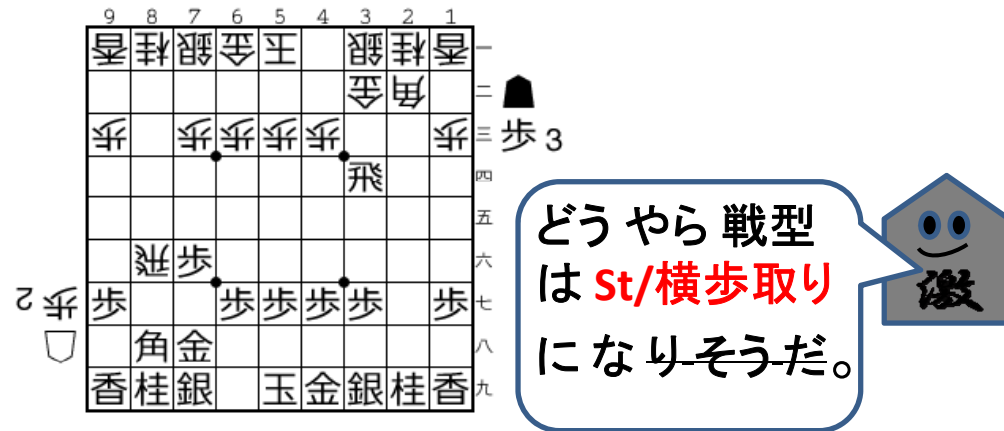


# Shogi Commentary Generation



Hiroataka Kameko (Univ. of Tokyo)  
Mori Shinsuke (Kyoto Univ.)

# Goal

Given as an input  
a game state

Generate  
commentaries



戦型は  
St/横歩取り  
になりった。



# Approach

1. Automatic **Symbol Grounding**

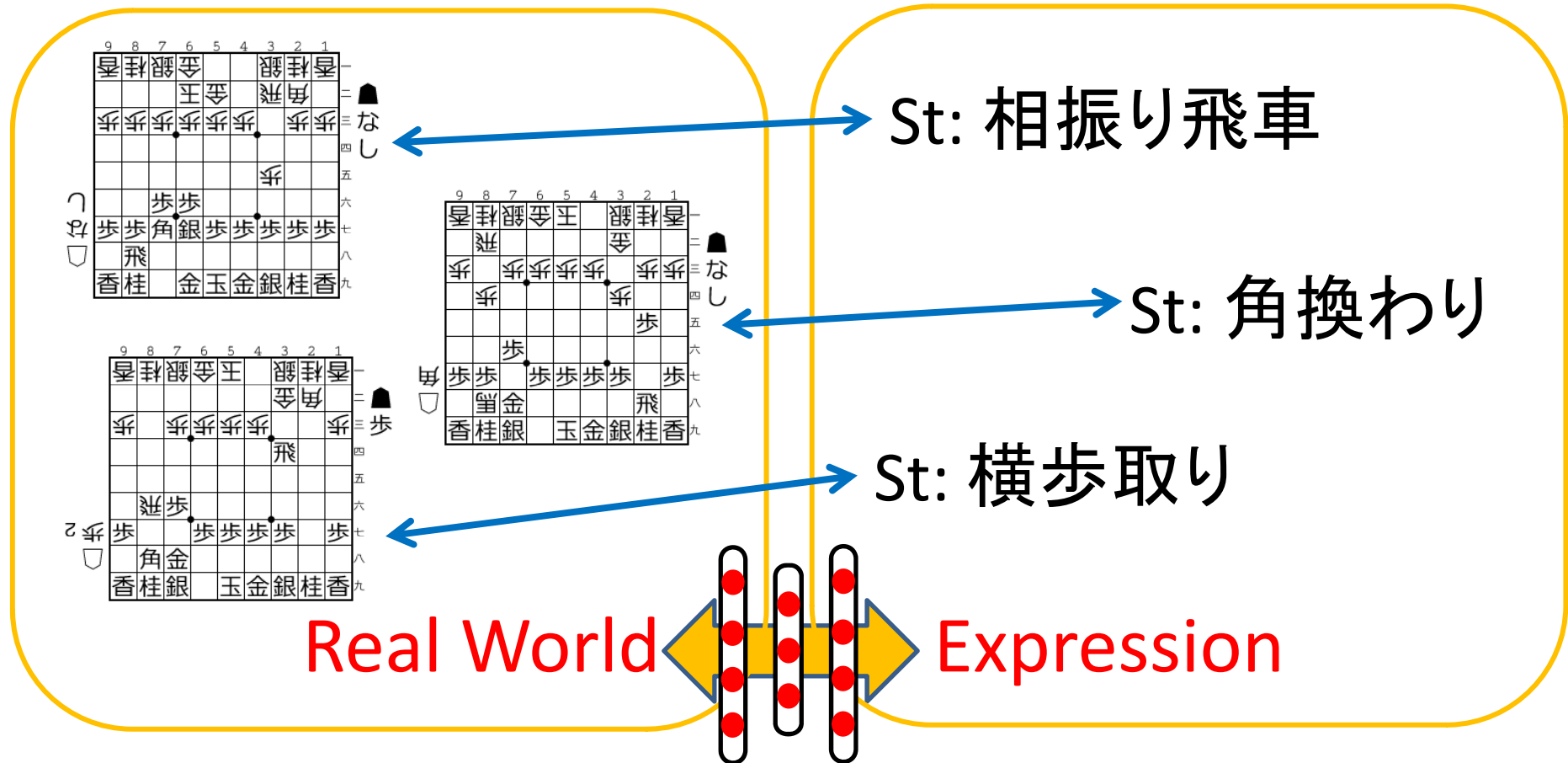
[Kameko et al., IPSJ2014 (in Japanese)]

