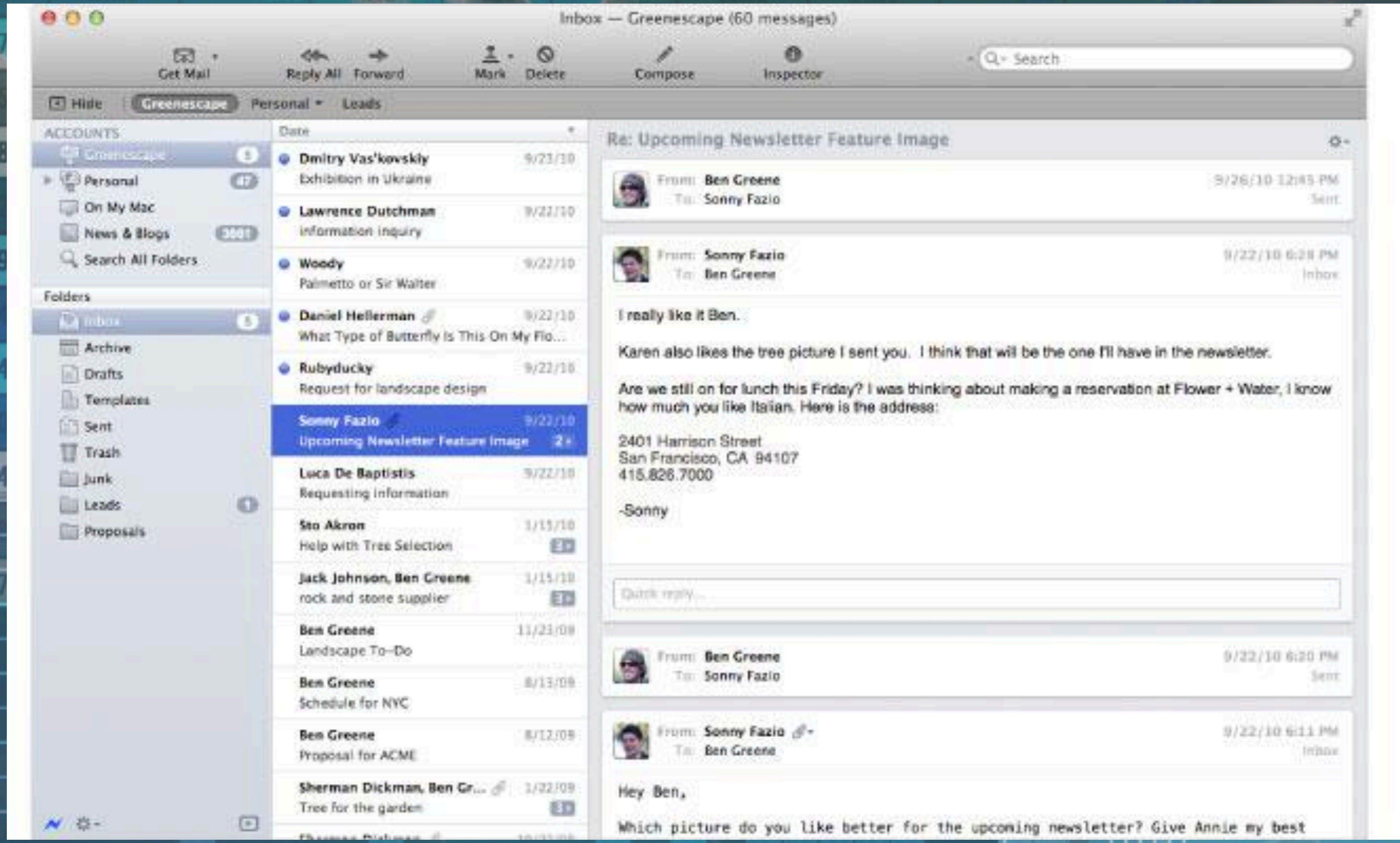
Reinforcement Learning for Knowledge Graph Reasoning

Knowledge Connexions Conference 2020

Victoria Lin
Senior Research Scientist, Salesforce AI Research

Dec 2, 2020

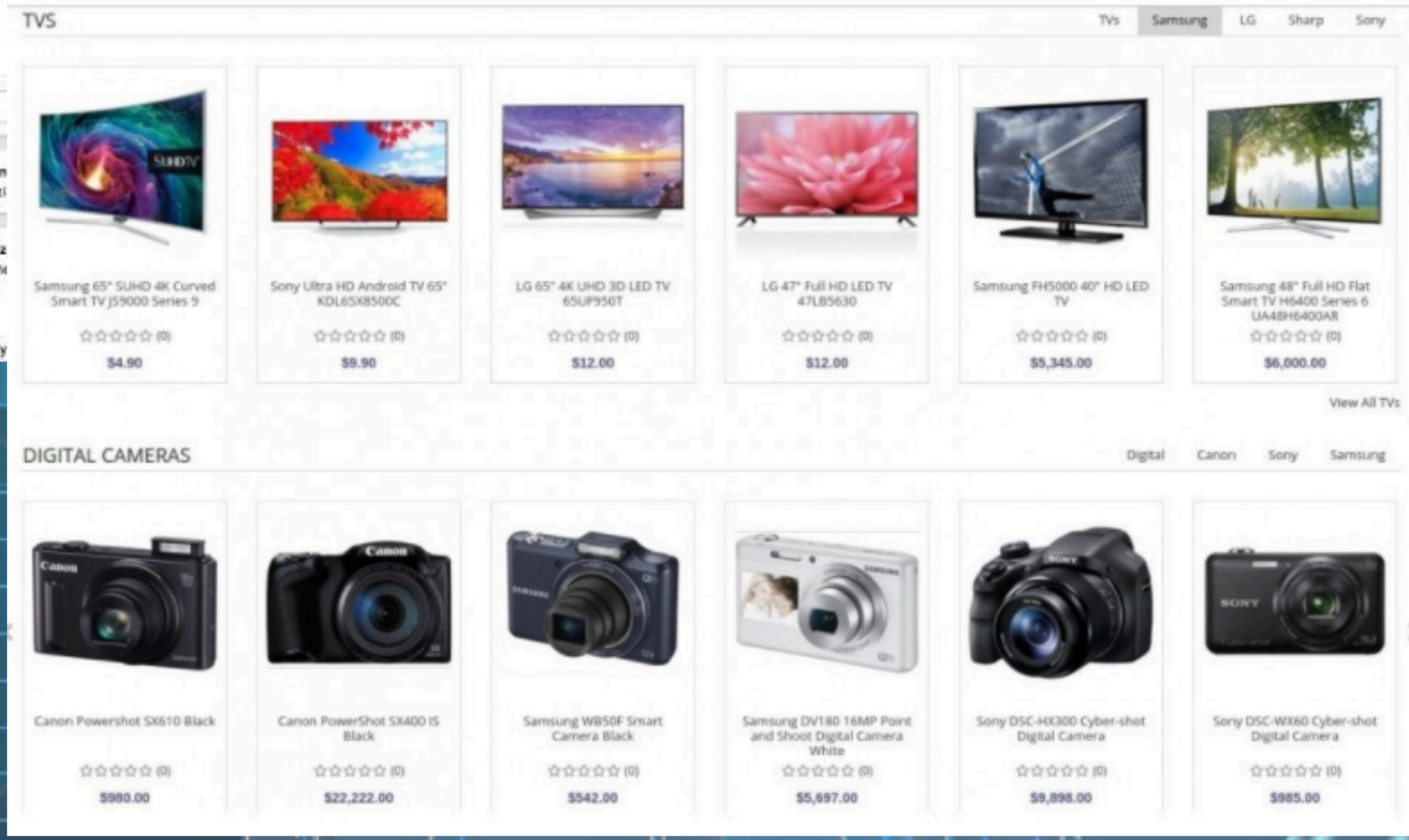
@VictoriaLinML



Contact graph

Knowledge Graphs (KGs) are powerhouses for many downstream AI applications.

