

MeCab を用いた漢文形態素解析器のための品詞の問題

守岡 知彦

1 品詞（素性）階層

MeCab では4階層の品詞（素性）階層を使って bigram を用いたパラメータ推定が行えるが、ここでどのような品詞（素性）階層を用いるかが問題となる。

そこで、品詞（素性）階層を評価するために、それぞれの体系に基づき、コーパスと辞書を記述し、各コーパスを互いに学習用コーパスとテストデータとして認識実験を行い、全組合せの認識精度を計算した。

この実験には、M (69 語)、K (68 語)、R (320 語)、W (248 語)、J (363 語) というコーパスを用いた。なお、M は雑多な文例、K は典型的な構文例、R は三国志呉書列伝よりの抜粋、W と J は旧唐書東夷伝の項目であり、地理的な記述が多く、特に J には外来語が多い。

1.1 大品詞

学習データ \ 入力	M	K	R	W	J
M (大品詞/品詞)	100	90/82	90/88	92/89	94/91
M (大品詞なし)	100	87	90	91	93
K (大品詞/品詞)	91/90	100	92/89	94/90	95/93
K (大品詞なし)	90	100	95	88	91
R (大品詞/品詞)	100/99	90/85	100	94/91	97/94
R (大品詞なし)	94	87	100	90	92
W (大品詞/品詞)	100/100	93/90	92/90	100	98/97
W (大品詞なし)	96	90	94	100	93
J (大品詞/品詞)	97/94	94/88	93/91	97/95	100
J (大品詞なし)	91	88	93	94	100

表 1: 大品詞がある場合とない場合の大品詞と品詞の F 値（意味素性あり）

大品詞を設けない場合の品詞の認識精度は、意味素性（サブ品詞）を付ける場合、概ね、大品詞を設けた場合の大品詞の認識精度と品詞の認識精度の中間ぐらいの値を取るといえるが、学習用コーパスと入力データの文体の差が大きい場合、大品詞を設けた方がロバスト性が高まると考えられる。

一方、意味素性を付けない場合、大品詞を設けない場合の品詞の認識精度は、大品詞を設けた場合に比べて下がりやすいといえる。このことから、大品詞を設けた場合のロバスト性の向上効果がかうかがえる。

学習データ \ 入力	M	K	R	W	J
M (大品詞/品詞)	100	89/84	87/84	90/83	88/85
M (大品詞なし)	100	83	85	83	85
K (大品詞/品詞)	85/81	100	84/80	88/86	84/82
K (大品詞なし)	81	100	80	84	77
R (大品詞/品詞)	96/92	84/76	99/99	91/85	89/84
R (大品詞なし)	87	76	100	82	82
W (大品詞/品詞)	94/93	87/79	87/84	99/99	94/93
W (大品詞なし)	91	78	84	100	93
J (大品詞/品詞)	93/90	87/79	93/90	92/88	99/99
J (大品詞なし)	91	76	87	87	100

表 2: 大品詞がある場合とない場合の大品詞と品詞の F 値 (意味素性なし)

よって、意味素性を付与したコーパス・辞書を用意でき、適切な学習用コーパスを用意できる場合においては、大品詞は必ずしも必要ないといえるが、開発途上の段階では大品詞を設けた方が安全であるかもしれない。

1.2 意味素性

学習データ \ 入力	M	K	R	W	J
M	100	89/84	87/84	90/83	88/85
K	85/81	100	84/80	88/86	84/82
R	96/92	84/76	99/99	91/85	89/84
W	94/93	87/79	87/84	99/99	94/93
J	93/90	87/79	93/90	92/88	99/99

表 3: 名詞・動詞ともに意味素性がない場合の大品詞と品詞の F 値

学習データ \ 入力	M	K	R	W	J
M	100	88/79	87/83	88/83	88/83
K	88/82	100	87/83	90/86	89/87
R	95/89	90/82	99/99	93/89	93/89
W	97/97	93/87	90/86	100	95/92
J	94/86	96/91	94/90	95/90	100

表 4: 動詞の意味素性がない場合の大品詞と品詞の F 値

コーパスに意味素性を付与した方が付与しない場合よりも認識精度は向上する。また、部分的に意味素性を付与する場合、動詞に意味素性を付与し名詞の意味素性を省略する方が、名詞に意味素

