^n(\Sigma^*)$. Formally:

$$L \subseteq \mathcal{P}^n(\Sigma^*).$$

Thus, an n-superhyperlanguage is a *set of* $(n-1)$*-superhyperwords*, generalizing the concept of a hyperlanguage to n-th level power sets of words.

Example 4.5 (n-Superhyperlanguage Example)**.** Now consider we want to organize these categories into higher-level groupings. For instance, an n-superhyperlanguage involves iterating the powerset construction multiple times.

- At the first level ($n = 1$), we have hyperwords (sets of words).
- At the second level ($n = 2$), we have sets of hyperwords, i.e., a hyperlanguage.
- At the third level ($n = 3$), we consider sets of hyperlanguages, and so forth.

Let:

$$\mathcal{H}_1 = \{H_1, H_2, H_3\}, \quad \mathcal{H}_2 = \{H_4, H_5, \dots\}$$

where each $\mathcal{H}_i$ is a hyperlanguage. Now, an n-superhyperlanguage with $n = 2$ (often called a superhyperlanguage) could be something like:

$$\mathcal{L}_2 = \{\mathcal{H}_1, \mathcal{H}_2\} \subseteq \mathcal{P}^2(\Sigma^*).$$

In a real-life context, think of this as a hierarchical classification scheme:

- Words represent individual items, such as ingredients, methods, or categories.
- Hyperwords represent thematic clusters of these items (e.g., groupings by similar meaning or usage).
- A hyperlanguage is a collection of such clusters, potentially representing the entire categorization of a domain at one level (e.g., all ingredient sets or all technique sets).
- An n-superhyperlanguage constructs even higher strata of organization, enabling the management of multiple domains and meta-level categories of these clusters.

This hierarchical approach, though conceptual, can reflect real-life complexities where we not only have sets of items but also need to organize sets of these sets at multiple meta-levels.

Theorem 4.6. *For any integer $n \geq 1$, an n-superhyperword generalizes the notion of a hyperword, and an n-superhyperlanguage generalizes the notion of a hyperlanguage.*

Proof. Recall that a hyperword is defined as an element of $\mathcal{P}(\Sigma^*)$, and a hyperlanguage is defined as a subset of $\mathcal{P}(\Sigma^*)$. By construction:

$$\mathcal{P}^1(\Sigma^*) = \mathcal{P}(\Sigma^*),$$

so a hyperword is a 1-superhyperword, and a hyperlanguage is a 1-superhyperlanguage.

For $n > 1$, an n-superhyperword is an element of:

$$\mathcal{P}^n(\Sigma^*) = \mathcal{P}(\mathcal{P}^{n-1}(\Sigma^*)).$$

When $n = 1$, we have $\mathcal{P}^1(\Sigma^*) = \mathcal{P}(\Sigma^*)$, which are exactly the hyperwords. Thus, any n-superhyperword for $n > 1$ belongs to a higher-level power set and can be seen as a collection of $(n-1)$-superhyperwords. Since each $(n-1)$-superhyperword can be traced down to lower levels of iteration until ultimately reaching $\mathcal{P}(\Sigma^*)$ (the hyperwords), the n-superhyperword concept strictly extends that of a hyperword to more complex, iterated structures.

Similarly, a hyperlanguage is a subset of $\mathcal{P}(\Sigma^*)$. By definition, an n-superhyperlanguage is a subset of $\mathcal{P}^n(\Sigma^*)$:

$$L \subseteq \mathcal{P}^n(\Sigma^*).$$

For $n = 1$, we obtain exactly the definition of a hyperlanguage. For $n > 1$, an n-superhyperlanguage is a collection of $(n-1)$-superhyperwords, each of which is one level more complex than a hyperword. This iterative construction therefore generalizes a hyperlanguage to higher levels, where instead of sets of hyperwords, we consider sets of sets of $(n-1)$-superhyperwords, and so forth.

In conclusion, the n-superhyperword and n-superhyperlanguage structures arise naturally by iterating the power set operation multiple times. Since hyperwords and hyperlanguages correspond to the $n = 1$ case, increasing n yields increasingly higher-order generalizations of these concepts. □

Theorem 4.7. *Every hyperword and hyperlanguage can be represented by an appropriate hyperstructure. More precisely:*

1. *Let Σ be a finite alphabet, and let Σ^* be the set of all finite words over Σ. A* hyperword *is an element of $\mathcal{P}(\Sigma^*)$, and a* hyperlanguage *is a subset of $\mathcal{P}(\Sigma^*)$. Both hyperwords and hyperlanguages can be modeled using hyperstructures of suitable form.*
2. *Similarly, for any $n \geq 1$, every n-superhyperword (an element of $\mathcal{P}^n(\Sigma^*)$) and every n-superhyperlanguage (a subset of $\mathcal{P}^n(\Sigma^*)$) can be represented by an n-superhyperstructure.*

Proof. Consider the base set $S = \Sigma^*$. A hyperstructure is defined as:

$$\mathcal{H} = (\mathcal{P}(S), \circ) = (\mathcal{P}(\Sigma^*), \circ),$$

where $\circ$ is an operation defined on subsets of Σ^*.

By definition, a *hyperword* is an element of $\mathcal{P}(\Sigma^*)$. Thus, each hyperword $W \subseteq \Sigma^*$ is simply an element of the ground set $\mathcal{P}(\Sigma^*)$ of the hyperstructure $\mathcal{H}$. In other words, hyperwords correspond directly to the elements of $\mathcal{P}(S)$ in the hyperstructure.

A *hyperlanguage* H is a subset of $\mathcal{P}(\Sigma^*)$. Observe that $H \subseteq \mathcal{P}(\Sigma^*)$ means $H \in \mathcal{P}(\mathcal{P}(\Sigma^*))$, i.e., H is an element of the second power set of Σ^*. If we set $S' = \mathcal{P}(\Sigma^*)$, then:

$$\mathcal{P}(S') = \mathcal{P}(\mathcal{P}(\Sigma^*)) \quad \text{and} \quad H \in \mathcal{P}(S').$$

Thus, by considering a hyperstructure whose base set is $S' = \mathcal{P}(\Sigma^*)$:

$$\mathcal{H}' = (\mathcal{P}(S'), \circ) = (\mathcal{P}(\mathcal{P}(\Sigma^*)), \circ),$$

we find that any hyperlanguage H is an element of the ground set of $\mathcal{H}'$. Hence, hyperlanguages can be represented within a hyperstructure constructed at the second power set level.

In summary, hyperwords correspond to elements of a hyperstructure defined on Σ^*, and hyperlanguages correspond to elements of a hyperstructure defined on $\mathcal{P}(\Sigma^*)$.

The notion of n-superhyperwords and n-superhyperlanguages generalizes this construction to higher levels of iterated power sets. For $n \geq 1$, we define:

$$\mathcal{P}^0(\Sigma^*) = \Sigma^*, \quad \mathcal{P}^{k+1}(\Sigma^*) = \mathcal{P}(\mathcal{P}^k(\Sigma^*)) \text{ for all } k \geq 0.$$

An n-superhyperword is an element of $\mathcal{P}^n(\Sigma^*)$, and an n-superhyperlanguage is a subset of $\mathcal{P}^n(\Sigma^*)$.

Consider the n-superhyperstructure:

$$\mathcal{SH}_n = (\mathcal{P}_n(S), \circ),$$

where $S = \Sigma^*$ and $\mathcal{P}_n(\Sigma^*) = \mathcal{P}^n(\Sigma^*)$.

By construction, $\mathcal{P}^n(\Sigma^*)$ serves as the ground set of the n-superhyperstructure. Thus:

$$\mathcal{SH}_n = (\mathcal{P}^n(\Sigma^*), \circ).$$

Since an n-superhyperword is an element of $\mathcal{P}^n(\Sigma^*)$, it directly corresponds to an element of the ground set of $\mathcal{SH}_n$. Likewise, an n-superhyperlanguage is a subset of $\mathcal{P}^n(\Sigma^*)$, hence an element of $\mathcal{P}(\mathcal{P}^n(\Sigma^*))$. By replacing the base set S with $\mathcal{P}^n(\Sigma^*)$ and constructing a hyperstructure at the next level, we ensure that n-superhyperlanguages can also be represented by a suitable $(n+1)$-level construction if needed.

Therefore, n-superhyperwords and n-superhyperlanguages naturally align with the concept of n-superhyperstructures, generalizing the relationship established for hyperwords and hyperlanguages. □

4.2 Natural HyperLanguage Processing and n-superhyperlanguage Processing

We now define Natural Hyperlanguage Processing, which extends NLP to operate on hyperlanguages rather than languages.

Definition 4.8 (Natural Hyperlanguage Processing (NHP))**.** Let Σ be a finite alphabet, and let $\mathcal{H} \subseteq \mathcal{P}(\Sigma^*)$ be a hyperlanguage (a set of sets of words).

A Natural Hyperlanguage Processing system is a tuple:

$$\mathcal{N}^{HL} = (\Sigma, \mathcal{H}, \mathcal{P}^{HL}, \mathcal{M}^{HL}, \mathcal{T}^{HL}),$$

where:

1. Σ: A finite alphabet.
2. $\mathcal{H} \subseteq \mathcal{P}(\Sigma^*)$: A hyperlanguage.
3. $\mathcal{P}^{HL} : \mathcal{H} \to [0,1]$: A probability model assigning probabilities to *hyperwords* $H \in \mathcal{H}$.
4. $\mathcal{M}^{HL} : \mathcal{H} \to O$: A mapping function transforming each hyperword $H \in \mathcal{H}$ into a structured output $o \in O$.
5. $\mathcal{T}^{HL} : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$: A similarity measure defined between pairs of hyperwords.

Theorem 4.9. *Natural Hyperlanguage Processing (NHP) generalizes Natural Language Processing (NLP).*

Proof. Consider an NHP system $\mathcal{N}^{HL} = (\Sigma, \mathcal{H}, \mathcal{P}^{HL}, \mathcal{M}^{HL}, \mathcal{T}^{HL})$ where $\mathcal{H} \subseteq \mathcal{P}(\Sigma^*)$.

If we restrict $\mathcal{H}$ so that every hyperword is a singleton set, i.e., for every $H \in \mathcal{H}$, $H = \{w\}$ for some $w \in \Sigma^*$, then there is a bijection between hyperwords in $\mathcal{H}$ and words in a language $\mathcal{L} \subseteq \Sigma^*$.

Under this restriction:

$$\mathcal{H} \cong \mathcal{L}, \quad \text{with } H = \{w\} \leftrightarrow w.$$

In this case, $\mathcal{N}^{HL}$ reduces to:

$$(\Sigma, \mathcal{L}, \mathcal{P}^{HL}, \mathcal{M}^{HL}, \mathcal{T}^{HL}),$$

which is structurally identical to the NLP definition $(\Sigma, \mathcal{L}, \mathcal{P}, \mathcal{M}, \mathcal{T})$.

Thus, NLP is a special case of NHP, proving that NHP generalizes NLP. □

We further generalize to n-superhyperlanguages.

Definition 4.10 (Natural n-Superhyperlanguage Processing (NnSHP))**.** Let Σ be a finite alphabet, and let $\mathcal{H}^{(n)} \subseteq \mathcal{P}^n(\Sigma^*)$ be an n-superhyperlanguage.

A Natural n-Superhyperlanguage Processing system is a tuple:

$$\mathcal{N}^{(n)} = (\Sigma, \mathcal{H}^{(n)}, \mathcal{P}^{(n)}, \mathcal{M}^{(n)}, \mathcal{T}^{(n)}),$$

where:

1. Σ: A finite alphabet.
2. $\mathcal{H}^{(n)} \subseteq \mathcal{P}^n(\Sigma^*)$: An n-superhyperlanguage.
3. $\mathcal{P}^{(n)} : \mathcal{H}^{(n)} \to [0,1]$: A probability model assigning probabilities to n-superhyperwords.
4. $\mathcal{M}^{(n)} : \mathcal{H}^{(n)} \to O$: A mapping function from n-superhyperwords to structured outputs.

