

将棋解説の自動生成のための 局面からの特徴語生成

亀甲 博貴¹ 浦 晃¹ 三輪 誠²

鶴岡 慶雅¹ 森 信介³ 近山 隆¹

¹ 東京大学

² マンチェスター大学

³ 京都大学

2013. 11. 09

背景

- 将棋ソフトが**プロ棋士に比肩**する強さ
 - 将棋ソフトに手を読ませながら観戦

+

- 注目の対局はプロ棋士が解説
 - 解説を聞いて**局面の理解が深まる**


||


- 将棋ソフトが解説できたら楽しい！！


将棋の解説生成例


[金子, 2012]


- GPS将棋の読み筋・評価値をTwitterで提供 (@gpsshogi)
- 詰み周りで多様な解説
- 指し手の狙いを解説 (Null Move探索)


 **gpsshogi** @gpsshogi 3月24日
[(145) ▲ 5八歩] 771 △ 1七飛成▲ 6五角△ 2二金▲ 3六桂△ 3三銀引▲ 4四歩△ 2三玉▲ 1三歩成△ 同龍▲ 4三歩成△ 4二銀 (28s)
開く


 **gpsshogi** @gpsshogi 3月24日
[(145) ▲ 5八歩]
* 王手を回避して, 後手玉は▲ 1三金(以下)の詰めろ.
開く

 **gpsshogi** @gpsshogi 3月24日
[(144) △ 1八飛] 840 ▲ 5八歩△ 1七飛成▲ 6五角△ 3三銀引▲ 4三金△ 同銀▲ 同角成(詰めろ)△ 2二金打▲ 3二銀(詰めろ)△ 1四龍▲ 2六桂 (45s) / 571 ▲ 6八桂△ 1七飛成▲ 6五角△ 3三飛
開く

 **gpsshogi** @gpsshogi 3月24日
[(144) △ 1八飛]
* 応手によっては詰がある.
開く

 **gpsshogi** @gpsshogi 3月24日
[(143) ▲ 1七香] 687 △ 1八飛▲ 5八歩(詰めろ)△ 1七飛成▲ 6五角△ 2二金▲ 4三金△ 同銀▲ 同角成(詰めろ)△ 3三金打▲ 4四馬△ 同金 (28s)
開く

 **gpsshogi** @gpsshogi 3月24日
[(143) ▲ 1七香]
* 後手玉は▲ 1三歩成(以下)の詰めろ.
開く

 **gpsshogi** @gpsshogi 3月24日
[(142) △ 1七金] 613 ▲ 同香(詰めろ)△ 1八飛▲ 9七玉(詰めろ)△ 1七飛成▲ 6五角△ 9五歩▲ 3二金(詰めろ)△ 9六歩▲ 8八玉△ 3二銀▲ 同角成 (300s)
開く

@gpsshogiの解説文の特徴

- GPS将棋 (= トップクラスの将棋ソフト) の探索
 - 次の一手はそれなりに当たる
 - 探索結果は概ね有用
- **自然言語による情報**が少ない
 - 用意された**テンプレート**に結果を埋めて出力
 - 情報はほぼ探索結果のみ
 - -> (コンピュータ)将棋に疎い人には理解が難しい

目的

- 局面を**自然言語**で解説できないか？

The diagram shows a Go board with columns numbered 1-9 from right to left and rows numbered 1-9 from top to bottom. The board contains several pieces: King (王) at 4-1, Silver (銀) at 6-2, Gold (金) at 5-2, Knight (桂) at 1-1 and 9-1, Pawn (歩) at 1-3, 2-3, 3-3, 4-3, 5-3, 6-3, 7-3, 8-3, 9-3, 6-4, 7-4, 8-4, 9-4, 6-5, 7-5, 8-5, 9-5, 6-6, 7-6, 8-6, 9-6, 6-7, 7-7, 8-7, 9-7, 6-8, 7-8, 8-8, 9-8, 6-9, 7-9, 8-9, 9-9. A black stone is at 3-2. A speech bubble on the left points to the board and contains the text "矢倉模様". A speech bubble on the right contains the text "次は▲3七銀として4六銀3七桂型を目指す".