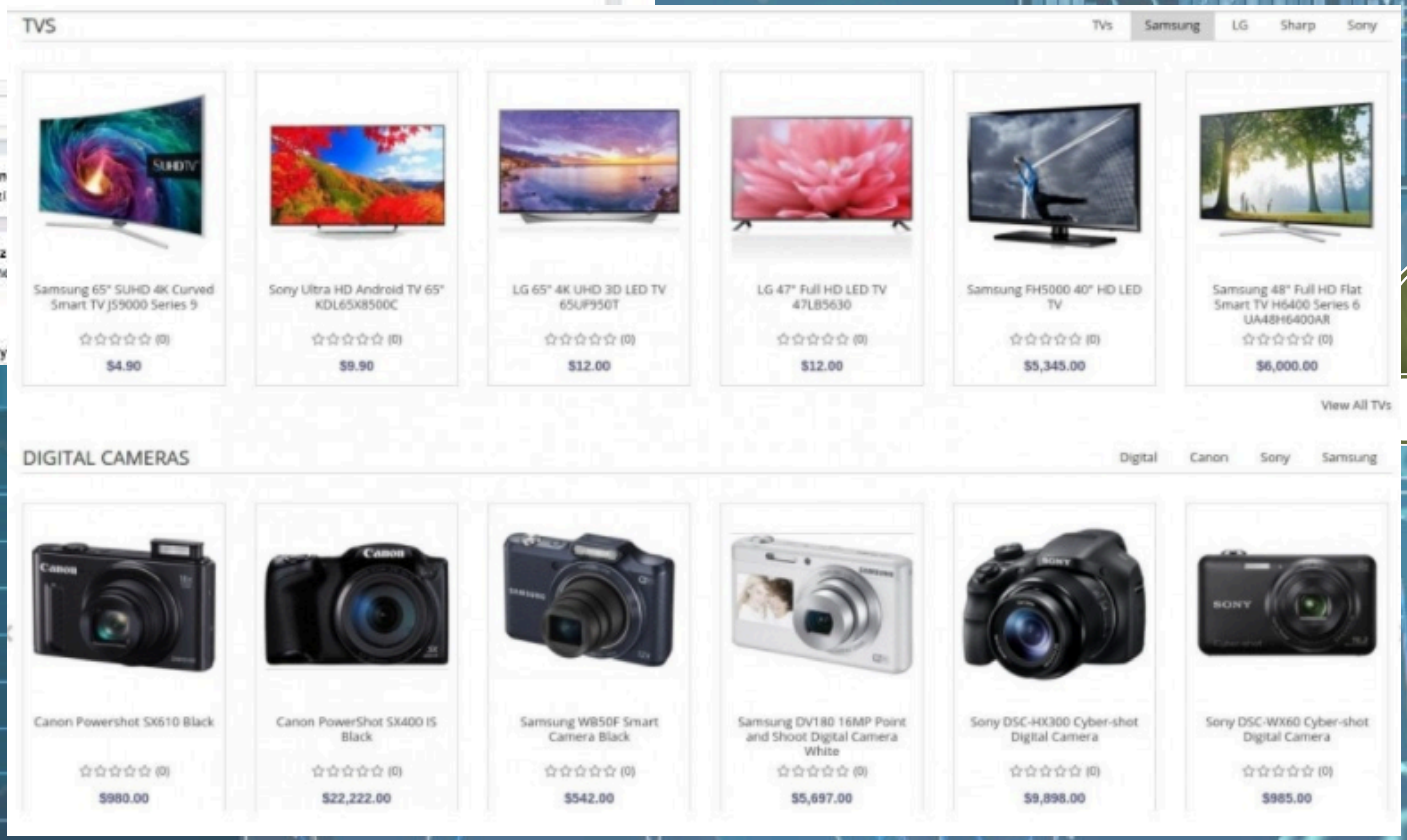
Contact graph



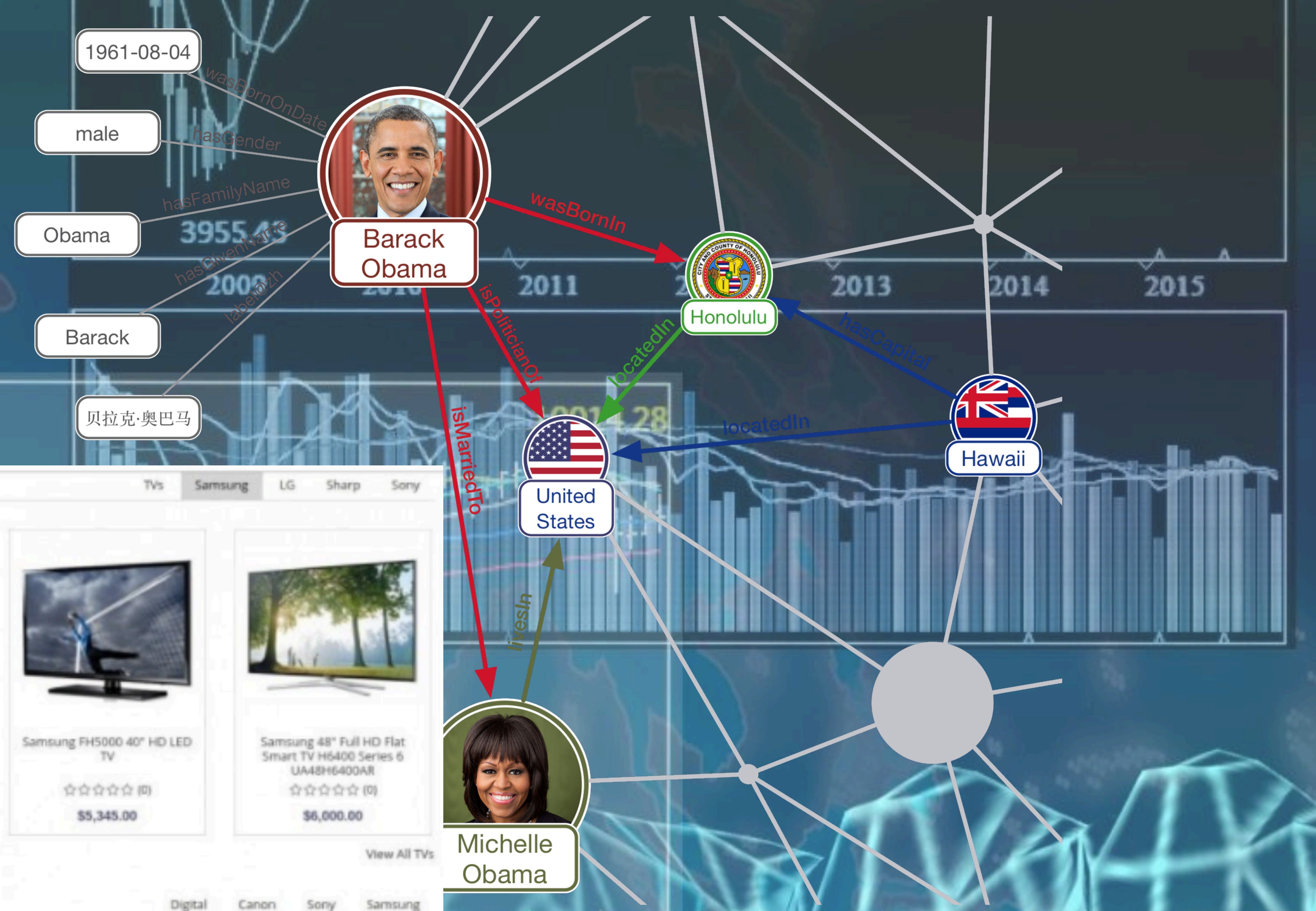
Product graph

Knowledge Graphs (KGs) are powerhouses for many downstream AI applications.

Contact graph



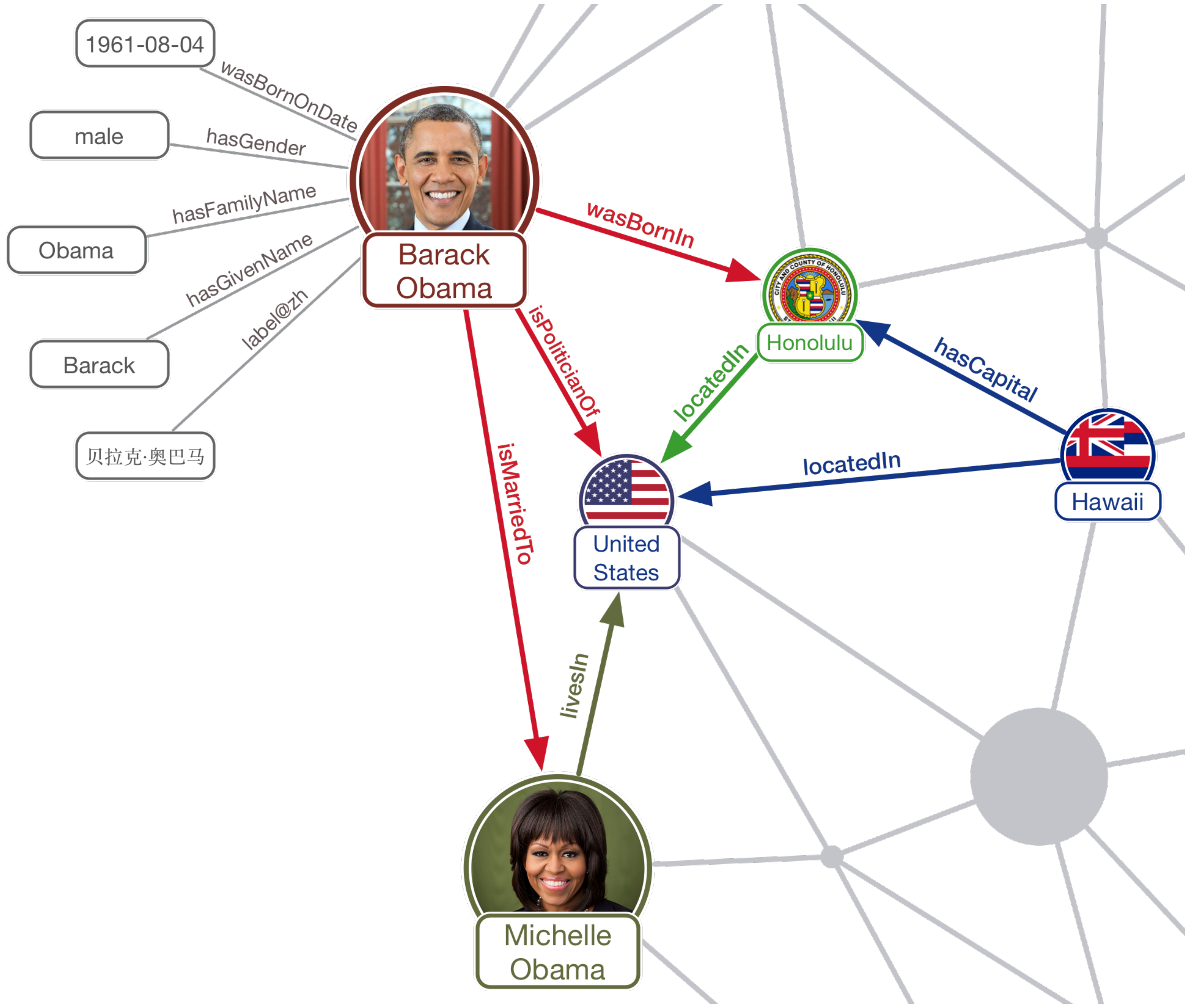
Product graph



Open-domain KG

Knowledge Graphs (KGs) are powerhouses for many downstream AI applications.

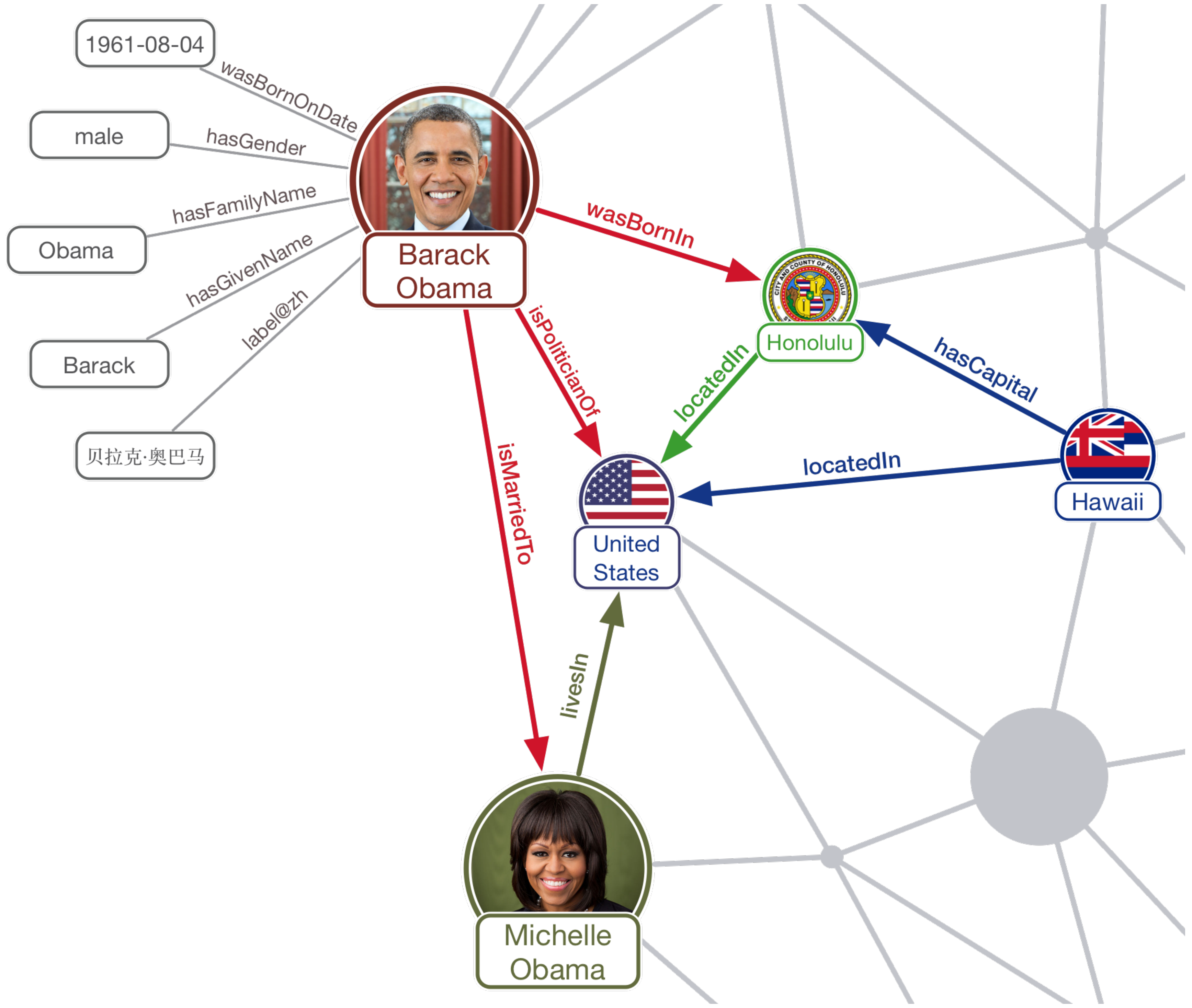
Knowledge Graph Search



Search:

Look up for facts that exist in the knowledge graph via a formal query language (e.g. SPARQL)

