A Theory of Link Prediction via Relational Weisfeiler-Leman on Knowledge Graphs

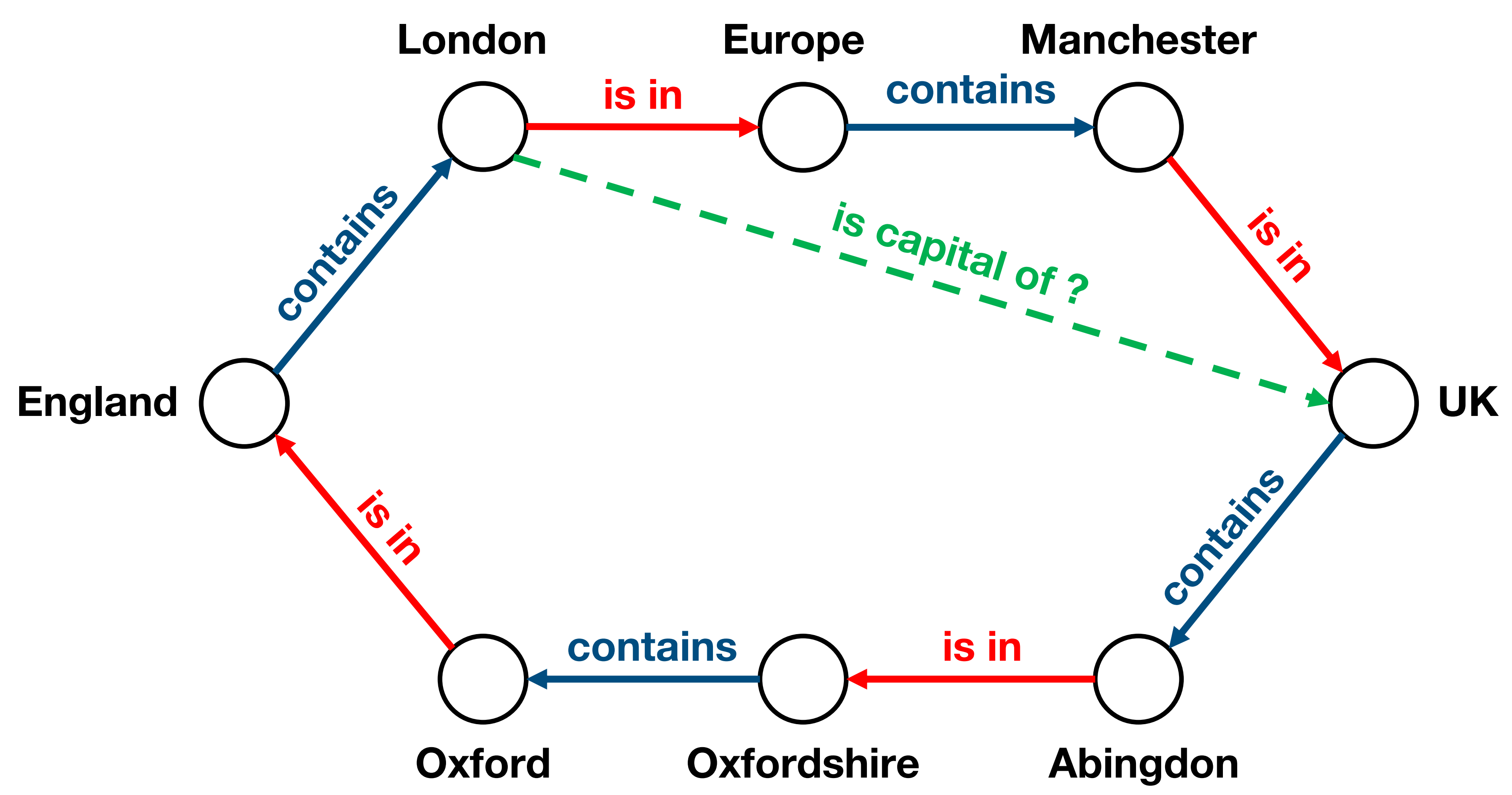
Xingyue Huang, Miguel Romero, İsmail İlkan Ceylan, Pablo Barceló

UNIVERSITY OF OXFORD

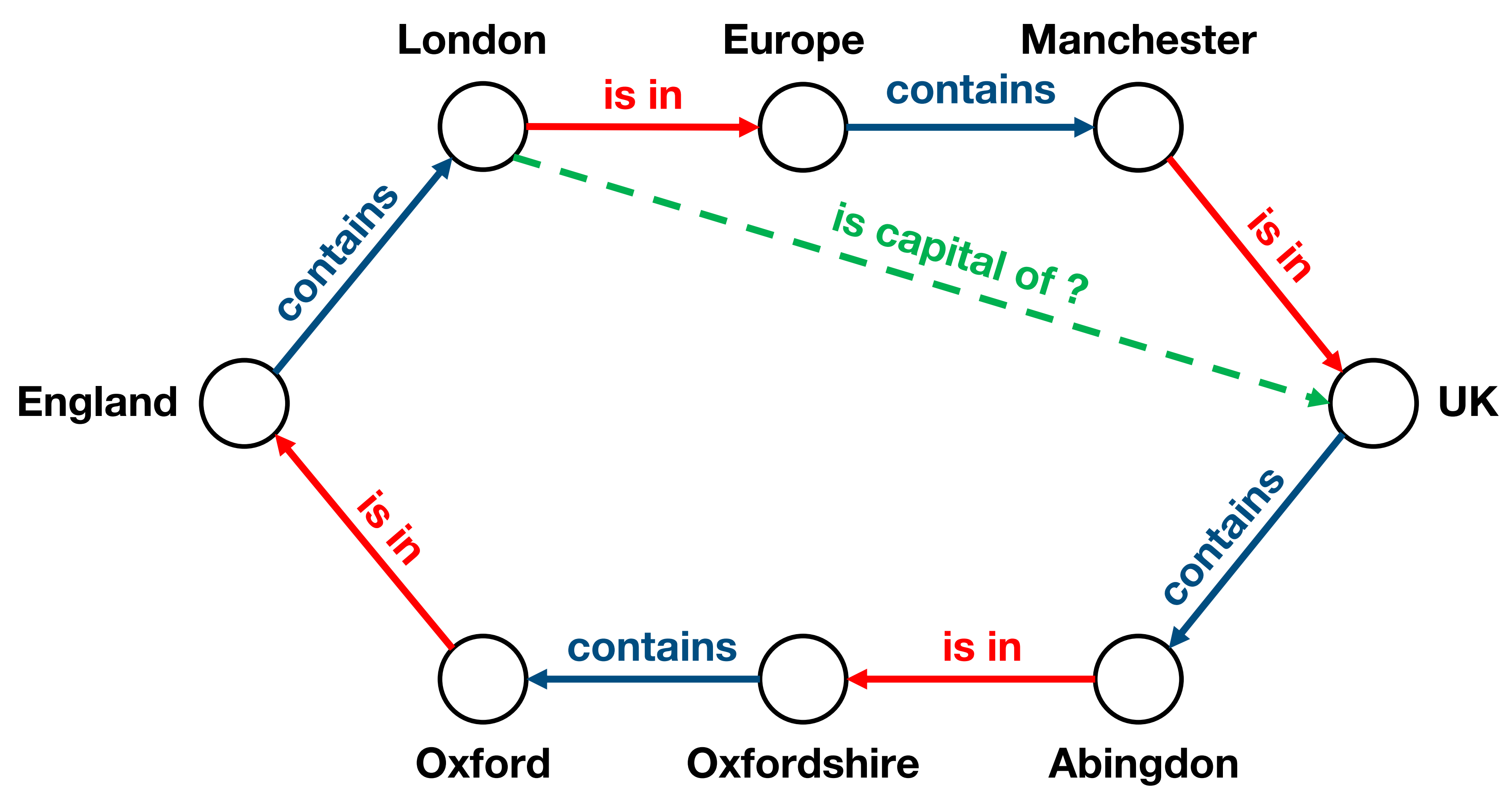
CENTRO NACIONAL DE INTELIGENCIA ARTIFICIAL

NEURAL INFORMATION PROCESSING SYSTEMS

Link Prediction on Knowledge Graphs

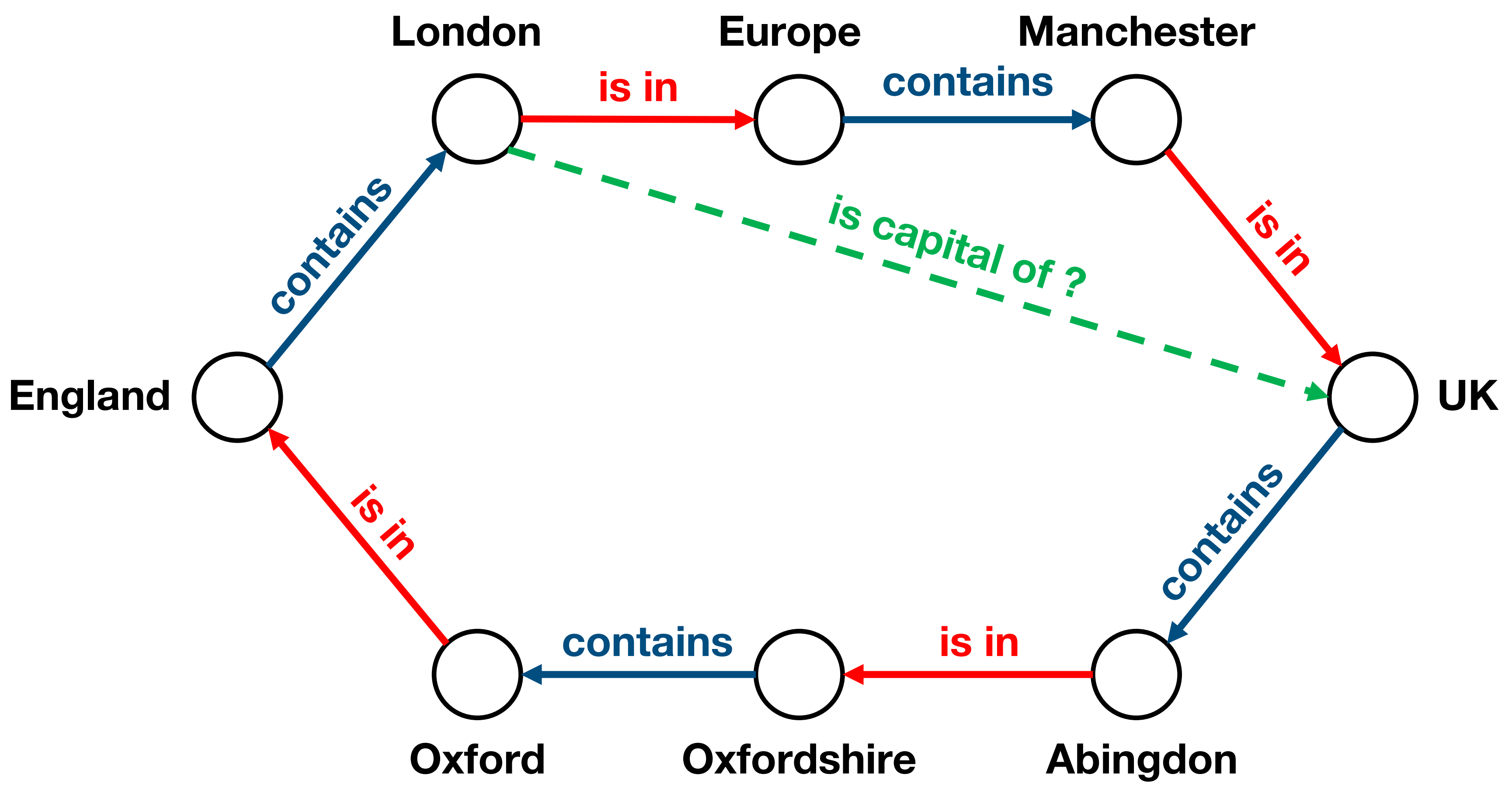


Link Prediction on Knowledge Graphs



Knowledge graph is a graph with edges labelled with relation types.

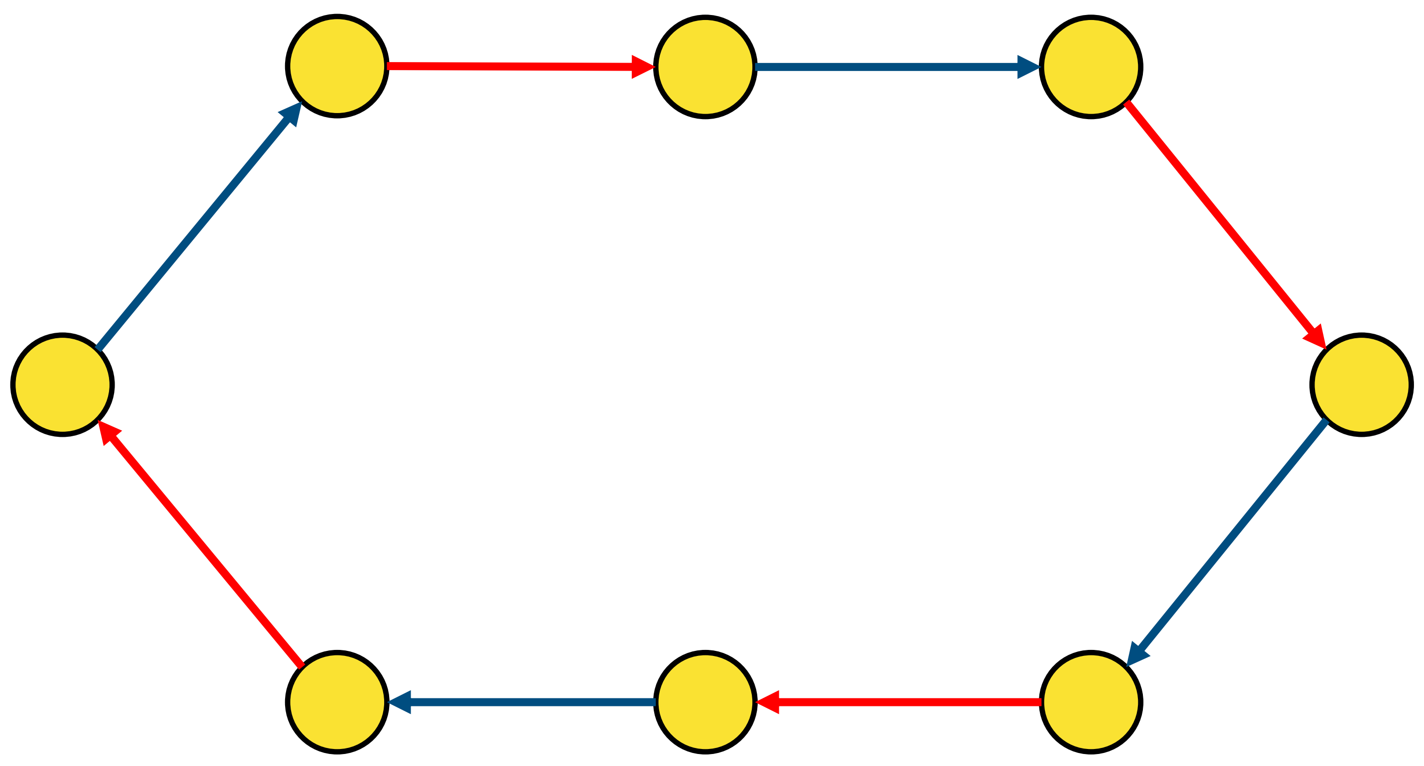
Link Prediction on Knowledge Graphs



Knowledge graph is a graph with edges labelled with relation types.

