

糖尿病日常咨询对话数据标注规范

Annotation Schema

版本日期：2022-10-05

杨锦锋，侯欣彤，高欣宇

版本修改规则：

- (1) 当前确定的内容为黑色文字
- (2) 当前删除的内容用红色和删除线表示
- (3) 新版本修改的内容用绿色文字显示
- (4) 产生新版本时，上一版本绿色文字如果是确定内容就恢复为黑色文字
- (5) 暂时不确定的内容用黄色作为文字底色
- (6) 标注示例用蓝色文字
- (7) 产生新版本时，文件名命名规则为“annotation schema – 20211119.docx”，同时修改第一页的版本日期信息。

规范优化的方向：

- (1) 规范可以明确的解释，没有歧义，没有模棱两可
- (2) 规范易于操作
- (3) 不同的人能得到相同的标注结果

标注原则

- (1) 如果一个句话有多个标注方案，使用相对简单的标注方案

- (2) 尽量按照字面意思标注，不做引申，不做推理

我有凌晨三点测过也是 7 点多到 8 点的样子

→ inform (时间=凌晨三点，血糖值=7 点多到 8 点)

- (3) 标注表达主要意思的语句

可以喝点青钱柳吗，前段时间吃东西没有注意，有点身高

→ ask-for-sure (药品=青钱柳) 后面的不标注

请问糖尿病能吃海鲜吗？平时在生活中应该注意些什么？

→ inform(疾病=糖尿病，状态=阳性)

ask-for-sure(饮食名=海鲜)

“平时在生活中应该注意些什么？”这句话就不标注

完善记录

2021-12-03

- (1) 在治疗 domain 里增加检查项槽位
- (2) ask-whether 合并到 ask-for-sure 中

2021-12-09

- (1) ask-for-sure 标注示例完善

2021-12-22

- (1) act label 增加一个 uncertain
- (2) 给饮食、运动、行为、治疗都增加一个 slot: 效果

2022-1-3

- (1) 为 advice-not、ask-for-sure、uncertain、general-explanation 增加示例
- (2) 为 deny() 三种标注情况做了整理
- (3) 放弃花括号表示组合关系的方式, 采用使用多个相同的 act label, 组合关系由一个 act label 的 arguments 来表示。
- (4) 部分格式修改

2022-1-18

- (1) 增加检查值的 slot (后期把检查项和检查值从治疗分出来)
- (2) 标注系统: <http://dialogdata.geniusdt.com:2022/>

2022-1-28

- (1) 对于一个槽位多个槽值的情况, 给出适用于标注系统的等价标注形式, 参见第 11 页

2022-03-30

- (1) 问题领域的状态槽位暂时不标注, 降低标注负担

2022-04-16

- (1) 补充 ask-how 的标注说明和标注样例

2022-04-30

- (1) 删除每个 domain 里的“影响问题” slot, 这个修改不在标注系统里, 到时在导出的数据里统一把“影响问题”替换为“问题” domain 的“症状” slot。
- (2) 给 general-advice 和 general-explanation 标注 domain, 以便更好的指导 policy 和 NLG

2022-10-05

- (1) 删除“影响问题” slot

1 Domain and Slots

Domain 和对应的 slots 是对糖尿病健康管理领域重要信息的结构化表示，可以作为医生在指导患者进行健康管理时的信息获取目标和信息传达目标，同时也限定对话内容的信息涉及范围。后续可进一步基于 domain 和 slots 设计数据库表，存储结构化的咨询信息。

