

1. The Bracketing Guidelines for the Penn Chinese Treebank v 3.0
University of Pennsylvania, 2000
- 2.

一、句法树库加工总则

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、句法功能标记集

The Bracketing Guidelines for the Penn Chinese

Treebank

	ADJP	
	ADVP	
	CLP	
	CP	
	DNP	

	DP	
	DVP	
	FRAG	
	IP	
	LCP	
	LST	
	NP	
	PP	
	PRN	
	QP	
	UCP	
	VP	
	VCD	
	VCP	
	VNV	A-not-A A-one-A
	VPT	V-de-R, V-bu-R
	VRD	
	VSB	

四、短语标注详细说明

1 形容词短语 (ADJP)

JJ

(ADJP (ADVP (AD)))

(ADJP (JJ)))

(ADJP (JJ)

(JJ))

2 副词短语 (ADVP)

AD
(ADVP (AD))

(ADVP (AD)
(AD))

3 量词短语 (CLP)

M
(QP (CD 50)
(CLP (M /)))

4 限定词短语 (DP)

QP NP DP NP
(NP (DP (DT))
(NP (NN)))

(NP (DP (DT))
(NP (NN)))

5 “的”字名词短语 (DNP)

[+]
ADJP (NP (DNP (ADJP (JJ))
(DEG))
(NN)))

DP (NP (DNP (DP (DT))
(QP (CD))
(CLP (M)))
(DEG))
(ADJP (JJ))
(NP (NN)))

QP (NP (DNP (QP (CD)
 (CLP (M)))
 (DEG))
 (NP (NN)
 (NN)))

NP (NP (DNP (NP (DP(DT)
 (NP (NN)))
 (DEG))
 (NP (NN)))

PP (NP (DNP (PP (P)
 (NP (NP (NP (NP (NR)
 (ADJP (JJ))
 (NP (NN)))
 (QP (OD)
 (CLP (M)))
 (NP (NN)
 (CC)
 (NN)))
 (NP (NN)
 (NN)))
 (DEG))
 (NP (NN))))

LCP (NP (DNP (LCP (NP (NN)
 (LC))
 (DEG))
 (NP (NN)))

6 “地”字动词短语 (DVP)

[+]

(DVP (IP (NP (NN))
 (VP (VA)))
 (DEV))

7 片段 (FRAG)

(FRAG (PU))
 (VV))
 (PU))

8 方位短语 (LCP)

(LCP (VP (NP (NT))
 (VP (VV)))
 (LC))

(PP (P))
 (LCP (VP (VV))
 (LC)))

9 列表标记 (LST)

LST

1

LST

(IP (LST (CD 2)
 (PU))
 (VP (VV))
 (NP (NN))))

2

(IP (PP (P))
 (NP (NN)))
 (PU))

(NP (NT))
 (NP (NP (NN)
 (NN))
 (ADJP (JJ))
 (NP (NN)))
 (VP (VC))
 (PU)))
 (IP (LST (PU))
 (IP (IP (NP (NN))
 (VP (ADVP (AD))
 (VP (VV)))))
 (PU))
 (IP (NP (NN))
 (VP (PP (P))
 (NP (DP (DT))
 (NP (M)))))
 (VP (VV))
 (NP (NN))))))))
 (PU))
 (IP (LST (PU))
 (NP (NN)
 (NN)
 (NN))
 (VP (VV)
 (NP (ADJP (JJ))
 (NP (NN)))))
 (PU))

10 名词短语 (NP)

(NP (ADJP (JJ))
 (NP (NN)))

10.1 NP 的最低级别

1.

(VP (VV)
 (NP (NN))

(NP (NN))

2. -

(NP (NN)
 (NN))

3.

(NP (NN)
 (CC)
 (NN))

4. [+]

(NP (NR)
 (NN))

NR,NN

(NP (NP (NR)
 (NR))
 (NP (NN)
 (NN)
 (NN)))

5.

