

# 分布意味論と形式意味論の合流

谷中 瞳

東京大学, 理化学研究所

2025/08/26

人工知能基本問題研究会 (SIG-FPAI)@富山

# 自己紹介

- 東京大学大学院 情報理工学系研究科  
コンピュータ科学専攻 谷中研究室  
准教授(卓越研究員)  
理化学研究所 革新知能統合研究センター  
説明可能AIチーム チームディレクター
- 研究分野: 自然言語処理、計算言語学
- 研究キーワード: 意味解析、自然言語推論
- 近著: 「ことばの意味を計算するしくみ」講談社
  - ・ことばの意味を形式的にとらえる計算言語学
  - ・ことばの意味を統計的にとらえる自然言語処理ことばの意味を計算する2つのアプローチの可能性と課題を整理  
<https://www.kspub.co.jp/book/detail/5369845.html>



生成AIを作る人使う人はもちろん、  
「ことばを数学する」と聞いて心がざわつくような  
すべての人に届けたい。

言語処理学会前会長 乾健太郎先生 推薦!  
講談社

# 私たちがことばの意味について判断できること

「太郎しか泳がなかった」という文に対して…

- 「太郎だけが泳いだ」は同じ意味(同義)
- 「太郎だけが泳がなかった」は反対の意味(反義)
- 「太郎だけが平泳ぎをした」と近い意味(類義)
- 「太郎しかが泳いだ」はおかしな意味(容認不可能)



母語話者は文を与えられたとき、その文が容認可能である(文として理解できる)か否かを、即座に判断できる  
容認可能な文については、その意味も即座に理解できる

# 推論: ことばの意味に関する重要な判断

- 推論(inference)
  - ある命題(前提, premise)から別の命題(仮説, hypothesis)を導くこと
    - 命題(proposition): 真偽を問えるもの(文の意味)
- 前提が仮説を(意味論的に)含意する  
= 前提が真である状況のもとで、仮説が常に真となる

日本語自然言語推論(NLI)データセットJaNLI

[Yanaka&Mineshima,2021]

<https://huggingface.co/datasets/hpprc/janli>

前提P 子供が走っている猫を見ている

仮説H<sub>1</sub> 猫が走っている 含意

仮説H<sub>2</sub> 子供が走っている 非含意



# ことばの意味に関する原理: 合成性

- **合成性原理** (principle of compositionality) [Frege,1884]  
言語表現の意味はその構成素をなす表現の意味と、それらの結合方法によって定まるという原理
- 次の二つの文は同じ単語からなるが、意味は異なる
  - 子供が走っている猫を見ている
  - 猫が走っている子供を見ている
- 言語表現の意味は、語の意味の組み合わせだけでなく、統語構造による影響を受ける



# ことばの意味に関する考え方1: 使用説

[Wittgenstein, 1953]

ことばの意味に関する様々な事実を説明するような理論が、言語哲学では議論されてきた。その中の一つの説が**使用説**

- ことばの意味とは、そのことばの使用法
- ことばの意味がわかるとは、そのことばの使用法がわかるということ

L. Wittgenstein, 2009, *Philosophical Investigations*

(鬼界彰夫訳, 哲学探究, 2020):

「意味」という語が使われる多くの場合に—すべての場合ではないとしても—この語は次のように説明することができる。すなわち、語の意味とは、言語におけるその使われ方である、と。

# 分布意味論(distributional semantics)

分布仮説に基づいて、語の意味からことばの意味をとらえる  
**分布仮説**(distributional hypothesis) : 語の意味はその語の周辺に現れる語(文脈, context)によって定まる

Zellig Harris, 1954:

“oculist and eye-doctor… occur in almost the same environments. … If A and B have almost identical environments we say that they are synonyms.”

J. R. Firth, 1957:

“You shall know a word by the company it keeps!”

(使用説に対する)批判: 合成性(人はこれまで見たことのない文も、文の構成素と統語構造にしたがってその文を理解できる)をどのように説明するか?

# ことばの意味に関する考え方2: 真理条件説

[Davidson, 1967]

- ことばの意味とは、ことばの**真理条件** (truth condition, ことばがどのような状況では真となり、どのような状況では偽となるか)のこと
- ことばの意味がわかるとは、そのことばの真理条件がわかるということ

例) 文「東京は晴れ、かつ、埼玉は雨」の真理条件

この文が真であるのは、

「東京は晴れ」と「埼玉は雨」がどちらも真のときであり、

この文が偽であるのは、

「東京は晴れ」と「埼玉は雨」の少なくとも一方が偽のとき

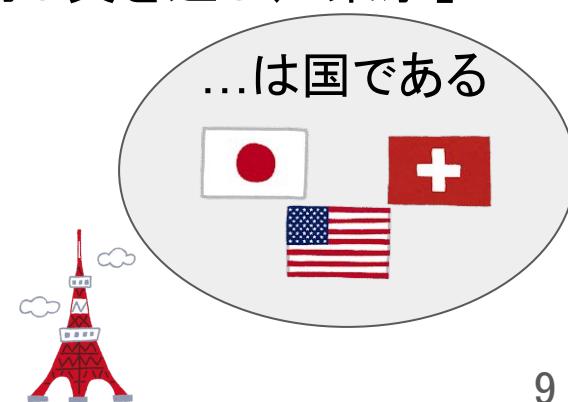
# 形式意味論(formal semantics)

- **真理条件的意味論** (truth-conditional semantics)に基づいて、文の意味からことばの意味をとらえる
- 定名詞句の意味を指示対象(個体)、述語の意味を指示対象の集合(個体に対し真理値を返す関数)として解釈し、語の意味から合成的に文の意味を説明する

例) 文「日本は国である」は真、文「東京は国である」は偽  
定名詞句「日本」「東京」: 個体を指示する  
述語「…は国である」: 「日本」や「スイス」という個体に対し真を返し、「東京」という個体に対し偽を返す関数  
(述語に個体を代入することで、文の真理値が定まる)

(真理条件説に対する)批判:

疑問文や命令文など真理値をもたない文の存在



# 意味表現、意味表示 (semantic representation)

分布意味論と形式意味論の合流：

ことばの意味を計算可能な形式（**意味表現、意味表示**）で表し、  
**推論**によって意味表現の妥当性を検証する

**自然言語処理** (Natural Language Processing) :

分布意味論に基づいてベクトル表現でことばの意味を分析

**計算言語学** (Computational Linguistics) :

形式意味論に基づいて論理表現などでことばの意味を分析

# ことばの意味を計算する2つのアプローチ

アプローチ	自然言語処理	計算言語学
意味の理論	分布意味論(使用説)	形式意味論(真理条件説)
意味の単位	語の意味	文の意味
意味表現	ベクトル表現	論理表現など

# 自然言語処理: 分布意味論に基づくアプローチ

- 大量のテキスト(コーパス)から統計的に文脈に基づいて語の意味を学習し、ベクトルで語の意味を表す
  - 伝統的な方法としては、前後に現れる語の共起頻度を要素とする行列を作り、単語ベクトルを構成
  - 意味が似ている語は、似たベクトルで表される

議論: 大規模言語モデル(LLM)は分布意味論に基づく?

