Dependency Parsing

With a Dynamic Oracle

Jonatan Byman, Filip Edberger,
Victor Eriksson & David Norman

LINKÖPINGS
UNIVERSITET

Syntactic Parsing

- Analysis of syntactic structure of natural language
- Part of speech tagging

  Pre-existing part of the project and will not be evaluated.

- Dependency parsing

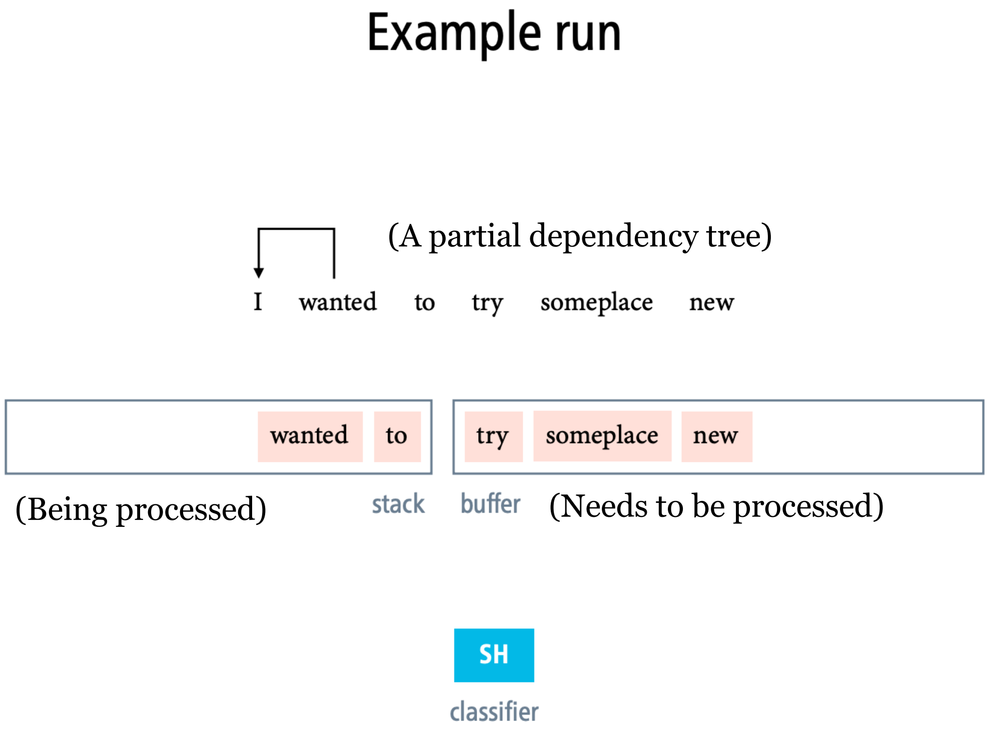
  Main part of the project.

Dependency parsing

- Baseline system
  - Arc Standard Algorithm
  - Non decomposable[1]
  - Static Oracle
- Extended system
  - Arc Hybrid Algorithm[2]
  - Dynamic Oracle
  - Decomposable[1]

1. https://aclanthology.org/Q13-1033.pdf
2. https://aclanthology.org/P11-1068.pdf

Configuration



Example run

(A partial dependency tree)

I wanted to try someplace new

wanted to

try someplace new

(Being processed) stack

buffer (Needs to be processed)

SH

classifier

https://www.ida.liu.se/~TDDE09/lectures/unit4/nlp-2024-405.pdf

Arc-Standard

- The shift transition (SH)
- The left-arc transition (LA)
- The right-arc transition (RA) ("Reversed" LA)

Arc-Hybrid

- (SH) Same as Arc-Standard
- (LA) Creates a dependency from the word in the buffer to the word on the stack. Removes the word from the stack
- (RA) Same as Arc-Standard

