

总结

黄飞虎

2021.2.3

学习情况

看了李宏毅的视频， 完成了实验八和实验九

```
cn vocab size: 3805
en vocab size: 3922
training dataset size: 18000
validation dataset size: 500
testing dataset size: 2636
```

```
train [4800] loss: 3.111, Perplexity: 22.442
val [4800] loss: 3.518, Perplexity: 33.722, bleu score: 0.419
train [5100] loss: 2.813, Perplexity: 16.667
val [5100] loss: 3.498, Perplexity: 33.055, bleu score: 0.428
train [5400] loss: 2.956, Perplexity: 19.221
val [5400] loss: 3.479, Perplexity: 32.422, bleu score: 0.428
train [5700] loss: 2.791, Perplexity: 16.301
val [5700] loss: 3.488, Perplexity: 32.708, bleu score: 0.438
train [6000] loss: 2.842, Perplexity: 17.156
val [6000] loss: 3.483, Perplexity: 32.557, bleu score: 0.437
```

```
test loss: 3.594911128966553, bleu_score: 0.41896599218356556
```

```
test_output.txt - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
[(' <BOS>', 'do', 'n't', 'underestimate', 'my', 'power', '.'], ['不要', '低估', '我', '的', '力量', '.'], ['不要', ' <UNK>', '我', '的', '力量', '.'])
[(' <BOS>', 'could', 'you', 'please', 'talk', 'a', 'bit', 'louder', '?', 'i', 'ca', 'n't', 'hear', 'very', 'well', '.'], ['你', '可以', '说', '我', ' <UNK>', '好', '吗', '吗'], ['你', '能', '大声', '点', '讲', '吗', '?', '我', '听', '不', ' <UNK>', '.'])
[(' <BOS>', 'now', 'is', 'the', 'time', 'to', 'act', '.'], ['现在', '就要', '的', '.'], ['现在', '是', '行动', '的', '时候', '了', '.'])
[(' <BOS>', 'i', 'was', 'just', 'talking', 'about', 'tom', '.'], ['我', '只是', '汤姆', '汤姆', '.'], ['我', '仅仅', '是', '在', '和', 'Tom', '交谈', '.'])
[(' <BOS>', 'you', 'have', 'n't', 'changed', 'at', 'all', '.'], ['你', '没', '有', ' <UNK>', '.'], ['你', '真的', '一点', ' <UNK>', '.'])
[(' <BOS>', 'this', 'is', 'a', 'beautiful', 'flower', '.'], ['这是', '是', ' <UNK>', '的', '的', '.'], ['这是', '一朵', '美丽', '的', '花', '.'])
[(' <BOS>', 'how', 'much', 'do', 'you', 'want', 'in', 'terms', 'of', 'salary', '?'], ['你', '有', '多少', ' <UNK>', '多少', '大', '?'], ['您', '想要', '多少', '工资', '?'])
[(' <BOS>', 'she', 's', 'got', 'a', 'good', 'eye', 'for', 'paintings', '.'], ['她', '是', '个', '好', '的', '厨师', '.'], ['她', '对', ' <UNK>', '有', '很', '好', '的', ' <UNK>', '.'])
[(' <BOS>', 'everyone', 'comes', 'to', 'tom', 'for', 'advice', '.'], ['所有人', '人', '都', '同意', '汤姆', '.'], ['每个', '人', '都', '请', ' <UNK>', '.'])
[(' <BOS>', 'one', 'of', 'the', 'girls', 'was', 'left', 'behind', '.'], ['所有', '们', '被', ' <UNK>', '.'], ['这些', '女孩', ' <UNK>', '其中', '一个', '被', '留下', '来', '了', '.'])
[(' <BOS>', 'i', 'got', 'it', 'free', '.'], ['我', '没关', '它', '了', '.'], ['我', '是', '免费', '得到', '的', '.'])
[(' <BOS>', 'i', 'do', 'n't', 'believe', 'the', 'child', 'came', 'to', 'tokyo', 'alone', '.'], ['我', '不', '记得', ' <UNK>', ' <UNK>', '.'], ['我', '不', '相信', '孩子', '能', '独自', '到', '东京', '.'])
[(' <BOS>', 'there', 'is', 'n't', 'anyone', 'in', 'the', 'room', '.'], ['房间里', '房间里', '房间', '.'], ['房间里', '没有', '任何人', '.'])
[(' <BOS>', 'i', 'will', 'have', 'finished', 'reading', 'this', 'novel', 'by', 'the', 'time', 'you', 'come', 'tomorrow', '.'], ['我', '明早', '在', '这里', ' <UNK>', ' <UNK>', '你', '.'], ['你', '明天', '来', '的', '时候', ' ', ' ', '我', '就', '看', '完', '这本', '小说', '了', '.'])
[(' <BOS>', 'i', 'like', 'may', 'the', 'best', 'of', 'all', 'the', 'months', '.'], ['我', '想要', ' <UNK>', '最', '最', '最', '的', '. ', '.'], ['所有', '月份', '中', '我', '最', '喜欢', '五月', '.'])
```

```
? checkpoint_10.pth
? checkpoint_20.pth
? checkpoint_30.pth
? checkpoint_40.pth
? checkpoint_50.pth
? checkpoint_60.pth
? checkpoint_70.pth
? checkpoint_80.pth
? checkpoint_90.pth
? checkpoint_100.pth
? last_checkpoint.pth
```

```
epoch [89/100], loss:6.56356
epoch [90/100], loss:6.53214
epoch [91/100], loss:6.51322
epoch [92/100], loss:6.48728
epoch [93/100], loss:6.46248
epoch [94/100], loss:6.43906
epoch [95/100], loss:6.41354
epoch [96/100], loss:6.39340
epoch [97/100], loss:6.36944
epoch [98/100], loss:6.35316
epoch [99/100], loss:6.32759
epoch [100/100], loss:6.30697
Latents Shape: (8500, 4096)
```

