A Qualia-binding Semantics of Compounds

James Pustejovsky

Brandeis University

November 26, 2024

Outline

Noun-Noun Compounds in Natural Language

- **Definition:** Noun-Noun compounds are lexical items formed by combining two nouns, e.g., "apple tree," "chicken soup."
- **Characteristics:**
  - Can express a wide range of semantic relations.
  - Often ambiguous without context.
- **Challenges:**
  - Understanding the underlying semantic relation.
  - Modeling the compositional semantics effectively.
- **Applications:**
  - Natural Language Processing (NLP).
  - Lexical semantics and cognitive linguistics.

Generative Lexicon Theory

- **Key Concepts:**
  - Lexical items have rich internal structures.
  - Meanings are generated dynamically through composition.
- **Qualia Structure:** Provides a structured representation of a word's meaning.
- **Dynamic Event Structure:** Subatomic representation of events and state change.
- **Dense Compositional Operations:** Type coercion, co-composition, and subselection provides richer compositionality.

Qualia Roles

- **Constitutive (C):** Material, parts, and components.
- **Formal (F):** Distinguishing properties of the entity.
- **Telic (T):** Purpose and function.
- **Agentive (A):** Factors involved in the creation or bringing about of the entity.
- **Example:** For "book":
  - **C:** Pages, cover.
  - **F:** Printed material.
  - **T:** Reading.
  - **A:** Writing, publishing.

Analyzing Noun-Noun Compounds

- **Semantic Relations:**
  - *Modifier-Head*: "chicken soup" (soup made of chicken).
  - *Complementation*: "student loan" (loan to a student).
  - *Possession*: "child's toy" (toy belonging to a child).
- **Ambiguity:** Same compound can have multiple interpretations.
- **Importance of Context:** Contextual clues help disambiguate meanings.

Qualia Modification in Noun-Noun Compounds

- **Lee et al. (2010) Survey:**
  - Explores how Qualia roles are modified in Noun-Noun compounds across languages.
  - Identifies patterns in qualia modification.
- **Findings:**
  - Certain Qualia roles are more frequently involved.
  - Cross-linguistic similarities and differences exist.
- **Implications:**
  - Understanding qualia modification aids in semantic interpretation.
  - Can inform computational models for NLP.

Holographic Reduced Representations (HRR)

- **Concept:** HRR uses high-dimensional vectors to represent symbols and their relations.
- **Key Operations:**
  - *Binding*: Circular convolution to combine vectors.
  - *Bundling (Superposition)*: Vector addition to combine multiple vectors.
  - *Unbinding*: Circular correlation to retrieve components.
- **Advantages:**
  - Efficient representation of complex structures.
  - Robustness to noise.

Binding and Bundling in Semantics

- **Binding:**
  - Encodes relationships between concepts.
  - E.g., $\mathbf{z} = \mathbf{x} \otimes \mathbf{y}$.
- **Bundling:**
  - Combines multiple pieces of information.
  - E.g., $\mathbf{z} = \mathbf{x} + \mathbf{y}$.
- **Application to Qualia Roles:**
  - Bind Qualia roles to their fillers.
  - Bundle all role-filler pairs to represent the full semantics.

Representing Qualia Structures with HRR

- **Role Vectors:** $\mathbf{r}_{\mathsf{C}}, \mathbf{r}_{\mathsf{F}}, \mathbf{r}_{\mathsf{T}}, \mathbf{r}_{\mathsf{A}}$.
- **Filler Vectors:** Specific to each noun.
- **Binding Role and Filler:**

$$\mathbf{v}_{\text{role}} = \mathbf{r}_{\text{role}} \otimes \mathbf{f}_{\text{role}}$$

- **Bundling All Roles:**

$$\mathbf{V}_{\text{noun}} = \sum_{\text{role}} \mathbf{v}_{\text{role}}$$

- **Example:** For "book":

$$\mathbf{V}_{\text{book}} = \mathbf{r}_{\mathsf{C}} \otimes \mathbf{f}_{\text{pages}} + \mathbf{r}_{\mathsf{F}} \otimes \mathbf{f}_{\text{printed}} + \dots$$