Knowledge Graph Reasoning



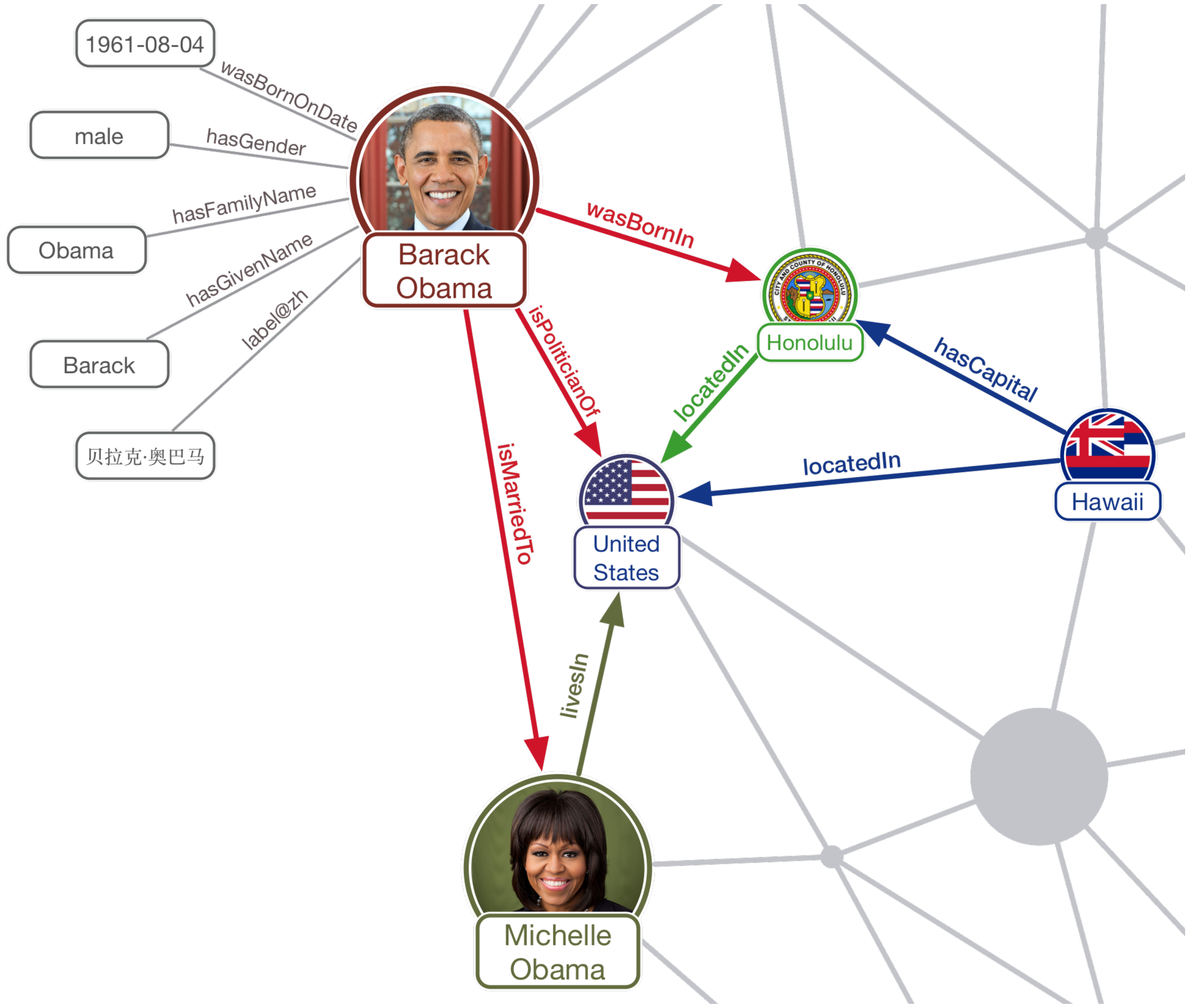
Inference/Reasoning:

Derive additional knowledge given existing facts

Barack Obama $\xrightarrow{\textit{WasBornIn}}$ Honolulu $\xleftarrow{\textit{HasCapital}}$ Hawaii

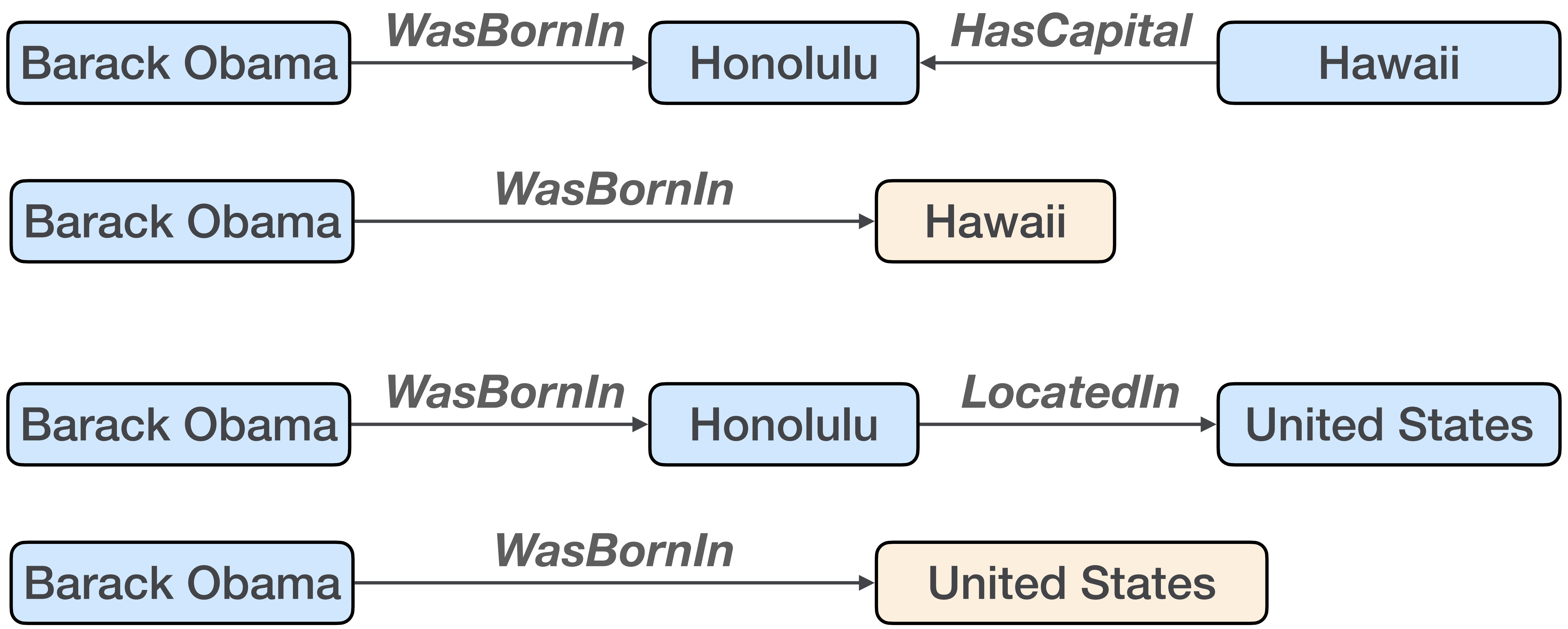
Barack Obama $\xrightarrow{\textit{WasBornIn}}$ Hawaii

Knowledge Graph Reasoning

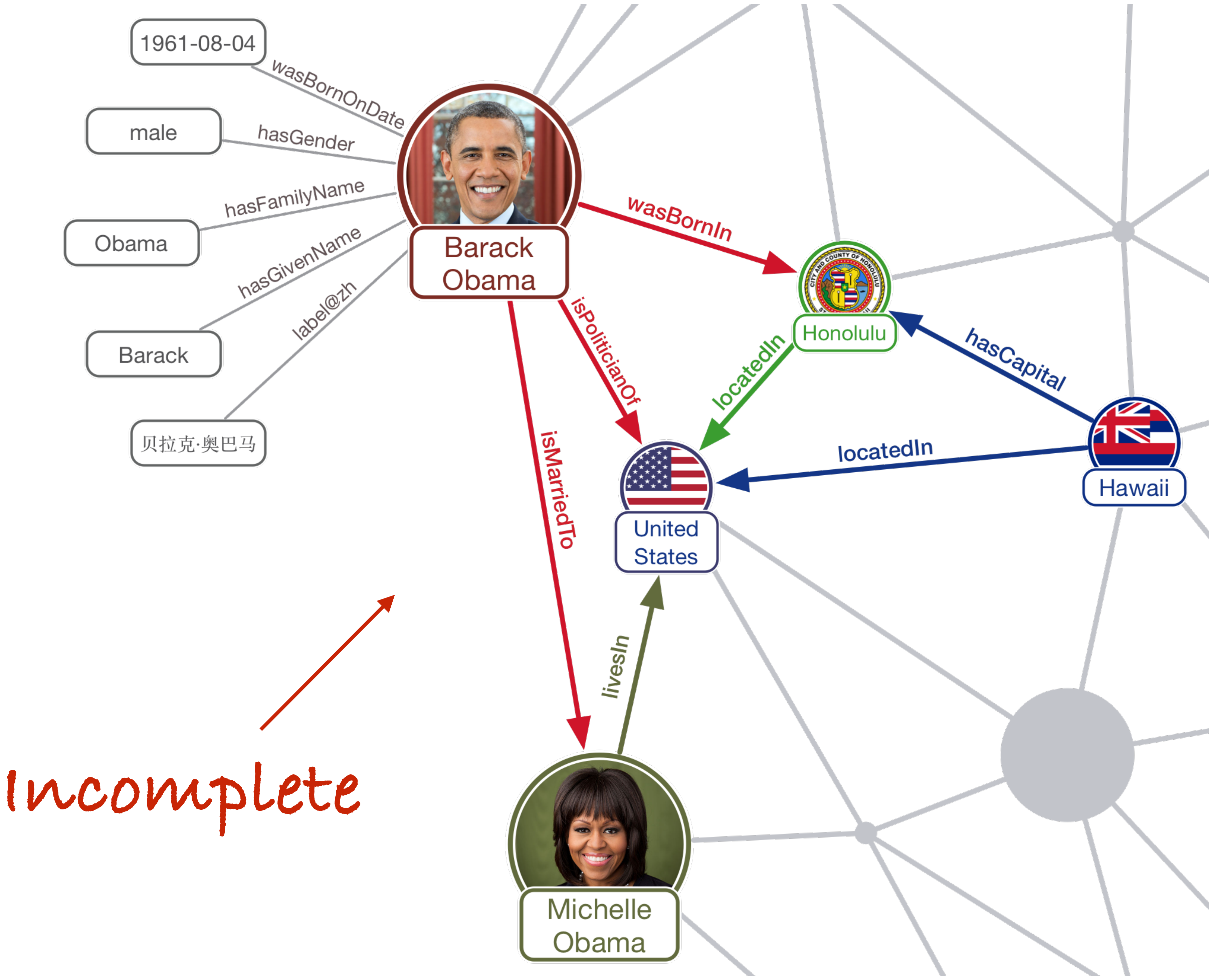


Inference/Reasoning:

Derive additional knowledge given existing facts



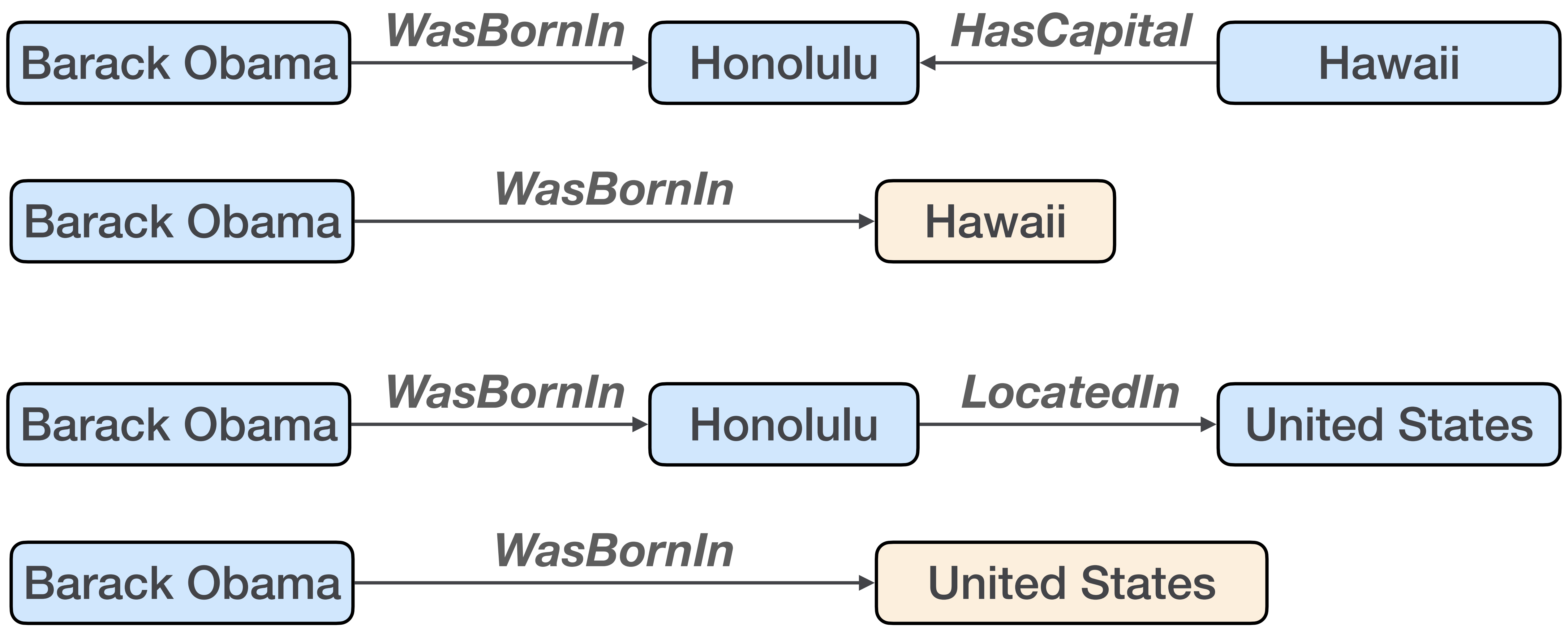
Knowledge Graph Reasoning



Reasoning is a core problem for KGs as most often it is impossible to curate and store all facts in a KG.

