

**Woldia University Institute of Technology**

**School of Computing**

**Department of Computer Science**

# Natural Language Processing Group Assignment

**Group Members:**

**No. Name ID**

1. **ASHENAFI Kasa 1300402**
2. **KIDUS DESALEGN 1301690**
3. **MAHLET LEMMA 1301945**
4. **Nani Naeto 1302333**
5. **SHERIF ABDELA 1302653**
6. **YEABTSEGA BEMNET 1303048**
7. **YOHANNIS GETACHE 1303101**

**Submitted to: Mr. Girma. B**

### Submission Date: May 2025

### 6.1: Representing Meaning in NLP

**What is Meaning Representation in NLP?**

**Answer:** Meaning representation in Natural Language Processing (NLP) refers to the formal encoding of semantic content from natural language into structured formats that computers can process and understand. It serves as a crucial bridge between human language and machine-processable representations.

**Key aspects of meaning representation include:**

**1.Purpose:**

* 1. Converts natural language into unambiguous, structured formats
  2. Enables computer systems to process and reason about meaning
  3. Facilitates automated inference and logical reasoning

1. **Characteristics:**
   1. **Unambiguous:** Each representation has a single, clear interpretation.
   2. **Verifiable:** Can be checked for consistency and correctness.
   3. **Inferential:** Supports drawing logical conclusions.
   4. **Canonical:** Similar meanings have similar representations.
   5. **Verifiability:** The ability to determine the truth value of the representation by checking a knowledge base.
   6. **Inference and Variables:** The capacity to draw valid conclusions based on the meaning representations and background knowledge.
   7. **Expressiveness:** The ability to represent a wide range of subject matter.

**2: Approaches to Representing Meaning**

Let's examine primary approaches to meaning representation, including foundational and more modern techniques:

#### 1. First-Order Predicate Logic (FOPL)

**Definition:** FOPL is a formal logical system that represents meaning through predicates, arguments, and logical operators. It allows for precise and unambiguous representation, making it suitable for formal reasoning.

**Components:**

* **Constants:** Specific entities (e.g., john, dog).
* **Variables:** Placeholders for entities (e.g., x, y).
* **Predicates:** Properties or relations (e.g., Dog(x), Has(dog, tail)).
* **Quantifiers:**
  + **Universal (∀):** "for all" (e.g., ∀x Dog(x) → Animal(x) - All dogs are animals).
  + **Existential (∃):** "there exists" (e.g., ∃x Dog(x) - There exists a dog).

**Example:**

* **Basic representation:** Dog(x) → Animal(x) // All dogs are animals Has(dog, tail) // Dogs have tails
* **Complex representation:** ∀x (Dog(x) → Animal(x)) // For all x, if x is a dog, then x is an animal. Has(fido, tail) // Fido has a tail.

#### 2. Semantic Networks

**Definition:** Semantic Networks represent knowledge as a graph, where nodes represent concepts or entities, and edges represent relationships between them. They offer an intuitive visual representation of knowledge.

**Components:**

* **Nodes:** Concepts, objects, or entities (e.g., dog, animal, tail).
* **Edges (Arcs):** Labeled relationships between nodes (e.g., is-a, has-part, can).

**Example:**

**Graphical representation:**

(Dog) --is-a--> (Animal)

(Dog) --has--> (Tail)

(Dog) --can--> (Bark)

#### 3. Frame-Based Representation (Frames)

**Definition:** Frames are structured representations that capture stereotypical situations or concepts. They consist of a name for the frame and a set of "slots" or "roles" that describe the elements involved in that situation.

**Components:**

* **Frames:** Common situations or scenarios (e.g., "Commercial\_Transaction").
* **Frame Elements:** Roles played by different participants in a frame (e.g., Buyer, Seller, Goods, Price).
* **Lexical Units (LUs):** Pairs of words and their meanings within a specific frame.

**Example:** A "Commercial\_Transaction" frame might have slots like:

* Buyer: Person or entity purchasing.
* Seller: Person or entity selling.
* Goods: Item being transacted.
* Price: Cost of the goods.
* Payment\_Method: How the payment is made.

**Application:** FrameNet is a prominent lexical resource built on frame semantics, mapping words to semantic frames and their associated roles.