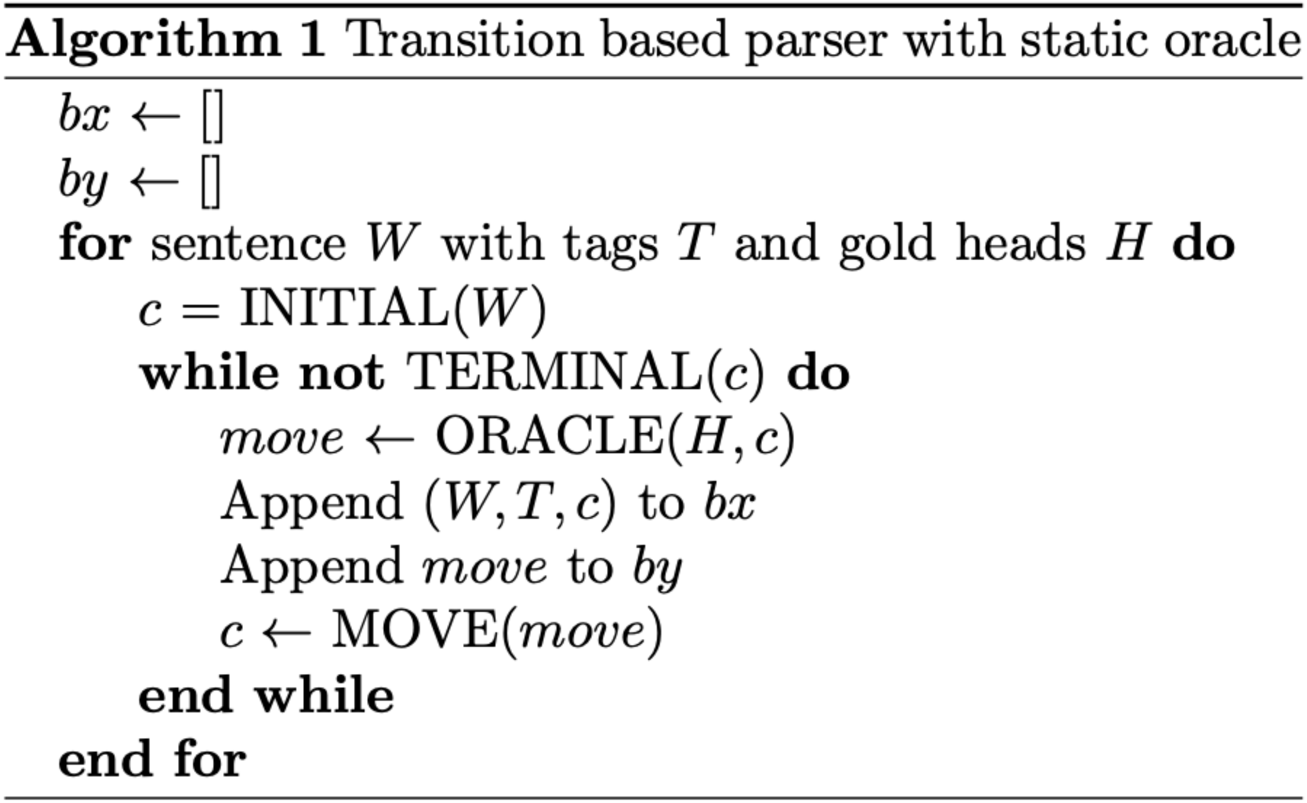
https://www.ida.liu.se/~TDDE09/lectures/unit4/nlp-2024-405.pdf
https://www.ida.liu.se/~TDDE09/commons/NLP-2021-45.pdf

Oracles

- Provides the parsing algorithm with transitions
- Static Oracle
  - Provides the gold-standard transition
  - Deterministic
- Dynamic Oracle
  - Provides a fixed set of correct transitions
  - Cost - function[1]
  - Non-Deterministic