1.1 问题（Problem）

患者的疾病、症状、以及诱发和被诱发的症状。问题及相关实体属性详情如表 1 所示。

表 1 问题及相关实体属性详情

Slot	Description	Values	Examples
疾病	语句中提及的疾病	从语句中识别	昨天开始没吃，我是妊娠糖尿病，但是已经生完四个月了；
症状	语句中提及的主观症状	从语句中识别，可以是多值，考虑规范化	林医生尿黄有泡沫，可是查了尿常规都还正常
症状部位	语句中提及的症状发生部位	从语句中识别，可以是多值，考虑规范化	对血管没有影响
持续时长	疾病或症状的持续时长		
状态	描述疾病或者症状是否正在发生，只对疾病和主观症状进行识别。参考“医学对话临床发现阴阳性判别任务” ¹	阴性、阳性、其他	
血糖值	语句中泛泛地提及血糖，没有明确是什么时候的血糖值	从语句中识别，可能是数字，也可能是高、低等的粗略判断	空腹高一般什么引起的；[高] 今天空腹 7•3 中午餐二 6•4 没吃药。 我想问下为什么我吃药空腹都没下过六？
时间	血糖值测试的时间点	从语句中识别，可能是餐前餐后或者夜间，也可能是类似于昨天前天(默认为当天)，可以没有值	空腹高一般什么引起的 今天空腹 7•3 中午餐二 6•4 没吃药。

1.1.1 血糖值和时间的槽位

- 血糖值的表达示例：
空腹高一般什么引起的
今天空腹 7•3 中午餐二 6•4 没吃药。
今天空腹 7•3 中午餐二 6•4 没吃药。
我想问下为什么我吃药空腹都没下过六？
吃了糙米餐一 9.0 餐二 9.4；

¹ <http://cips-chip.org.cn/2021/eval1>

吃了糙米餐一 9.0 餐二 9.4;
餐二都可以在 7•8 以下;
可是我餐后都可以达标
餐二都可以在 7•8 以下
早上餐二 6.5 中午就 9.1
早上餐二 6.5 中午就 9.1
今天空腹 7•3 中午餐二 6•4 没吃药。
今天空腹 7•3 中午餐二 6•4 没吃药。
我晚上餐后不高;
就是夜间的血糖太低了, 到第二天早上血糖反弹性的升高
就是夜间的血糖太低了, 到第二天早上血糖反弹性的升高
哪您其他阶段的血糖怎样? 您的餐后血糖还可以了
餐三一般是 5•8

血糖值的可能表达方式:

高、不高、低、升高、9.0、9.4、没下过六、达标、不能正常、7 以下、5-7、可以、6 点多、8 左右、没有超过七、不高于 7.8、

● 血糖测试时间的表达示例:

空腹高一般什么引起的
今天空腹 7•3 中午餐二 6•4 没吃药。
今天空腹 7•3 中午餐二 6•4 没吃药。
我想问下为什么我吃药空腹都没下过六?
吃了糙米餐一 9.0 餐二 9.4;
吃了糙米餐一 9.0 餐二 9.4;
餐二都可以在 7•8 以下;
可是我餐后都可以达标
餐二都可以在 7•8 以下
早上餐二 6.5 中午就 9.1
早上餐二 6.5 中午就 9.1
我晚上餐后不高;
就是夜间的血糖太低了, 到第二天早上血糖反弹性的升高
就是夜间的血糖太低了, 到第二天早上血糖反弹性的升高
哪您其他阶段的血糖怎样? 您的餐后血糖还可以了
餐三一般是 5•8
前天中午还只有 5•5

1.1.2 状态槽位

状态的四个值的判断标准

- (1) 阳性(pos): 症状/疾病明确是存在, 并且是患者本人的。
- (2) 阴性(neg): 症状/疾病明确不存在的, 并且这个表达是与患者本人相关的;
- (3) 其他(other): 未知的标注其他, 一般指用户没有回答、不知道或者回答不明确/模棱两可不好推断、或不是患者本人的情况, 或者是无实际意义的, 一般是医生的解释说是