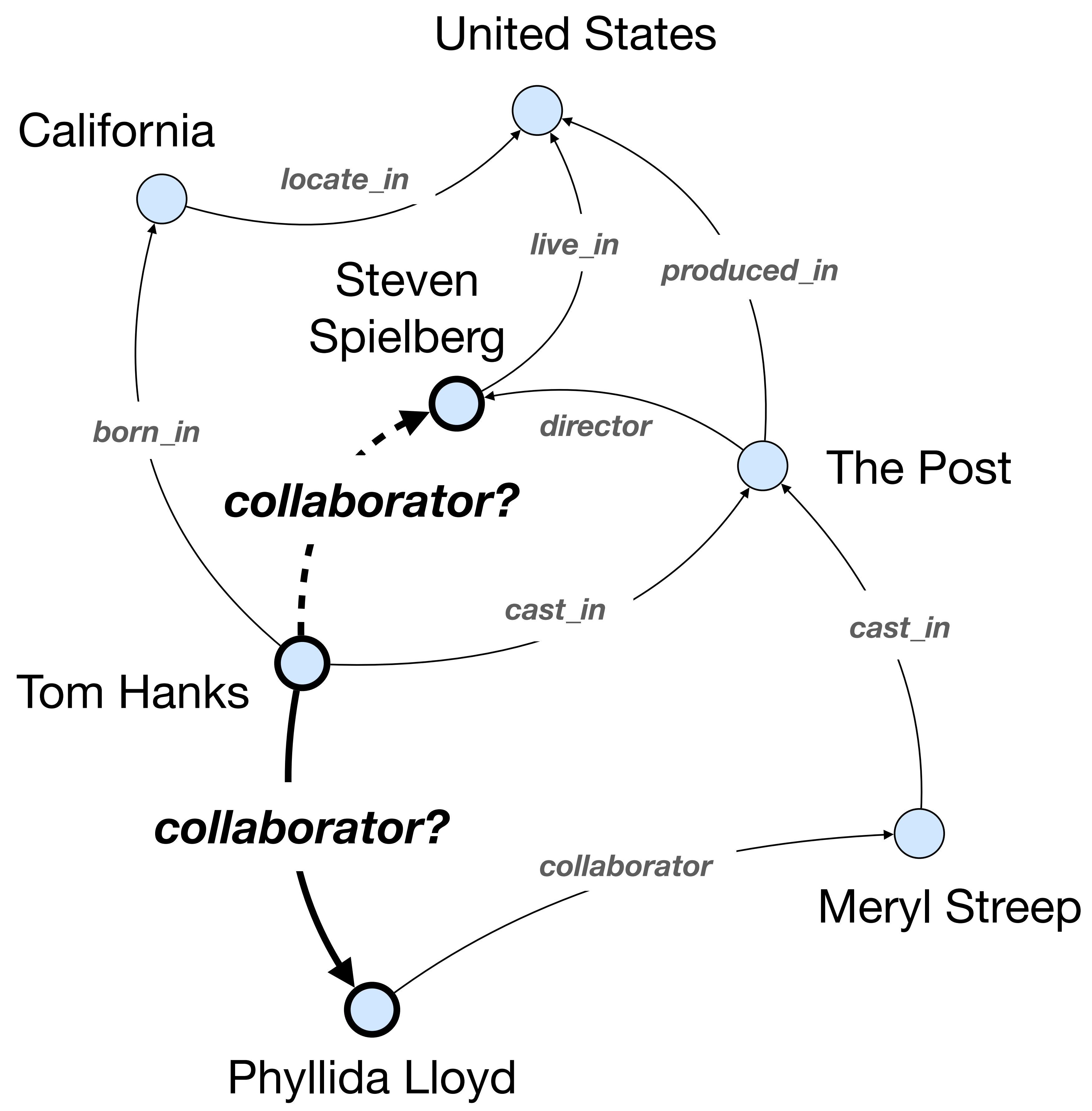
Inference/Reasoning:

Derive additional knowledge given existing facts



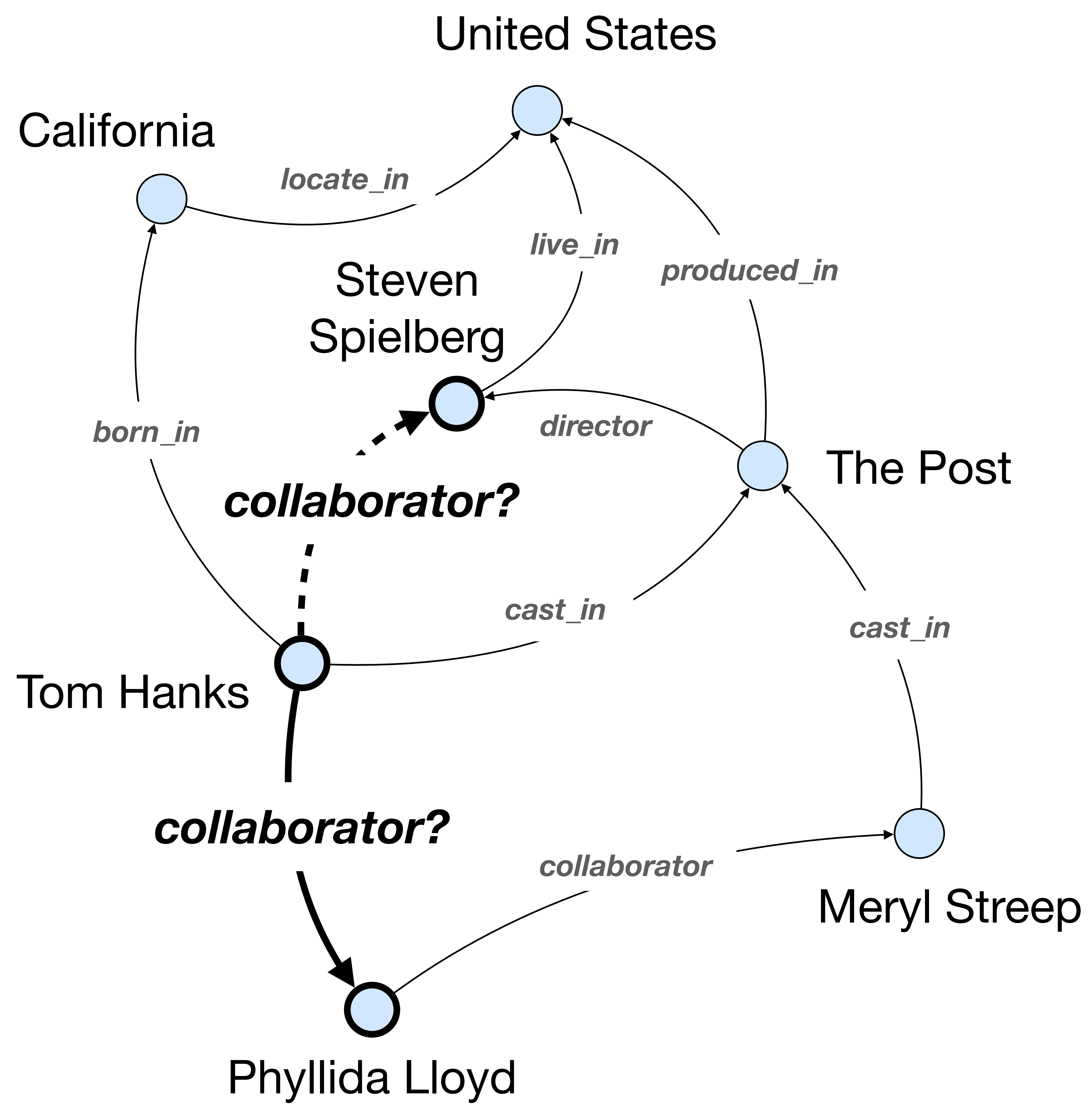
Knowledge Graph Reasoning

- Knowledge Graph Embeddings
- Path Ranking Algorithm (PRA) RL
- Sequential Decision Making RL

