# 题目：End To End dialogue systems

**姓名：叶志豪**

**学号：3115006392**

**手机：15626124528**

**导师：蔡瑞初**

**研究方向：NLP，chatbot，dialogue systems**

**摘要**

要求简明扼要，阐述清楚研究的意义，研究的主要内容，及创新性（300-400字）：

人机对话系统（chatbot）是将机器视为一个认知主体的人机双向信息交换系统，是实现人机交互的一种方式。根据人机对话系统的功能不同，可以分为两个大类：开放域的对话系统（Open Domain Dialog System）和领域任务型的对话系统(Task Oriented Dialog System)。开放域对话系统不针对某一特定领域，目的不是完成某项特定的任务，而是满足用户的聊天或者开放百科问答等需求，其代表性产品就是闲聊机器人，比如微软小冰；领域任务型对话系统则是针对具体的应用领域，比如餐厅预订、公交线路查询、设备控制等，以完成一项具体的领域任务为目标，代表性产品有苹果的Siri和微软的Cortana。

在开放域对话系统方面，有基于检索和生成式的两种方式生成对话系统，而在开放领域中，多轮对话中如何记忆上文中的逻辑和信息，是一个比较大的问题。另一方面，在任务型的对话系统中，目前主要是分模块的解决方式有一定的效果,但是，这样的方式需要分开来训练各个模块，同时需要大量的特定领域的人工干预（domain-specific handcrafting）,在模型泛化方面有很大的局限，在进行任务对话的时候，缺乏鲁棒性。同时随着技术的发展，如何构建个性化，更加吸引用户的对话系统也是一个较大的问题。随着深度学习的发展，end to end 模型在很多方面有所成就，如机器翻译，文本挖掘和图像识别等。但在多轮对话系统（Multi-turn dialogue systems），个性化对话系统（Personalized dialogue systems）方面，还有很大的提升空间，结合上面的问题，和我本人的研究，本报告将会主要讲述end to end模型在开放域和领域任务型的多轮对话历史建模和和个性化建模的应用和研究。

**报告正文**

参照以下提纲撰写，要求内容翔实、清晰，层次分明，标题突出。**请勿删除或改动下述提纲标题及括号中的文字。**（4000-8000字）：

1. **项目的立项依据**（研究意义、国内外研究现状及发展动态分析，需结合科学研究发展趋势来论述科学意义；或结合国民经济和社会发展中迫切需要解决的关键科技问题来论述其应用前景。附主要参考文献目录）；