一般知识，和病人当前的状态条件独立不具有标注意义，及有些检查项带疾病名称的，识别的疾病（乙肝 五项/ 乙肝 抗体），药品名中出现的“疾病”不标注。

标注的时候只考虑上文和当前语句，不考虑下文。如果一个症状在当前语句中判断为阴性，在下文中判断为阳性，则下文的判断会更新当前这个判断，这个更新在状态跟踪中完成。

标注示例：

最近感觉好累，气喘 → 症状=[累，气喘]，状态=阳性

糖尿病急剧消瘦是什么原因？ → 症状=消瘦，状态=阳性

没有过低血糖 → 症状=低血糖，状态=阴性

诺和龙对肝脏有损害，对消化道有刺激症状，比如腹痛、腹泻、恶心呕吐、便秘

→ 影响症状=[腹痛、腹泻、恶心、呕吐、便秘]，状态=其他

User: 低血糖有什么伤害？ → 症状=[累，气喘]，状态=阳性

Doctor: 会出现脑部眩晕 → 影响症状=脑部眩晕，状态=其他

1.2 饮食（Diet）

患者的日常饮食信息，吃的什么，吃多少，饮食主要成分及其成分含量，饮食的影响。饮食及相关实体属性详情如表 2 所示。

表 2 饮食及相关实体属性详情

Slot	Description	Values	Examples
饮食名	对“吃什么”的表达。 有些是菜名，有些是具体的食材，有些是模糊的说法	从语句中识别的字符序列	我吃 莲藕排骨 还喝 汤 ； 吃的 杂粮馒头 ； 南瓜 含糖量比较多； 午餐 吃的多了吧；
时间	饮食的时间点	从语句中识别的字符序列	今天 中午 四两熟米饭，炒白菜肉都没吃
饮食量	对“吃多少”的表达，大多是模糊的说法，因为糖尿病患者常见的问题是吃多了	从语句中识别的字符序列	这个考虑你喝的汤有点 多 ； 午餐吃的 多 了吧；
成分	饮食中的营养成分	可以从语句中识别，再做规范化	南瓜含 糖 量比较多； 主要是糙米的含的 维生素 比较多
成份量	饮食中某种营养成分多少的表达，通常比较模糊	可以从语句中识别，再做规范化	主要是糙米的含的 维生素 比较 多 ； 南瓜含糖量比较 多
效果	描述饮食带来的效果	从语句中识别	

1.3 行为（Behavior）

患者日常行为的信息，包括行为名，行为的时间，频率。行为及相关实体属性详情如表 3 所示。

表 3 行为及相关实体属性详情

Slot	Description	Values	Examples
行为名	语句中提及行为的表达，语义上比较宽泛。运动、吃药、少吃、测血糖、抽烟、喝酒这类笼统的表达都可视为行为	从语句中识别	建议你 运动 加上，适量的 少吃 ； 林医生，血糖需要每天都 测 吗？ 我已经 吃药 一个半月 睡眠不好 也会影响血糖
频率	语句中描述行为的频率	从语句中识别	一般建议 一周测两次 ； 我现在停药，但是 天天 都会测。
时间	语句中表达行为发生的时间	从语句中识别	我 现在 停药，但是天天都会测。
持续时长	语句中表达行为持续的时长	从语句中识别	我已经吃药 一个半月
效果	描述行为带来的效果	从语句中识别	但是经过锻炼控制的 不错

1.4 运动（Sport）

明确表达的一种运动项目。运动及相关实体属性详情如表 4 所示。

表 4 运动及相关实体属性详情

Slot	Description	Values	Examples
运动名	语句中提及运动项目的表达，	从语句中识别	然后早上 1 个小时 羽毛球 晚上 2 个小时 羽毛球
频率	语句中描述运动的频率	从语句中识别	
时间	语句中表达运动发生的时间	从语句中识别	
持续时长	语句中表达运动持续的时长	从语句中识别	晚上 2 个小时 羽毛球
强度	语句中描述运动的强度	从语句中识别	现在天冷了 打羽毛球都 不怎么出汗 了
效果	聊天涉及到的运动效果	从语句中识别	