5. $\mathcal{T}^{(n)} : \mathcal{H}^{(n)} \times \mathcal{H}^{(n)} \to \mathbb{R}$: A similarity measure on n-superhyperwords.

Theorem 4.11. *Natural n-Superhyperlanguage Processing (NnSHP) generalizes both NLP and NHP.*

Proof. By definition, an n-superhyperlanguage $\mathcal{H}^{(n)} \subseteq \mathcal{P}^n(\Sigma^*)$.

For $n = 1$, we have $\mathcal{H}^{(1)} \subseteq \mathcal{P}(\Sigma^*)$, which is a hyperlanguage. Thus, an N1SHP system:

$$\mathcal{N}^{(1)} = (\Sigma, \mathcal{H}^{(1)}, \mathcal{P}^{(1)}, \mathcal{M}^{(1)}, \mathcal{T}^{(1)})$$

coincides with an NHP system:

$$\mathcal{N}^{HL} = (\Sigma, \mathcal{H}, \mathcal{P}^{HL}, \mathcal{M}^{HL}, \mathcal{T}^{HL}).$$

Hence, NHP is a special case of NnSHP at $n = 1$.

From Theorem 4.9, we know NHP generalizes NLP. Since NnSHP generalizes NHP, it also generalizes NLP. Concretely, by setting $n = 1$ and then restricting hyperwords to singletons, we recover the NLP scenario.

Thus, NnSHP includes both NHP and NLP as special cases, proving that NnSHP generalizes both NLP and NHP. □

Question 4.12. Is it possible to define a Natural Plithogenic n-Superhyperlanguage Processing? What potential applications could it have?

Question 4.13. What are the properties of Fuzzy n-Superhyperlanguage, Neutrosophic n-Superhyperlanguage, Fuzzy Hyperlanguage, Neutrosophic Hyperlanguage, and Plithogenic Hyperlanguage? Additionally, what are their potential applications and operations?

4.3 Large Hyperlanguage Model and Large SuperhyperLanguage Model

We define the *Large HyperLanguage Model* and the *Large SuperHyperLanguage Model* as theoretical generalizations of the Large Language Model. These models are generalizations achieved by incorporating the concepts of HyperLanguage and SuperHyperLanguage. Future studies are expected to explore computational experiments, practical implementation methods, and diverse applications of these models.

Definition 4.14 (Large Hyperlanguage Model (LHLM))**.** Let Σ be a finite alphabet, and let $\mathcal{H} \subseteq \mathcal{P}(\Sigma^*)$ be a hyperlanguage. A *Large Hyperlanguage Model (LHLM)* is a probabilistic model that assigns probabilities to hyperwords (elements of $\mathcal{H}$) and supports processing tasks analogous to those of an LLM, but at the hyperword level. Formally, an LHLM is defined as:

$$\mathcal{M}_{\text{LHLM}} = (\Sigma, \mathcal{H}, \mathcal{P}^{HL}, \mathcal{T}^{HL}, \mathcal{G}^{HL}, \mathcal{O}^{HL}),$$

where:

1. Σ: A finite alphabet of tokens.
2. $\mathcal{H} \subseteq \mathcal{P}(\Sigma^*)$: A hyperlanguage, i.e., a set of hyperwords.
3. $\mathcal{P}^{HL} : \mathcal{H} \to [0, 1]$: A probability distribution over hyperwords $H \in \mathcal{H}$. For any hyperword $H = \{w_1, w_2, \dots, w_k\}$, we define:
$$\mathcal{P}^{HL}(H) = P_\theta(H),$$
where P_θ is parameterized by θ and may factorize over the words in H or utilize more complex dependencies.
4. $\mathcal{T}^{HL}$: The training procedure, adjusting parameters θ to fit observed collections of hyperwords drawn from $\mathcal{H}$.
5. $\mathcal{G}^{HL}$: The model architecture (e.g., a hyperword-level Transformer) that processes sets of words simultaneously or in a structured manner.

6. O^{HL}: The output space, consisting of hyperwords or structured objects derived from hyperwords.

Theorem 4.15. *A Large Hyperlanguage Model (LHLM) generalizes a Large Language Model (LLM).*

Proof. A Large Language Model (LLM) $\mathcal{M}_{\text{LLM}} = (\Sigma, \mathcal{P}, \mathcal{T}, \mathcal{G}, O)$ assigns probabilities to words (elements of Σ^*).

Consider an LHLM $\mathcal{M}_{\text{LHLM}} = (\Sigma, \mathcal{H}, \mathcal{P}^{HL}, \mathcal{T}^{HL}, \mathcal{G}^{HL}, O^{HL})$ where $\mathcal{H} \subseteq \mathcal{P}(\Sigma^*)$.

If we restrict every hyperword $H \in \mathcal{H}$ to be a singleton set, i.e., $H = \{w\}$ for some $w \in \Sigma^*$, then there is a one-to-one correspondence between hyperwords and individual words. Under this restriction:

$$\mathcal{H} \cong \mathcal{L}, \quad \text{with } H = \{w\} \leftrightarrow w.$$

Thus, $\mathcal{P}^{HL}$ reduces to $\mathcal{P}$, $\mathcal{T}^{HL}$ reduces to $\mathcal{T}$, $\mathcal{G}^{HL}$ reduces to $\mathcal{G}$, and O^{HL} reduces to O, recovering the exact structure of an LLM.

Hence, LLMs are a special case of LHLMs, proving that LHLM generalizes LLM. □

Definition 4.16 (Large n-Superhyperlanguage Model (LnSHM))**.** Let Σ be a finite alphabet, and let $\mathcal{H}^{(n)} \subseteq \mathcal{P}^n(\Sigma^*)$ be an n-superhyperlanguage. A *Large n-Superhyperlanguage Model (LnSHM)* is defined analogously, but operates over n-superhyperwords. Formally:

$$\mathcal{M}_{\text{SH}}^{(n)} = (\Sigma, \mathcal{H}^{(n)}, \mathcal{P}^{(n)}, \mathcal{T}^{(n)}, \mathcal{G}^{(n)}, O^{(n)}),$$

where:

1. Σ: A finite alphabet.
2. $\mathcal{H}^{(n)} \subseteq \mathcal{P}^n(\Sigma^*)$: An n-superhyperlanguage.
3. $\mathcal{P}^{(n)} : \mathcal{H}^{(n)} \to [0,1]$: A probability distribution over n-superhyperwords.
4. $\mathcal{T}^{(n)}$: The training process to learn parameters θ from data structured as n-superhyperwords.
5. $\mathcal{G}^{(n)}$: The model architecture capable of processing n-superhyperwords.
6. $O^{(n)}$: The output space, potentially consisting of even higher-order structures derived from n-superhyperwords.

Theorem 4.17. *A Large n-Superhyperlanguage Model (LnSHM) generalizes both LHLM and LLM.*

Proof. An LnSHM operates over an n-superhyperlanguage $\mathcal{H}^{(n)} \subseteq \mathcal{P}^n(\Sigma^*)$.

1. From LnSHM to LHLM: For $n = 1$, an n-superhyperlanguage reduces to a hyperlanguage. Thus, setting $n = 1$:

$$\mathcal{H}^{(1)} \subseteq \mathcal{P}(\Sigma^*),$$

the LnSHM becomes an LHLM. Hence, LHLMs are a special case of LnSHMs.

2. From LHLM to LLM: From Theorem 4.15, we know that LHLM generalizes LLM. Since LnSHM generalizes LHLM, it follows by transitivity that LnSHM also generalizes LLM.

Therefore, LnSHM includes both LHLM and LLM as special cases. □

Question 4.18. Is it possible to perform model extensions using hypergeometric probability [172, 239]?

Question 4.19. Is it possible to define a Large Plithogenic n-Superhyperlanguage Model? What potential applications could it have?

5 Future Direction: Hyperprobability and n-SuperHyperprobability

This section outlines the future directions of this research. Studies on probability spaces [37,63,127,201,206] are closely related to the findings presented in this work. Here, we define the concepts of Hyperprobability [42] and n-SuperHyperprobability. Future research is expected to investigate their potential applications to the diverse language models and probabilistic frameworks discussed in this paper.

Notation 5.1. *Let Ω be a finite sample space, and let $\mathcal{P}(\Omega)$ denote the powerset of Ω, i.e., the set of all subsets of Ω.*

Definition 5.2 (Hyperprobability Space)**.** (cf. [38–42]) A *hyperprobability space* is a triple $(\Omega, \mathcal{P}(\Omega), P_H)$, where:

- Ω is the sample space.
- $\mathcal{P}(\Omega)$ is the set of events (subsets of Ω).
- $P_H : \mathcal{P}(\Omega) \to [0, 1]$ is a hyperprobability measure satisfying the following properties:
  1. $P_H(\emptyset) = 0$ and $P_H(\Omega) = 1$,
  2. For any disjoint $A, B \in \mathcal{P}(\Omega)$:
  $$P_H(A \cup B) = P_H(A) + P_H(B).$$

To generalize hyperprobability to higher orders, we recursively define the powerset operation and construct higher-level probability spaces.

Definition 5.3 (n-SuperHyperprobability Space)**.** 1. Define the n-th powerset recursively:

$$\mathcal{P}^0(\Omega) := \Omega, \quad \mathcal{P}^{k+1}(\Omega) := \mathcal{P}(\mathcal{P}^k(\Omega)) \quad \text{for all } k \geq 0.$$

2. An *n-superhyperprobability space* is a triple $(\Omega, \mathcal{P}^n(\Omega), P_{SH})$, where:
   - Ω is the sample space.
   - $\mathcal{P}^n(\Omega)$ is the set of n-th level events.
   - $P_{SH} : \mathcal{P}^n(\Omega) \to [0, 1]$ is the n-superhyperprobability measure satisfying:
     (a) $P_{SH}(\emptyset) = 0$ and $P_{SH}(\mathcal{P}^n(\Omega)) = 1$,
     (b) For any disjoint $A, B \in \mathcal{P}^n(\Omega)$:
     $$P_{SH}(A \cup B) = P_{SH}(A) + P_{SH}(B).$$

Remark 5.4. The definitions of *Hyperprobability* and *n-SuperHyperprobability* extend classical probability theory by leveraging powerset structures. To ensure their mathematical validity:

- *Consistency with Classical Probability:* For $n = 0$, P_{SH} reduces to a classical probability measure on Ω, satisfying $P_{SH}(\emptyset) = 0$, $P_{SH}(\Omega) = 1$, and additivity.
- *Iterative Construction:* The recursive definition of $\mathcal{P}^n(\Omega)$ ensures that each level adheres to the axioms of probability, extending the additivity property to higher-order sets.
- *Higher-Order Events:* The inclusion of n-th powerset structures allows the modeling of layered uncertainties, providing a mathematically rigorous framework for higher-order probabilistic reasoning.

Thus, the definitions are consistent with the axioms of probability and are mathematically robust.

Example 5.5 (Probability Example 1: Hyperprobability for Weather Events)**.** Let $\Omega = \{\text{Sunny}, \text{Rainy}, \text{Cloudy}\}$ represent possible weather states.

- The powerset $\mathcal{P}(\Omega)$ consists of all subsets of Ω, e.g., $\{\emptyset, \{\text{Sunny}\}, \{\text{Rainy}\}, \{\text{Sunny}, \text{Rainy}\}, \dots\}$.

- A hyperprobability measure P_H might assign:

$$P_H(\{\text{Sunny}\}) = 0.5, \quad P_H(\{\text{Rainy}\}) = 0.3, \quad P_H(\{\text{Sunny, Rainy}\}) = 0.2.$$

- The additivity property ensures $P_H(\emptyset) = 0$ and $P_H(\Omega) = 1$.

Example 5.6 (Probability Example 2: n-SuperHyperprobability for Risk Assessment)**.** Consider a financial system where $\Omega = \{\text{Low Risk, Moderate Risk, High Risk}\}$.

