Implementing a Logical Inference System for Japanese Comparatives

Aug. 5

Yosuke Mikami, Daiki Matsuoka, Hitomi Yanaka

NLI Involving Comparatives

- It requires understanding of **quantities** and **comparative relations**

- These are difficult for the deep-learning based approach to handle accurately

P1: John is heavier than Bob.

P2: Bob is heavier than 70 kg.

⇨

H: John is heavier than 70 kg.

Approach Based on Compositional Semantics

- **Compositional semantics**

  - Derives the meaning of a phrase from the meanings of its parts

  - The meanings are represented as **logical formulas**

    (called semantic representations)

- Some NLI systems utilize compositional semantics
  - **ccgcomp** [Haruta+ 2022] **(for English comparatives)**
  - ccg2lambda [Mineshima+ 2015, 2016, Martínez-Gómez+ 2017]
  - ccgtemp [Sugimoto+ 2022]
  - LangPro [Abzianidze 2015]
  - FraCoq [Bernardy and Chatzikyriakidis 2017]
  - MonaLog [Hu+ 2020]

Challenges in handling Japanese comparatives

- Ccgcomp can handle **English comparatives**, but …

- **Japanese** has several unique properties related to comparatives

  - Absence of comparative morphemes

  - Approximate interpretation of equatives

  - Expressions with presupposition

  - …

⇒ **It is difficult to apply ccgcomp directly to Japanese comparatives**

Our Research Question

How can we implement **a logical inference system** based on compositional semantics, focusing on **Japanese comparatives**?

- Framework

  - **Combinatory Categorial Grammar (CCG)** [Steedman 2000]

    - Used for syntactic and semantic parsing

  - **Degree semantics** [Cresswell 1976, Klein 1980]

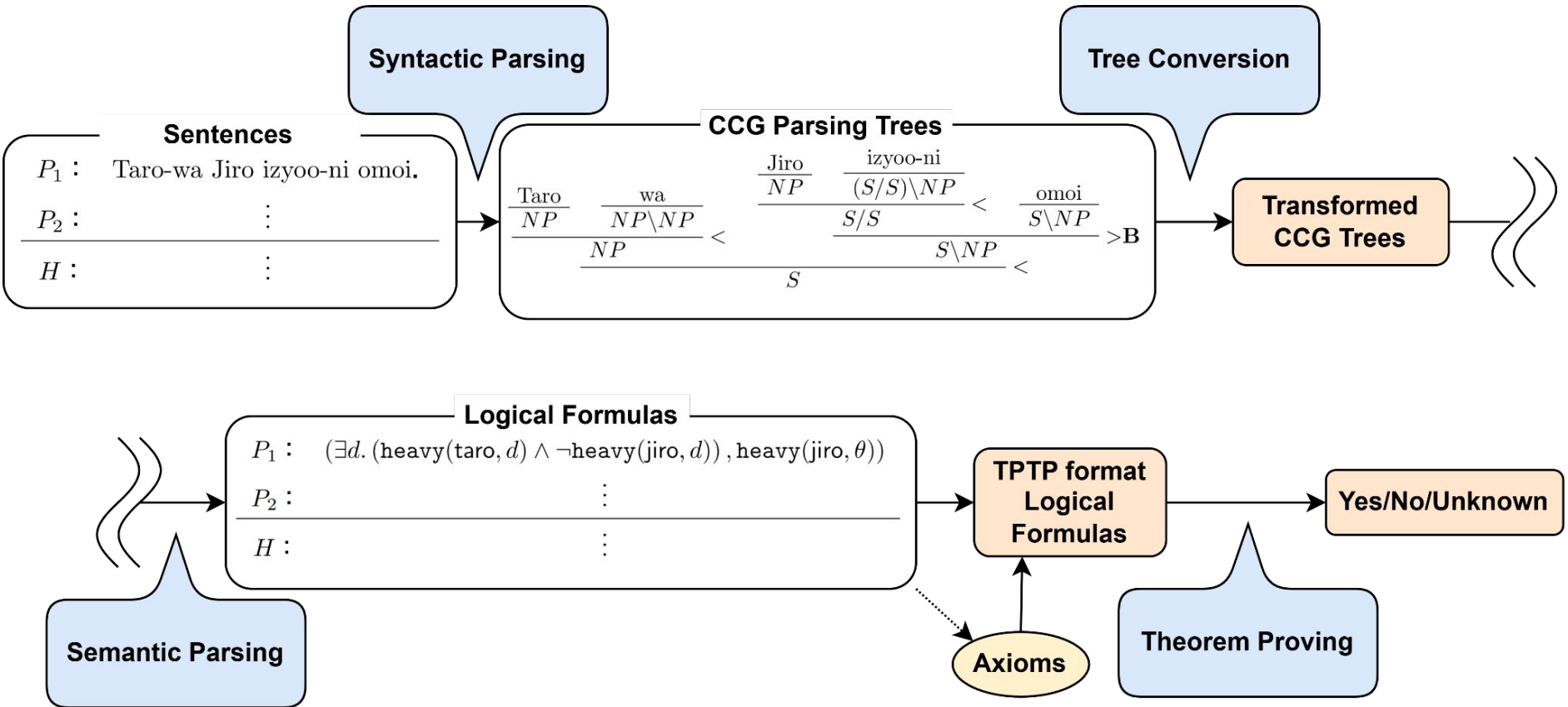   - Used for expressing the meaning of adjectives and comparatives