Link prediction is to predict missing link or relation on pairs of nodes.

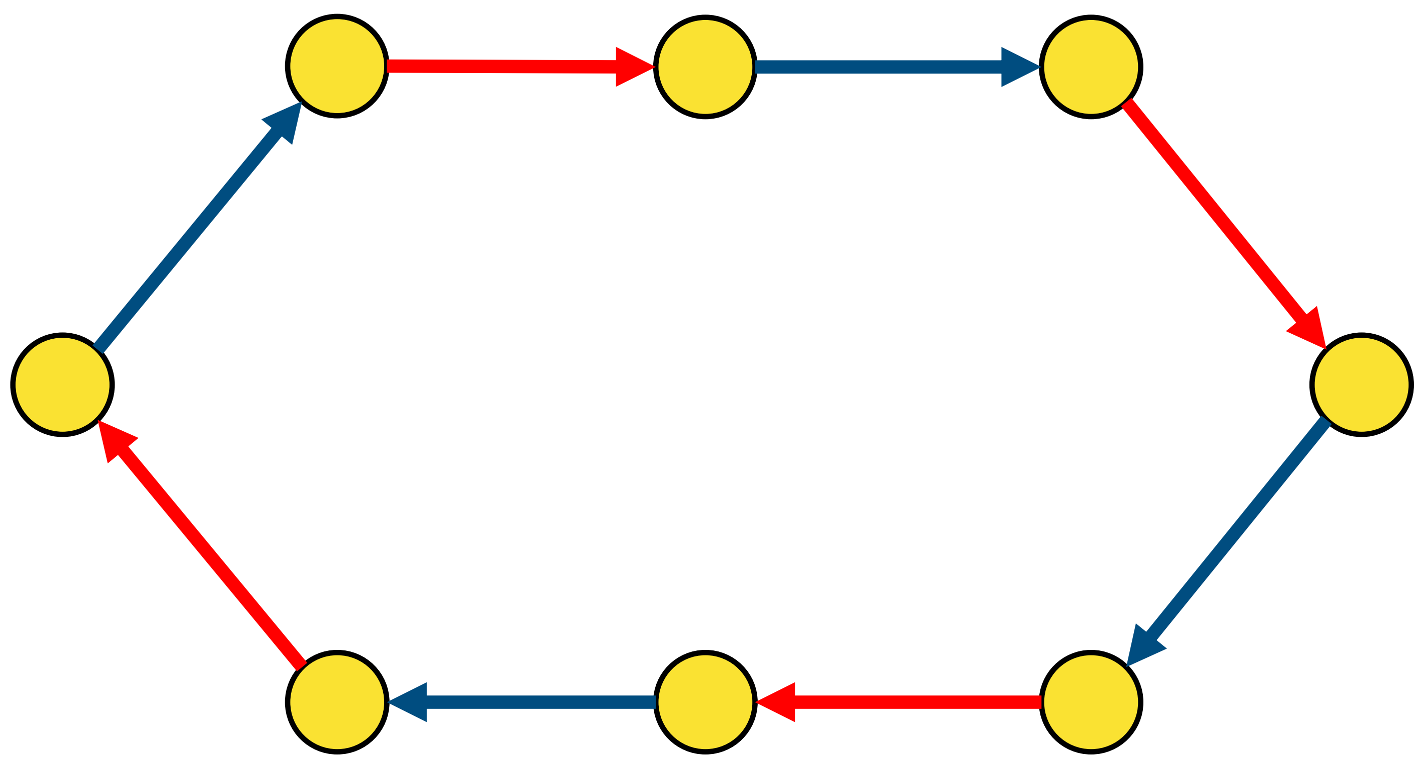
Relational Message Passing Neural Networks



$$\boxed{\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v}$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) |\, w \in \mathcal{N}_r(v), r \in R\}\!\!\}), \textsc{Read}(\{\!\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\!\})\right),$$

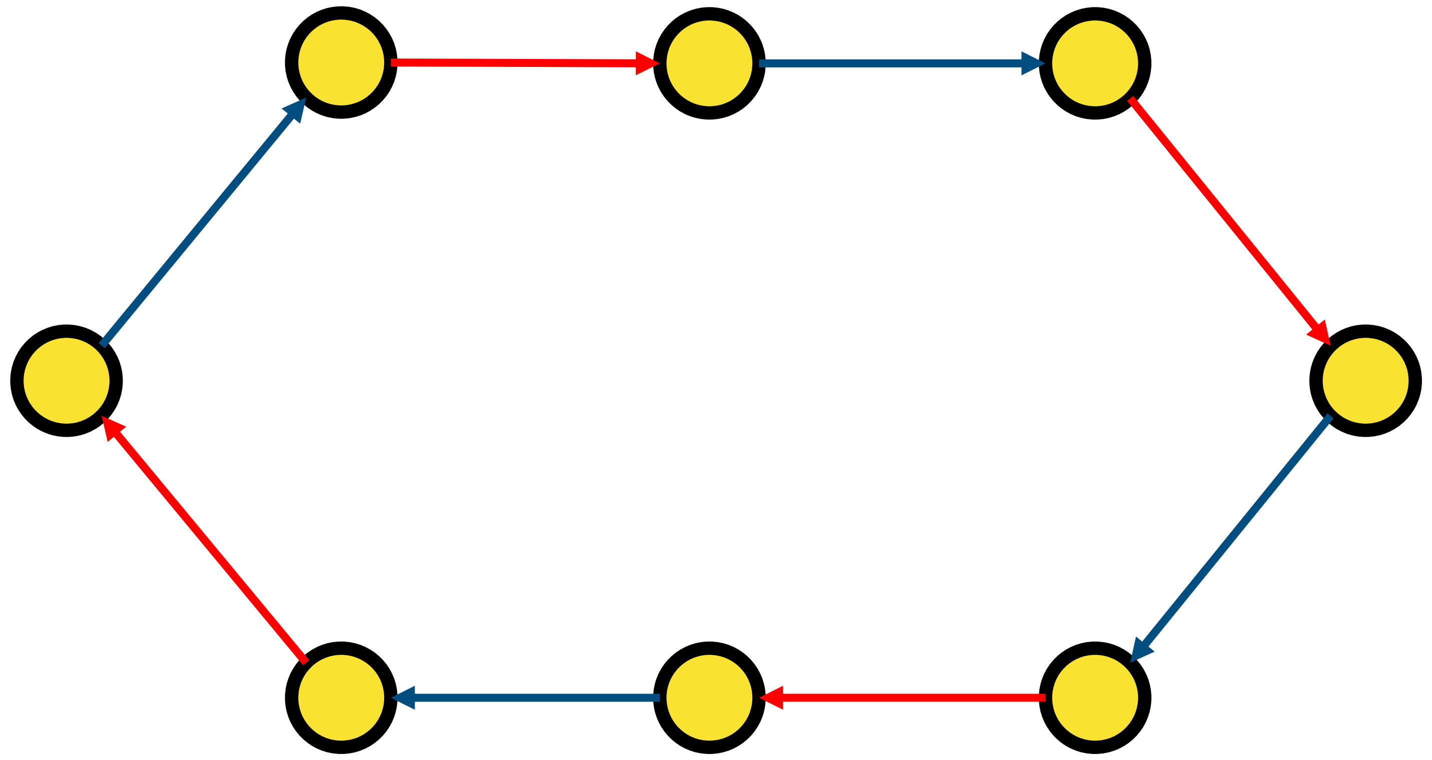
Relational Message Passing Neural Networks



$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \boxed{\textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) \mid w \in \mathcal{N}_r(v), r \in R\}\!\}),} \textsc{Read}(\{\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\})\right),$$

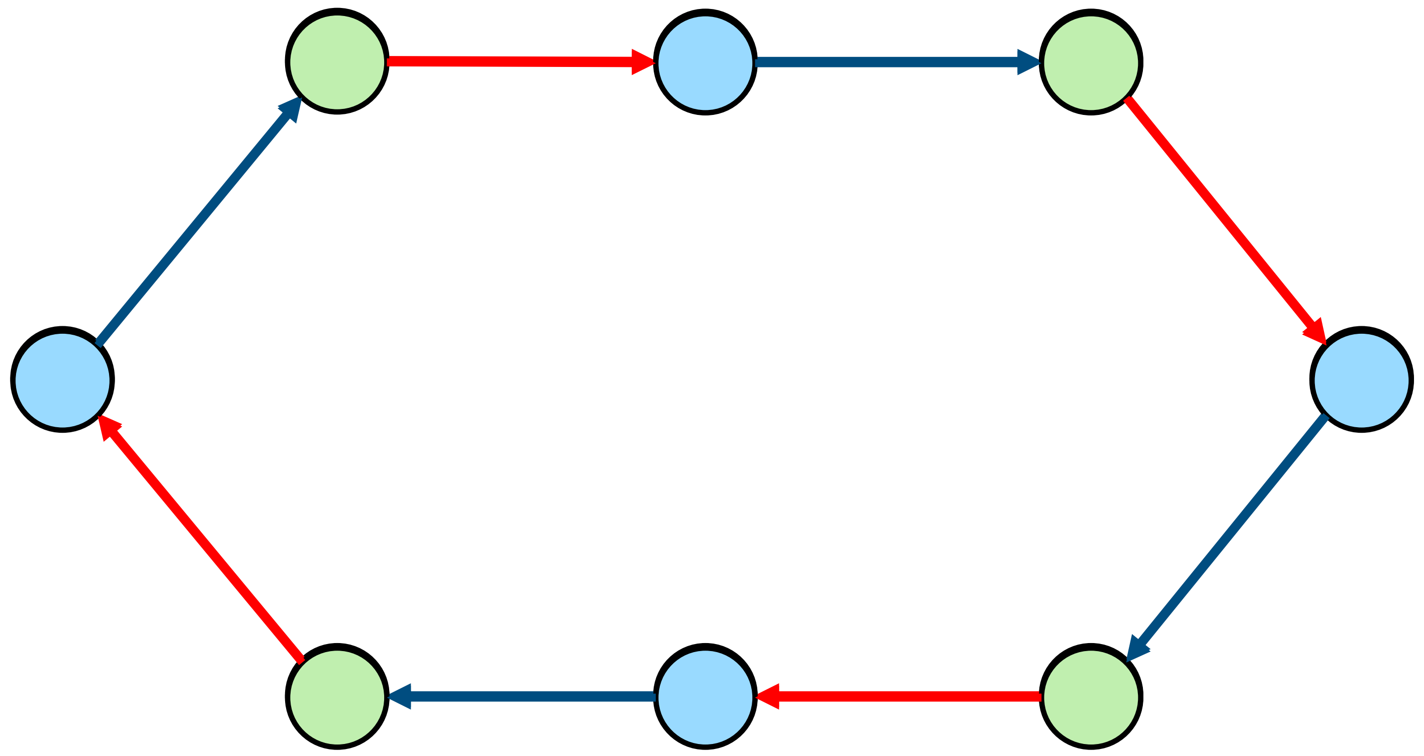
Relational Message Passing Neural Networks



$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) \mid w \in \mathcal{N}_r(v), r \in R\}\!\}), \boxed{\textsc{Read}(\{\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\})}\right),$$