(NP (NT 2012)
 (NT 05)
 (NT 02))

(NP (NR)
 (NR))

10.2 NP 修饰语

1. QP

(NP (QP (CD 30)
(CLP (M)))
(ADJP (JJ))
(NP (NN)))

2. DP

(NP (DP (DT))
(NP (NN)))

3. ADJP

(NP (ADJP (JJ))
(NP (NN)))

4. DNP

(NP (DNP (NP (NN))
(DEG))
(NP (NN)))

5.

(1) NP

(NP (CP (IP (VP (NP (NT))
(VP (VE))
(NP (NN))
(NN)))))
(DEC))
(NP (NN))
(NN)))

(2) PP

(a)

(NP (CP (IP (NP (PN))
(VP (VV)))

(DEC))
(NP (NN)))

(b)

(NP (CP (IP (NP (NN))
(VP (VPT (VV)
(AD)
(AS))
(NP (ADJP (JJ))
(NP (NN)))))
(DEC))
(NP (NN))))))))

(c)

(IP (NP (NR)
(NN)
(NN))
(VP (VC)
(CP (IP (NP (NN))
(NN))
(VP (VV)
(NP (NN)
(NN)
(NN))))
(DEC))
(NP (NN)
(NN))))

(3)

(NP (CP (IP (VP (VV)
(NP (NN))))
(DEC)))

(4)

(a) NP

(NP (CP (IP (NP (NN))
(VP (VV)))
(NP (NN)))

(b) PP

(NP (CP (IP (NP (NN))
(VP (VP (VA)))
(NP (NN)))

6.

(IP (NP (NN))
(VP (VP (VV))
(VP (VA)))
(PU ,)
(NP (DP (DT))
(NP (NN)))

11 介词短语 (PP)

(PP (P)
(LCP (VP (VV))
(LC)))

12 插入语 (PRN)

PRN

PRN

(IP (PU ")
(NP (PN))

(VP (VV))
 (PU ")
 (PU)
 (PRN (IP (NP (NR))))
 (VP (VV)))
 (PU)
 (IP (PU "
 (NP (PN))
 (VP (VV))
 (VP (VV)))
 (PU ""))
 (PU .))

13 数量短语 (QP)

NP

(NP (NP (NN))
 (QP (CD 3.0)
 (CLP (M mm))))

14 非对等并列短语 (UCP)

(NP (UCP (ADJP (JJ))
 (PU)
 (NP (NN)))
 (NP (NN)))

15 动词短语 (VP)

15.1 并列动词复合(VCD)

(VP (VCD (VV))
 (VV)))

15.2 动词结果复合(VRD)

(VP (VRD (VV)
(VV)))

15.3 [修饰语+中心词](VSB)

(VP (VSB (VV)
(VV)))

15.4 [VV + VC] (VCP)

(VP (VCP (VV)
(VC)))

15.5. A-不-A, A-一-A (VNV)

(VNV (VV)
(AD)
(VV))

15.6 V-不-V, V-得-V (VPT)

(VPT (VV)
(AD)
(VV))

15.7 动词核心结构

1. [+ /]

(VP (VV)
(AS)
(IP (NP (ADJP (JJ)
(NP (NN)
(VA)))

2.

$VCD + VRD: (VRD (VCD (VV \quad)) (VV \quad))$
 $(VV \quad)$

$VCD + V-ASP: (VCD (VV \quad)) (VV \quad)$
 $(AS \quad)$

$VRD + V-ASP: (VRD (VA \quad)) (VV \quad)$
 $(AS \quad)$

3.

$(VP (VV \quad))$
 $(CC \quad)$
 $(VV \quad)))$

4.