学習データ \ 入力	M	K	R	W	J
M	100	90/84	92/90	92/88	93/91
K	88/86	100	88/86	90/88	93/91
R	96/95	87/81	100	92/87	95/90
W	96/94	87/79	92/89	100	97/96
J	94/94	88/84	96/93	93/92	100

表 5: 名詞の意味素性がない場合の大品詞と品詞の F 値

学習データ \ 入力	M	K	R	W	J
M	100	90/82	90/88	92/89	94/91
K	91/90	100	92/89	94/90	95/93
R	100/99	90/85	100	94/91	97/94
W	100/100	93/90	92/90	100	98/97
J	97/94	94/88	93/91	97/95	100

表 6: 名詞・動詞ともに意味素性がある場合の大品詞と品詞の F 値

性を付与し動詞の意味素性を省略するよりも効果的であるといえる。

2 動詞の品詞階層

2.1 連結、対象、存在、否定等

連結 (copula) 動詞や存在、否定等を表す動詞である。表 7 にそのリストを示す。

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	連結	存在	爲
v	動詞	対象	構成、目的	爲
v	動詞	存在	(対象が)	有
v	動詞	存在	(場所に)	在
v	動詞	否定	存在、禁止	無

表 7: 連結、存在、否定等

論点

- 「爲」はひとつのカテゴリーにまとめるべき?

2.2 性質、状態等

性質や状態状態等を表す動詞である。表 8 にそのリストを示す。

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	性質	形容 (比較対象に)	如、猶、類
v	動詞	状態	滞在 (場所に)	留
v	動詞	状態	滞在 (人を)	留
v	動詞	状態	日常、滞在、居住	居、留
v	動詞	状態	位置、水上 (場所に)	勒、泛
v	動詞	状態	付着、依存 (~に)	附、依
v	動詞	状態	付着	

表 8: 性質、状態等

論点

- もしかして形容詞にすべき?

2.3 事象等

事象等を表す動詞である。表 9 にそのリストを示す。

到達 (achievement)・瞬間動詞 + 無意志動詞?

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	事象	死去	死
v	動詞	事象	遭遇	會

表 9: 性質、状態等

2.4 習慣・嗜好等

習慣・嗜好等を表す動詞である。表 10 にそのリストを示す。

状態 (state) 動詞 + 意志動詞?

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	習慣	嗜好 (~を)	好
v	動詞	習慣	表敬 (~を)	敬
v	動詞	習慣	交流	通

表 10: 習慣・嗜好等

2.5 関係

関係を表す動詞である。表 11 にそのリストを示す。

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	関係	付着、依存	依、縁、随

表 11: 関係・理由等

2.6 授与・教示・質問等

授与・教示・質問等を表す動詞である。目的語を 2 つ取り得る。表 12 にそのリストを示す。

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	授与	(相手に物を)	遺
v	動詞	設置	授与 (相手に)	益
v	動詞	行為	授与 (相手に)	餽
v	動詞	行為	授与 (物を)	貢、獻、賜
v	動詞	教示	(知識を)	教

表 12: 授与・教示・質問等

2.7 使役

いわゆる使役動詞である。目的語を 2 つ取る。表 13 にそのリストを示す。

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	使役	命令、派遣	令、使、遣

表 13: 使役

2.8 意志・態度等

意志や態度等を表す動詞である。表 14 にそのリストを示す。

活動 (action) ・継続動詞 + 意志動詞?

論点

- 「保証」、「守備」は「習慣」・「嗜好」等と同じカテゴリーにすべき?

2.9 感情、認識、計画、評価等

感情、認識、計画、評価等の認知活動を表す動詞である。モダリティに相当? 表 15 にそのリストを示す。

活動 (action) ・継続動詞 + 意志動詞?

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	意志	了承、同意 従属	承、從、附
v	動詞	態度	尊大、侮蔑、敵対、慰撫、赦免	貴、輕、撫、敵、舍
v	動詞	態度	保証、守備	保、守
v	動詞	謙敬	表敬	
v	動詞	謙敬	謙讓	

表 14: 意志・態度等

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	思惟	感情、計画、評価、認識	(頭の中で思ったこと)

表 15: 感情、計画、評価、認識等

2.10 身体動作、直接的行為等

身体動作や直接的な行為等を表す動詞である。表 14 にそのリストを示す。
(活動動詞 or 到達・瞬間動詞) + 意志動詞?