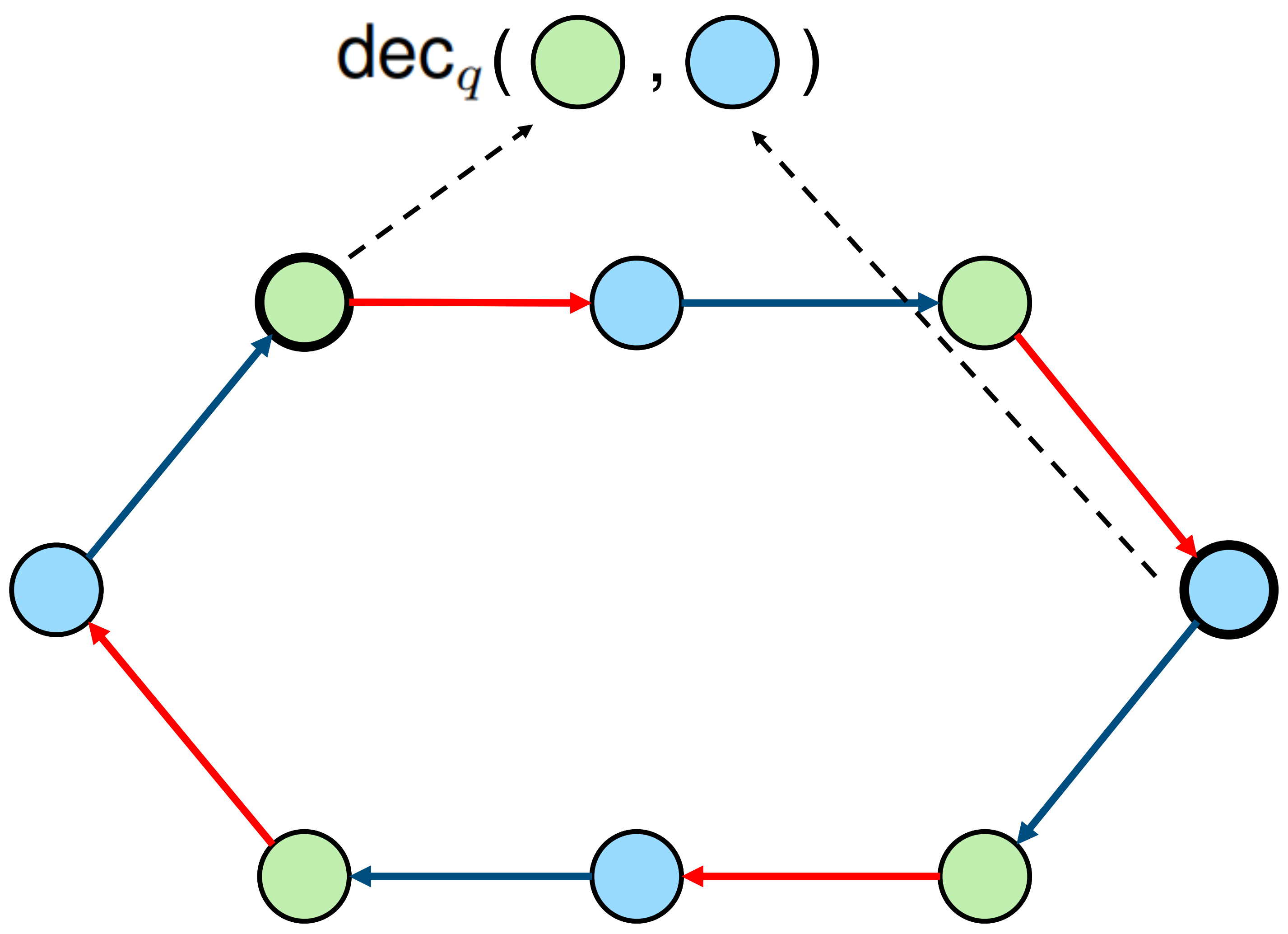
Relational Message Passing Neural Networks



$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) \mid w \in \mathcal{N}_r(v), r \in R\}\!\!\}), \textsc{Read}(\{\!\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\!\})\right),$$

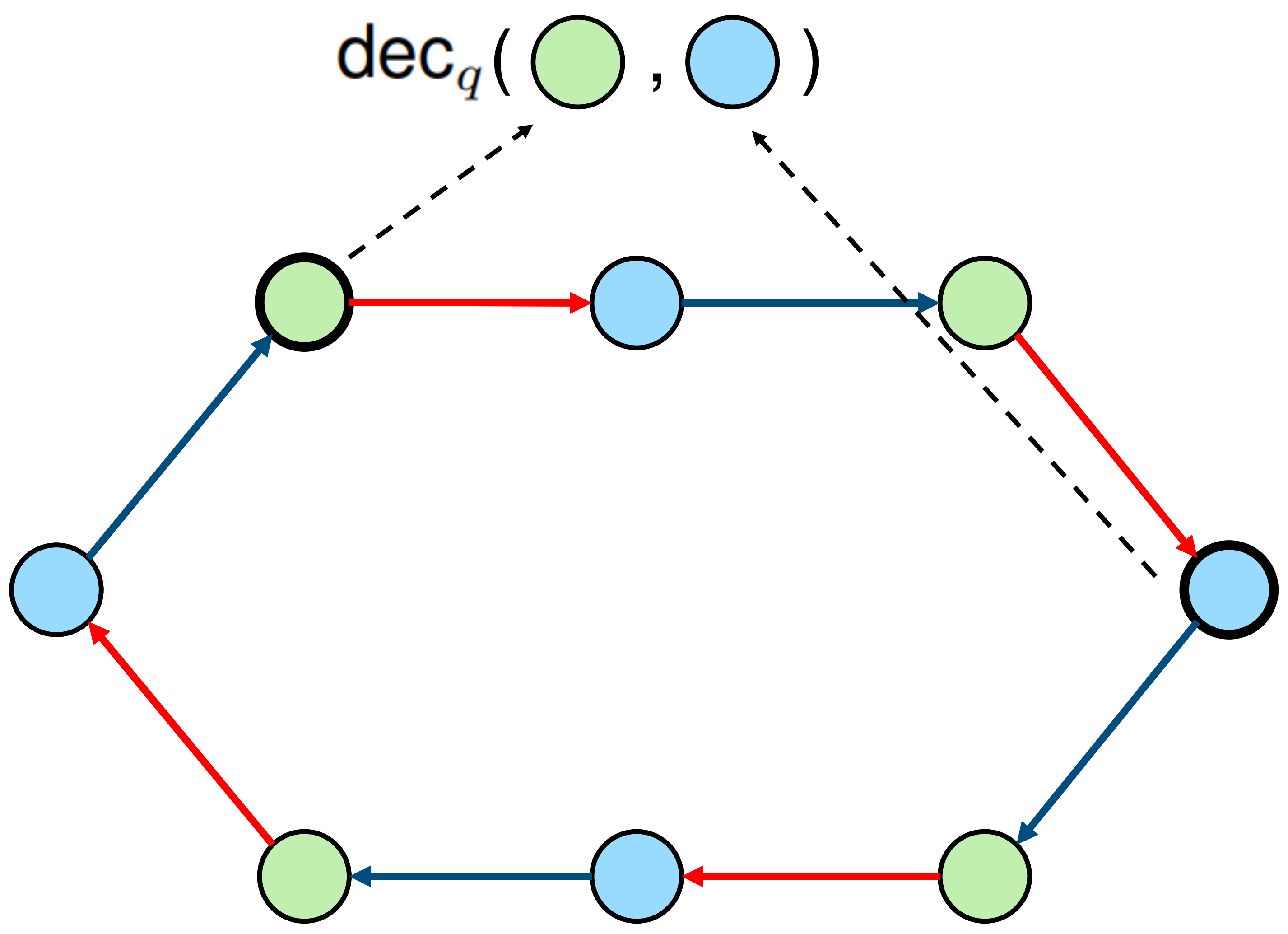
Relational Message Passing Neural Networks



$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$
$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) |\ w \in \mathcal{N}_r(v), r \in R\}\!\!\}), \textsc{Read}(\{\!\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\!\})\right),$$

Relational Message Passing Neural Networks

R-MPNNs relies on binary decoder for link prediction.

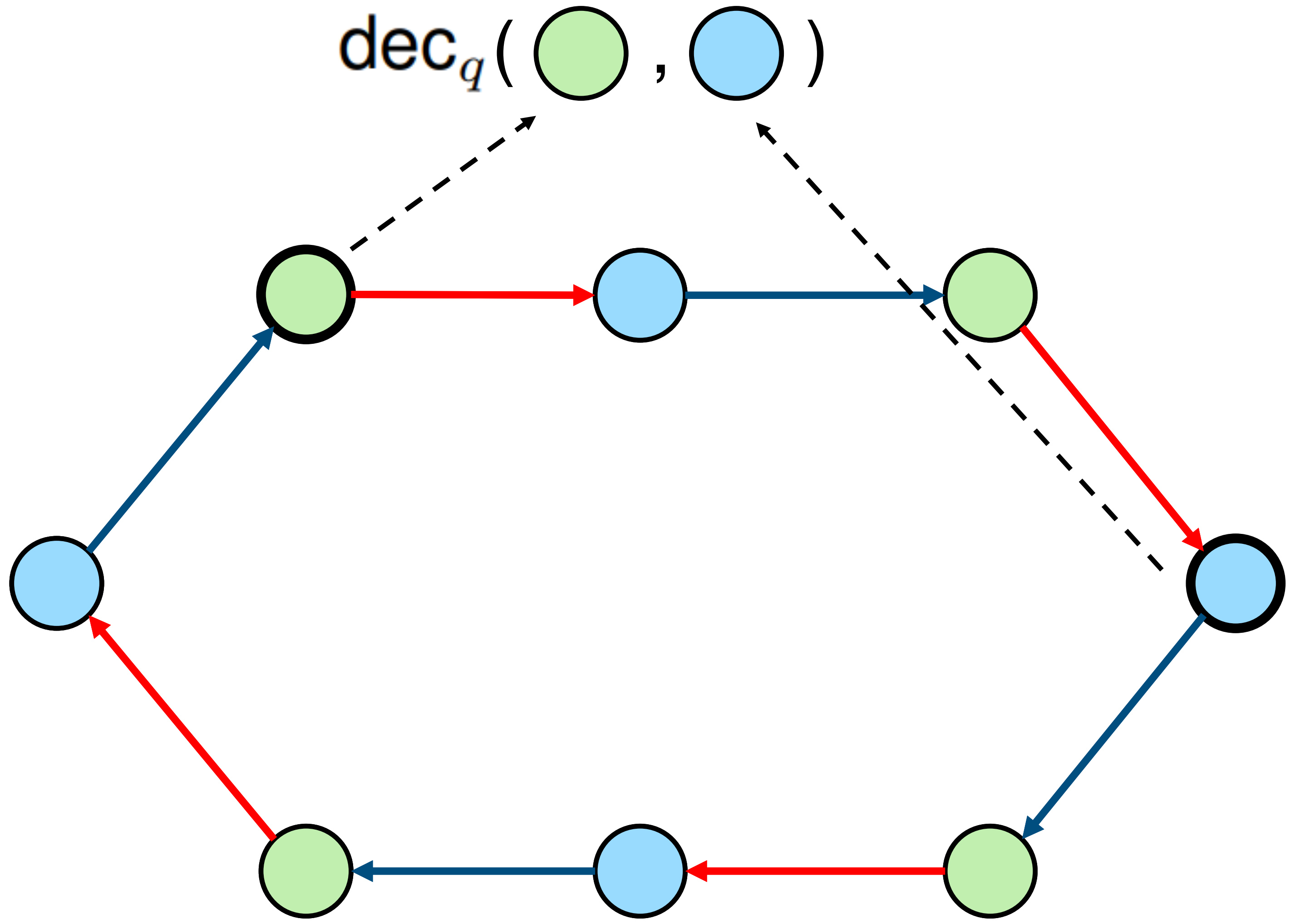


$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) \mid w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\})\right),$$

Relational Message Passing Neural Networks

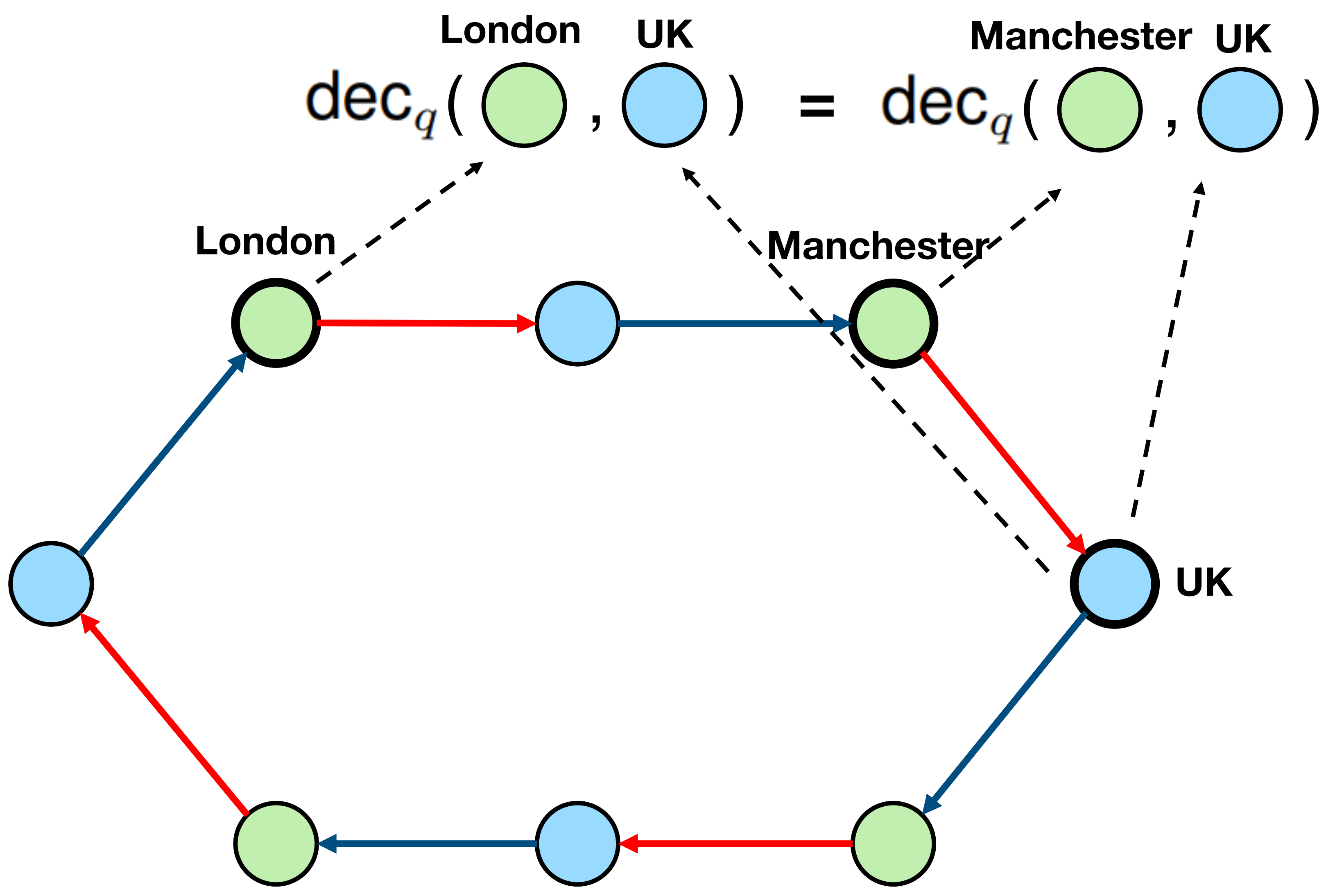
R-MPNNs relies on binary decoder for link prediction.