$(IP (NP (NN \quad))$
 $(VP (NP (ADJP (JJ \quad))$
 $(NP (NN \quad)))))$

NP QP VP

15.8 VP 附加成分

1. ADVP

$(IP (ADJP (JJ \quad))$
 $(NP (NN \quad))$
 $(VP (ADVP (AD \quad))$
 $(ADVP (AD \quad))$
 $(VP (VV \quad))))$

2. NP()

(IP (NP (NP (NP (NP (NR))
(ADJP (JJ))
(NP (NN)))
(NP (NN)))
(NP-PN (NR)))
(VP (NP (NT))
(PP (P))
(LCP (IP (VP (VV))

NP VP

(VP (NP (ADJP (JJ))
(NP (NN)))
(VP (VV)))

3. QP

(IP (NP (NR))
(VP (ADVP (AD))
(QP (ADVP (ADVP))
(QP (CD))
(CLP (M)))
(VP (VV))

4. PP

(PP (P))
(LCP (NP (CP (IP (NP (NP (NR))
(NP (NP (NR))
(NR))
(NP (NN)))
(NP (DP (DT))
(NN (NN)))
(NP (NN)))

(VP (VV))
 (DEC))
 (NP (NN))
 (LC))
 (VP (VV))
 (NP (NN))))

5. CP

(IP (NP (DNP (NP (NR))
 (DEG))
 (NP (NN))
 (CC))
 (NN)))
 (VP (CP (ADVP (CS))
 (IP (VP (PP (P))
 (NP (NR))
 (VP (VV))
 (NP (ADJP (JJ))
 (NP (NN))))))))
 (VP (ADVP (AD))
 (VP (VV))
 (VP (VE))
 (NP (NN))))))))

6. IP

(IP (NP (NR))
 (VP (IP (NP- (NN))
 (VP (VV))
 (ADVP (AD))
 (VP (VV))
 (VP (VV))
 (NP (NN))))))))

7. DVP

(IP (NP (PN))
 (VP (DVP (IP (NP))

(VP (VV))
 (DEV))
 (VP (VV))
 (NP (NN))))

8. LCP

(VP (LCP (IP (NP (DNP (NP (NT))
 (DEG))
 (QP (CD))
 (NP (NN) (NN)))
 (VP (VV)))
 (LC))
 (PU))
 (ADVP (AD))
 (VP (VE))

15.9 从属子句

1.

(IP (NP (NN))
 (VP (ADVP (AD))
 (ADVP (AD))
 (VP (VV))
 (IP (NP (NR))
 (VP (NP (NT))
 (PP (P))
 (NP (NN))
 (NN))
 (NN)))
 (VP (VV))
 (NP (QP (CD))
 (ADJP (JJ))
 (NP (NN)))))))
 (PU))

(NP (NN)))))

3.

PP

(VP (PP (P)

(NP (NN)))

(ADV (AD))

(VP (VV))

4. MSP

[VP1/PP+MSP+VP2]

MSP

VP2

VP

(VP (MSP)

(VP (VP (VV))

(NP (NN)))

五、子句标注详细说明

1 以曲折成分为核心的简单从句(IP)

IP

NP VP

NP ,

1.1 简单陈述句

(IP (NP (NN)

(NN))

(VP (VA)))

1.2 被动句

(IP (NP (NN))
 (VP (BA)
 (IP (NP (NN))
 (VP (VP (VV))
 (NP (NN))

1.3 不定式

(IP (VP (VV)
 (IP (VP (VV)
 (NP (NN)))))

1.4 状语从句

IP

(IP (NP (NT))
 (PU)
 (IP (NP (NV)
 (CC)
 (NV))
 (VP (ADVP (CS))
 (ADVP (AD))
 (VP (VV)
 (NP (NT))))))
 (PU)
 (ADVP (AD))
 (NP (NV)
 (CC)
 (NV))

2.

(IP (IP (NP (NN))
 (VP (VV)
 (NP (NR))))
 (VP (ADVP (AD))
 (VP (VA)))))

3.

