

Dependency parsing (係り受け解析)

Naoaki Okazaki

[okazaki at ecei.tohoku.ac.jp](mailto:okazaki@ecei.tohoku.ac.jp)

<http://www.chokkan.org/>

<http://twitter.com/#!/chokkanorg>

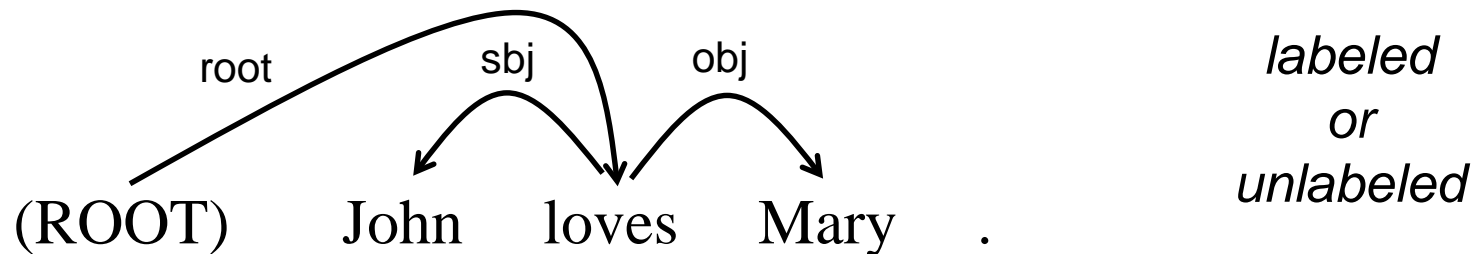
#nlptohoku

Acknowledgements

- Portions of this material are from:
 - Nivre and Kübler (2006). Dependency Parsing. *Tutorial at Coling-ACL 2006*.
 - 長尾真編 (1996). 自然言語処理. 岩波講座ソフトウェア科学(15).

Dependency (依存, 係り受け)

- Analyze syntactic structures of sentences using relationships between two words



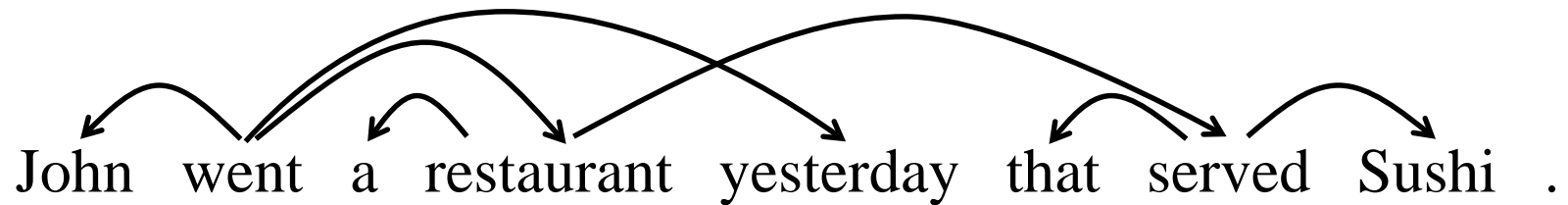
- $\text{subj}(\text{loves}, \text{John})$: "John" is a nominal subject of the verb "loves"
- $\text{obj}(\text{loves}, \text{Mary})$: "Mary" is a direct object of the verb "loves"

type *head* *dependent*
 governor *modifier*

- Fit to languages with scrambling (語順の入れ替え)

