

SpaCE2022 – task2 标注任务工作指南要点（詹卫东 2022-06-21）

核心问题	答案
为何标注	(1) 描述句子的空间信息有什么问题 (2) 为计算机空间语义理解提供负例数据
如何标注	(1) 通读：一句一句通读全段 (2) 找错：找到有信息异常点的小句（标点句） (3) 分类：对异常情况进行分类 (4) 转写：把异常点用S、P、E要素方式呈现出来 (5) 验证：看看（4）是否符合（2）的原意
如何判断	(1) 基于个人自然语感的判断（不自然、费解，即异常） (2) 基于日常生活常识的推理（合乎常情，无关严密逻辑）
如何优化	(1) 形式整齐优于形式杂乱 (2) 多原因优于单原因 (3) “归入正常”优于“为异常强找理由”
怎分高下	(1) 空间信息描述清晰准确是最高标准 (2) 不一定能做到最好，但千万别做最差：避免严重的信息描述错误（跨句远距离选错要素成分，填入“的”“的话”等助词性成分，弄错S、P、E要素的填写位置，等等，均为严重错误）

分类	含义	示例
搭配不当	句中有两个表示空间信息的语言成分不能组合	把头放在地面边
语义冲突	句中有两处空间信息存在矛盾、冲突	孩子们睡在蒙古包里，他们一醒来，就跳起来跑进蒙古包找妈妈
不符合常识或背景信息	句中有一处文本片段的空间语义违反常识或者违反句子的背景信息	南川河、北川河、西川河从北、北、西三方汇入
并无异常	一段话空间语义正常	圆明园在北京西北部，是一座举世闻名的皇家园林。
不适合用于计算机空间语义理解	一段话有错字、句子整体无空间方位义、句中存在敏感信息	由谁去充任那些苦难的角色？又有谁去体现这世间的幸福，骄傲和快乐？

空间语义三要素	S（空间实体）	P（空间方位）	E（空间事件）
含义	一个实体是空间实体，当且仅当句中描述该实体的空间方位信息	空间实体方位是句中描述空间实体S的空间方位信息的语言成分	空间义相关事件是表达空间实体S位于P的方式、目的等的核心谓词成分
示例1	奖状	在墙上	贴
示例2	戒指里面刻着他的名字 [_P 戒指里面] [_E 刻] 着 [_S 他的名字]		
示例3	她从不坐贵宾席 [_S 她] 从不 [_E 坐] [_P 贵宾席]		
示例4	大客车未遵守信号灯指示就右转弯由西向东行驶 [_S 大客车] 未遵守信号灯指示就 [_P 右转弯由西向东] [_E 行驶]		

句子	异常类型	异常片段					
她胳膊上少了那只篮子，她把它塞到那个女学生座位四面了	搭配不当	text1			text2		
		塞到			四面		
孩子们睡在蒙古包里，他们一醒来，就跳起来跑进蒙古包找妈妈	语义冲突	S1	P1	E1	S2	P2	E2
		孩子们	在蒙古包里	睡	孩子们	进蒙古包	跑
打开西宁的卫星遥感地图，南川河、北川河、西川河从北、北、西三方汇入，在城区形成湟水河，自西向东穿城而过。	不符合常识或背景信息	S			P		E
		南川河、北川河、西川河			从北、北、西		汇入

空间语义要素标注原则	要点	具体要求
原则1： 要素完整原则	勿遗漏信息	(1) S、P、E尽量齐备 (2) S里填写的成分清楚完整；P里填写的成分清楚完整；E里填写的成分清楚完整
原则2： 要素邻近原则	勿舍近求远	(1) S、P、E尽量在一个标点句范围内选取，非必要，不跨出当前标点句。 (2) 如果跨标点句选S、P、E成分，一定是实义成分，而非功能成分（比如：介词不需要也不能跨标点句选取）；一定是离当前标点句最近的成分。
原则3： 语义简明原则	勿冗词赘语	(1) S、P、E能清楚表达空间信息即可，不需要填写修饰性成分。 (2) E中动词词组如果包含表达空间信息的修饰性成分，应该填入。
原则4： 平行比对原则	勿形义错配	(1) 两组<S, P, E> 尽量保持无空间信息冲突的成分形式一样，有空间信息冲突的成分形式有别 (2) S1 – S2, P1 – P2, E1 – E2 并行对照
原则间关系	(1) 如果有多个候选标注方案，那么，符合原则最多的方案最优 (2) 原则的优先级顺序是：原则1 > 原则4 > 原则3 > 原则2	

①	句子	异常类型	异常片段			
	她胳膊上少了那只篮子，她把它塞到那个女学生座位四面了	搭配不当	text1		text2	
			塞到		四面	
		不符合常识或背景信息	S	P		E
			它	到那个女学生座位四面		塞

②	句子	异常类型	异常片段			
	核桃被他从左扔到左，从右扔到左	搭配不当	text1		text2	
			从左		扔到左	
		不符合常识或背景信息	S	P		E
			核桃	从左到左		扔

③	句子	异常类型	异常片段			
	许多村民是坐在草垛上、在村口的路灯上	搭配不当	text1		text2	
			坐		在村口的路灯上	
		不符合常识或背景信息	S	P		E
			村民	在村口的路灯上		坐

Task2 标注工作指南

https://2030nlp.github.io/Sp22AnnoOL/task2_guide.html

Task2 标注工作常见问题说明

https://2030nlp.github.io/Sp22AnnoOL/task2_complement.html

Task2 标注需注意问题及常见错误示例

<https://docs.qq.com/doc/DSXVNdW9hemlJRVB5?>

Task2_双人标注不一致_原因分析

<https://docs.qq.com/doc/DSXVxTGdvRUpOSHVy?>

SpaCE2022 标注工作文档全集

<https://2030nlp.github.io/Sp22AnnoOL/menu>