

蒙古语言文字数字资源建设与共享工程

信息处理用蒙古文相关标准

MGC/01-08

信息技术 传统蒙古文语义角色分类体系及其 标记集

Information technology — Classification system and its marking set of traditional
Mongolian semantic roles

内蒙古自治区民族事务委员会

内蒙古大学

2016年11月

前 言

本标准是依据《现代蒙古语语义框架研究》、《面向信息处理的蒙古语动词语义研究》等专著和《面向信息处理的蒙古语简单谓动句句模研究》、《基于语料库的蒙古语简单句施事和当事语义角色特征研究》等博士学位论文为基础，规定了信息处理用的蒙古文语义角色分类体系及其标记集。

本标准由内蒙古自治区民族事务委员会提出。

本标准起草单位：内蒙古大学。

本标准起草人：额尔敦朝鲁、德·萨日娜、富涛、包晓荣、斯琴夫、苏布德、娜荷芽

信息技术 传统蒙古文语义角色分类体系及其标记集

1 范围

本标准规定了信息处理用蒙古文语义角色标记集。
本标准适用于蒙古文的各种语料库。

2 术语和定义

下列术语和定义适用于本标准。

2.1

语义角色 semantic roles

受谓词支配的或相关的体词性成分。

2.2

主体 subject

在句子语义结构中动作行为、状态、心理活动、关系的主体或被说明的主题。

2.3

施事 agent

有意发起的动作行为的主体。

2.4

当事 parties

自然发生的动作或变化的主体。

2.5

主题 topic

表明或描写性质、特征、状态的主题。

2.6

起事 source

关系双方的起方。即在表关系的动核结构中被说明的主体。

2.7

感事 experiencer

与心理动词相联系各类心理活动的体验者、感受者或经验者。

2.8

表事 reflection

性状动词所联系的主体。

2.9

共事 dative

与施事协同或共同完成动作行为的主体。

2.10

使事 causation

使施事发起动作行为的主体。

2.11

客体 object

动词所联系的客体动元，即主体作用于动词后动核所支配的对象。

- 2.12
受事 patient
受有意发起动作行为影响的直接承受者。
- 2.13
涉事 concerned
不受有意发起动作行为影响的承受者。
- 2.14
与事 receiver
动作行为的交接或针对对象。
- 2.15
成事 result
动作的结果或成果，是动作发生后产生或出现的客体。
- 2.16
转事 transformation
主体的变换或替换主体的对象。
- 2.17
释事 explanation
解释受事、涉事的语义成分。
- 2.18
比事 comparison
与主体作比较的对象。
- 2.19
任事 role
担任某种角色的语义成分。
- 2.20
止事 goal
关系双方的止方。即在表关系的动核结构中说明起事的语义成分。
- 2.21
邻体 complement
补充说明动作行为的语义成分。
- 2.22
时间 time
动作行为或状态发生、进行、结束的时间。
- 2.23
处所 location
动作行为或状态发生、进行、结束的空间、场所。
- 2.24
方向 direction
动作行为所指向的方向。
- 2.25
工具 instrument
动作行为发出时所使用的工具、材料。
- 2.26
原因 reason