矢倉模様

次は▲3七銀として
4六銀3七桂型を目指す

- 局面を解説する文章の生成モデルを作成する

発表の流れ

- 背景・目的
- **提案手法**
- 評価
- まとめ

提案手法

- 人間による解説を元に機械学習
- 学習に用いる**解説文の選択**
- 解説文に出てきそうな**単語を予測**
 - 3層パーセプトロンによる単語の予測
- **直前の単語**と組み合わせで1単語予測->文生成
 - 直前2語を用いて生成

教師とする解説文

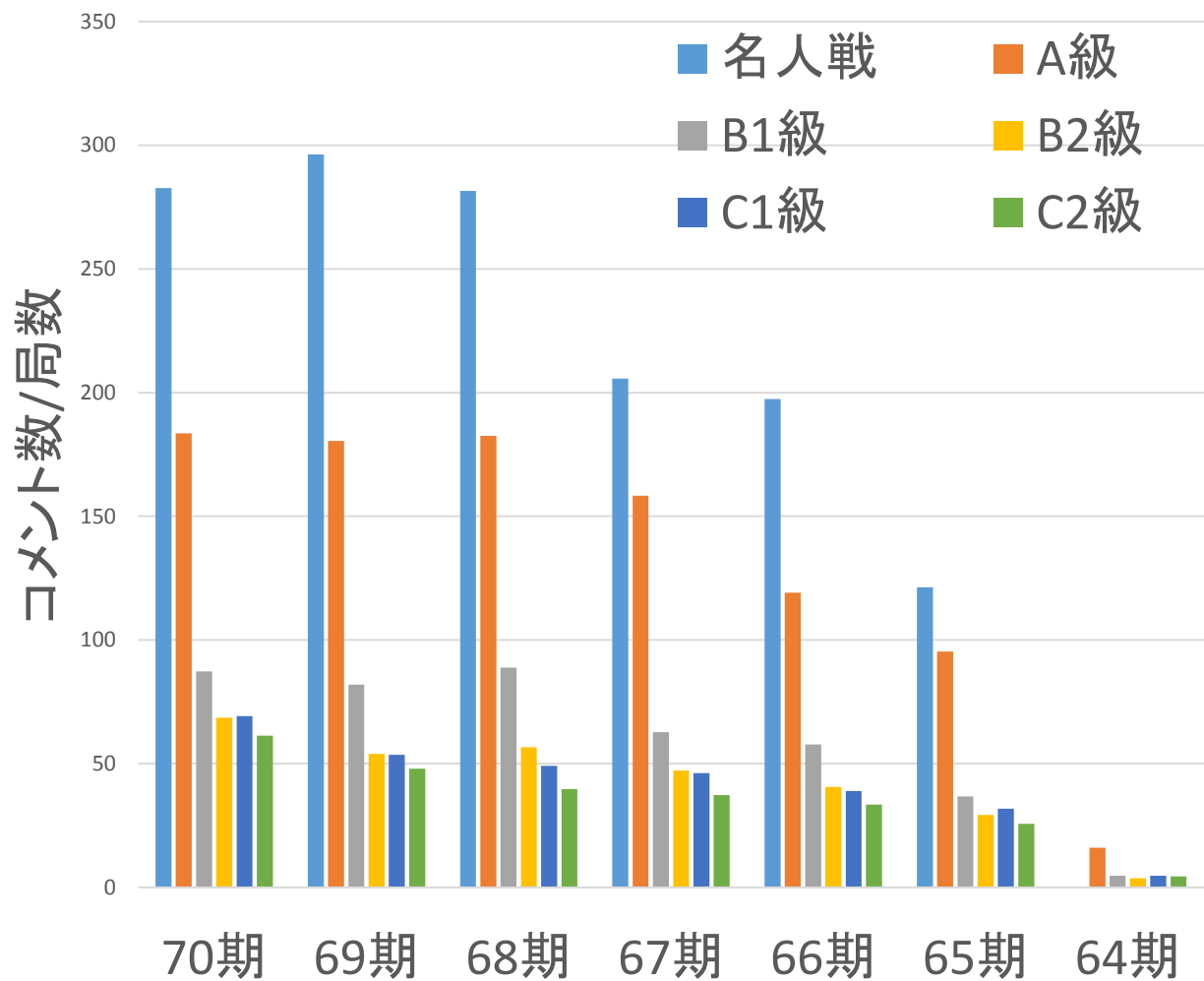
- 順位戦＋名人戦の棋譜速報
 - 名人戦棋譜速報<http://www.meijinsen.jp/>
 - 名人戦と順位戦の解説付き棋譜
 - KIF形式
- そのまま使うには問題点が多い
 - 学習の対象外とすべき解説コメント
 - クラス毎のコメント量のムラ

