

RLオータムセミナー2021

RLAS2021最終課題

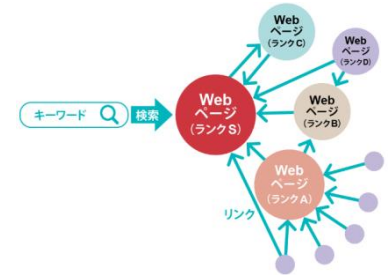
チャットボットの強化学習への適用における一考察

-問題設定とレコメンドエンジンとの比較から-

提出日: 2022年1月11日
放送大学 太田 博三

1. 概要①

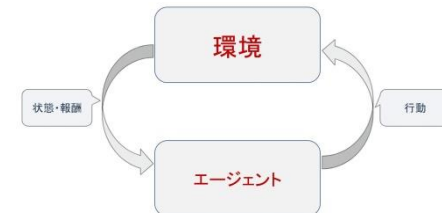
1) 従来のレコメンドエンジンを念頭に、自然言語処理のチャットボットの強化学習への効果的な適用を模索した.



2) 従来、対話応答の短い発話では、いきなりのエンドユーザーの意図(Intent)は汲み取りずらいのが課題である.



3) そこで、強化学習を用いて、且つ効果の上がる問題設定を検討した.



1. 概要②

4) 強化学習は直接的に観測されないため, レコメンドエンジンのようにはゆかないが, FAQなどでその発話範囲を閉じて, 質問の意図を特定しつつ, それに対してスコアリングするアプローチする「Self-Improving」なチャットボットを考察した。

5) DQNで総当たり戦となるため, 多くのケースには対応できないが, おおよその場合は発話の意図に沿って対応できることを確認した。



2. 先行研究

1) 対話エージェントの応答生成

Jiwei Li et.al(2016)ら "Deep Reinforcement Learning for Dialogue Generation"

対話エージェントの応答生成が強化学習の典型であった。報酬モデルと方策勾配法とで、2者間の仮想エージェントを想定し、対話をシミュレートするものであった。

2) 包括的なチャットボット・モデルの確立

Iulian V. Serban et.al(2017)ら "A Deep Reinforcement Learning Chatbot"

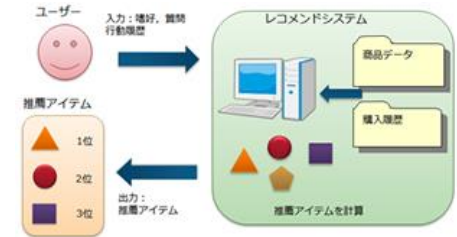
言語生成に向けたBoWモデルからテンプレート・ベースやトピックスベースでのモデルなど、様々な側面からA/Bテストを通じて、最適な応答選択を可能とした研究がなされている。

3) 目標達成志向のチャットボットへの取り組み

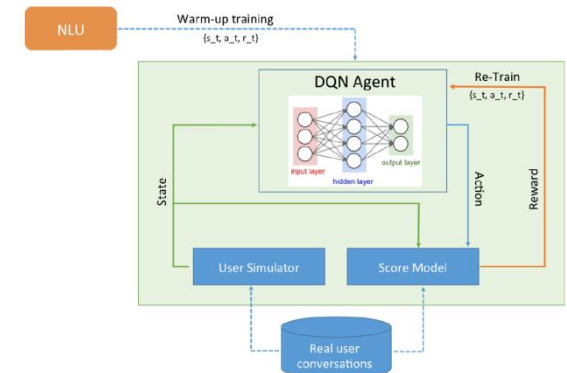
Donghoon Ham et.al(2020) "End-to-End Neural Pipeline for Goal-Oriented Dialogue Systems using GPT-2"

3.問題設定

1) 強化学習の問題設定は，直接的なレコメンドエンジンに比べて，間接的にならざるおえない．



2) ここで，FAQを想定し，スコアリングによるしきい値で，応答選択を効率化した．



4.実装面の考察

- 主要な3つのUnitで構成されている.