- At $n = 1$, hyperprobabilities might assign likelihoods to events such as:

$$P_H(\{\text{Low Risk, Moderate Risk}\}) = 0.6, \quad P_H(\{\text{High Risk}\}) = 0.4.$$

- At $n = 2$, the second powerset $\mathcal{P}^2(\Omega)$ includes sets of hyperprobabilities, such as:

$$\mathcal{P}^2(\Omega) = \{\{\{\text{Low Risk}\}, \{\text{Moderate Risk}\}\}, \dots\}.$$

- An $n = 2$ superhyperprobability measure P_{SH} might evaluate:

$$P_{SH}(\{\{\{\text{Low Risk}\}, \{\text{Moderate Risk}\}\}\}) = 0.7.$$

Example 5.7 (Probability Example 3: Quantum State Representation)**.** In quantum mechanics, let $\Omega = \{\psi_1, \psi_2, \psi_3\}$ represent quantum states.

- At $n = 1$, hyperprobabilities might describe probabilities of quantum state superpositions:

$$P_H(\{\psi_1, \psi_2\}) = 0.8, \quad P_H(\{\psi_3\}) = 0.2.$$

- At $n = 2$, P_{SH} can model probabilities over sets of such probabilities, capturing nested uncertainties in quantum measurements.

Question 5.8. What definitions would emerge if this concept were applied to language models, natural language processing, neural networks, and AI? Additionally, would any improvements be observed?

Funding

No external financial support was provided for this study.

Acknowledgments

We sincerely thank all individuals whose guidance and encouragement have been instrumental in the successful completion of this research. We also extend our gratitude to readers for their interest in this work. Additionally, we acknowledge the authors of the cited references, whose foundational contributions have greatly enriched our study.

Data Availability

This paper is purely theoretical and mathematical. Consequently, no data analysis was performed. Future researchers are encouraged to explore related data analyses or empirical investigations as necessary.

Ethical Approval

This study focuses exclusively on theoretical and mathematical research, involving no human participants or animals.

Conflicts of Interest

The authors declare that there are no conflicts of interest concerning the publication of this study.

Disclaimer

This study centers on theoretical advancements and has not undergone practical testing or application. Future studies are encouraged to validate and refine the methods through empirical research. While efforts have been made to ensure accuracy and proper citation, unintentional errors or omissions may occur. Readers are advised to independently verify the referenced materials. The interpretations and views expressed herein are solely those of the authors and do not necessarily reflect the perspectives of their affiliated institutions.

References

[1] Mujahid Abbas, Ghulam Murtaza, and Florentin Smarandache. *Basic operations on hypersoft sets and hypersoft point*. Infinite Study, 2020.

[2] Mohamed Abdel-Basset, Mohamed El-Hoseny, Abduallah Gamal, and Florentin Smarandache. A novel model for evaluation hospital medical care systems based on plithogenic sets. *Artificial intelligence in medicine*, 100:101710, 2019.

[3] Mohamed Abdel-Basset, Mai Mohamed, Mohamed Elhoseny, Le Hoang Son, Francisco Chiclana, and Abdel Nasser H. Zaied. Cosine similarity measures of bipolar neutrosophic set for diagnosis of bipolar disorder diseases. *Artificial intelligence in medicine*, 101:101735, 2019.

[4] Zhao Aiwu and Guan Hongjun. Fuzzy-valued linguistic soft set theory and multi-attribute decision-making application. *Chaos Solitons & Fractals*, 89:2–7, 2016.

[5] Muhammad Akram, A Nagoor Gani, and A Borumand Saeid. Vague hypergraphs. *Journal of Intelligent & Fuzzy Systems*, 26(2):647–653, 2014.

[6] Muhammad Akram, Danish Saleem, and Talal Al-Hawary. Spherical fuzzy graphs with application to decision-making. *Mathematical and Computational Applications*, 25(1):8, 2020.

[7] Muhammad Akram, Shumaiza, and José Carlos Rodríguez Alcantud. Multi-criteria decision making methods with bipolar fuzzy sets. *Forum for Interdisciplinary Mathematics*, 2023.

[8] José Carlos R Alcantud, Azadeh Zahedi Khameneh, Gustavo Santos-García, and Muhammad Akram. A systematic literature review of soft set theory. *Neural Computing and Applications*, 36(16):8951–8975, 2024.

[9] Ebrahem Ateatullah Algehyne, Muhammad Lawan Jibril, Naseh A Algehainy, Osama Abdulaziz Alamri, and Abdullah Khaled J Alzahrani. Fuzzy neural network expert system with an improved gini index random forest-based feature importance measure algorithm for early diagnosis of breast cancer in saudi arabia. *Big Data Cogn. Comput.*, 6:13, 2022.

[10] M Irfan Ali, Feng Feng, Xiaoyan Liu, Won Keun Min, and Muhammad Shabir. On some new operations in soft set theory. *Computers & Mathematics with Applications*, 57(9):1547–1553, 2009.

[11] Mumtaz Ali, Luu Quoc Dat, Le Hoang Son, and Florentin Smarandache. Interval complex neutrosophic set: Formulation and applications in decision-making. *International Journal of Fuzzy Systems*, 20:986 – 999, 2017.

[12] Mumtaz Ali and Florentin Smarandache. Complex neutrosophic set. *Neural Computing and Applications*, 28:1817–1834, 2016.

[13] Lilas Alrahis and Ozgur Sinanoglu. Graph neural networks for hardware vulnerability analysis-can you trust your gnn? In *2023 IEEE 41st VLSI Test Symposium (VTS)*, pages 1–4. IEEE, 2023.

[14] Reza Ameri and Morteza Norouzi. On commutative hyperrings. *International journal of algebraic hyperstructures and its applications*, 1(1):45–58, 2014.

[15] Shahzaib Ashraf, Saleem Abdullah, Tahir Mahmood, Fazal Ghani, and Tariq Mahmood. Spherical fuzzy sets and their applications in multi-attribute decision making problems. *J. Intell. Fuzzy Syst.*, 36:2829–2844, 2019.

[16] Krassimir T. Atanassov. Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 20:87–96, 1986.

[17] Krassimir T. Atanassov. On interval valued intuitionistic fuzzy sets. *Interval-Valued Intuitionistic Fuzzy Sets*, 2019.

[18] Krassimir T Atanassov. Circular intuitionistic fuzzy sets. *Journal of Intelligent & Fuzzy Systems*, 39(5):5981–5986, 2020.

[19] Krassimir T. Atanassov. New topological operator over intuitionistic fuzzy sets. 2023.

[20] Michael Francis Atiyah and Ian G. MacDonald. Introduction to commutative algebra. 1969.

[21] Ansar Aynetdinov and Alan Akbik. Semscore: Automated evaluation of instruction-tuned llms based on semantic textual similarity. *ArXiv*, abs/2401.17072, 2024.

[22] Breck Baldwin and Krishna Dayanidhi. Natural language processing with java and lingpipe cookbook. 2014.

[23] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the dangers of stochastic parrots: Can language models be too big. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pages 610–623, 2021.

[24] Claude Berge. *Hypergraphs: combinatorics of finite sets*, volume 45. Elsevier, 1984.

[25] Steven Bird, Ewan Klein, and Edward Loper. *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.", 2009.

[26] Pranab Biswas, Surapati Pramanik, and Bibhas Chandra Giri. Single valued bipolar pentapartitioned neutrosophic set and its application in madm strategy. 2022.

[27] Sid Black, Stella Biderman, Eric Hallahan, Quentin G. Anthony, Leo Gao, Laurence Golding, Horace He, Connor Leahy, Kyle McDonell, Jason Phang, Michael Martin Pieler, USVSN Sai Prashanth, Shivanshu Purohit, Laria Reynolds, Jonathan Tow, Benqi Wang, and Samuel Weinbach. Gpt-neox-20b: An open-source autoregressive language model. *ArXiv*, abs/2204.06745, 2022.

[28] FERNANDEZ A BLANCO, Maria R Casals, Ana Colubi, Norberto Corral, BARZANA M GARCIA, MA Gil, RODRIGUEZ G GONZALEZ, MT López, M Montenegro, MA Lubiano, et al. Random fuzzy sets: a mathematical tool to develop statistical fuzzy data analysis. 2013.

[29] Borzoo Bonakdarpour and Sarai Sheinvald. Automata for hyperlanguages. *arXiv preprint arXiv:2002.09877*, 2020.

[30] Borzoo Bonakdarpour and Sarai Sheinvald. Finite-word hyperlanguages. *Information and Computation*, 295:104944, 2023.

[31] Igor Borja. An introduction to automata design with tikz's automata library. *TUGboat*, 2023.

[32] Alain Bretto. Hypergraph theory. *An introduction. Mathematical Engineering. Cham: Springer*, 1, 2013.

[33] Eric Brill. Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging. *Computational linguistics*, 21(4):543–565, 1995.

[34] S Broumi and Tomasz Witczak. Heptapartitioned neutrosophic soft set. *International Journal of Neutrosophic Science*, 18(4):270–290, 2022.

[35] Said Broumi and Florentin Smarandache. Several similarity measures of neutrosophic sets. *Infinite Study*, 410(1), 2013.

[36] Said Broumi, Selcuk Topal, Assia Bakali, Mohamed Talea, and Florentin Smarandache. A novel python toolbox for single and interval-valued neutrosophic matrices. 2020.

[37] Jacob Burbea and C. Radhakrishna Rao. Entropy differential metric, distance and divergence measures in probability spaces: A unified approach. *Journal of Multivariate Analysis*, 12:575–596, 1982.

[38] M. Burgin. *Functional Algebra and Hypercalculus in Infinite Dimensions: Hyperintegrals, Hyperfunctionals and Hyperderivatives*. Functional Algebra and Hypercalculus in Infinite Dimensions, 2017.

[39] M. Burgin and A. Krinik. Introduction to conditional hyperprobabilities. *Integration*, 2(3):285, 2011.

[40] Mark Burgin. Integrating random properties and the concept of probability. *Integration*, 3(2):137–155, 2012.

[41] Mark Burgin and Alex Krinik. Hyperexpectations and random variables without expectations. *Integration*, 3(3):245, 2012.

[42] Mark Burgin and Alex C. Krinik. Probabilities and hyperprobabilities. In *8th Annual International Conference on Statistics, Mathematics and Related Fields, Conference Proceedings*, pages 351–367, Honolulu, Hawaii, 2009.

[43] Humberto Bustince and P Burillo. Vague sets are intuitionistic fuzzy sets. *Fuzzy sets and systems*, 79(3):403–405, 1996.

[44] Jia Syuen Chai, Ganeshsree Selvachandran, Florentin Smarandache, Vassilis C Gerogiannis, Le Hoang Son, Quang-Thinh Bui, and Bay Vo. New similarity measures for single-valued neutrosophic sets with applications in pattern recognition and medical diagnosis problems. *Complex & Intelligent Systems*, 7:703–723, 2021.

[45] Juanjuan Chen, Shenggang Li, Shengquan Ma, and Xueping Wang. m-polar fuzzy sets: an extension of bipolar fuzzy sets. *The scientific world journal*, 2014(1):416530, 2014.

[46] Liuxin Chen, Yutai Wang, and Dongmei Yang. Picture fuzzy z-linguistic set and its application in multiple attribute group decision-making. *J. Intell. Fuzzy Syst.*, 43:5997–6011, 2022.

[47] Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Pondé, Jared Kaplan, Harrison Edwards, Yura Burda, Nicholas Joseph, Greg Brockman, Alex Ray, Raul Puri, Gretchen Krueger, Michael Petrov, Heidy Khlaaf, Girish Sastry, Pamela Mishkin, Brooke Chan, Scott Gray, Nick Ryder, Mikhail Pavlov, Alethea Power, Lukasz Kaiser, Mohammad Bavarian, Clemens Winter, Philippe Tillet, Felipe Petroski Such, David W. Cummings, Matthias Plappert, Fotios Chantzis, Elizabeth Barnes, Ariel Herbert-Voss, William H. Guss, Alex Nichol, Igor Babuschkin, Suchir Balaji, Shantanu Jain, Andrew Carr, Jan Leike, Joshua Achiam, Vedant Misra, Evan Morikawa, Alec Radford, Matthew M. Knight, Miles Brundage, Mira Murati, Katie Mayer, Peter Welinder, Bob McGrew, Dario Amodei, Sam McCandlish, Ilya Sutskever, and Wojciech Zaremba. Evaluating large language models trained on code. *ArXiv*, abs/2107.03374, 2021.