[Enyan+2024]

- 訓練データの規模は分布意味論の定義に含まれない
- 指示チューニング(タスクの指示と入出力ペアの学習)は学習時のデータ分布を変える可能性はあるが、ことばの意味を言語使用の分布に基づいて表すという分布意味論の考え方から外れるものではない

## 語の意味から文の意味の計算へ：言語モデル

文(単語列)  $w_1, w_2, \dots, w_{i-1}$  の次に続く単語  $w_i$  の  
出現確率(確からしさ)  $P(w_1, w_2, \dots, w_i)$  を計算するモデル

$$P(\text{今日}, \text{の}, \text{天気}, \text{は}, \text{GPT}) = 0.0000003$$

$$P(\text{今日}, \text{の}, \text{天気}, \text{は}, \text{パンダ}) = 0.0000007$$

$$P(\text{今日}, \text{の}, \text{天気}, \text{は}, \text{晴れ}) = 0.0000127$$

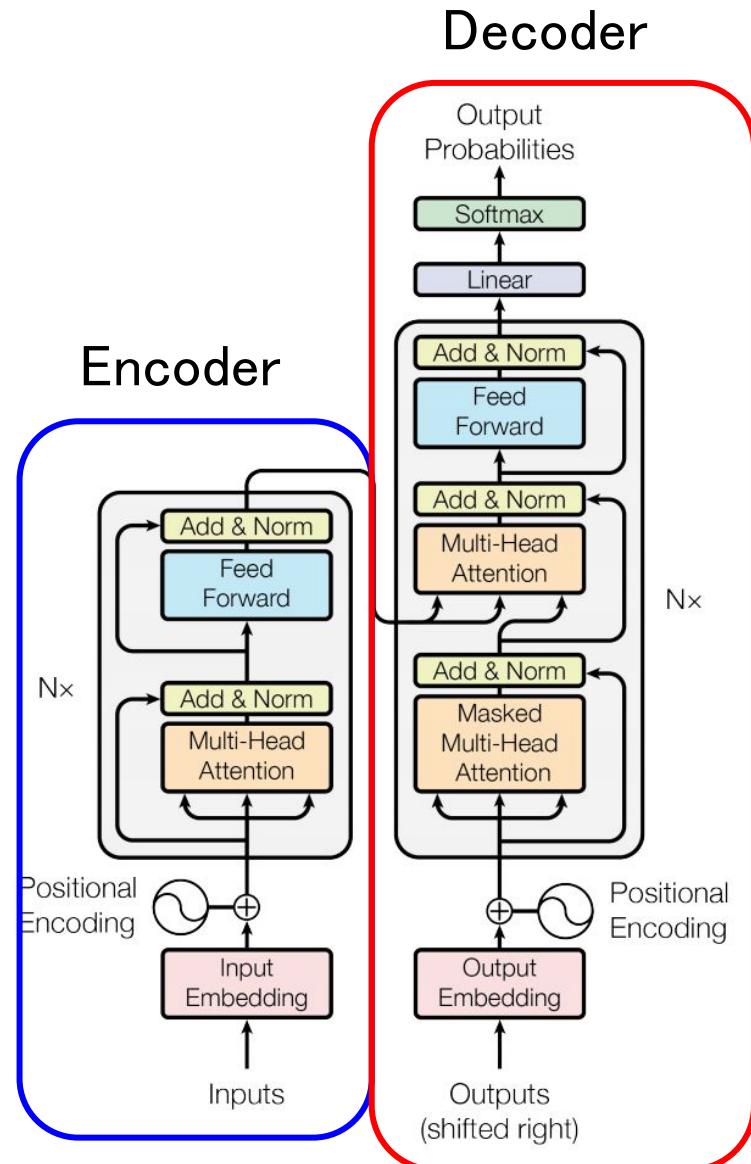
出現確率が高い文を自然な文として生成

→ 今日,の,天気,は,晴れ

# 大規模言語モデルの根幹: Transformer

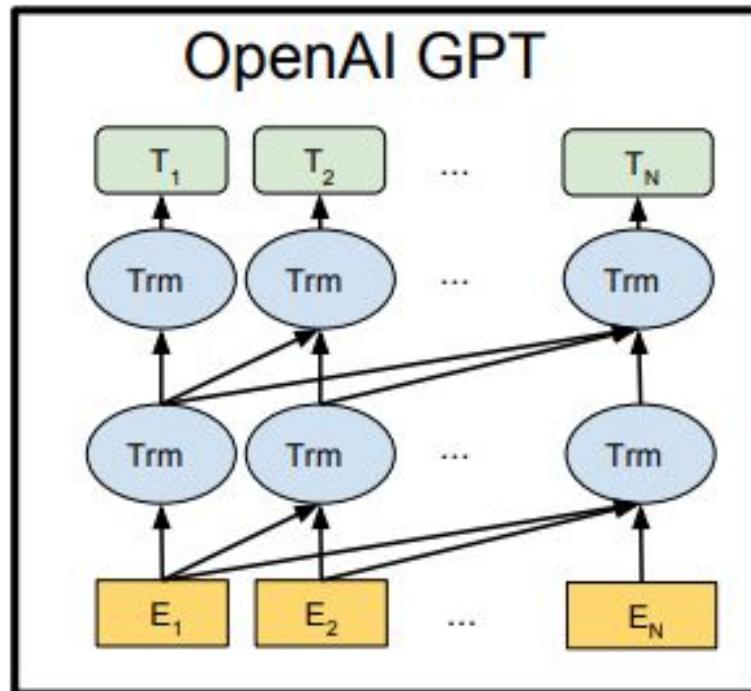
[Vaswani+ 2017]

- Attentionに基づく Encoder–Decoder (系列変換) モデル
- **Attention**: 入力系列中の重要な情報(どの単語に注目するか)を直接用いるしくみ
- **Encoder–Decoder**: 単語列を埋め込みベクトルに変換する Encoder モデルと、Encoder のベクトルを受け取り 1 単語ずつ生成する Decoder モデル
- 計算処理を並列化でき、計算効率が上がり、大規模言語モデル (LLM) の誕生へ



# GPT (Generative Pre-trained Transformer) [Radford+2018]

元祖・大規模コーパスによる事前学習に基づく大規模言語モデル  
Transformer Decoderモデルに基づく言語モデルで、単語列の次に続く  
単語の確率を計算



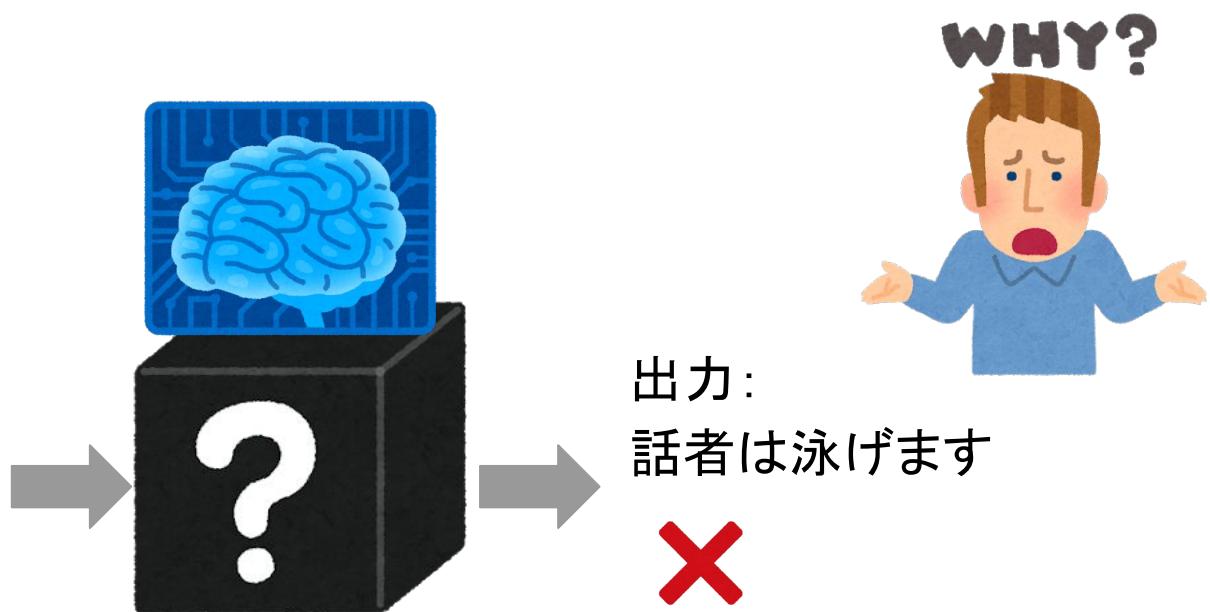
図は[Devlin+2019]から引用

# 分布意味論と形式意味論の合流1： 形式意味論に基づく言語モデルの分析

# 言語モデルの解釈性(interpretability)の問題

ベクトルによる意味表現やベクトル表現を用いた推論は原則ブラックボックスであり、どのようにことばの意味をとらえているのか説明が難しい

入力:  
「私は泳げなくないわけで  
はない」と発言した  
話者は泳げますか？



分布意味論と形式意味論の合流 1: 言語モデルの分析に形式意味論の知見を活用できなか?