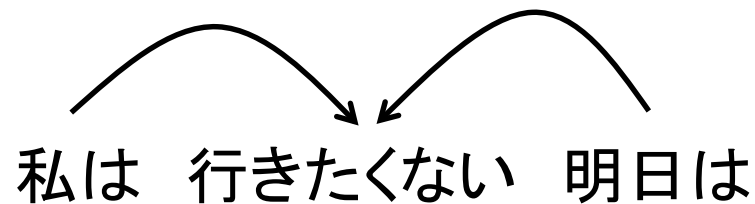
Non-projective dependency tree

- With crossing edges

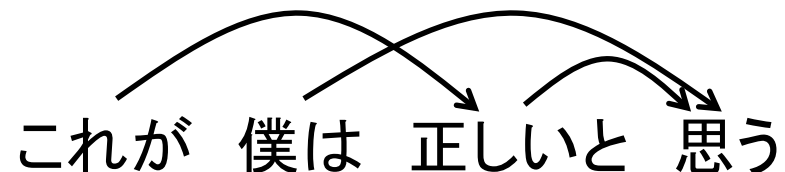


Japanese dependency grammar (黒橋・長尾, 1996)

- Dependency analysis has been explored extensively in Japanese, which allows scrambling (語順の入れ替え) and case ellipsis (格要素の省略)
- Reversed direction of dependencies (traditionally)
- Mostly right direction only
 - Exception:



- Most Japanese written sentences are non-projective
 - Exception:

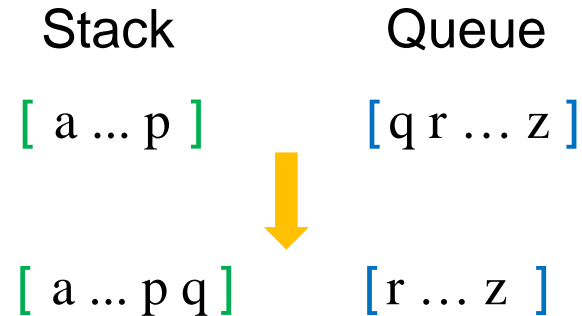


Arc-standard transitions

- Three actions

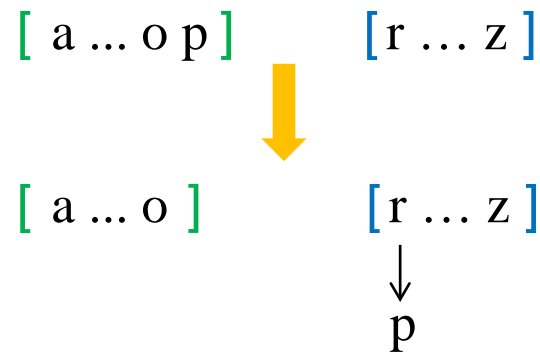
- Shift

- Pop a token from the queue
- Push the token into the stack



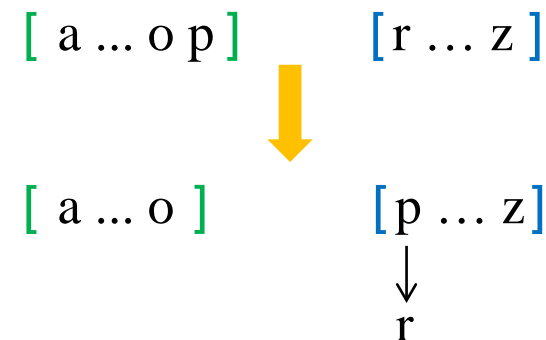
- Left-arc

- Pop a token from the stack
- Add an edge
 - From: the top token in the queue
 - To: the popped token



- Right-arc

- Pop a token from the stack
- Add an edge
 - From: the popped token
 - To: the top token in the queue



0: Initial state

Stack

Queue

[ROOT] [Economic news had little effect on financial markets .]

1: Shift

Stack

Queue

[ROOT Economic] [news had little effect on financial markets .]

2: Left-arc (nmod)

Stack

Queue

[ROOT]

[news had little effect on financial markets .]

↓ nmod

Economic

3: Shift

Stack

[ROOT news]

↓ nmod

Economic

Queue

[had little effect on financial markets .]

4: Left-arc (sbj)

Stack

Queue

[ROOT]

[had little effect on financial markets .]