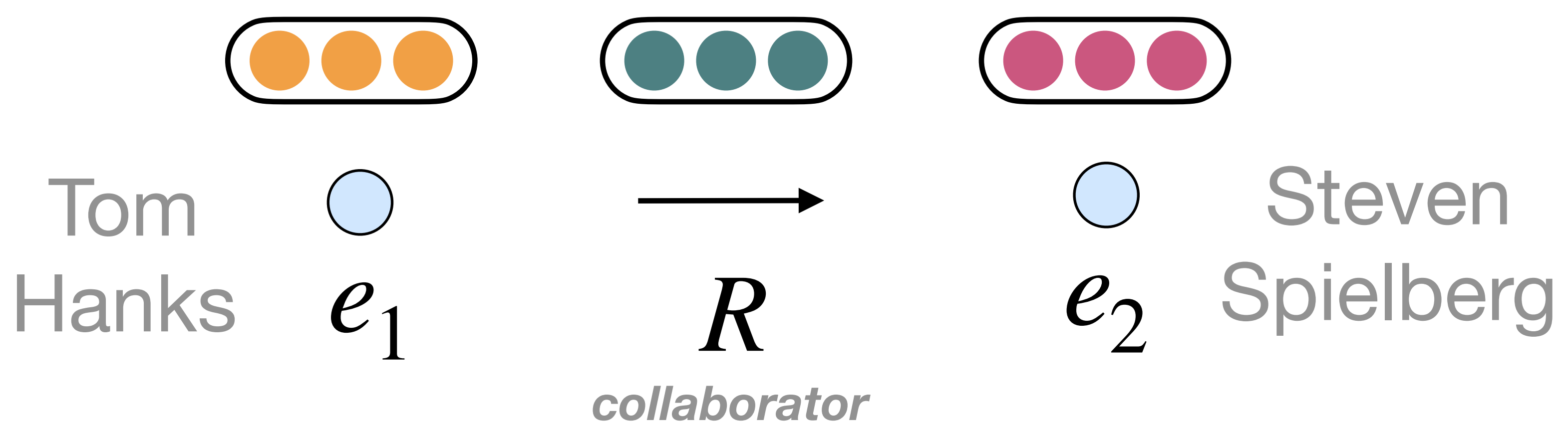


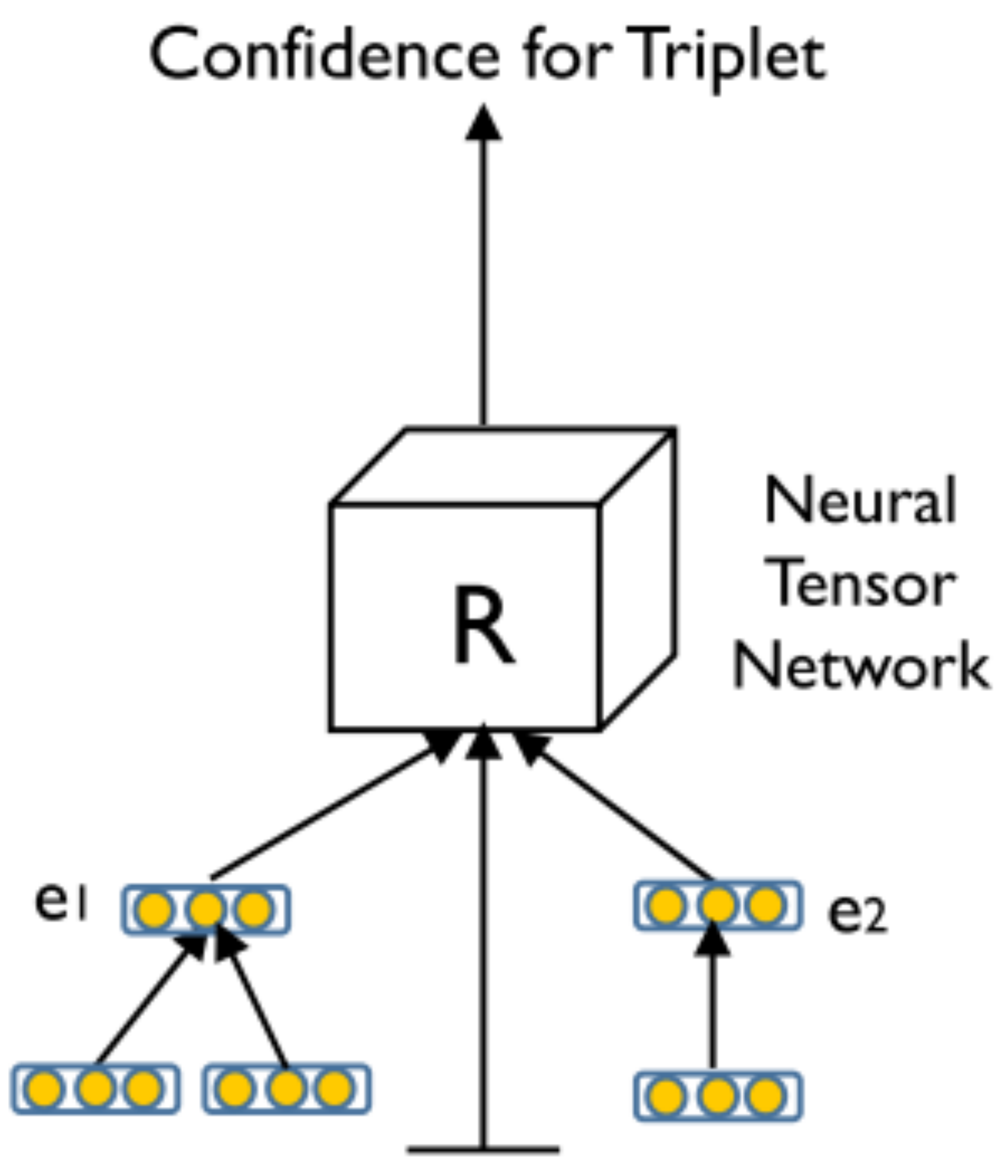
Knowledge Graph Reasoning

- Knowledge Graph Embeddings
- Path Ranking Algorithm (PRA)
- Reinforcement Learning



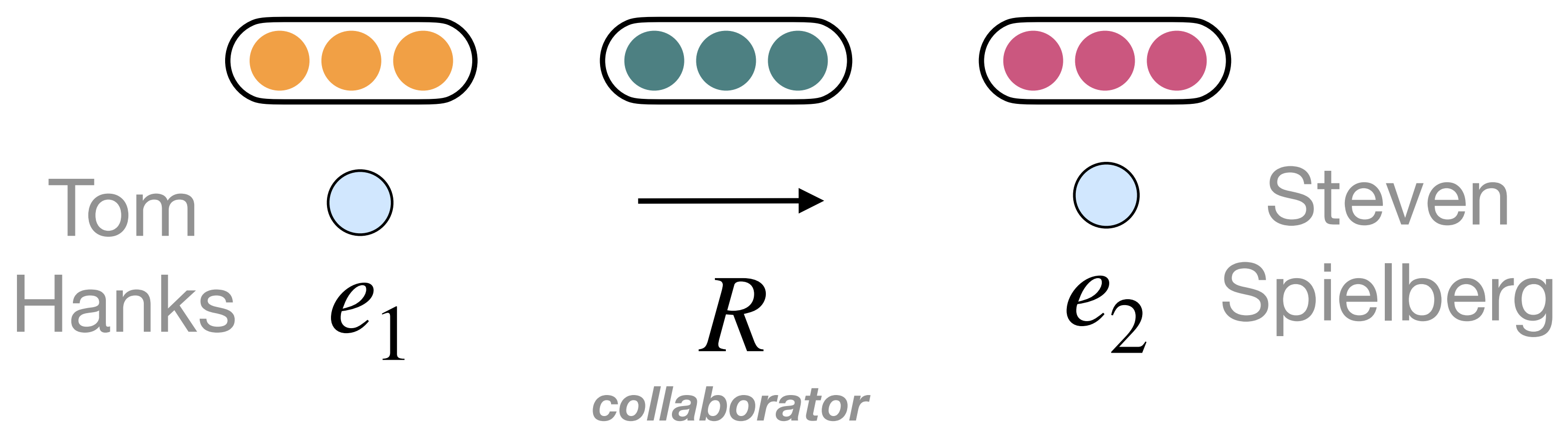
Knowledge Graph Embeddings





NTN (Socher et. al. 2013), DistMult (Yang et al. 2015), ComplEx (Trouillon et al. 2016), ConvE (Dettmers et al. 2018).

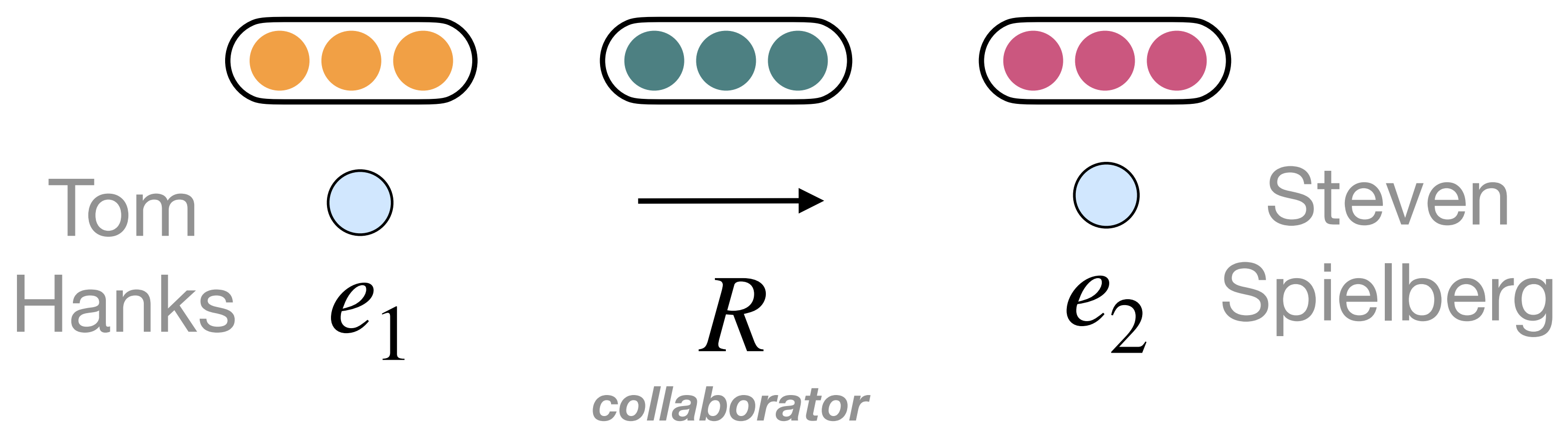
Knowledge Graph Embeddings



Model	Scoring Function $\psi_r(\mathbf{e}_s, \mathbf{e}_o)$	Relation Parameters	Space Complexity
SE (Bordes et al. 2014)	$\left\lVert \mathbf{W}_r^L \mathbf{e}_s - \mathbf{W}_r^R \mathbf{e}_o \right\rVert_p$	$\mathbf{W}_r^L, \mathbf{W}_r^R \in \mathbb{R}^{k \times k}$	$\mathcal{O}(n_e k + n_r k^2)$
TransE (Bordes et al. 2013a)	$\lVert \mathbf{e}_s + \mathbf{r}_r - \mathbf{e}_o \rVert_p$	$\mathbf{r}_r \in \mathbb{R}^k$	$\mathcal{O}(n_e k + n_r k)$
DistMult (Yang et al. 2015)	$\langle \mathbf{e}_s, \mathbf{r}_r, \mathbf{e}_o \rangle$	$\mathbf{r}_r \in \mathbb{R}^k$	$\mathcal{O}(n_e k + n_r k)$
ComplEx (Trouillon et al. 2016)	$\langle \mathbf{e}_s, \mathbf{r}_r, \mathbf{e}_o \rangle$	$\mathbf{r}_r \in \mathbb{C}^k$	$\mathcal{O}(n_e k + n_r k)$
ConvE	$f(\mathrm{vec}(f([\overline{\mathbf{e}_s}; \overline{\mathbf{r}_r}] * \omega))\mathbf{W})\mathbf{e}_o$	$\mathbf{r}_r \in \mathbb{R}^{k'}$	$\mathcal{O}(n_e k + n_r k')$

Convolutional 2D Knowledge Graph Embeddings. Dettmers et al. 2018.