[1] https://aclanthology.org/Q13-1033.pdf

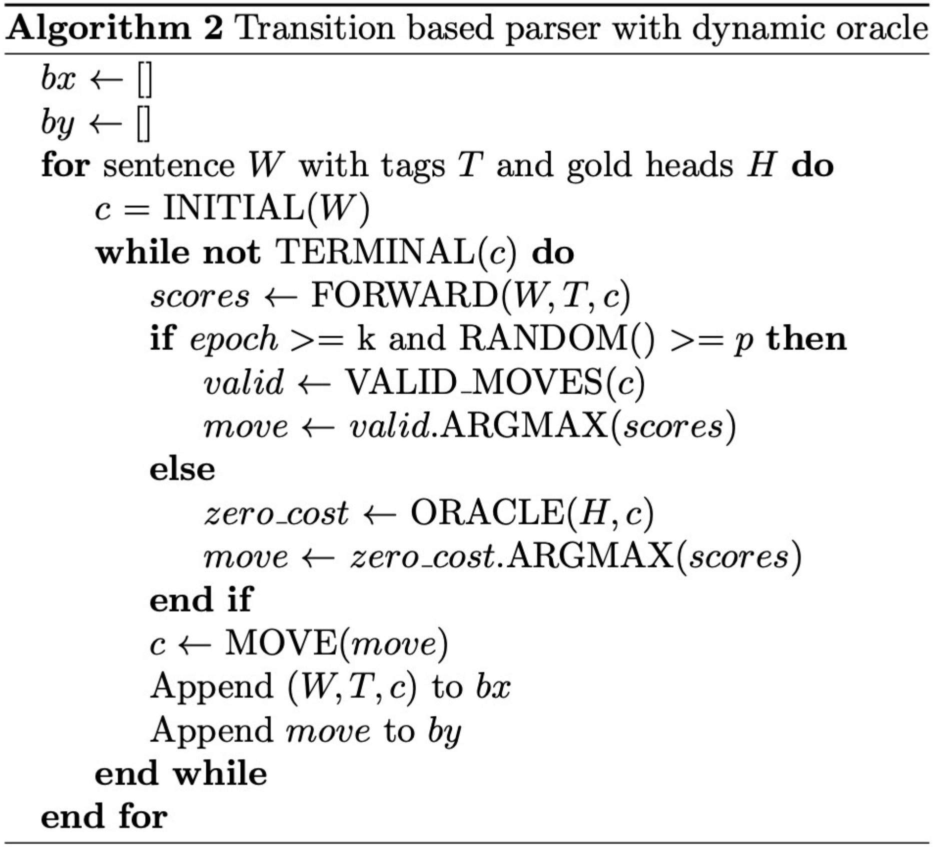
Psudo Code for the training loop
- Static oracle



Algorithm 1 Transition based parser with static oracle

```
$bx \leftarrow []$
$by \leftarrow []$
for sentence $W$ with tags $T$ and gold heads $H$ do
    $c$ = INITIAL($W$)
    while not TERMINAL($c$) do
        $move \leftarrow$ ORACLE($H, c$)
        Append ($W, T, c$) to $bx$
        Append $move$ to $by$
        $c \leftarrow$ MOVE($move$)
    end while
end for
```

Pseudocode for the training loop - Dynamic oracle



Algorithm 2 Transition based parser with dynamic oracle

```
$bx \leftarrow []$
$by \leftarrow []$
for sentence $W$ with tags $T$ and gold heads $H$ do
    $c$ = INITIAL($W$)
    while not TERMINAL($c$) do
        $scores \leftarrow$ FORWARD($W, T, c$)
        if $epoch >=$ k and RANDOM() $>= p$ then
            $valid \leftarrow$ VALID_MOVES($c$)
            $move \leftarrow valid$.ARGMAX($scores$)
        else
            $zero\_cost \leftarrow$ ORACLE($H, c$)
            $move \leftarrow zero\_cost$.ARGMAX($scores$)
        end if
        $c \leftarrow$ MOVE($move$)
        Append ($W, T, c$) to $bx$
        Append $move$ to $by$
    end while
end for
```

Results

- Average UAS over 10 runs with different random seeds
    - Comparison with similar work done by others
- Effect of p and k
    - Comparison with similar work done by others

10 runs with different seed

- Baseline system
  Arc Standard with static oracle
  - EN:
    - Avg UAS = 0.6621
    - σ = 0.0041
  - SV:
    - Avg UAS = 0.6706
    - σ = 0.0035

10 runs with different seed

- Arc Hybrid
  Static Oracle
  - EN:
    - Avg UAS = 0.6249
    - σ = 0.0037
  - SV:
    - Avg UAS = 0.6587
    - σ = 0.0023