1.5 治疗（Treatment）

患者日常采取的治疗，通常包括药品名、处置名，还可以包括剂量、适应症。治疗及相关实体属性详情如表 5 所示。

表 5 治疗及相关实体属性详情

Slot	Description	Values	Examples
药品	有关药品的称谓	从语句中识别	林医生 胰岛素 打腰

			部一圈都可以吗？ 建议你阿卡波糖吃两粒
用药量	用药量的表达，有时是模糊的表达	从语句中识别	建议你阿卡波糖吃两粒
用药(治疗)频率	用药频率或者治疗频率	从语句中识别	吃阿卡波糖一天三次一次1粒
时间	治疗或者用药时间		
持续时长	治疗或者用药时长		
适应症	药品针对的主要问题，可以关联到症状	从语句中识别，考虑规范化，可以是多值	
药品类型	药品的用途类型	从语句中识别，考虑规范化，可以是多值	螺旋霉素是消炎药
治疗名	药品以外的治疗手段或者保健方法、或者检查手段/项目	从语句中识别	换不同的地方注射
部位	治疗施加的部位，有时也是笼统的说法	从语句中识别	换不同的地方注射； 我现在都是在腹部两边换着打
检查项	聊天涉及到的检查项目名称	从语句中识别	没有其他症状吧。 严格尿常规吧和24h蛋白尿。看看有没有尿蛋白 周日我做的检查糖化血红蛋白6.4还可以吧 查一下糖化血红蛋白
检查值	检查项对应的检查值	从语句中识别，并且与检查项同时出现	周日我做的检查糖化血红蛋白6.4还可以吧
效果	聊天涉及到的治疗效果	从语句中识别	吃了两天二甲胃受不了就不吃了，后来自己降下来了

1.6 基本信息 (profile)

基本信息的实体属性有身高、体重、年龄、性别、既往史。主要标记患者自身的身体信息。基本信息及相关实体属性详情如表6所示。

表6 基本信息及相关实体属性详情

Slot	Description	Values	Examples
身高	语句中提及身高的表达，	从语句中识别	100斤 151cm
体重	语句中提及体重的表达，	从语句中识别	您体重是？ 100斤 151cm