[48] Shyi-Ming Chen. Measures of similarity between vague sets. *Fuzzy sets and Systems*, 74(2):217–223, 1995.

[49] Eli Chien, Chao Pan, Jianhao Peng, and Olgica Milenkovic. You are allset: A multiset function framework for hypergraph neural networks. *ArXiv*, abs/2106.13264, 2021.

[50] KR1442 Chowdhary and KR Chowdhary. Natural language processing. *Fundamentals of artificial intelligence*, pages 603–649, 2020.

[51] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, et al. Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research*, 24(240):1–113, 2023.

[52] Ronan Collobert and Jason Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, pages 160–167, 2008.

[53] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. Natural language processing (almost) from scratch. *Journal of machine learning research*, 12:2493–2537, 2011.

[54] Bui Cong Cuong and Pham Van Hai. Some fuzzy logic operators for picture fuzzy sets. *2015 Seventh International Conference on Knowledge and Systems Engineering (KSE)*, pages 132–137, 2015.

[55] Suman Das, Rakhal Das, and Binod Chandra Tripathy. Topology on rough pentapartitioned neutrosophic set. *Iraqi Journal of Science*, 2022.

[56] Brian A. Davey and Hilary A. Priestley. Introduction to lattices and order. 1990.

[57] B Davvaz, A Dehghan Nezhad, and A Benvidi. Chemical hyperalgebra: Dismutation reactions. *Match-Communications in Mathematical and Computer Chemistry*, 67(1):55, 2012.

[58] Bijan Davvaz and Azam Salasi. A realization of hyperrings. *Communications in Algebra®*, 34(12):4389–4400, 2006.

[59] Supriya Kumar De, Ranjit Biswas, and Akhil Ranjan Roy. Some operations on intuitionistic fuzzy sets. *Fuzzy sets and Systems*, 114(3):477–484, 2000.

[60] José R González de Mendívil. A generalization of myhill–nerode theorem for fuzzy languages. *Fuzzy Sets and Systems*, 301:103–115, 2016.

[61] José R González de Mendívil and Jose R Garitagoitia. Fuzzy languages with infinite range accepted by fuzzy automata: Pumping lemma and determinization procedure. *Fuzzy Sets and Systems*, 249:1–26, 2014.

[62] Somen Debnath. *Introduction to Hyper Neutrosophic Set*. Infinite Study, 2020.

[63] Freddy Delbaen. Coherent risk measures on general probability spaces. 2002.

[64] Irfan Deli, Mumtaz Ali, and Florentin Smarandache. Bipolar neutrosophic sets and their application based on multi-criteria decision making problems. *2015 International Conference on Advanced Mechatronic Systems (ICAMechS)*, pages 249–254, 2015.

[65] Irfan Deli, Yusuf Subas, Florentin Smarandache, and Mumtaz Ali. Interval valued bipolar neutrosophic sets and their application in pattern recognition. *viXra*, 2016.

[66] Keith Devlin. *The joy of sets: fundamentals of contemporary set theory*. Springer Science & Business Media, 1994.

[67] Giuseppe Di Maio and Somashekhar Naimpally. Comparison of hypertopologies. 1991.

[68] Reinhard Diestel. *Graph theory*. Springer (print edition); Reinhard Diestel (eBooks), 2024.

[69] Ehsan Doostmohammadi, Oskar Holmstrom, and Marco Kuhlmann. How reliable are automatic evaluation methods for instruction-tuned llms? *ArXiv*, abs/2402.10770, 2024.

[70] Danny Driess, Fei Xia, Mehdi SM Sajjadi, Corey Lynch, Aakanksha Chowdhery, Brian Ichter, Ayzaan Wahid, Jonathan Tompson, Quan Vuong, Tianhe Yu, et al. Palm-e: An embodied multimodal language model. *arXiv preprint arXiv:2303.03378*, 2023.

[71] Shervin Le Du, Senaida Hernández Santana, and Giannicola Scarpa. A gentle introduction to quantum natural language processing. *ArXiv*, abs/2202.11766, 2022.

[72] Mateusz Dubiel, Yasmine Barghouti, Kristina Kudryavtseva, and Luis A. Leiva. On-device query intent prediction with lightweight llms to support ubiquitous conversations. *Scientific Reports*, 14, 2024.

[73] Yanxia Dui and Hongchun Hu. Social media public opinion detection using multimodal natural language processing and attention mechanisms. *IET Inf. Secur.*, 2024, 2024.

[74] Nancy El-Hefenawy, Mohamed A Metwally, Zenat M Ahmed, and Ibrahim M El-Henawy. A review on the applications of neutrosophic sets. *Journal of Computational and Theoretical Nanoscience*, 13(1):936–944, 2016.

[75] Ralph Elliott, John RW Glauert, JR Kennaway, and Ian Marshall. The development of language processing support for the visicast project. In *Proceedings of the fourth international ACM conference on Assistive technologies*, pages 101–108, 2000.

[76] Shahzad Faizi etc. Decision making with uncertainty using hesitant fuzzy sets. *International Journal of Fuzzy Systems*, 20:93 – 103, 2017.

[77] Bahram Farhadinia. Correlation for dual hesitant fuzzy sets and dual interval-valued hesitant fuzzy sets. *International Journal of Intelligent Systems*, 29, 2014.

[78] Alviano Aditya Fauzi and Dewi Kusrini. Analysis of japanese adverbs in japanese drama suizokukan gaaru. In *Fifth International Conference on Language, Literature, Culture, and Education (ICOLLITE 2021)*, pages 60–66. Atlantis Press, 2021.

[79] Yifan Feng, Haoxuan You, Zizhao Zhang, R. Ji, and Yue Gao. Hypergraph neural networks. In *AAAI Conference on Artificial Intelligence*, 2018.

[80] Hadar Frenkel and Sarai Sheinvald. Realizable and context-free hyperlanguages. *arXiv preprint arXiv:2209.10306*, 2022.

[81] Caro Fuchs, Simone Spolaor, Marco S. Nobile, and Uzay Kaymak. pyfume: a python package for fuzzy model estimation. *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–8, 2020.

[82] Takaaki Fujita. Note for neutrosophic incidence and threshold graph. *SciNexuses*, 1:97–125, 2024.

[83] Takaaki Fujita. A review of fuzzy and neutrosophic offsets: Connections to some set concepts and normalization function. In *Advancing Uncertain Combinatorics through Graphization, Hyperization, and Uncertainization: Fuzzy, Neutrosophic, Soft, Rough, and Beyond (Third Volume)*. Biblio Publishing, 2024.

[84] Takaaki Fujita. Short note of supertree-width and n-superhypertree-width. *Neutrosophic Sets and Systems*, 77:54–78, 2024.

[85] Takaaki Fujita. Superhypergraph neural networks and plithogenic graph neural networks: Theoretical foundations. *arXiv preprint arXiv:2412.01176*, 2024.

[86] Takaaki Fujita. Uncertain labeling graphs and uncertain graph classes (with survey for various uncertain sets). 2024.

[87] Takaaki Fujita. *Advancing Uncertain Combinatorics through Graphization, Hyperization, and Uncertainization: Fuzzy, Neutrosophic, Soft, Rough, and Beyond*. Biblio Publishing, 2025.

[88] Takaaki Fujita. A comprehensive discussion on fuzzy hypersoft expert, superhypersoft, and indetermsoft graphs. *Neutrosophic Sets and Systems*, 77:241–263, 2025.

[89] Takaaki Fujita. Exploration of graph classes and concepts for superhypergraphs and n-th power mathematical structures. 2025.

[90] Takaaki Fujita. Review of some superhypergraph classes: Directed, bidirected, soft, and rough. 2025.

[91] Takaaki Fujita and Smarandache Florentin. Some graph parameters for superhypertree-width and neutrosophictree-width. In *Advancing Uncertain Combinatorics through Graphization, Hyperization, and Uncertainization: Fuzzy, Neutrosophic, Soft, Rough, and Beyond (Third Volume)*. Biblio Publishing, 2024.

[92] Takaaki Fujita and Florentin Smarandache. Antipodal turiyam neutrosophic graphs. *Neutrosophic Optimization and Intelligent Systems*, 5:1–13, 2024.

[93] Takaaki Fujita and Florentin Smarandache. A concise study of some superhypergraph classes. *Neutrosophic Sets and Systems*, 77:548–593, 2024.

[94] Takaaki Fujita and Florentin Smarandache. Fundamental computational problems and algorithms for superhypergraphs. In *Advancing Uncertain Combinatorics through Graphization, Hyperization, and Uncertainization: Fuzzy, Neutrosophic, Soft, Rough, and Beyond (Second Volume)*. Biblio Publishing, 2024.

[95] Takaaki Fujita and Florentin Smarandache. Introduction to upside-down logic: Its deep relation to neutrosophic logic and applications. In *Advancing Uncertain Combinatorics through Graphization, Hyperization, and Uncertainization: Fuzzy, Neutrosophic, Soft, Rough, and Beyond (Third Volume)*. Biblio Publishing, 2024.

[96] Takaaki Fujita and Florentin Smarandache. A review of the hierarchy of plithogenic, neutrosophic, and fuzzy graphs: Survey and applications. In *Advancing Uncertain Combinatorics through Graphization, Hyperization, and Uncertainization: Fuzzy, Neutrosophic, Soft, Rough, and Beyond (Second Volume)*. Biblio Publishing, 2024.

[97] Takaaki Fujita and Florentin Smarandache. A short note for hypersoft rough graphs. *HyperSoft Set Methods in Engineering*, 3:1–25, 2024.

[98] Takaaki Fujita and Florentin Smarandache. Study for general plithogenic soft expert graphs. *Plithogenic Logic and Computation*, 2:107–121, 2024.

[99] Takaaki Fujita and Florentin Smarandache. Uncertain automata and uncertain graph grammar. *Neutrosophic Sets and Systems*, 74:128–191, 2024.

[100] Luyu Gao, Aman Madaan, Shuyan Zhou, Uri Alon, Pengfei Liu, Yiming Yang, Jamie Callan, and Graham Neubig. Pal: Program-aided language models. *ArXiv*, abs/2211.10435, 2022.

[101] Yue Gao, Yifan Feng, Shuyi Ji, and Rongrong Ji. Hgnn+: General hypergraph neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45:3181–3199, 2022.

[102] Harish Garg. Linguistic interval-valued pythagorean fuzzy sets and their application to multiple attribute group decision-making process. *Cognitive Computation*, 12(6):1313–1337, 2020.

[103] David Gems. The hyperfunction theory: an emerging paradigm for the biology of aging. *Ageing research reviews*, 74:101557, 2022.

[104] Saeed Ghahramani. *Fundamentals of probability*. CRC Press, 2024.

[105] Masoud Ghods, Zahra Rostami, and Florentin Smarandache. Introduction to neutrosophic restricted superhypergraphs and neutrosophic restricted superhypertrees and several of their properties. *Neutrosophic Sets and Systems*, 50:480–487, 2022.

[106] Jayanta Ghosh and Tapas Kumar Samanta. Hyperfuzzy sets and hyperfuzzy group. *Int. J. Adv. Sci. Technol*, 41:27–37, 2012.

[107] Pika Yestia Ginanjar. Enhancing students motivation in japanese conversation (online course): through cm. *IZUMI*, 9(1):31–39, 2020.

[108] Yoav Goldberg. *Neural network methods in natural language processing*. Morgan & Claypool Publishers, 2017.

[109] S Gomathy, D Nagarajan, S Broumi, and M Lathamaheswari. *Plithogenic sets and their application in decision making*. Infinite Study, 2020.

[110] Georg Gottlob, Nicola Leone, and Francesco Scarcello. Hypertree decompositions and tractable queries. In *Proceedings of the eighteenth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 21–32, 1999.