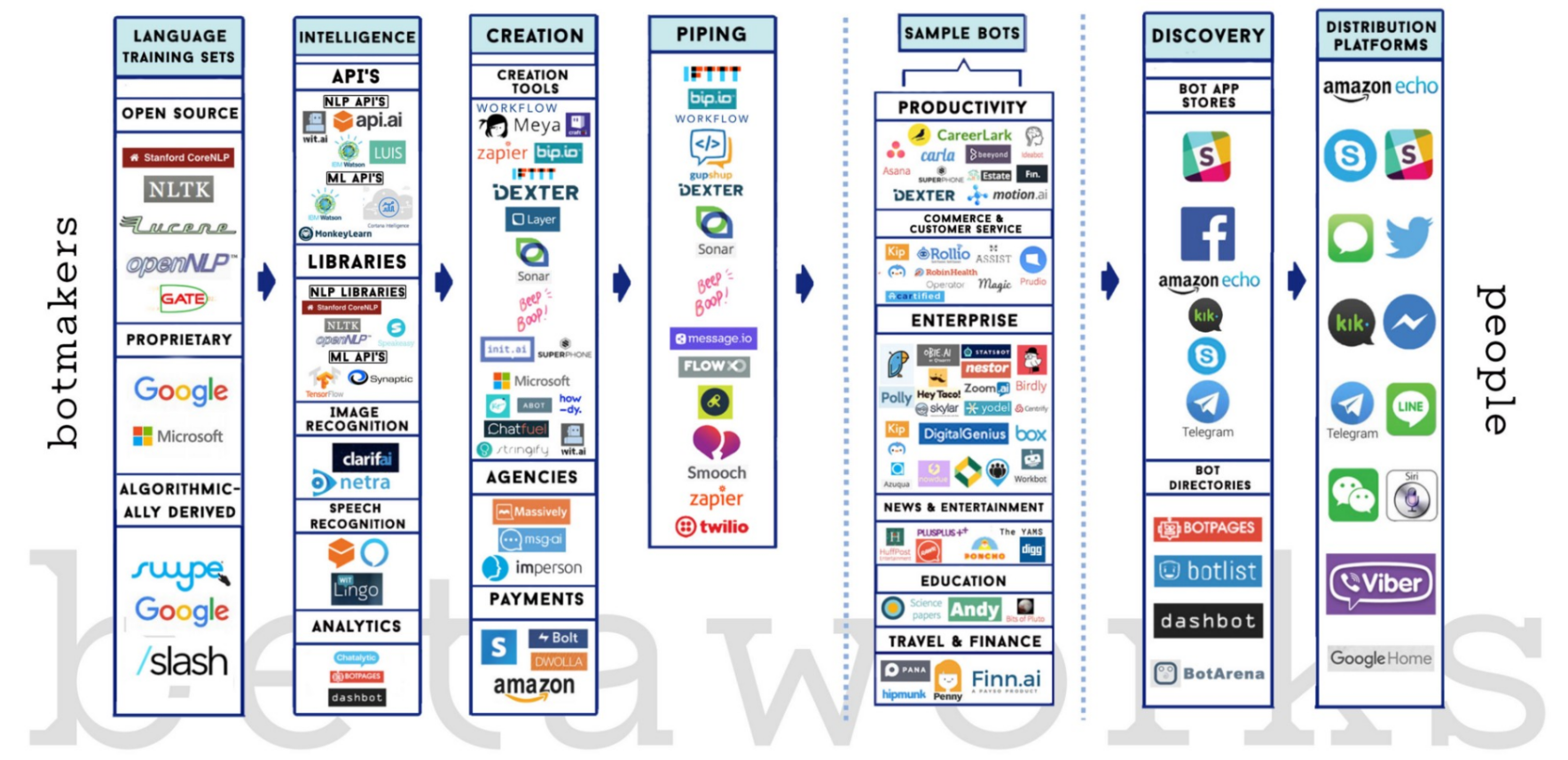
1.1 研究意义

最早的对话系统仅仅处理简单的交流任务，比如提供飞行旅行信 （Hempel, 2008）。当今的对话系统已经dialogue system越来越多的出现在更复杂的任务中，比如，智能环境（Intelligent Environments, Heinroth & Minker, 2013）, 车载系统（Geutner etal., 2002），个人助手（Janarthanam et al., 2013）， 智能家居（Krebber et al.,2004），及人机交互（Foster et al., 2014）等。对话系统涉及的另外一个重要课题是用户建模（Andrade et al., 2014），这个也直接影响到每一个用户的个人体验。 在工业界，聊天机器人已经开始形成规模，各个领域的数百家公司正在开发它们。聊天机器人可应用的领域包括：在线客服，私人助理，娱乐工具等等，也就是说，研究对话系统（dialogue system）的重要性是不言而喻的。现在的对话系统（dialogue system）中很大的问题在于上下文不连贯，毫无逻辑性，和不能很好的个性化和更加的有趣，也只有解决这些问题，dialogue system的能力才能更加的提升，更好的为人类服务。

1.2 国内外研究现状及发展动态分析

语言是人类的特殊能力，（对话系统）dialogue systems是 AI 领域的一个高难度题目，因为它涉及多个语言相关的子领域，如语音识别、语音合成、语言理解、语义表示、对话管理、语言生产、情感建模、以及多模交互。Grand View Research 公司做的一项研究调查报告显示2012年对话系统的市场市值大约为 3.5 亿美金，该公司预测到 2020 年会增长 31.7%。同时该项报告还指出，考虑到市场以及经济效应，人工客服领域将是未来 （对话系统）dialogue systems的一个大方向。