Prominent examples are RGCN [1] and CompGCN [2].

$$
\begin{aligned}
\boldsymbol{h}_v^{(0)} &= \boldsymbol{x}_v \\
\boldsymbol{h}_v^{(t+1)} &= \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) \mid w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\})\right),
\end{aligned}
$$

Relational Message Passing Neural Networks

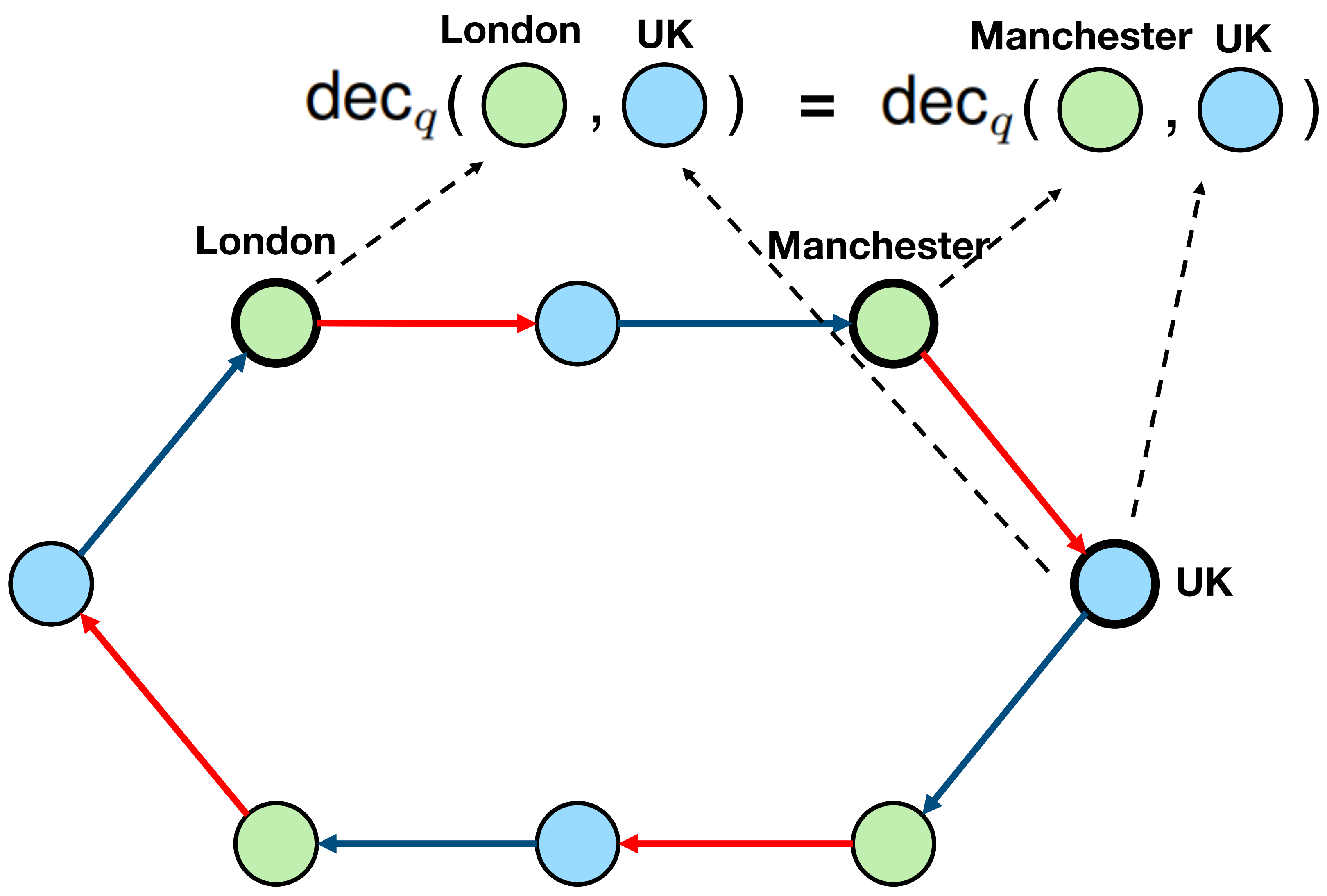


$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) |\, w \in \mathcal{N}_r(v), r \in R\}\!\!\}), \textsc{Read}(\{\!\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\!\})\right),$$

Relational Message Passing Neural Networks

R-MPNNs are at most as powerful as *relational local 1-WL test* [3].

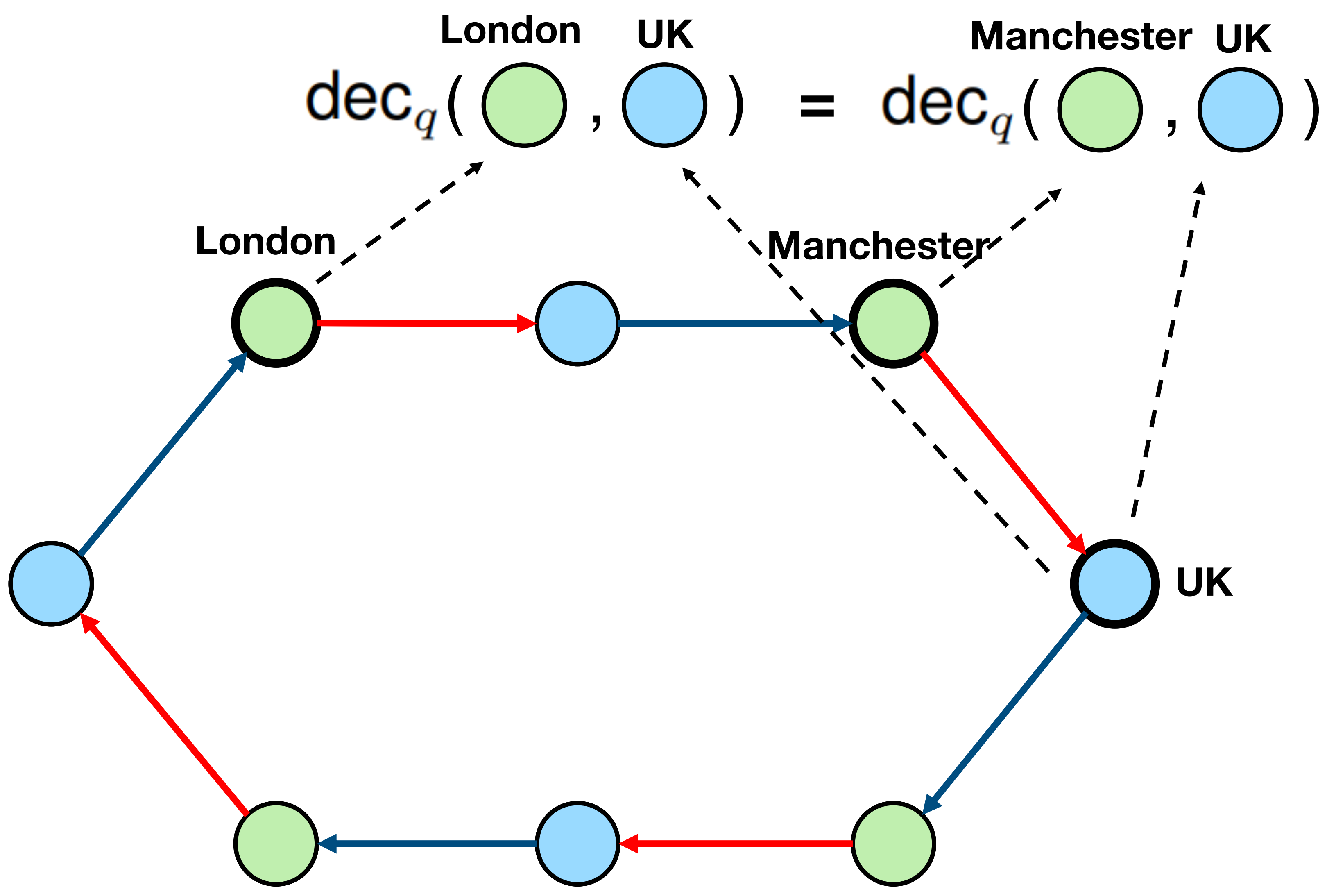


$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) \mid w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\})\right),$$

Relational Message Passing Neural Networks

R-MPNNs are at most as powerful as *relational local 1-WL test* [3]

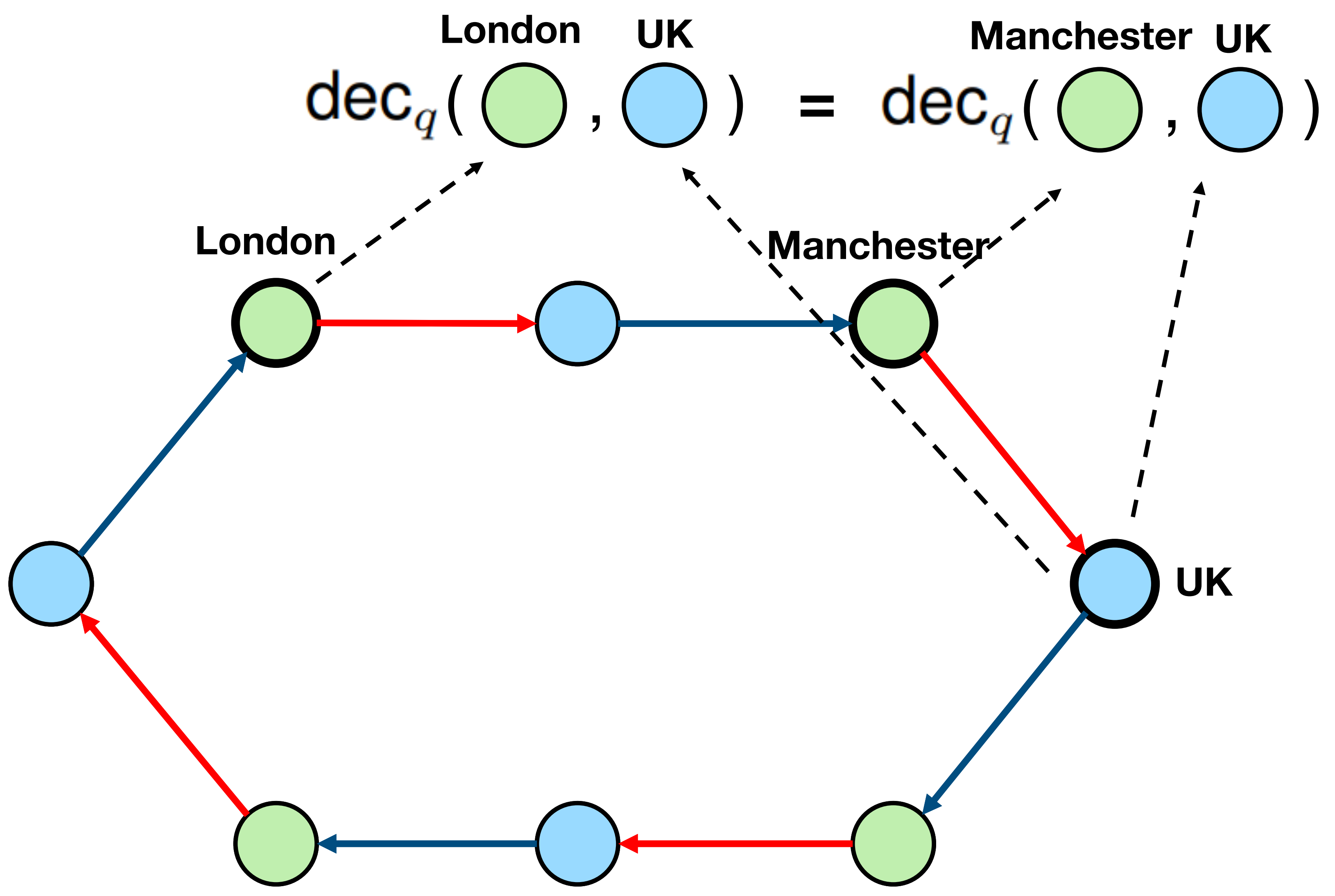


Higher order method is computationally prohibitive.

$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) |\ w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\})\right),$$

Relational Message Passing Neural Networks



R-MPNNs are at most as powerful as *relational local 1-WL test* [3]

Higher order method is computationally prohibitive.

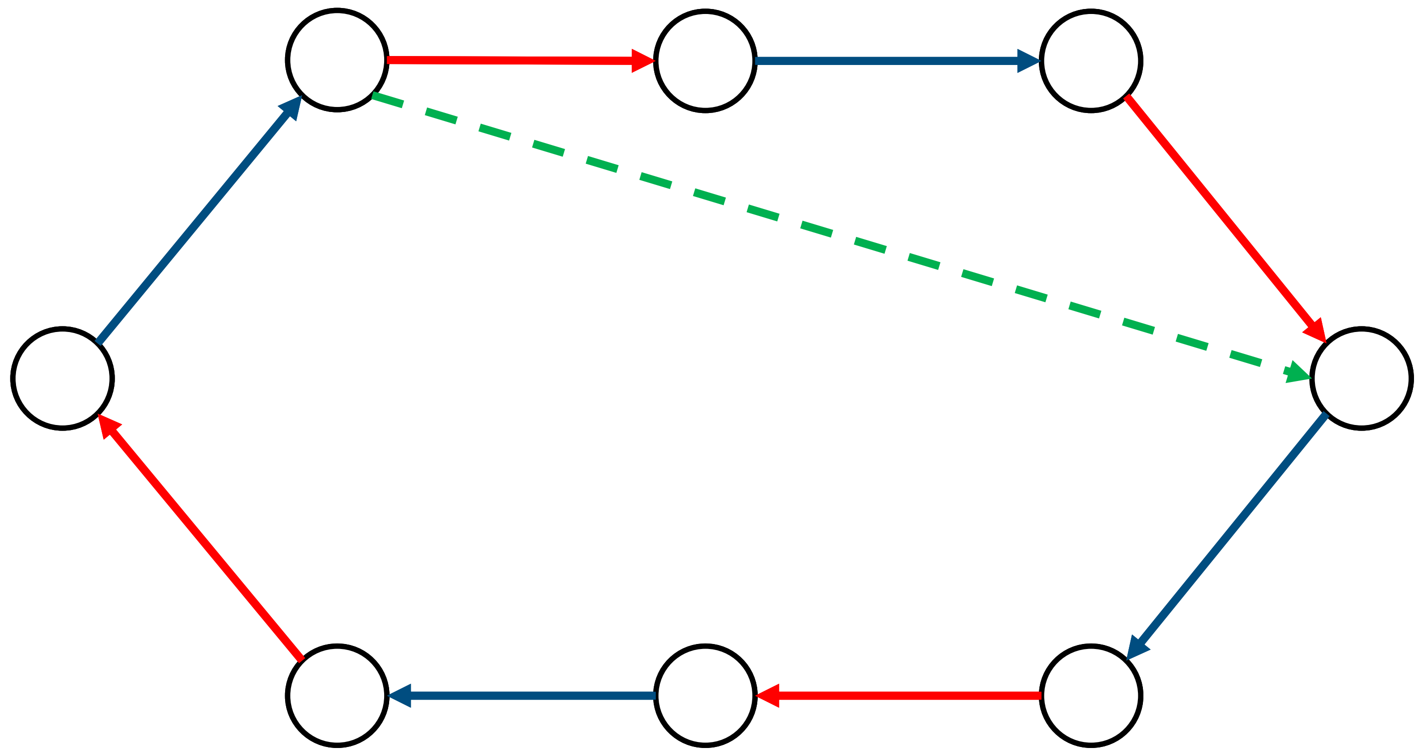
$$\boldsymbol{h}_v^{(0)} = \boldsymbol{x}_v$$

$$\boldsymbol{h}_v^{(t+1)} = \textsc{Upd}\left(\boldsymbol{h}_v^{(f(t))}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_w^{(t)}) \mid w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_w^{(t)} \mid w \in V\}\!\})\right),$$

What is a good trade off between expressivity and scalability?