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	動作	身体、道具	(直接的な動作)

表 16: 身体動作、直接的行為等

2.11 達成等

達成等を表す動詞である。表 17 にそのリストを示す。

2.12 移動等

移動を表す動詞である。表 18 にそのリストを示す。

2.13 設置等

設置を表す動詞である。表 19 にそのリストを示す。
(瞬間動詞 or 達成動詞) + 意志動詞?

2.14 情動的行動

情動的行動を表す動詞である。表 20 にそのリストを示す。
(瞬間動詞 or 達成動詞) + 意志動詞?

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	達成	完遂	成
v	動詞	移動	到達 (場所に)	至

表 17: 達成等

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	移動	出発、前進、往路、渡来	出、前、行、往、渡
v	動詞	移動	前進 (行程を)	行
v	動詞	移動	帰還、退去	歸、還、去
v	動詞	移動	帰還 (場所に)	歸、還
v	動詞	移動	逃走、離脱	走、逃、脱
v	動詞	移動	追尾	追
v	動詞	移動	来訪	來
v	動詞	移動	来訪 (場所に)	入、來

表 18: 移動等

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	設置	開設 (dst を)	設、置
v	動詞	設置	結合 (src を)	束
v	動詞	設置	分割	分
v	動詞	設置	分割 (src を)	分
v	動詞	設置	増加 (内容を)	増
v	動詞	設置	被覆 (src を)	蔽
v	動詞	設置	置換 (src に)	代
v	動詞	設置	人 (場所・地位に)	置
v	動詞	行為	改変	改
v	動詞	行為	改変 (内容を)	改
v	動詞	行為	結合 (src を)	併
v	動詞	設置	上方	戴
v	動詞	衣裳	装着	佩、服
v	動詞	名称	命名	題
v	動詞	行為	製作	作

表 19: 設置等

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	情報	明示 (内容を)	明
v	動詞	情報	欺瞞	欺
v	動詞	行為	言明 (言葉を)	曰、云
v	動詞	行為	言明 (内容を)	宣
v	動詞	行為	言明 (内容を)	表
v	動詞	行為	言明 (文書を)	奉
v	動詞	行為	言明 (指令を)	敕
v	動詞	行為	言明 (相手に)	謂
v	動詞	行為	通告 (相手に)	告
v	動詞	文書	記述 (文章を)	屬
v	動詞	行為	読解 (書物を)	讀

表 20: 情報の行動

2.15 对人的行動

对人的行動を表す動詞である。表 21 にそのリストを示す。

(瞬間動詞 or 達成動詞) + 意志動詞?

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	行為	推薦 (人を)	薦
v	動詞	行為	拔擢 (人を)	擢
v	動詞	行為	侮蔑 (人を)	辱
v	動詞	行為	殺害 (人を)	殺
v	動詞	行為	捕縛 (人を)	禽
v	動詞	行為	叱責 (人を)	罰
v	動詞	行為	軍 (人・国・勢力を)	討、、征、攻
v	動詞	達成	勝利 (対象を)	破
v	動詞	行為	軍 (軍を)	舉
v	動詞	行為	叱責	呵、叱
v	動詞	行為	召喚	召
v	動詞	行為	付着	從
v	動詞	行為	人	引

表 21: 对人的行動

2.16 行為

大品詞	品詞	意味素性 1	意味素性 2	例
v	動詞	役職	就任、着任	仕、就
v	動詞	役職	経歴 (役職を)	
v	動詞	役職	着任 (役職に)	拜、為
v	動詞	行為	役職 (任務を)	務
v	動詞	行為	監察 (対象を)	檢察
v	動詞	行為	外交	朝貢
v	動詞	行為	訴訟	
v	動詞	行為	探索 (対象を)	求、探
v	動詞	行為	入手 (物を)	得、梟
v	動詞	行為	購入 (物を)	市
, v	動詞	行為	所持 (~を)	持
v	動詞	行為	対立 (争点を)	争
v	動詞	行為	委任 (場所・仕事を)	委
v	動詞	行為	派遣 (~に)	出
v	動詞	行為	農業 (田畑を)	耕
v	動詞	行為	実行 (行為を)	為
v	動詞	行為	実行	致、為
v	動詞	行為	対応	對
v	動詞	行為	救助	救
v	動詞	行為	解放	放
v	動詞	行為	選択	撰
v	動詞	行為	学習	
v	動詞	行為	提示	
v	動詞	行為	自首	

表 22: その他動詞