微信智言

演讲者 牛成

内容页

- 腾讯小微
- 对话系统
- 案例介绍
- 期待与展望

腾讯小微

并不是一个单一的对话模型，而是专门为合作伙伴设计的，以对话为核心能力的，易于接入的开放式平台

腾讯云小微

连接人与应用

源自微信 & 腾讯云

直接提供对多个高频对话场景的完整支持

用户随心自定义场景，实时生效

精准的语义理解能力为所有场景服务

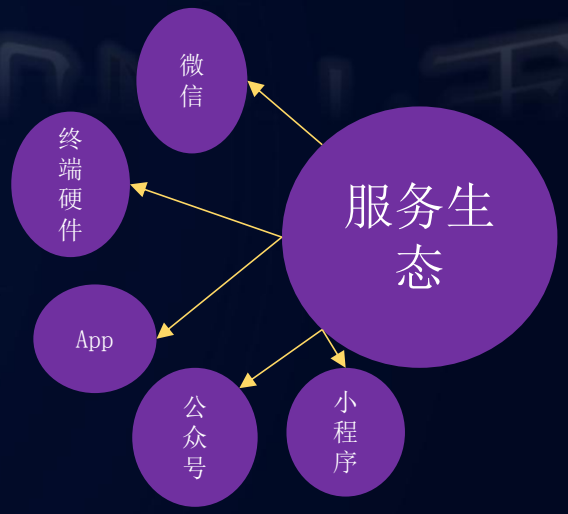
来自腾讯生态伙伴的深度支持

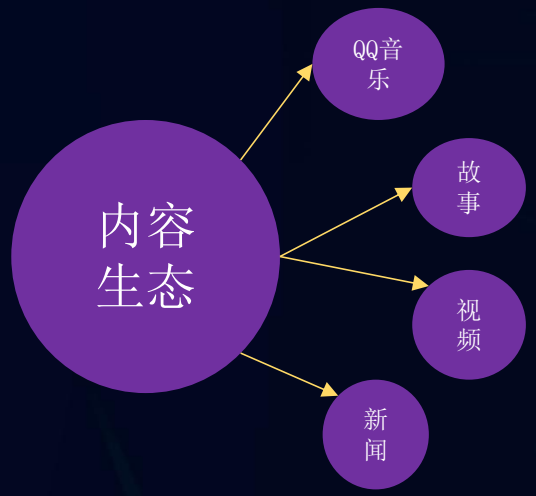
腾讯云小微

腾讯云小微

小微

无微不至



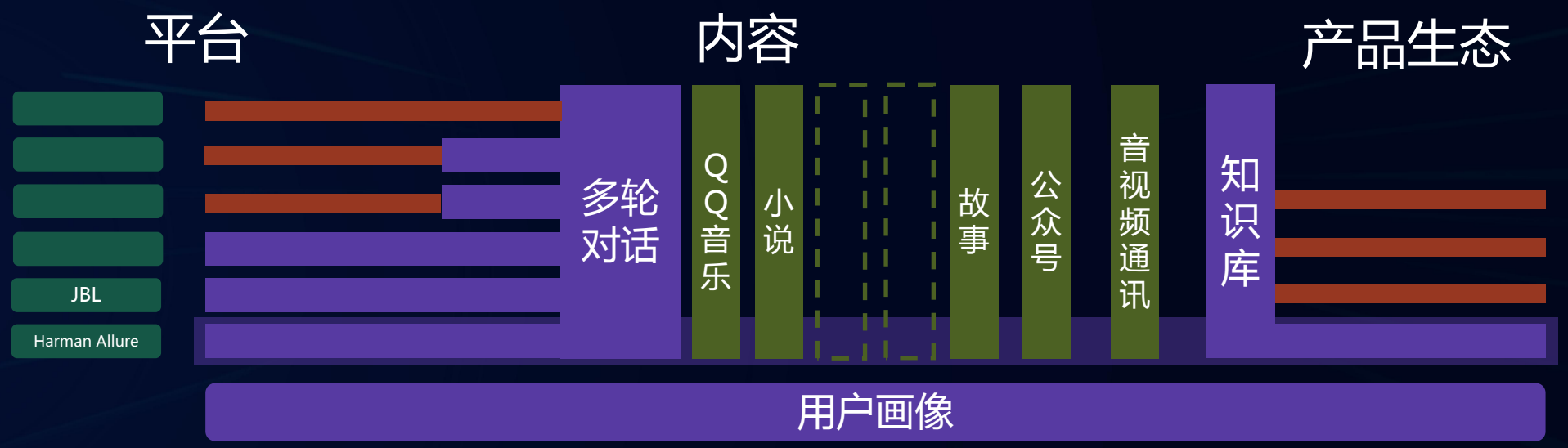


腾讯小微

对话系统由来已久，可以分为3个阶段。最早从1966年就有人开始尝试基于规则的对话系统。到了2010年左右，数据驱动系统兴起，包括信息检索技术和机器学习模型，尤其是机器翻译模型开始被用在对话系统上。近年，人们开始尝试，用对话系统真正去解决现实中的问题，即使还只是一小部分。