# 言語モデルは合成性や体系性を獲得しているか

- **合成性** (compositionality)
  - 人はこれまで見たことのない文も、文の構成素と統語構造にしたがつてその文を理解できる
- **体系性** (systematicity) [Fodor and Pylyshyn, 1988]
  - 人はある処理ができれば、関連した処理もできる
  - 80–90年代にコネクショニズム、現在でいうニューラルネットを批判する文脈で提起された概念

John loves Annという文を理解できる人は、Ann loves Johnという文も理解できる。

どちらか一方だけ理解できる人は変？



## 単調性推論(monotonicity)

量化表現や否定表現などの単調性に基づき、文中の構成素を意味的に上位 / 下位の構成素に置き換えた文と、元の文との含意関係が成立する推論現象 [van Benthem,1983]

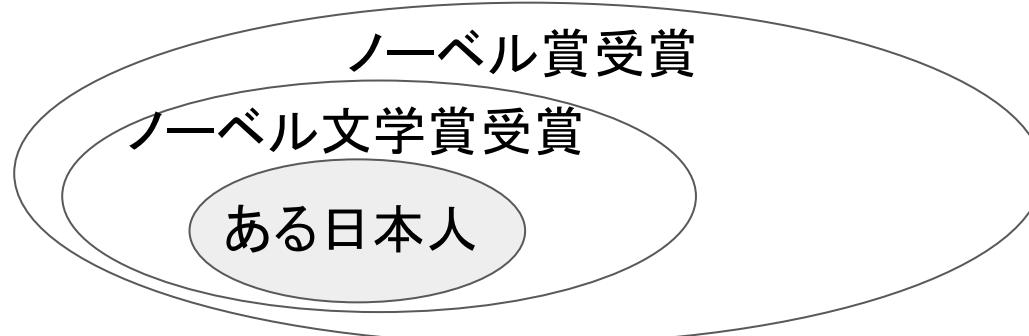
上方含意(Upward entailing)↑ :

意味的に下位の構成素から上位の構成素への含意関係が成立

前提P: ある日本人がノーベル文学賞↑を受賞した

仮説H: ある日本人がノーベル賞を受賞した

含意



# 単調性推論(monotonicity)

下方含意(Downward entailing) ↓ :

意味的に上位の構成素から下位の構成素への含意関係が成立

P: 今年日本人はノーベル賞↓を受賞しなかった

H<sub>1</sub>: 今年日本人はノーベル文学賞を受賞しなかった

含意

H<sub>2</sub>: 今年日本人はノーベル文学賞と化学賞を受賞しなかった

含意

単調性推論の特性:

- 量化表現や否定表現などの**単調性**と構成素の置き換えで含意関係が定まる
- 上位・下位の置き換えのしかた も様々な種類がある  
(語彙の置換・追加・削除)

日本人

ノーベル賞受賞

ノーベル文学賞受賞

# 文脈自由文法を用いた単調性推論データの自動構築

## 1. 文脈自由文法を用いて量化表現を含む前前提文を生成

生成規則  $N \rightarrow \{\text{dogs}\}$ ,  $IV \rightarrow \{\text{ran}\}$ ,  $TV \rightarrow \{\text{chased}\}$ ,  $Q \rightarrow \{\text{some}\}$ ,

$NP \rightarrow Q \ N \mid Q \ N \ Sbar$ ,  $S \rightarrow NP \ IV$ ,  $Sbar \rightarrow \text{which } TV \ NP$  から生成される文

Some dogs ran (再帰的規則の適用回数n=1のとき)

Some dogs which chased some dogs ran (n=2のとき)

Some dogs which chased some dogs which chased some dogs ran (n=3のとき)

## 2. 単調性に従い構成素の表現を置き換えて仮説文を生成

P: Some dogs ran      H: Some animals ran      含意

## 3. 前提文と仮説文を入れ換えて新たな推論ペアを生成

P': Some animals ran      H': Some dogs ran 非含意

形式意味論における多様な構文・意味に関する知見が、  
自然言語の評価・学習データセットの構築に活用できる

# 単調性推論に基づく言語モデルの体系性の評価

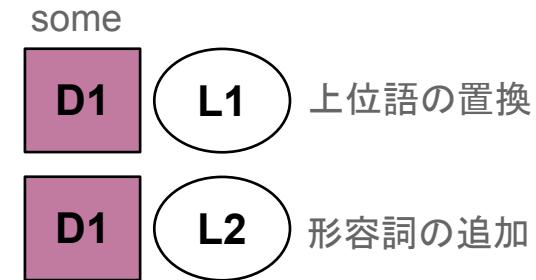
[Yanaka+2020]

**学習** (i. の量化表現の種類を段階的に増やしデータを追加)

i. 量化表現1種類 × 置き換え全種類の組み合わせ

P: **Some** dogs ran. H: **Some** animals ran. 含意

P: **Some** wild animals ran. H: **Some** animals ran. 含意



ii. 量化表現全種類 × 置き換え1種類の組み合わせ

P: **A** dog ran. H: **An** animal ran 含意

P: **No** animal ran. H: **No** dog ran 含意



**評価** 量化表現 × 置き換えの未知の組み合わせ

P: **A** wild animal ran. H: **An** animal ran. 含意

P: **No** animal ran. H: **No** wild animal ran. 含意



# 単調性推論に基づく言語モデルの体系性の評価

**学習** (i. の量化表現の種類を段階的に増やしデータを追加)

i. 量化表現1種類 × 置き換え全種類の組み合わせ

P: **Some** dogs ran. H: **Some** animals ran. 含意

P: **Some** wild animals ran. H: **Some** animals ran. 含意

ii. 量化表現全種類 × 置き換え1種類の組み合わせ

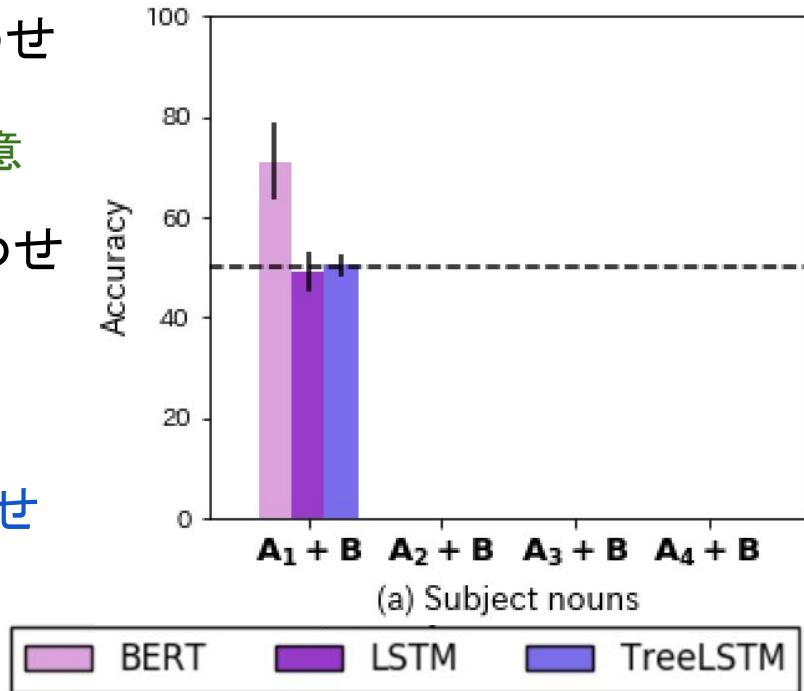
P: **A** dog ran. H: **An** animal ran. 含意