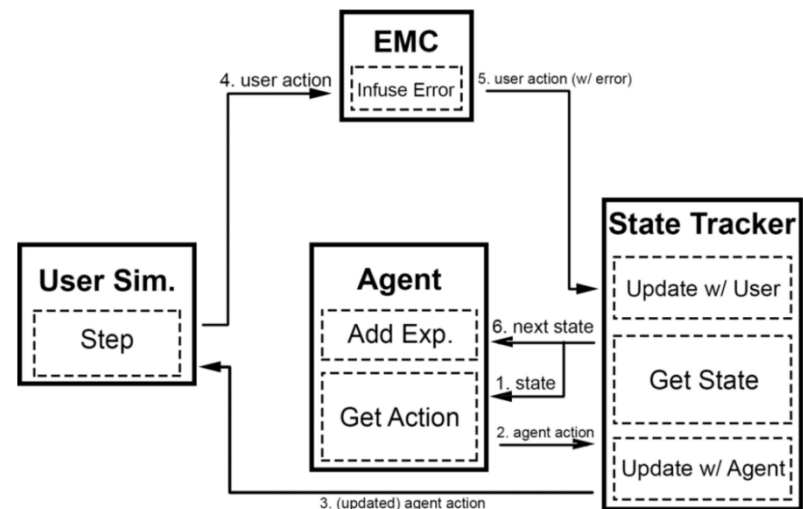
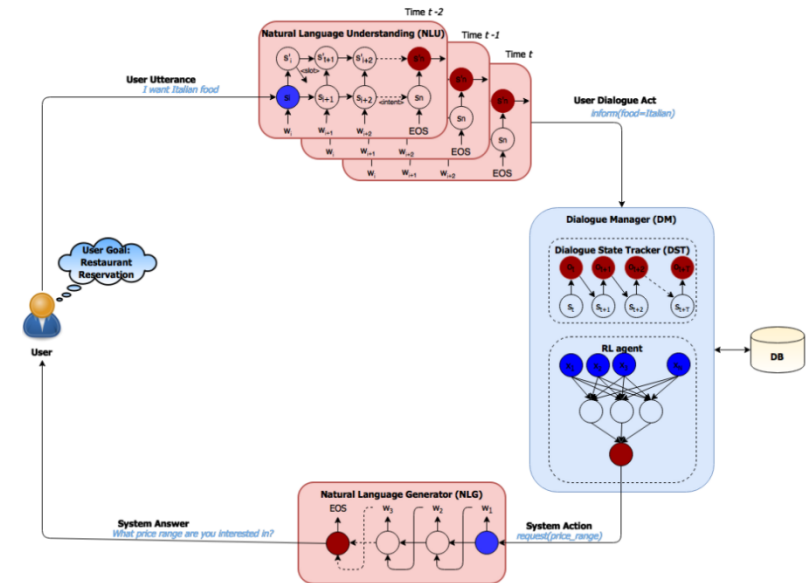
- 1) Dialogue Manager(DM),
- 2) Natural Language Generator(NLG)
- 3) Dialogue State Tracker(DST)

①UserのGoal:レストランの予約をすること

②NLUコンポーネントで,
User Utterance(イタリア料理が食べたい)

③Dialogue Managerコンポーネントで
User-Dialogue Act(イタリア料理)を処理し,
RL AgentコンポーネントでDB化される.