不只是一个对话系统，而是一个四位一体的平台组

不只是一个对话系统，而是一个四位一体的平台组



腾讯小微

小微

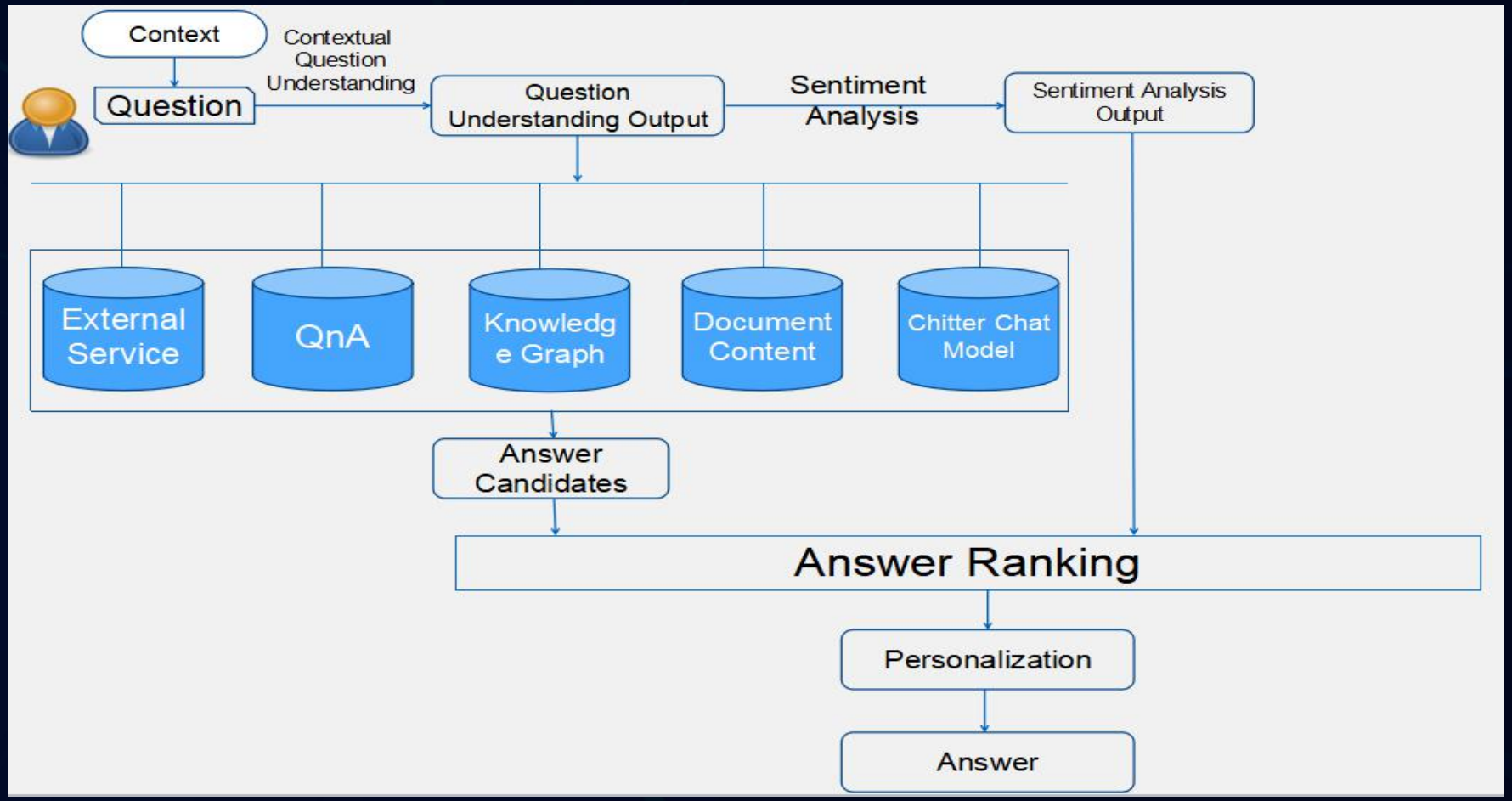
语义理解
系统

精准的单意图识别和多意图处理

对话流程中的关键词(槽位)处理

多轮对话管理（系统自带与用户自定义）

用户问题处理流程


Context
Contextual Question Understanding
Question
Question Understanding Output
Sentiment Analysis
Sentiment Analysis Output
External Service
QnA
Knowledge Graph
Document Content
Chitter Chat Model
Answer Candidates
Answer Ranking
Personalization
Answer