(IP (IP (NP (NR))
 (VP (PP (P)
 (NP (NN))))
 (VP (VV))))
 (NP (NN))
 (VP (VV))))

2 以补语为核心的从句(CP)

CP

2.1 关系子句

(IP (NP (CP (IP (NP (PN))
 (VP (VV))))
 (DEC))
 (NP (NP (NN))
 (ADJP (JJ))
 (NP (NN)))))
 (VP (ADVP (AD))
 (PP (P)
 (NP (NP (NR))
 (NP (NN)))))
 (VP (VV)))))

(IP (NP (CP (IP (VP (VV)
 (NP (ADJP (JJ))
 (NP (NN))))))
 (NP (NN))
 (VP (VV))))

2.2 从属子句

1. IP CP

(CP (IP (NP (NR))
 (PU))
 (NP (DP (DT))
 (NP (NN))))
 (VP (ADVP (AD))
 (ADVP (AD))
 (ADVP (AD))
 (VP (VV))))
 (SP))
 (PU))

2. CS ADVP

((IP (CP (ADVP (CS))
 (IP (NP (DNP (NP (NN))
 (DEG))
 (NP (NN))))
 (VP (ADVP (AD))
 (VP (VV))))))
 (PU))
 (ADVP (AD))

3. CS

(IP (NP (CP (IP (VP (VV)
 (NP (NR))

(NN)

(DEC))

(NP (NN))

(VP (ADVP (CS))

(VP (VV)

(NP (NN))))

2.3 问句

1

(CP (IP (NP (PN))

(VP (VC)

(NP (NR))))

(SP)

(PU ?))

2 [VP+]

(CP (IP (NP (PN))

(VP (VV)

(AS))

(SP)

(PU ?))

六、结构语法关系

1 补充关系

1.1 动词短语

1 (NP) (IP)

(IP

(NP (NN) (NN))

(VP (VV) (AS))

(NP (QP (CD))

(CLP (M))))

(NP (NN)))

(IP

(NP

(NP-PN (NR))

(ADJP (JJ))

(NP (NN)))

(VP (VV))

(IP

(NP (NR))

(VP (VV))

(NP

(NP (NN))

(PP (P))

(NP (NR)))

(NP (NN))))

(IP (NP (NR))

(VP (VV))

(IP (PU ")

(NP (NR))

(VP (VV))

(AS)))

(PU ")))

2

(IP (NP (NR))

(VP (VV))

(VP (ADVP (AD))

(VP (VV))

(NP (DNP (NP (NR))

(DEG))

(NP (NN))))

3 -IO

(IP (NP (NN))

(NN))
 (VP (VV)
 (NP-IO (ADJP (JJ))
 (NP (NN))))
 (NP (ADJP (JJ))
 (NP (NN))))))

4

(VP (VV)
 (NP (NN))
 (QP (CD)
 (CLP (M)))))

(VP (VV)
 (AS)
 (QP (CD)
 (CLP (M)))))

1.2 介词短语

(NP) IP - LCP

1) NP

(PP (P)
 (NP (DNP (NP (NN))
 (DEG))
 (NP (NN)))))

2) LCP

(PP (P)
 (LCP (NP (DNP (NP (NR))
 (DEG))
 (NP (NP (NR))
 (NP (NN))
 (NN)))))))
 (LC)))

3) IP

(PP (P)
(IP (NP (NN))
(VP (VV))))

1.3 介词-方位结构

- [(NP)/ (IP)+ (LC)]

1) NP LC

(VP (PP (P)
(LCP (NP (DNP (NP (NR))
(DEG))
(NP (NP (NR))
(NP (NN))
(NN))))
(LC)))

2) IP LC

(PP (P)
(LCP (IP (NP)
(VP (VV)
(NP (NN)
(NN))
(LC)))))

1.4 复句

1

(IP (CP (ADVP (CS))
(CP (IP (NP (DNP (LCP (NP (NN))
(LC))
(DEG))
(NP (PU)) (NN)) (PU)))
(VP (VC)
(NP-PRD (CLP (M))
(NP (PU)) (NN)) (PU))))

(SP))

2 SP IP

(CP (IP (NP (DT)
 (NN))
 (VP (ADVP (AD))
 (ADVP (AD))
 (ADVP (AD))
 (VP (VV)))))
 (SP)
 (PU))

