

# Natural Language Processing, Computational Linguistics (Miyao Group)

Homepage <https://myNlp.is.s.u-tokyo.ac.jp/>  
E-mail [yusuke@is.s.u-tokyo.ac.jp](mailto:yusuke@is.s.u-tokyo.ac.jp)

- Create theories and technologies for human language ability on computers
- Build a better world by applying natural language processing technologies

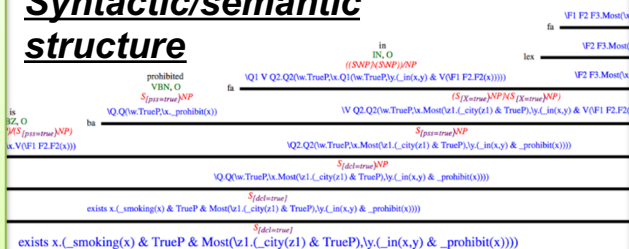
## Syntactic/semantic parsing

- Compute syntactic/semantic structures of sentences
- Parse sentences into syntactic trees
  - Compose semantic structures along syntactic trees
  - Syntax/semantics-aware language models

### Semantic inference

*Smoking is prohibited in most cities.*  
→ *Smoking is allowed in some cities.*

### Syntactic/semantic structure



## Dialogue system

- Exchange information with a computer using natural language
- Integrated application of natural language analysis/generation
  - Context-dependent semantic analysis and text generation
  - Verbalizing non-linguistic information

## Natural language

## Grounding

- Connect language with the real world
- Ground language on images, numerical data, database, etc.
  - Integrated understanding of linguistic/non-linguistic data

*Smoking is prohibited in most cities.*

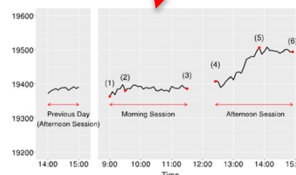
*Where was the author of Edo Hyakkei born?*

*Nikkei continues to fall. The closing price of the morning session decreases by 5 yen to 19,386 yen.*