2. Automatic **Skeleton Compilation**

[Mori et al., INLG2014]

3. Combine Module 1 and 2

# Symbol Grounding Training



- Three layer perceptron
- About 300,000 positions with commentaries (with lots of noisy commentaries ...)

# Available Resources

2014年2月8日 第7回朝日杯将棋オープン戦決勝  
羽生善治三冠 対 渡辺明二冠

2014/02/08 14:30  
第7回朝日杯将棋オ  
東京・有楽町朝日ホ  
持ち時間:各40分

残り時間 終局  
▲0時間00分  
△0時間00分

▼羽生善治

- \*110 △同 馬
- \*111 ▲同 歩
- \*112 △4九飛成
- 113 ▲7七玉
- \*114 △9九角
- \*115 ▲8八銀
- \*116 △同角成
- \*117 ▲同 玉

棋譜保存  
自動更新しない

この手を指したあと、渡辺が額に手をやって天をあおいだ。  
「△9九角に駒を使うわけがわからないので▲8八銀で...ん、これはいきなり詰みと言うことですか。△同角成▲同玉△7九銀と追って、これは詰んでいるように見えます。(佐藤康光九段)

- Position and Commentary
- Junisen C1, C2
  - 2003～2013
  - 3,786 match
  - 212,066 sent.

# NLP tool 1

## Word Segmentation

- KyTea (Pointwise) with the distributed model
  - 91.60 (Recall: 92.20, Precision: 91.00)
- Adaptation to commentary
  - Dictionary (7,209 move exp., 532 player names)
  - Full annotation
    - 3,299 sent., 24,966 words
  - Partial annotation
    - 3,000 sent. (1 s-NE for each sent)
  - 97.69 (Recall: 98.04, Precision: 97.34)

# NLP tool 2

## Shogi Named Entity Recognition

- Defined 23 Shogi named entity classes (s-NE)
- Full annotation
  - 3,299 sent., 24,966 words
- Partial annotation
  - 3,000 sent. (1 s-NE for each sent)
  - About 90%