対象外とするコメントの例

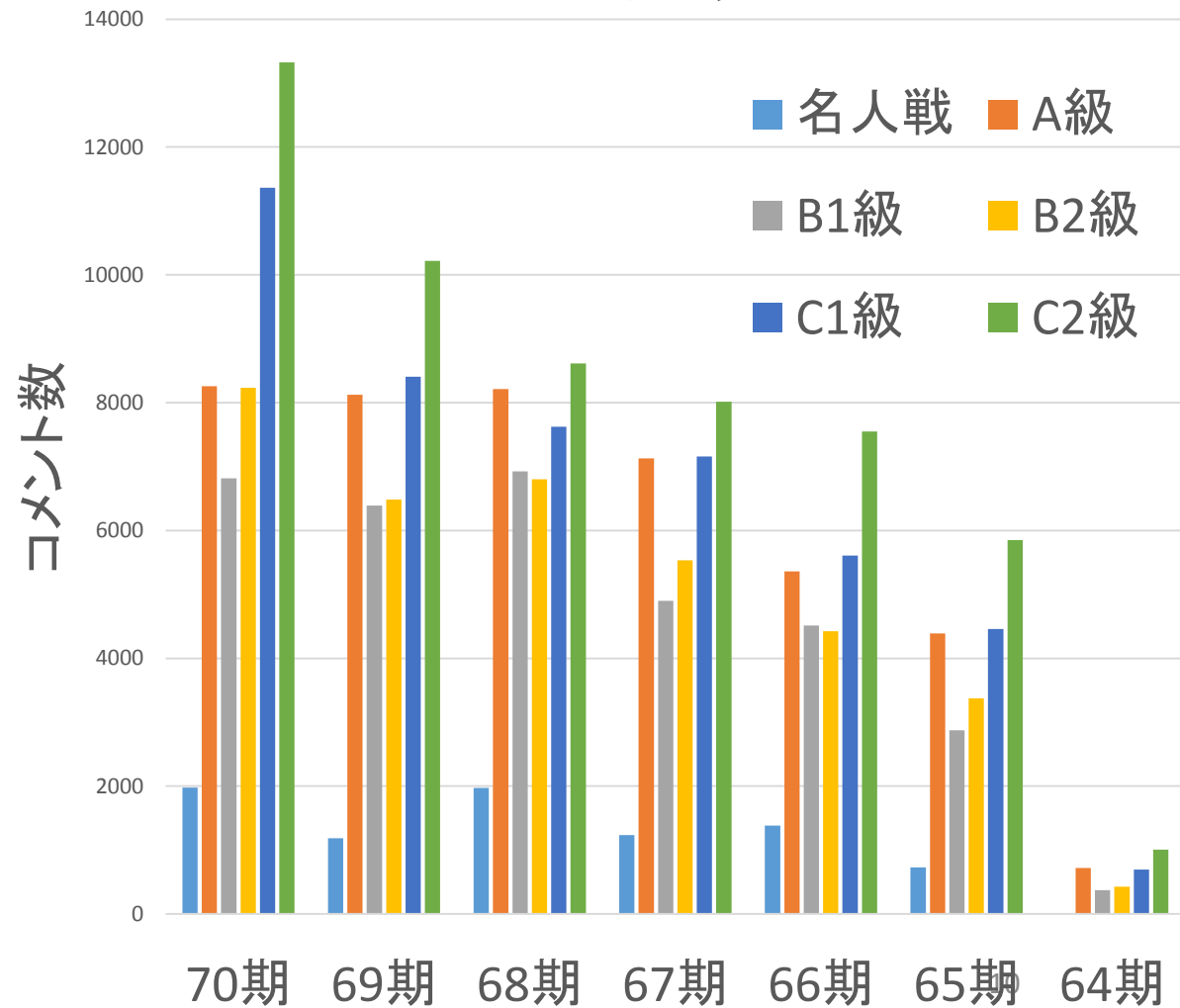
- 局面の解説に直接は関係しないコメント
ここまでの消費時間は～～
- 将棋と関係ないコメント
昼食(おやつ・夕食)の注文は～～
対局場所は～～