论文情况

任务型对话系统平台的设计与实现

本文的工作

一个任务型对话系统通常主要包含两个模块：一是命令解析模块，二是对话管理模块。

首先本课题从任务型对话系统纵向的维度出发，重点考察命令解析模块的算法准确性，考察了采用基于语义槽填充的命令解析方法。

另外从任务型对话系统横向扩展性的维度出发，考察了任务型对话系统中，命令解析模块和对话管理模块实现的通用流程。随后本文设计并实现了一个任务型对话系统服务平台，用于提供开发者创建自己的基于自然语言交互的任务型对话系统。

任务型对话系统的框架

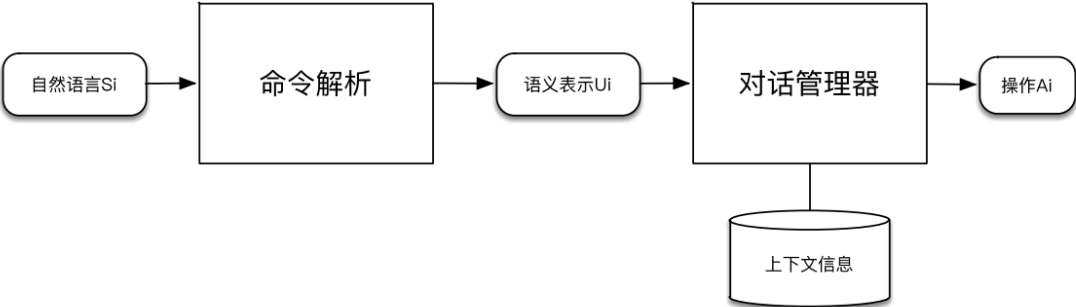


图 1-1 任务型对话系统的工作流程



图 1-2 命令解析模块

表 1-1 任务型对话系统的模块划分

模块名称	子模块名称	含义
命令解析	意图识别	识别用户输入的自然语言语句的意图
	语义槽填充	识别用户输入的自然语言语句中与意图相关的语义槽信息
对话控制	意图内的对话控制	在一个意图之内的对话轮次，控制机器人与用户的自然语言交互。
	意图间的对话控制	在不同意图之间，控制机器人与用户的自然语言交互。