# s-NE Class (1/2)

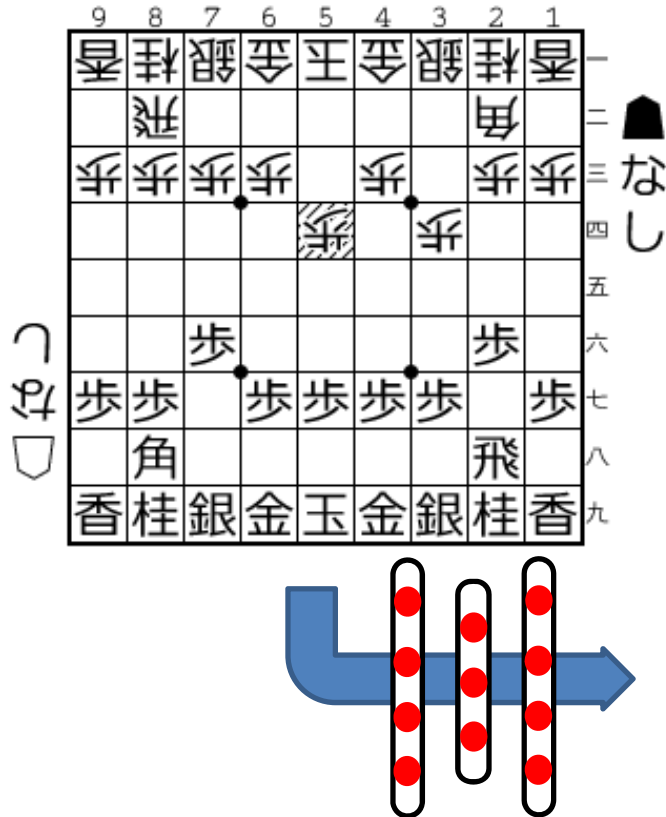
- Hu: human 人 (対局者や解説者を含む人, 羽生四冠, 先手,...)
- Tu: turn 手番 (先手, 後手, ▲, △)
- Po: position 位置 (81通りのみ)
- Pi: piece 駒 (成り駒を含め有限)
- Ps: piece specifier 駒の指定 (指し手が曖昧な場合の指定子)
- Mc: move compliment 動きの明確化 (成, 不成)
- Pa: piece attribute 駒の属性 (道, 利き, 頭, ...)
- Pq: piece quantity 駒の数 (1枚, 切れ)
- Re: region 盤面の領域 (中央, 駒台, 4筋, 3段目, ...)
- Ph: phase 対局の進行 (序盤, 中盤, 終盤, ...)
- Ai: aim 狙い
- St: strategy 戦型



# s-NE Class (2/2)

- Ca: castle 囲い
- Ce: castle eval 囲いの評価
- Me: move eval 指し手評価
- Mn: move name 指し手別名
- Ee: eval element 評価要素 (部分の評価のみ)
- Ev: eval 形勢評価 (局面全体の判断のみ)
- Ti: time 時間
- Ac: player action 対局者が主語の述語
- Ap: piece action 駒が主語の述語
- Ao: other action その他が主語の述語
- Ot: other 他の重要概念