*A cat is playing with a ball.*



**Image**



**Numerical data**



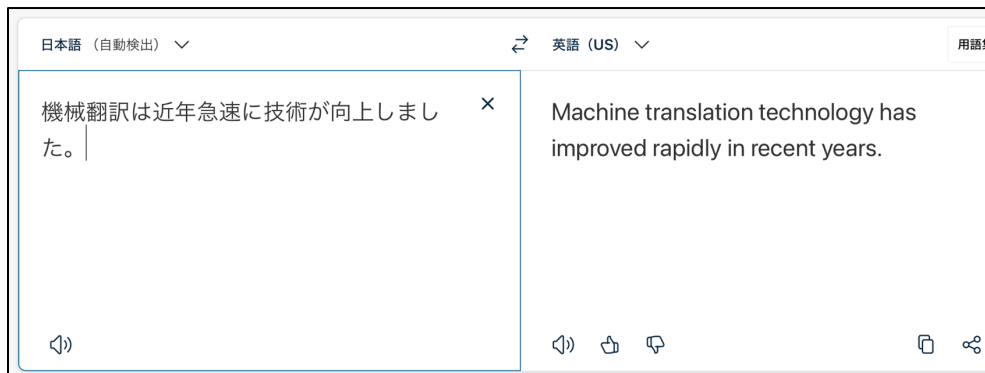
## Knowledge graph

Refer to the slide for students' research

# What is Natural Language Processing (NLP)?

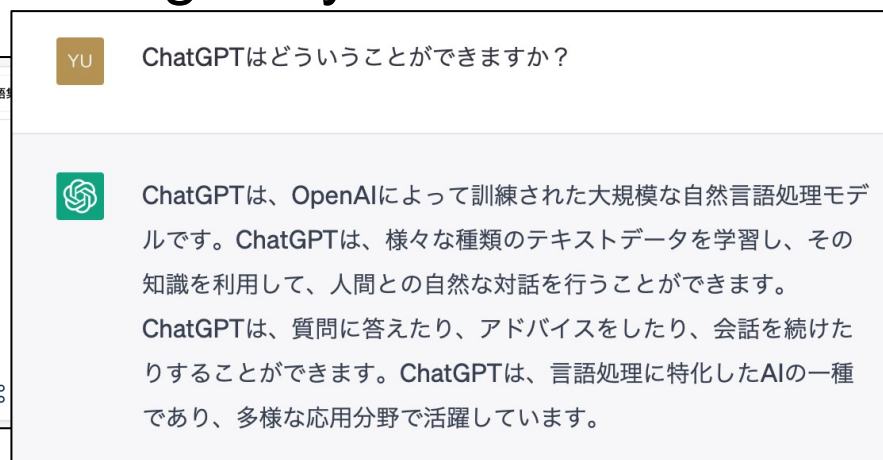
- Explore computational models of human linguistic ability (understanding/generating natural language)
  - Semantic understanding, reasoning, thinking, communication
  - Formal theory and technologies for natural language understanding
- Support human activity by NLP technologies
- Target any problem where input and/or output is natural language