[111] Georg Gottlob, Nicola Leone, and Francesco Scarcello. Hypertree decompositions: A survey. In *Mathematical Foundations of Computer Science 2001: 26th International Symposium, MFCS 2001 Mariánské Lázne, Czech Republic, August 27–31, 2001 Proceedings 26*, pages 37–57. Springer, 2001.

[112] Georg Gottlob and Reinhard Pichler. Hypergraphs in model checking: Acyclicity and hypertree-width versus clique-width. *SIAM Journal on Computing*, 33(2):351–378, 2004.

[113] Hongjun Guan, Shuang Guan, and Aiwu Zhao. Intuitionistic fuzzy linguistic soft sets and their application in multi-attribute decision-making. *J. Intell. Fuzzy Syst.*, 31:2869–2879, 2016.

[114] Muhammad Gulistan, Naveed Yaqoob, Ahmed Elmoasry, and Jawdat Alebraheem. Complex bipolar fuzzy sets: An application in a transport's company. *J. Intell. Fuzzy Syst.*, 40:3981–3997, 2021.

[115] Fatma Kutlu Gündogdu and Cengiz Kahraman. A novel vikor method using spherical fuzzy sets and its application to warehouse site selection. *J. Intell. Fuzzy Syst.*, 37:1197–1211, 2019.

[116] Fatma Kutlu Gündogdu and Cengiz Kahraman. Spherical fuzzy sets and spherical fuzzy topsis method. *J. Intell. Fuzzy Syst.*, 36:337–352, 2019.

[117] Daya Guo, Qihao Zhu, Dejian Yang, Zhenda Xie, Kai Dong, Wentao Zhang, Guanting Chen, Xiao Bi, Yu Wu, Y. K. Li, Fuli Luo, Yingfei Xiong, and Wenfeng Liang. Deepseek-coder: When the large language model meets programming - the rise of code intelligence. *ArXiv*, abs/2401.14196, 2024.

[118] Patrick Haller, Ansar Aynetdinov, and A. Akbik. Opiniongpt: Modelling explicit biases in instruction-tuned llms. In *North American Chapter of the Association for Computational Linguistics*, 2023.

[119] Mohammad Hamidi, Florentin Smarandache, and Elham Davneshvar. Spectrum of superhypergraphs via flows. *Journal of Mathematics*, 2022(1):9158912, 2022.

[120] Mohammad Hamidi and Mohadeseh Taghinezhad. *Application of Superhypergraphs-Based Domination Number in Real World*. Infinite Study, 2023.

[121] Michael A Harrison. *Introduction to formal language theory*. Addison-Wesley Longman Publishing Co., Inc., 1978.

[122] Jun-Yan He, Yifan Wang, Lijun Wang, Huchuan Lu, Jun-Yan He, Jinpeng Lan, Bin Luo, and Xuansong Xie. Multi-modal instruction tuned llms with fine-grained visual perception. *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 13980–13990, 2024.

[123] Wei He and Yiting Dong. Adaptive fuzzy neural network control for a constrained robot using impedance learning. *IEEE Transactions on Neural Networks and Learning Systems*, 29:1174–1186, 2018.

[124] R. Hema, R. Sudharani, and M. Kavitha. A novel approach on plithogenic interval valued neutrosophic hypersoft sets and its application in decision making. *Indian Journal Of Science And Technology*, 2023.

[125] Nasimeh Heydaribeni, Xinrui Zhan, Ruisi Zhang, Tina Eliassi-Rad, and Farinaz Koushanfar. Hypop: Distributed constrained combinatorial optimization leveraging hypergraph neural networks. *ArXiv*, abs/2311.09375, 2023.

[126] Doeko Homan. Very basic set theory. 2023.

[127] Dug Hun Hong and Seok Yoon Hwang. Correlation of intuitionistic fuzzy sets in probability spaces. *Fuzzy Sets Syst.*, 75:77–81, 1995.

[128] Dug Hun Hong and Chul Kim. A note on similarity measures between vague sets and between elements. *Information sciences*, 115(1-4):83–96, 1999.

[129] Zhaoxi Hong, Yixiong Feng, Y. Wang, Bingtao Hu, Zhifeng Zhang, Haoyang Qiu, and Jianrong Tan. Quality characteristic extraction for complex products with multi-granular fuzzy language based on the triple bottom lines of sustainability. *Comput. Ind. Eng.*, 167:107980, 2022.

[130] John E Hopcroft and Jeffrey D Ullman. *Formal languages and their relation to automata*. Addison-Wesley Longman Publishing Co., Inc., 1969.

[131] John E. Hopcroft and Jeffrey D. Ullman. Introduction to automata theory, languages and computation. 1979.

[132] Chao Hu, Ruishi Yu, Binqi Zeng, Yu Zhan, Ying Fu, Quan Zhang, Rongkai Liu, and Heyuan Shi. Hyperattack: Multi-gradient-guided white-box adversarial structure attack of hypergraph neural networks. *ArXiv*, abs/2302.12407, 2023.

[133] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.

[134] Yining Hua, Vi Nguyen, Liqin Wang, Dinah Foer, and Li Zhou. Identifying transgender and gender diverse individuals in electronic health records: A context-aware natural language processing approach. In *American Medical Informatics Association Annual Symposium*, 2022.

[135] Jing Huang and Jie Yang. Unignn: a unified framework for graph and hypergraph neural networks. *ArXiv*, abs/2105.00956, 2021.

[136] Lei Huang, Weijiang Yu, Weitao Ma, Weihong Zhong, Zhangyin Feng, Haotian Wang, Qianglong Chen, Weihua Peng, Xiaocheng Feng, Bing Qin, et al. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 2023.

[137] JE Humphreys. On the hyperalgebra of a semisimple algebraic group. In *Contributions to Algebra*, pages 203–210. Elsevier, 1977.

[138] Satham Hussain, Jahir Hussain, Isnaini Rosyida, and Said Broumi. Quadripartitioned neutrosophic soft graphs. In *Handbook of Research on Advances and Applications of Fuzzy Sets and Logic*, pages 771–795. IGI Global, 2022.

[139] Isao Imai. *Applied hyperfunction theory*, volume 8. Springer Science & Business Media, 2013.

[140] Riccardo Andrea Izzo, Gianluca Bardaro, and Matteo Matteucci. Btgenbot: Behavior tree generation for robotic tasks with lightweight llms. *ArXiv*, abs/2403.12761, 2024.

[141] Muhammad Naveed Jafar, Muhammad Haris Saeed, Kainat Muniba Khan, Faten S. Alamri, and Hamiden A. Wahed Khalifa. Distance and similarity measures using max-min operators of neutrosophic hypersoft sets with application in site selection for solid waste management systems. *IEEE Access*, PP:1–1, 2022.

[142] Sirus Jahanpanah and Roohallah Daneshpayeh. On derived superhyper be-algebras. *Neutrosophic Sets and Systems*, 57(1):21, 2023.

[143] Sirus Jahanpanah and Roohallah Daneshpayeh. An outspread on valued logic superhyperalgebras. *Facta Universitatis, Series: Mathematics and Informatics*, pages 427–437, 2024.

[144] Thomas Jech. *Set theory: The third millennium edition, revised and expanded*. Springer, 2003.

[145] CheonSu Jeong. Fine-tuning and utilization methods of domain-specific llms. *ArXiv*, abs/2401.02981, 2024.

[146] Pu Ji, Jian qiang Wang, and Hong yu Zhang. Frank prioritized bonferroni mean operator with single-valued neutrosophic sets and its application in selecting third-party logistics providers. *Neural Computing and Applications*, 30:799–823, 2018.

[147] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12):1–38, 2023.

[148] Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, and Yongjun Wang. Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4), 2017.

[149] Mariano Jiménez, Mar Arenas Parra, Amelia Bilbao-Terol, and Maria Victoria Rodríguez. Linear programming with fuzzy parameters: An interactive method resolution. *Eur. J. Oper. Res.*, 177:1599–1609, 2007.

[150] Calvin Jongsma. Basic set theory and combinatorics. *Undergraduate Texts in Mathematics*, 2019.

[151] Juan juan Peng, Xin Ge Chen, Xiao Kang Wang, Jian Qiang Wang, Qingqing Long, and Lv Jiang Yin. Picture fuzzy decision-making theories and methodologies: a systematic review. *International Journal of Systems Science*, 54:2663 – 2675, 2023.

[152] Juan juan Peng, Jian qiang Wang, and Xiao hui Wu. Novel multi-criteria decision-making approaches based on hesitant fuzzy sets and prospect theory. *Int. J. Inf. Technol. Decis. Mak.*, 15:621–644, 2016.

[153] Jaiung Jun. Algebraic geometry over hyperrings. *Advances in Mathematics*, 323:142–192, 2018.

[154] Young Bae Jun, Kul Hur, and Kyoung Ja Lee. Hyperfuzzy subalgebras of bck/bci-algebras. *Annals of Fuzzy Mathematics and Informatics*, 2017.

[155] Young Bae Jun, Seok-Zun Song, and Seon Jeong Kim. N-hyper sets. *Mathematics*, 6(6):87, 2018.

[156] Dan Jurafsky and James H. Martin. Book reviews: Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. In *International Conference on Computational Logic*, 2000.

[157] Xiaojun Kang, Xinchuan Li, Hong Yao, Dan Li, Bo Jiang, Xiaoyue Peng, Tiejun Wu, Shihua Qi, and Lijun Dong. Dynamic hypergraph neural networks based on key hyperedges. *Information Sciences*, 616:37–51, 2022.

[158] Faruk Karaaslan and Fatih Hunu. Type-2 single-valued neutrosophic sets and their applications in multi-criteria group decision making based on topsis method. *Journal of Ambient Intelligence and Humanized Computing*, 11:4113 – 4132, 2020.

[159] Abdullah Kargın, Florentin Smarandache, and Memet Şahin. *New Type Hyper Groups, New Type SuperHyper Groups and Neutro-New Type SuperHyper Groups*. Infinite Study, 2023.

[160] Sanjeev Kumar Karn, Rikhiya Ghosh, Polim Kusuma, and Oladimeji Farri. shs-nlp at radsum23: Domain-adaptive pre-training of instruction-tuned llms for radiology report impression generation. In *Workshop on Biomedical Natural Language Processing*, 2023.

[161] Mohamed Khaled. Topics in algebra. *AMS/MAA Textbooks*, 2021.

[162] Ahmed Mostafa Khalil, Shenggang Li, Harish Garg, Hong xia Li, and Sheng quan Ma. New operations on interval-valued picture fuzzy set, interval-valued picture fuzzy soft set and their applications. *IEEE Access*, 7:51236–51253, 2019.

[163] Robert L. Kohl. An introduction to algebraic topology. 2016.

[164] Marc Krasner et al. A class of hyperrings and hyperfields. *International Journal of Mathematics and Mathematical Sciences*, 6(2):307–311, 1983.

[165] Sanjay Kukreja, Tarun Kumar, Amit Purohit, Abhijit Dasgupta, and Debashis Guha. A literature survey on open source large language models. *Proceedings of the 2024 7th International Conference on Computers in Management and Business*, 2024.

[166] Hon Keung Kwan and Yaling Cai. A fuzzy neural network and its application to pattern recognition. *IEEE transactions on Fuzzy Systems*, 2(3):185–193, 1994.

[167] Young-Jou Lai and Ching-Lai Hwang. Fuzzy mathematical programming: Methods and applications. 1995.

[168] Gwénolé Lecorvé and John D. Kelleher. Introduction to the special issue on cross-modal and multimodal natural language processing. In *ICON*, 2022.

[169] Oliver Urs Lenz, Chris Cornelis, and Daniel Peralta. Fuzzy-rough-learn 0.2: a python library for fuzzy rough set algorithms and one-class classification. *2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–8, 2022.

[170] Reiner Lenz and Alan C. Bovik. Group theory. 2007.