年龄	语句中提及年龄的表达，	从语句中识别	
性别	语句中提及性别的表达，	从语句中识别	
既往史	语句中提及既往史的表达，	从语句中识别	

1.7 六个领域的关联关系

六个领域的关联关系如图 1 所示。

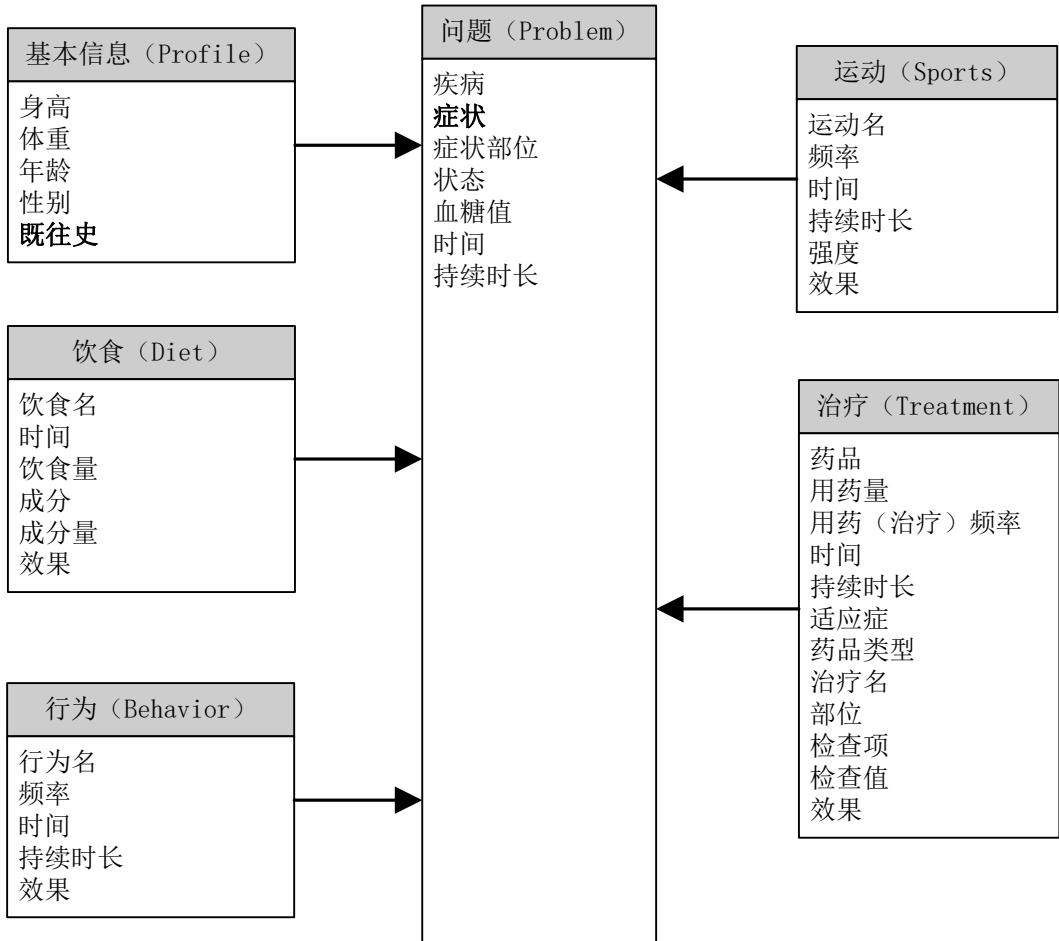


图 1 六个领域的关联关系

2 Act labels

Act labels 表示语句的动作，是要告知对方还是要询问对方，或者是建议什么。

Act label 与对话的角色相关，有些 act label 表示用户的动作，有些是表示医生的动作，有些是用户和医生都具有的动作。

slot-value 作为 act 的 arguments，act 和 slot-value 一起表示说话者每句话的意图。有的 act 可以没有 slot-value。Slot 的标注形式是 domain 和 slot 组合到一起构成 slot 的名称，即 domain.slot。后续标注的时候注意下。

标注方式可以是这样的形式：

Act (slot1= value1, slot2= value2,)

一句话可以有多个 act，多个相同的 act 可以合并一个。

(1) 用同一种 act label 来表示组合关系

另外，有些 slot 是搭配出现的，比如血糖值和时间，这两者的对应关系需要保持，因此需要记录多个 slot 的组合关系，可以多次使用同一种 act label 来表示组合关系。比如下面的标注表达的是用户想告知对方四项信息，并且 slot1= value1 和 slot2= value2 是对应的，slot3= value3 和 slot4= value4 是对应的，这个对应关系不能乱。

inform (slot1= value1, slot2= value2)

inform (slot3= value3, slot4= value4)

(2) 一个 slot 多个 value 的情况

在下面的标注示例中，采取的方案是用[]来表示多个值，比如：

Act(slot=[value1, value2])

这种方式在标注系统里不支持，可以采取标注系统支持的等价形式，如下：

Act(slot= value1, slot= value2)

示例：

建议你运动加上，适量的少吃 → advice(行为=[运动, 少吃])

等价形式：advice(行为=运动，行为=少吃)

2.1 inform (患者)

患者向医生告知信息，这是最常见的一种 act

示例：

我吃莲藕排骨还喝汤，血糖升高 → Inform(症状=血糖升高，饮食名=[莲藕排骨，汤])

早上餐二 6.5 中午就 9.1

→ inform(血糖值=6.5，时间=早餐二);

inform(血糖值=9.1，时间=午餐二)

吃了糙米餐一 9.0 餐二 9.4

→ inform(饮食名=糙米);

inform(血糖值=9.0，时间=餐一);

inform(血糖值=9.2，时间=餐二);

2.2 advice (医生)

医生表达对患者的建议或者推荐，医生对患者的指导也标注为 advice

示例：

建议你运动加上，适量的少吃 → *advice(行为=[运动,少吃])*

在低血糖的时候吃一些饼干 → *advice(症状=低血糖, 饮食=饼干)*

建议你阿卡波糖吃两粒 → *advice(药物名=阿卡波糖, 用药量=两粒)*

换不同的地方注射 → *advice(治疗名=注射, 部位=不同的地方)*

2.3 advice-not (医生)

医生不建议患者的事情

示例：

不建议喝酒吸烟 → *advice-not(行为=[喝酒, 吸烟])*

这个建议你还是正常的吃吧，不要在糙米了 → *advice-not(饮食名=糙米)*

还是别喝可乐，含糖量高 → *advice-not(饮食名=可乐)*

2.4 accept (患者)

患者表示接收医生的建议、解释、诊断等，是对 *advice* 和 *advice-not*、*explanation* 的回应，没有 arguments。

示例：

Doctor: 这个您停药观察一段时间看看

User: 我再停段时间看看。 → *accept()*

2.5 explanation (医生)

医生回答患者的提问，主要是两方面：解答患者的疑问，或者解释某个诊断或推荐的理由，也可以表示诊断。

示例：

主要是糙米的含的维生素比较多 → *explanation(饮食=糙米, 成分=维生素, 分量=多)*

午餐吃的多了吧 → *explanation(饮食名=午餐, 饮食量=多)*

这个主要考虑是反弹效应 → *explanation(症状=反弹效应)*

不是，这个考虑你喝的汤有点多 → *explanation(饮食名=汤, 饮食量=多)*

这个低了，您吃的什么药物？ → *explanation(血糖-血糖值=低)*

2.6 ask-for-sure (患者、医生)

