

自然言語理解ユニットタスク構築に向けた 読解タスクの分析

NLP若手の会 第11回シンポジウム ポスター発表 P32
白浜(和歌山), 2016年08月30日

菅原朔(東大), 相澤彰子(NII)
sakus@is.s.u-tokyo.ac.jp

■ 研究の概要

❖ 背景・大目的

文章読解などの高度な自然言語理解を実現するシステムを詳細に設計・評価するために、システムが最低限解ける必要のある問題群（ユニットタスクと呼称）を細分化して構築し、そのタスク群を解かせつつ開発を進める（ユニットテストと同様の態度）

❖ 現状の課題

自然言語理解の能力を構成する要素が何であるかを網羅的に定義することが難しい

❖ 課題解決のためのアプローチ

仮に整理した能力のセットを用いて既存の読解タスクをアノテーションすることで、読解のプロセスを漏れなく構成できるかどうか検証する。能力のセットに何が足りないか・何を細分類しなければならないか検討する

→ それぞれの能力を実現する指針で既存システムを改善して、読解タスクの精度を向上させる（という話にしたい）

John picked up the apple.
John went to the office.
John went to the kitchen.
John dropped the apple.
Q: Where was the apple before
the kitchen?
A: office

想定されるユニットタスクの例
bAbl tasks (Weston+ 2015) より

■ 読解に必要とされる能力の分類と MCTest (MC160 dev. set) のアノテーション結果と既存手法の精度

Skills	Freq.	Descriptions or examples	Smith ¹ no RTE	Smith ¹ RTE
List, enumeration	11.7%	事物や状態の追跡・保持	78.6%	71.4%
Mathematical operations	4.2%	四則演算、図形の理解	20.0%	20.0%
Coreference resolution	57.5%	共参照解析	65.2%	69.6%
Logical reasoning	0.0%	帰納、演繹、条件文、量化子	-	-
Analogy	0.0%	修辞技法的な Trope, 暗喩や直喩など	-	-
Spatiotemporal relations*	28.3%	事象の時間・空間的な関係	70.6%	76.5%
Causal relations*	18.3%	明示的に表現される因果関係	63.6%	68.2%
Complex sentences*	15.8%	等位接続、縦続接続	52.6%	68.4%
Commonsense reasoning	49.2%	辞書的知識、定性的知識、状態の変化	59.3%	64.4%
Linguistic knowledge*	10.0%	修辞技法的な Scheme, 構文や句読点など	50.0%	50.0%
-	-	(Accuracy in all 120 questions)	67.5%	70.0%

*: 「それを理解していること」の意 / ¹ Smith+ 2015, A Strong Lexical Matching Method for the Machine Comprehension Test

- ❖ 「その問いを解くためにどの能力が必要か」をマルチレベルでアノテーション
- ❖ 能力は、「節レベルの事象の内容は理解できること」を前提として、「節間の関係を理解するもの」として定義
- ❖ ただし Complex sentences は、同一文中の複数の節の関係を理解する能力という意味で特徴的（時間・空間・因果関係は独立した能力として区別）
- ❖ Linguistic knowledge は、修辞的な言語表現を解釈する能力という意味で例外的
- ❖ システムの差異に対して、期待する能力が向上しているかを比較することが可能になる

■ アノテーション例（単独の問いに5つの能力が必要）

ID: MC160.dev.29 (1) multiple:

C1: The princess climbed out the window of the high tower and climbed down the south wall when her mother was sleeping.

C2: She wandered out a good ways.

C3: Finally she went into the forest where there are no electric poles but where there are some caves.

Q: Where did the princess wander to after escaping?

A: Forest

- ❖ Coreference resolution:
 - She in C2 = She in C3 = the princess in C1
- ❖ Temporal relation:
 - the actions in C1 → wandered out ... in C2
 - went into ... in C3
- ❖ Complex sentence and linguistic knowledge (ellipsis):
 - C1 = the princess climbed out ... and [the princess] climbed down ...
- ❖ Commonsense reasoning:
 - escaping in Q ⇒ the actions in C1
 - wandered out in C2 and went into the forest in C3 ⇒ wander to the forest in Q and A

■ 議論・まとめ

- ❖ 問題あたりに必要な能力の数は、120問に対する割合で 0: 9.2%, 1: 27.5%, 2: 30.0%, 3: 26.7%, 4: 5.8%, 5: 0.8% であった。5つ必要とする問い(1/120問)が左記
- ❖ Logical reasoning と Analogy は MCTest の文章が易しいために出現しなかったと考えられる
 - 論説文など、やや高度な文章の読解タスクをアノテーションする (cf. QA4MRE (2013), SQuAD (2016))
- ❖ Commonsense reasoning の下位分類が課題であり、辞書的な知識とイベント的な知識などの区別（とそれに基づいてユニットテストが作成できること）が必要
- ❖ 既存手法2つの差 (RTE の有無) で causal relations / commonsense reasoning などの成績の向上が期待されるが、データ数が少ないために有意な差が出ていない
 - 現在、MCTest の他のデータセットについてアノテーションを継続

■ 今後の展望

- ❖ 提案する能力の分類に合わせたユニットタスクを作成
- ❖ 提案する能力を強化する形で実際に既存手法を改良