## Machine translation



DeepL

## Dialogue system



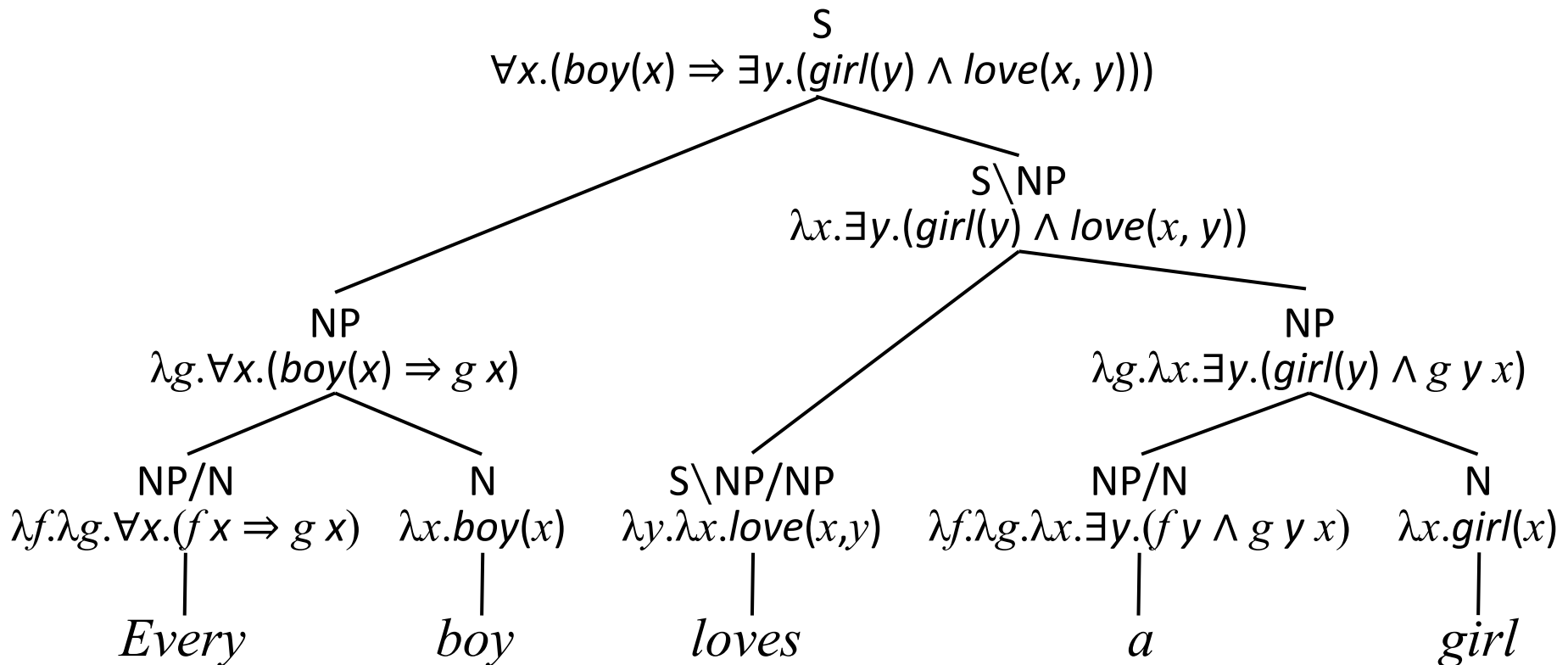
ChatGPT

# Research in our group

- Fundamental theories and technologies for natural language, with particular interests in:
  - **Syntactic/semantic parsing:** analyze the structure of sentences and compute their semantics
  - **Grounding:** Ground natural language semantics on the real world
  - **Dialogue system:** Achieve natural language communication applying the above technologies
- Students research their own topics on natural language, not limited to the above
  - Refer to “Students’ research” for instances

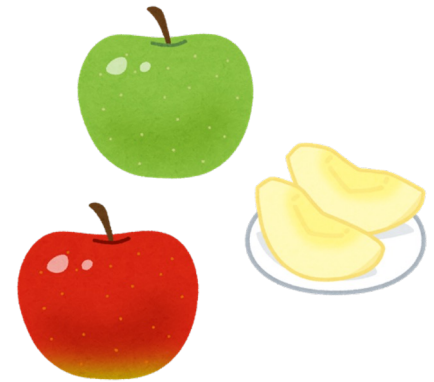
# Syntactic/semantic parsing

- Computing parse tree & semantics of natural language
- Fundamental problem to compute the **internal structure** of language
- Compute a parse tree of a sentence, and compose its semantic representation



# Grounding

- Connect natural language (our brain) to the real world (outside)
  - What is “apple”?
  - I’d like to visit “Hongo”
- Develop semantic representations and technologies for grounding

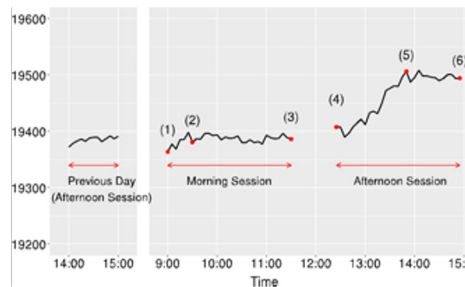


*A cat is playing with a ball.*



**Vision**

*Nikkei continues to fall. The closing price of the morning session decreases by 5 yen.*



**Numerical data**

*Where was the author of Edo Hyakkei born?*



**Knowledge graph**

# Dialogue system

- Communicate with a computer using natural language
- Integrated application that requires a variety of NLP technologies
  - Syntactic/semantic analysis, context understanding, language model, grounding
- Explore NLP technologies via the real-world application

Knowledge graph



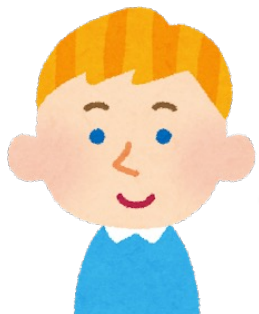
```
SELECT DISTINCT ?uri WHERE {  
  res:Edo_Hyakkei dbo:author ?x .  
  ?x dbo:birthplace ?uri .  
}
```

I went to an art gallery.

Nice! Something interesting?

Like Edo Hyakkei

Really?  
Its author is from near my home.



# Students' research

- Students' research is on broad topics about natural language, while it is recommended to explore the nature of natural language rather than focusing on applications
- Theories/technologies for NLP
  - Multilingual syntactic parsing
  - Syntax-aware language models
  - Formal theory of argumentation structure
  - Controlling automatic text generation
- Analyzing natural language using computer science
  - Analyzing universal properties of natural language syntax
  - Analyzing emergent language
- Grounding
  - Linking images and databases
  - Integrated analysis of natural language and math formulas
- Dialogue systems
  - Goal-oriented dialogue systems based on large language models