

言語知識の近似値として 言語モデルは利用できるか

大谷直輝

(東京外国語大学)

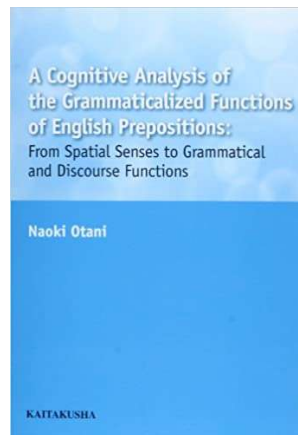
自己紹介

大谷直輝

- 所属: 東京外国語大学

専門:

- 英語学・言語学(認知言語学・構文文法・コーパス言語学)
- 今回、チュートリアルを担当しましたので、興味がある方は資料をご参照ください。



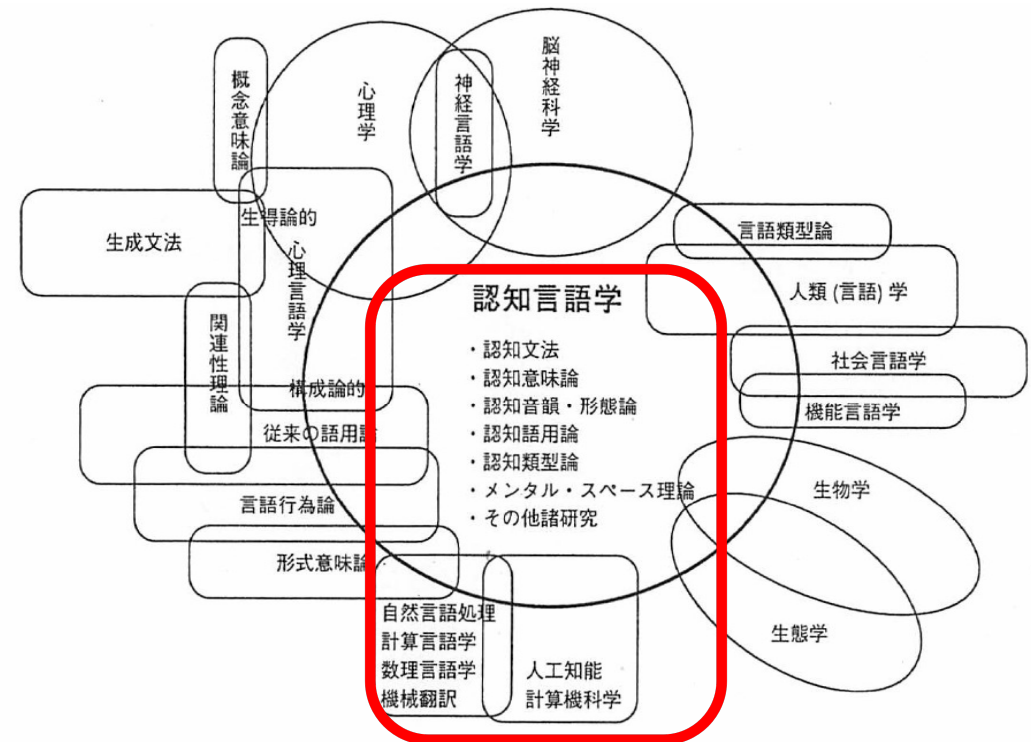
はじめに

タイトル

- 言語知識の近似値として言語モデルは利用できるか

タイトルの背後にある問題意識

- 言語学を行う究極的な目的は「**人間の言語知識(あるいは能力)はどのようなものであるかを探ること**」であるが、言語知識(心のありようの一部)を探るヒントとして、言語モデルをはじめとしたNLPの知見を利用できないか。



辻「認知言語学の見取り図」(『言語』27(11)より) 改変

言語学者としての問題意識

言語学者としての問題意識

- 人間が持つ言語知識(能力)とはどのようなものかを知りたい。
 - どのような知識(能力)があれば、人間は言語を用いた活動ができるか知りたい。
-
- (1) A: Can you pass me the salt?
B: Yes, I can!
 - (2) A: Can you pass me the salt?
B: Sure, here you are.
-
- (3a) Answer the phone.
(3b) I want you to answer the phone.
(3c) Will you answer the phone?
(3d) Can you answer the phone?
(3e) Would you answer the phone?
(3f) Could you answer the phone?
(3g) I was wondering if you could answer the phone.