- Arc Hybrid (k=0, p=0.9)
  Dynamic Oracle
  - EN:
    - Avg UAS = 0.6300
    - σ = 0.0058
  - SV:
    - Avg UAS = 0.6731
    - σ = 0.0080

Comparison with literature

- Paper by Yoav Goldberg and Joakim Nivre (2013)[1]

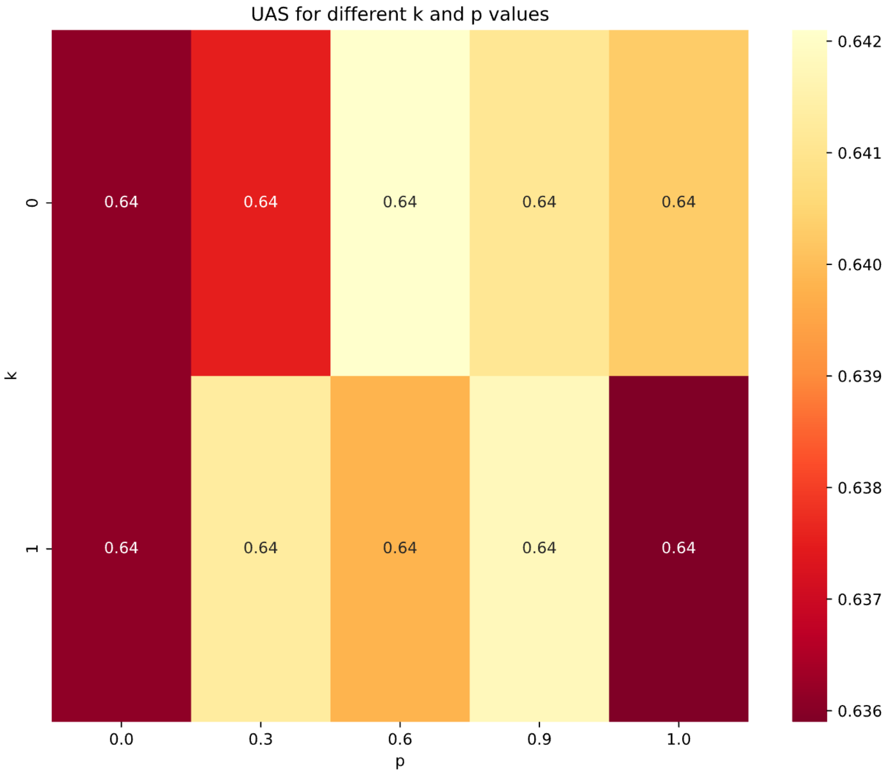
  - Avarage UAS, Hybrid parser, Dynamic oracle

- Theirs (EN):
  - Static: 86,43
  - Dynamic: 87,62
  - ca 1,35 %

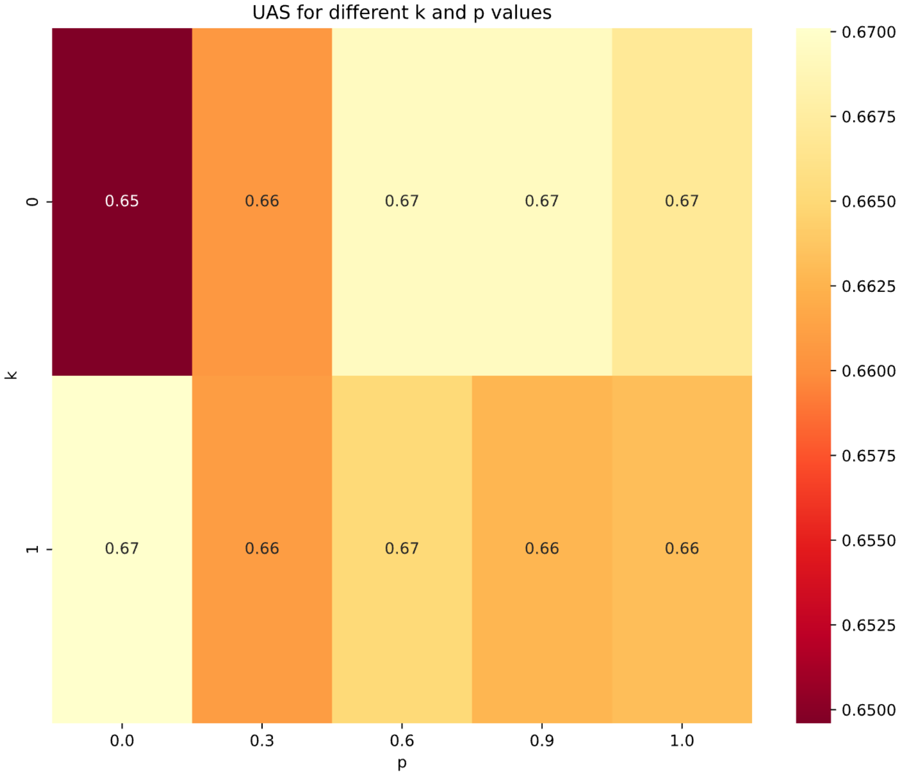
- Ours (EN):
  - Static: 62,49
  - Dynamic: 63,00
  - ca 0.81 %