# Characteristic Word List



## Perceptron Output

St/中飛車/1.82078

Re/5筋/1.94516

Ac/突/2.15669

Ac/使/2.35015

Tu/先手/2.71036

St/ゴキゲン中飛車/3.19339

Ca/から/3.78502

Ca/矢倉/3.81759

Ac/指/3.86289

Ac/位を取らせ/4.28983

Ac/警戒/4.5293

Pa/道/4.63135

Ot/小倉/4.66681

Ee/濃厚/4.82272

Ot/上村/4.84702

Re/6筋/4.94942

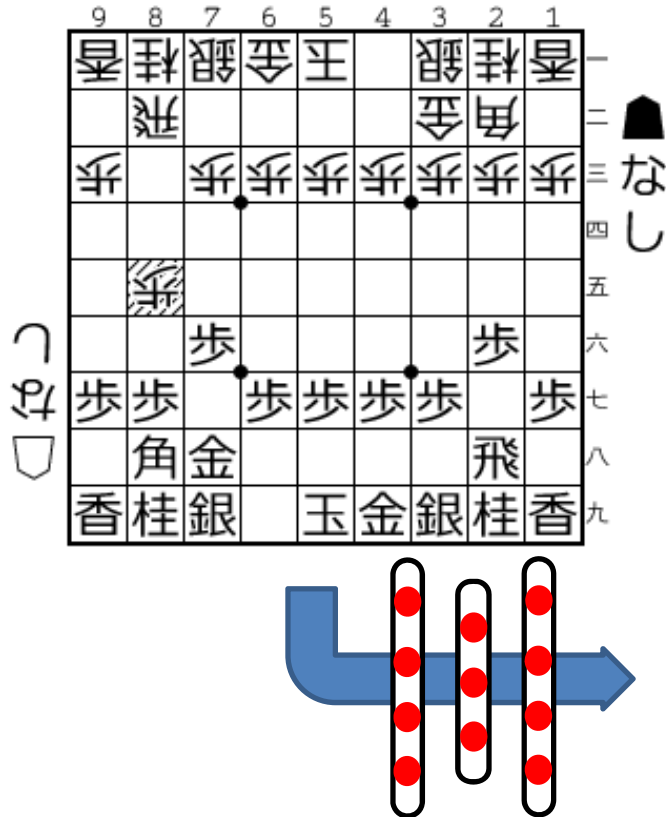
Ot/森下/4.98288

Ee/得意/5.02482

St/向飛車/5.06046

PN/山崎隆之七段/5.103

# Characteristic Word List



## Perceptron Output

Pi/角/0.664741

Ai/換わ/1.20699

St/横歩取り/1.81199

St/相掛か/2.57532

Ot/消費/2.74976

Pi/飛車/3.53324

Ac/突/3.68507

Ac/指/3.73438

St/居飛車/3.75575

St/横歩/4.13743

Ac/引/4.22017

Ac/上が/4.24186

St/一手損角/4.24559

Ee/手広/4.57861

Ac/早/4.62582

Ac/明示/4.78106

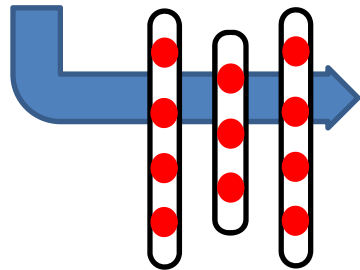
Ac/伸ば/4.81008

Ee/使/4.81875

Ca/穴熊/4.84669

Ac/用意/4.87284

# Characteristic Word List



## Perceptron Output

Ai/交換/1.59779

Ac/上が/1.8875

Ca/矢倉/2.75859

Pi/角/3.46435

Tu/先手/3.68673

St/力戦居飛車/3.72252

Ac/目指/3.84657

Ee/意味がな/4.10097

St/振り飛車/4.15819

Pi/歩/4.35746

Ca/から/4.36748

Pi/飛車/4.60326

Mn/打ち/4.80395

Ac/避け/4.89056

St/横歩取り/5.0024

St/中飛車/5.17757

Ac/振/5.55577

St/居飛車/5.56065

Ai/換わ/5.60267

St/相居飛車/5.61964

# From the **Characteristic Word List** to a Sentence

## 1. **Skeleton Pool** Compilation

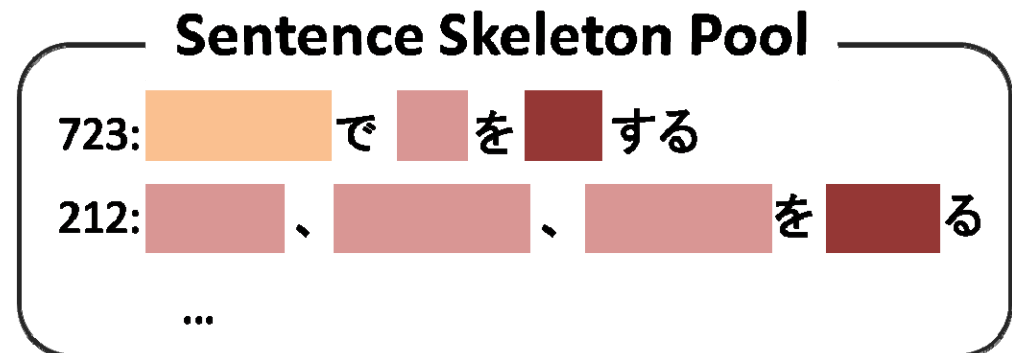
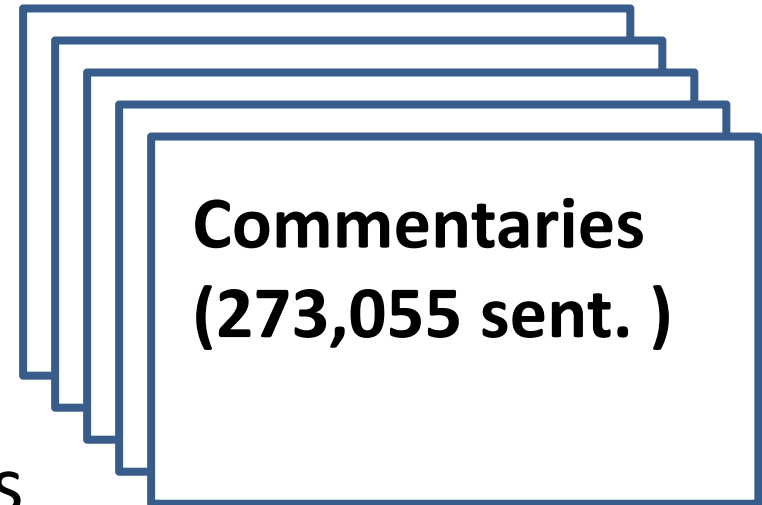
1. Run WS on huge comment examples
2. Run NER on the WS result
3. Replace s-NEs with the type labels