Outline

- Overview of Our Proposed System

- Linguistic Properties Unique to Japanese

- Experiment

- Results & Discussion

Outline

- **Overview of Our Proposed System**

- Linguistic Properties Unique to Japanese

- Experiment
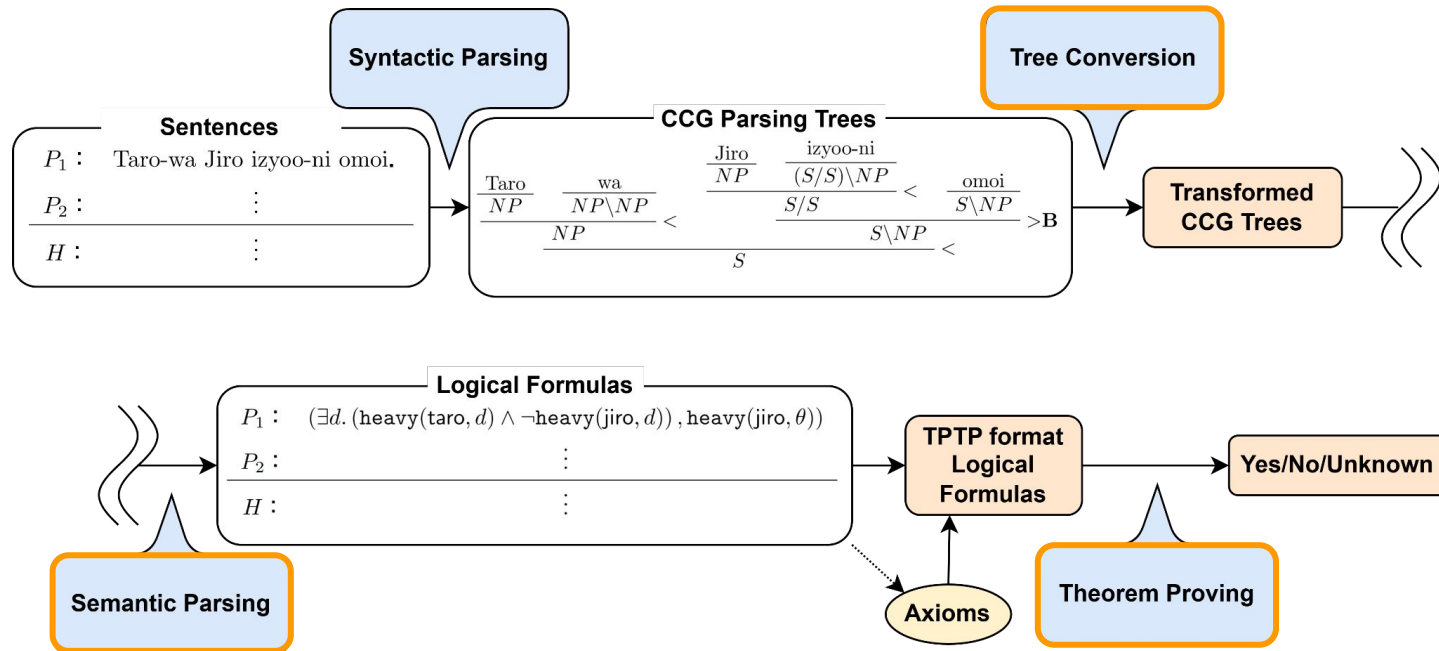
- Results & Discussion

Overview of Our System (ccg-jcomp)