#### 4. Conceptual Dependency (CD)

**Definition:** Developed by Roger Schank, Conceptual Dependency theory represents meaning using a small set of primitive actions and states, aiming for a canonical representation that abstracts away from surface linguistic variations.

**Primitive Actions Example:**

* PTRANS: Physical transfer of an object (e.g., "John went to the store").
* MTRANS: Mental transfer of information (e.g., "Mary told John a secret").
* ATTEND: Focusing of attention (e.g., "John looked at the book").

**Application:** CD was used in early NLP systems for tasks like machine translation and question answering, demonstrating how a limited set of primitives could represent diverse sentences.

#### 5. Abstract Meaning Representation (AMR)

**Definition:** AMR is a graph-based semantic representation that captures the "who did what to whom, when, and where" of a sentence. It focuses on the predicate-argument structure and simplifies sentences into a more abstract, canonical form.

**Key Concepts:**

* **Nodes:** Represent concepts (word senses, events, entities).
* **Edges:** Represent relations between concepts, including semantic roles (e.g., ARG0 for agent, ARG1 for patient) and general semantic relations (e.g., time, location).
* **Polarity:** Indicates positive, negative, or neutral sentiment.

**Example:** For the sentence "John quickly ran to the store," the AMR graph would represent:

(r / run-01

:ARG0 (j / John)

:ARG1 (s / store)

:manner (q / quick))

Here, run-01 is the main action, John is the agent (ARG0), store is the destination (ARG1), and quick describes the manner.

**Applications:** AMR has found applications in various NLP tasks, including information extraction, summarization, machine translation, and question answering, due to its ability to capture core semantic meaning.

### Conclusion

Meaning representation is fundamental to NLP as it provides the foundation for machines to understand and process natural language. Each approach offers distinct advantages and addresses different aspects of semantic complexity:

* **FOPL:** Excels in formal reasoning and precise, unambiguous representation, ideal for logical inference.
* **Semantic Networks:** Provide intuitive visualization and easy modification, useful for representing hierarchical or associative knowledge.
* **Frames:** Capture stereotypical situations and events, aiding in understanding domain-specific knowledge.
* **Conceptual Dependency:** Aim for a universal set of primitives to represent actions, facilitating cross-linguistic understanding.
* **AMR:** Focuses on predicate-argument structure for broad-coverage semantic parsing, suitable for a wide range of NLP applications.

The choice between these approaches depends on specific requirements such as:

* Complexity of knowledge to be represented
* Type of reasoning needed
* Need for human readability
* Computational efficiency requirements
* The specific NLP task (e.g., question answering, machine translation, information extraction)

### Challenges in Meaning Representation

* **Ambiguity:** Natural language is inherently ambiguous (lexical, syntactic, semantic), making it challenging to derive a single, unambiguous meaning representation.
  + **Polysemy:** A word having multiple related meanings (e.g., "bank" as a financial institution or a river bank).
  + **Homonymy:** Words with the same spelling/pronunciation but unrelated meanings (e.g., "bat" as an animal or sports equipment).
  + **Word Sense Disambiguation (WSD):** The crucial task of identifying the correct meaning of a word based on its context.
* **Vagueness:** While not leading to multiple representations, vagueness (e.g., "Turkish food") can make it difficult to determine precise meaning, requiring common-sense reasoning.
* **Expressiveness vs. Computability:** Balancing the ability to represent highly complex and nuanced meanings with the computational efficiency of processing those representations. Highly expressive representations can be computationally expensive.
* **Contextual Understanding:** Capturing the full meaning of text often requires understanding the broader context, including real-world knowledge, discourse, and pragmatics, which are difficult to formalize.
* **Knowledge Acquisition Bottleneck:** Creating large-scale, comprehensive knowledge bases manually for meaning representation is time-consuming and labor-intensive.

### References

1. Jurafsky, D., & Martin, J. H. (2023). Speech and Language Processing (3rd ed. draft). [Speech and Language Processing (3rd ed. draft)](https://web.stanford.edu/~jurafsky/slp3/" \t "_blank)
2. Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
3. Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media.
4. Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., ... & Palmer, M. (2013). Abstract Meaning Representation for Sembanking. In Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse (pp. 178-186). [AMR for Sembanking](https://aclanthology.org/W13-1723.pdf" \t "_blank)