P: **No** animal ran. H: **No** dog ran. 含意

**評価** 量化表現 × 置き換えの未知の組み合わせ

P: **A** wild animal ran. H: **An** animal ran. 含意

P: **No** animal ran. H: **No** wild animal ran. 含意



BERTは量化表現と置き換えの未知の組み合わせからなる推論  
に対して一部汎化、LSTMやTreeLSTMは汎化しない

# 単調性推論に基づく言語モデルの体系性の評価

**学習** (i. の量化表現の種類を段階的に増やしデータを追加)

i. 量化表現1種類 × 置き換え全種類の組み合わせ

P: **Some** dogs ran. H: **Some** animals ran. 含意

P: **Some** wild animals ran. H: **Some** animals ran. 含意

ii. 量化表現全種類 × 置き換え1種類の組み合わせ

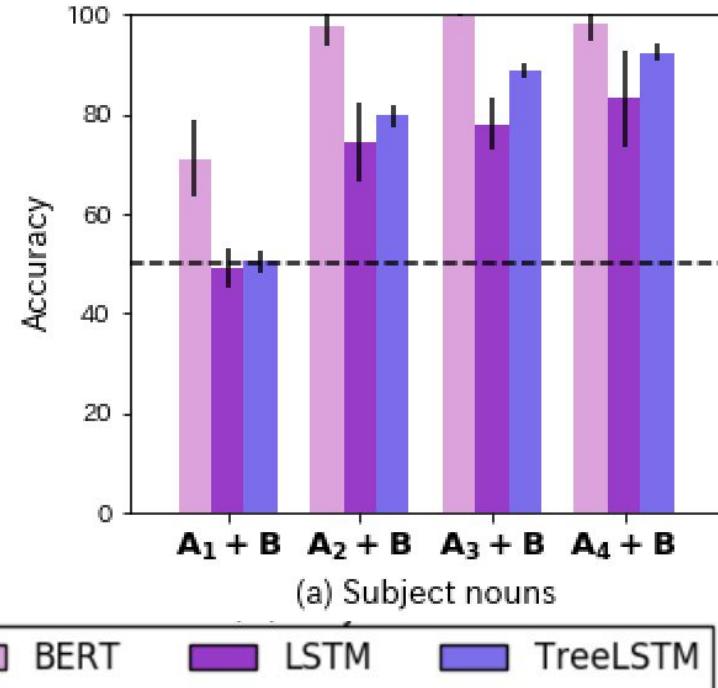
P: **A** dog ran. H: **An** animal ran. 含意

P: **No** animal ran. H: **No** dog ran. 含意

**評価** 量化表現 × 置き換えの未知の組み合わせ

P: **A** wild animal ran. H: **An** animal ran. 含意

P: **No** animal ran. H: **No** wild animal ran. 含意



BERTは量化表現と置き換えの未知の組み合わせからなる推論  
に対して一部汎化、LSTMやTreeLSTMは汎化しない

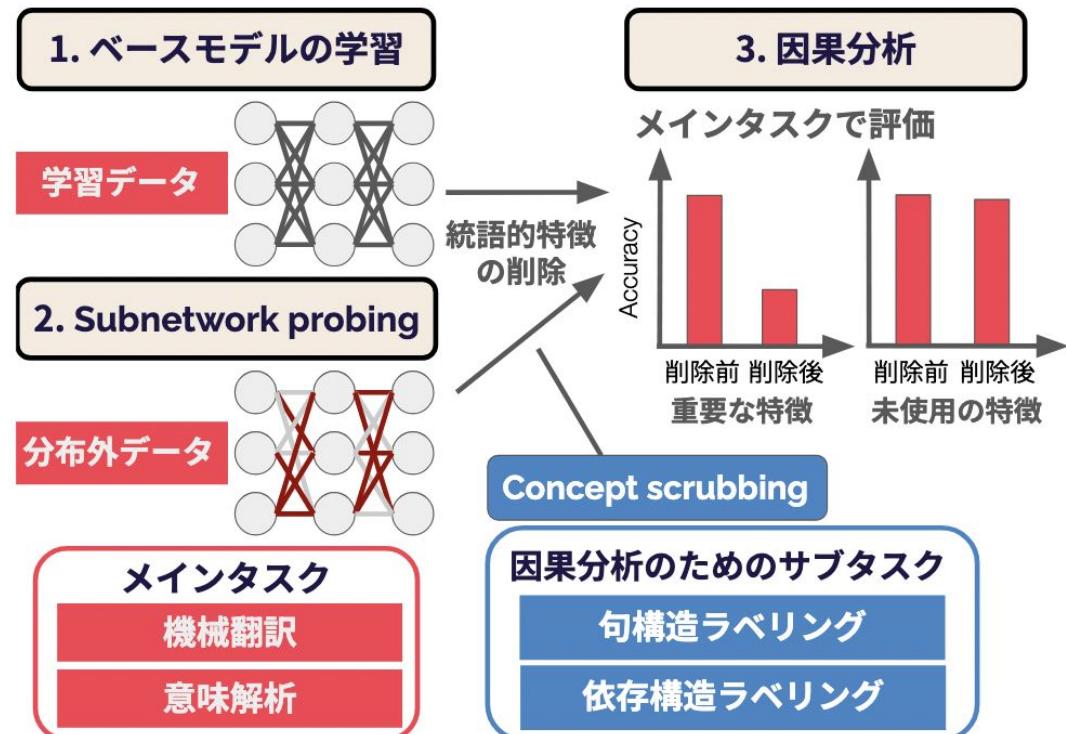
i.のデータを段階的に増やしていくとLSTMやTreeLSTMも性能向上:データ拡張による部分的な改善の可能性

# 合成性に関するTransformerの内部機序の分析

[Kumon and Yanaka, NAACL2025]

合成的汎化に寄与するサブネットワークの探索と、統語情報の活用に関する因果分析により、Transformerの内部機序を分析

- 合成的汎化能力の高いサブネットワークの存在を実証的に発見
- 一方、見つかったサブネットワークは統語情報を用いた合成的な解法だけでなく、非合成的な解法にも依存する傾向



# 分布意味論と形式意味論の合流2： 意味解析と推論

# 分布仮説はことばの意味に対して常に成り立つか

否定、量化、数量、時制、比較、照応などの機能語(文法的な役割を果たす語↔内容語)の意味は分布仮説に従うのか、LLMで体系的に捉えられるのか自明でない

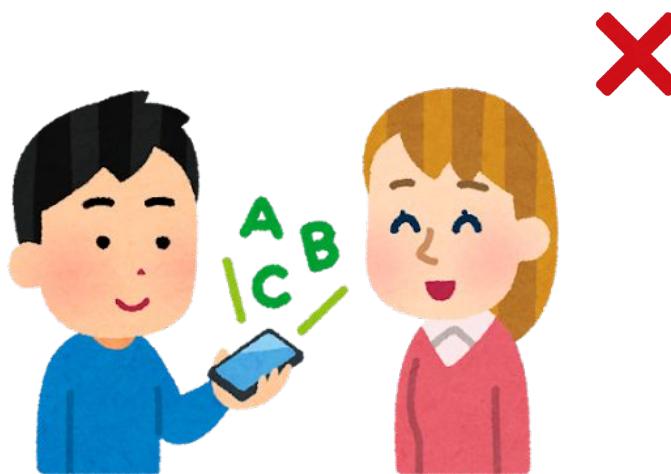
## 翻訳の例

日本語

私は泳げなくないわけではない

英語

Not that I can't swim.



