Modeling Noisy-Channel Language Processing with Incremental and Approximate Probabilistic Inference

Thomas Hikaru Clark (thclark@mit.edu), Jacob Vigly, Roger Levy, Edward Gibson — MIT Brain and Cognitive Sciences

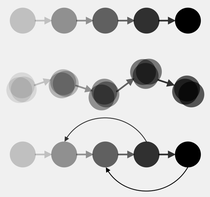
Background

Comprehension of Anomalous Utterances as Rational (Bayesian) Inference [1] [2] :

Utterance u : *the boy licked the ball .* **Intended meaning** w : ?

$$P(w \mid u) = \frac{P(w) \cdot P(u \mid w)}{\sum_{w'} P(w') \cdot P(u \mid w')}$$

- Comprehenders are sensitive to **prior probability** of intended meanings and …
- **Likelihood** of different **error operations** (e.g. insertions vs. deletions)
- **Limitations of current accounts**: level of analysis, cognitive plausibility, fine-grained predictions for arbitrary inputs



Our Goal

An implemented, algorithmic-level account of noisy-channel language comprehension that uses **incremental** and **approximate** inference and supports **reanalysis**.

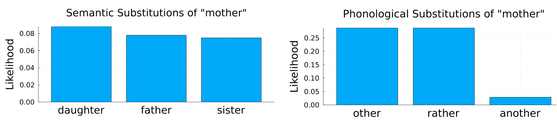
Generative Model and Inference Algorithm

Generative Model

▶ Language Model Prior (GPT-2, restricted vocabulary)

$$P(w_t \mid w_1, w_2, \ldots, w_{t-1})$$

▶ Error Model Likelihood (Normal, Substitutions, Deletions, Insertions)

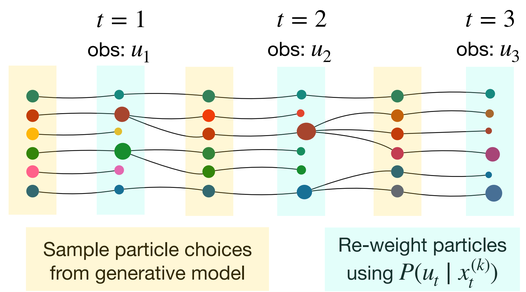


$$P(a_t \mid \texttt{action prior})$$
$$P(u_t \mid \mathbf{w}, a_t)$$

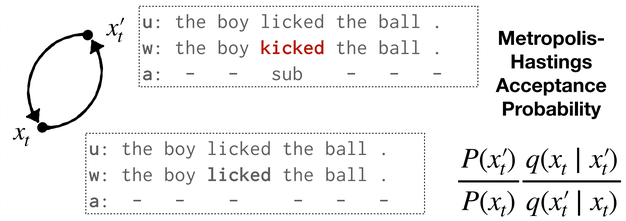
Incremental, Approximate Inference

Noisy-Channel Comprehension ≈ posterior inference for latent variables given observed utterance $u_1 \ldots u_t$

Particle Filter: Each **particle** $x_t^{(k)}$ represents a hypothesis about the intended sentence and errors



Reanalysis as Monte Carlo Rejuvenation



Metropolis-Hastings Acceptance Probability

$$\frac{P(x_t')}{P(x_t)} \frac{q(x_t \mid x_t')}{q(x_t' \mid x_t)}$$

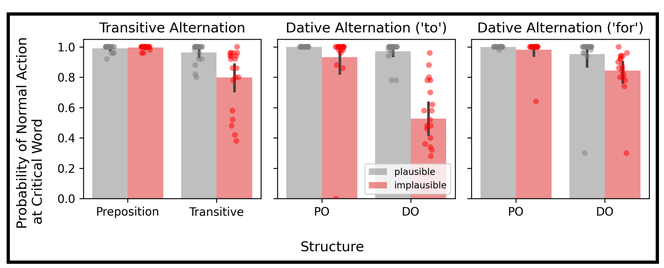
- Provides a mechanism for reanalysis of earlier commitments
- Revised words are not necessarily incrementally surprising words
- Prediction: readers may regress to regions where rejuvenation acceptance is more likely

Model Results: Qualitative Match to Prior Human Data

1. Coarse Inferences

Condition	Sentence
DO-Plausible	The mother gave the daughter the candle .
DO-Implausible	The mother gave the candle the daughter .
PO-Plausible	The mother gave the candle to the daughter .
PO-Implausible	The mother gave the daughter to the candle .
Trans-Implausible	The lasagna defrosted the microwave .
Intrans-Implausible	The microwave defrosted in the lasagna .