コメント量のムラ

1局あたりのコメント数



コメント総数

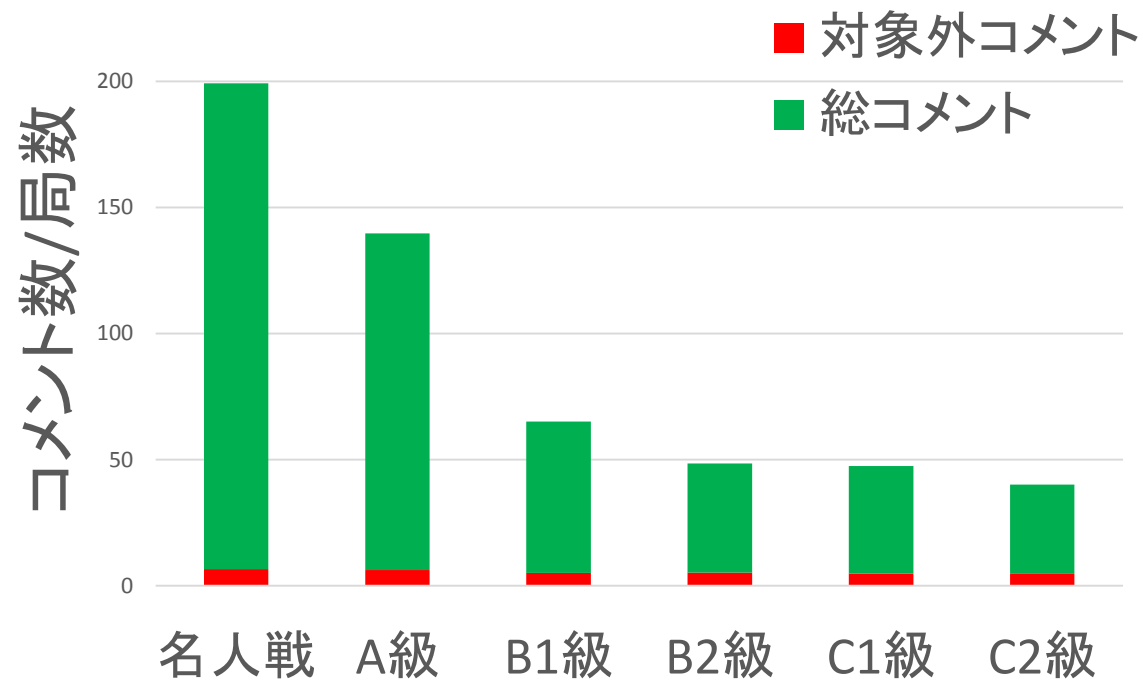


下位クラスの対象外コメント量

- パターンマッチで取れる対象外コメントだけでも多い

ここまでの消費時間は
残り時間は
昼食の注文は
夕食の注文は

- 相対的に**対象外コメントの占める割合が高い**
 - 簡単に取り除ける4種類だけでC2級は**1割超**



対象外コメントへの対策

- 下位クラスの棋譜を全て無視するだけではダメ
 - 上位クラス中の対象外コメントが残ってしまう
 - 下位クラス中の対象コメントを捨ててしまう



- コメントが**有用であるかどうかの判別**を図る

戦型に言及する文の判別

- 含まれる単語を元に**戦型に言及する文**かそうでないかを判別

- *横歩取りになりそうだ* -> 戦型に言及する文
- *金にヒモを付けて受けた* -> 戦型に言及しない文

「矢倉」「穴熊」「中飛車」->戦型に言及？

判別モデル

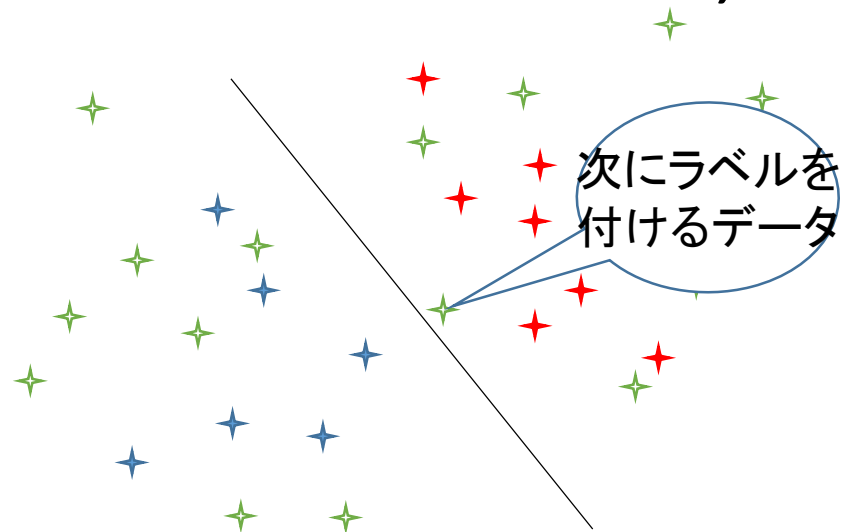
- 平均化パーセプトロン
- 特徴: 各単語が文中にいくつ現れたか
 - 形態素解析器「**KyTea**」により単語に切り分け
 - 指し手表現を「MOVE」に置換して統一
 - ▲5八金右, △8八角成, 4六, etc.

形態素解析器KyTea [Neubigら, 2011]

- 日本語文を単語(形態素)に分ける
 - *先手陣の矢倉が完成した*
-> *先手 / 陣 / の / 矢倉 / が / 完成 / し / た*
- 「MeCab」に比べ**分野適応**に優れる
 - 将棋に関する300文を用いてモデル学習
 - 学習前: *中飛 / 車 / 対 / 四 / 間 / 飛車*
 - 学習後: *中飛車 / 対 / 四間 / 飛車*

学習データ

- 学習データを**能動学習**で選ぶ
 - 最初に50文ラベル付け->学習
 - クラス間スコアの小さいものにラベル付け->学習
 - 最初の50文を合わせて1,000文のラベル付け



