

Switching Head Tail Funnel UNITER による

対象物体および配置目標に関する指示文理解と物体操作

是方諒介（慶應義塾大学大学院 理工学研究科 開放環境科学専攻）

要旨

高齢化が進行する現代社会において、在宅介助者不足に対する解決策として、被介助者を物理的に支援可能な生活支援ロボットへの期待が高まっている。しかし、人間からの自然言語による指示をロボットが理解する能力についてはいまだ不十分である。本研究では物体の把持および配置に関する指示文を生活支援ロボットが理解し実行するための手法の構築を目的とする。例えば「皿の左側にある黄色いボトルを、上に何も置かれていない椅子まで移動させる」という指示文が与えられたときに、環境中を撮影した画像からボトルおよび椅子を特定し、把持および配置動作を行うことが求められる。本タスクにおいて、指示文の対象物体および配置目標を短時間で予測することは重要である。単純な手法では、対象物体候補と配置目標候補に関するすべての組合せについて推論を行うため多くの推論回数を要し、リアルタイム性で実用面に問題がある。本論文では、言語理解における対象物体および配置目標の探索に必要な推論回数を削減することが可能な Switching Head-Tail Funnel UNITER を提案する。Switching head-tail 機構を導入することで、対象物体および配置目標について、単一モデルで個別に予測可能にする。結果として、ALFRED を基にした新規データセットにおいて提案手法がベースライン手法を言語理解精度で上回った。また、実機ロボットを用いて実環境における実験も行い、同様に提案手法がベースライン手法を言語理解精度で上回った。さらに、把持および配置動作を 90%以上の成功率で実行可能であることを示した。