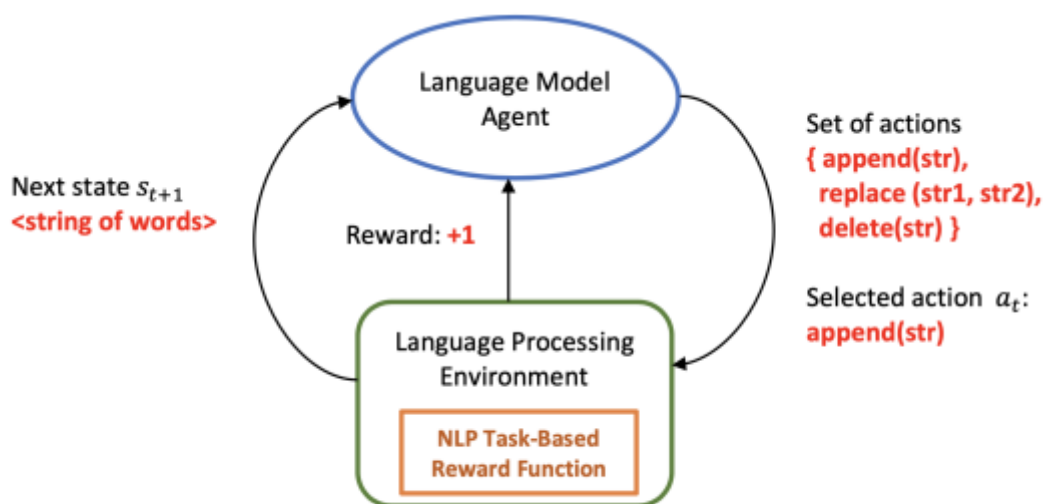


# Survey on reinforcement learning for language processing-hoho

本文阐述了强化学习与自然语言处理技术的结合点。一些自然语言处理问题可以被建模为马尔科夫决策过程，因此这些问题可以很好的用强化学习的方式去解决

## 研究方法

一个比较通用的模型如图：



应用场景：

1. 句法解析

如：譬如有一组句法生成语法：

$$\begin{aligned}
 S &\rightarrow 0A1 \\
 A &\rightarrow 0A1 \mid B \\
 B &\rightarrow \#
 \end{aligned}$$

Figure 2: Grammar  $G_1$  with 4 production rules.

任务：将实际输入字符串00#11转换为语法树

可以通过强化学习来建模流程：



- 环境状态反馈给Agent为当前语法树，
- Agent选择某条语法规则作为行为作用与环境
- 环境反馈奖励给Agent，并产生新了语法树结构

结果如下：

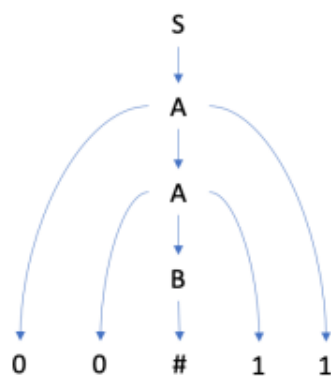
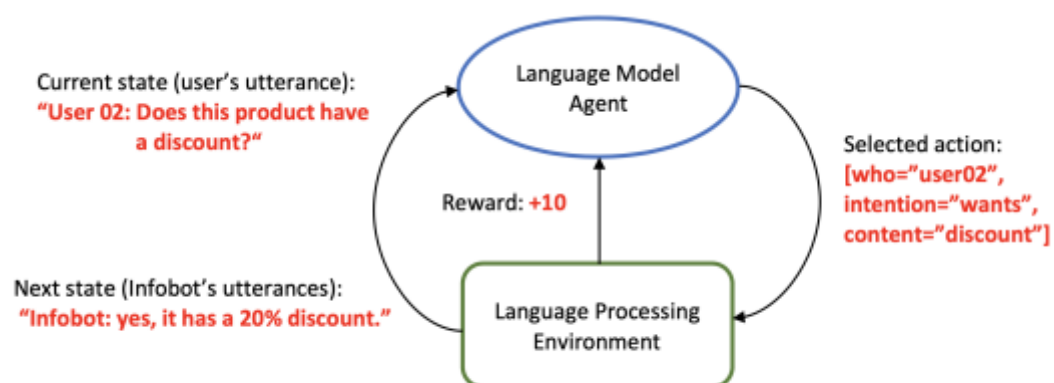


Figure 3: Parse tree of string 00#11 generated from grammar  $G_1$ .

## 2. 自然语言理解

类似：



## 3. 自然语言生成



实际例子：文本改写

- 环境输入某字符串s作为状态
- 智能体根据字典库，选择adding/deleting等动作在字符串s上添加词或删除词
- 下一个状态就是adding/deleting之后的新字符串，可见状态的转移概率是完全确定的
- 奖励函数可以定义为前后两个字符串的相似度

## 研究前景

- Recognition of the user's input
- Internal representation learning
- Exploitation of domain knowledge
- Exploitation of embodiment
- Language evolution
- Word embeddings
- Intelligent conversational systems
- Assessment of conversational systems
- Document-editing RL Assistants