What We Mainly Implemented

1. Rules to modify the output of the CCG parser
2. Semantic representations assigned to a CCG tree
3. Axioms for Japanese comparatives

Outline

- Overview of Our Proposed System

- **Linguistic Properties Unique to Japanese**

- Experiment

- Results & Discussion

Linguistic Properties Unique to Japanese

- We handled the following properties

  - Prop 1: **Absence of comparative morphemes**

  - Prop 2: **Approximate interpretation of equatives**

  - Prop 3: **Expressions with presupposition**

Prop 1: Absence of Comparative Morphemes

- English has comparative morphemes such as *-er, more*

- Japanese has no such morphemes

  - Taro is <u>heavy</u>　　　　　→　　　Taro-wa <u>omoi</u>.

  - Taro is <u>heavier</u> than Jiro.　　→　　Taro-wa Jiro yori <u>omoi</u>.

Same surface form

Prop 1: Solution (1/2)

- English has comparative morphemes such as *-er, more*

- Japanese has no such morphemes

  - Taro is <u>heavy</u>   →   Taro-wa **θ yori** <u>omoi</u>.

  - Taro is <u>heavier</u> than Jiro.   →   Taro-wa Jiro yori <u>omoi</u>.

Same surface form

Prop 1: Solution (2/2)

- We insert an unpronounced symbol (**empty category**) *cmp* when there is no standard of comparison

⇒ **We can uniformly handle the meaning of sentences with & without the standard of comparison**

Prop 1: Solution (Example)

$$\frac{\dfrac{\text{Jiro}}{NP} \quad \dfrac{\text{yori (than)}}{(S/S)\backslash NP}}{\dfrac{S/S}{: \lambda S.S(\lambda A\,x.\exists d.(A(x,d) \wedge \neg A(\mathsf{jiro},d)))}} <$$

$$\frac{\text{omoi (heavy)}}{S\backslash NP} \quad : \lambda Q\,N.Q(\lambda x.N(\mathsf{heavy},x))$$

$$\frac{S\backslash NP}{: \lambda Q.Q(\lambda x.\exists d.(\mathsf{heavy}(x,d) \wedge \neg\mathsf{heavy}(\mathsf{jiro},d)))} > \mathbf{B}_{\times}$$

Semantic composition of "Jiro yori omoi"

$$\frac{\dfrac{cmp}{S/S} \quad : \lambda S.S(\lambda A\,x.A(x,\theta)) \qquad \dfrac{\text{omoi (heavy)}}{S\backslash NP} \quad : \lambda Q\,N.Q(\lambda x.N(\mathsf{heavy},x))}{\dfrac{S\backslash NP}{: \lambda Q.Q(\lambda x.\mathsf{heavy}(x,\theta))}} > \mathbf{B}_{\times}$$

Semantic composition of "cmp omoi"

Prop 2: Approximate Interpretation of Equatives

- Japanese equatives have an approximate interpretation

- English equatives

  - John is as heavy as Bob.

  - Interpretation: John is **at least** as heavy as Bob.

- **Japanese equatives**

  - Taro-wa Jiro to onaji kurai-no omosa-da. (Taro is as heavy as Jiro)

  - Interpretation: Taro's weight is **close to** Jiro's.

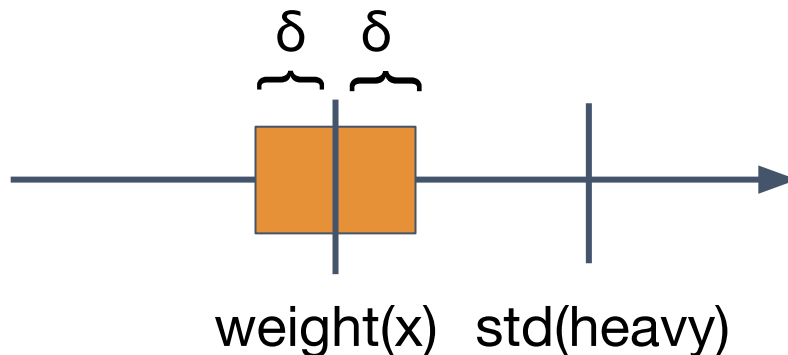- **Can be true even when Taro's weight is slightly less than Jiro's.**