适用于四种情况的标注：

(1) 患者询问医生并以期获取“确定”答案的问题

比如患者询问医生“能不能”的问题，能不能吃什么，能不能做什么，或者找医生确认某个行为可不可以做，期待医生的肯定或者否定回答。*Ask-for-sure* 的 arguments 包含的 slot-value 就是要问的问题本身，如果 arguments 包含多个 slot-value，表示这个多个 slot-value 联合起来是否同时成立。如果是患者已经发生的，就用 *inform* 标注已经发生的 arguments。

示例：

糖人建议吃糙米吗 → *ask-for-sure(饮食名=糙米)*

我吃莲藕排骨还喝汤，血糖升高，是不是不能喝呀？ → *ask-for-sure(饮食名=汤)*

哦，少喝还是可以的，对吧？ → *ask-for-sure(行为=少喝)*

那我早晚运动也是可以的吗？ → *ask-for-sure(行为=运动, 时间=早晚)*

好，我怎么听说糙米含糖低呀 → *ask-for-sure* (饮食名=糙米，成分=糖，分量=低)
 林医生，血糖需要每天都测吗 → *ask-for-sure* (行为=测血糖，频率=每天)
 林医生 胰岛素打腰部一圈都可以吗？ → *ask-for-sure* (药品名=胰岛素，部位=腰部)
 吃维生素 B1B12 维生素 C 会升血糖吗 →
 ask-for-sure (药品名=[维生素 B1, B2, 维生素 C], 症状=升血糖，状态=否)
 我听说老是在腹部打 脂肪容易萎缩 → *ask-for-sure* (治疗=打，部位=腰部，症状=脂肪萎缩，状态=其他)
 他的脚很干燥还起皮。是糖尿病足吗 → *inform* (症状=[脚很干燥，起皮]，状态=阳性)
ask-for-sure (疾病=糖尿病足)

一个细节：

slot-value 是作为 *ask-for-sure* 的 arguments 还是作为 *inform* 的 arguments，比如下面的例子。

请问糖尿病能吃海鲜吗？ → *ask-for-sure* (疾病=糖尿病，状态=其他，饮食名=海鲜)

请问糖尿病能吃海鲜吗？ → *inform* (疾病=糖尿病，状态=阳性)

ask-for-sure (饮食名=海鲜)

在这个例子中，倾向于认为患者已经患有糖尿病。如果是患者已经发生的，就用 *inform* 标注已经发生的 arguments。同时，对 *ask-for-sure* 来说，arguments 越少越简单，所以倾向于第二种标注方式。

(2) 医生的追问，确认是否有

用户：空腹血糖 50，餐后没量

医生：是妊娠期吗？ → *ask-for-sure* (时间=妊娠期)

(3) 患者想从多个选项中确认某一个

患者询问的问题中有多个不确定的选项，询问的目的是想确认某个选项，arguments 就是要确认的选项。这种情况通常是一个 slot 对应多个值，表示或的关系

那是早餐前还 2 小时后 → *ask-for-sure* (时间=[早餐前、两小时后])

等价形式：*ask-for-sure* (时间=早餐前，时间=两小时后)

(4) 医生想从多个选项中确认某一个

也可表示医生询问患者的目的或者行为

请问您是想要控制血糖，还是想辅助减肥？ → *ask-for-sure* (行为=[控制血糖、辅助减肥])

等价形式：*ask-for-sure* (行为=控制血糖，行为=辅助减肥)

2.7 ask-why (患者)

患者询问“为什么”的问题，通常表达对某种已发生事情的不解，希望医生给出解释。*Ask-why* 没有 arguments，和 *inform* 组合到一起，用 *inform* 的 arguments 来表示要问的主题，这些主题是已经发生了的，所以用 *inform* 表示。

示例：

为什么平常血糖正常，吃了糙米餐一 9.0 餐二 9.4

→ *inform* (血糖值=正常，时间=平常)；

inform (饮食名=糙米)；

inform (血糖值=9.0，时间=餐一)；

inform (血糖值=9.2，时间=餐二)；

ask-why()

