Graph Distance as Surprise: Free Energy Minimization in Knowledge Graph Reasoning

Gaganpreet Jhajj Fuhua Lin

School of Computing and Information Systems
Athabasca University, Canada

NORA Workshop @ NeurIPS 2025

Motivation: From Neuroscience to AI

The Free Energy Principle (FEP)

Biological systems minimize surprise by maintaining accurate world models

Recent breakthrough: Murphy et al. (2024) showed that syntactic operations minimize surprise through tree depth structures

Can we extend this to knowledge graphs?

The Challenge

Unlike syntactic trees, knowledge graphs are:

- Directed graphs with cycles – not simple trees
- Multiple paths between nodes
- Complex semantic relationships

We need a framework that:

- Handles cycles naturally
- Maintains theoretical grounding in FEP
- Works for general graphs

Our Approach: Graph Distance as Surprise

Key Insight: Use shortest-path distance to measure surprise

$$S_{\text{geo}}(e \mid C) = \begin{cases} \min_{c \in C} d_{\mathcal{G}}(c, e) & \text{if path exists} \\ \alpha & \text{otherwise} \end{cases}$$

where $d_{\mathcal{G}}(c, e)$ is computed via breadth-first search (BFS)

- Shorter distances $\rightarrow$ higher probability $\rightarrow$ lower surprise
- Disconnected entities $\rightarrow$ high surprise (α)

Why This Works: Three Justifications

1. **Proper Generalization**
   - For trees, recovers Murphy's tree depth exactly
2. **Least-Action Principle**
   - Shortest paths minimize cumulative cost
   - Aligns with active inference
3. **Computational Grounding**
   - In GNNs, k message-passing iterations $=$ k-hop neighborhoods
   - Minimizing iterations $=$ minimizing distance

Bonus: Cycles handled naturally (BFS uses visited sets)

Example: Canadian Prime Ministers

Query: "Who is the Prime Minister?"
Context: Canada

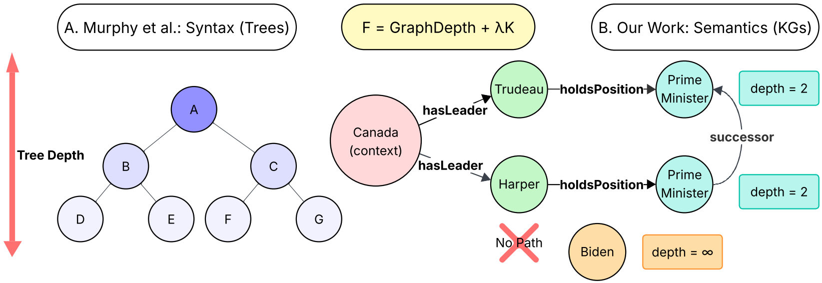
Results

- Trudeau: distance $= 2 \rightarrow$ Low surprise
- Harper: distance $= 2 \rightarrow$ Low surprise
- Biden: distance $= \infty \rightarrow$ High surprise

Framework correctly identifies both Canadian PMs as plausible while rejecting the US president!

Cycle between Trudeau $\leftrightarrow$ Harper handled naturally.

Example: Canadian Prime Ministers



Connection to Free Energy Principle

Under FEP, agents minimize:

$$F = -\log P(o, s) - H[Q(s)]$$

Our framework:

- KG serves as the agent's generative model
- Shorter distances $\rightarrow$ higher probability
- S_{geo} implements the surprise term
- Compatible with active inference

Future Work & Positioning

Honest Positioning

This is work-in-progress proposing one research direction.
Different, perhaps more elegant ideas may emerge!

Next Steps:

- Empirical validation on FB15k-237, YAGO
- Compare to human semantic similarity judgments
- Integration with existing KG reasoning systems
- Extension to temporal knowledge graphs
- Application to GNN depth selection

Contributions

1. **Novel connection** between FEP from neuroscience and KG reasoning in AI
2. **Extension** of surprise minimization from trees to general graphs with cycles
3. **Theoretical framework** for distance-based reasoning in KG systems
4. **Computational grounding** linking to GNNs and active inference

References

[1] K. Friston, “The free-energy principle: a unified brain theory?,” *Nat Rev Neurosci*, vol. 11, no. 2, pp. 127–138, Feb. 2010,

[2] E. Murphy, E. Holmes, and K. Friston, “Natural language syntax complies with the free-energy principle,” *Synthese*, vol. 203, no. 5, p. 154, May 2024,

[3] T. Parr, G. Pezzulo, and K. J. Friston, *Active inference: the free energy principle in mind, brain, and behavior*. MIT Press, 2022.

[4] T. Kipf, “Semi-supervised classification with graph convolutional networks,” arXiv preprint arXiv:1609.02907, 2016.