Conditional Message Passing Neural Networks (C-MPNNs)

Conditional Message Passing Neural Networks

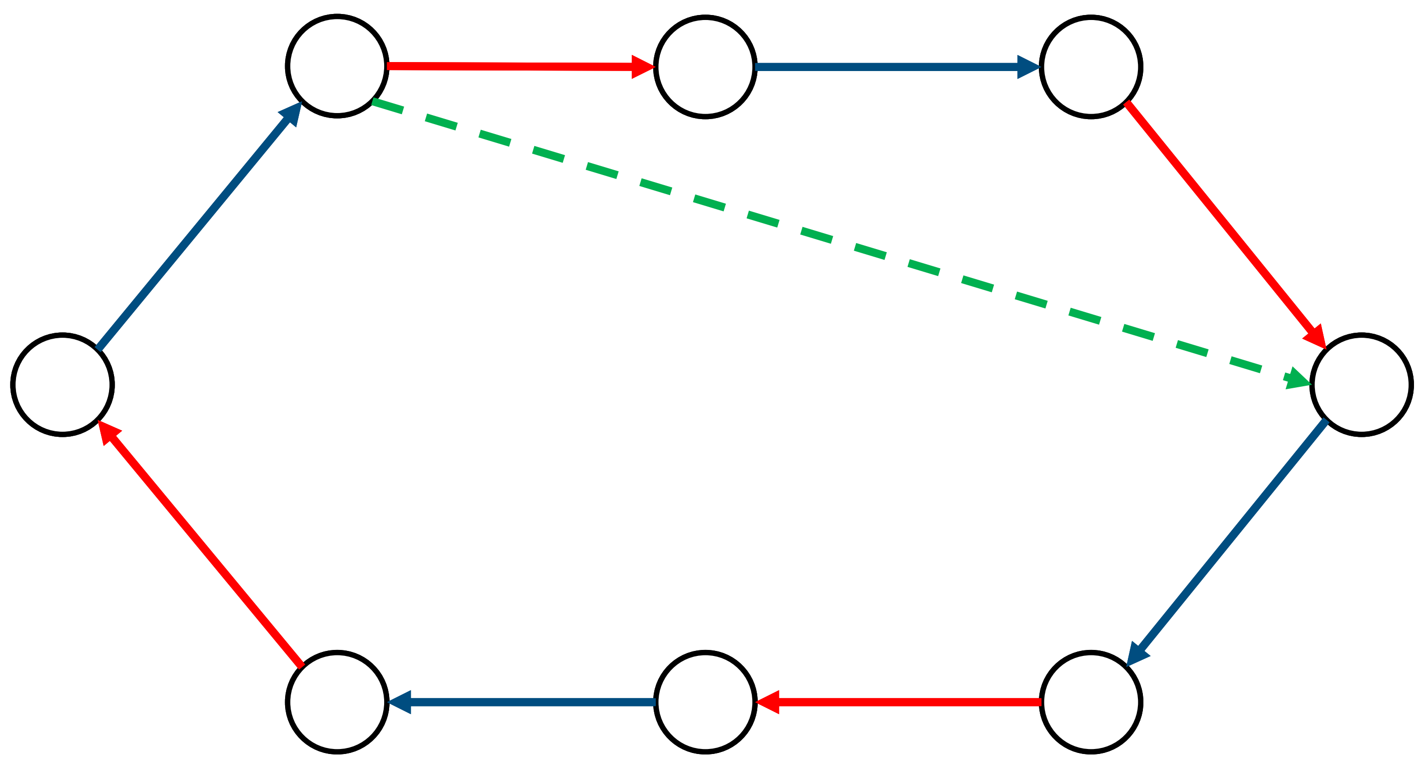


$$\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}(\boldsymbol{h}_{v|u,q}^{f(t)}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) | \; w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\})),$$

Conditional Message Passing Neural Networks

NBFNet [4] locally computes pairwise representations by conditioning.

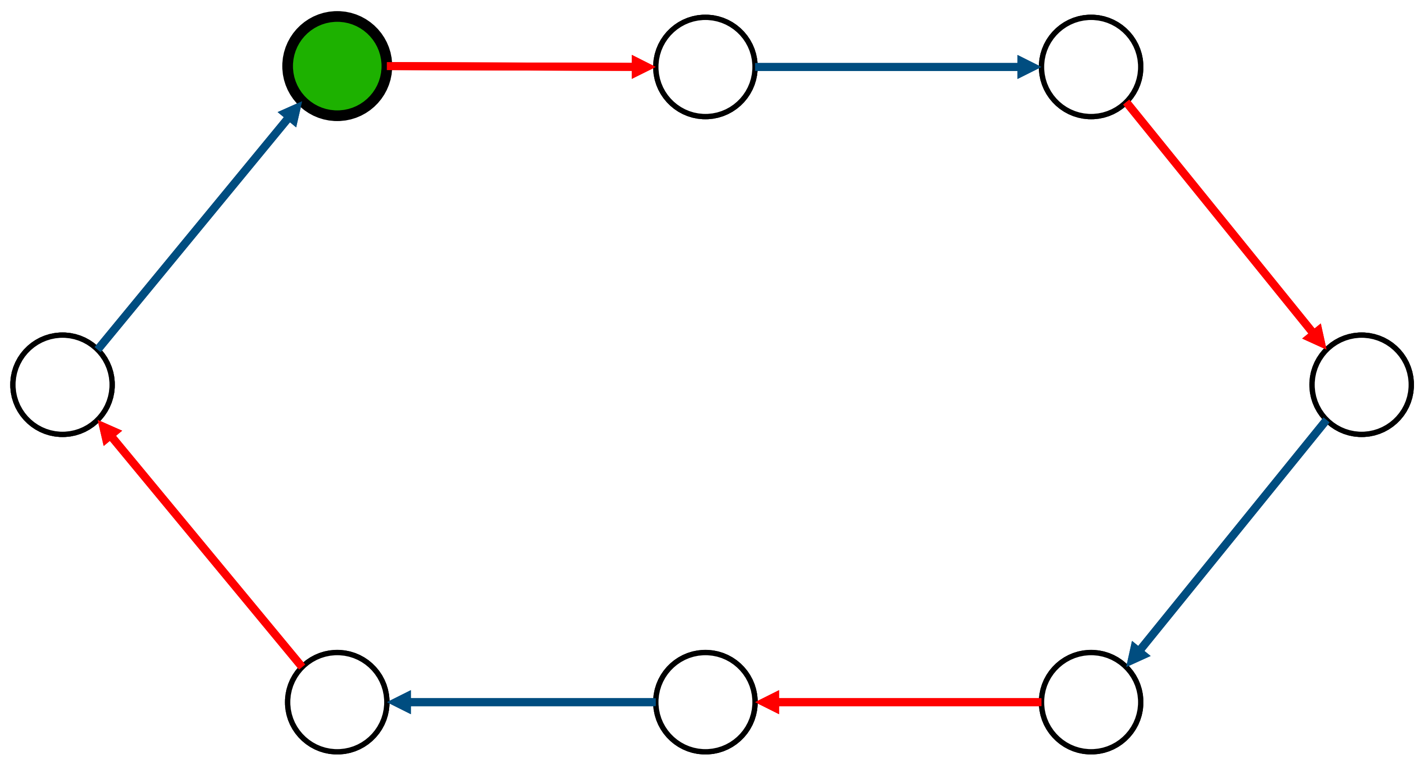


$$\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}(\boldsymbol{h}_{v|u,q}^{f(t)}, \textsc{Agg}(\{\!\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) |\; w \in \mathcal{N}_r(v), r \in R\}\!\!\}), \textsc{Read}(\{\!\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\!\})),$$

Conditional Message Passing Neural Networks

NBFNet [4] locally computes pairwise representations by conditioning.

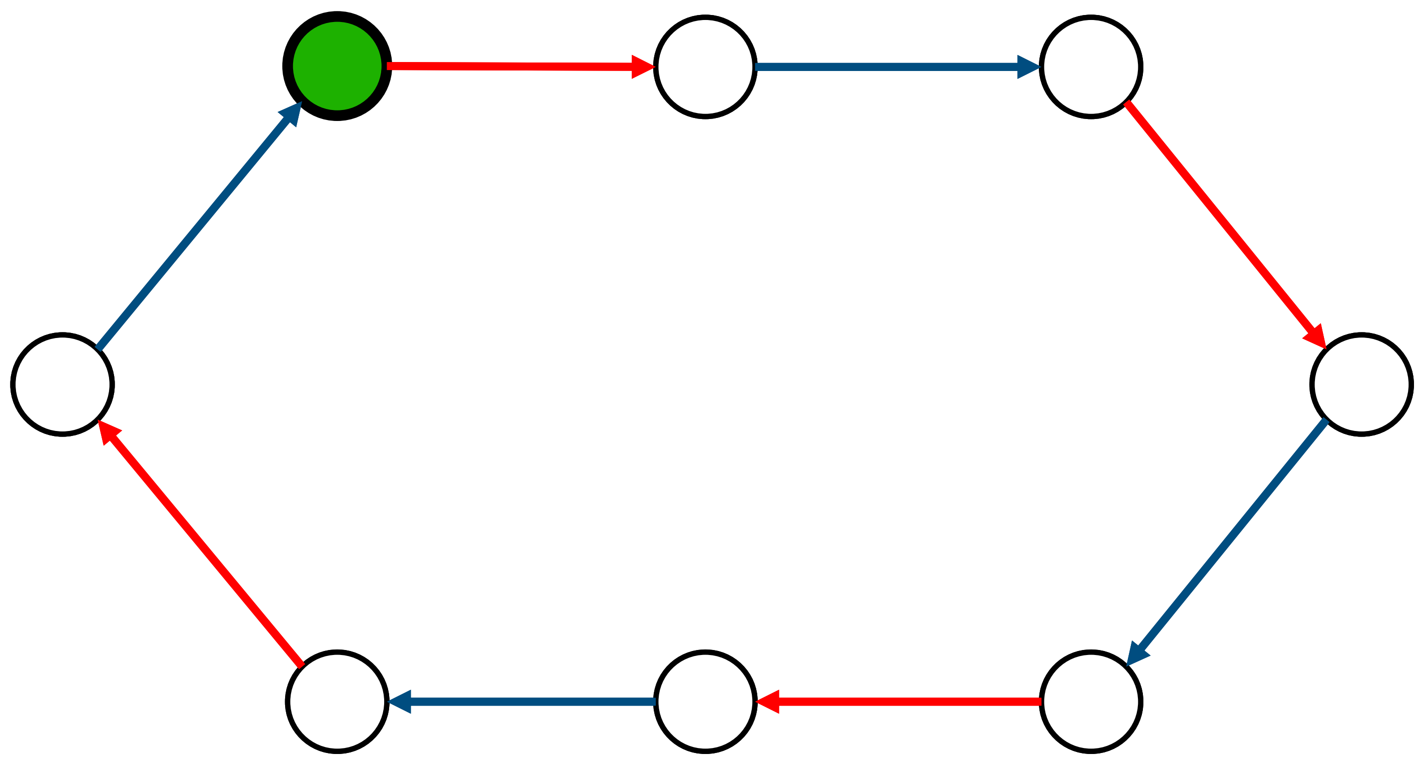


$$\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}(\boldsymbol{h}_{v|u,q}^{f(t)}, \textsc{Agg}(\{\!\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) | \; w \in \mathcal{N}_r(v), r \in R\}\!\!\}), \textsc{Read}(\{\!\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\!\})),$$

Conditional Message Passing Neural Networks

NBFNet [4] locally computes pairwise representations by conditioning.