Interpreting Noun-Noun Compounds

- **Compound Representation:**

$$\mathbf{C} = \mathbf{V}_{N_1} \star \mathbf{V}_{N_2}$$

  - $\star$: Combination operation (e.g., binding or bundling).
- **Semantic Interpretation:**
  - Determine which Qualia roles are involved.
  - Use role-specific operations to combine meanings.
- **Example:** "Chicken soup"
  - Likely involves **C** role (soup made of chicken).
  - Combine $\mathbf{V}_{\text{chicken}}$ and $\mathbf{V}_{\text{soup}}$ accordingly.

Type Theory Basics

- **Types:** Abstract representations of the kinds of values.
  - E.g., e: entities, t: truth values.
- **Functions:** Mapping from one type to another.
  - E.g., $e \rightarrow t$: functions from entities to truth values (properties).
- **Typed Lambda Calculus:** Formal system for representing functions and their applications.

Applying Type Theory to Qualia Roles

- **Qualia Roles as Functions:**
  - Each role maps an entity to its role-specific property.
  - E.g., C : $e \rightarrow e$.
- **Compound Semantics:**
  - Interpret compounds as function applications.
  - E.g., For "chicken soup":

$$\text{C(soup)(chicken)}$$

- **Compositionality:**
  - The meaning of the compound derives from the meanings of its parts and their combination.

Combining HRR and Type Theory

- **Unified Framework:**
  - Use HRR for representation.
  - Use Type Theory for formal interpretation.
- **Semantic Composition:**
  - Binding corresponds to function application.
  - Bundling corresponds to combining properties.
- **Example Interpretation:**

$$\mathbf{C} = \mathbf{V}_{N_2} + (\mathbf{r}_{\text{role}} \otimes \mathbf{V}_{N_1})$$

  - Represents modifying N_2 with N_1 via a specific Qualia role.

Formalizing the Example: "Apple Pie"

- **Qualia Structures:**
  - $\mathbf{V}_{\text{apple}} = \sum \mathbf{r}_i \otimes \mathbf{f}_i$
  - $\mathbf{V}_{\text{pie}} = \sum \mathbf{r}_j \otimes \mathbf{f}_j$
- **Semantic Relation:**
  - Likely involves **C** role (pie made of apples).
- **Compound Representation:**

$$\mathbf{C} = \mathbf{V}_{\text{pie}} + (\mathbf{r}_{\text{C}} \otimes \mathbf{V}_{\text{apple}})$$

- **Type-Theoretic Interpretation:**
  - C(pie)(apple)
  - The constitutive role of "pie" is filled by "apple."

Type Theory Basics

- **Simple Types:**
  - e: Entity type.
  - t: Truth value type.
- **Functional Types:**
  - If α and β are types, then $\alpha \to \beta$ is a function type.
- **Higher-Order Functions:**
  - Functions can take functions as arguments or return functions.
  - Example: $(\alpha \to \beta) \to \gamma$.

Qualia Roles as Typed Functions

- **Representing Qualia Roles:**
  - CON : $e \to e \to t$
  - TEL : $e \to (e \to t)$
  - FORM : $e \to P(e)$ (where $P(e)$ is the set of properties of e)
  - AGENT : $e \to e \to t$
- **Example:**
  - For pie $\in e$, CON(pie) : $e \to t$
  - CON(pie)(apple) $=$ True if "apple" is a constituent of "pie."