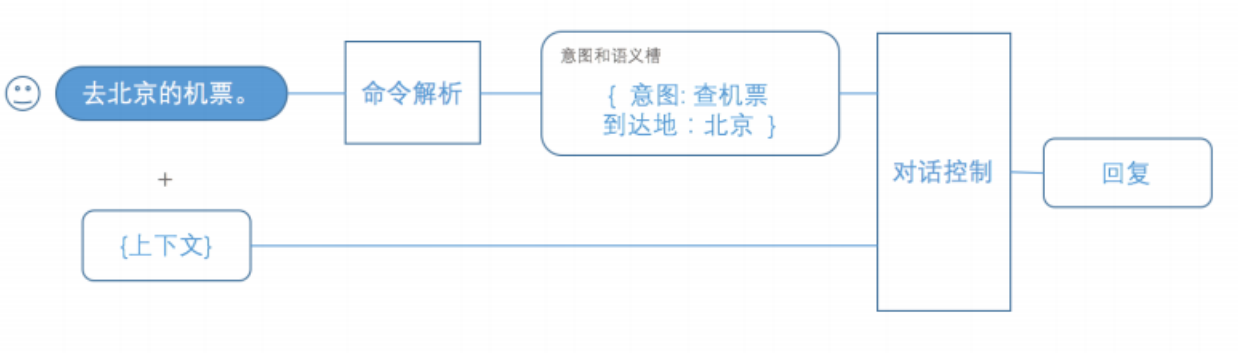


图 1-3 命令解析和对话管理模块

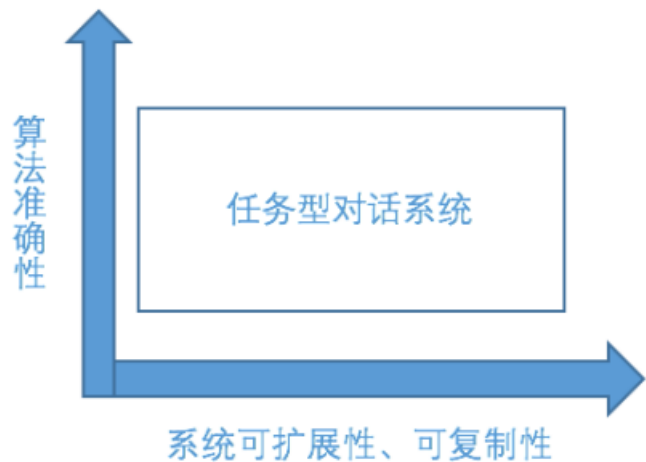


图 1-4 本文研究的两个维度

从任务型对话系统纵向维度进行考虑，研究重点放在命令解析的算法 准确性。重点考察了命令解析任务中一个关键的任务——语义槽填充。

命令解析任务的传统工作流程是：首先识别出用户输入的语句所属的领域；接着系统根据识别得到的领域，进一步识别出该输入所属的意图；最后，系统 根据识别得到的领域或者意图，识别出输入语句的语义表示。其中，基于语义 槽填充（slot filling）的方法是命令解析的重要方法之一。

表 2-1 意图定义

领域	意图
出行	查机票
	查火车
	订机票
	订火车
	查公交
	查地铁
饮食	叫车
	查饭店
住宿	查酒店
娱乐	查歌曲
	查小说
	查电影

表 2-2 查机票意图的语义槽定义

语义槽类型	说明
from_date	出发日期
to_date	到达日期
from_city	出发城市
to_city	到达城市
airline	航空公司

表 2-3 语义槽填充结果

语义槽类型	出发日期	出发城市	到达城市	航空公司
语义槽值	明天	北京	上海	南方航空

在命令解析模块中，其中关键的任务是指令的语义槽识别。我们可以将语义槽识别的任务视为序列标注任务。

目前在序列标注方面提出的方法主要分为三个类别：1）基于规则的方法；2）基于传统机器学习的方法；3）基于深度学习的方法。

本文使用的基线模型为条件随机场模型（CRF），特征模板如右图所示。

表 2-5 本文使用的 CRF 特征模板

特征模版	特征模版类型	特征模版说明
1={c-2}	Unigram	当前词之前两个的词
2={c-1}	Unigram	当前词之前的词
3={c-0}	Unigram	当前词
4={c+1}	Unigram	当前词之后的词
5={c+2}	Unigram	当前词之后两个的词
6={c-2}-{c-1}	Bigram	当前词前两个词的组合
7={c-1}-{c-0}	Bigram	当前词前一个词和当前词的组合
8={c-0}-{c+1}	Bigram	当前词和当前词之后的词的组合
9={c+1}-{c+2}	Bigram	当前词之后两个词的组合
10={p-2}	Unigram	当前词之前两个的词的词性
11={p-1}	Unigram	当前词之前的词的词性
12={p-0}	Unigram	当前词的词性
13={p+1}	Unigram	当前词之后的词的词性
14={p+2}	Unigram	当前词之后两个的词的词性
15={p-2}-{p-1}	Bigram	当前词之前两个词的词性组合
16={p-1}-{p-0}	Bigram	当前词之前的词和当前词的词性组合
17={p-0}-{p+1}	Bigram	当前词和当前词之后的词的词性组合
18={p+1}-{p+2}	Bigram	当前词之后两个词的词性组合