# LLMが課題とする様々な推論現象

## 事実性推論 (veridical inference) [Ross&Pavlick19][Yanaka+24]

P: 太郎は二郎が来たと気づいていた	H: 二郎が来た	含意
P: 太郎は二郎が来たと信じていた	H: 二郎が来た	非含意

## 時間推論 (temporal inference) [Thukral+21][杉本+2024]

P: 車は2時間で到着した	H: 車は3時間以内に到着した	含意
P: 車は2時間で到着した	H: 車は3時間で到着した	非含意

## 比較表現 (comparative) [Haruta+20]

P: 太郎は二郎より多くの本を買った	H: 太郎は何冊かの本を買った	含意
P: 太郎は二郎より多くの本を買った	H: 二郎は何冊かの本を買った	非含意

分布意味論と形式意味論の合流 2: 形式意味論に基づいてことばの意味を解析し推論できないか？

# 形式意味論の構成

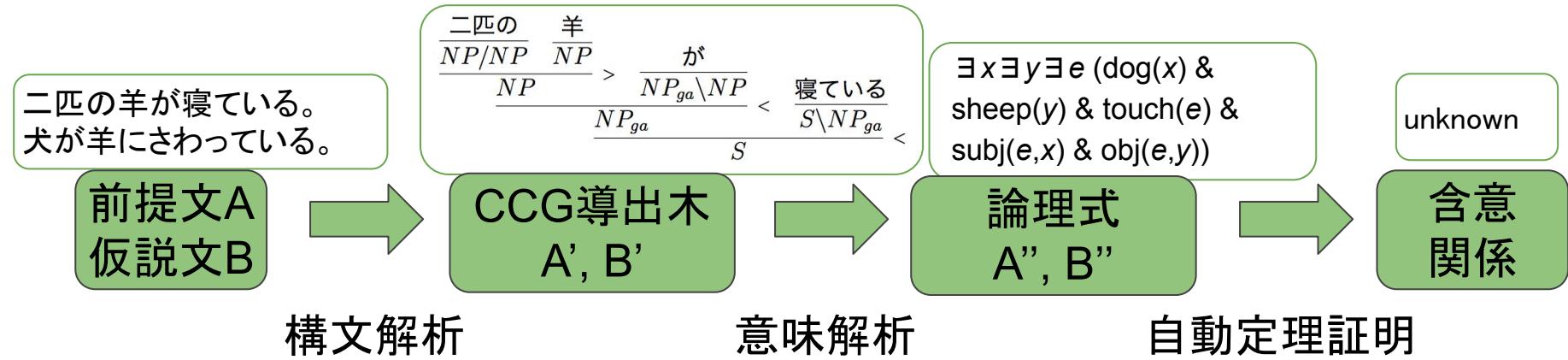
[戸次 (2017)「形式意味論」, 人工知能学事典]

1. 自然言語の意味とは何か、という問題について何らかの立場を採り、意味を表示するための**形式言語を定義** する
2. 自然言語の**統語論**について、何らかの立場を採る
3. 与えられた言語の与えられた文もしくは談話について、2. の**統語論が与える構造にしたがって、1. による意味表示**（意味表現）**を計算**する機構を与える
4. その言語の個々の文または談話について、3. によって**予測**される意味と、母語話者の言語直観と**を照合**することにより経験的検証を行う。誤りは1.2.3. に遡り(通常は3. を)修正する

# 形式意味論に基づく意味解析・論理推論システム

## ccg2lambda[Mineshima+2015][Yanaka+2018] :

# 組合せ範疇文法に基づく意味解析・論理推論システム



関連する意味解析・論理推論システム:LangPro[Abzianidze+2017],

MonaLog[Hu+2020], lightblue[Tomita+2024],

**NeuralLog**[Zhen+2021], **Hy-NLI**[Kalouli+2020], **UDepLambda**[Reddy+2017],

Boxer[Bos,2015]など

# 組合せ範疇文法[Steedman, 2000]

- Combinatory Categorial Grammar(CCG)  
統語構造と意味解釈の対応が明示的かつ簡潔な文法理論
- 語の語彙的・文法的な性質を**統語範疇**として辞書に記述し、言語普遍的な性質を少数の**統語規則**で記述する**語彙化文法**の一つ
- 統語範疇の定義：
  - $S, NP, N$ は統語範疇である。これを**基底範疇**という。
  - $X$ と $Y$ が統語範疇ならば、 $X/Y$ と $Y\neq X$ も統語範疇である。これを**関数型範疇**という。
- $S$ (文),  $NP$ (名詞句),  $N$ (名詞)

# 関数適用規則と関数合成規則

CCGの代表的な統語規則(組合せ規則)

- 関数適用規則

$$\frac{X/Y : f \quad Y : a}{X : fa} > \qquad \frac{Y : a \quad X \setminus Y : f}{X : fa} <$$

- 関数合成規則

$$\frac{X/Y : f \quad Y/Z : g}{X/Z : \lambda x. f(gx)} >_{\mathbf{B}} \qquad \frac{Y \setminus Z : g \quad X \setminus Y : f}{X \setminus Z : \lambda x. f(gx)} <_{\mathbf{B}}$$

# ccg2lambdaを用いた含意関係の証明:(i) 構文解析

文A: Some cats are runningと文B: Some animals are runningの意味表示を構文解析・意味解析により導出し、含意関係を示す

A: *Some cats are running.*

$$\frac{\frac{\text{Some}}{NP/N} \quad \frac{\text{cats}}{N} > \frac{\text{are}}{(S \setminus NP)/(S \setminus NP)} \quad \frac{\text{running}}{S \setminus NP}}{\frac{NP}{S}} <$$

B: *Some animals are running.*

$$\frac{\frac{\text{Some}}{NP/N} \quad \frac{\text{animals}}{N} > \frac{\text{are}}{(S \setminus NP)/(S \setminus NP)} \quad \frac{\text{running}}{S \setminus NP}}{\frac{NP}{S}} <$$

# 範疇文法と論理の関係: ランベック計算

ランベック計算[Lambek, 1958]: 理論言語学における統語構造の導出過程(導出木)と、記号論理学における論理式の証明過程(証明図)を同型と見なせることを提唱

範疇文法	論理
統語範疇	命題(論理式)
基底範疇	原始命題
関数型範疇	条件命題
統語規則	推論規則
関数適用規則	含意除去則
導出木	証明図

# 範疇文法と論理の関係: ランベック計算

範疇文法の関数適用規則は、前提が現れる順序を気にする論理の含意除去則に対応

- 関数適用規則

$$\frac{X/Y : f \quad Y : a}{X : fa} > \quad \frac{Y : a \quad X \setminus Y : f}{X : fa} <$$

- 含意除去則

前提  $\frac{X/Y \quad Y}{X} /E$  結論  $\frac{Y \quad Y \setminus X}{X} \setminus E$

ランベック計算の $X/Y$ は $X \leftarrow Y$ 、 $Y \setminus X$ は $Y \rightarrow X$ と同じ意味  
(CCGの $X \setminus Y$ はランベック計算の $Y \setminus X$ に対応する点に注意)

# CCGに基づく構文解析と意味解析

- 辞書:語の統語範疇と意味表示を同時に指定(語彙項目)
- 組合せ規則:統語構造の構成方法と意味合成の計算方法を同時に指定
- 様々なCCG構文解析器が現在も研究されている
  - 英語:C&C[Clark and Curran, 2007], HoLCCG[Yamaki+, 2023]
  - 日本語:depccg[Yoshikawa+, 2017], lightblue[Tomita+, 2024]

## 辞書の例

語	統語範疇	意味表示
John	$NP$	$John$
love	$(S \setminus NP)/NP$	$\lambda y. \lambda x. love(x, y)$

語の意味は  
ラムダ式で記述  
(後で説明!)

