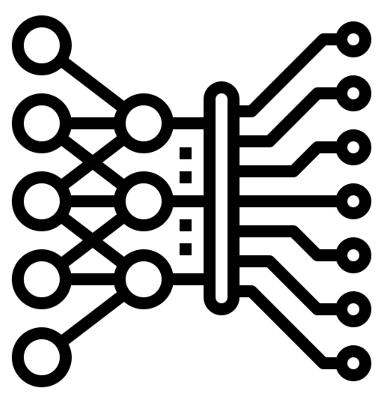
DeepMind

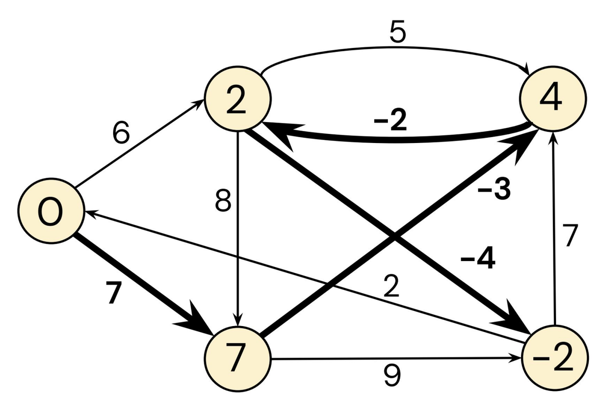
Persistent Message Passing

Heiko Strathmann, Mohammadamin Barekatain, Charles Blundell, Petar Veličković

Algorithmic Reasoning with GNNs



Neural networks



Algorithms

- \+ Operate on **raw** inputs
- \+ Generalise on **noisy** conditions
- \+ Models **reusable** across tasks
- \- Require **big data**
- \- Unreliable when **extrapolating**
- \- Lack of **interpretability**

- \+ Trivially **strongly** generalise
- \+ **Compositional** (subroutines)
- \+ Guaranteed **correctness**
- \+ **Interpretable** operations
- \- Inputs must match **spec**
- \- Not **robust** to task variations

Non-Markovian Inputs

GNNs typically overwrite their internal states in every step

- Previous applications: only latest version of the data relevant
- Here: queries require knowledge of previous versions of the data
- This places significant "pressure" on internal states

Current GNNs cannot explicitly memorize their past states

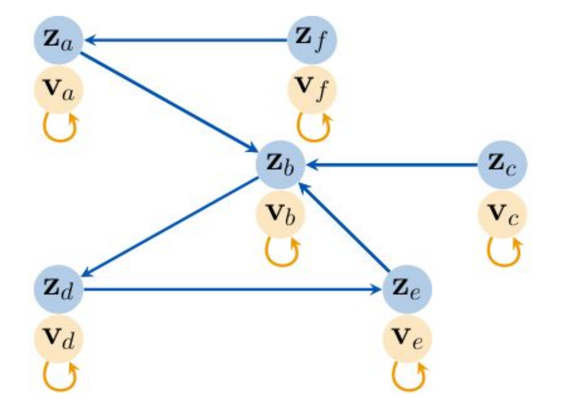
Persistent Message Passing

Idea: endow GNNs with an explicit, persistent, memory

- Instead of overwriting nodes -> persisting nodes
- Paired with an efficient relevance / query mechanism
- Effectively an **episodic memory** of past computation

Strong performance on non-Markovian tasks

Input Encoding



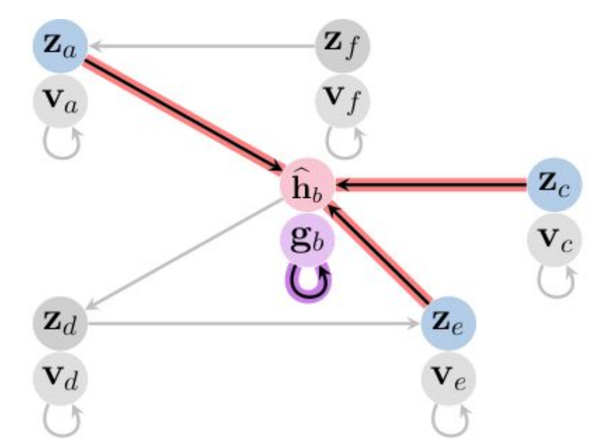
Initial connectivity

$$\Pi^{(0)} = \Pi$$
$$\Lambda^{(0)} = I$$

Node representations

$$v_j^{(t)} = f_{\text{relevance}}\left(\texttt{time\_stamp}(j), t, h_j^{(t-1)}\right)$$
$$z_j^{(t)} = f_{\text{operation}}\left(\texttt{expand}(E^{(t)}, s^{(t)}), h_j^{(t-1)}\right)$$