3層パーセプトロンによる単語予測

- 解説文中に現れるであろう単語の予測
 - 入力: 局面の特徴(「激指」の評価関数の特徴)
 - 出力: 解説文中の単語の有無

単語の予測ベクトル

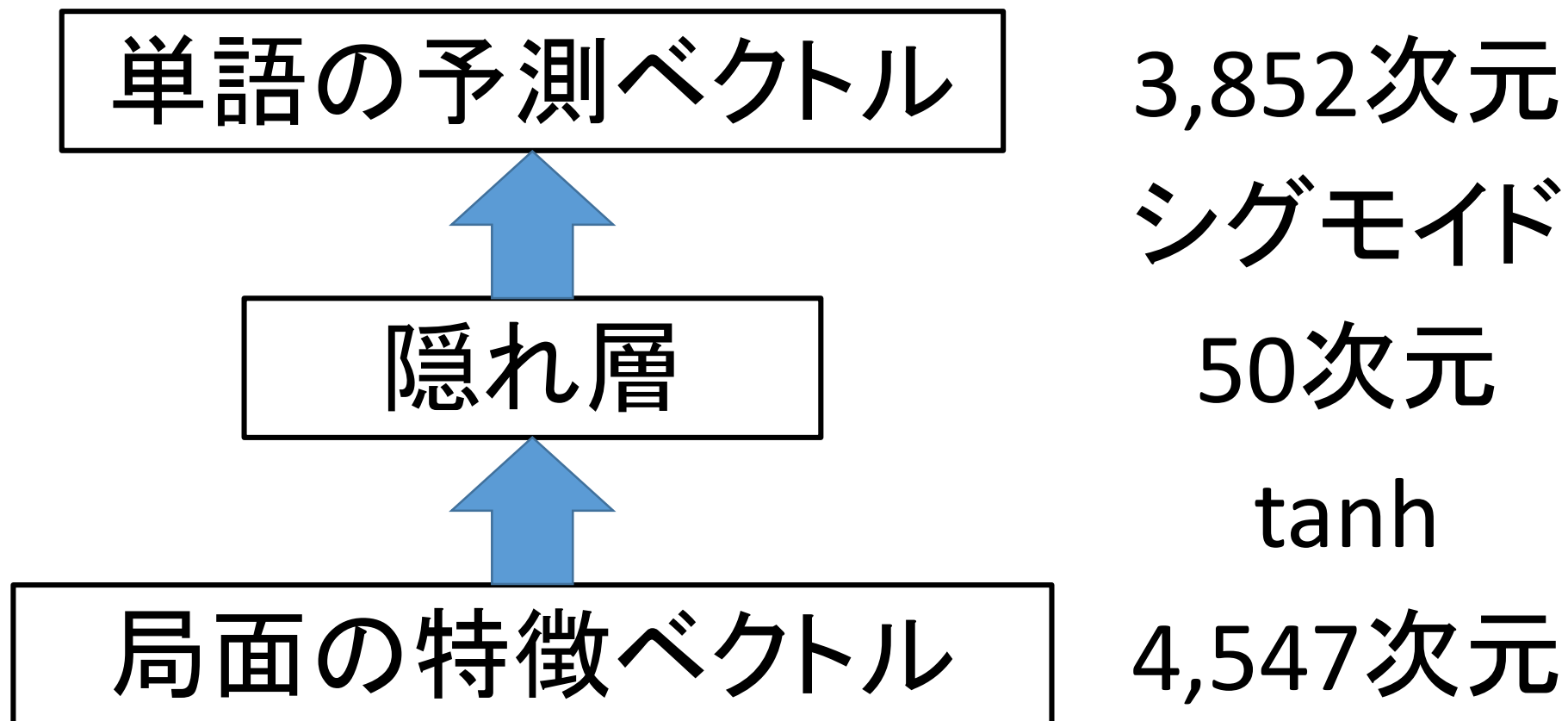
隠れ層

局面の特徴ベクトル

| | | | |
|----|-------|-----------|--------------|
| 採用 | 0.775 | 型 | 0.219 |
| 矢倉 | 0.219 | 横歩 | 0.822 |
| 取り | 0.789 | 戦法 | 0.577 |