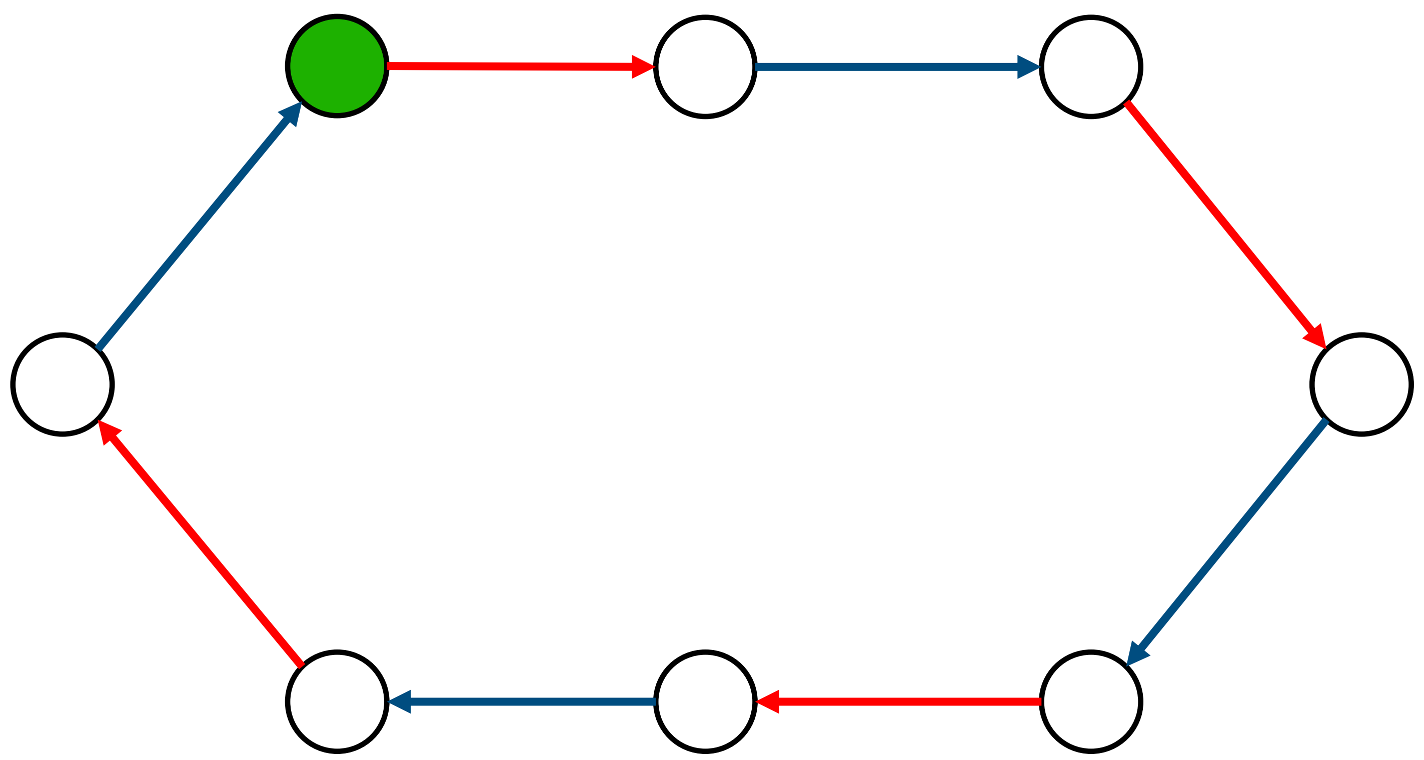


The initialization function must satisfy target node distinguishability.

$$\boxed{\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)}$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}(\boldsymbol{h}_{v|u,q}^{f(t)}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) | \; w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\})),$$

Conditional Message Passing Neural Networks

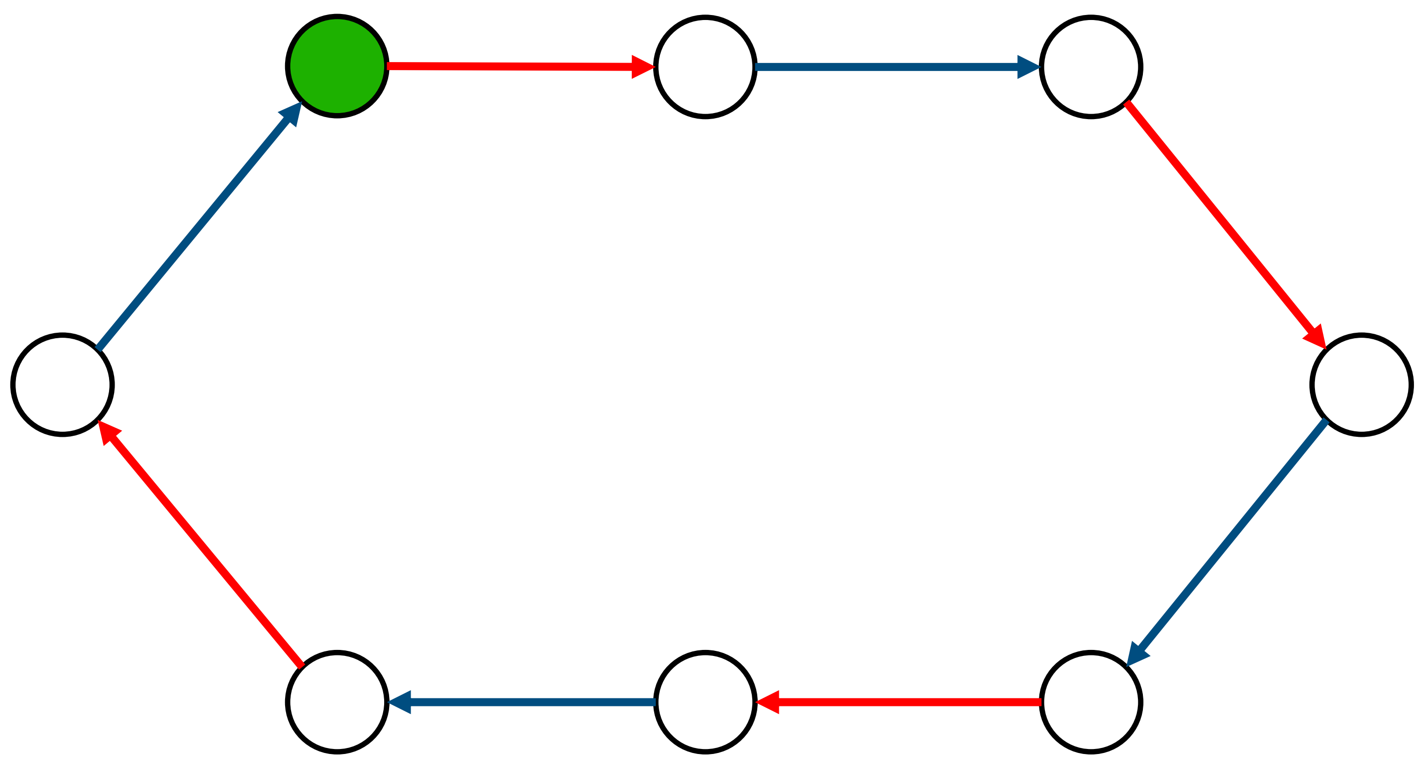


$$\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}\left(\boxed{\boldsymbol{h}_{v|u,q}^{f(t)},} \textsc{Agg}(\{\!\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) \mid w \in \mathcal{N}_r(v), r \in R\}\!\!\}), \textsc{Read}(\{\!\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\!\})\right),$$

Conditional Message Passing Neural Networks

The history function shows which historical self-representation we choose to update.

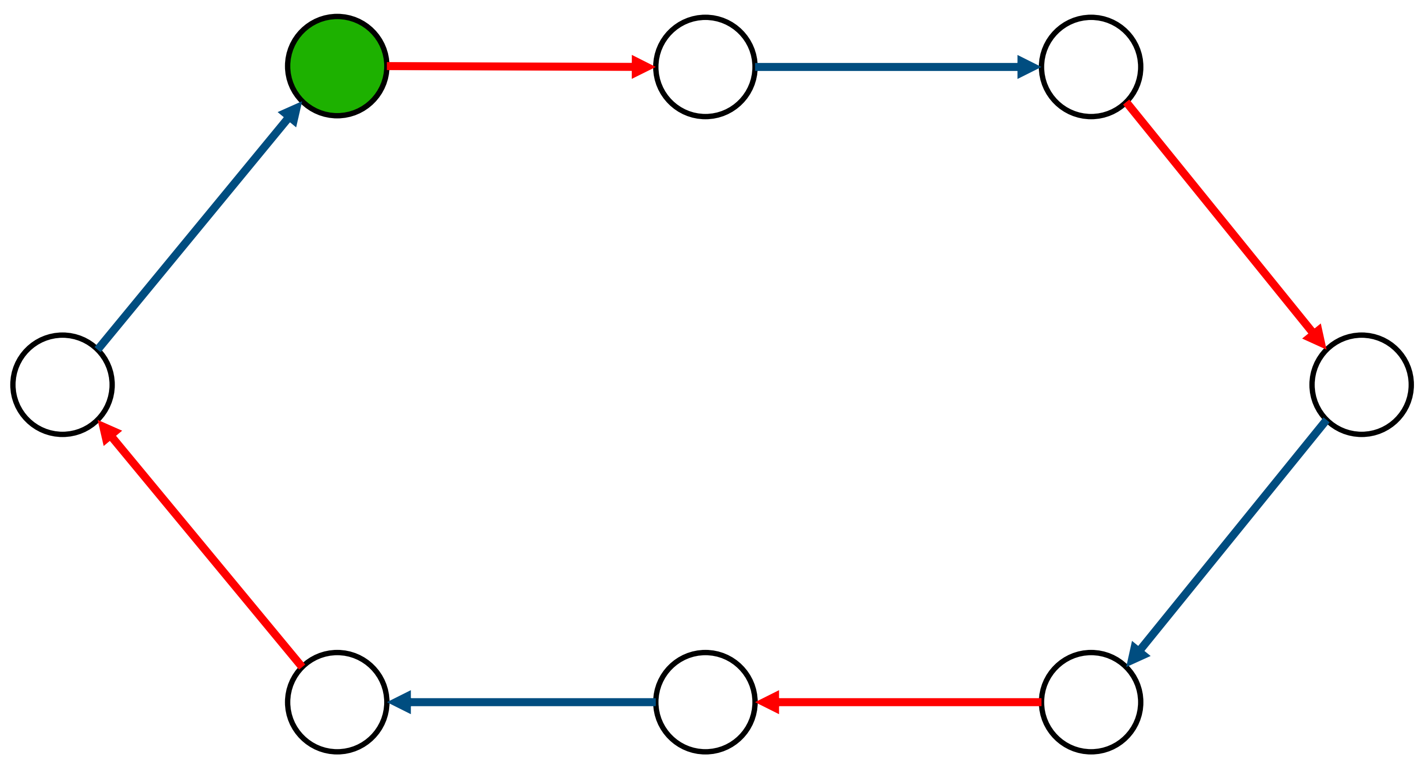


$$\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}(\boxed{\boldsymbol{h}_{v|u,q}^{f(t)}}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) \mid w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\})),$$

Conditional Message Passing Neural Networks

The history function shows which historical self-representation we choose to update.

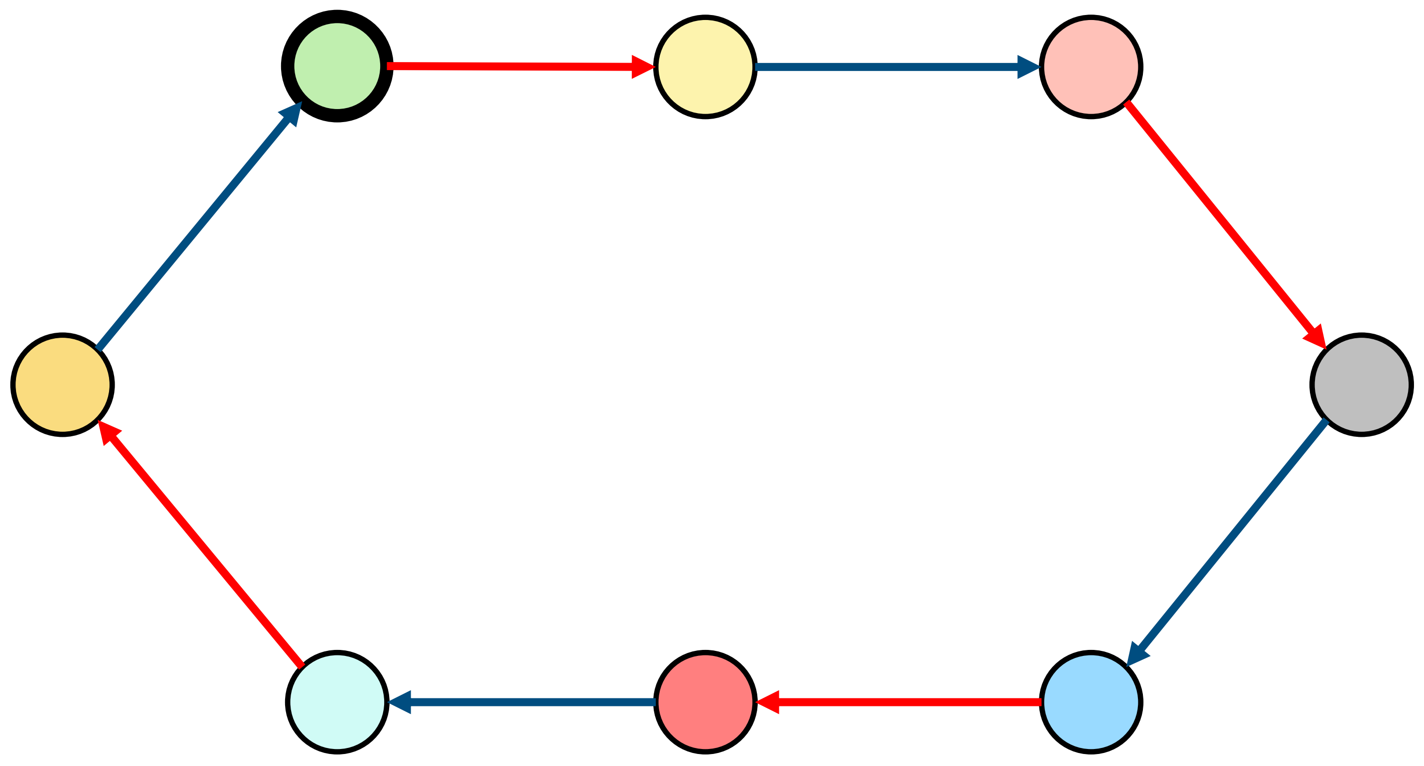


We prove that the choice of the historic function is irrelevant in theoretical expressiveness.

$$\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}(\boxed{\boldsymbol{h}_{v|u,q}^{f(t)},} \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) | \; w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\})),$$