实验模型

1.循环神经网络

2.双向长短期记忆神经网络

3.加入条件随机场层的双向长短期记忆神经网络

实验

数据集：本文使用ATIS (Airline Travel Information System) 数据集作为实验的数据集。

标注方法：使用经典的 IOB(in/out/begin)标记法，例如对于句子“查一下明天从北京到上海 的机票”，将原句子分词后，标注为：

表 2-6 IOB 标注方法

分词结果	查	明天	从	北京	到	上海	的	机票
标签	O	B-from_date	O	B-from_city	O	B-to_city	O	O

实验结果

评价指标：本实验采用准确率（Precision，P）、召回率（Recall，R）以及 F1 值对模型 的性能进行评价。

$$P = \frac{\text{正确识别的语义槽个数}}{\text{语义槽总数}}$$

$$R = \frac{\text{正确识别的语义槽个数}}{\text{识别出的语义槽总数}}$$

$$F = \frac{2 \times P \times R}{(P + R)}$$

表 2-7 实验结果

	开发集 F1	测试集 F1
CRF	94.45	90.93
RNN	97.67	94.47
Bi-LSTM	97.99	94.87
Bi-LSTM-CRF	98.35	95.22

任务型对话系统平台的设计

案例展示

表 4-2 语义槽定义示例

意图名称	语义槽参数			
	语义槽名称	实体类型	是否必须	描述
查询机票	from_city	geoloc	√	出发地
	to_city	geoloc	√	到达地
	from_date	date	√	出发时间
查询火车票	from_city	geoloc	√	出发地
	to_city	geoloc	√	到达地
	from_date	date	√	出发时间
	railway_type	railway_type	✗	车型

