

1. GB/T 13715-92
2. 973
3. The Segmentation Guidelines for the Penn Chinese Treebank 3.0
University of Pennsylvania, 2000

一、分词总则

#

1		8
2	()	
3		
4		

二、词类标记集

The Part-Of-Speech Tagging Guidelines for the Penn Chinese
Treebank

	NT	
	NR	
	NN	
	VC	
	VE	
	VV	
	VA	
	JJ	
	CD	
	OD	
	DT	
	M	/
	AD	/
	PN	/
	P	/
	LC	/
	CC	
	CS	
	DEC	/ (de5)
	DEG	/ (de5)
	DER	(de5)
	DEV	(de5)
	SP	
	AS	
	ETC	
	MSP	
	IJ	
	ON	
	LB	(bei4)
	SB	(bei4)
	BA	ba3

	FW	
	PU	

三、特殊词性分词规范

1. 专业术语

#JJ #NN

#JJ #NN

#VV

#VV

#NN

#NN

#NN

#NN

#NN

2. 时间名词

[1]

#NT

#NT

[2]

1997 #NT 3 #NT 19 #NT

#NT 2 #NT 18 #NT 35 #NT

[3]

#NT

#NT

#NT

#NT

[4]

08:35:28#NT

2003-03-29#NT

3. 地名

表示地理区域的名称

[1]

#NR

#NR

#NR

[2]

#NR

#NR

4. 团体机构名

包括团体、机构、组织的专有名称

[1]

#NR

[2]

#NR

#NN

#NN

5. 量词

表示事物的单位或动词作的量。包括，常和名词连用的名量词，有个体量词（位、辆、张），度量词（克、千米），复合量词（人次、架次、吨公里），不定量词（点、些）；以及动词量词（次、回、趟）和时量词（天、小时）等

[1]

3#CD #M #NN

36.2#CD #M

[2]

/ #M

Mmol/l #M

6. 定位词

[1]

#JJ #NN

#JJ #NN

#LC #VV

#LC #VV

[3]

#LC

#LC

JJ

7. 标点符号

#NN :#PU 1#CD +#PU

T#NN :#PU 36.5#CD

8. 界限词素

#JJ

#NT

四、组合词性分词规范

1. [名词+名词]

1.1 并列关系

[1]

#NN

#NN

[2]

#NN #NN #NN

#NN #NN #NN

1.2 定中关系

[1] 2 4

1

#NN

#NN

#NN

#NN

#NN

#NN

#NN

[2]

2

#NN #NN

#NN #NN

#NN

2. [名词+动词]

[1]

#VA
#NN { #AD} #VA

[2]

#NN
#NN
#NN

3. [名词+形容词]

#NN #VA
#NN #VA

4. [名词+定位词]

[1]

[2]

[3]

#NN
#NT

5. [动词+名词]

5.1 动宾关系

[1] 2 4

1

#NN
#NN
#NN
#NN
#NN

[2]

2

#NN #NN

5.2 述宾关系

+

#VV

#VV

#VV

#VV

#VV

#VV #NN

#VV #NN

#VV #NN

#VV #NN

#VV #NN

#VV #NN

6. [动词+动词]

6.1 并列关系

#VV

#VV

#VV #VV #VV

6.2 状中关系

[1] 2

#VV

[2]

#VV #VV

#VV #VE

#VV #VV

[3]

#VV #AD #VV

#VV #DER #VV

[4]

#VV #VV

#VV #VV

#VV

#VV #VV

7. [动词+形容词]

#VV

#VV #VA

#VV #VA

8. [动词+介词]

#VV #P

#VV

9. [形容词+名词]

[1] [+]

#NN

#JJ #NN

#JJ #NN

#JJ #JJ #NN

#JJ #NN

[2] 4

#JJ #NN

#JJ #NN

[3] [+]

#JJ

#JJ

10. [形容词+动词]

#VV

#VV

#VV

11. [形容词+形容词]

#JJ #JJ #NN

#JJ #JJ #NN

#NN #NN

12. [数词+名词]

[1] [+]

#NN

#OD #NN

[2] [+]

#JJ

#NN

#DT #NN

[3] [+]

30 #CD

[4]

#CD #NN

#CD #NN

13. [数词+量词]

#AD 3.0#CD mm#M
#NN 120#CD /#PU 70#CD mmHg#M
#CD #M
#DT #M
#DT #M #NN

14. [副词+动词]

#AD #VV
#AD #VV
#AD #VV
#AD #VV
#AD #VV
#AD #VV
#AD #VE
#AD #VV
#AD #VE
#AD #VC

15. [副词+形容词]

#AD #VA
#AD #VA

16. [代词+名词]

#NN
#NN

17. [代词+定位词]

#NN

#NN

18. [否定词+形容词]

#VA

#VA

#VA

#VA

#VA

#VA

#VA

19. [多字词+前/后缀]

[1] +
#NN
#NN

[2]
#NN
#JJ
#JJ
#NN #VV #NN

五、通用分词规范

1

2

3

1