# 意味表示の型

**型つきラムダ計算** : モンタギュー意味論[Montague, 1973]にはじまる形式意味論の、標準的な記述言語。文法理論と意味理論のシンプルで統一的な計算体系の一つ

## 意味表示の型の定義

- $e$ と $t$ は型である。 $e$ は個体(entity)、 $t$ は真理値(type)の型を表す。
- $\alpha$ と $\beta$ が型ならば $\alpha \rightarrow \beta$ も型である。
- それ以外は型ではない。

例	語	意味表示	型
	John	<i>John</i>	$e$
	run	$\lambda x. run(x)$	$e \rightarrow t$

# 型に基づく意味合成

意味合成では、各構成素の意味表示と型、それらの合成によって得られる意味表示と型が整合している必要がある

例: John runs の意味合成

runs

型:  $e \rightarrow t$

意味表示:  $\lambda x. run(x)$

John

型:  $e$

意味表示: *John*

John runs

型:  $t$

意味表示:  $(\lambda x. run(x))(John) = run(John)$

β簡約で意味合成を行う

# 統語範疇と意味表示の型の対応づけ

- CCGの統語範疇と組合せ規則は統語構造の構成方法と意味合成の計算方法を同時に指定している
- 統語範疇から意味表示の型への写像を定義する

$$S^* = t$$

$$NP^* = e$$

$$N^* = e \rightarrow t$$

$$(Y/X)^* = (X)^* \rightarrow (Y)^*$$

$$(X\backslash Y)^* = (Y)^* \rightarrow (X)^*$$

例: 自動詞

$$(S \backslash NP)^* = NP^* \rightarrow S^* = e \rightarrow t$$

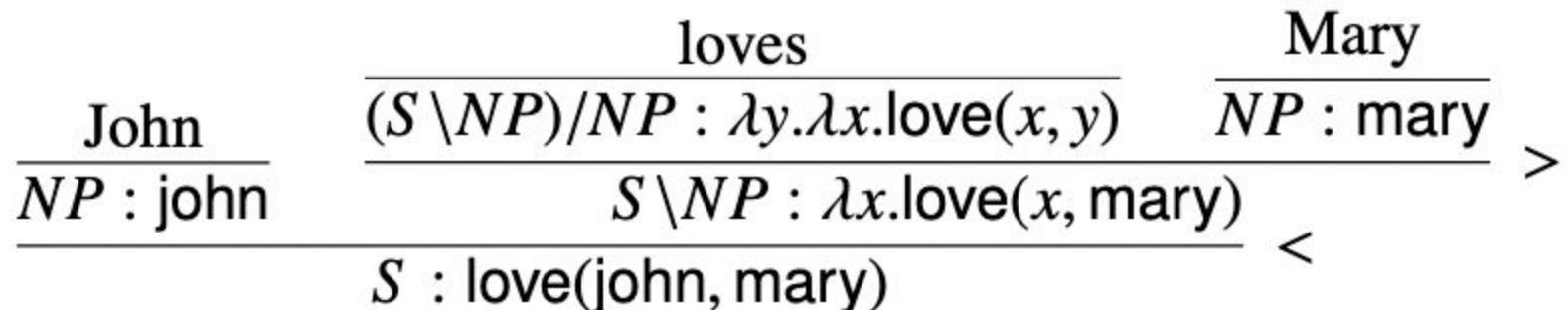
他動詞  $((S \backslash NP)/NP)^* = NP^* \rightarrow (S \backslash NP)^* = e \rightarrow e \rightarrow t$

# CCGと型つきラムダ計算に基づく意味合成:まとめ

## 語彙項目の例

語	統語範疇	意味表示	意味表示の型
John	$NP$	$John$	$e$
love	$(S \setminus NP)/NP$	$\lambda y. \lambda x. love(x, y)$	$e \rightarrow e \rightarrow t$

## John loves Mary のCCG導出木と意味合成の例



## (ii) CCGに基づく意味解析

文A: Some cats are runningと文B: Some animals are runningの意味表示を構文解析・意味解析により導出し、含意関係を示す

A: *Some cats are running.*

$$\frac{\begin{array}{c} \text{Some} \\ \hline \lambda F G. \exists x. (F(x) \wedge G(x)) \\ NP/N \end{array} \quad \frac{\begin{array}{c} \text{cats} \\ \hline \lambda x. \text{cat}(x) \\ N \end{array}}{>} \quad \frac{\begin{array}{c} \text{are} \\ \hline \lambda X. X \\ (S \setminus NP)/(S \setminus NP) \end{array}}{>} \quad \frac{\begin{array}{c} \text{running} \\ \hline \lambda Q. Q(\lambda x. \exists e. (\text{run}(e) \wedge (\text{subj}(e) = x))) \\ S \setminus NP \end{array}}{>}}{\lambda G. \exists x. (\text{cat}(x) \wedge G(x)) \\ NP} < \frac{\lambda Q. Q(\lambda x. \exists e. (\text{run}(e) \wedge (\text{subj}(e) = x)))}{\lambda Q. Q(\lambda x. \exists e. (\text{run}(e) \wedge (\text{subj}(e) = x))) \\ S \setminus NP} < \\
 \frac{}{\exists x \exists e. (\text{cat}(x) \wedge \text{run}(e) \wedge (\text{subj}(e) = x)) \\ S}$$

B: *Some animals are running.*

$$\frac{\begin{array}{c} \text{Some} \\ \hline \lambda F G. \exists x. (F(x) \wedge G(x)) \\ NP/N \end{array} \quad \frac{\begin{array}{c} \text{animals} \\ \hline \lambda x. \text{cat}(x) \\ N \end{array}}{>} \quad \frac{\begin{array}{c} \text{are} \\ \hline \lambda X. X \\ (S \setminus NP)/(S \setminus NP) \end{array}}{>} \quad \frac{\begin{array}{c} \text{running} \\ \hline \lambda Q. Q(\lambda x. \exists e. (\text{run}(e) \wedge (\text{subj}(e) = x))) \\ S \setminus NP \end{array}}{>}}{\lambda G. \exists x. (\text{animal}(x) \wedge G(x)) \\ NP} > \frac{\lambda Q. Q(\lambda x. \exists e. (\text{run}(e) \wedge (\text{subj}(e) = x)))}{\lambda Q. Q(\lambda x. \exists e. (\text{run}(e) \wedge (\text{subj}(e) = x))) \\ S \setminus NP} > \\
 \frac{}{\exists x \exists e. (\text{animal}(x) \wedge \text{run}(e) \wedge (\text{subj}(e) = x)) \\ S}$$

### (iii) 定理証明器を用いた自動推論

文間の含意関係を定理証明器(Coq, Vampireなど)で自動証明

- 常識的知識や世界知識は公理を用いて補完
- 証明過程を観測でき、解釈性がある

前提文A: *Some cats are running.* 仮説文B: *Some animals are running.*

① A1:  $\exists e_1 \exists x_1 (\text{cat}(x_1) \wedge \text{run}(e_1) \wedge (\text{subj}(e_1)=x_1))$

B1:  $\exists e_2 \exists x_2 (\text{animal}(x_2) \wedge \text{run}(e_2) \wedge (\text{subj}(e_2)=x_2))$



存在量化子(ヨ)の除去

② A2:  $\text{cat}(x_1) \wedge \text{run}(e_1) \wedge (\text{subj}(e_1)=x_1)$

B2:  $\text{animal}(x_2) \wedge \text{sing}(e_2) \wedge (\text{subj}(e_2)=x_2)$



連言(∧)の除去

③ A3:  $\text{cat}(x_1)$ , A4:  $\text{run}(e_1)$ , A5:  $\text{subj}(e_1)=x_1$

B3:  $\text{animal}(x_2)$ , B4:  $\text{run}(e_2)$ , B5:  $\text{subj}(e_2)=x_2$



変数の单一化

公理  $\forall x. \text{cat}(x) \rightarrow \text{animal}(x)$  補完

サブゴールの除去

④ A3:  $\text{cat}(x_1)$ , A4:  $\text{run}(e_1)$ , A5:  $\text{subj}(e_1)=x_1$   
~~B3:  $\text{animal}(x_2)$ , B4:  $\text{run}(e_2)$ , B5:  $\text{subj}(e_2)=x_2$~~