[171] Irene Li, Keen You, Xiangru Tang, Yujie Qiao, Lucas Huang, Chia-Chun Hsieh, Benjamin Rosand, and Dragomir R. Radev. Ehrkit: A python natural language processing toolkit for electronic health record texts. *ArXiv*, abs/2204.06604, 2022.

[172] Gerald J Lieberman and Donald Bruce Owen. *Tables of the hypergeometric probability distribution*, volume 3. Stanford University Press Stanford, CA, 1961.

[173] Jiaju Lin, Haoran Zhao, Aochi Zhang, Yiting Wu, Huqiuyue Ping, and Qin Chen. Agentsims: An open-source sandbox for large language model evaluation. *ArXiv*, abs/2308.04026, 2023.

[174] Chengyuan Liu, Shihang Wang, Yangyang Kang, Lizhi Qing, Fubang Zhao, Changlong Sun, Kun Kuang, and Fei Wu. More than catastrophic forgetting: Integrating general capabilities for domain-specific llms. In *Conference on Empirical Methods in Natural Language Processing*, 2024.

[175] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9):1–35, 2023.

[176] Yu-Ting Liu, Yang-Yin Lin, Shang-Lin Wu, Chun-Hsiang Chuang, and Chin-Teng Lin. Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network. *IEEE Transactions on Neural Networks and Learning Systems*, 27:347–360, 2016.

[177] Robert Lowen and Mark Sioen. Proximal hypertopologies revisited. *Set-Valued Analysis*, 6:1–19, 1998.

[178] Yao Lu, Shang Liu, Qijun Zhang, and Zhiyao Xie. Rtllm: An open-source benchmark for design rtl generation with large language model. *2024 29th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pages 722–727, 2023.

[179] Roberto Lucchetti. Hypertopologies and applications. *Recent Developments in Well-Posed Variational Problems*, pages 193–209, 1995.

[180] Xuejiao Ma, Yu Jin, and Qingli Dong. A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting. *Appl. Soft Comput.*, 54:296–312, 2017.

[181] Tahir Mahmood, Kifayat Ullah, Qaisar Khan, and Naeem Jan. An approach toward decision-making and medical diagnosis problems using the concept of spherical fuzzy sets. *Neural Computing and Applications*, pages 1–13, 2019.

[182] Pradip Kumar Maji, Ranjit Biswas, and A Ranjan Roy. Soft set theory. *Computers & mathematics with applications*, 45(4-5):555–562, 2003.

[183] Rama Mallick and Surapati Pramanik. Pentapartitioned neutrosophic set and its properties. *Neutrosophic Sets and Systems*, 35:49, 2020.

[184] Wasim Akram Mandal. Bipolar pythagorean fuzzy sets and their application in multi-attribute decision making problems. *Annals of Data Science*, 10:555–587, 2021.

[185] Christopher D Manning. *Foundations of statistical natural language processing*. The MIT Press, 1999.

[186] Christopher D. Manning and Hinrich Schütze. Book reviews: Foundations of statistical natural language processing. In *International Conference on Computational Logic*, 1999.

[187] Christopher D Manning, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. The stanford corenlp natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, pages 55–60, 2014.

[188] Gonzalo Martínez, José Alberto Hernández, Javier Conde, Pedro Reviriego, and Elena Merino-Gómez. Beware of words: Evaluating the lexical diversity of conversational llms using chatgpt as case study. *ACM Transactions on Intelligent Systems and Technology*, 2024.

[189] Manoj Mathew, Ripon Kumar Chakrabortty, and Michael J. Ryan. A novel approach integrating ahp and topsis under spherical fuzzy sets for advanced manufacturing system selection. *Eng. Appl. Artif. Intell.*, 96:103988, 2020.

[190] William McClure. *Using Japanese: A guide to contemporary usage*. Cambridge University Press, 2000.

[191] Josie McCulloch. Fuzzycreator: A python-based toolkit for automatically generating and analysing data-driven fuzzy sets. *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–6, 2017.

[192] Lingyu Meng and Liangqun Li. Time-sequential hesitant fuzzy set and its application to multi-attribute decision making. *Complex & Intelligent Systems*, 8(5):4319–4338, 2022.

[193] Rada Mihalcea, Hugo Liu, and Henry Lieberman. Nlp (natural language processing) for nlp (natural language programming). In *Conference on Intelligent Text Processing and Computational Linguistics*, 2006.

[194] Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, 267:1–38, 2019.

[195] Mai Mohamed and Asmaa Elsayed. A novel multi-criteria decision making approach based on bipolar neutrosophic set for evaluating financial markets in egypt. *Multicriteria Algorithms with Applications*, 2024.

[196] Dmitriy Molodtsov. Soft set theory-first results. *Computers & mathematics with applications*, 37(4-5):19–31, 1999.

[197] Mitsuo Morimoto. *An introduction to Sato's hyperfunctions*, volume 129. American Mathematical Soc., 1993.

[198] M Myvizhi, Ahmed M Ali, Ahmed Abdelhafeez, and Haitham Rizk Fadlallah. *MADM Strategy Application of Bipolar Single Valued Heptapartitioned Neutrosophic Set*. Infinite Study, 2023.

[199] Prakash M Nadkarni, Lucila Ohno-Machado, and Wendy W Chapman. Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18(5):544–551, 2011.

[200] Somashekhar Naimpally. All hypertopologies are hit-and-miss. *Applied General Topology*, 3(1):45–53, 2002.

[201] Joseph Naor and Moni Naor. Small-bias probability spaces: efficient constructions and applications. In *Symposium on the Theory of Computing*, 1990.

[202] Daniel Nettle. Social scale and structural complexity in human languages. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367:1829 – 1836, 2012.

[203] Gia Nhu Nguyen, Le Hoang Son, Amira S Ashour, and Nilanjan Dey. A survey of the state-of-the-arts on neutrosophic sets in biomedical diagnoses. *International Journal of Machine Learning and Cybernetics*, 10:1–13, 2019.

[204] Erik Nijkamp, Bo Pang, Hiroaki Hayashi, Lifu Tu, Haiquan Wang, Yingbo Zhou, Silvio Savarese, and Caiming Xiong. Codegen: An open large language model for code with multi-turn program synthesis. In *International Conference on Learning Representations*, 2022.

[205] Giorgio Nordo, Saeid Jafari, Arif Mehmood, and Bhimraj Basumatary. A python framework for neutrosophic sets and mappings. *ArXiv*, abs/2404.05735, 2024.

[206] Ferdinand Österreicher and Igor Vajda. A new class of metric divergences on probability spaces and its applicability in statistics. *Annals of the Institute of Statistical Mathematics*, 55:639–653, 2003.

[207] Simon Ott, Konstantin Hebenstreit, Valentin Li'evin, Christoffer Egeberg Hother, Milad Moradi, Maximilian Mayrhauser, Robert Praas, Ole Winther, and Matthias Samwald. Thoughtsource: A central hub for large language model reasoning data. *Scientific Data*, 10, 2023.

[208] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35:27730–27744, 2022.

[209] Gerhard Paass and Sven Giesselbach. Foundation models for natural language processing: Pre-trained language models integrating media. *Foundation Models for Natural Language Processing*, 2023.

[210] Shabaz Patel, Hassan Kane, and Rayhan Patel. Building domain-specific llms faithful to the islamic worldview: Mirage or technical possibility? *ArXiv*, abs/2312.06652, 2023.

[211] Vasile Patrascu. Penta and hexa valued representation of neutrosophic information. *arXiv preprint arXiv:1603.03729*, 2016.

[212] Zdzis law Pawlak. Rough sets. *International journal of computer & information sciences*, 11:341–356, 1982.

[213] Zdzislaw Pawlak. Rough set theory and its applications to data analysis. *Cybernetics & Systems*, 29(7):661–688, 1998.

[214] Zdzis law Pawlak. Rough sets and intelligent data analysis. *Information sciences*, 147(1-4):1–12, 2002.

[215] Zdzis law Pawlak. *Rough sets: Theoretical aspects of reasoning about data*, volume 9. Springer Science & Business Media, 2012.

[216] Zdzislaw Pawlak, Jerzy Grzymala-Busse, Roman Slowinski, and Wojciech Ziarko. Rough sets. *Communications of the ACM*, 38(11):88–95, 1995.

[217] Zdzislaw Pawlak, Lech Polkowski, and Andrzej Skowron. Rough set theory. *KI*, 15(3):38–39, 2001.

[218] Zdzis law Pawlak and Andrzej Skowron. Rudiments of rough sets. *Information sciences*, 177(1):3–27, 2007.

[219] Zdzislaw Pawlak, S. K. Michael Wong, Wojciech Ziarko, et al. Rough sets: probabilistic versus deterministic approach. *International Journal of Man-Machine Studies*, 29(1):81–95, 1988.

[220] Zhiliang Peng, Wenhui Wang, Li Dong, Yaru Hao, Shaohan Huang, Shuming Ma, and Furu Wei. Kosmos-2: Grounding multimodal large language models to the world. *arXiv preprint arXiv:2306.14824*, 2023.

[221] Soujanya Poria, Ong Yew Soon, Bing Liu, and Lidong Bing. Affect recognition for multimodal natural language processing. *Cognitive Computation*, 13:229 – 230, 2020.

[222] Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. Stanza: A python natural language processing toolkit for many human languages. *ArXiv*, abs/2003.07082, 2020.

[223] Jian qiang Wang, Jia ting Wu, Jing Wang, Hong yu Zhang, and Xiao hong Chen. Interval-valued hesitant fuzzy linguistic sets and their applications in multi-criteria decision-making problems. *Inf. Sci.*, 288:55–72, 2014.

[224] Jian qiang Wang, Yu Yang, and Lin Li. Multi-criteria decision-making method based on single-valued neutrosophic linguistic maclaurin symmetric mean operators. *Neural Computing and Applications*, 30:1529 – 1547, 2016.

[225] Atiqe Ur Rahman, Muhammad Haris Saeed, and Hamiden Abd El-Wahed Khalifa. Multi-attribute decision-making based on aggregations and similarity measures of neutrosophic hypersoft sets with possibility setting. *Journal of Experimental & Theoretical Artificial Intelligence*, 36:161 – 186, 2022.

[226] George Rahonis. Fuzzy languages. In *Handbook of Weighted Automata*, pages 481–517. Springer, 2009.

[227] Nadesan Ramaruban. *Commutative hyperalgebra*. University of Cincinnati, 2014.

[228] M Ramya, Sandesh Murali, and R.Radha. Bipolar quadripartitioned neutrosophic soft set. 2022.

[229] Jinjun Rao, Bo Li, Zhen Zhang, Dongdong Chen, and Wojciech Giernacki. Position control of quadrotor uav based on cascade fuzzy neural network. *Energies*, 2022.

[230] Victor J. Rayward-Smith. Hypergrammars: an extension of macrogrammars. *Journal of Computer and System Sciences*, 14(1):130–149, 1977.

[231] Lisbeth J Reales-Chacón, María E Lucena de Ustáriz, Francisco J Ustáriz-Fajardo, Andrea C Peñafiel Luna, Gabriela J Bonilla-Ayala, Pablo Djabayan-Djibeyan, José L Erazo-Parra, and Mónica A Valdiviezo-Maygua. Study of the efficacy of neural mobilizations to improve sensory and functional responses of lower extremities in older adults with diabetic peripheral neuropathy using plithogenic n-superhypergraphs. *Neutrosophic Sets and Systems*, 74:1–12, 2024.

[232] Stefano Crespi Reghizzi, Luca Breveglieri, and Angelo Morzenti. *Formal languages and compilation*. Springer, 2013.

[233] Akbar Rezaei, Tahsin Oner, Tugce Katican, Florentin Smarandache, and N Gandotra. *A short history of fuzzy, intuitionistic fuzzy, neutrosophic and plithogenic sets*. Infinite Study, 2022.

[234] Judith Roitman. *Introduction to modern set theory*, volume 8. John Wiley & Sons, 1990.

[235] Azriel Rosenfeld. Fuzzy graphs. In *Fuzzy sets and their applications to cognitive and decision processes*, pages 77–95. Elsevier, 1975.

