第四章 动机驱动的认知对话系统建模

本文研究的长远目标是实现一个具有与人类相当的认知能力的对话系统，而这样的对话系统不仅需要第二章介绍的深层知识表示体系的支持，还需要有一个能处理情感动机以及理解用户潜在意向，并依据这些理解来做出对话决策的认知对话管理体系。因此本文深入研究符合人类自然对话的情感模型和语言语义学理论，旨在构建一个具有认知功能的可灵活扩展的动机驱动的对话管理机制。本章在第三章介绍的Psi模型的理论基础上，结合言语行为理论的指导，并借助概率逻辑网络的强大推理功能，提出了一个动机驱动的对话控制模型。下面将具体介绍该模型的各个模块的理论基础和设计方案，并指出在该对话控制模型上如何实现一个认知对话系统。

4.1 认知对话系统的概念模型

在介绍本文提出的动机驱动的认知对话控制模型之前，我们先给出本文研究的认知对话系统（以下简称“CogDial”）的概念模型和框架。目前该系统只能处理英文对话，接下来章节中的对话的例子基本按照英文的习惯举的例子，可能不太适合中文的习惯，但是该系统框架中所采用的技术和理论基础都是完全适合中文的。

简单地说，CogDial的系统流程可以描述如下，此系统将接收到的人类话语转化成基于超图的抽象逻辑表达形式，概率逻辑网PLN会将这这些逻辑表达形式与系统的动机和情感状态，以及系统的知识库等结合起来，进行一定的解析和逻辑推理，从而得到想要回应的表达内容的逻辑表达形式，最后将这些逻辑表达式再转化成自然语言回应给谈话对象。概括地说，CogDial主要采用了以下几项技术：

* 基于动机驱动的情感模型Psi及其相关机制的宏观规划（Macroplanning）。该模块以言语行为理论为指导，主要用于以下两方面内容：1）在哪个时间点发生哪些言语行为，2）使用哪些内容去封装这些言语行为。宏观规划更多的是从语用和推理方面出发，淡化语义或语法（或者音韵词汇）等的限制。
* 基于概率逻辑网络的概率逻辑推理技术，用于实现Psi上的动机和行为的自适应规范化
* 将自然语言语句通过句法分析和语义逻辑关系抽取转换成基于超图的抽象逻辑形式表示的自然语言理解技术（这将在本文的第五章中介绍）。
* 将经过宏观规划得到的以超图形式表示的要表达的抽象内容进行微观规划（Microplanning）和表层生成技术最终转换成自然语言语句（这将在本文的第六章中介绍）。

• 在实现模式上，我们将上述两模块都封装在“言语行为规划器（Speech Act Schema）”中，以 Context & Procedure → Goal的蕴涵式形式表示，这些言语行为规划器将用于：

– 在动机驱动的情感模型控制机制下，结合当前的语境，激活相应的言语行为规划器

* 收集相关内容，生成能与微观规划衔接的Atoms集合

– 调用一系列的认知机制（主要包括用于概率逻辑推理机制PLN，和注意力分配机制ECAN [96]），选择相关的Atoms发送到微观规划器中用于表层生成自然语言语句。

CogDial的初步预期目标主要是实现面向游戏角色和人形机器人的对话系统，但同样的研究思路也完全能应用在基于文本的对话系统，例如智能对话搜索接口。我们分阶段来实现这样的系统，以一个相对小而简单的系统为起点，通过不断改进和完善，以及结合系统本身的自学习和认知能力，最终实现一个在一定程度上接近人类水平的系统。目前，我们的系统还没有达到最终的水平，但基本框架已经到位，并用在汉森机器人公司的类人机器人Sophia[[1]](#footnote-0)上 。

4.2 目标驱动的对话控制模型

本节我们着重讨论本文提出的动机驱动的认知对话控制模型。该模型

(对话系统框架图放这里)

In our cognitive approach to dialogue, dialogue control occurs via the combination of multiple cognitive processes on a shared set of knowledge representations. The dynamics of dialogue control is guided by a motivational system, based on Dorner's Psi model, that chooses actions and allocates system resources according to a set of goals defined over multiple time scales.