## 2. Sentence Generation

1. Select the most suitable **skeleton**
2. Fill s-NE slots with the **characteristic words**

# Skeleton Pool Compilation

1. Define Shogi NE (s-NE)
2. Prepare training data
3. Train NER
4. Run NER on commentaries
5. Replace s-NE with type symbol



# Sentence Generation



**Sentence Skeleton Pool**

723:   で   を   する

212:   、   、   を   る

...

Perceptron Output

<p><b>St/中飛車/1.82078</b></p> <p>Re/5筋/1.94516</p> <p>Ac/突/2.15669</p> <p>Ac/使/2.35015</p> <p>Tu/先手/2.71036</p> <p>St/ゴキゲン中飛車/3.19339</p> <p>Ca/から/3.78502</p> <p>Ca/矢倉/3.81759</p> <p>Ac/指/3.86289</p> <p>Ac/位を取らせ/4.28983</p>	<p>Ac/警戒/4.5293</p> <p>Pa/道/4.63135</p> <p>Ot/小倉/4.66681</p> <p>Ee/濃厚/4.82272</p> <p>Ot/上村/4.84702</p> <p>Re/6筋/4.94942</p> <p>Ot/森下/4.98288</p> <p>Ee/得意/5.02482</p> <p>St/向飛車/5.06046</p> <p>PN/山崎隆之七段/5.103</p>
--	--

Fill all the **skeletons** and choose the most suitable one

**中飛車になった。**

+5.5492 = +4.64 + +2.50 + -1.59

**83 1 Stになった。 St:中飛車(+2.50)**

中飛車となった。

+5.5810 = +4.62 + +2.50 + -1.54

# Sentence Generation (detailed log)

Fill all the skeletons and choose the most suitable one

Generated Sentence: s

Total score =  $-\log P(s) + \sum -\log P(w_i) / n + -\log \text{Freq}(t)$

Skeleton: t Characteristic word list: w1, w2, ...

中飛車になった。

+5.5492 = +4.64 + +2.50 + -1.59

83 1 St になった。 St:中飛車(+2.50)

中飛車となった。

+5.5810 = +4.62 + +2.50 + -1.54

72 1 St となった。 St:中飛車(+2.50)

戦型は中飛車だった。

+6.6693 = +4.17 + +2.50 + -0.00

1 1 戦型は St だった。 St:中飛車(+2.50)

ふわりと角を突いた。

+7.0780 = +2.28 + +5.50 + -0.70

7 1 ふわりと Pi を Ac いた。 Pi:角(+8.01), Ac:突(+2.98)

久保は中飛車。

+7.0891 = +4.59 + +2.50 + -0.00

1 1 久保は St。 St:中飛車(+2.50)



