

コンピューターは文章が理解できるか？

何が機械読解の問題を簡単にしているのか？

菅原朔（東京大学）・相澤彰子（NII）

どんな研究？

- 機械のために作られた文章題は本当に人間らしい言葉の理解が実現できるほど難しいものなのか？
- 問題文と文脈文の性質を使って問題の難易度を区別する手法の提案

何がわかる？

- 既存のデータセットには簡単な問題が多く、機械は見かけより賢くない？
- 機械に人間の言語が「ちゃんとわかっている」ようにするために開発に使うデータをしっかり作るのは難しい

状況設定

機械読解タスクと背景

機械に文章読解を学習させ解かせるタスク。近年では数万問程度の大規模なデータセットが多数提案されていますが、ほとんどの問題はクラウドソーシング等で作られています。こうした問題で高い精度を出せるとき「ちゃんとわかっている」と言えるでしょうか？

SQuAD 2.0 (2018) というデータセットでは、機械の精度が人間を超えていたりします 😊 😊 😊

例題

2014年11月、ソニーはハッカーの標的となり、ハッカーは注目を集める映画プロジェクトに関しての極秘の電子メールの詳細を公開しました。これらの中には、スペクターの制作についてのソニーの不満を表現するメモが含まれました。(Wikipedia: 007 スペクター (映画) より)

問題文：いつハッカーはソニー・ピクチャーズのメールシステムに侵入しましたか？

研究内容

2014年11月、ソニーはハッカーの標的となり、ハッカーは注目を集める映画プロジェクトに関しての極秘の電子メールの詳細を公開しました。[...]

問題文：いつハッカーはソニー・ピクチャーズのメールシステムに侵入しましたか？

【観察】

- 「いつ」と訊かれているものの時間を表す単語が「2014年11月」しかない！
- もし他に時間を表す単語があっても「ソニー」「メール」「ハッカー」を見るだけで答えられそう！

【提案手法】

次のふたつのルールで問題を分類

1. 問題文の最初の数語だけで正答できる ★
2. 問題文と表現が近い文に答えがある ★

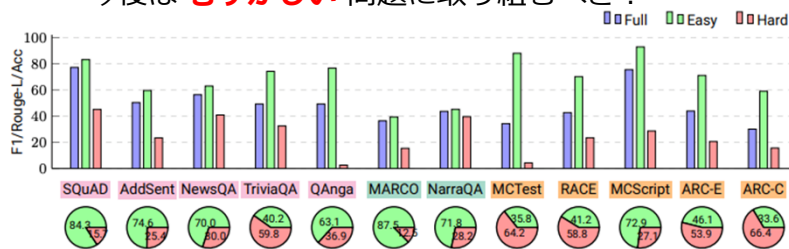
どちらかを満たす → 「やさしい」

どちらも満たさない → 「むずかしい」

【分析：単純なルールで分類に成功！】

- 機械の精度：
やさしい > むずかしい
- 単語の一致を見て解ける問題の割合：
やさしい > むずかしい
- 推論が必要な問題の割合：
むずかしい > やさしい

→ 今後は むずかしい 問題に取り組むべき？



参考文献

Saku Sugawara, Kentaro Inui, Satoshi Sekine, Akiko Aizawa, "What Makes Reading Comprehension Questions Easier?", in *Proceedings of EMNLP*, pp.4208-4219, 2018.
Pranav Rajpurkar, Robin Jia, Percy Liang, "Know What You Don't Know: Unanswerable Questions for SQuAD", in *Proceedings of ACL*, pp.784-789, 2018.