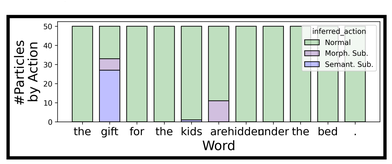
- [1] found more non-literal interpretations for implausible DO and Transitive conditions
- Extract "literal interpretations" from model as inferred "normal" action for a_t at start of second noun phrase
- Model shows same qualitative pattern as human participants

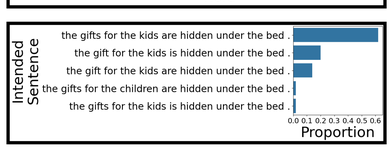


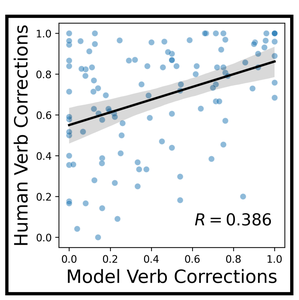
2. Fine-Grained Inferences

Condition	Sentence
Singular-Singular-Plural	The gift for the kid are hidden under the bed .
Singular-Plural-Plural	The gift for the kids are hidden under the bed .
Singular-Singular-Plural	The location of the star are recorded by astronomers .
Singular-Plural-Plural	The location of the stars are recorded by astronomers .

- [2] found more verb corrections when the subject phrase had higher prior
- For each item, extract model posterior for actions at subject and at verb
- Model preference for verb correction correlates with human preference — without fitting model parameters to human data

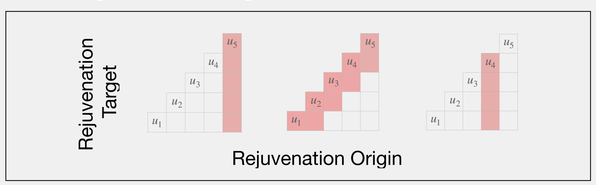


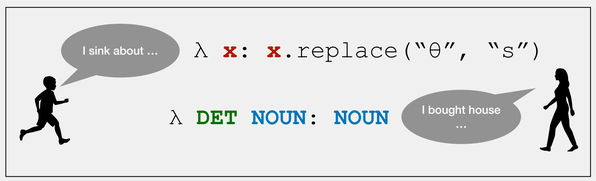


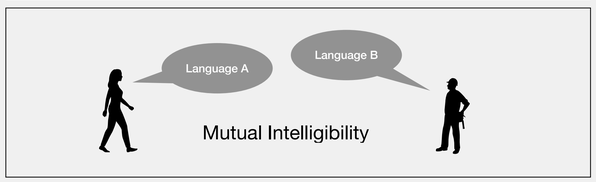


Future Work

- Model different reanalysis strategies [4]
- Learning new error operations [5]
- L2 comprehension as Noisy-Channel inferences [3]
- Modeling the effect of cognitive resource constraints







References: [1] Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. [2] Qian, P., & Levy, R. P. (2023). Comprehenders' Error Correction Mechanisms are Finely Calibrated to Language Production Statistics. [3] Futrell, R., & Gibson, E. (2017). L2 processing as noisy channel language comprehension. [4] Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. [5] Ellis, K., Wong, C., Nye, M., Sablé-Meyer, M., Morales, L., Hewitt, L., Cary, L., Solar-Lezama, A., & Tenenbaum, J. B. (2021). DreamCoder: Bootstrapping inductive program synthesis with wake-sleep library learning. [6] Levy, R., Reali, F., & Griffiths, T. L. (2008). Modeling the effects of memory on human online sentence processing with particle filters.