↓ sbj

news

↓ nmod

Economic

5: Shift

Stack

[ROOT had]

↓ subj
news

↓ nmod
Economic

Queue

[little effect on financial markets .]

6: Shift

Stack

[ROOT had little]

↓ subj
news

↓ nmod
Economic

Queue

[effect on financial markets .]

7: Left-arc (nmod)

Stack

[ROOT had]

↓ subj
news

↓ nmod
Economic

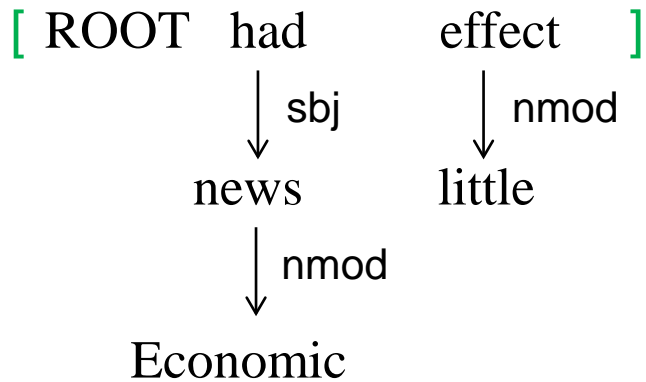
Queue

[effect on financial markets .]

↓ nmod
little

8: Shift

Stack



Queue

[on financial markets .]

9: Shift

Stack

[ROOT had effect on]

↓ subj
news

↓ nmod
Economic

↓ nmod
little

Queue

[financial markets .]

10: Shift

Stack

[ROOT had effect on financial]

↓ subj
news

↓ nmod
Economic

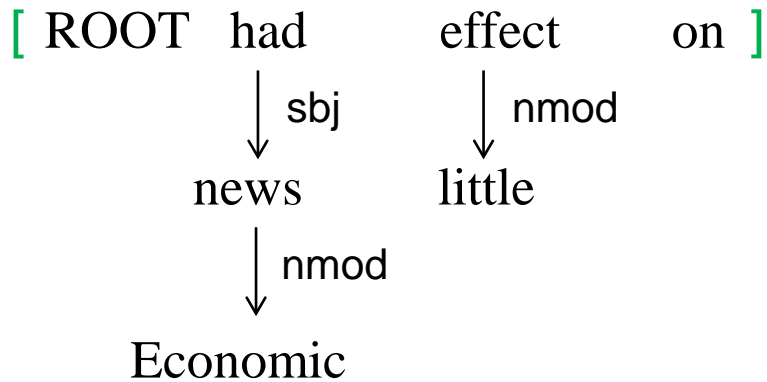
↓ nmod
little

Queue

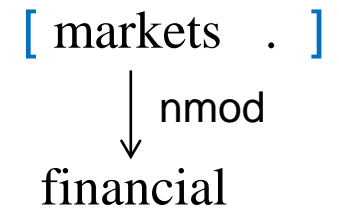
[markets .]

10: Left-arc (nmod)

Stack

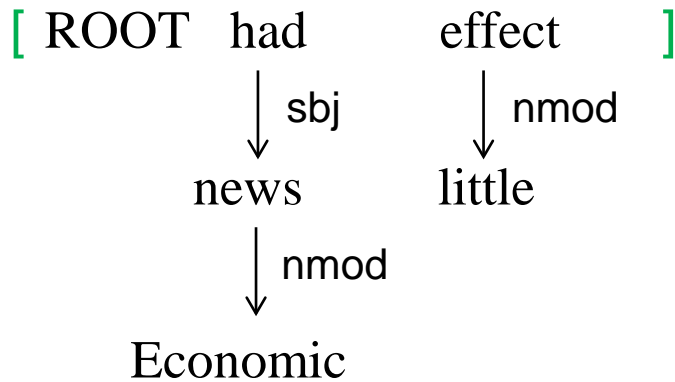


Queue

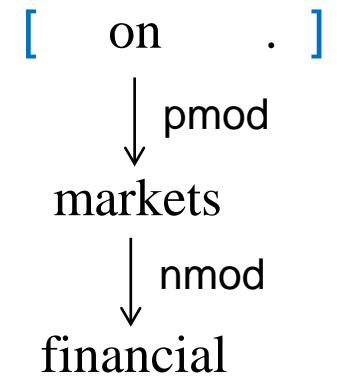


11: Right-arc (pmod)