下面是chatbot市场一览图，也充分的说明了在工业界的chabot的发展情况。



从上图可以看出，chatbot（dialogue systems）在人们生活中可以扮演非常重要的角色，而且越来越深入人们的生活，各大公司都在打造自己的bot，为了让用户体验更好，个性化的服务是关键的一环，这也说明研究个性化chatbot（dialogue systems）的意义，另外上图也说明了chatbot（dialogue systems）的趋势是越来越好的，也是更有益于人类的。

参考文献

李林琳 & 赵世奇. 对话系统任务综述与基于POMDP的对话系统. PaperWeekly.

Hempel, T. (2008).Usability of speech dialogue systems: Listening to the target audience.Springer.

Heinroth, T., &Minker, W. (2013). Introducing spoken dialogue systems into IntelligentEnvironments. New York: Springer.http://dx.doi.org/10.1007/978-1-4614-5383-3

Geutner, P., Steffens, F.,& Manstetten, D. (2002). Design of the VICO spoken dialogue system:Evaluation of user expectations by Wizard-of-Oz experiments. Proceedings of the3rd International Conference on Language Resources and Evaluation (LREC‘02),Canary Islands.

Janarthanam, S., Lemon,O., Liu, X., Bartie, P., Mackaness, W., & Dalmas, T. (2013). Amultithreaded conversational interface for pedestrian navigation and questionanswering. Proceedings of the 14th Annual Meeting of the Special Interest Groupon Discourse and Dialogue (SIGDIAL), 151–153.

Krebber, J. Möller, S.,Pegam, R., Jekosch, U., Melichar, M., & Rajman, M. (2004). Wizard-of-Oztests for a dialog system in smart homes. Paper presented at the Joint CongressCFA/DAGA ’04, Strasbourg.

Foster, M. E., Giuliani,M., & Isard, A. (2014). Task-based evaluation of context-sensitivereferring expressions in human-robot dialogue.Language, Cognition andNeuroscience, 29(8), 1018–1034. <http://dx.doi.org/10.1080/01690965.2013.855802>

Andrade, A. O., Pereira,A. A., Walter, S., Almeida, R., Loureiro, R., Compagna, D., & Kyberd, P. J.(2014). Bridging the gap between robotic technology and health care. BiomedicalSignal Processing and Control, 10,65–78.http://dx.doi.org/10.1016/j.bspc.2013.12.009

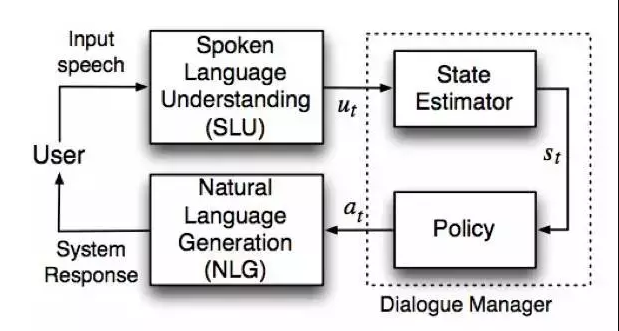
Allen, J. (1995).Naturallanguage understanding. Redwood City, CA: The Benjamin Cummings.