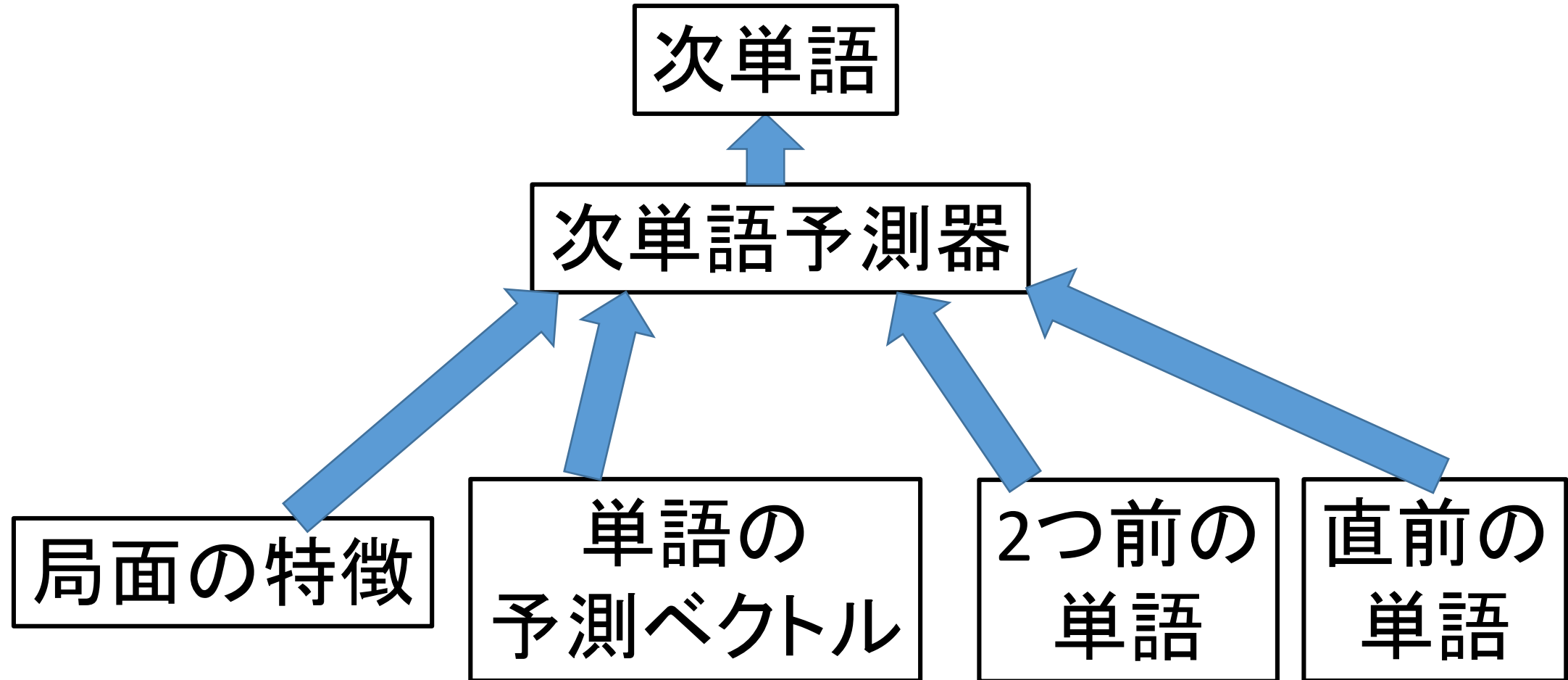
3層パーセプトロンの設計



文の生成

- 次の1語予測を繰り返し、解説文を生成する
 - 単語の予測ベクトルと直前2単語を元に次の単語を予測
- 単語予測をパーセプトロンにより行う

パーセプトロンによる文生成



パーセプトロンによる文生成

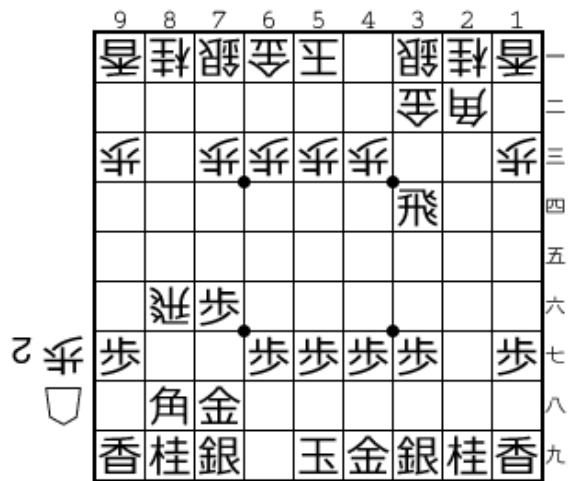
横歩

横歩