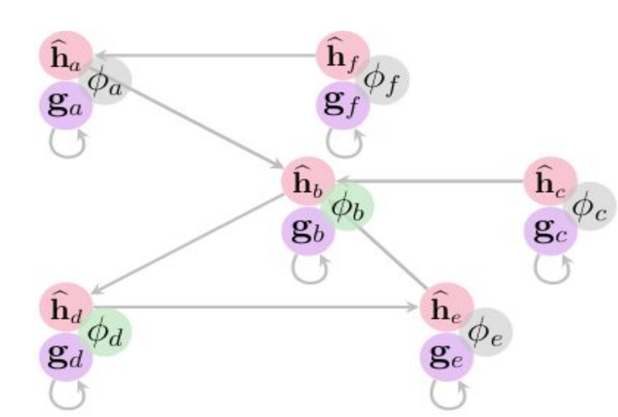
Messages



Next-step candidates
$$\hat{h}_j^{(t)} = U\left(z_j^{(t)}, \max_{\Pi_{jj'}^{(t-1)}=1} M\left(z_{j'}^{(t)}, z_j^{(t)}\right)\right)$$

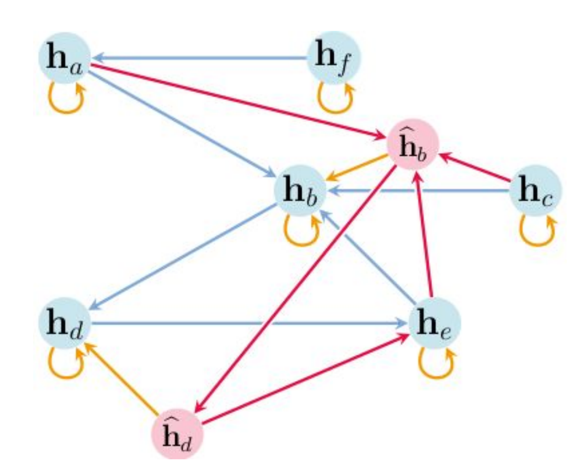
Relevance latents
$$g_j^{(t)} = U\left(v_j^{(t)}, \max_{\Lambda_{jj'}^{(t-1)}=1} M\left(v_{j'}^{(t)}, v_j^{(t)}\right)\right)$$

Masking



$$\mu_j^{(t)} = \psi_{\text{relevance}}(g_j^{(t)}) \in \{0,1\}$$
$$\phi_j^{(t)} = \psi_{\text{persistency}}(\hat{h}_j^{(t)} \cdot \mu_j^{(t)}) \in \{0,1\}$$

Persistency



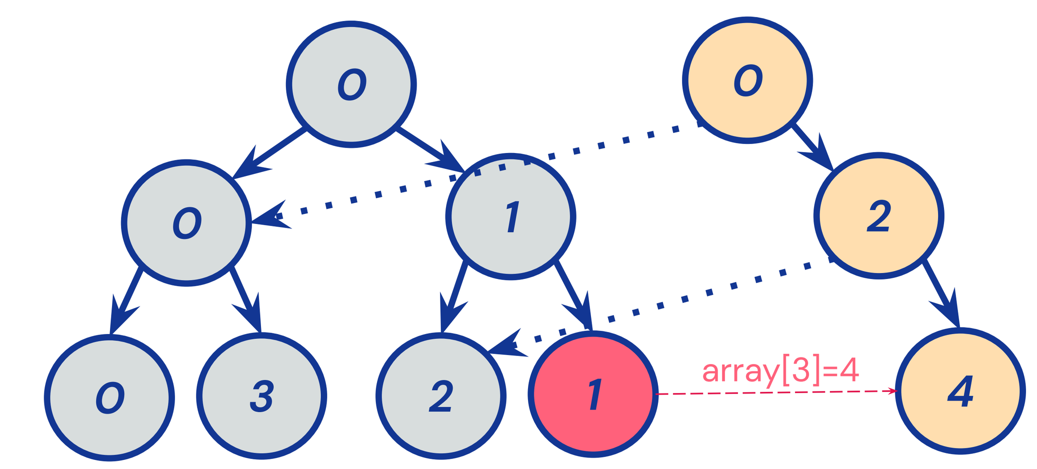
- "Persist" selected nodes
- Original nodes untouched
- Track connectivity