1.5 “的”字名词短语

NP PP LCP QP

ADJP DNPs NP NPs

1 NP

(VP (VV)
 (AS)
 (NP-OBJ (DNP (NP (DP (DT))
 (NP (NN))
 (DEG))
 (NP (NN)
 (NN))

2 PP

(PP (P)
 (NP (DNP (PP (P)
 (NP-PN (NR))
 (DEG))
 (NP (NN)
 (PU)
 (NN))))

3 LCP

(NP (DNP (LCP (NP (QP (CD))
 (NP (NN)))
 (LC))
 (DEG))
 (NP (ADJP (JJ))
 (NP (NN)))))

4 QP

(NP (DNP (QP (ADJP (JJ))
 (QP (CD 20) (M))
 (DEG))
 (NP (NN))))

5 ADJP

(NP (DNP (ADJP (JJ))
 (DEG))
 (ADJP (JJ))
 (NP (NN))))

1.6 “地”字动词短语

(DVP)

VP NP IP DVP VP

1 VP

(VP (DVP (VP (VA))
 (DEV))
 (VP (VV))))

2 NP

(VP (DVP (NP (ADJP (JJ))
 (NP (NN))
 (DEV))
 (VP (VV))))

2 IP

(VP (DVP (IP (NP (NN)
 (VP (VA)))
 (DEV))
 (VP (VV)))

2 附加关系

IP

CP

2.1 复句

(IP (CP (ADVP (CS))
 (IP (NP (NP (NR))
 (CC)
 (NP (NN)
 (NN)))

(IP (CP (ADVP (AD)) (CS))
 (IP (NP (NR))
 (VP (VV))
 (VP (VV)))))

ADVP

ADVP

VP

(IP (IP (NP (NT))
 (PU)
 (NP (NN)
 (CC)
 (NN))
 (VP (ADVP (CS))
 (ADVP (AD))
 (VP (VV))

(NP (NT))))

2.2 单句

1

(IP (PP (PP))
 (NP (NN)))
 (NP (NT))
 (NP (NN))
 (NN))

2

(IP (NP (NR))
 (NP (NR))
 (VP (VV))
 (PU .))

(IP (NP (NN))
 (NP (NR))
 (VP (VV))
 (NP (NN)))
 (PU .))

3

IP

CS

CP

CP

(IP (CP (ADVP (CS))
 (IP (NP (DP (DT))
 (NP (NN)))
 (VP (ADVP (AD))
 (ADVP (AD))
 (VP (VA))))
 (ADVP (AD))

(NP (DP (DT))
 (CLP (M)))
 (NP (NN)))
 (VP (ADVP (AD))
 (VP (VA)))
 (PU .))

CS

IP

(IP (IP (NP (NR))
 (VP (ADVP (CS))
 (NP (NT))
 (VP (VV))
 (NP (NP (ADJP (JJ))
 (NP (NN)))
 (NP (NN))))))))
 (PU))
 (ADVP (AD))
 (NP (NN))
 (VP (ADVP (AD))
 (VP (VV)))
 (PU))

4 NP IP

(IP (NP (NT))
 (NP (NP (NR))
 (NP (NN))
 (NN))
 (NN)))
 (VP (VV))
 (QP (CD))))

2.3 名词短语

1) ADJP

(NP (ADJP (JJ))

(NP (NN)))

2) QP

(NP (QP (CD))

(CLP (M)))

(NP (NN)))

3) NP

(NP (NP (NR))

(ADJP (JJ))

(NP (NN)))

4) DNP

(NP (DNP (QP (ADJP (JJ))

(QP (CD 20))

(CLP (M)))

(DEG))

(NP (NN)))

5) DP

(NP (DP (DT))

(CLP (M)))

(NP (NN)))

6) CP

(NP (CP (IP (VP (ADVP (AD))

(VP (VA)))

(DEC))

(NP (NN)))

()

7) CP

(NP (CP (IP (NP (NN))

(NN)))

(VP (VV))

(NP (DP (DT))

(CLP (M))
(NP (NN))))

(DEC))

(NP (NN))