2．**项目的研究内容、研究目标，以及拟解决的关键科学问题**（此部分为重点阐述内容）；

2.1 研究内容

由对话系统可以分为两个大类，那么下面将分为领域任务型和开放域两大类来报告研究内容：

2.1.1领域任务型的对话系统(Task Oriented Dialog System)：

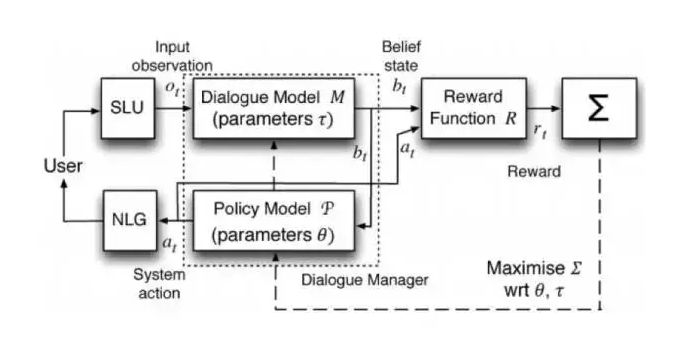
对于一个领域任务型（task oriented） 的对话系统，可以分为以下三个部分，也就是SLU，NLG，Dialogue Manager，口语理解模块（SLU）将语言转换成抽象语义表示，即用户对话行为，而后系统更新其内部状态，然后系统通过决策规则确定系统行为，最后语言生成模块（NLG）将系统行为转化为自然语言字符串。其中，状态变量 包含跟踪对话过程的变量，以及表示用户需求的属性值（又称为slots）。

其中每个部分部分都有不同的方法和模型才解决。举个例子，比如：SLU系统涉及从自然语言询问中识别说话者的意图和提取语义构成，这两项任务常被称为意图识别和槽位填充。意图识别和槽位填充通常是分别进行的。意图推测可以理解为语义表达分类的问题，比较流行的分类方法有例如支持向量机（SVMs）和深度神经网络。槽位填充可以理解为序列标注任务。解决序列标注问题流行的方法包括最大熵马尔可夫模型（MEMMs），条件随机场(CRFs)以及循环神经网络(RNNs)。由于在此主要是讲end to end 的模型，所以不详细报告各个部分分开模型方法的解决方式。

领域任务型的对话系统(Task Oriented Dialog System)在最近几年的end to end的模型效果慢慢提高，且非end to end 的模型的需要domain-specific handcrafting，及其迁移和鲁棒性的缺陷。现在将结合效果比较好的end to end模型，结合其优缺点，来报告研究内容。

基于Memory Network在QA chatbot领域充分展示了优势，而Bordes and Weston (2016)（MemN2N based dialog system）将其应用在了领域任务型的对话系统(Task Oriented Dialog System)，应用Memory Network的优势在于不用独立涉及 SLU,DST,DPL，也就是直接的端到端的一种训练，在其缺点在于很难结合知识库。同时在该论文中的实验是基于模板检索的，也就是存在迁移差的问题。而Wen et al. (2016b,a) 提出的End-to-end training of modular dialogue 在system中能较好的结合知识库，也就是在项目只需要较少的人为干预即可，在其缺点在于DST需要独立的训练，虽说是端到端的训练（end to end via gradient descent。）但其实还是有不能直接的定义为end to end的模型。此外Williams and Zweig (2016)提出相近的模型，可以同时用SL和RL的训练方式训练，领域专家需要提供少量对话样本（第一篇中21个对话样本，对话平均长度7个turns），可以自动学习对话状态（ dialogue states）。但是缺点同样是需要较多的人为的干预，比如需要人为的制定动作掩码（action mask）等。以上的end to end模型，各有优缺点，有许多上升的空间，本项目主要将在怎么减少人为干预而达到较好效果，和怎么提高模型泛化能力。

其次，在领域任务型的对话系统(Task Oriented Dialog System)的关键的DM（dialogue management）模块，有不同的实现方式，值得一提的是基于POMDP的形式，新型的对话管理系统基于部分可观察的马尔可夫决策过程（POMDPs）（Williams & Young, 2007），该方法假定对话过程是马尔可夫决策过程，也就是说，对话初始状态是, 每一个后续状态用转移概率来表示2017-06-30 13-54-17屏幕截图。 状态变量2017-06-30 13-56-03屏幕截图 是无法直接观察到的，它代表了对用户需求理解的不确定程度。系统把SLU的输出看作是一个带噪音的基于用户输入的观察值，这个观察值的概率为2017-06-30 13-54-29屏幕截图，这里的转移概率和生成概率用恰当的随机统计模型表示，又称为对话模型 M，而每个步骤中采取哪个行动则由另一个随机模型控制，该模型称之为对话策略 P。在对话过程中，每一步还需要一个回报函数来体现理想中的对话系统特性。对话模型 M 和对话策略 P 的优化是通过最大化回报函数的期望来实现的，该过程可以通过直接用户交互在线训练，也可以利用离线的语料库训练。