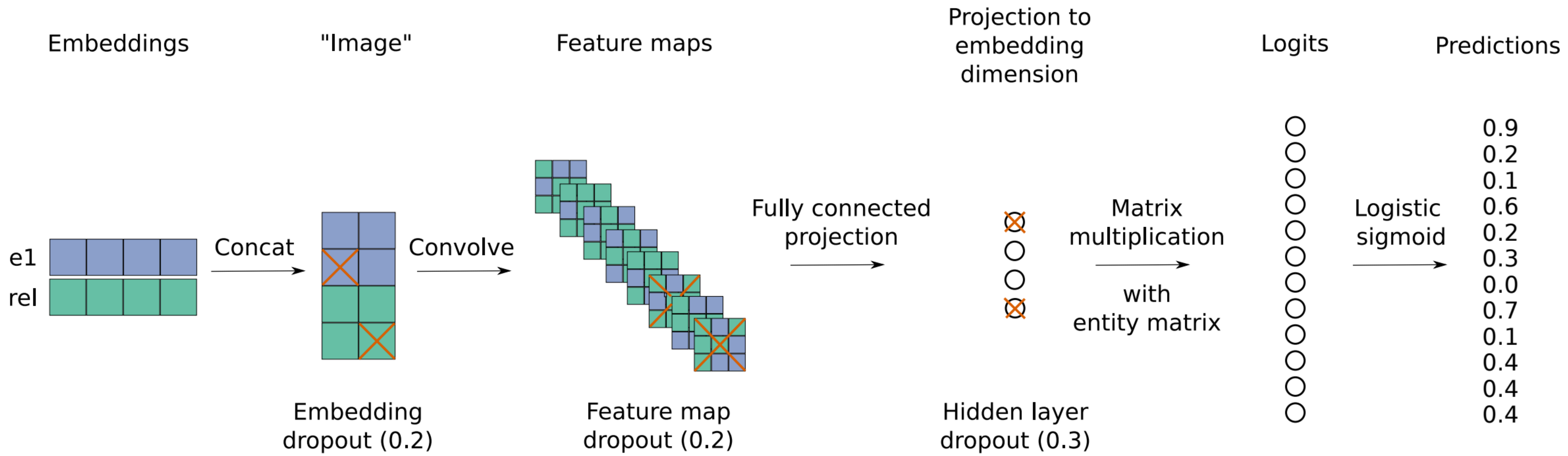
Knowledge Graph Embeddings



Model	Scoring Function $\psi_r(\mathbf{e}_s, \mathbf{e}_o)$	Relation Parameters	Space Complexity
SE (Bordes et al. 2014)	$\left\| \mathbf{W}_r^L \mathbf{e}_s - \mathbf{W}_r^R \mathbf{e}_o \right\|_p$	$\mathbf{W}_r^L, \mathbf{W}_r^R \in \mathbb{R}^{k \times k}$	$\mathcal{O}(n_e k + n_r k^2)$
TransE (Bordes et al. 2013a)	$\left\| \mathbf{e}_s + \mathbf{r}_r - \mathbf{e}_o \right\|_p$	$\mathbf{r}_r \in \mathbb{R}^k$	$\mathcal{O}(n_e k + n_r k)$
DistMult (Yang et al. 2015)	$\langle \mathbf{e}_s, \mathbf{r}_r, \mathbf{e}_o \rangle$	$\mathbf{r}_r \in \mathbb{R}^k$	$\mathcal{O}(n_e k + n_r k)$
ComplEx (Trouillon et al. 2016)	$\langle \mathbf{e}_s, \mathbf{r}_r, \mathbf{e}_o \rangle$	$\mathbf{r}_r \in \mathbb{C}^k$	$\mathcal{O}(n_e k + n_r k)$
ConvE	$f(\text{vec}(f([\overline{\mathbf{e}_s}; \overline{\mathbf{r}_r}] * \omega))\mathbf{W})\mathbf{e}_o$	$\mathbf{r}_r \in \mathbb{R}^{k'}$	$\mathcal{O}(n_e k + n_r k')$

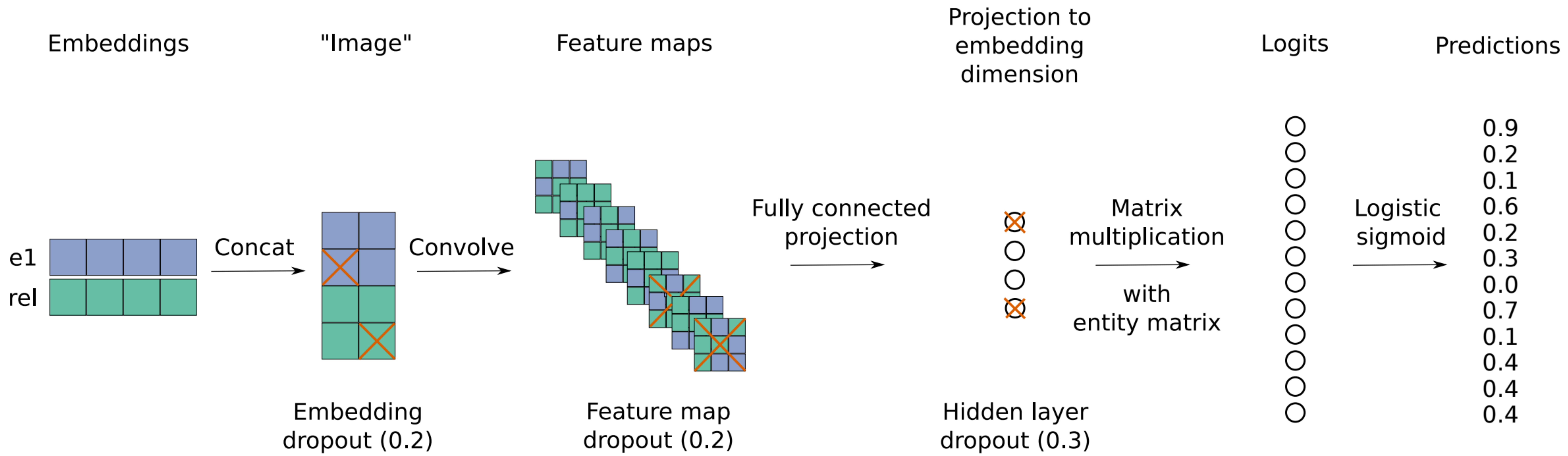
Convolutional 2D Knowledge Graph Embeddings. Dettmers et al. 2018.

ConvE: Convolutional 2D Knowledge Graph Embeddings



Convolutional 2D Knowledge Graph Embeddings. Dettmers et al. 2018.

ConvE: Convolutional 2D Knowledge Graph Embeddings

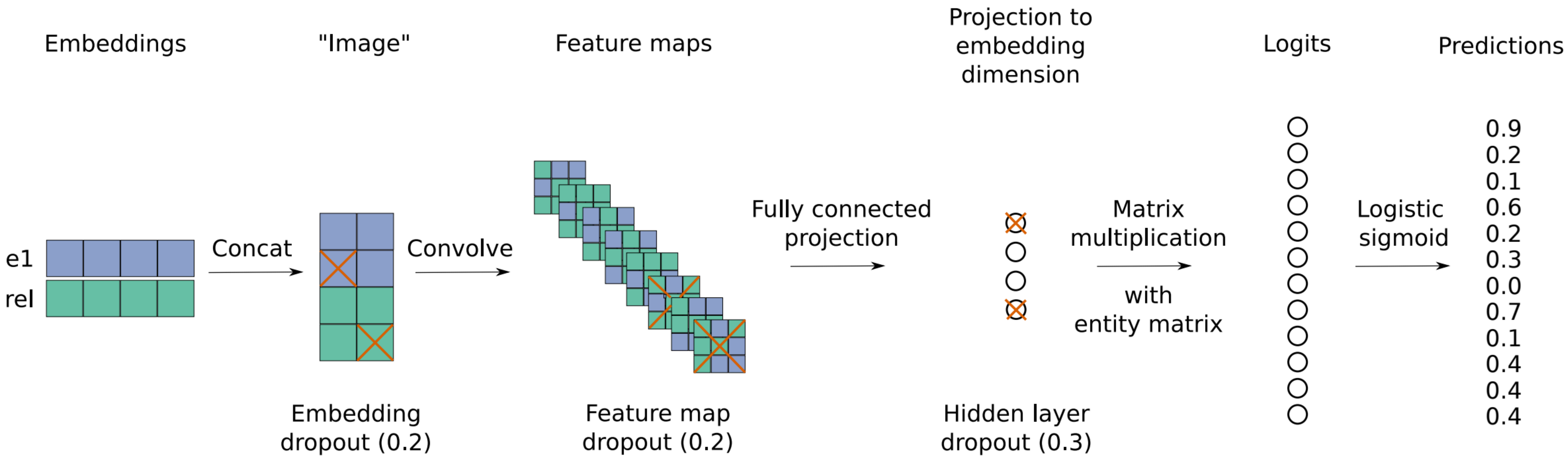


💡 Training:

- Score facts observed in the partial KG higher than those unobserved
- Sample negatives

Convolutional 2D Knowledge Graph Embeddings. Dettmers et al. 2018.

ConvE: Convolutional 2D Knowledge Graph Embeddings

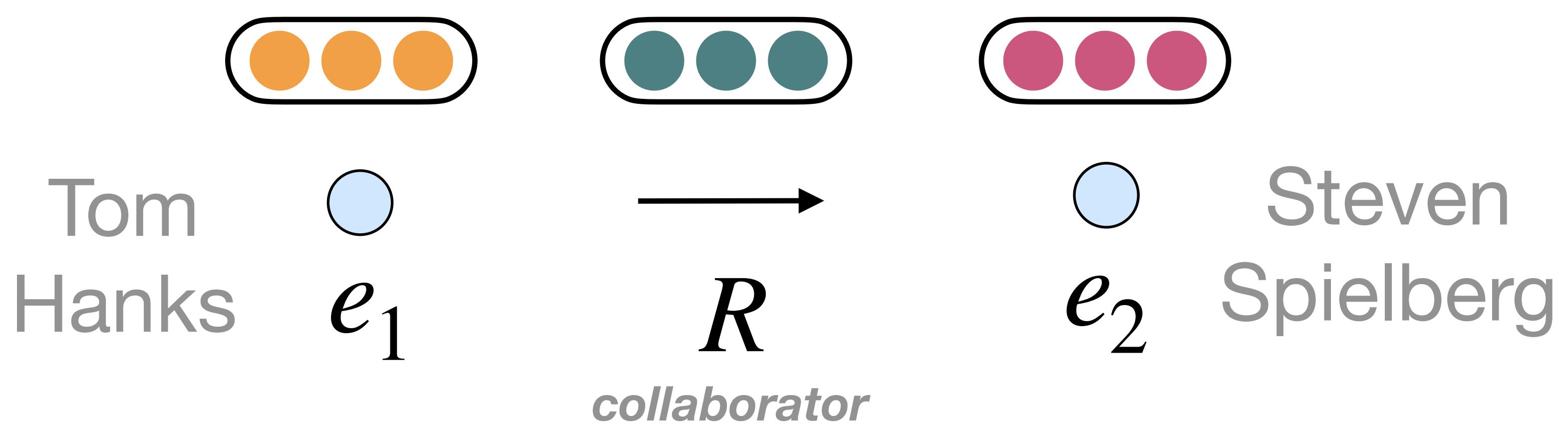


💡 SOTA model proposed in 2018

Better scoring functions were proposed since. Please refer to the literature for more details.

Convolutional 2D Knowledge Graph Embeddings. Dettmers et al. 2018.

Knowledge Graph Embeddings



Highly accurate &
Efficient

	Acc
WordNet	86.2
Freebase	90.0

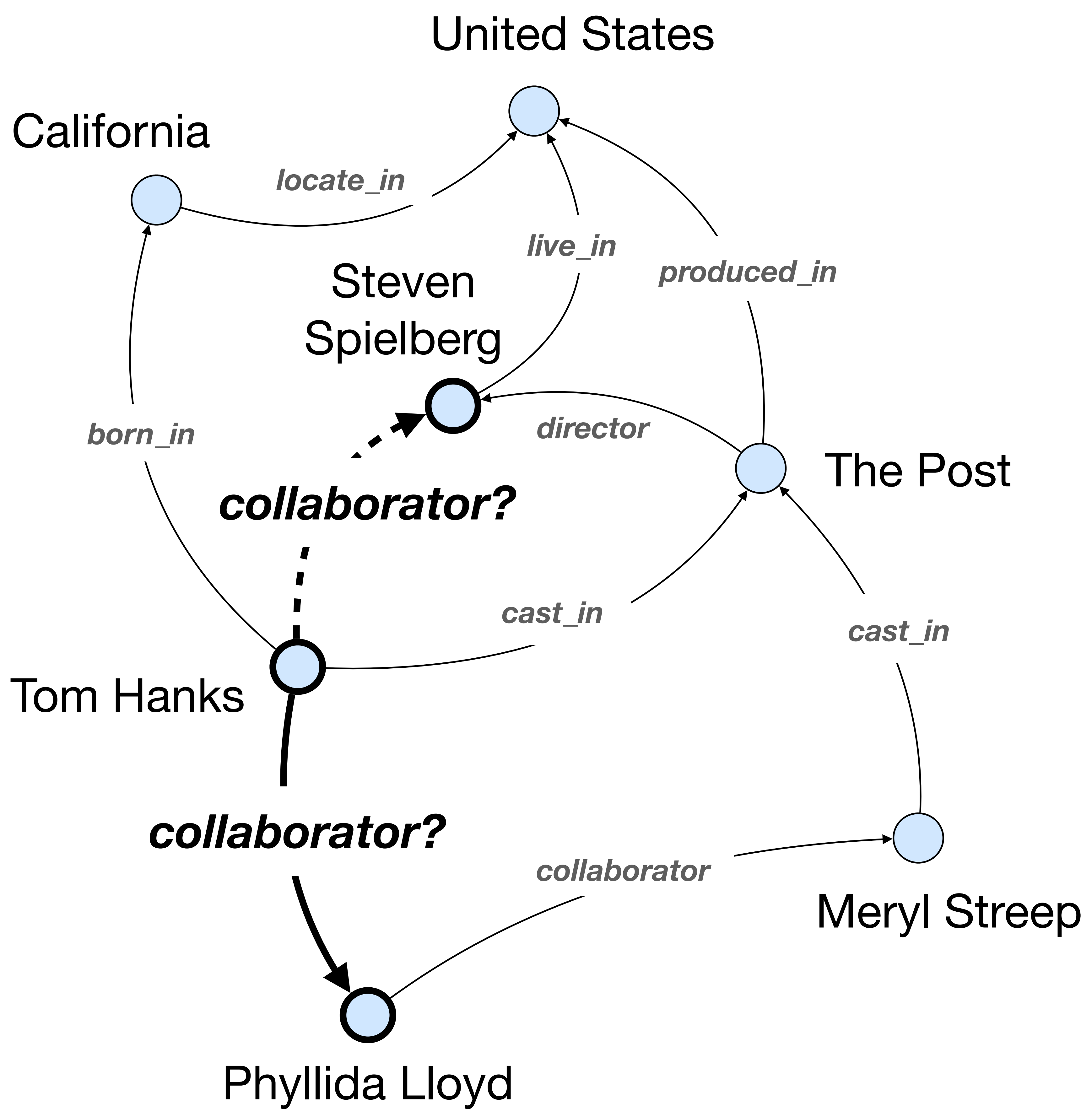
Tab 1. NTN KB fact inference performance on the WordNet and Freebase benchmarks (Socher et. al. 2013)

- Lack interpretability
- Does not perform well for rare/unseen entities

Convolutional 2D Knowledge Graph Embeddings. Dettmers et al. 2018.