# 最近の研究：日本語比較表現の推論[Mikami+NALOMA2025]

日本語意味論テストセットJSeM[Kawazoe+2021]の比較表現セクション71件においてGPT-4oの正答率76.0%よりも高い正答率84.5%を達成

前提を伴う比較表現の推論を扱えるよう、文全体の意味表示と前提の意味表示の組で表すmultidimensional semantics[Potts,2004]を導入

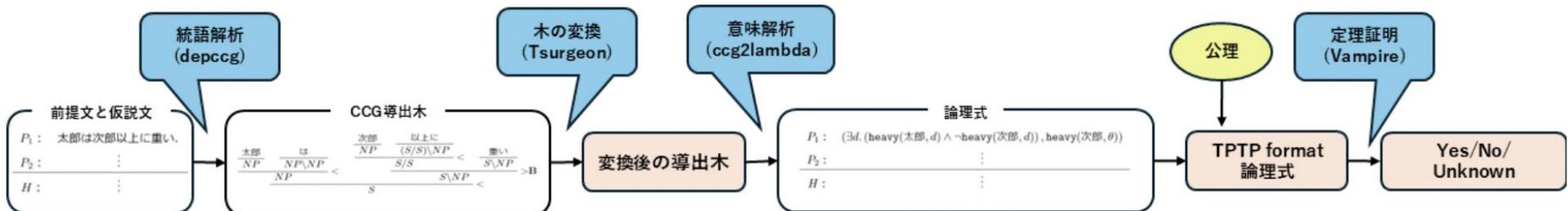
P: 太郎は次郎以上に重い

( $\exists d. (\text{heavy}(\text{太郎}, d) \wedge \neg \text{heavy}(\text{次郎}, d)),$   
 $\text{heavy}(\text{次郎}, \theta))$

H: 次郎は重い

( $\text{heavy}(\text{次郎}, \theta),$   
 $\top)$

含意



# ことばの意味を計算する2つのアプローチ

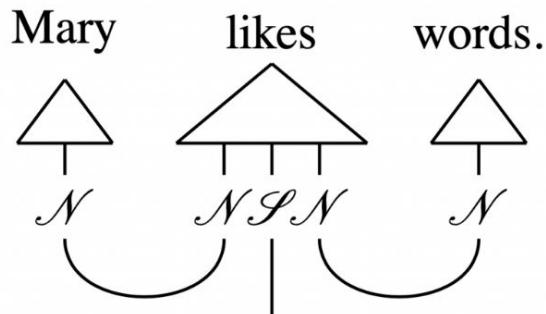
分野	自然言語処理	計算言語学
意味の理論	分布意味論(使用説)	形式意味論(真理条件説)
意味の基準	文脈	文
意味表現	ベクトル表現	論理表現など
利点	<ul style="list-style-type: none"><li>・連続値・離散値問わず</li><li>end-to-endで出力</li><li>・内容語の意味の計算</li></ul>	<ul style="list-style-type: none"><li>・推論過程を解釈しやすい</li><li>・未知のデータに適応可</li><li>・機能語の意味の計算</li></ul>
課題	<ul style="list-style-type: none"><li>・機能語の意味の計算</li><li>・未知のデータへの適応</li><li>・中身がブラックボックス</li></ul>	<ul style="list-style-type: none"><li>・内容語の意味の計算</li><li>・連続値の扱い</li><li>・頑健なパイプライン処理</li></ul>

アプローチを組み合わせて相補的に課題を解決する可能性  
(cf. Neuro-Symbolic AI)

# 合成性を満たす分布意味論: 合成分布意味論

Compositional Distributional Semantics (CDS) [Coecke,2010]

- 文の統語構造を考慮して語の意味ベクトルを合成
  - 他動詞*like*のような項を取る語の意味を、項と同じベクトル空間の元ではなく、ベクトル空間のテンソル積の元として表し、テンソルの縮約で項関係を考慮して意味合成
  - 型つきラムダ計算との対応で述べると: 関数の代わりにテンソル積、関数適用の代わりにテンソル縮約を用いる
- 実装では量子計算が用いられる(量子言語処理、QNLP)
  - DisCoPy: pythonツールキット <https://discopy.org/>



# LLMの演繹推論能力の改善

記号論理学の知見に基づいて演繹推論のデータを大量に自動構築し  
学習を行うこと[Morishita+2023,2024]や、テキストから論理式への翻訳を行  
わせてから推論を行うようプロンプトを与えること[Ozeki+2024]によって、  
LLMの演繹推論能力が向上

→同様のアプローチで、LLMの  
自然言語の演繹推論能力を  
どこまで改善できるか？

```
## Input
Premise 1: Some A are B.
Premise 2: All B are C.
Hypothesis: All A are C.

## Translation into predicate logic
Premise 1:  $\exists x(Ax \wedge Bx)$ 
Premise 2:  $\forall x(Bx \rightarrow Cx)$ 
Hypothesis:  $\exists x(Ax \rightarrow Cx)$ 

## Reasoning
[Explain your reasoning for the answer]

## Answer
[Your answer must be one word: "entailment", "contradiction", or "neither"]

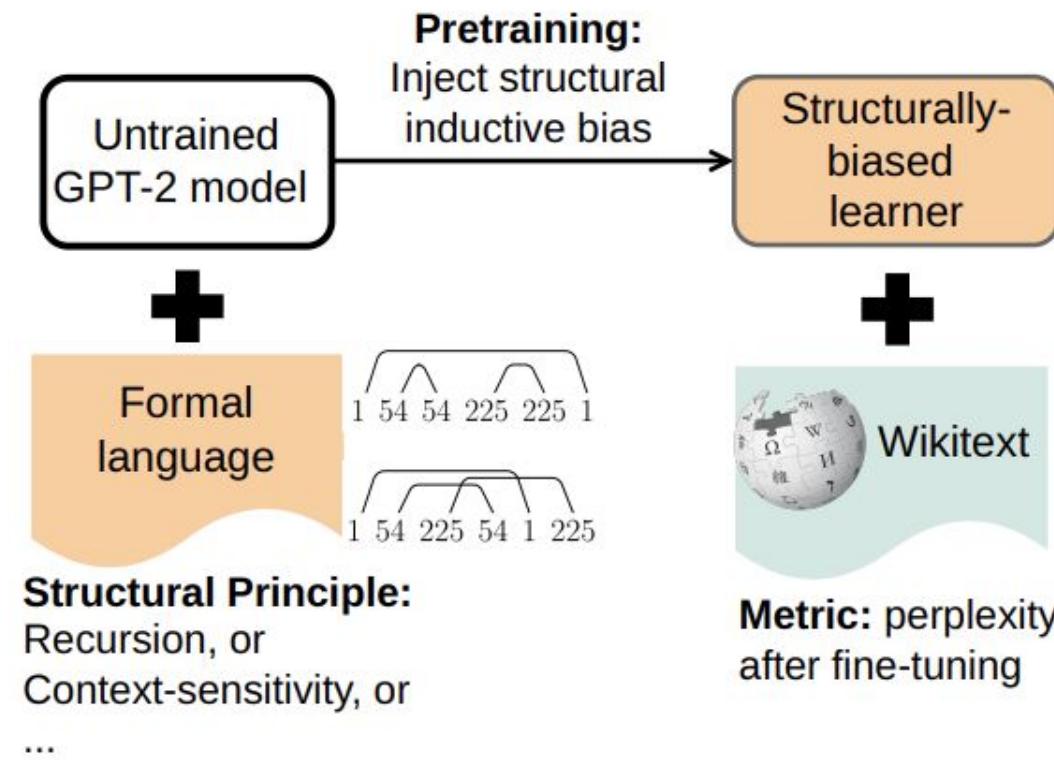
## Input
Premise 1: One friend of Taro is a friend of Paul.
Premise 2: All of Paul's friends are German.
Hypothesis: All of Taro's friends are German.

## Translation into predicate logic
```