Prop 2: Solution

- Introduce the following semantic representation (simplified version)

$$\left|\text{weight}(\text{taro}) - \text{weight}(\text{jiro})\right| < \delta$$

- Add a new axiom for δ and the standard degree to be heavy

$$\forall x.\text{weight}(x) - \delta < \text{std}(\text{heavy}) \Leftrightarrow \text{weight}(x) + \delta < \text{std}(\text{heavy})$$

weight(x)   std(heavy)

Prop 3: Expressions with Presupposition

- Japanese comparatives have expressions with presupposition
  [Kubota 2012, Hayashishita 2007]

- **Presupposition**: not affected by <u>entailment-canceling operators</u>
  [Potts 2015]

  - **"John knows Bob ran"** presupposes that **Bob ran**

  - **"John <u>does not</u> know Bob ran"** also presupposes that **Bob ran**

- In Japanese comparatives…

  - **"Taro-wa Jiro izyoo-ni omoi"** presupposes that **Jiro is heavy**

    Taro-Top Jiro than   heavy

Prop 3: Solution

- Multidimensional semantics [Karttunen+ 1979]

- Expresses the meaning as a pair <at-issue content, presupposition>

Taro-wa Jiro
izyoo-ni omoi ⇨ $\langle \underbrace{\exists d.\ \text{heavy}(\text{taro}, d) \wedge \neg\text{heavy}(\text{jiro}, d)}_{\text{Taro is heavier than Jiro}}, \underbrace{\text{heavy}(\text{jiro}, \theta)}_{\text{Jiro is heavy}} \rangle$

- Negated sentence ⇒ only negate the at-issue content

Taro-wa Jiro
izyoo-ni omoi
**toiu-wake-de
-wa-nai** ⇨ $\langle \underbrace{\neg\exists d.\ \text{heavy}(\text{taro}, d) \wedge \neg\text{heavy}(\text{jiro}, d)}_{\text{Taro is \textbf{not} heavier than Jiro}}, \underbrace{\text{heavy}(\text{jiro}, \theta)}_{\text{Jiro is heavy}} \rangle$

Outline

- Overview of Our Proposed System

- Linguistic Properties Unique to Japanese

- Experiment

- Results & Discussion

Experiment

- **JSeM** [Kawazoe+ 2017] (comparatives section)

  - For evaluation of semantic theories and processing systems for Japanese

  - We changed problems involving temporal and modal expressions

- The number of problems is **71**

- We use **GPT-4o**, **Swallow 8B**, and **Swallow 70B** as baselines

  - Swallow: the continual pre-trained model for Japanese

Outline

- Overview of Our Proposed System

- Linguistic Properties Unique to Japanese

- Experiment

- **Results & Discussion**

Results

- Accuracy of each system

- **ccg-jcomp outperformed all of the baseline models**

Majority	GPT-4o	Swallow 8B	Swallow 70B	**ccg-jcomp**
.592	.774	.549	.712	**.845**

Feature of GPT-4o & Swallow 70B

They tended to answer the problems involving **presupposition** incorrectly

Premise: Taro-wa Hanako **izyoo-ni** hayaoki da.
(Taro is an earlier riser than Hanako.)

Hypothesis: Hanako-wa hayaoki da.
(Hanako is an early riser.)

Gold Answer: Entailment

Model Answer: Neutral

Error Analysis (ambiguity)

- JSeM contains problems with ambiguity

> "ITEL-wa APCOM-no keiyaku-yori ooku-no chuumon-o kakutoku-sita"
> (ITEL won more orders than APCOM's contracts.)

- "APCOM-no keiyaku" (APCOM's contracts) has two meanings
    - Contracts that APCOM won
    - Contracts that ITEL won from APCOM

- To handle this, we need
    - To add a new semantic representation to "yori"
    - System for distinguishing the two interpretations

Error Analysis (adjective-like predicates)

- Japanese has predicates that behave like an adjective

> "Kono boo-wa ano boo-yori magatte-i-ru."
> (This stick is more bent than that stick.)

- "Magar-u" (bend) is a verb, but "magatte-i-ru" (bent) behaves like an adjective

- ccg-jcomp treats it as a verb and failed to parse the sentence correctly.