Knowledge Graph Reasoning

- Knowledge Graph Embeddings
- Path Ranking Algorithm (PRA) RL
- Reinforcement Learning



Path Ranking Algorithm

- Identifying relation paths connecting two entities (e_1, e_2) in the KG and use the paths as features for predicting new relations.

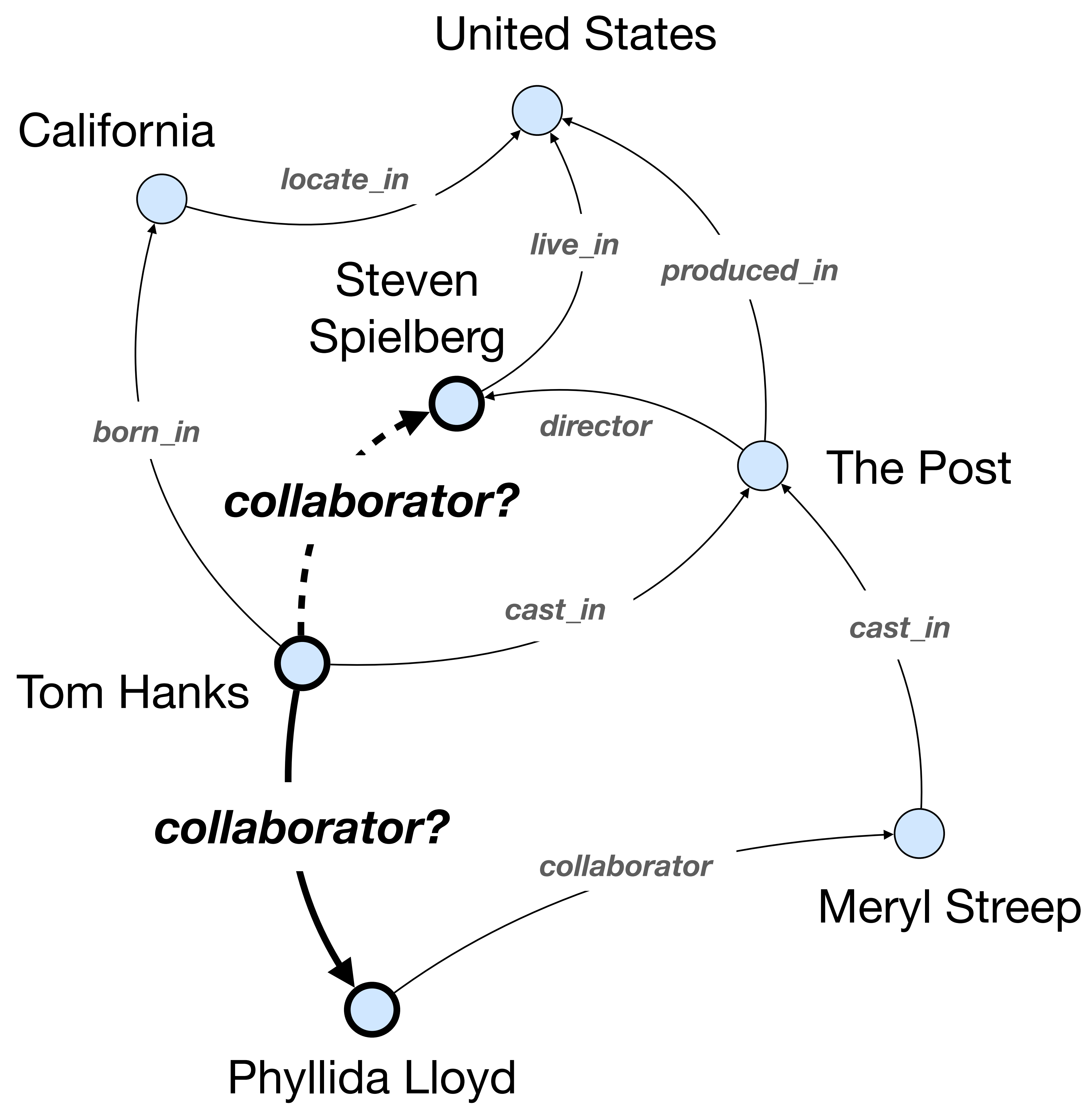
$$P = R_1 \ldots R_\ell$$

Tom Hanks *CastIn* $\exists e$ *Director* S.S.

Tom Hanks *CastIn* $\exists e$ *ProducedIn* $\exists e$ *LiveIn'* S.S.

Tom Hanks *BornIn* $\exists e$ *LocatedIn* $\exists e$ *LiveIn'* S.S.

...



Relational Retrieval Using a Combination of Path-Constrained Random Walks. Lao and Cohen 2010.

Random Walk Inference and Learning in A Large-Scale Knowledge Base. Lao and Cohen 2011.

Finding Inference Paths

- Exhaustive (Lao and Cohen 2010)
  - Obtaining all paths connecting e_1 and e_2 (dynamic programming)

Relational Retrieval Using a Combination of Path-Constrained Random Walks. Lao and Cohen 2010.

Random Walk Inference and Learning in A Large-Scale Knowledge Base. Lao and Cohen 2011.

Finding Inference Paths

- Exhaustive (Lao and Cohen 2010)
  - Obtaining all paths connecting e_1 and e_2 (dynamic programming)
- Data-driven (Lao and Cohen 2011)
  - Identifying only paths that are potentially useful for an inference task
  - Any node e visited during path search must be supported by at least a fraction α of seed entities s_i seen during training
  - Any path P must retrieve at least one target entity t_i on the training set

s_i supports e iff. the random walk probability between s_i and e is greater than 0

Relational Retrieval Using a Combination of Path-Constrained Random Walks. Lao and Cohen 2010.

Random Walk Inference and Learning in A Large-Scale Knowledge Base. Lao and Cohen 2011.

Finding Inference Paths

- Reinforcement Learning (Xiong et al. 2018)
  - Learn a policy based agent to sample the most informative paths between e_1 and e_2
  - Starting from e_1, the agent uses a policy network to pick the most promising relation to extend its path at each step until it reaches the target entity e_2, or has reached a maximum number of search steps.
  - Hybrid reward

$$r_{\text{GLOBAL}} = \begin{cases} +1, & \text{if the path reaches } e_{target} \\ -1, & \text{otherwise} \end{cases} \quad r_{\text{EFFICIENCY}} = \frac{1}{length(p)} \quad r_{\text{DIVERSITY}} = -\frac{1}{|F|}\sum_{i=1}^{|F|} cos(\mathbf{p}, \mathbf{p}_i)$$

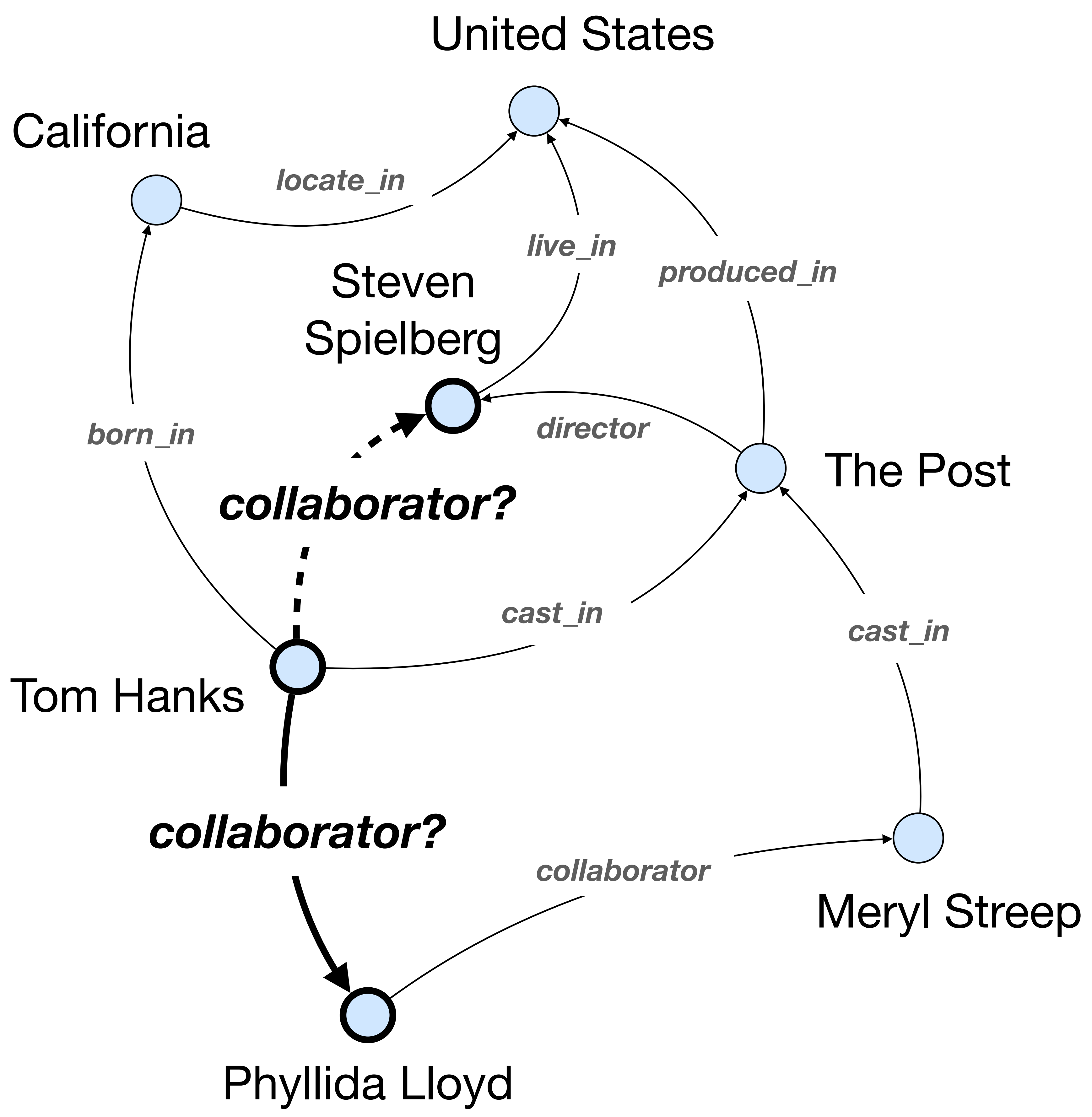
  - Supervised policy learning and retraining with reward

DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning. Xiong et al. 2018.