8) NP

(NP (NP (NP (NR))

(NP (NN))

(NP (NR))

9) IP

(PP (PP))

(NP (DP (DT))

(CLP (M))

(IP (NP (NN))

(VP (VV))

(NP (NR))))

(NP (NN))))

2.4 动词短语

1)

(IP (NP (PN))

(VP (ADVP (AD))

(VP (VV))

(NP (NN))))

2) DP QP

(IP (NP (NN))

(VP (DP (DT))

(CLP (M))

(VP (VV))

(NP (NN))

(QP (CD 5))

(CLP (M))))

3)

(IP (NP (NR))
 (VP (NP (NN))
 (VP (VV))
 (AS))))

4) DVP

(IP (NP (NR))
 (VP (DVP (VP (VA))
 (DEV))
 (VP (VV))
 (AS))
 (NP))))

5) PP

(IP (NP (NR))
 (VP (PP (P))
 (NP (NN)))
 (VP (VV))
 (AS))
 (NP))))

2.5 介词短语

:

(PP (ADVP (AD))
 (PP (P))
 (NP (NT))))

2.6 量词短语

:

(QP (ADVP (AD))
 (QP (CD))
 (CLP (M))))

2.7 量词短语

:

(VP (VV
 (QP (CD 107
 (CLP (NP (NR
 (CLP (M))))

(NP (QP (CD
 (CLP (ADJP (JJ
 (CLP (M))))
 (NP (NN))))

2.8 形容词短语

:
 (ADJP (ADVP (AD
 (ADJP (JJ))))

(ADJP (NP (NP (NR
 (NP (NN))))
 (ADJP (JJ))))

3 并列关系

UCP

3.1 词并列

1
 (IP (NP (NN
 (CC
 (NN))
 (VP (NP (NT
 (VP (VV))))

(NP (NP (NR
 (NP (NN))))

(PU,)
(NR)
(PU,)
(NR)
(ETC))
(NP))

2

(VP (CC)
(VA)
(CC)
(VV))

3

(QP (CD)
(CC)
(CD))

3.2 短语并列

(NP (NP (NR))
(NP (NP (NN)))
(CC)
(NP (NP (PN))
(NP (NN)))
(NP (NN)))
(NP (NP (NP (NP (NR)))
(NP (NN)))
(PU)
(NP (NP (NR)))
(NP (NN)))
(ETC))
(NP (NN))
(ADJP (JJ))
(NP (NN)))

3.3 子句并列

1)

```
(IP (IP (NP (NP (NR   ))
                (NP (NN   )))
    (VP (VV   )
        (IP (VP (BA   )
                (IP (NP (NP (PN   ))
                        (PP (P   )
                            (NP (NP (NR   ))
                                (NP (NN   )
                                    (NN   ))))
                            (NP (NN   )))
                        (VP (PP (P   )
                            (QP (CD 150)
                                (M   )))
                        (VP (VRD (VV   )
                            (VV   ))
                        (QP (CD 200)
                            (M   ))))))))
    (PU   )
    (CC   )
    (IP (VP (VV   )
        (IP (NP- (NN   ))
            (VP (VV   ))))
    (PU   ))
```

2)

```
(IP (IP (NP (ADJP (JJ   ))
                (NP (NN   )))
    (VP (NP (QP (CD   ))
        (NP (NN   ))))
    (PU   )
    (IP (NP (NP (ADJP (JJ   ))
        (NP (NN   )))
        (NP (ADJP (JJ   ))
```

(NP (NN
 (VP (QP (CD
 (CLP (M
 (PU)
 (IP (VP (ADVP (AD
 (PP (P
 (NP (NT
 (NN
 (VP (VV
 (QP (CD
 (CC
 (CD
 (PU))

4 概要表示

1

(XP X
 (YP)
 (ZP)
 ...)

(XP (YP)
 (ZP)
 ...
 X)

2 :

:
 (XP (YP)
 (ZP)
 ...
 (XP))

3

(XP {CONJ}

(XP)

{CONJ}

(XP)

...)