基于 POMDP 的对话系统融合了两个核心观点：置信状态跟踪和增强学习。这两个方面可以放在同一个框架下学习。与传统方式相比，该方法具有如下优点：1.置信状态为语音识别噪声提供了更好的鲁棒性（Williams & Yong, 2007）2. 通过保存各个状态的置信分布，系统可以并行的追踪各种对话路径，它不是贪婪的选择当前最优解而是综合考虑各种状态的全局解。3. 显式的表达状态和行为能将回报函数和状态行为对关联起来。但另一方面在实践中运用 POMDP 并不容易，有许多实际问题需要解决。对话系统中的状态行为空间巨大，求解这个空间需要复杂的算法和软件。实时的贝叶斯推理也非常难，完整的 POMDP 的学习策略是不可实现的，因此必须利用近似法求解。优化基于 POMDP 的 对话系统的最直接方式是通过直接用户对话。但是，通常难以找到足够数量的用户帮助训练系统，所以实践中常常通过用户模仿器的方式来对参数模型进行优化。

然后就是个性化的问题，Kaixiang Mo等人（Mo K et al，,2016）提出一种迁移学习基于POMDP的模型，将群体作为源领域，而个体作为目标领域，在解决POMDP的问题中，提出由两部分构成的Q-function，一部分为通用的Q-function，另一部分为个性化的Q-function，这样来避免负迁移的产生。其优点在于较好的将整体的知识个性化到个体上。

最后顺便一提的是由于现在的task oriented dialogue system的评测标准没有很好的Intrinsic（内部） metric，现在的Intrinsic（内部） metric评测标准主要是success rate, average reward, average turns，同时在数据集方面，对task oriented dialogue system的数据较少，Li, Xuijun等人（Li, Xiujun, et al,2016）提出**User Simulator在一定程度上解决了这个问题。在这个方面还需要研究，包括评测标准（**[evaluate](http://dict.youdao.com/w/evaluate/" \l "keyfrom=E2Ctranslation)metric**）**和数据集。

2.1.2开放域的对话系统（Open Domain Dialog System）

对于开放域的对话系统，分为检索式和生成两种方式，一种是基于生成式，即机器人“自己说话”，一个字一个字创作出回复语句来。另外一种是基于检索式，即机器人“转发”别人的话。从互联网大家的话语中寻找到合适的回答予以转发。在对话达到较多的轮数是，就很容易的出现前后不连贯，上下难衔接，因此在进行检索的时候考虑历史信息则显得尤为重要。那么如何让机器人理解对话历史信息从而聪明地进行回复呢？微软亚洲研究院(Wu Y, Wu W, Zhou M, et al,2016)提出了一种SMN模型是基于检索式的对历史对话建模较为好的模型，但是其在上下连贯性，跳转，顺承还是有待提高。这也是多轮对话在开放域的对话系统（Open Domain Dialog System）所需要解决的。

最后，对于对话系统（Dialog System），现在更加注重的个性化和怎么才能更加的吸引用户。这个这是一个比较有趣的研究的点，Xing C等人（Xing C，et al.2017）提出的模型能生成与主题相关相关的对话，而清华大学计算机系朱小燕、黄民烈老师团队研发的 ECM的模型（Zhou H et al,2017）能生成有感情的对话都是属于这类对话系统。

2.2 研究目标

首先是在领域任务型的对话系统(Task Oriented Dialog System)方面，实现其生成更加的完善同时需要更少的人工干预。同时提高其迁移性，减少领域专家知识，比如基于人为设计好的特定领域应答模板。RL在这一方面有比较大的优势，下面将会试着解决怎么讲RL算法很好的结合 对话系统（dialogue system）。