# 形式言語を用いた事前学習

[Papadimitriou and Jurafsky,2023][Hu+.2025]

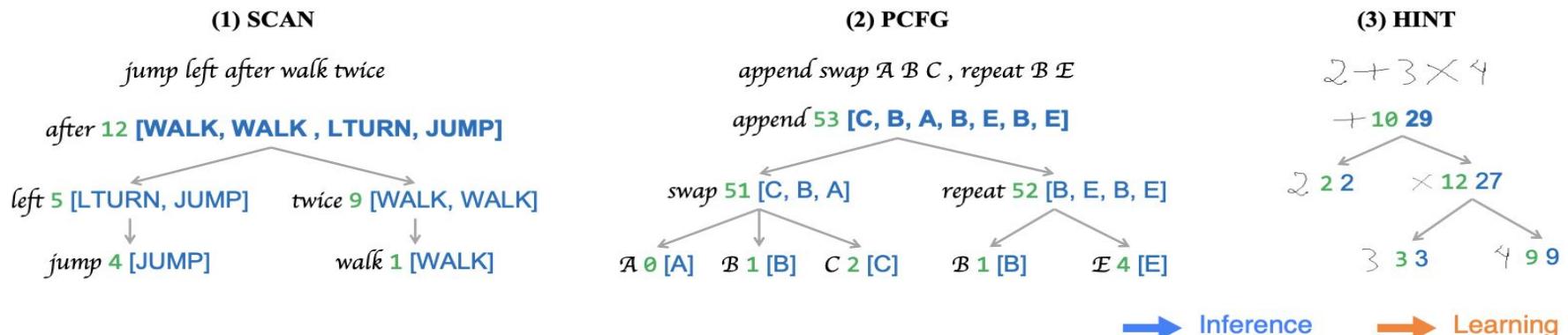
形式言語で再帰や文脈依存性を考慮し合成したデータを事前学習に用いることで、モデルの統語的な帰納的バイアスを制御して与えられる可能性



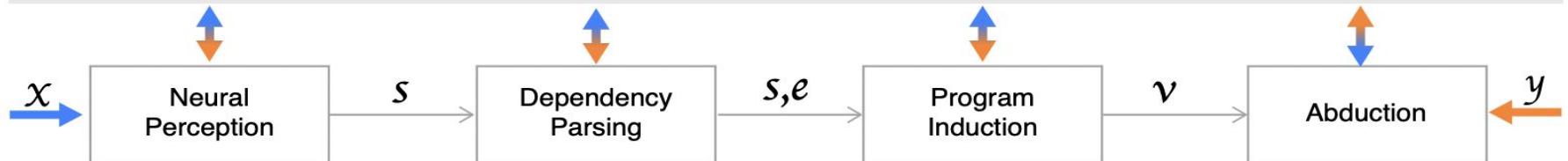
# Neuro-Symbolicアプローチによる体系性の改善

[Li+2024]

syntax-semantics interfaceを考慮した内部表現を設計し、  
SCAN[Lake and Baloni,2018]などの人工的な合成的汎化タスクで性能  
向上

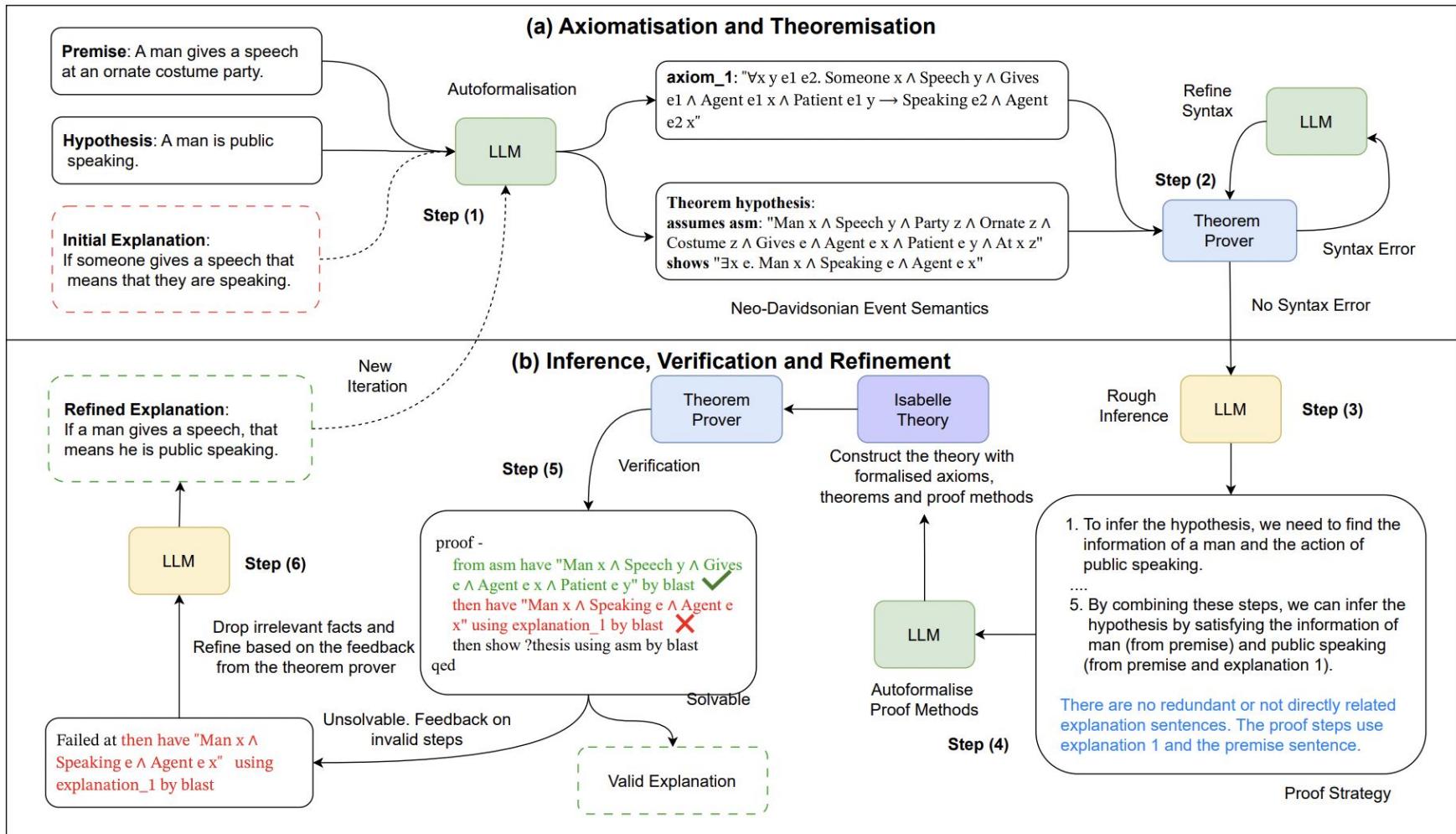


Grounded Symbol System  
 $\mathcal{T} = \langle (x, s, v), e \rangle$



# LLMと定理証明器の融合による、説明の検証と改良

[Quan+2024]



# おわりに：ことばの意味を多面的にみてみよう

- LLMの発展に伴い、他分野への自然言語処理応用や社会実装の幅が広がる
- 分布意味論と形式意味論の合流：
  - LLMの分析による言語理論の再検証
  - LLMの課題解決に向けた言語理論の活用
- 人がどのようにことばの意味を計算しているのかもまだ十分に解明されていない。ことばの意味の研究は、様々な分野と立場が関わる学際的な研究領域。自ら様々な立場から検討し反証する形で相互尊重することで、ことばの意味についての「真理」にさらに一步近づける可能性

ご清聴ありがとうございました！