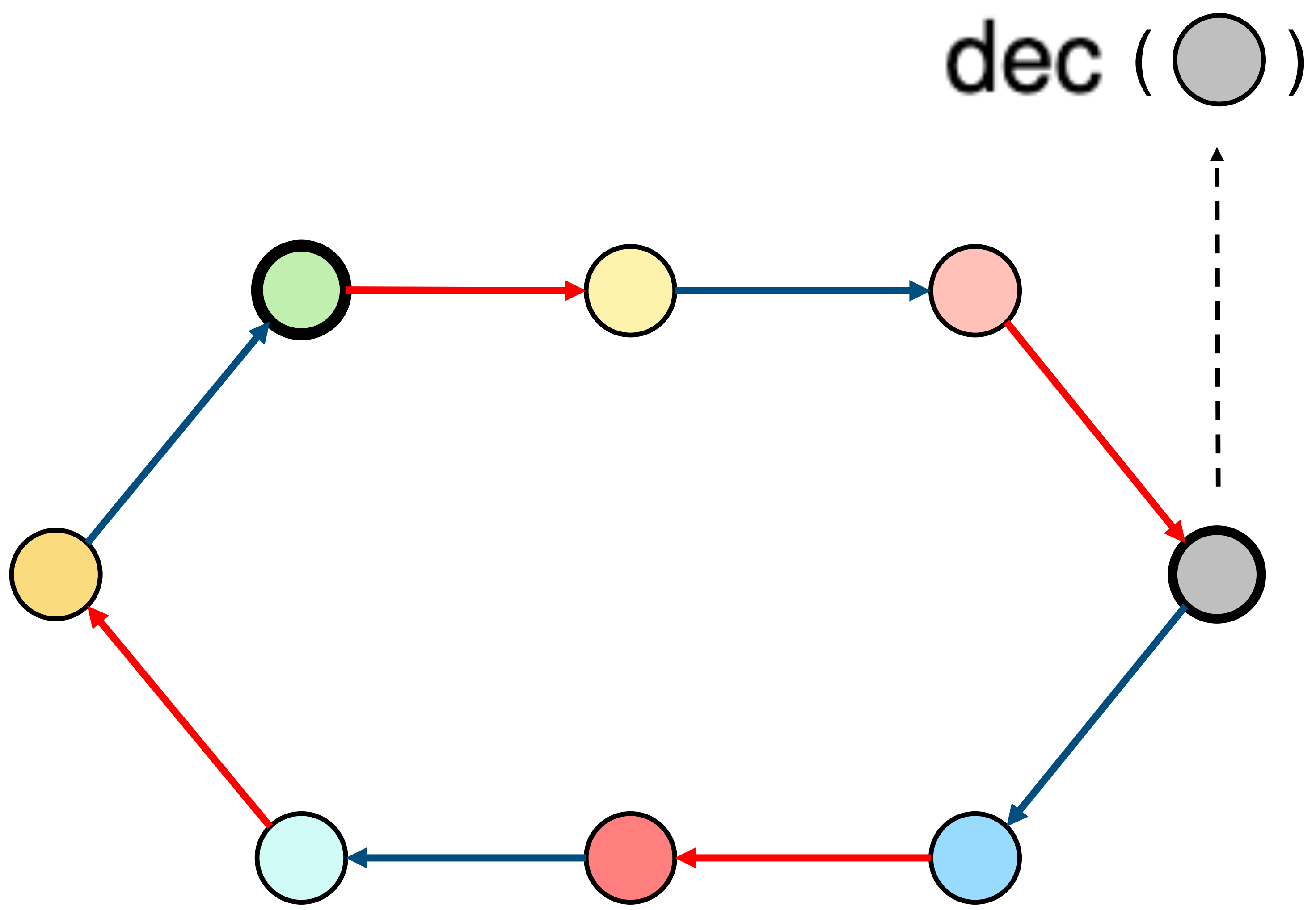
Conditional Message Passing Neural Networks



$$\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}(\boldsymbol{h}_{v|u,q}^{f(t)}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) | \; w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\})),$$

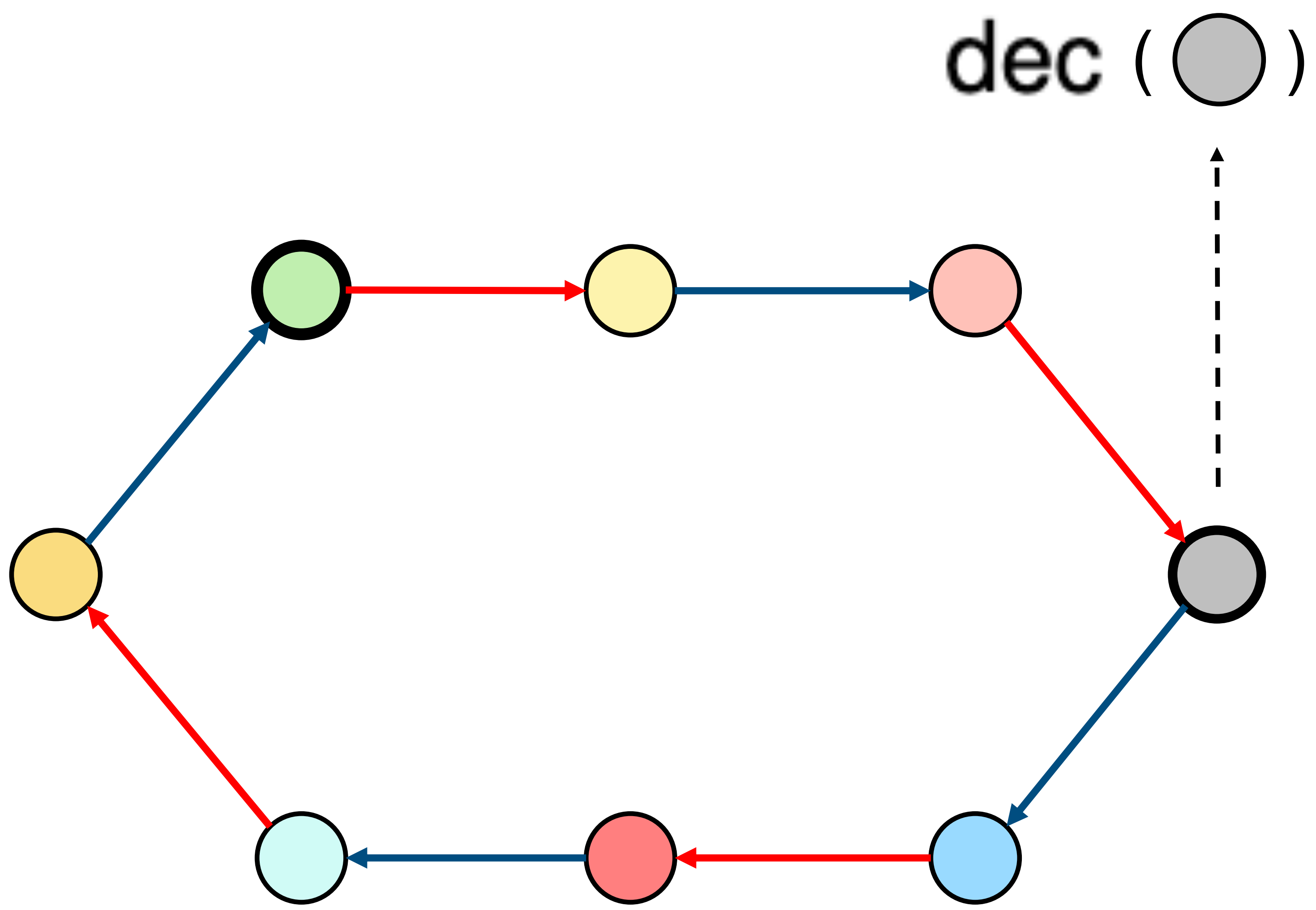
Conditional Message Passing Neural Networks



$$\boldsymbol{h}^{(0)}_{v|u,q} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}^{(t+1)}_{v|u,q} = \textsc{Upd}(\boldsymbol{h}^{f(t)}_{v|u,q}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}^{(t)}_{w|u,q}, \boldsymbol{z}_q) \mid w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}^{(t)}_{w|u,q} \mid w \in V\}\!\})),$$

Conditional Message Passing Neural Networks

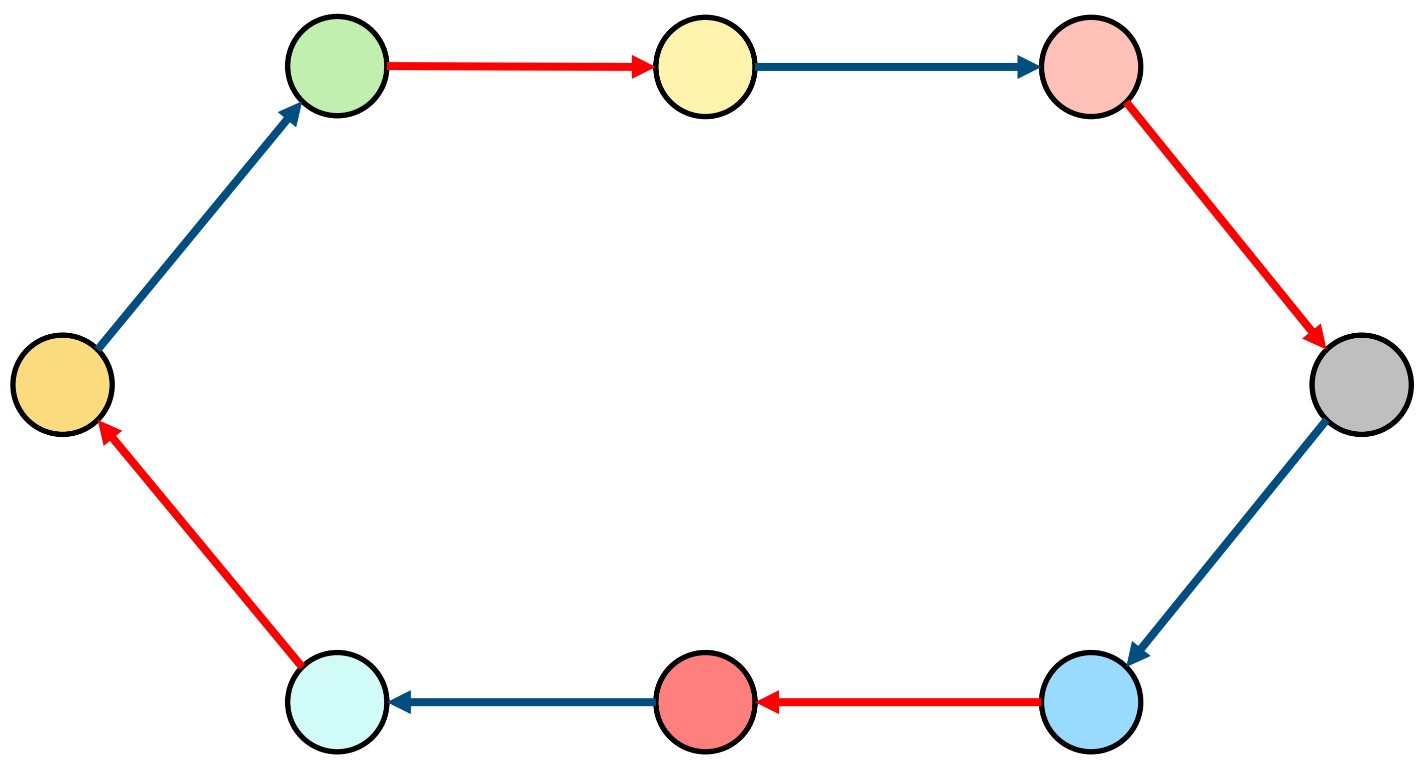


C-MPNNs relies on unary decoder for link prediction.

$$\boldsymbol{h}_{v|u,q}^{(0)} = \textsc{Init}(u, v, q)$$

$$\boldsymbol{h}_{v|u,q}^{(t+1)} = \textsc{Upd}(\boldsymbol{h}_{v|u,q}^{f(t)}, \textsc{Agg}(\{\!\{\textsc{Msg}_r(\boldsymbol{h}_{w|u,q}^{(t)}, \boldsymbol{z}_q) | \; w \in \mathcal{N}_r(v), r \in R\}\!\}), \textsc{Read}(\{\!\{\boldsymbol{h}_{w|u,q}^{(t)} \mid w \in V\}\!\})),$$

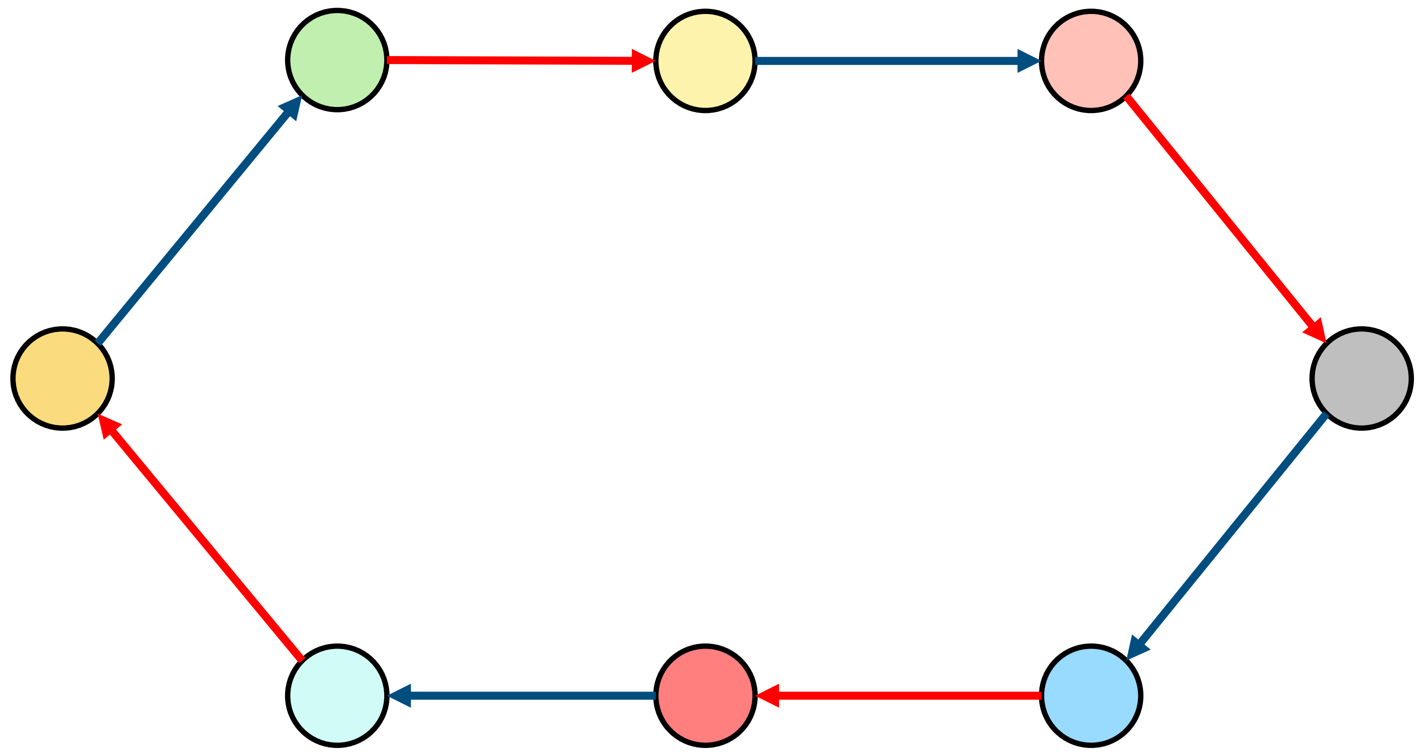
Expressiveness of C-MPNNs



$$\mathsf{rawl}_2^{(0)}(u,v) = \eta(u,v),$$
$$\mathsf{rawl}_2^{(t+1)}(u,v) = \tau\big(\mathsf{rawl}_2^{(t)}(u,v), \{\!\{(\mathsf{rawl}_2^{(t)}(u,w), r) \mid w \in \mathcal{N}_r(v), r \in R)\}\!\}\big),$$