- To handle this, we need to give an exceptional representation to "te-i-ru"

Summary

We proposed **ccg-jcomp**, a logical inference system for **Japanese comparatives**

- Focused on some linguistic properties **unique to Japanese** related to comparatives

- **Results**: Ccg-jcomp showed **higher accuracy** than several LLMs

Future Work

Our system has several limitations

- **Scalability**

  - To handle phenomena other than comparatives, we need to add other

    frameworks. ⇒ not trivial

- **Small number of problems**

  - The number of problems is a bit small to see the robustness of our system

    and to evaluate the LLMs in few-shot settings.

Thank You for Listening!

Our system is available on https://github.com/ynklab/ccg-jcomp

References (1/2)

- Koji Mineshima, Pascual Mart́ınez-Ǵomez, Yusuke Miyao, and Daisuke Bekki. Higher-order logical inference with compositional semantics. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 2055–2061, 2015.
- Koji Mineshima, Ribeka Tanaka, Pascual Mart́ınez Ǵomez, Yusuke Miyao, and Daisuke Bekki. Building compositional semantics and higher-order inference system for a wide-coverage japanese ccg parser. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 2236–2242, 2016.
- Tomoki Sugimoto and Hitomi Yanaka. Compositional semantics and inference system for temporal order based on Japanese CCG. In Samuel Louvan, Andrea Madotto, and Brielen Madureira, editors, Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, pages 104–114, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.acl-srw.10. URL https://aclanthology.org/2022.acl-srw.10/.
- Izumi Haruta, Koji Mineshima, and Daisuke Bekki. Implementing natural language inference for comparatives. Journal of Language Modelling, 10(1):139–191, Nov. 2022. doi: 10.15398/jlm.v10i1.294. URL https://jlm.ipipan.waw.pl/index.php/JLM/article/view/294.
- Lasha Abzianidze. 2015. A tableau prover for natural logic and language. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 2492–2502, Lisbon, Portugal. Association for Computational Linguistics.

References (2/2)

- Jean-Philippe Bernardy and Stergios Chatzikyriakidis. 2017. A type-theoretical system for the FraCaS test suite: Grammatical framework meets coq. In Proceedings of the 12th International Conference on Computational Semantics (IWCS) — Long papers.
- Hai Hu, Qi Chen, Kyle Richardson, Atreyee Mukherjee, Lawrence S. Moss, and Sandra Kuebler. 2020. MonaLog: a lightweight system for natural language inference based on monotonicity. In Proceedings of the Society for Computation in Linguistics 2020, pages 334–344, New York, New York. Association for Computational Linguistics
- Mark Steedman. 2000. The Syntactic Process. MIT press.
- Max J Cresswell. The semantics of degree. In Montague grammar, pages 261–292. Elsevier, 1976.
- Ewan Klein. A semantics for positive and comparative adjectives. Linguistics and philosophy, 4:1–45, 1980.
- Lauri Karttunen and Stanley Peters. Conventional Implicature. In Presupposition, pages 1–56. Brill, 1979.
- Ai Kawazoe, Ribeka Tanaka, Koji Mineshima, and Daisuke Bekki. An inference problem set for evaluating semantic theories and semantic processing systems for japanese. In Mihoko Otake, Setsuya Kurahashi, Yuiko Ota, Ken Satoh, and Daisuke Bekki, editors, New Frontiers in Artificial Intelligence, pages 58–65, Cham, 2017. Springer International Publishing. ISBN 978-3-319-50953-2.

Problems we removed (temporal expression)

Premise 1: Taro-wa Hanako-yori 30 pun hayaku kitaku-sita.
(Taro came home 30 minutes earlier than Hanako)

Premise 2: Hanako-wa 6 ji tyodo-ni kitaku-sita.
(Hanako came home exactly 6 o'clock.)

Hypothesis: Taro-wa 5 ji han-ni kitaku-sita.
(Taro came home at 5:30.)

Gold Answer: Entailment

Problems we changed (modals)

Premise: Taro-wa Hanako izyoo-ni hayaoki kamo-sire-nai.

(Taro may be an earlier riser than Hanako.)

Hypothesis: Hanako-wa hayaoki-da.

(Hanako is an early riser.)

Gold Answer: Entailment

Premise: Taro-wa Hanako izyoo-ni hayaoki toiu-wake-de-wa-nai.

(Taro is not an earlier riser than Hanako.)

Hypothesis: Hanako-wa hayaoki-da.

(Hanako is an early riser.)

Gold Answer: Entailment