然后是个性化迁移方面，可以找到更加具有创新性的模型，提高个性化 对话系统（dialogue system）的能力。

最后是开放域的对话系统（Open Domain Dialog System）方面，同样的，也是试着解决对历史状态的建模，提出较好的方式解决对话不连贯等问题

总之，尽自己最大的努力，在对话系统（dialogue system）方面研究可以有所收获。

2.3 拟解决的关键科学问题

对于开放域的对话系统，解决怎么更好的对历史对话建模，使其对话更加的连贯。

对于领域任务型的对话系统(Task Oriented Dialog System)：端到端的对话系统模型（End-to-end dialogue systems model） 现在还不能实现在最少的专家知识( human knowledge)，现在还是需要加入比如response模板，Entity Input，Entity Output等的专家知识。

另外，现在的对话系统（dialogue systems）很多并没有实现个性化的设置，在之前的的文章（Mo K et al，,2016）有用迁移学习（transfer learning）的方式结合的个性话，但是效果从其实验来看还是很有待提高。

最后，在领域迁移方面的缺陷也是一个大问题，比如在飞机订票的自动客服如何很好的迁移到其他领域，比如迁移到订火车票的自动客服，这个也是有待于解决的问题。

3．**拟采取的研究方案及可行性分析**（包括研究方法、技术路线、实验手段、关键技术等说明）；

3.1 研究方案

研究方法：首先在研究这个项目时候，首先是看论文，每周坚持看相关方向的论文，并做论文的总结，查看那些方法的可行性和自己是否可以有其他方式的创新，若有些必要复现的论文则进行复现，并比较结果。然后，定时写总结并和其他同学讨论，可以是工作室的同学，也可以在网上的平台比如paperweekly进行讨论。

在技术路线第一要把深度学习的基本模型包括CNN，RNN，GAN和RL的有关模型方法都掌握，当然这一步我基本完成。第二是结合dialogue应该学习当下比较流行的模型，比如seq2seq框架，比如attention在领域任务型的对话系统(Task Oriented Dialog System)的应用，其中在当下有很多基于seq2seq的开放域的对话系统（Open Domain Dialog System）的论文，这些普遍存在chatbot回复单一，上下文不接的问题。

就是GAN in chatbot的应用，也就是GAN 在离散序列上直接的传梯度是无效上的（尽管可以用连续的Word2vec，但是词典的离散的，也就是微调所谓连续的Word2vec，但还是有可能得到一个词典不存在的词，如果强行近似就会很容易到达一个局部最优而导致的效果不好），第一个用GAN做了open domain dialogue system，主要的灵感就是GAN就像图灵测试般，只有人类和机器生成的对话无法区分的时候，才是最好的时候。

在RL算法或者是VAE也有在chatbot众多的应用，这个有待学习和发掘！

当然，我会更加的倾向于GAN和RL在chatbot的应用。首先是GAN 在nlp中有越来越多的应用（Yu, Lantao, et al.2017，Li, Jiwei, et al. 2017，Kusner,&Matt J.2016），同时GAN和RL中的actor-critic methods有很多的相似之处（Pfau, D., & Vinyals, O. 2016），那么若GAN可以用于open domain的chatbot上，同时actor-critic也可用于生文本的应用（Bahdanau, Dzmitry, et al. 2016），并解决了偏差爆炸（**exposure bias）的问题**，受上面两者的启发，是否可以之间将actor-critic结合seq2seq框架应用于chatbot呢？这个有待尝试。

最后是个性化对话系统构建的技术方案，现在有文章写道讲迁移学习应用于个性化的对话系统中，这具有很大的启发性，之后会着重对于迁移学习的学习，怎么提高正迁移和避免负迁移，这些都是下来要做的工作。