Stack



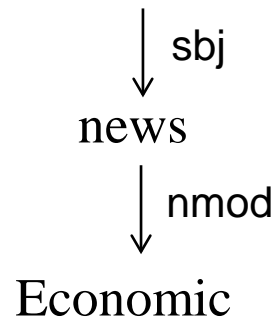
Queue



12: Right-arc (nmod)

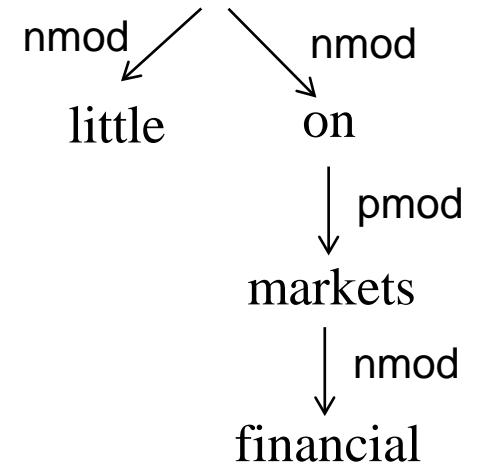
Stack

[ROOT had]



Queue

[effect .]

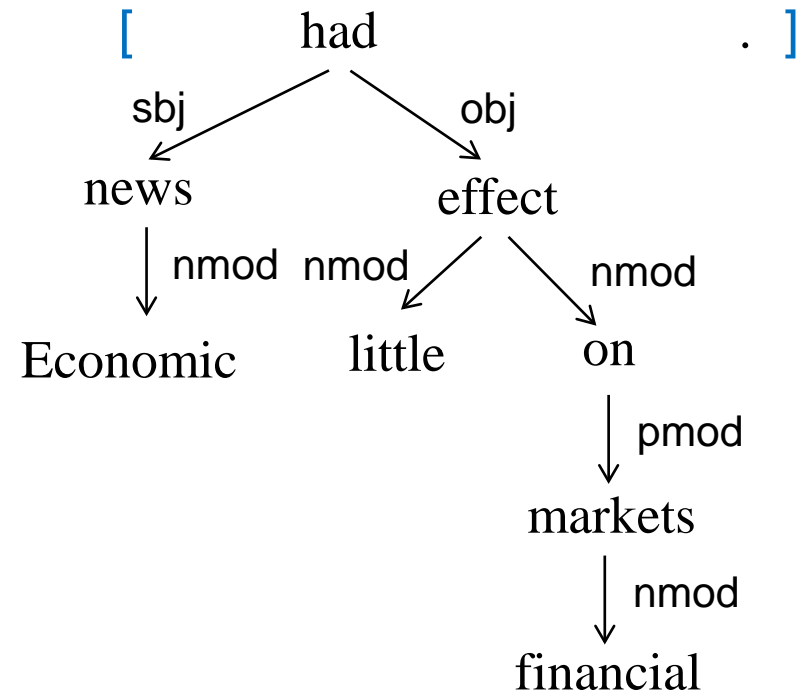


13: Right-arc (obj)

Stack

[ROOT]