Task: Dynamic Range Minimum Queries

Compute (historic) minimum across ranges in continuously updated array

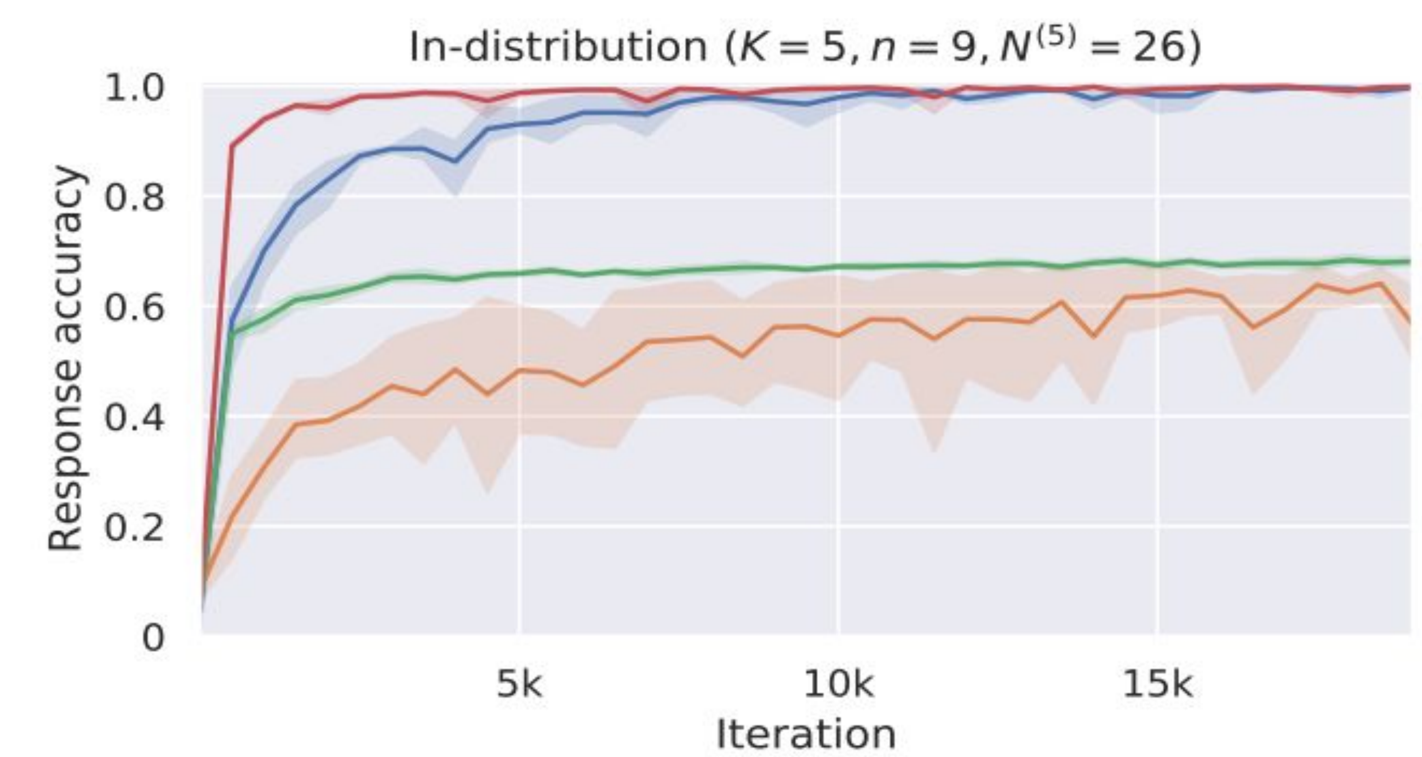
Solved by Segment Trees

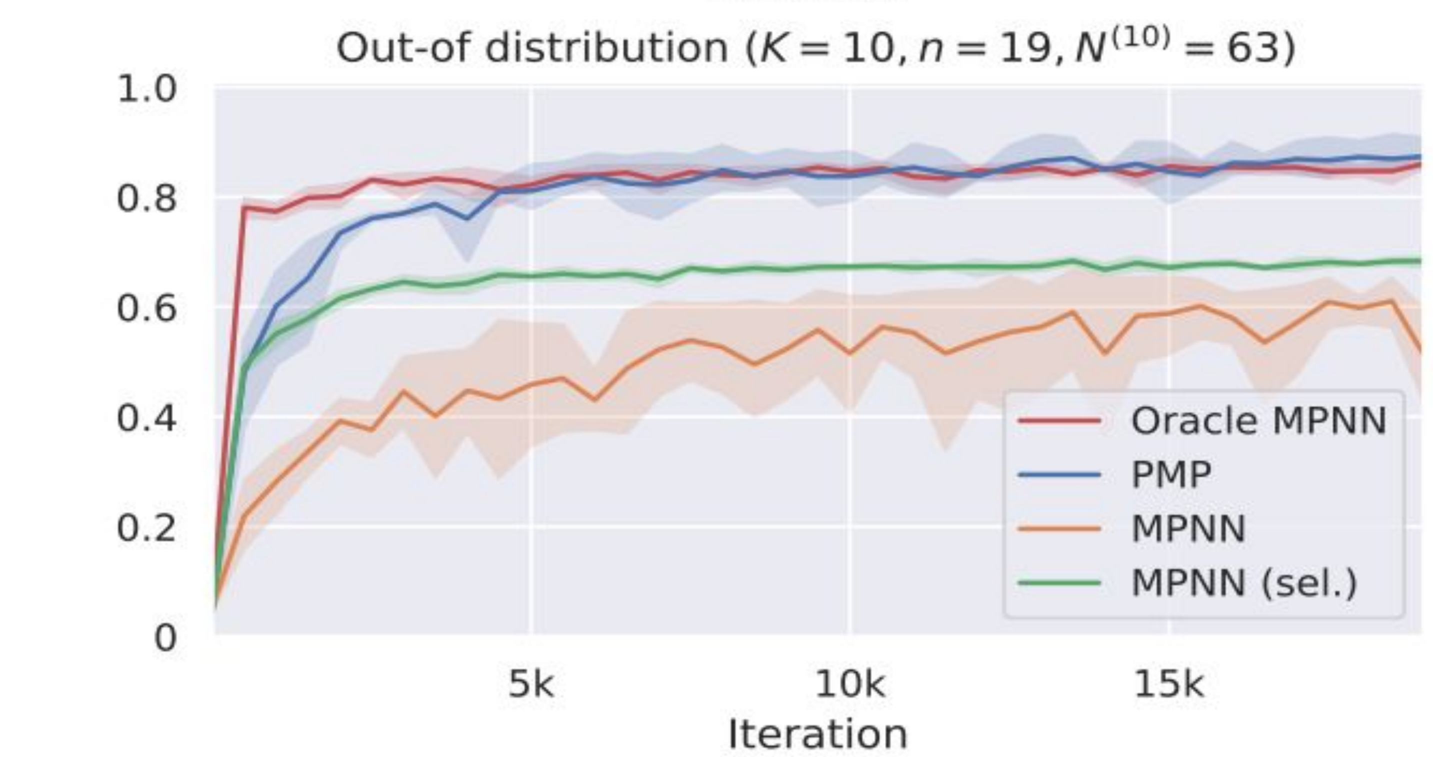
- O(log n) query time
- Leaf nodes store array
- Intermediate nodes store minimum of children



Results

- **PMP generalizes out-of-distribution to 2x larger test inputs**
- **Significantly outperforms GNNs which overwrite states**
- Supervised on response (binary encoding), ground-truth masks, and node-values





https://arxiv.org/abs/2103.01043