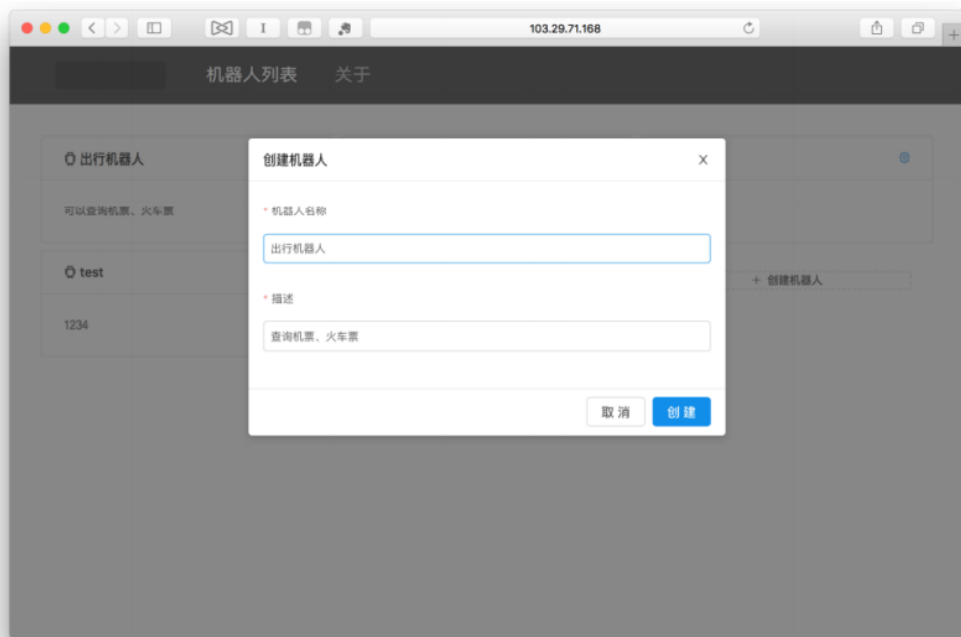


图 4-11 创建出行对话系统

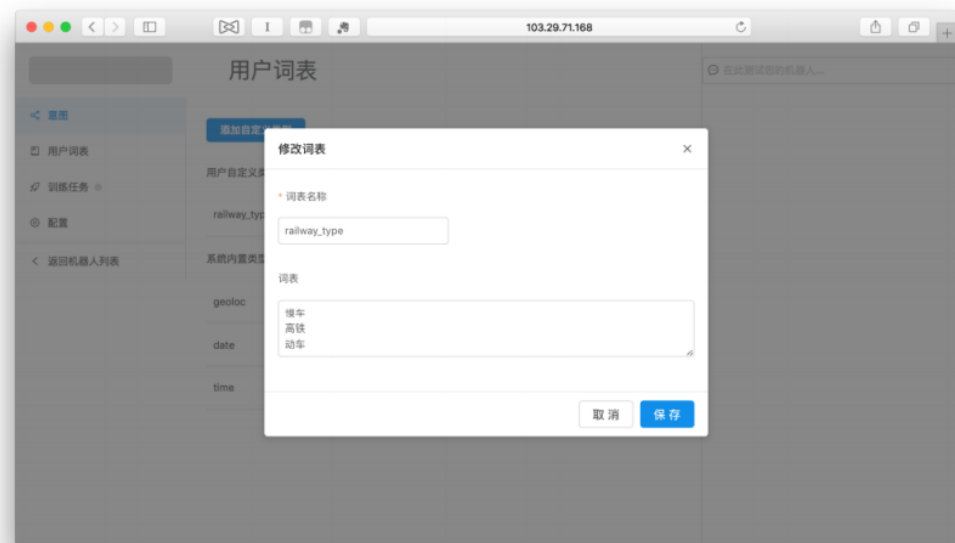


图 4-12 添加用户自定义实体

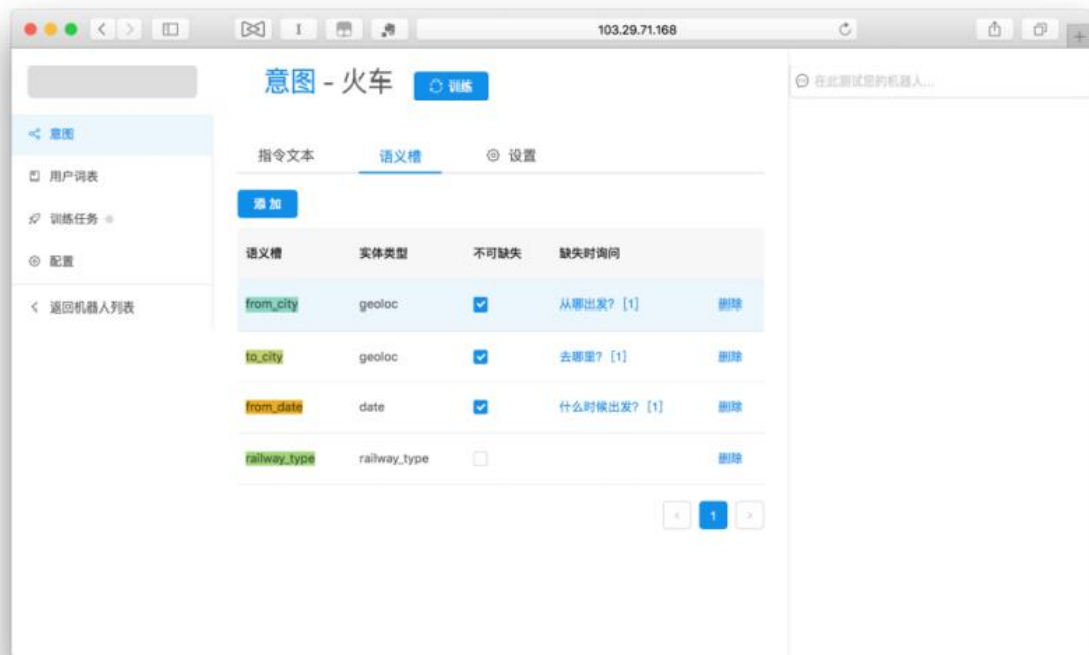


图 4-13 火车票意图的语义槽定义

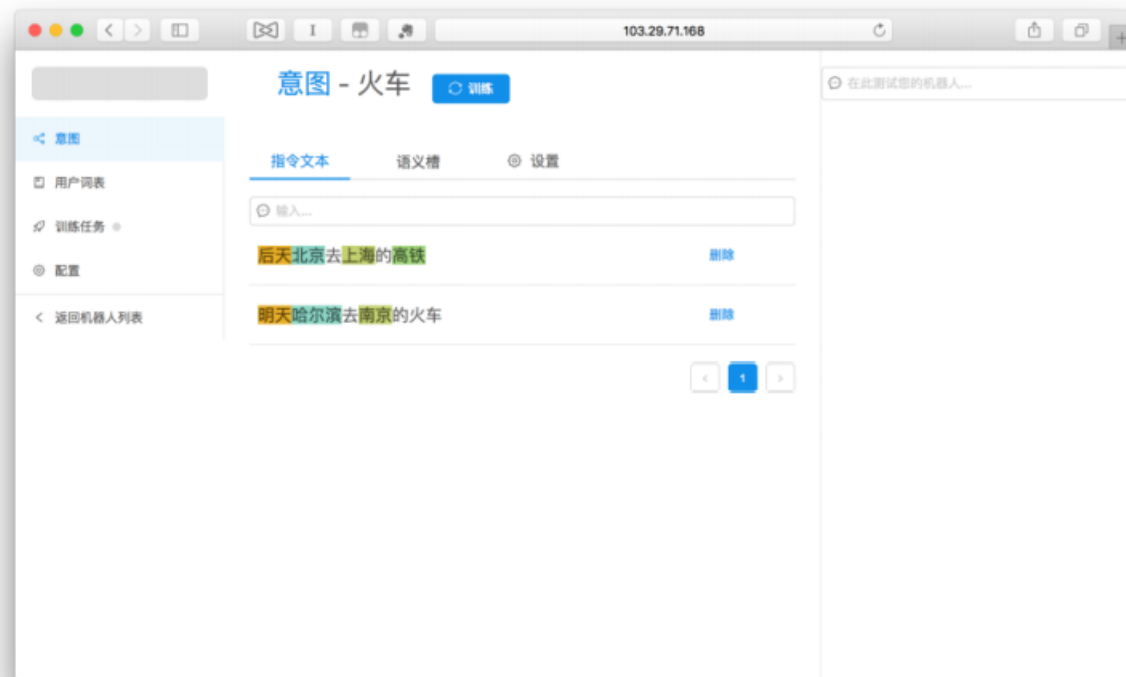