Lambda Calculus Representation

- **Qualia Role Functions:**
  - Constitutive Role:

$$\text{CON} = \lambda x_e.\lambda y_e.\text{constitutes}(y, x)$$

  - Telic Role:

$$\text{TEL} = \lambda x_e.\lambda y_e.\text{purpose}(x, y)$$

- **Type Assignments:**
  - CON : $e \to e \to t$
  - TEL : $e \to e \to t$
- **Variables and Constants:**
  - x, y: Variables of type e
  - $\text{apple}, \text{pie}, \text{student}, \text{loan}$: Constants of type e

Interpreting "Apple Pie"

- **Semantic Representation:**

$$\text{ApplePie} = \text{CON(pie)(apple)}$$

- **Derivation:**

$$\begin{aligned} & \text{CON(pie)(apple)} \\ = & (\lambda x.\lambda y.\text{constitutes}(y, x))(\text{pie})(\text{apple}) \\ = & (\lambda y.\text{constitutes}(y, \text{pie}))(\text{apple}) \\ = & \text{constitutes(apple, pie)} \end{aligned}$$

- **Interpretation:**
  - The expression evaluates to True if "apple" constitutes "pie."
- **Type Checking:**
  - constitutes : $e \times e \rightarrow t$
  - constitutes(apple, pie) : t

Interpreting "Student Loan"

- **Semantic Representation:**

$$\text{StudentLoan} = \text{TEL(loan)(student)}$$

- **Derivation:**

$$\begin{aligned} & \text{TEL(loan)(student)} \\ = & (\lambda x.\lambda y.\text{purpose}(x, y))\text{(loan)(student)} \\ = & (\lambda y.\text{purpose}(\text{loan}, y))\text{(student)} \\ = & \text{purpose}(\text{loan}, \text{student}) \end{aligned}$$

- **Interpretation:**
  - The expression evaluates to True if the purpose of "loan" is "student."
- **Type Checking:**
  - purpose : $e \times e \rightarrow t$
  - $\text{purpose}(\text{loan}, \text{student}) : t$

Model-Theoretic Framework

- **Domain D:** Set of entities in the model.
- **Interpretation Function I:** Assigns meanings to constants and predicates.
  - $I(\text{apple}), I(\text{pie}), I(\text{student}), I(\text{loan}) \in D$
  - $I(\text{constitutes}) \subseteq D \times D$
  - $I(\text{purpose}) \subseteq D \times D$
- **Truth Conditions:**
  - $\text{constitutes}(a, p)$ is True iff $(a, p) \in I(\text{constitutes})$
  - $\text{purpose}(l, s)$ is True iff $(l, s) \in I(\text{purpose})$

Interpreting Compounds in the Model

- **Example: "Apple Pie"**
  - $I(\text{ApplePie}) = \text{True iff } (I(\text{apple}), I(\text{pie})) \in I(\text{constitutes})$
- **Example: "Student Loan"**
  - $I(\text{StudentLoan}) = \text{True iff } (I(\text{loan}), I(\text{student})) \in I(\text{purpose})$
- **Satisfaction in the Model:**
  - The compound expression is satisfied if the corresponding relation holds in the model.

Representing Semantics with HRR

- **Role and Filler Vectors:**
  - Assign high-dimensional vectors to roles and fillers.
  - $\mathbf{r}_{\text{CON}}, \mathbf{r}_{\text{TEL}}$: Role vectors.
  - $\mathbf{f}_{\text{apple}}, \mathbf{f}_{\text{pie}}$: Filler vectors.
- **Binding Operation:**
  - $\mathbf{v}_{\text{role\_filler}} = \mathbf{r}_{\text{role}} \otimes \mathbf{f}_{\text{filler}}$
  - Binding represents function application.
- **Compound Representation:**

$$\mathbf{C} = \mathbf{V}_{N_2} + \mathbf{r}_{\text{role}} \otimes \mathbf{f}_{N_1}$$

  - Combines the meaning of N_2 modified by the role filled by N_1.