我也不知道。可是我餐后都可以达标，餐二都可以在 7•8 以下，有时候可以在 7 以下，为什么空腹就是不能正常？

→ *inform* (时间=餐后, 血糖值=达标);
inform(时间=餐二, 血糖值=7.8 以下);
inform(时间=有时候, 血糖值=7 以下);
inform(时间=空腹, 血糖值=不正常);
ask-why()

恩恩，早上餐二 6.5 中午就 9.1，找不到原因

→ *inform*(血糖值=6.5, 时间=早餐二);
inform(血糖值=9.1, 时间=午餐二);
ask-why ()

林医生 空腹高一般什么引起的 → *inform*(血糖值=高, 时间=空腹); *ask-why* ()

2.8 ask-for (患者、医生)

患者询问的问题有明确的目标，期望的答案是具体的事物，比如要吃什么，做什么治疗，什么时间，等等。**Arguments** 就是要询问的目标信息，以及与目标信息相关的限定的信息。目标信息的值用问号表示，通常作为最后一个 **argument**。也可以用于医生追问患者的问题。

示例：

患者询问示例：

多长时间呢？ → *ask-for*(行为=频率=?)
低血糖有什么伤害？ → *ask-for* (症状=低血糖, 影响症状=?)
低血糖的数值是多少 → *ask-for* (症状=低血糖, 血糖值=?)
我想问问停药几天会反弹 → ; *ask-for* (行为=停药, 影响症状=反弹, 持续时长=?)

医生追问示例：

晚餐大概吃多少呢 → *ask-for* (饮食名=晚餐, 饮食量=?)
这个低了，您吃的什么药物？ → *ask-for*(药物名=?)
您好，发现糖尿病多长时间了呢，89 是空腹血糖吗？ → *ask-for*(持续时长=? , 身高=? 体重=?)

2.9 ask-how (患者、医生)

患者询问医生该如何做、或者泛泛的什么状态，也可表示医生问患者的行为、饮食等情况，**arguments** 表示与问题相关的信息，**slot** 是有 **value** 的，表示具体的行为、治疗、饮食是如何实施的。也可以没有 **arguments**，目的是引出医生的建议

示例：

您一般睡眠怎么样？ → *ask-how*(行为名=睡眠)
这个不错，说说您是怎样控制的 → *ask-how*(行为名=控制)
血糖控制的怎样，注射胰岛素是会长胖的 → *ask-how*(行为名=血糖控制)
饭后有点高呢，饭后怎么控制 → *ask-how*(行为名=控制, 时间=饭后)
餐后您的运动怎样 → *ask-how*(行为名=运动, 时间=餐后)

User: 那**怎么办** → *ask-how()*

Doctor: 建议你调整你晚上的用药量

2.10 assure (患者、医生)

医生表示肯定，用于标注医生对 *ask-for-sure* 的回答，没有 arguments

示例：

User: 哦，少喝还是可以的，对吧？

Doctor: 是的 → *assure()*

User : 糖人建议吃糙米吗

Doctor: 是可以的 → *assure()*

User: 林医生 胰岛素打腰部一圈都可以吗？

Doctor: 是的 → *assure()*

2.11 deny (患者、医生)

适用于三种情况的标注：

(1) 医生表示否定，用于标注医生对 *ask-for-sure* 的回答，没有 arguments

示例：

User : 吃维生素 B1B12 维生素 C 会升血糖吗

Doctor: 这个是不会的 → *deny()*

User: 不会生个傻瓜残疾之类的吗啊？

Doctor: 不会 → *deny()*

(2) 可用于对 *ask-for* 的标注，当医生对 *ask-for* 问题的回答是 none，比如“没有”、“不需要”等。

对血管没有影响 → *deny()*

(3) 也用于用户的否定表示

Doctor: 做过糖耐量检查没有？

User: 没有做，就这个没做，现在已经 39 周了，马上就生了，就是羊水有点多，让我控制血糖 → *deny()*

2.12 uncertain (患者)