[1] https://aclanthology.org/Q13-1033.pdf

Effect of the hyperparemters p and k



ENGLISH, PARSER TAGS, 2 EPOCHS



SWEDISH, PARSER TAGS, 2 EPOCHS

Comparison with literature

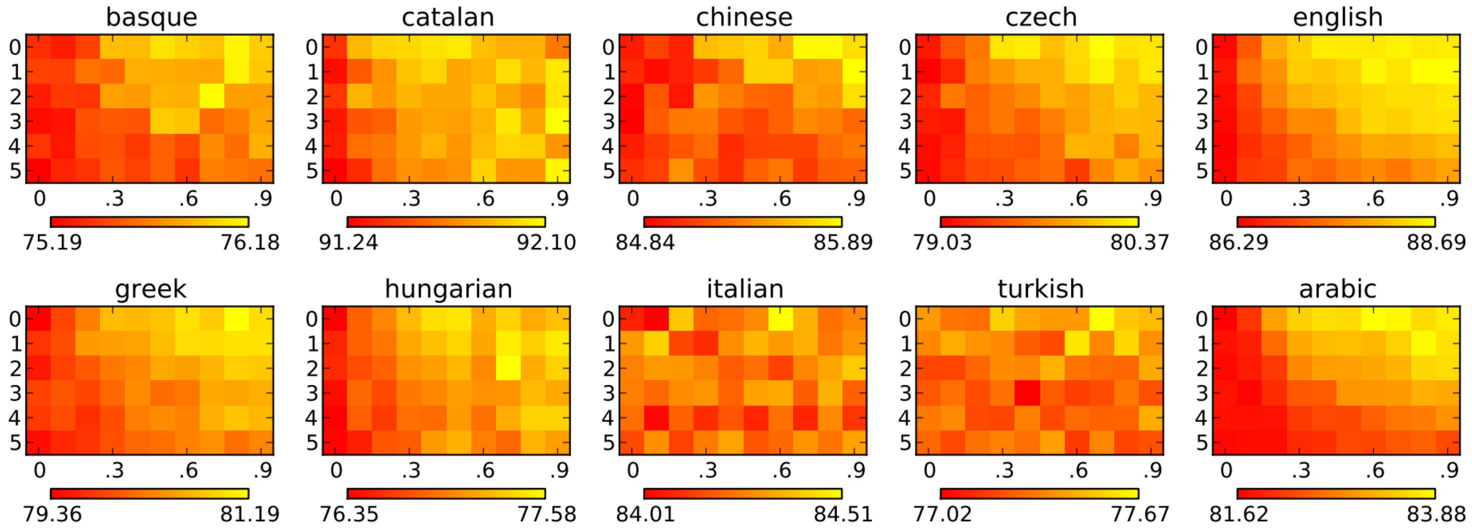


Figure 1: Effect of k (y axis) and p (x axis) values on parsing accuracies for the arc-eager system on the various CoNLL-2007 shared-task languages. Each point is an average UAS of 5 runs with different seeds. The general trend is that smaller k and higher p are better.

Thanks for listening

LINKÖPINGS UNIVERSITET