# Result

Match name ( <u>Winner</u> )	% appropriate commentaries
2011年7月19日 第24期竜王戦決勝トーナメント <u>久保利明二冠</u> 対 橋本崇載八段	8/92 (8sentences /92moves)
2014年9月8日 第27期竜王戦挑戦者決定三番勝負 第3局 羽生善治名人 対 <u>糸谷哲郎六段</u>	9/95
2014年9月18日 第62期王座戦五番勝負 第2局 <u>羽生善治王座</u> 対 豊島将之七段	7/108
2014年8月5日～8月6日 第55期王位戦七番勝負 第3局 羽生善治王位 対 木村一基八段	12/178


# Result

Match name ( <u>Winner</u> )	% Good
2011年7月19日 <u>久保二冠</u> 対 橋本八段	8.7%
2014年9月8日 羽生名人 対 <u>糸谷六段</u>	9.5%
2014年9月18日 <u>羽生王座</u> 対 豊島七段	6.5%
2014年8月5,6日 羽生王位 対 木村八段	6.7%

- Appropriate: 7.6% (Very strict criteria)
  - Almost 99% grammatically correct
  - 30% correct if we allow time lag
- Want to raise it to about 50%

# Error Analysis

## 1. Characteristic word estimation

- 40%
- Use more information from 
- Refer to the game trees

## 2. Shogi NER

- 30%
- More data

## 3. Skeleton selection

- 20%
- Try more theoretical scoring functions

# In the near future

Navigation: ◀ ◁ ▷ ▶ ▶▶ ▶▶▶ 反転 Help Kifu for Flash 設定

▽羽生善治



▽羽生善治

9	8	7	6	5	4	3	2	1	
車	銀	龍	將	將		將	龍	車	一
	車					將	行	行	二
歩		歩	歩	歩	歩	歩	歩	歩	三
	歩								四
		歩				歩			五
歩	歩		歩	歩	歩	歩		歩	六
	角	金					飛		七
香	桂	銀		王	金	銀	桂	香	八
									九

▲豊島将之



▲豊島将之

\* = 開始局面 =  
 \* 1 ▲7六歩  
 \* 2 △8四歩  
 \* 3 ▲7八金  
 \* 4 △3二金  
 \* 5 ▲2六歩  
 \* 6 △8五歩  
 \* 7 ▲7七角  
 \* 8 △3四歩  
 \* 9 ▲8八銀  
 \* 10 △7七角成  
 \* 11 ▲同 銀

棋譜保存

自動更新しない



says...

Pi:角 Ai:換わりになった。

駒音2 再現速度普通 文字拡大 最終駒 2字駒2

終局	消費時間	5時間00分	残り時間
▲豊島将之	4時間55分	<div style="width: 98%;"></div>	0時間05分
△羽生善治	4時間19分	<div style="width: 72%;"></div>	0時間41分

# Reference

- **対数線形言語モデルを用いた将棋解説文の自動生成**
  - 亀甲 博貴, 三輪 誠, 鶴岡 慶雅, 森 信介, 近山 隆
  - 情報処理学会論文誌, Vol.55, No11, 2014.
    - Characteristic word extraction and commentary generation by a word 3-gram model
    - Limited to the opening
- **FlowGraph2Text: Automatic Sentence Skeleton Compilation for Procedural Text Generation**
  - Shinsuke Mori, Hirokuni Maeta, Tetsuro Sasada, Koichiro Yoshino, Atsushi Hashimoto, Takuya Funatomi, Yoko Yamakata
  - In proc. of INLG, 2014.
    - Automatic skeleton compilation by using NER
    - Sentence generation from a recipe flow graph
- **将棋解説文のグラウンディングのための指し手表現と局面状態の対応付け**
  - 亀甲 博貴, 三輪 誠, 鶴岡 慶雅, 森 信介
  - 情報処理学会ゲームプログラミングワークショップ, 2014. (優秀論文賞)
    - Commentary tree
    - Sentence generation based on commentary trees

# Reference

- **KyTea**

- <http://www.phontron.com/kytea/index-ja.html>
- [Neubig+, LREC, 2010 ], [Neubig+, ACL, 2011], [Mori+, InterSpeech, 2012]
- [森+, 情処論, 2011], [森+, 自然言語処理, 2011]

- **PWNER**

- <http://plata.ar.media.kyoto-u.ac.jp/mori/research/topics/NER/>
- [Mori+, CwC, 2012]