Queue



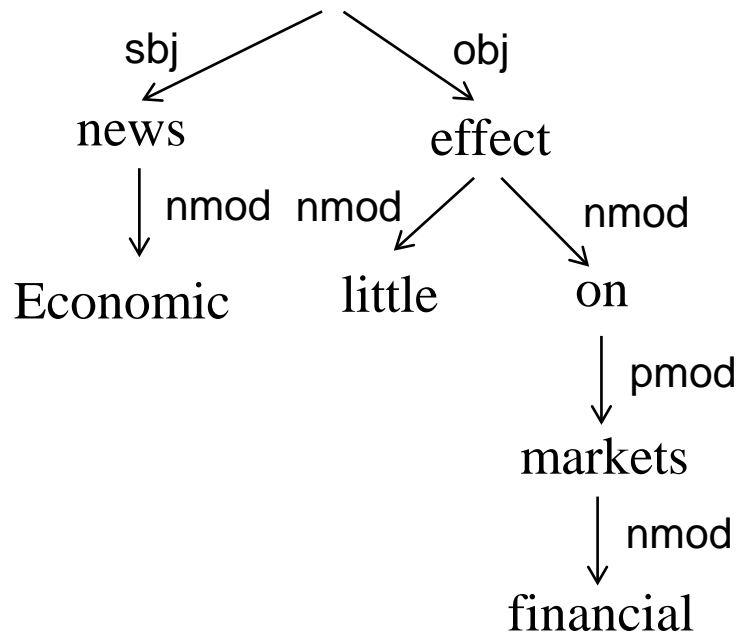
14: Shift

Stack

Queue

[ROOT had]

[.]

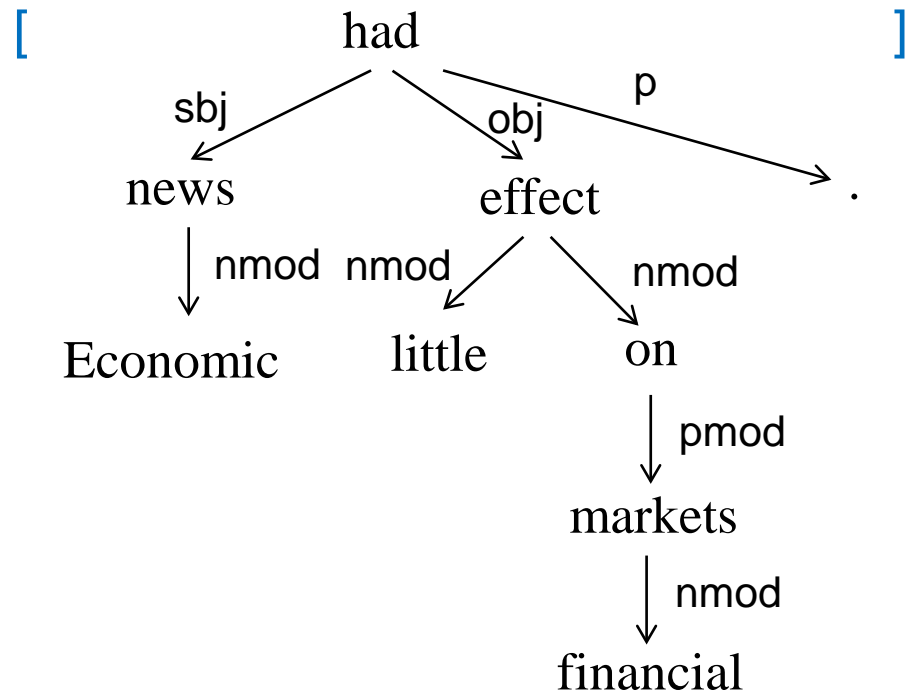


15: Right-arc (p)

Stack

[ROOT]

