Lab 4. Assignment guidelines & Building dependency parser 1

LING-581-Natural Language Processing 1

Instructor: Hakyung Sung

September 18, 2025

Assignment guidelines

Overview

1. Background research brief (Due: 10/10, Friday)
2. Final project proposal (Due: 11/07, Friday)
3. Assignment (Due: 11/21, Friday)
4. Final paper (Due: 12/11, Thursday)

Overview

1. Background research brief (Due: 10/10, Friday)
2. Final project proposal (Due: 11/07, Friday)
3. Assignment (Due: 11/21, Friday)
4. Final paper (Due: 12/11, Thursday)

1. Background Research Brief

- **Topic / Area**
  - One sentence stating the focus
  - 3-5 keywords
- **Research Question / Problem**
  - 1-2 sentences clearly stating the core question or hypothesis
- **Mini Annotated Bibliography**
  - **Full citation** (use a consistent style (e.g., APA, *`acl_natbib`*)
  - **1-sentence contribution** (key finding/idea)
  - **Relevance** (why it might be matter for your project) (optional)
  - **Method/Data** (e.g., corpus, model, experiment) (*optional*)
- No lecture on October 9th (Thursday); individual group meeting (10 mins); Please aim to have a working draft ready for your group check-in.

2. Final project proposal

- Each group should submit a developed version of the background research brief.
- This proposal extends and deepens what you wrote—same project, more concrete plans for the final project.
- This is a graded assignment: 50 points total (10% of final grade).

Assignment page: Final project proposal

3. Assignment

11/6	Background research presentation (1, 2)
11/11	Background research presentation (3, 4)
11/13	Background research presentation (5, 6)
11/18	Background research presentation (7, 8)
11/20	Background research presentation (9)

- Each group should listen to and learn other students' research interests by actively listening to 8 presentations (not your own group) and write summaries.
- This is a graded assignment: 50 points total (10% of final grade).

Assignment page: Assignment

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?
    1. Introduction

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?
    1. Introduction
    2. Related work (*your findings from the background research go here)

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?
    1. Introduction
    2. Related work (*your findings from the background research go here)
    3. Method/Approach

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?
    1. Introduction
    2. Related work (*your findings from the background research go here)
    3. Method/Approach
    4. Experiments (including data, setup)

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?
    1. Introduction
    2. Related work (*your findings from the background research go here)
    3. Method/Approach
    4. Experiments (including data, setup)
    5. Results

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?
    1. Introduction
    2. Related work (*your findings from the background research go here)
    3. Method/Approach
    4. Experiments (including data, setup)
    5. Results
    6. Conclusion

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?
  1. Introduction
  2. Related work (*your findings from the background research go here)
  3. Method/Approach
  4. Experiments (including data, setup)
  5. Results
  6. Conclusion
  7. Limitations, Ethical consideration, Acknowledgments, References...

4. Final presentation/project (TBA)

11/25	Final presentation (1, 2, 3)
11/27	**Thanksgiving break (No class)**
12/2	Final presentation (4, 5, 6)
12/4	Final presentation (7, 8, 9)

- Each group present 20 mins?
- Paper format: ACL-short paper format (4 pages), familiarity with LaTex/Overleaf?
  1. Introduction
  2. Related work (*your findings from the background research go here)
  3. Method/Approach
  4. Experiments (including data, setup)
  5. Results
  6. Conclusion
  7. Limitations, Ethical consideration, Acknowledgments, References...

More to be announced/questioned-answered:

11/4	Final project discussion, Q&A

Building dependency parser

Goal

Train a dependency parser on the provided CoNLL-formatted training data and generate predictions—heads (*HEAD*) and dependency labels (*DEPREL*)—for the test set.

Overview

- **Write your own parser** using the stack, buffer, and shift/reduce mechanisms discussed in class (more guidelines on the Colab)

Overview

- **Write your own parser** using the stack, buffer, and shift/reduce mechanisms discussed in class (more guidelines on the Colab)
- **Use an existing library/toolkit** to train a parser (you'll need to research options).

Overview

- **Write your own parser** using the stack, buffer, and shift/reduce mechanisms discussed in class (more guidelines on the Colab)
- **Use an existing library/toolkit** to train a parser (you'll need to research options).
- **Propose another reasonable method** not covered in class.

Overview

- **Write your own parser** using the stack, buffer, and shift/reduce mechanisms discussed in class (more guidelines on the Colab)
- **Use an existing library/toolkit** to train a parser (you'll need to research options).
- **Propose another reasonable method** not covered in class.
- **Tip.** If it feels difficulty, vibe-coding is encouraged; make sure your work is reproducible (fixed seeds, documented environment) and well documented/commented.

Data

Use the provided **train** / **dev** / **test** splits.

Link: *https://github.com/hksung/Fall25_PythonTutorial/tree/main/corpus/dep_parse*

Data Format (10 tab-separated fields)

1. *ID* (1-based)
2. *FORM* (word)
3. *LEMMA* (empty)
4. *CPOSTAG* (coarse POS)
5. *POSTAG* (fine POS)
6. *FEATS* (often `-`)
7. *HEAD* (0 for root; `-` if unknown)
8. *DEPREL* (label; `-` if unknown)
9. *DEPS* (often `-`)
10. *MISC* (often `-`)

Timeline

- **Week 4 (planning/check-in):** In small groups, briefly discuss parser designs or candidate libraries. Submit a short process write-up on MyCourses (a concise description is enough). *Submission earns full credit for this checkpoint.*
- **Week 5 (submission):** We'll continue during Lab 5 in class. Please submit the outputs listed in the **Deliverables** section by the following Friday (9/26).