Expressiveness of C-MPNNs

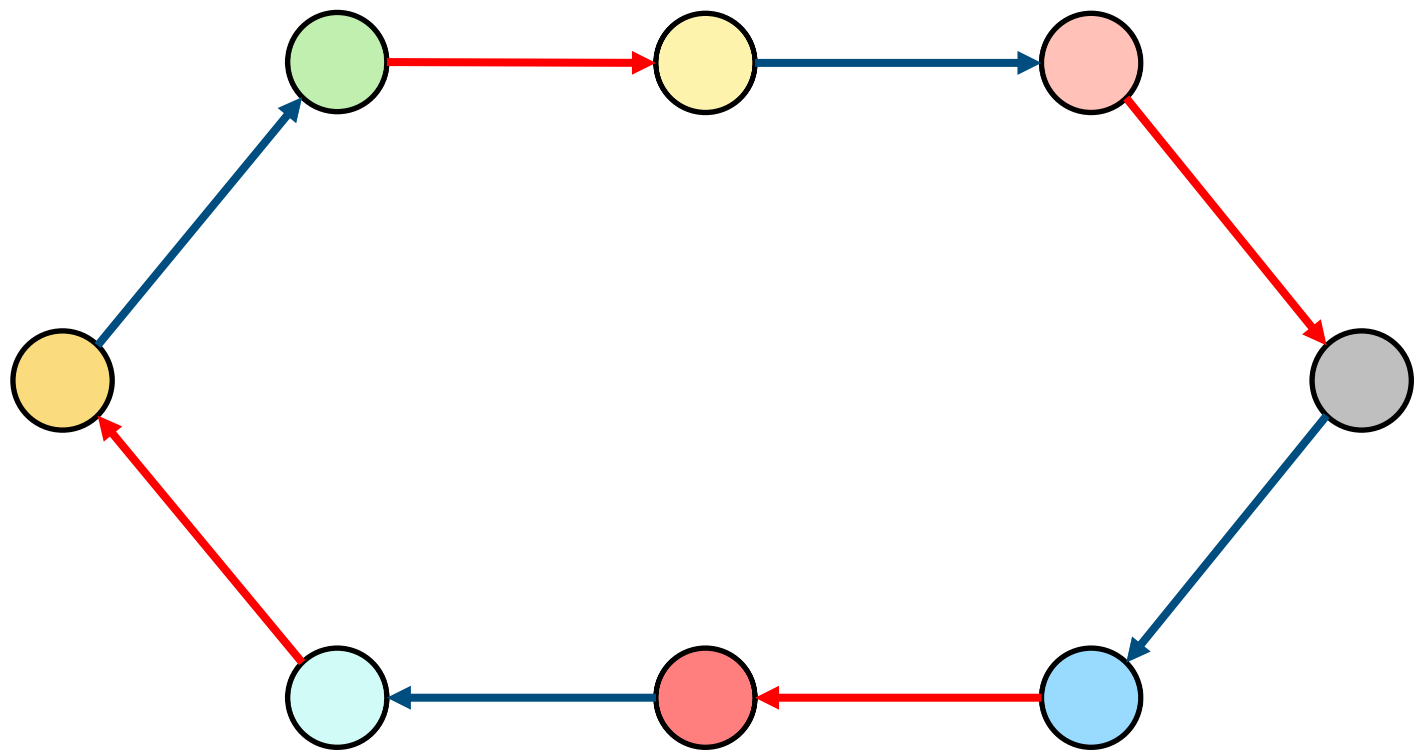
C-MPNNs are at most as expressive as relational asymmetric local 2-WL (rawl_2).



$$\mathsf{rawl}_2^{(0)}(u,v) = \eta(u,v),$$
$$\mathsf{rawl}_2^{(t+1)}(u,v) = \tau\big(\mathsf{rawl}_2^{(t)}(u,v), \{\!\{(\mathsf{rawl}_2^{(t)}(u,w), r) \mid w \in \mathcal{N}_r(v), r \in R)\}\!\}\big),$$

Expressiveness of C-MPNNs

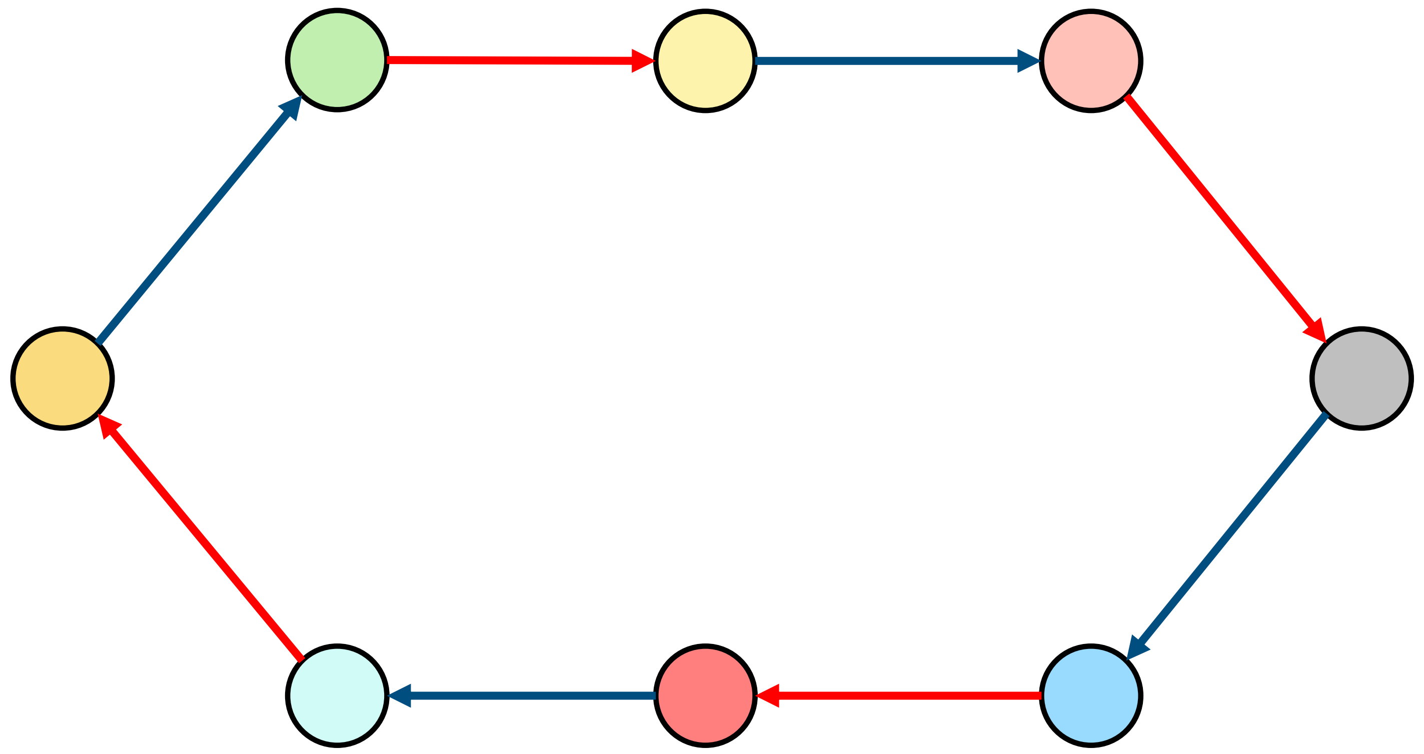
C-MPNNs are at most as expressive as relational asymmetric local 2-WL (rawl_2).



There exists a C-MPNN (even without readout) that achieves the same express power of rawl_2.

$$\text{rawl}_2^{(0)}(u,v) = \eta(u,v),$$
$$\text{rawl}_2^{(t+1)}(u,v) = \tau\big(\text{rawl}_2^{(t)}(u,v), \{\!\{(\text{rawl}_2^{(t)}(u,w), r) \mid w \in \mathcal{N}_r(v), r \in R)\}\!\}\big),$$

Logical Characterization of C-MPNNs

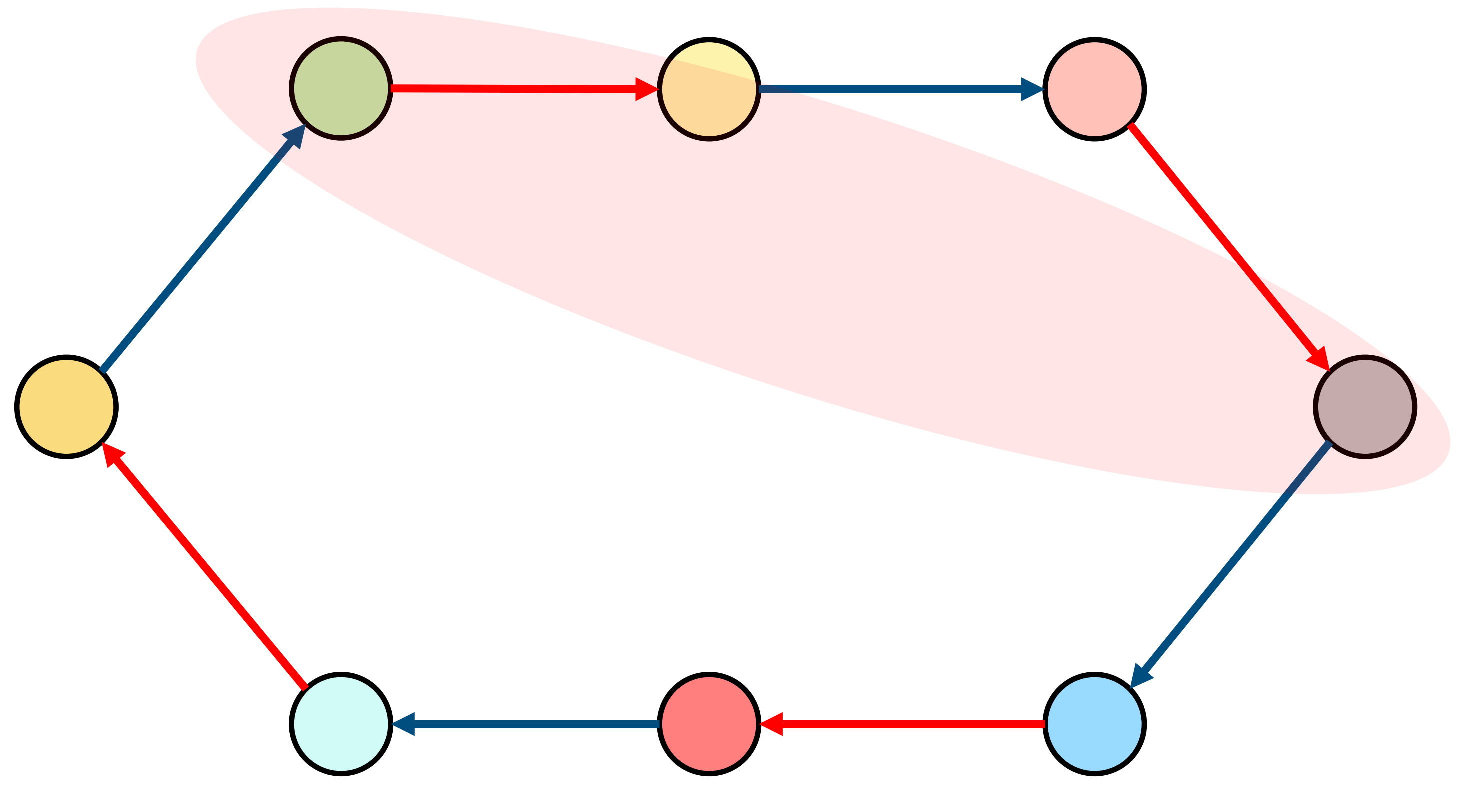


C-MPNNs (without readout) can uniformly express precisely functions in $\mathsf{rFO}^3_{\mathrm{cnt}}$.

C-MPNNs (with readout) can uniformly express all functions in $\mathsf{erFO}^3_{\mathrm{cnt}}$.

Summary and Outlook

Pairwise representation: C-MPNNs encodes pairwise node representations conditioned on source node.



Expressiveness results: C-MPNNs can match the expressive power of relational asymmetric local 2-WL, and logical characterizations.

Experimental validation: Experimental analysis is carried out to verify the impact of model choices to validate our theoretical findings.

Thank you!

Selected References

[1] Michael Sejr Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In ESWC, 2018.

[2] Shikhar Vashishth, Soumya Sanyal, Vikram Nitin, and Partha Talukdar. Composition-based multi-relational graph convolutional networks. In ICLR, 2020.

[3] Pablo Barceló, Mikhail Galkin, Christopher Morris, and Miguel Romero. Weisfeiler and leman go relational. In LoG, 2022.

[4] Zhaocheng Zhu, Zuobai Zhang, Louis-Pascal Xhonneux, and Jian Tang. Neural bellman-ford networks: A general graph neural network framework for link prediction. In NeurIPS, 2021.