Queue



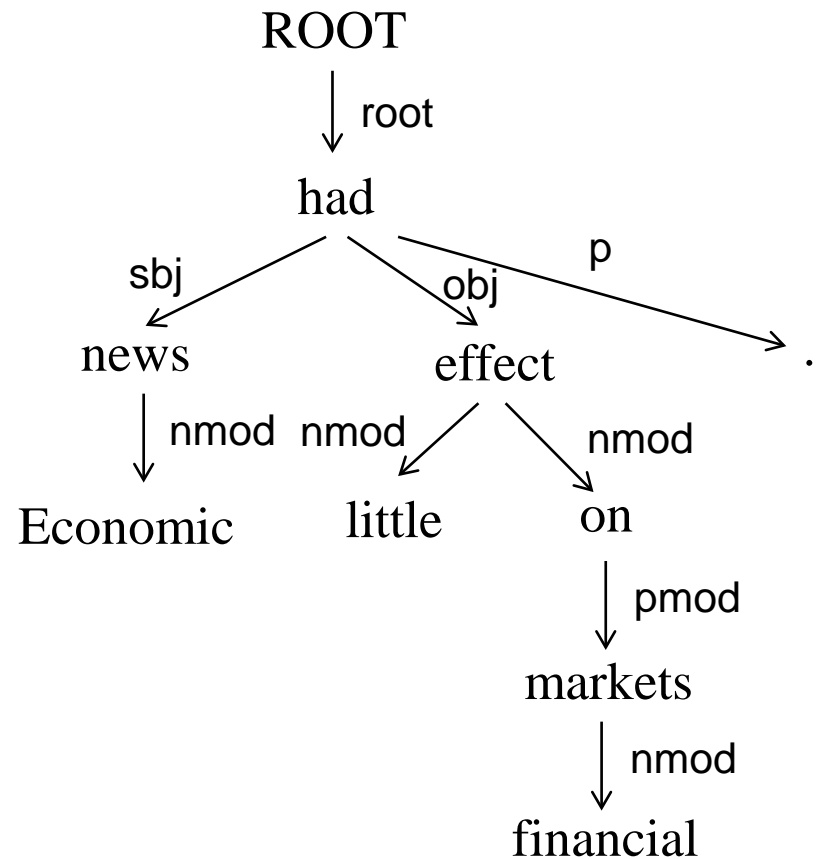
16: Right-arc (root)

Stack

[]

Queue

[]



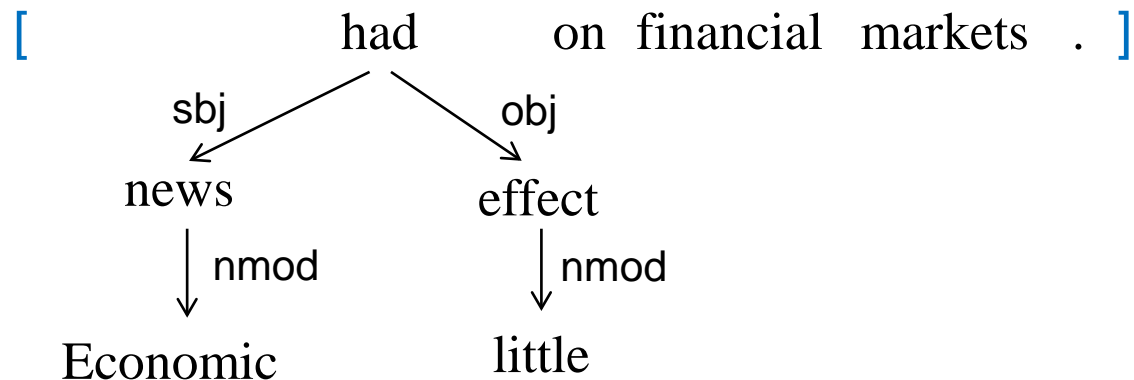
Shift-reduce conflicts

- What if we chose “Right-arc (obj)” at step 8?

Stack

[ROOT]

Queue



- We cannot attach the prepositional phrase *on financial markets* to *effect* (NP attachment) but only to *had* (VP attachment)!

Summary

- Two grammar formalisms and parsing algorithms
 - Constituency and dependency
- The process of obtaining a tree from a sentence
 - (P)CKY and shift-reduce algorithm
- Difficulty in parsing
 - Attachment and coordination
 - Time complexity and space requirement