意图分类

意图反映用户需求，在对话系统中，每一个意图都与**现实世界中的某个需求**点对应。

用户实际需求

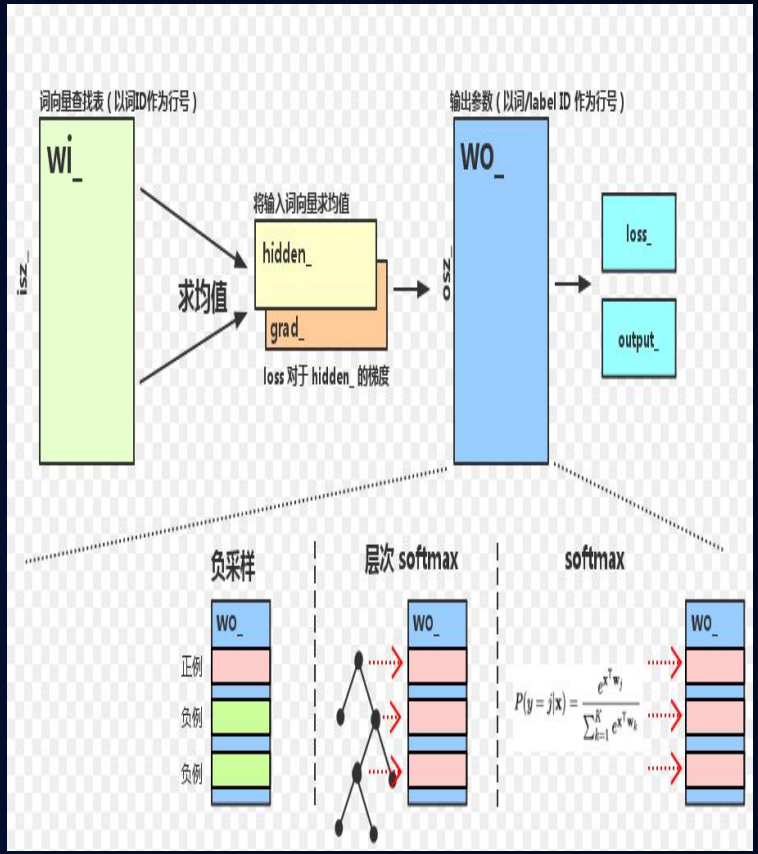
明天天气怎么样
下午三点打球
腾讯股票价格
放首歌听听
停止播放
下一首

多级意图

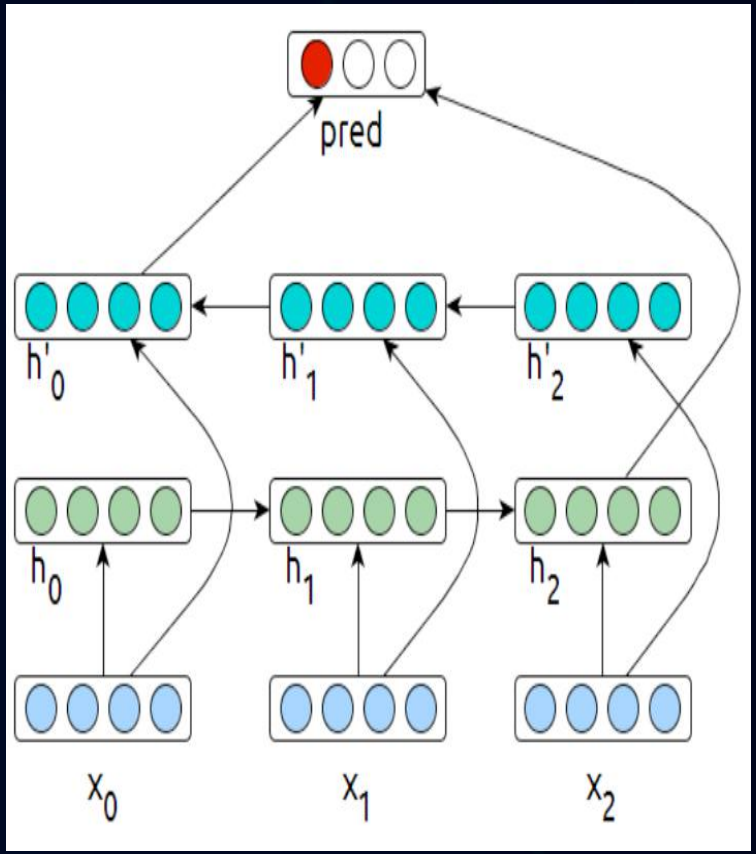
天气
提醒
股票
音乐
停止
下一个
指令

系统模型结构

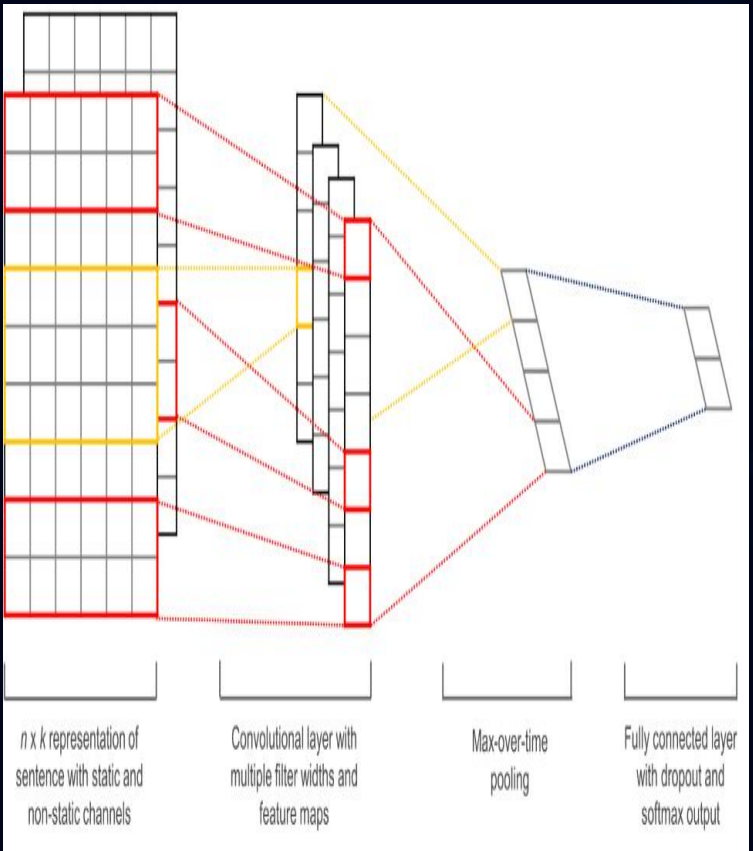
Typical Models for Intent Classification