4.2.1 针对对话控制的Psi的配置

Psi模型给我们提供了动机驱动的情感模型建模的理论指导，鉴于该模型面向心理学领域，要将其应用到对话控制领域，除了本文第3.3节中介绍的将Psi模型表示成CogDial所采用的知识表示库Atomspace中使其能作用于知识推理和学习之外，还需要将Psi模型中的动机和调节因子等参数配置成更符合认知对话系统的需求形式。本小节就具体介绍CogDial如何配置Psi中的各项参数。

基于Psi的基本框架，我们选择了以下的动机需求用来指导我们的认知对话系统中的行为选择：

* 社交需求（Affiliation）：与他人互动，希望被其他成员接纳的需求。取悦谈话对象可以视为此目标的一个特例
* 能力需求（Competence）：通过对话达到某种目标的需求，衡量言语表达方式的指标
* 确定性需求（Certainty）：智能体对自身语境的了解需求，特别是对目标的了解需求
* 新颖性需求（Novelty）：维持智能的对话而不是简单重复的问答多好。

CogDial中的对话控制主要通过不同的需求来选择相应的言语行为规划器，因此，在选择言语行为规划器前，我们需要知道当前状态下上述每种需求的被满足的程度。鉴于目前的CogDial系统的有限能力，为了能更好地实现一个面向集成认知体系的智能会话系统框架，我们将上述的几项需求做了更具体化的定义：

\begin{itemize}

\item 社交需求（Affiliation）：我们对该需求进行了下面三种满足程度：

\begin{itemize}

\item 当对话系统正在与人或者其他智能体进行会话时，该需求在一定程度上被满足；

\item 当系统正与多个人或智能体进行会话时，该需求被满足的程度提升；

\item 当该系统的会话内容都属于积极向上的时候（通过情感分析技术实现），该需求被满足的程度达到最高。

\end{itemize}

\item 能力需求（Competence）：此项需求需要通过OpenCog来评估。简单来说，对每一个目标，OpenCog记录着会话系统在过去完成该项目标的程度，然后根据当前的不同目标所占的权重，我们可以通过以下公式估算该需求被满足的程度（首先针对每一个目标，计算目标权重\*能达到该目标的概率，然后求总）。当然计算目标被完成的程度，还应该考虑实现目标的语境等因素，目前我们的系统更注重构建一个智能会话系统框架，因此在语境无关的假设下来衡量目标被实现的程度。

\item 确定性需求（Certainty）：如果会话系统正在和一个陌生人对话，或者系统无法理解大量被提及的单词或概念，那么当前的确定性需求被满足的程度就会降低。如果系统获取了新的可靠知识，那么该需求被满足的程度就会增加。

\item 新颖性需求（Novelty）：我们定义了以下几种方式来增加智能绘画系统对新鲜度需求的满足程度值：

\begin{itemize}

\item 多和不同的人类或智能体会话；

\item 谈论新的话题或引入新的词汇和概念；

\item 学习新的可靠信息（ OpenCog推理得出的新的置信度高的知识存储的载体原子（Atoms））

\end{itemize}

\end{itemize}

基于上述目标需求的智能会话系统，除了需要结合前文所述的OpenPsi的框架理论，以及本文描述的自然语言理解和生成的相关工具之外，还需要有以下模块：

\begin{itemize}

\item 制定一组能被特定目标需求激活的对话控制程序

\item 建立目标和相应去实现目标的行为之间的关联，可通过相关规则来实现，也可以通过强化学习来实现，我们系统框架采用两种结合的方法，但目前的系统还是以规则关联为主。

\end{itemize}

1. http://www.hansonrobotics.com/robot/sophia/ [↑](#footnote-ref-0)