次単語予測器

(文頭)

(文頭)



| | | | |
|----|-------|----|-------|
| 採用 | 0.775 | 型 | 0.219 |
| 矢倉 | 0.219 | 横歩 | 0.822 |
| 取り | 0.789 | 戦法 | 0.577 |

パーセプトロンによる文生成

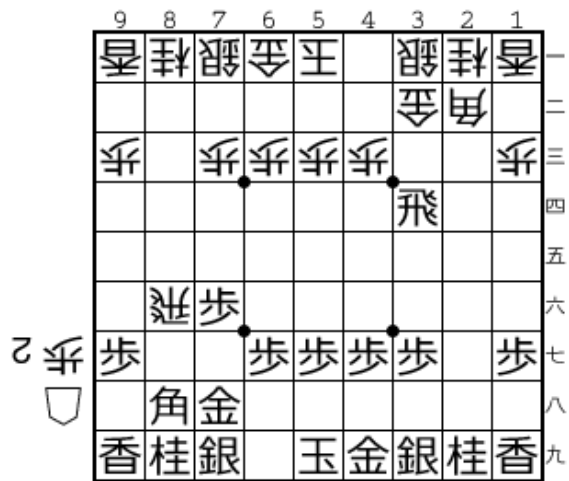
横歩 取り

取り

次単語予測器

(文頭)