[236] Sheldon M Ross. *Applied probability models with optimization applications*. Courier Corporation, 2013.

[237] Sheldon M Ross. *Introduction to probability models*. Academic press, 2014.

[238] Grzegorz Rozenberg and Arto Salomaa. *Handbook of Formal Languages: Volume 3 Beyond Words*. Springer Science & Business Media, 2012.

[239] Rovshan G Sadygov and John R Yates. A hypergeometric probability model for protein identification and validation using tandem mass spectral data and protein sequence databases. *Analytical chemistry*, 75(15):3792–3798, 2003.

[240] Muhammad Haris Saeed, Muhammad Rayees, and Atiqe Ur Rahman. Refined pythagorean fuzzy sets: Properties, set-theoretic operations and axiomatic results. 2023.

[241] Il'ya Yur'evich Samonenko. The number of regular languages, recognized by group hyperautomata. *Intelligent systems. Theory and applications*, 22(2):113–122, 2018.

[242] P Sathya, Nivetha Martin, and Florentine Smarandache. Plithogenic forest hypersoft sets in plithogenic contradiction based multi-criteria decision making. *Neutrosophic Sets and Systems*, 73:668–693, 2024.

[243] Ipek Baris Schlicht, Defne Altiok, Maryanne Taouk, and Lucie Flek. Pitfalls of conversational llms on news debiasing. *ArXiv*, abs/2404.06488, 2024.

[244] Gulfam Shahzadi, Muhammad Akram, Arsham Borumand Saeid, et al. An application of single-valued neutrosophic sets in medical diagnosis. *Neutrosophic sets and systems*, 18:80–88, 2017.

[245] Francina Shalini. Trigonometric similarity measures of pythagorean neutrosophic hypersoft sets. *Neutrosophic Systems with Applications*, 2023.

[246] Lumin Shao. Fuzzy language in literature and translation. 2023.

[247] Utsav Sharma, Bing-Yue Wu, Sai Rahul Dhanvi Kankipati, Vidya A. Chhabria, and Austin Rovinski. Openroad-assistant: An open-source large language model for physical design tasks. *2024 ACM/IEEE 6th Symposium on Machine Learning for CAD (MLCAD)*, pages 1–7, 2024.

[248] Xuan Shen, Zhenglun Kong, Changdi Yang, Zhaoyang Han, Lei Lu, Peiyan Dong, Cheng Lyu, Chih hsiang Li, Xuehang Guo, Zhihao Shu, Wei Niu, Miriam Leeser, Pu Zhao, and Yanzhi Wang. Edgeqat: Entropy and distribution guided quantization-aware training for the acceleration of lightweight llms on the edge. *ArXiv*, abs/2402.10787, 2024.

[249] Xiaolong Shi, Saeed Kosari, Hossein Rashmanlou, Said Broumi, and S Satham Hussain. Properties of interval-valued quadripartitioned neutrosophic graphs with real-life application. *Journal of Intelligent & Fuzzy Systems*, 44(5):7683–7697, 2023.

[250] George K. Sidiropoulos, Kyriakos D. Apostolidis, Nikolaos Damianos, and George A. Papakostas. Fsmpy: A fuzzy set measures python library. *Inf.*, 13:64, 2022.

[251] Prem Kumar Singh. Complex plithogenic set. *International Journal of Neutrosophic Sciences*, 18(1):57–72, 2022.

[252] Pushpinder Singh. Correlation coefficients for picture fuzzy sets. *J. Intell. Fuzzy Syst.*, 28:591–604, 2015.

[253] Seema Singh, DS Hooda, and SC Malik. Fuzzy rough information measures and their applications. *Int J Comput Sci Inf Technol (IJCSITY)*, 9(4):1–17, 2021.

[254] Ahmed Sleem, Mohamed Abdel-Baset, and Ibrahim Mahmoud El-Henawy. Pyivns: A python based tool for interval-valued neutrosophic operations and normalization. *SoftwareX*, 12:100632, 2020.

[255] F. Smarandache. Introduction to superhyperalgebra and neutrosophic superhyperalgebra. *Journal of Algebraic Hyperstructures and Logical Algebras*, 2022.

[256] Florentin Smarandache. Neutrosophic overset, neutrosophic underset, and neutrosophic offset. similarly for neutrosophic over-/under-/offlogic, probability, and statisticsneutrosophic, pons editions brussels, 170 pages book, 2016.

[257] Florentin Smarandache. A unifying field in logics: Neutrosophic logic. In *Philosophy*, pages 1–141. American Research Press, 1999.

[258] Florentin Smarandache. Neutrosophic set-a generalization of the intuitionistic fuzzy set. *International journal of pure and applied mathematics*, 24(3):287, 2005.

[259] Florentin Smarandache. n-valued refined neutrosophic logic and its applications to physics. *Infinite study*, 4:143–146, 2013.

[260] Florentin Smarandache. Degrees of membership¿ 1 and¡ 0 of the elements with respect to a neutrosophic offset. *Neutrosophic Sets and Systems*, 12:3–8, 2016.

[261] Florentin Smarandache. *Plithogenic set, an extension of crisp, fuzzy, intuitionistic fuzzy, and neutrosophic sets-revisited*. Infinite study, 2018.

[262] Florentin Smarandache. Plithogeny, plithogenic set, logic, probability, and statistics. *arXiv preprint arXiv:1808.03948*, 2018.

[263] Florentin Smarandache. n-superhypergraph and plithogenic n-superhypergraph. *Nidus Idearum*, 7:107–113, 2019.

[264] Florentin Smarandache. *Extension of HyperGraph to n-SuperHyperGraph and to Plithogenic n-SuperHyperGraph, and Extension of HyperAlgebra to n-ary (Classical-/Neutro-/Anti-) HyperAlgebra*. Infinite Study, 2020.

[265] Florentin Smarandache. History of superhyperalgebra and neutrosophic superhyperalgebra (revisited again). *Neutrosophic Algebraic Structures and Their Applications*, page 10, 2022.

[266] Florentin Smarandache. Interval-valued neutrosophic oversets, neutrosophic undersets, and neutrosophic offsets. *Collected Papers. Volume IX: On Neutrosophic Theory and Its Applications in Algebra*, page 117, 2022.

[267] Florentin Smarandache. *Introduction to the n-SuperHyperGraph-the most general form of graph today*. Infinite Study, 2022.

[268] Florentin Smarandache. Operators on single-valued neutrosophic oversets, neutrosophic undersets, and neutrosophic offsets. *Collected Papers*, 9:112, 2022.

[269] Florentin Smarandache. *Practical applications of IndetermSoft Set and IndetermHyperSoft Set and introduction to TreeSoft Set as an extension of the MultiSoft Set*. Infinite Study, 2022.

[270] Florentin Smarandache. *The SuperHyperFunction and the Neutrosophic SuperHyperFunction (revisited again)*, volume 3. Infinite Study, 2022.

[271] Florentin Smarandache. Decision making based on valued fuzzy superhypergraphs. 2023.

[272] Florentin Smarandache. Foundation of the superhypersoft set and the fuzzy extension superhypersoft set: A new vision. *Neutrosophic Systems with Applications*, 11:48–51, 2023.

[273] Florentin Smarandache. *New types of soft sets "hypersoft set, indetermsoft set, indetermhypersoft set, and treesoft set": an improved version*. Infinite Study, 2023.

[274] Florentin Smarandache. New types of topologies and neutrosophic topologies. *Neutrosophic Systems with Applications*, 1:1–3, 2023.

[275] Florentin Smarandache. New types of topologies and neutrosophic topologies (improved version). *Neutrosophic Sets and Systems*, 57(1):14, 2023.

[276] Florentin Smarandache. *SuperHyperFunction, SuperHyperStructure, Neutrosophic SuperHyperFunction and Neutrosophic SuperHyperStructure: Current understanding and future directions*. Infinite Study, 2023.

[277] Florentin Smarandache. Foundation of revolutionary topologies: An overview, examples, trend analysis, research issues, challenges, and future directions. 2024.

[278] Florentin Smarandache. Foundation of superhyperstructure & neutrosophic superhyperstructure. *Neutrosophic Sets and Systems*, 63(1):21, 2024.

[279] Florentin Smarandache. Superhyperstructure & neutrosophic superhyperstructure, 2024. Accessed: 2024-12-01.

[280] Florentin Smarandache and Mohamed Abdel-Basset. *Optimization Theory Based on Neutrosophic and Plithogenic Sets*. Academic Press, 2020.

[281] Florentin Smarandache and NM Gallup. Generalization of the intuitionistic fuzzy set to the neutrosophic set. In *International Conference on Granular Computing*, pages 8–42. Citeseer, 2006.

[282] Florentin Smarandache, WB Kandasamy, and K Ilanthenral. Applications of bimatrices to some fuzzy and neutrosophic models. 2005.

[283] Florentin Smarandache and Nivetha Martin. *Plithogenic n-super hypergraph in novel multi-attribute decision making*. Infinite Study, 2020.

[284] Florentin Smarandache, Memet Şahin, Derya Bakbak, Vakkas Uluçay, and Abdullah Kargın. *Neutrosophic SuperHyperAlgebra and New Types of Topologies*. Infinite Study, 2023.

[285] Florentin Smarandache, A. Saranya, A. Kalavathi, and S. Krishnaprakash. Neutrosophic superhypersoft sets. *Neutrosophic Sets and Systems*, 77:41–53, 2024.

[286] Seok-Zun Song, Seon Jeong Kim, and Young Bae Jun. Hyperfuzzy ideals in bck/bci-algebras. *Mathematics*, 5(4):81, 2017.

[287] Simone Spolaor, Caro Fuchs, Paolo Cazzaniga, Uzay Kaymak, Daniela Besozzi, and Marco S. Nobile. Simpful: A user-friendly python library for fuzzy logic. *Int. J. Comput. Intell. Syst.*, 13:1687–1698, 2020.

[288] Bo Stenstrom. *Rings of quotients: an introduction to methods of ring theory*, volume 217. Springer Science & Business Media, 2012.

[289] S Sudha, Nivetha Martin, and Florentin Smarandache. *Applications of Extended Plithogenic Sets in Plithogenic Sociogram*. Infinite Study, 2023.

[290] John Brendan Sullivan. Representations of the hyperalgebra of an algebraic group. *American Journal of Mathematics*, 100(3):643–652, 1978.

[291] Rina Supriatnaningsih, Rustono Rustono, Edi Astini, and Tatang Hariri. The use of japanese politeness markers by university students. In *Proceedings of the 5th International Conference on Science, Education and Technology, ISET 2019, 29th June 2019, Semarang, Central Java, Indonesia*, 2020.

[292] Eulalia Szmidt. *Distances and similarities in intuitionistic fuzzy sets*, volume 307. Springer, 2014.

[293] Jinjun Tang, Fang Liu, Wenhui Zhang, Ruimin Ke, and Yajie Zou. Lane-changes prediction based on adaptive fuzzy neural network. *Expert Syst. Appl.*, 91:452–463, 2018.

[294] Jinjun Tang, Fang Liu, Yajie Zou, Weibin Zhang, and Yinhai Wang. An improved fuzzy neural network for traffic speed prediction considering periodic characteristic. *IEEE Transactions on Intelligent Transportation Systems*, 18:2340–2350, 2017.

[295] Ross Taylor, Marcin Kardas, Guillem Cucurull, Thomas Scialom, Anthony S. Hartshorn, Elvis Saravia, Andrew Poulton, Viktor Kerkez, and Robert Stojnic. Galactica: A large language model for science. *ArXiv*, abs/2211.09085, 2022.

[296] Lev Telyatnikov, Maria Sofia Bucarelli, Guillermo Bernardez, Olga Zaghen, Simone Scardapane, and Pietro Lió. Hypergraph neural networks through the lens of message passing: A common perspective to homophily and architecture design. *ArXiv*, abs/2310.07684, 2023.