Fasttext



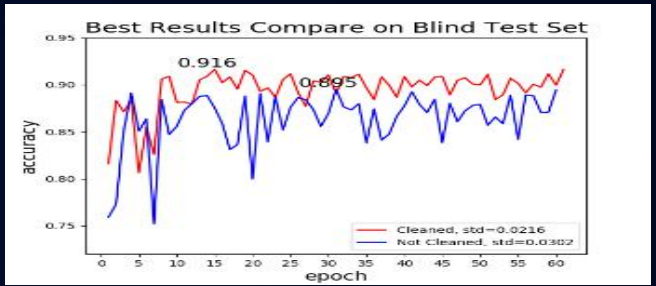
RNN for text classification

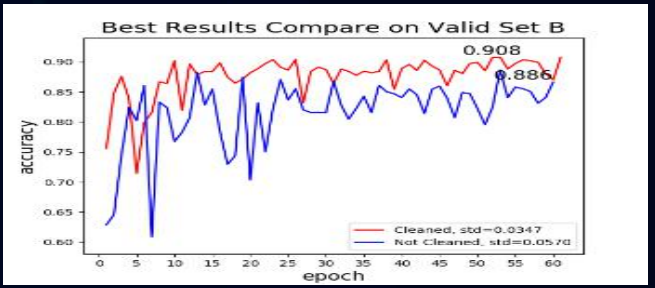


CNN for text classification

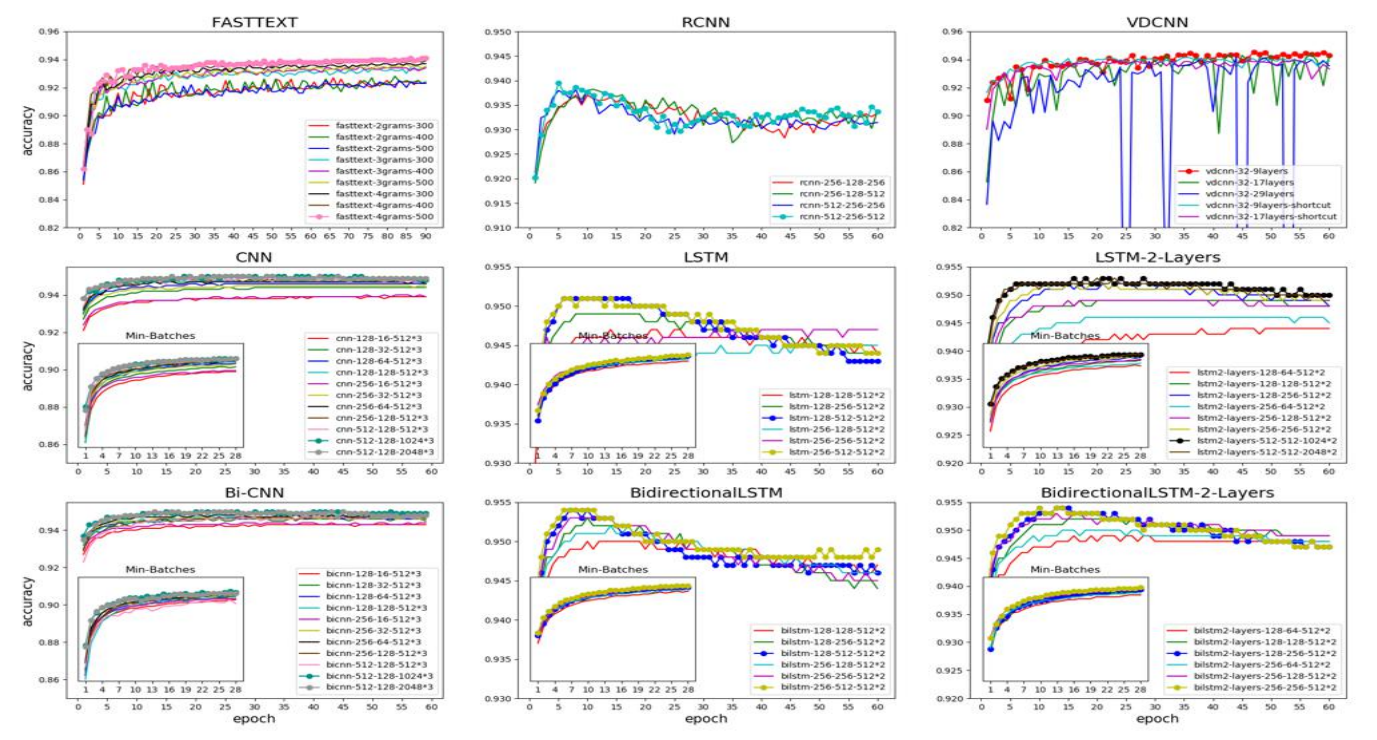
模型效果比对

数据准备
数据清洗





模型规模
参数调整



参数精调：

Pre-Train，Learning Rate，
Batch Size，Regularization

腾讯小微

用户问题指代消解

$$q' = H(q, C)$$

Input:
q: 当前Query
c: 上下文

Output:
q':指代消解后的Query

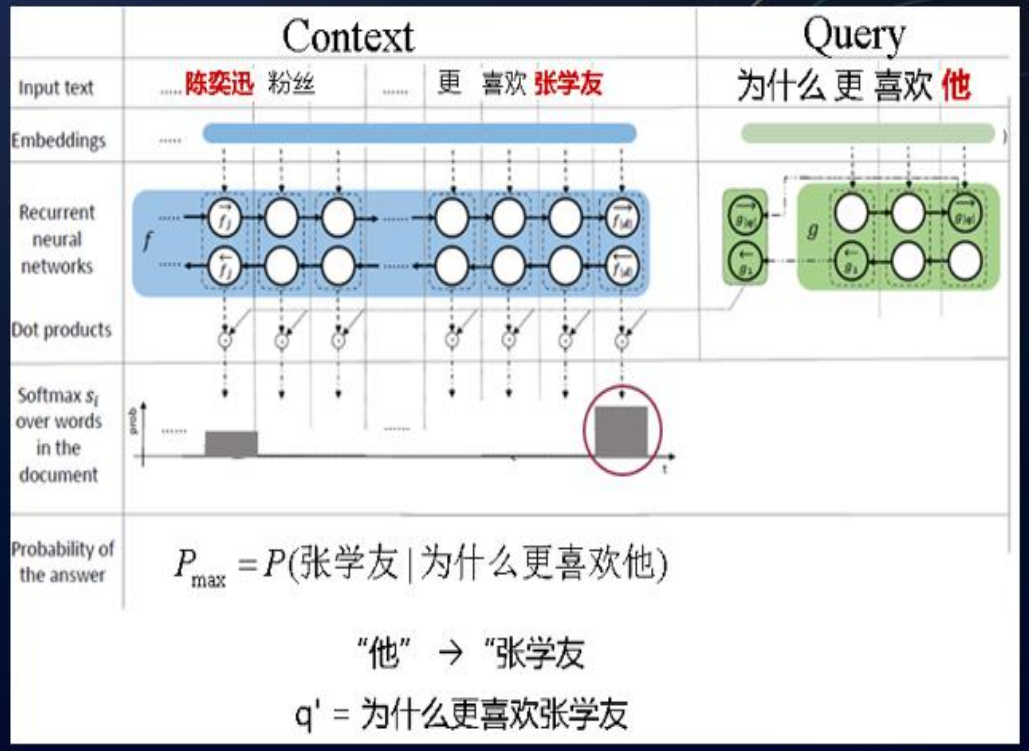
About 5% of the total queries

Examples:

C1: 你是陈奕迅粉丝吗?
C2: 更喜欢张学友
q：为什么更喜欢他？
q ‘:为什么更喜欢张学友

$C1$：你住哪儿？
$C2$：不二寺。
q ：那在哪儿？
q'：不二寺在哪儿？

- 100K training data
- Accuracy: 90%
- Majority of the errors are caused by the mistakes of entity tagging



DSTCT7 Flex Dataset

Example partial dialogue:

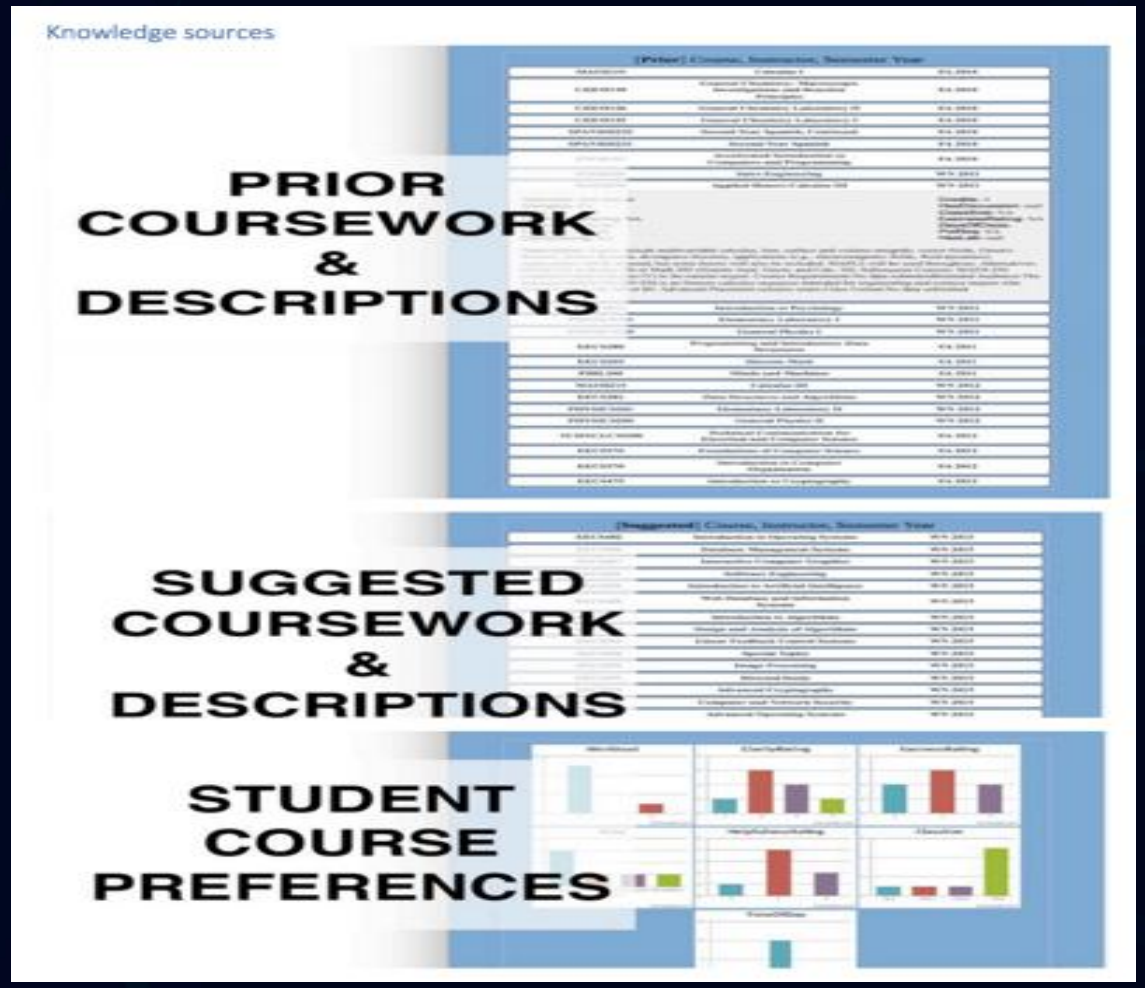
ADVISOR | Hi! What can I help you with?
STUDENT | Hello! I'm trying to schedule classes for next semester. Can you help me?
STUDENT | Hardware has been an interest of mine.
STUDENT | But I don't want too hard of classes
ADVISOR | So are you interested in pursuing Electrical or Computer Engineering?
STUDENT | I'm undecided
STUDENT | I enjoy programming but enjoy hardware a little more.
ADVISOR | Computer Engineering consists of both programming and hardware.
ADVISOR | I think it will be a great fit for you.
STUDENT | Awesome, I think that's some good advice.
STUDENT | What classes should I take to become a Computer Engineer?
ADVISOR | You haven't taken EECS 203, 280, and 270, so it may be in your best interest to take one or two of those classes next semester
STUDENT | Ok. Which of those is in the morning. I like morning classes

