

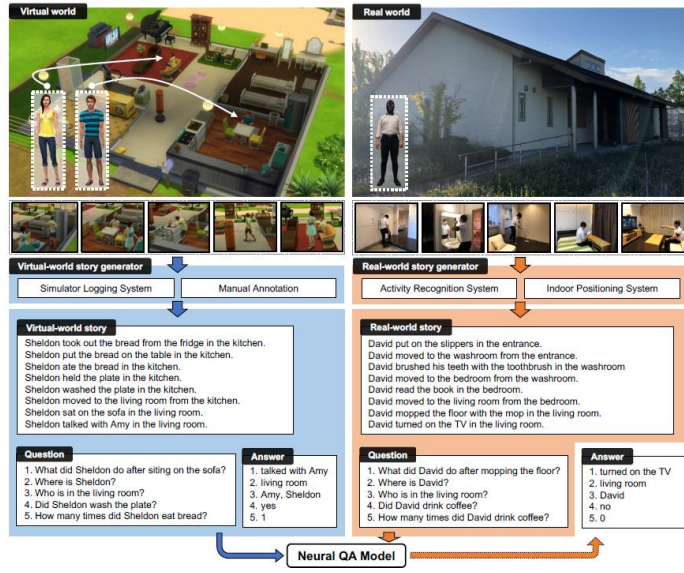
仮想世界のデータを用いたドメイン適応による実世界質問応答の高度化

～ 仮想世界で得た推論能力を、現実世界で活かす ～

宮西大樹 (国際電気通信基礎技術研究所) ACT-I先端研究フォーラム 2020



Sim2RealQA : Simulation to Real Question Answering



解決したい問題

- 日常生活で起きた出来事を問い合わせることができるシステムを作るためには、実世界の質問応答データが数多く必要になります
- 本研究では、ライフシミュレーターを用いることで日常生活に関する大量の質問応答データを作成し、このデータを用いて高精度に実世界の質問応答課題を解く方法を開発しました

どのような未来を切り拓くのか？

- 実世界質問応答の技術を利用すれば、「いつ、どこで、誰が、何をしたか」といった情報を簡単に把握できるようになります
- 人間の記憶支援・忘れ物や落とし物の検索・人の監視や見守り・ホームロボットといった実世界に根ざしたシステムの実現が期待できます

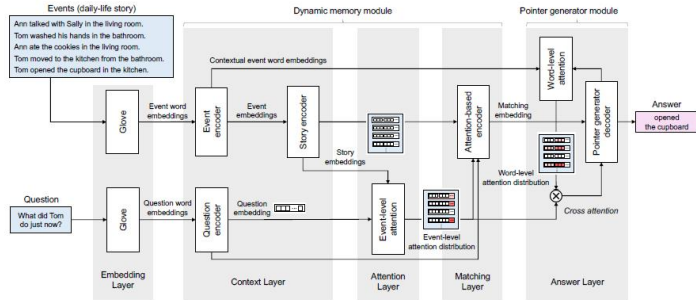
得られた成果

- 現実世界の質問応答の問題を高精度に解くために、仮想世界のデータを使う Sim2RealQA を提案しました
- 実際の家屋とライフシミュレーションゲーム内で収集した日常生活の行動ログをもとに現実・仮想双方の質問応答データセットを作成しました
- 現実世界の正解ラベルが取得困難な場合、Sim2RealQA の枠組みが実世界質問応答の問題を解く上で有効な手段であることを示しました

データ収集に利用した現実世界 & 仮想世界の環境

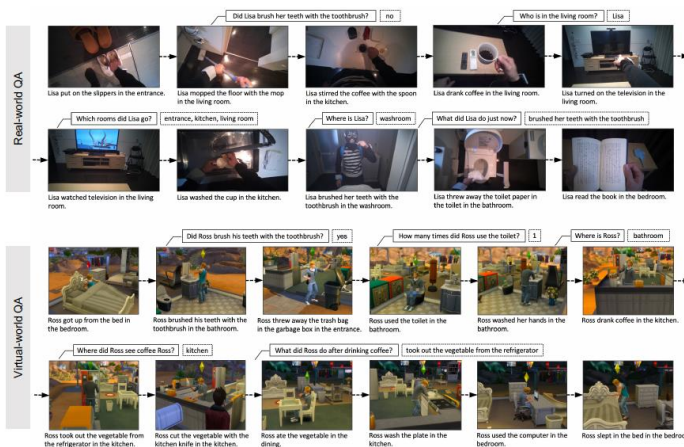


ドメイン横断質問応答モデル

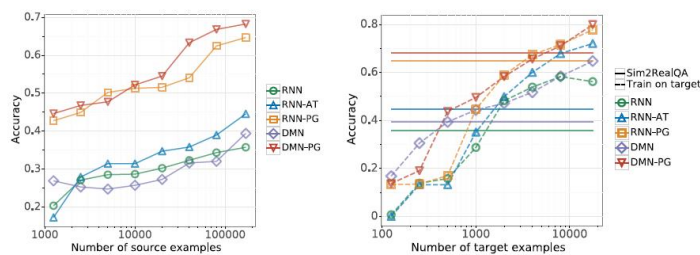


- 質問文に適合する出来事を文レベルの注意機構を使って推定
- 単語と文レベルの注意機構とコピー機構を組み合わせ、質問に適合する未知の解答を出力可能にした

現実世界と仮想世界の質問応答データの例



Sim2RealQAの実験結果



- 学習に使う仮想世界のデータ量が増えるほど、実世界質問応答の問題をより高精度に解けることがわかった
- 大量の仮想世界のデータで学習したモデルは、少量の現実世界のデータで学習したモデルを上回る性能を示すことがわかった