用户表示不确定，通常是对医生 *ask-for-sure* 的回应，可以有 arguments，也可以没有

示例：

才发现不知道 1 型还是 2 型，这几天没测血糖，4 号早上 95

→ *uncertain (疾病=[1 型糖尿病, 2 型糖尿病])*

不知道，就是单纯血糖高 → *uncertain()*

2.13 general-advice (医生)

医生给出的泛泛的健康指导，没有 arguments

示例：

这个您能做基因检查才可以确诊的 → *general-advice()*

血糖控制的好坏和饮食、运动、日常生活都息息相关，您吃药都没下过六，要多方面分析原因
→ *general-advice()*

如果通过饮食、运动和日常生活习惯调整确实可以稳定控制血糖的情况下，也是可以的
→ *general-advice()*

不同人情况会不同的，建议在专业医生指导下，规律服药 → *general-advice()*

为了在 NLG 阶段更好地引导对话生成，需要把 domain 作为 arguments

2.14 chitchat (患者、医生)

患者和医生之间的一些寒暄，或者不知道怎么标注的，或者与健康咨询不直接相关的，都标注为 chit-chat()，没有 arguments

示例：

凡事有利就有弊 → *chitchat()*

很慢要 1 小时才高一点 → *chitchat()* -- 不知道怎么标注就标 *chitchat*

我在网上看了下 如果双方都有糖尿病遗传几率在 5% 一方有几率更小
→ *chitchat()*

User: 也只是微微出点汗 打羽毛球真不错 比走路跑步好多了 打球还有意思还能运动
时间过的又快 → *chitchat()*

Doctor: 恩，建议你坚持下去 → *chitchat()*

User: 他快 40 岁了，喝什么茶好 → *chit-chat ()*

Doctor: 节制性生活，多休息 → *general-advice ()*

User: 他就是老口渴 → *chit-chat ()*

Doctor: 一般来说，花茶、红茶、绿茶，有茶多酚，降血脂，对心血管有益
→ *general-advice ()*

2.15 general-explanation (医生)

医生给出的泛泛的解释，没有 arguments

示例：

糖尿病足两侧足动脉波动减弱，证明血流不好，然后足部容易感染，触觉、温度觉减退，感染不容易控制，导致坏疽。 → *general-explanation()*

User: 尿常规是查什么呢

Doctor: 尿蛋白尿红白细胞，酮体等等尿比重尿胆原等等。 → *general-explanation()*

为了在 NLG 阶段更好地引导对话生成，需要把 domain 作为 arguments

2.16 Act label 对应的场景

基于以上对 act label 的定义，可以对 act label 描述出用户的对话场景和医生的对话场景分别如表 7，表 8 所示。

表 7 用户的 act label 和对话场景

场景类型	Act 标签	描述
用户告知	Inform	向医生告知现状
用户发问	Ask-for	询问具体的问题
	Ask-for-sure	确认某种情况是否可行
	Ask-how	询问怎么办
	Ask-why	询问为何
用户回应	accept	接收建议
	Assure、deny	明确的肯定或者否定
	uncertain	表达不确定
闲聊	chitchat	闲聊

表 8 医生的 act label 和对话场景

场景类型	Act 标签	描述
确切回答	Advice、advice-not	建议或者不建议什么
	Assure、Deny	明确的肯定或者否定
	Explanation	给出明确的解释
对用户追问	Ask-for	追问某种情况
	Ask-for-sure	确认是否有某种情况
	Ask-how	询问某种情况如何泛泛的建议
泛泛回应	General-advice	泛泛的建议
	General-explanation	泛泛的解释
	chitchat	闲聊

2.17 Act label 汇总

Act label 汇总信息如表 9 所示。

表 9 Act label 汇总信息

Act labels	Role	Arguments	Other
inform	User	Slot=value	
Advice	Doctor	Slot=value	
Advice-not	Doctor	Slot=value	
Accept	User	/	
Explanation	Doctor	Slot=value	
Ask-for-sure	User/ Doctor	Slot=value	
Ask-why	User	/	和 inform 配合起来
Ask-for	User/ Doctor	Slot=value, Slot=?	
Ask-how	User/ Doctor	Slot=value	
Assure	User/Doctor	/	

Deny	User/Doctor	/	
Uncertain	User	Slot=value	
General-advice	Doctor	/	
chitchat	User/ Doctor	/	
General-explanation	Doctor	/	