动作行为发生的原因。

2.27

目的 aim

动作行为所要达到的目的。

2.28

依据 accordance

动作行为所依从的标准、前提或基础。

2.29

范围 range

动作行为发生或出现的范围。

2.30

境况 circumstances

动作行为发生的境况。

2.31

所属 ownership

跟谓词有间接语义关系，当事所归属的对象。

2.32

饰体 decoration

对动作行为、状态起修饰性描写的成分。

2.33

性状 status

动作行为发起的过程、性质和条件。

2.34

数量 amount

动作行为、状态发生的数量。

2.35

情态 modality

表示动作行为情态的成分。

2.36

频度 frequency

动作行为、状态发生的次数。

2.37

定语 attribution

修饰名核的成分。

2.38

区别 distinction

表示名核类别的成分。

2.39

领事 possessive

表示名核领属范围的成分。

2.40

比喻 metaphor

比喻名核的成分。

2.41

性质 property

表示名核性质和特征的成分。

3 标记集的内容

本标准包括语义成分的标记符号。

4 标记代码的制定原则

本标准标记集包括两级标记。

4.1 《第一级标记》的命名方法

用大写英文字母标记第一级。如：

OGO (来自 ogolegdegci): 主体

TVS (来自 tvsvgci): 客体

在语料中无需标注第一级标记。

4.2 《第二级标记》的命名方法

用小写英文字母标记第二级。如：

sed (来自 sedub): 主题

ham (来自 hamtvrarci): 共事

在语料中只标注第二级标记。

5 标记集

以下是标记集内容。

语义角色标记代码见表 1。

表 1 语义角色标记代码

序号	例句	标记代码		语义角色名称	代码说明
		大类	小类		
01		OGO		主体	ogolegdegci
02	<u>abu</u> jöbsiyerbe		uid	施事	<u>uile</u> dugci
03	<u>borogan</u> arilba		uij	当事	<u>uile</u> jigci
04	<u>xödelmürilexü-bol</u> cixula ajił yum		sed	主题	<u>sedub</u>
05	<u>oi-yin böxö talabai-ni</u> man-u ulus-un nutug debisger-ün xagas-tay tengčen_e		ehl	起事	ehilegci
06	<u>ta</u> man-i sanan bayilguý		ser	感事	<u>sere</u> gci
07	<u>xeüxen orgil sünderle.jü</u> bayin a		hrg	表事	<u>haragv</u> lgci
08	<u>bi jirgal-tay</u> agulžaba		ham	共事	<u>hamt</u> vrarci
09	<u>xuwang sü ing_eji-ber-iyen</u> nige xulu čagan gulir jigelegülbe		uld	使事	<u>uile</u> dugulugci
10		TVS		客体	<u>tvsv</u> gci
11	tere <u>xömo_n</u> jõdoba		hur	受事	<u>hurte</u> gci
12	bi <u>nam-un-iyen jigaburi-yi</u> sonosba		ort	涉事	<u>ortog</u> ci
13	bi <u>teguñ-du</u> nemeri ögbe		edl	与事	<u>edle</u> gci
14	tere <u>nom-iyen</u> exilen bičibe		but	成事	<u>butug</u> ci

15	gutul-ni xarin <u>ayil-un hurim-un demčl bolba</u>		hvb	转事	<u>hvbisvgci</u>
16	eyimü-ecē malčid tegün-i <u>ejl nagur geju</u> nerelejey		tal	释事	<u>tayilbvrilagci</u>
17	bide xalagun dulagan ger-i <u>eji-yin-iyen</u> <u>elige-ber uligerledeg</u>		adl	比事	<u>adalidhagci</u>
18	oi-yin bóxó talabay-ni <u>man-u ulus-un nutug</u> <u>debisger-ün xagas-tai tengčen_e</u>		dag	止事	<u>dagvsvgci</u>
19	nam-ača li da nöxör-i <u>erxilegci-ber</u> tomilaba		egu	任事	<u>egurgelegci</u>
20		NOH		邻体	<u>nohocegci</u>
21	ja, <u>tabun cag-tu</u> aguljay_a		cag	时间	<u>cag</u>
22	bi <u>begejing-du</u> ociba		0r0	处所	<u>0r0n</u>
23	nidü čičim_e xarangguy-yi <u>nasi časi jušün</u>		jug	方向	<u>jug cig</u>
24	hómós <u>hele-ber ujel</u> sanag_a-ban solilcadag		bag	工具	<u>bagaji</u>
25	tere <u>jobaxu-dagan</u> oxilaba		sil	原因	<u>siltagan</u>
26	önir <u>xüseleng-ün tölöge</u> dabsin_a		j0r	目的	<u>j0rilg_a</u>
27	<u>nam-un jarčim-iyar</u> bey_e-ben čingg_a šagarday_a		und	依据	<u>undusulel</u>
28	<u>jarim tal a-bar</u> basa dögomčileju bolon_a		heb	范围	<u>hebciy_e</u>
29	<u>eng-ün bayidal-du</u> türüguin mör-tü-ni tölöblege-yin ner_e-yi bičin_e		0rc	境况	<u>0rcin</u>
30	<u>bars-tu jigür</u> urguba		hay	所属	<u>hariy_a</u>
31		CIM		饰体	<u>cimegci</u>
32	denglü <u>nigente</u> untarajay		cin	性状	<u>cinar bayidal</u>
33	<u>süme_e-yin jing</u> arban xoyar cohil_a		t0h	数量	<u>t0g_ahemjiy_e</u>
34	tede čimay <u>lab xoyar</u> xonog amaragulxu bayixuu		yan	情态	<u>yanjv handvlg_a</u>
35	<u>basa yamar nigen</u> yagum_a pas pas duugarun_a		dah	频度	<u>dahilta</u>
36		TOD		定语	<u>t0d0dhagci</u>
37	<u>ene tal a-yin</u> surxu bičig nayiragulba		ilg	区别	<u>ilgal</u>
38	<u>biden-u</u> ongoča xödelul_e		har	领事	<u>harivalal</u>
39	<u>ayungg_a metü</u> ayixubtur dagun garun_a		jui	比喻	<u>juirlel</u>
40	<u>uran buljumor</u> modon deger_e dagulan_a		0ns	性质	<u>0ncalig sinji</u>