Example: HRR Representation of "Apple Pie"

- **Vectors:**
  - $\mathbf{V}_{\text{pie}} = \sum_{\text{roles}} \mathbf{r}_{\text{role}} \otimes \mathbf{f}_{\text{role}}$
  - $\mathbf{r}_{\text{CON}}$, $\mathbf{f}_{\text{apple}}$
- **Compound Vector:**

$$\mathbf{C} = \mathbf{V}_{\text{pie}} + (\mathbf{r}_{\text{CON}} \otimes \mathbf{f}_{\text{apple}})$$

- **Interpretation:**
  - The constitutive role of "pie" is filled by "apple."
- **Retrieval:**
  - To retrieve the constituent, unbind:

$$\mathbf{f}_{\text{apple}} \approx \mathbf{C} \otimes \mathbf{r}_{\text{CON}}^{-1}$$

Formal Steps in Interpretation

1. **Assign Types:**
   - $N_1, N_2 \in e$
   - Qualia roles R as functions $e \to e \to t$
2. **Identify Relevant Role:**
   - Determine which role R is involved in the compound.
3. **Semantic Representation:**

$$\text{Compound} = R(N_2)(N_1)$$

4. **Lambda Calculus Application:**

$$\begin{aligned} &(\lambda x.\lambda y.R'(y, x))(N_2)(N_1) \\ = \; &R'(N_1, N_2) \end{aligned}$$

5. **Model Interpretation:**
   - Evaluate $R'(N_1, N_2)$ in the model.

Linking HRR and Formal Semantics

- **HRR Operations as Semantic Functions:**
  - Binding ($\otimes$) corresponds to function application.
  - Bundling ($+$) corresponds to aggregation of information.
- **Type-Theoretic Correspondence:**
  - $\mathbf{r}_{\text{role}}$: Represents function R.
  - $\mathbf{f}_{\text{filler}}$: Represents argument N_1.
  - $\mathbf{r}_{\text{role}} \otimes \mathbf{f}_{\text{filler}}$: Represents $R(N_1)$.
- **Semantic Composition:**
  - $\mathbf{C} = \mathbf{V}_{N_2} + (\mathbf{r}_{\text{role}} \otimes \mathbf{f}_{N_1})$
  - Reflects the modification of N_2 by N_1 via role R.

Challenges and Future Work

- **Role Selection:**
  - Determining the correct Qualia role for a compound can be challenging.
- **Contextual Information:**
  - Incorporating context to improve interpretation.
- **Learning Representations:**
  - Automatically learning role vectors and fillers from data.
- **Integration with NLP Systems:**
  - Applying the approach to real-world language processing tasks.

Conclusion

- **Noun-Noun Compounds:** Present semantic challenges due to ambiguity and variety of relations.
- **Qualia Roles:** Provide a structured way to represent lexical semantics.
- **HRR and Binding:** Offer computational tools to model complex semantic compositions.
- **Formal Semantics:** Type theory helps in formally interpreting the meanings.

Further Reading

- Bouillon, P., E. Ježek, C. Melloni, and A. Picton. *Annotating Qualia Relations in Italian and French Complex Nominals.* In LREC, pp. 1527-1532. 2012.
- Ježek, E., and Melloni, C. (2011). *Nominals, Polysemy, and Co-predication*. Journal of Cognitive Science, 12, 1-31.
- Johnston, M. and Busa, F. (1999) *Qualia structure and the compositional interpretation of compounds*. Breadth and depth of semantic lexicons, pp.167-187.
- Lee, C.-Y., Chang, C.-H., Hsu, W.-C., and Hsieh, S.-K. (2010). *Qualia Modification in Noun-Noun Compounds: A Cross-Language Survey*. ROCLING 2010 Poster Papers.
- Lenci, A., and Zamparelli, R. (2008). *Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space*.
- Plate, T. (2003). *Holographic Reduced Representation: Distributed Representation for Cognitive Structures*.
- Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press.
- Pustejovsky, J., and Ježek, E. (2008). *Semantic Coercion in Language: Beyond Distributional Analysis*. Italian Journal of Linguistics, 20(1), pp.175-208.