言語学の分析をする上での葛藤

言語知識を探るうえでの方法論的な問題点

- 人間が持つ言語知識は直接捉えることが難しいため扱いにくい。何かを言語知識の近似値としてみなす必要がある。
- 言語学では、文のミニマルペアの比較、コーパスにおける頻度や分布の変化などを用いて、言語知識のありようを探ろうとしてきた。ただし、言語だけを見ていて、言語を生成する役割を担う心の働きのことを語るには限界がある。
- そこで、その近似値として、認知科学の諸分野の知見、脳波の測定、心理実験、言語習得、動物のコミュニケーションなどに注目をする。
- そのような、近似値を捉える一つの方法として、言語モデルは利用できないか。利用できる場合、どのような研究が可能となるか。

学会中にあった前提が覆された出来事

- 自然言語処理は(認知)科学ではない工学の一種だ！
- そもそも、言語モデルを言語知識の近似値としてみなすこと自体できるのか！

学会に参加して感じたこと

言語学

- 人文系の言語学(philology)
- (認知)科学としての言語学(linguistics)

自然言語処理

- (応用的な)工学としての自然言語処理
- 認知科学としての自然言語処理

機械と人間の心の関係に興味を持っている参加者も結構いる。

- 人工知能を通して人間の脳の仕組みについて考えたい。(機械→心)
- ロボットの語彙概念習得に幼児の言語習得の考え方を利用する。(心→機械)

結論

- 学会への参加を通して、言語学者が、認知科学的(つまり、言語処理を通して、人間についての理解を深めるというよう)な考えを持って、言語処理学会に参加すると、学際的なネタがたくさん見つかるのではないか。

認知言語学と言語処理の共通点

私の立場 = 用法基盤モデル

- 言語知識は(文法も語彙も)言語使用の中で繰り返し現れるパターンからボトムアップで立ち現れる。

深層学習ベースの自然言語処理

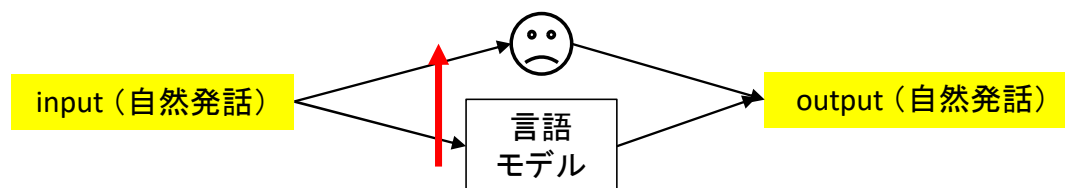
- 現在では、品詞解析, 構文解析を途中で挟まず、入力から直接出力を得るアプローチが主流。

深層学習ベースの自然言語処理と使用基盤モデルの類似点

- 発話された文(uttered sentences)が重要なインプットになる。
- 言語使用から立ち現れる言語知識を想定する。
- 規則は前もって与えられず、使用の中から学習する。
- 繰り返される言語使用からパターンが抽出されて言語知識となる
- 言語を理解・産出するための言語モデル(ニューラル・ネット)は脳を模倣して作られている。

認知言語学と言語処理の共通点

言語モデルは元々脳を模倣して作られたもの。さらに、inputが同じで、出来上がったものも似ているならば、その過程が似ていてもおかしくはないのでは？



こんなことはできないか？

1. 意味の重要な一部をなすとされる捉え方 (construal) を画像生成などで再現できないか。
2. 言語モデルの処理の過程にあり、特別な意味付けがなされていないものを有効活用できないか。
 - 単語や文を置き換えたベクトルや、それらをまとめたベクトルが、何らかの認知作用や言語学のメタ言語に対応をしないか？
3. 形は同じだが意味や機能が異なるというような構造を、クラスタ等を有効に利用することで、判別できないか？

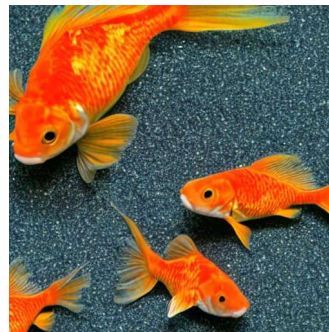
こんなことはできないか1？

語の意味の中に入り込んでいる話者の捉え方の問題

- 言語の中に入り込んだ主観的意味の裏付けAIによって生成された画像が使えないか？

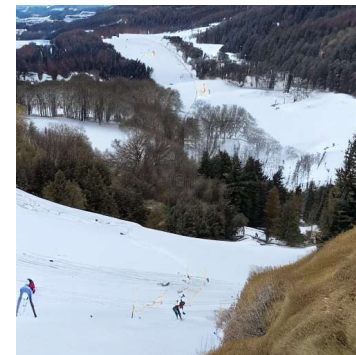
1. 百科事典的意味

- 金魚のイメージ



2. 視点:

- 上り坂 vs 下り坂



3. 概念が組み上がる過程の違い

- father's father vs grandfather



4. 心象

- swim the river vs swim across the river

Stable Diffusionで作成



AIによって生成された画像に捉え方が反映されていたら面白い

こんなことはできないか2？

単語や文を置き換えたベクトルの有効利用

- ベクトルは処理の過程にある、意味付けがなされていないもの
- 高次元のベクトルを、2~10程度に置き換えると、何らかの認知作用や言語学のメタ言語に対応をしないか？