横歩



| | | | |
|----|-------|----|-------|
| 採用 | 0.775 | 型 | 0.219 |
| 矢倉 | 0.219 | 横歩 | 0.822 |
| 取り | 0.789 | 戦法 | 0.577 |

パーセプトロンによる文生成

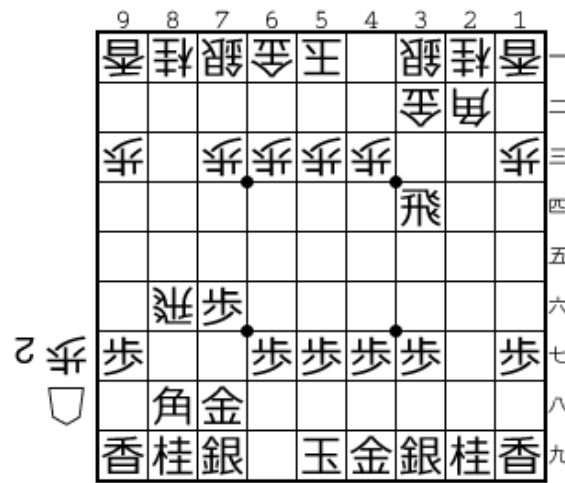
横歩 取り

(文末)

次単語予測器

横歩

取り



| | | | |
|----|-------|----|-------|
| 採用 | 0.775 | 型 | 0.219 |
| 矢倉 | 0.219 | 横歩 | 0.822 |
| 取り | 0.789 | 戦法 | 0.577 |

発表の流れ

- 背景・目的
- 提案手法
- **評価**
- まとめ

判別結果

- 約35万文中10,540文を「戦型に言及する」と判別
 - このうち200文について確認
 - > 162文(81%)は正しく戦型に言及する文
 - ランダムな500文中に戦型に言及する文は20文(4%)
 - 判別器が「戦型に言及する」としたものはある程度信頼できる
- 以降はこの10,540文を対象とする

単語の予測結果（一部）

| 単語 | Precision | Recall | F-score |
|-----|-----------|--------|---------|
| 穴熊 | 0.45 | 0.77 | 0.57 |
| 矢倉 | 0.29 | 0.49 | 0.36 |
| 換わり | 0.62 | 0.68 | 0.64 |
| 居飛車 | 0.09 | 0.01 | 0.02 |

- 出現単語の予測問題と考えると**低精度**

予測結果の考察

- 局面がそうならなかった際のコメント
 - 穴熊にするのがセオリーだが、ここでは左美濃
 - 穴熊を断念
- 局面とは関係ない知識からのコメント
 - 千日手局は穴熊だったが、指し直し局は矢倉模様
 - (棋士名)はあまり穴熊をやらない
- 単語だけで文の意味は表現できない

生成された解説文の例



予測ベクトル
 形 0.472
 システム 0.981

- *型型 MOVE 矢倉の駆け引きがある。*

生成された解説文の例



予測ベクトル

| | | | |
|----|---|-----|---|
| と | 1 | 損 | 1 |
| 角 | 1 | 換わり | 1 |
| 一手 | 1 | 閉ま | 1 |

• *MOVE* を目指す。

生成された解説文の例



予測ベクトル
 美濃 0.820

- 型から型歩型を MOVE 型から左美濃囲いを目指す可能性がある。

発表の流れ

- 背景・目的
- 提案手法
- 評価
- **まとめ**

まとめ

- 簡単な判別器で学習に用いる解説文を選択できた
- 解説文中のキーワードを当てることはできなかった
- 得られたベクトルを用いて解説文らしいものが一部の局面に対して得られた

今後の課題

- より**表現力豊かな**モデルや特徴セットの提案
 - 指し手の履歴, 読み筋, etc...
- **戦型に限らない**解説文生成
 - 読み筋・評価
- 解説文間の関係の考慮