Example Candidate set:

Twenty next utterance options, correct ones shown in bold:

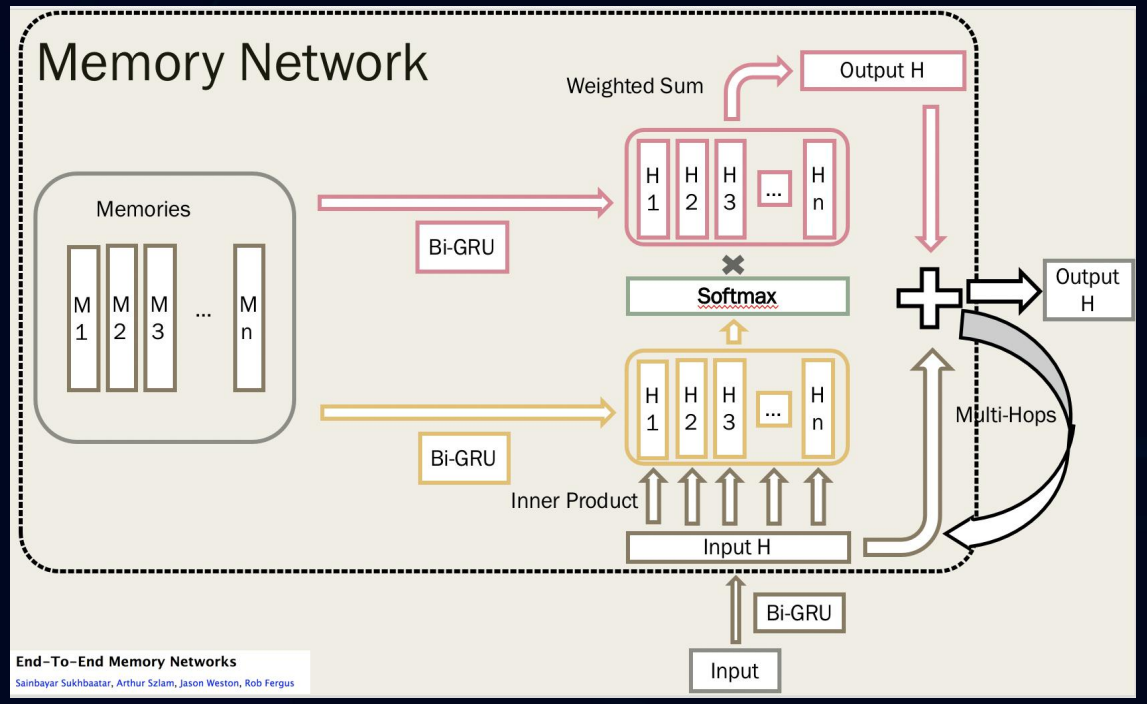
- *Is there anything else I can help answer?*
- ***They have not released the plans for next semester yet.***
- *Do you have an interest in this class?*
- *Do you find this class interesting?*
- *Does this course interest you?*
- *It wouldn't be smart to combine 381 with another EECS course, unless you like to stay up late*