言語学の考え方

- 言語の文法的な振舞いを動機づける重要な要因があるとされる
 - 有性性 (animacy) : 英語の代名詞
 - 他動性 (transitivity) : 高いと他動詞、低いと自動詞
 - 他動性は10のパラメーターからなるという理論 (Hopper and Thompson 1980)
 - 現実性 : 直説法と接続法
 - 複雑性と抽象性
- 少数に圧縮されたベクトルが、文法標識などのメタ言語や機能的動機づけ等と一致していると面白い

こんなことはできないか3？

クラスタの有効利用:

- すでに、意味変化や文法化の研究で使われ始めている。
- さらに応用をして、形は同じだが意味や機能が異なるというようなものを、自動で判別できないか？
- 用法の分類やカテゴリー化は言語学の基本なので、分類の妥当性にクラスタを使えないか？

例

- 両義的な構造
- 1. Flying planes can be dangerous (動名詞 vs 現在分詞)
 - 飛んでいる飛行機は危険だ
 - 飛行機を飛ばすことは危険だ
- 2. Tough構文
 - He is eager to please.
 - He is easy to please.
 - It is easy to please him.

注意点

チュートリアルSlackにあった質問

- **質問**: 仮に構文文法における何らかの仮説が、言語モデル等の学習過程で支持された場合、それで仮説が支持されたとみなすのでしょうか...? 何か考え方・立場? があれば知りたいです。
- **回答**: そうみなしてくれるといいのですが、恐らくそうはならないかと思います。おっしゃるように、ニューラルネットモデルの支持によって、理論の妥当性が多少増す程度かもしれませんが、理論を実証するためには、それ以外にも調査や実験を行い、それらが総合的に理論を支持していることを示す必要があるように思います。

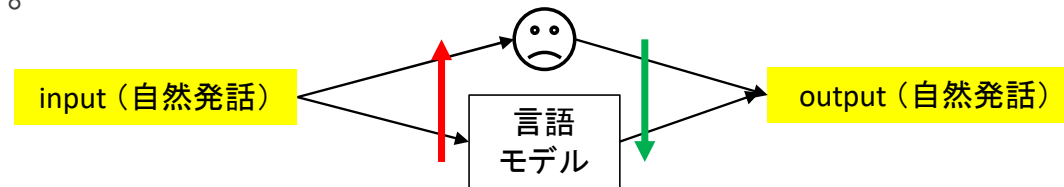
言語学に対する批判

- 本当に自然言語処理を生かそうと思うなら相手のことも知るべきだし、実際に手を動かすべき。
- 言語学者が言わなくても、自然言語処理の学者は言語学のことを十分に知っている。
- 言語学者は口だけ、言語の範囲を狭く限定する。
- 継続的に言語処理学会に参加するにはアップデートし続けることと、高い柔軟性が必要と感じた。

まとめ

まとめ

- 自然言語処理は人間の心から産出される言語を扱う学問。そのため、心のありようと無関係ではいられない。
- 言語モデルの様々な側面を言語知識や認知の処理に見立てることで、心を捉える手掛かりが得られないであろうか。



- 一方で、言語学から自然言語処理への貢献も考えられる。例えば、言語学者の知識は、自然言語のアノテーションの改良などに大いに役立つ可能性がある。
- 言語学も一枚岩ではなく、様々な立場がある(生得性重視vs 学習重視、用法基盤モデルvs辞書文法モデル)。言語学と自然言語処理の研究者が、問題意識を共有し、互いに対する理解を深め、相互に発展するテーマを見つけられるきっかけとなる場に、本ワークショップがなると嬉しい。

謝辞

ご清聴ありがとうございました。
質問・コメント・ご批判がありましたらお知らせください。

大谷直輝 (otani@tufs.ac.jp)

本発表は、科研費(19K00657)の助成を受けています。

中国語の部屋

中国語の部屋 (Searle, John (1980) Minds, Brains, and Programs より)

- ある部屋の中に、アルファベットしか理解できない人を閉じこめておく。この部屋には外部と紙きれのやりとりをするための小さい穴がひとつ空いており、この穴を通して英国人に1枚の紙きれが入れられる。そこには彼が見たこともない文字(=漢字)が並んでいる。英国人に見れば、それは「△☆□」といった記号の羅列にしか見えない。彼の仕事はこの記号の列に対して、新たな記号を書き加えてから、紙きれを外に返すことである。どういう記号の列に、どういう記号を付け加えればいいのか、それは部屋の中にある1冊のマニュアルの中に全て書かれている。例えば、「★△◎▽☆□」と書かれた紙片には「■@◎▽」と書き加えてから外に出せ、などと書かれている。彼はこの作業をただひたすら繰り返す。
- 部屋の外にいる人間は、小部屋の中には中国語を理解している人がいると考える。しかし、小部屋の中には漢字が読めず、作業の意味を理解していない英国人がいるだけである。それでも部屋の外部から見ると、中国語による対話が成立している。

wikipediaより

解釈

1. 人工知能における解釈: 強い(=人間の脳を模倣した)人工知能は検証不可能である。
2. 言語学における解釈: 統語論は意味論を含まない。(統語論の自律性)

「中国語の部屋」に基づくと、言語知識の近似値として言語モデルを利用する発想は起こらない。