[297] Shailja Thakur, Baleegh Ahmad, Hammond A. Pearce, Benjamin Tan, Brendan Dolan-Gavitt, Ramesh Karri, and Siddharth Garg. Verigen: A large language model for verilog code generation. *ACM Transactions on Design Automation of Electronic Systems*, 29:1 – 31, 2023.

[298] Vicenç Torra. Hesitant fuzzy sets. *International journal of intelligent systems*, 25(6):529–539, 2010.

[299] Vicenç Torra and Yasuo Narukawa. On hesitant fuzzy sets and decision. In *2009 IEEE international conference on fuzzy systems*, pages 1378–1382. IEEE, 2009.

[300] Mutsuko Tsuboi. Nationalism and ethnicity in the modern japanese context: Translation and ideology in the late 19th century. *CLINA*, 5(1):27, 2019.

[301] Natsuko Tsujimura. *An introduction to Japanese linguistics*. John Wiley & Sons, 2013.

[302] Vakkas Ulucay, Irfan Deli, and Mehmet Sahin. Similarity measures of bipolar neutrosophic sets and their application to multiple criteria decision making. *Neural Computing and Applications*, 29:739–748, 2018.

[303] Robert L Vaught. *Set theory: an introduction*. Springer Science & Business Media, 2001.

[304] Guillaume Verdon, Trevor McCourt, Enxhell Luzhnica, Vikash Singh, Stefan Leichenauer, and Jack Hidary. Quantum graph neural networks. *arXiv preprint arXiv:1909.12264*, 2019.

[305] Nikolai Konstantinovich Vereshchagin and Alexander Shen. *Basic set theory*. Number 17. American Mathematical Soc., 2002.

[306] Aishwarya Vijayan. A prompt engineering approach for structured data extraction from unstructured text using conversational llms. *Proceedings of the 2023 6th International Conference on Algorithms, Computing and Artificial Intelligence*, 2023.

[307] Thanh Vu, Dat Quoc Nguyen, Dai Quoc Nguyen, Mark Dras, and Mark Johnson. Vncorenlp: A vietnamese natural language processing toolkit. In *North American Chapter of the Association for Computational Linguistics*, 2018.

[308] Haibin Wang, Florentin Smarandache, Yanqing Zhang, and Rajshekhar Sunderraman. *Single valued neutrosophic sets*. Infinite study, 2010.

[309] Haibin Wang, Florentin Smarandache, Yanqinq Zhang, and Rajshekhar Sunderraman. Single valued neutrosophic sets. *viXra*, 2010.

[310] Meng Wang and Fanghui Hu. The application of nltk library for python natural language processing in corpus research. *Theory and Practice in Language Studies*, 2021.

[311] Yue Wang, Hung Le, Akhilesh Deepak Gotmare, Nghi D. Q. Bui, Junnan Li, and Steven C. H. Hoi. Codet5+: Open code large language models for code understanding and generation. In *Conference on Empirical Methods in Natural Language Processing*, 2023.

[312] Yuxin Wang, Quan Gan, Xipeng Qiu, Xuanjing Huang, and David Paul Wipf. From hypergraph energy functions to hypergraph neural networks. In *International Conference on Machine Learning*, 2023.

[313] Guiwu Wei and Hui Gao. The generalized dice similarity measures for picture fuzzy sets and their applications. *Informatica*, 29:107–124, 2018.

[314] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837, 2022.

[315] Xiangpeng Wei, Hao-Ran Wei, Huan Lin, Tianhao Li, Pei Zhang, Xingzhang Ren, Mei Li, Yu Wan, Zhiwei Cao, Binbin Xie, Tianxiang Hu, Shangjie Li, Binyuan Hui, Yu Bowen, Dayiheng Liu, Baosong Yang, Fei Huang, and Jun Xie. Polylm: An open source polyglot large language model. *ArXiv*, abs/2307.06018, 2023.

[316] Yuxiang Wei, Shiqi Wang, and Yun Li. Graph theory based machine learning for analog circuit design. In *2023 28th International Conference on Automation and Computing (ICAC)*, pages 1–6. IEEE, 2023.

[317] Anna Maria Wierzbicka. I and thou: Universal human concepts present as words in all human languages. *Russian Journal of Linguistics*, 2022.

[318] Yorick Wilks. Natural language processing - introduction to the special section. *Commun. ACM*, 39:60–62, 1996.

[319] Patrick Henry Winston. *Artificial intelligence*. Addison-Wesley Longman Publishing Co., Inc., 1992.

[320] Robert Wisbauer. *Foundations of module and ring theory*. Routledge, 2018.

[321] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pages 38–45, 2020.

[322] Shengqiong Wu, Hao Fei, Leigang Qu, Wei Ji, and Tat-Seng Chua. Next-gpt: Any-to-any multimodal llm. *arXiv preprint arXiv:2309.05519*, 2023.

[323] Zeshui Xu. Hesitant fuzzy sets theory. In *Studies in Fuzziness and Soft Computing*, 2014.

[324] Hongpeng Yan. Research on rule extraction method based on concept lattice of intuitionistic fuzzy language. *Applied Mathematics and Nonlinear Sciences*, 8:867 – 884, 2022.

[325] Rui Yang, Qingcheng Zeng, Keen You, Yujie Qiao, Lucas Huang, Chia-Chun Hsieh, Benjamin Rosand, Jeremy Goldwasser, Amisha D. Dave, Tiarnán D. L. Keenan, Emily Y. Chew, Dragomir R. Radev, Zhiyong Lu, Hua Xu, Qingyu Chen, and Irene Li. Ascle: A python natural language processing toolkit for medical text generation. 2023.

[326] Rui Yang, Qingcheng Zeng, Keen You, Yujie Qiao, Lucas Huang, Chia-Chun Hsieh, Benjamin Rosand, Jeremy Goldwasser, Amisha D. Dave, Tiarnan D.L. Keenan, Yu He Ke, Chuan Hong, Nan Liu, Emily Y. Chew, Dragomir R. Radev, Zhiyong Lu, Hua Xu, Qingyu Chen, and Irene Li. Ascle-a python natural language processing toolkit for medical text generation: Development and evaluation study. *Journal of Medical Internet Research*, 26, 2024.

[327] Xianjun Yang, Junfeng Gao, Wenxin Xue, and Erik Alexandersson. Pllama: An open-source large language model for plant science. *ArXiv*, abs/2401.01600, 2024.

[328] Xibei Yang, Dongjun Yu, Jingyu Yang, and Chen Wu. Generalization of soft set theory: from crisp to fuzzy case. In *Fuzzy Information and Engineering: Proceedings of the Second International Conference of Fuzzy Information and Engineering (ICFIE)*, pages 345–354. Springer, 2007.

[329] Jun Ye and Shigui Du. Some distances, similarity and entropy measures for interval-valued neutrosophic sets and their relationship. *International Journal of Machine Learning and Cybernetics*, 10:347 – 355, 2017.

[330] Qinghao Ye, Haiyang Xu, Guohai Xu, Jiabo Ye, Ming Yan, Yiyang Zhou, Junyang Wang, Anwen Hu, Pengcheng Shi, Yaya Shi, et al. mplug-owl: Modularization empowers large language models with multimodality. *arXiv preprint arXiv:2304.14178*, 2023.

[331] Jeonghee Yi, Tetsuya Nasukawa, Razvan Bunescu, and Wayne Niblack. Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In *Third IEEE international conference on data mining*, pages 427–434. IEEE, 2003.

[332] Shukang Yin, Chaoyou Fu, Sirui Zhao, Ke Li, Xing Sun, Tong Xu, and Enhong Chen. A survey on multimodal large language models. *National Science Review*, page nwae403, 2024.

[333] Hong yu Zhang, Jian qiang Wang, and Xiao hong Chen. An outranking approach for multi-criteria decision-making problems with interval-valued neutrosophic sets. *Neural Computing and Applications*, 27:615–627, 2016.

[334] Lotfi A Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.

[335] Lotfi A Zadeh. Biological application of the theory of fuzzy sets and systems. In *The Proceedings of an International Symposium on Biocybernetics of the Central Nervous System*, pages 199–206. Little, Brown and Comp. London, 1969.

[336] Lotfi A Zadeh. A fuzzy-set-theoretic interpretation of linguistic hedges. 1972.

[337] Lotfi A Zadeh. Fuzzy sets and their application to pattern classification and clustering analysis. In *Classification and clustering*, pages 251–299. Elsevier, 1977.

[338] Lotfi A Zadeh. Fuzzy sets versus probability. *Proceedings of the IEEE*, 68(3):421–421, 1980.

[339] Lotfi A Zadeh. Fuzzy logic, neural networks, and soft computing. In *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh*, pages 775–782. World Scientific, 1996.

[340] Lotfi A Zadeh. Fuzzy sets and information granularity. In *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh*, pages 433–448. World Scientific, 1996.

[341] Lotfi A Zadeh. A note on prototype theory and fuzzy sets. In *Fuzzy sets, fuzzy logic, and fuzzy systems: Selected papers by Lotfi A Zadeh*, pages 587–593. World Scientific, 1996.

[342] Team Glm Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Diego Rojas, Guanyu Feng, Hanlin Zhao, Hanyu Lai, Hao Yu, Hongning Wang, Jiadai Sun, Jiajie Zhang, Jiale Cheng, Jiayi Gui, Jie Tang, Jing Zhang, Juanzi Li, Lei Zhao, Lindong Wu, Lucen Zhong, Ming yue Liu, Minlie Huang, Peng Zhang, Qinkai Zheng, Rui Lu, Shuaiqi Duan, Shudan Zhang, Shulin Cao, Shuxun Yang, Weng Lam Tam, Wenyi Zhao, Xiao Liu, Xiaoyu Xia, Xiaohan Zhang, Xiaotao Gu, Xin Lv, Xinghan Liu, Xinyi Liu, Xinyue Yang, Xixuan Song, Xunkai Zhang, Yi An, Yifan Xu, Yilin Niu, Yuantao Yang, Yueyan Li, Yushi Bai, Yuxiao Dong, Zehan Qi, Zhaoyu Wang, Zhenyi Yang, Zhengxiao Du, Zhen-Ping Hou, and Zihan Wang. Chatglm: A family of large language models from glm-130b to glm-4 all tools. *ArXiv*, abs/2406.12793, 2024.

[343] Duzhen Zhang, Yahan Yu, Jiahua Dong, Chenxing Li, Dan Su, Chenhui Chu, and Dong Yu. Mm-llms: Recent advances in multimodal large language models. *arXiv preprint arXiv:2401.13601*, 2024.

[344] Hongyu Zhang, Jian qiang Wang, and Xiao hong Chen. An outranking approach for multi-criteria decision-making problems with interval-valued neutrosophic sets. *Neural Computing and Applications*, 27:615 – 627, 2015.

[345] Hongyu Zhang, Jianqiang Wang, and Xiaohong Chen. An outranking approach for multi-criteria decision-making problems with interval-valued neutrosophic sets. *Neural Computing and Applications*, 27:615–627, 2016.

[346] Qian-Sheng Zhang and Sheng-Yi Jiang. A note on information entropy measures for vague sets and its applications. *Information Sciences*, 178(21):4184–4191, 2008.

[347] Yingying Zhang. Approaches to multiple attribute group decision making under interval-valued pythagorean fuzzy sets and applications to environmental design majors teaching quality evaluation. *Int. J. Knowl. Based Intell. Eng. Syst.*, 27:289–301, 2023.

[348] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. A survey of large language models. *arXiv preprint arXiv:2303.18223*, 2023.

[349] Chen Zheng, Ke Sun, Hang Wu, Chenguang Xi, and Xun Zhou. Balancing enhancement, harmlessness, and general capabilities: Enhancing conversational llms with direct rlhf. *ArXiv*, abs/2403.02513, 2024.

[350] H-J Zimmermann. Fuzzy set theory and mathematical programming. *Fuzzy sets theory and applications*, pages 99–114, 1986.

[351] Hans-Jürgen Zimmermann. Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets and Systems*, 1:45–55, 1978.

[352] Hans-Jürgen Zimmermann. Applications of fuzzy set theory to mathematical programming. *Inf. Sci.*, 36:29–58, 1985.