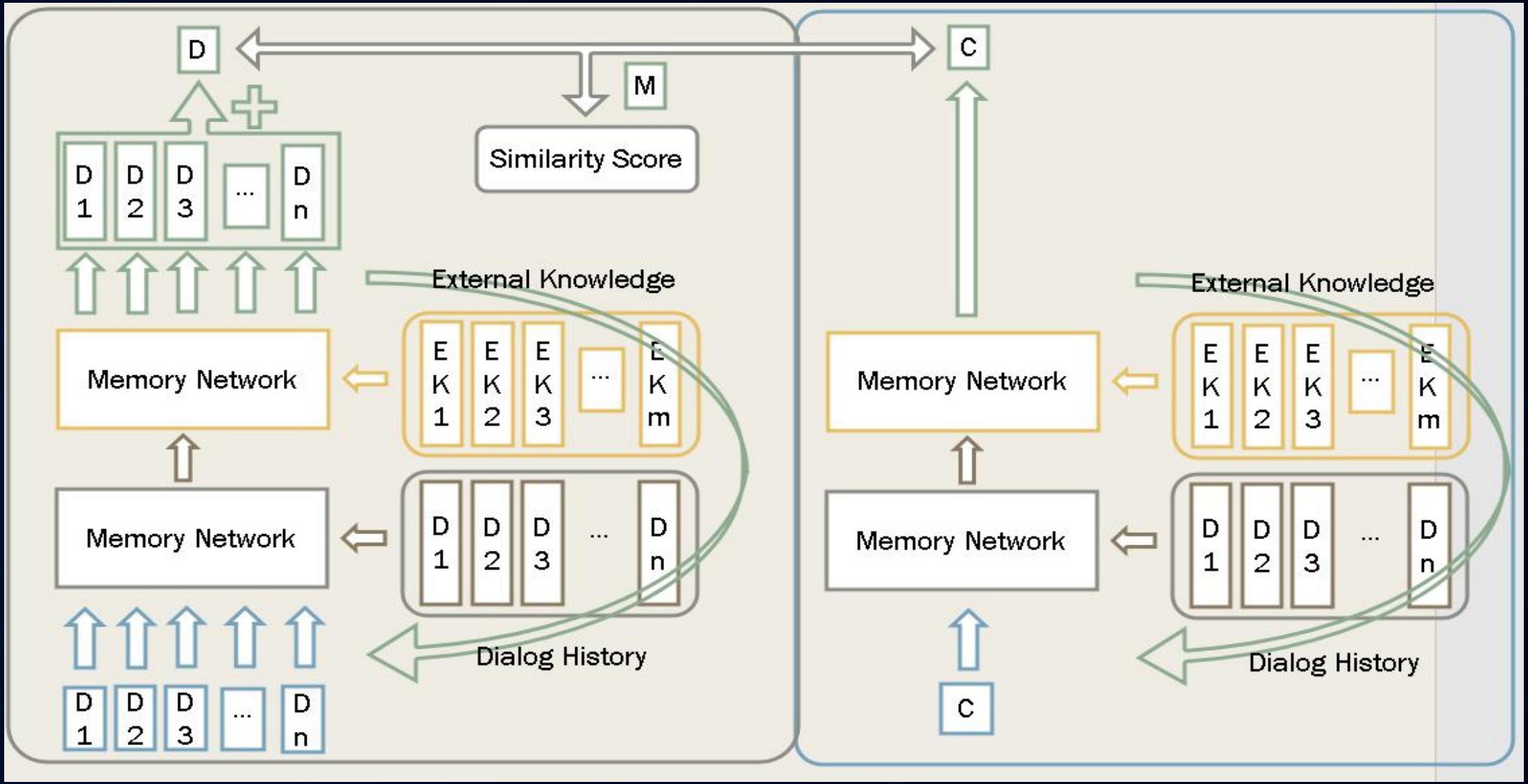
Knowledge sources

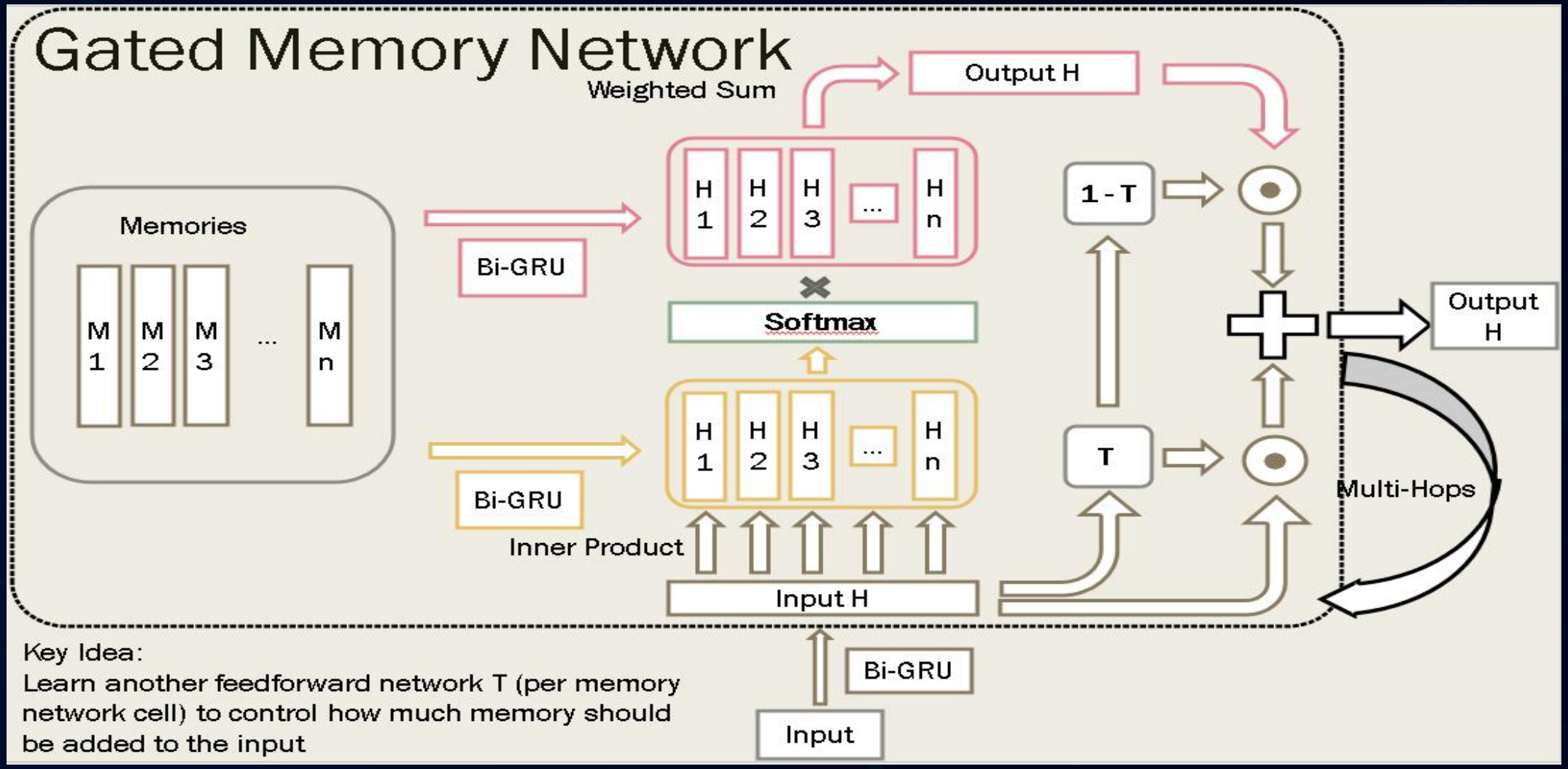


Structure of a memory network cell

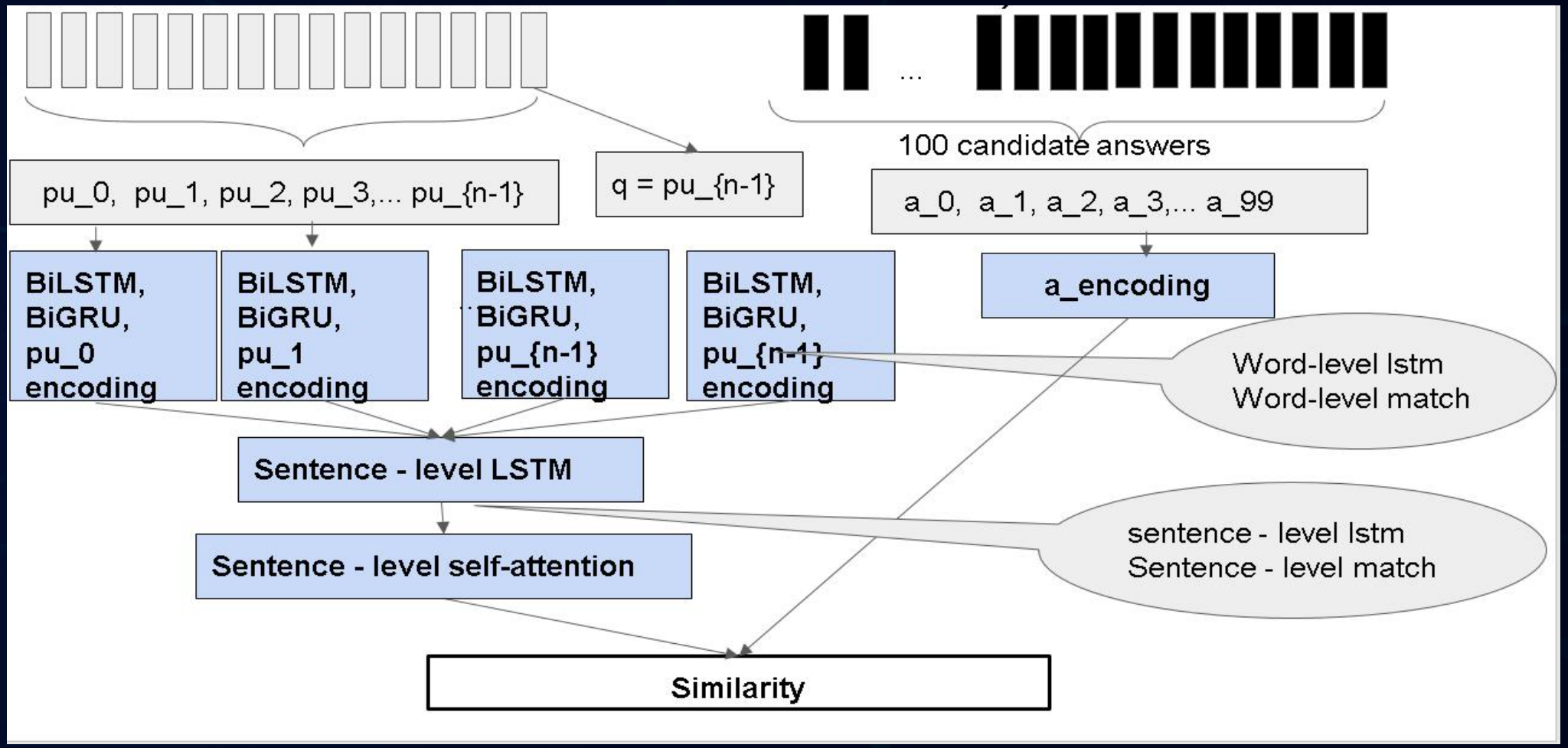
- slightly different from the original paper
  - using **Bidirectional GRU** instead of word embedding matrices
  - using different sets of weights for every memory network hop




D
M
C
Similarity Score
D 1
D 2
D 3
...
D n
External Knowledge
Memory Network
E K 1
E K 2
E K 3
...
E K m
Memory Network
D 1
D 2
D 3
...
D n
Dialog History
D 1
D 2
D 3
...
D n
External Knowledge
Memory Network
E K 1
E K 2
E K 3
...
E K m
Memory Network
D 1
D 2
D 3
...
D n
Dialog History
C


Gated Memory Network
Weighted Sum
Output H
Memories
M 1
M 2
M 3
...
M n
Bi-GRU
H 1
H 2
H 3
...
H n
Softmax
1 - T
T
Output H
Multi-Hops
Bi-GRU
Inner Product
Input H
Bi-GRU
Input
Key Idea:
Learn another feedforward network T (per memory network cell) to control how much memory should be added to the input

Baseline: Hierarchical LSTM



模型benchmark

- recall@10: probability of correct answers in top 10 results

Model	Recall@10
H-LSTM without Attention	0.52
H-LSTM with Attention	0.55
Memory Network	0.65
Gated Memory Network	0.67

案例介绍

Harman Allure

JBL Go Smart 2

海美迪小白盒V1

长安汽车

南方航空

招行客服

案例介绍

智能耳机

外交部12308助手

期待与展望

期待与展望

- **持续打通各硬件终端和软件终端**
  - 包括各种家居硬件，手机，车载；以及公众号，小程序，App
- **持续提供更加精准的对话理解的服务**
  - 即使是用户自定义场景，也会在启动阶段达到较好的理解能力
- **不断增长的内容资源**
  - 基于音乐，公众号，新闻，FM等

/
new trend
new technology
new application

/
Cloud + community
Developer conference