3.2 可行性分析

在领域任务型的对话系统(Task Oriented Dialog System)方面，前面分析的几篇文章都给出了不错的试验结果，比起用人为的规则式，end to end模型生成式模型的效果也可见一斑（Li, Xuijun, et al,2017），但是比起分模块的涉及方式，end to end模型还有很大的提升空间，但随着深度学习模型的成熟哦，也就更加肯定的可以认为深度end to end模型的可行性，然后在实验手段硬件方面，首先结合自己1080的GPU，当然，我也非常希望可以正真的**进入DMIR实验室，得到蔡老师的指导**，这样可以和实验室的同学更好交流，并可以使用实验室的服务器，可以加快实验的进程。在框架方面现在大部分都是用tensorflow，但是随着pytorch的优势性，也就是动态图调试，若有需要的时候也会应用。

4．**本项目的特色与创新之处**；

首先本项目的特色在于应用性强，不管是在工业界还是在学术界，对话系统（dialogue system）都是有极大的应用性。不像其他如机器阅读，GAN模型在图像上的研究，chatbot同时现在各大公司都有各种聊天系统或者自动客服，需求量很大，更加接近人们的生活，这更加说明了研究chatbot的意义。然后是创新之处，本项目将集中于结合RL算法和GAN，VAE在的应用，对于GAN，VAE它们在CV领域已经有许多的应用，但是在nlp中的应用还有待提高，将他们结合在nlp的具有一定的创新意义，另外RL在chatbot的应用也有其优缺点，而怎么将其优点，也就是生成更加“有营养”的对话，或者生成对话更加的连贯放大，同时避免其难以训练和高方差的缺点，也是一个值得研究的内容。最后是个性化的dialogue system现在还有很大的提升空间和研究空间，本项目也注重于个性化服务的对话系统的构建，这也是具有创新的地方。

参考文献

Antoine Bordes and Jason Weston. Learning end-to-end goal-oriented dialog. arXiv preprint arXiv:1605.07683, 2016.

Tsung-HsienWen, Milica Gasic, Nikola Mrksic, Lina M Rojas-Barahona,

Pei-Hao Su, Stefan Ultes, David Vandyke, and Steve Young. A network-

based end-to-end trainable task-oriented diaEnd To End task Completion neural dialogue systemslogue system. arXivpreprint arXiv:1604.04562, 2016.

Jason D Williams and Geoffrey Zweig. End-to-end lstm-based dialog

control optimized with supervised and reinforcement learning. arXiv

preprint arXiv:1606.01269, 2016.

Williams J D, Young S. Partially observable Markov decision processes for spoken dialog systems[J]. Computer Speech & Language, 2007, 21(2): 393-422.

Wu Y, Wu W, Zhou M, et al. Sequential Match Network: A New Architecture for chatbotMulti-turn Response Selection in Retrieval-based Chatbots[J]. arXiv preprint arXiv:1612.01627, 2016.

Zhou H, Huang M, Zhang T, et al. Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory[J]. arXiv preprint arXiv:1704.01074, 2017.

Xing C, Wu W, Wu Y, et al. Topic Aware Neural Response Generation[C]//AAAI. 2017: 3351-3357.

Mo K, Li S, Zhang Y, et al. Personalizing a Dialogue System with Transfer Learning[J]. arXiv preprint arXiv:1610.02891, 2016.

Yu, Lantao, et al. "SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient." *AAAI*. 2017.

Li, Jiwei, et al. "Adversarial learning for neural dialogue generation." *arXiv preprint arXiv:1701.06547* (2017).

Kusner, Matt J., and José Miguel Hernández-Lobato. "GANS for Sequences of Discrete Elements with the Gumbel-softmax Distribution." *arXiv preprint arXiv:1611.04051* (2016).

Pfau, D., & Vinyals, O. (2016). Connecting generative adversarial networks and actor-critic methods. *arXiv preprint arXiv:1610.01945*.

Li, Xuijun, et al. "End-to-end task-completion neural dialogue systems." *arXiv preprint arXiv:1703.01008* (2017).

Li, Xiujun, et al. "A user simulator for task-completion dialogues." *arXiv preprint arXiv:1612.05688* (2016).

Bahdanau, Dzmitry, et al. "An actor-critic algorithm for sequence prediction." *arXiv preprint arXiv:1607.07086* (2016).