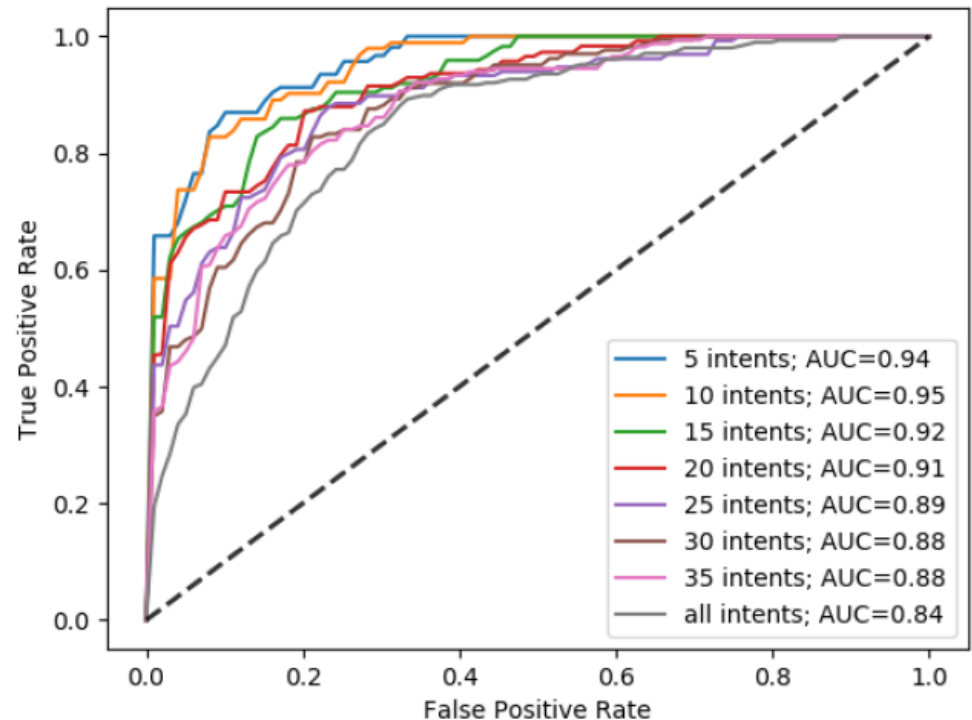
④このSystem Action(場所)を,
NLGコンポーネントで受け付けて,
価格の範囲をUserに問うもの.



5. 結果の考察

- Intent(意図)が5個から10個までは, 95%である.
- しかし, 25個以上になると90%を下回ってしまう.

Episode	Avg Reward:
600	0.07
1300	0.14
3500	0.58
6000	0.7
8000	0.812
100000	0.846
120000	0.864
160000	0.896
200000	0.903
240000	0.916
280000	0.925
300000	0.925
320000	0.929
360000	0.936
400000	0.944



6. まとめ

- 短めな発話でも、シーケンシャルな性質も加味できるようになった.
- 学習の初期段階の50%から75%に、20-30epochで到達できるメリットがある.
- その反面、FAQの意図設定とスロットの数との人為的なバランスが必要となる.

文献・参照先URL等

- A Deep Reinforcement Learning Chatbot
- <https://arxiv.org/pdf/1709.02349.pdf>
- Deep Reinforcement Learning for Dialogue Generation
- <https://arxiv.org/pdf/1606.01541.pdf>
- Building Advanced Dialogue Managers
- for Goal-Oriented Dialogue Systems
- <https://arxiv.org/pdf/1806.00780.pdf>
- Training a Goal-Oriented Chatbot with Deep Reinforcement Learning — Part I | by Max Brenner | Towards Data Science
- <https://towardsdatascience.com/training-a-goal-oriented-chatbot-with-deep-reinforcement-learning-part-i-introduction-and-dce3af21d383>
- Self-improving Chatbots based on Reinforcement Learning
- May 2019
- https://www.researchgate.net/publication/333203489_Self-improving_Chatbots_based_on_Reinforcement_Learning
- Self-improving Chatbots based on Deep Reinforcement Learning | by Debmalya Biswas | Towards Data Science
- <https://towardsdatascience.com/self-improving-chatbots-based-on-reinforcement-learning-75cca62debce>
- maxbren/GO-Bot-DRL: Goal-Oriented Chatbot trained with Deep Reinforcement Learning
- <https://github.com/maxbren/GO-Bot-DRL>
- MiuLab/TC-Bot: User Simulation for Task-Completion Dialogues
- <https://github.com/MiuLab/TC-Bot>
- https://github.com/otanet/RLAS2021_chatbot_DeepRL_20220111