图 4-14 为火车票意图添加并标注语料

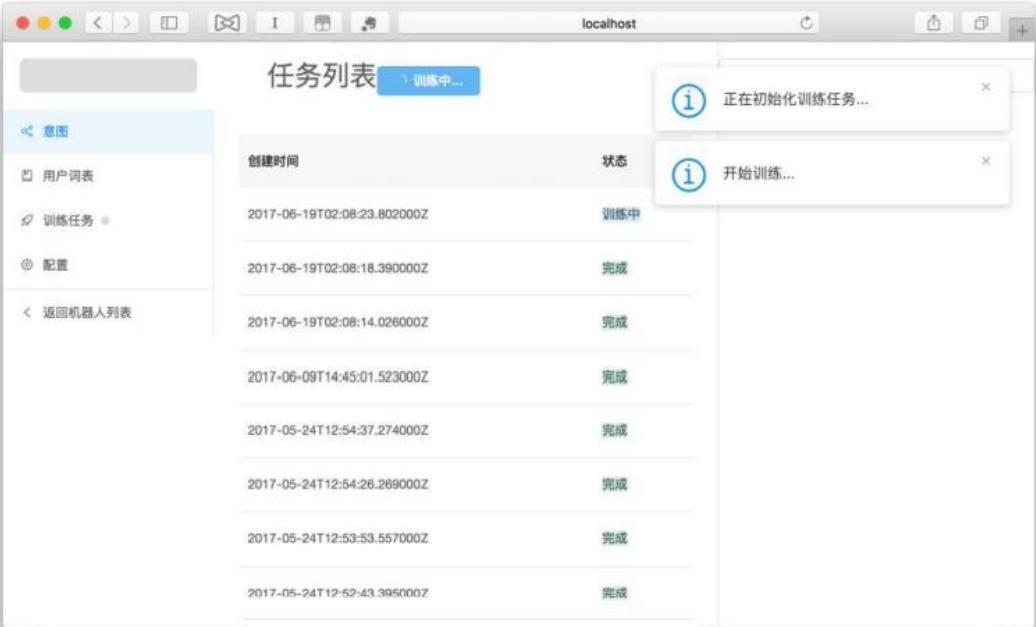


图 4-15 训练任务列表

表 4-3 调用结果示例

文本	
输入	查一下明天去北京的高铁
输出	从哪出发？
输入	后天从哈尔滨出发
输出	下面是[2017-06-21]从[哈尔滨]去[北京]的[高铁]



图 4-16 查询火车票测试示例

Thank you