Path Ranking Algorithm

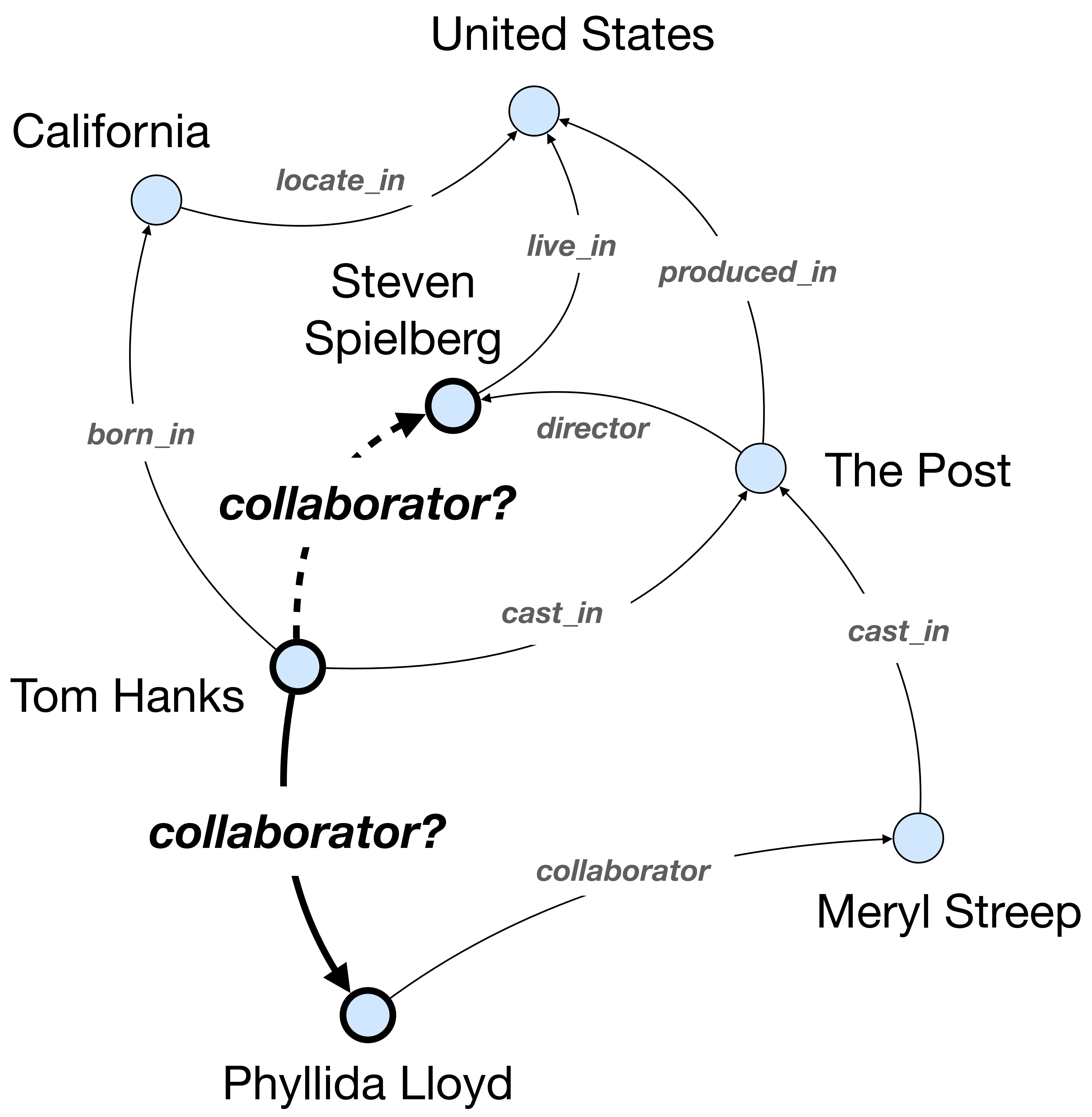
- Explainable
- Performs well for $e_s \xrightarrow{r_q?} e_t$ queries
- Can work for rare/unseen entities as reasoning is based on path features

- Inefficient for $e_s \xrightarrow{r_q} ?$ queries

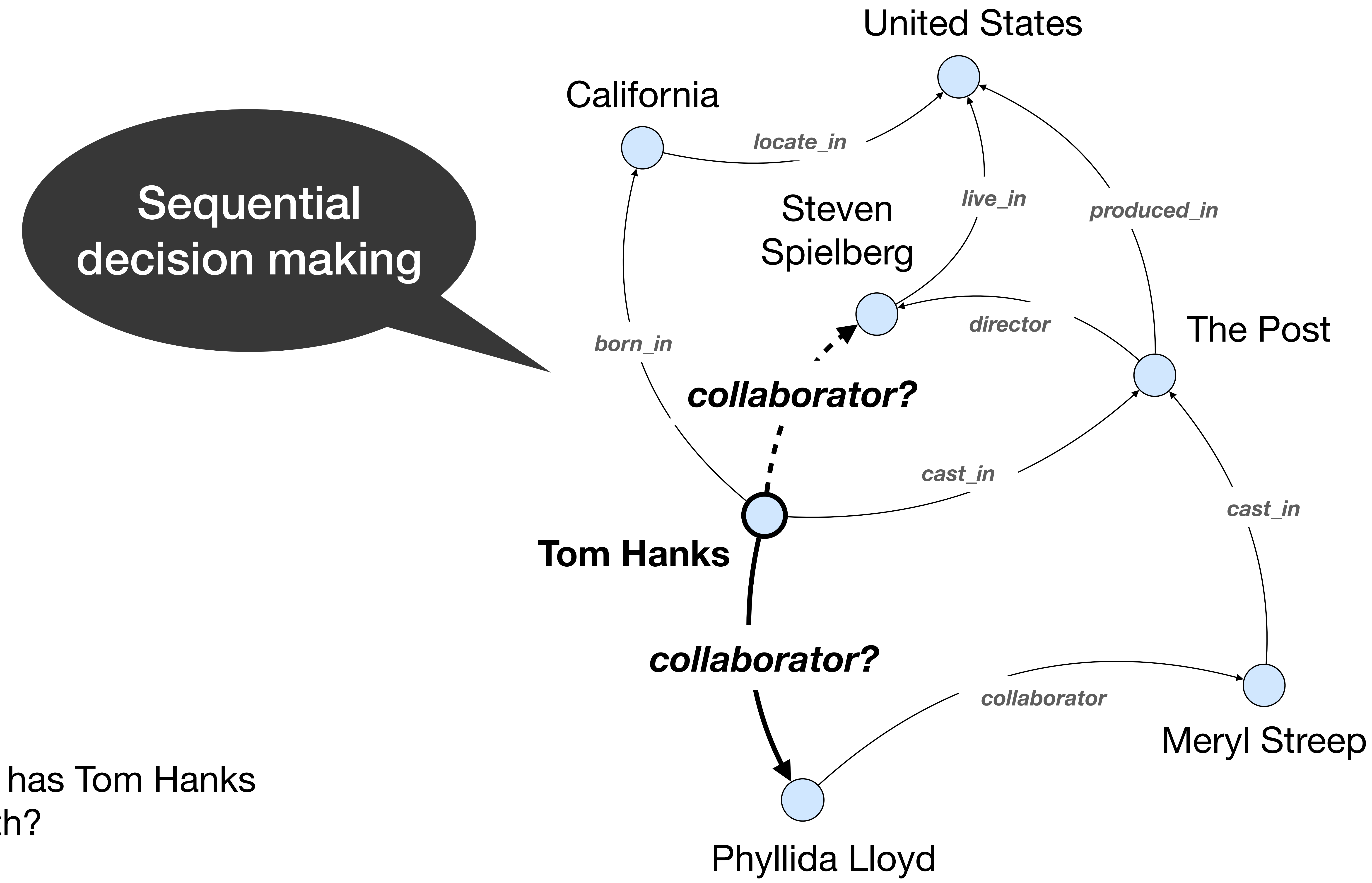


Knowledge Graph Reasoning

- Knowledge Graph Embeddings
- Path Ranking Algorithm (PRA)
- Sequential Decision Making RL



Sequential Multi-Hop Reasoning

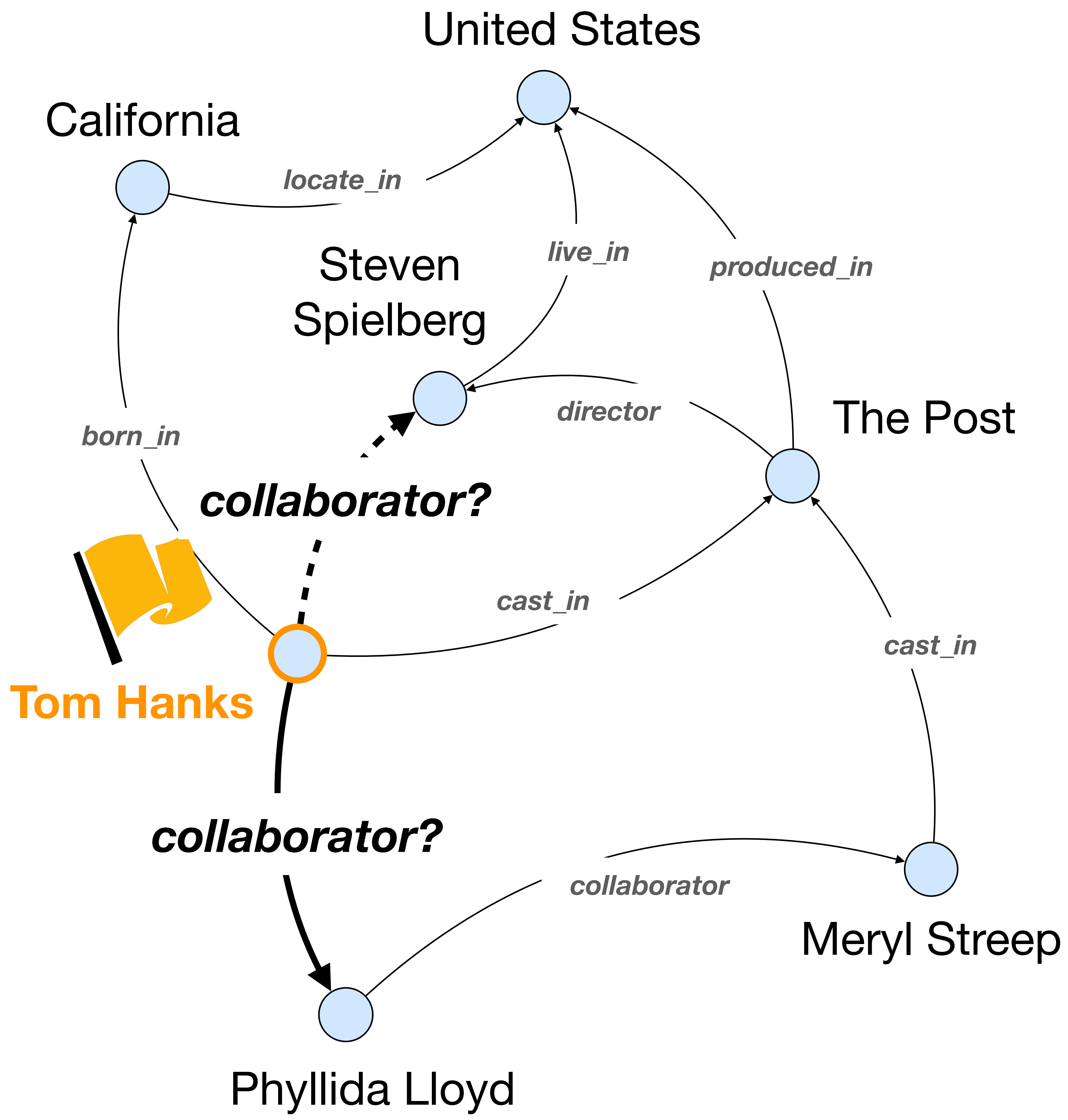


Which directors has Tom Hanks collaborated with?

Sequential Multi-Hop Reasoning

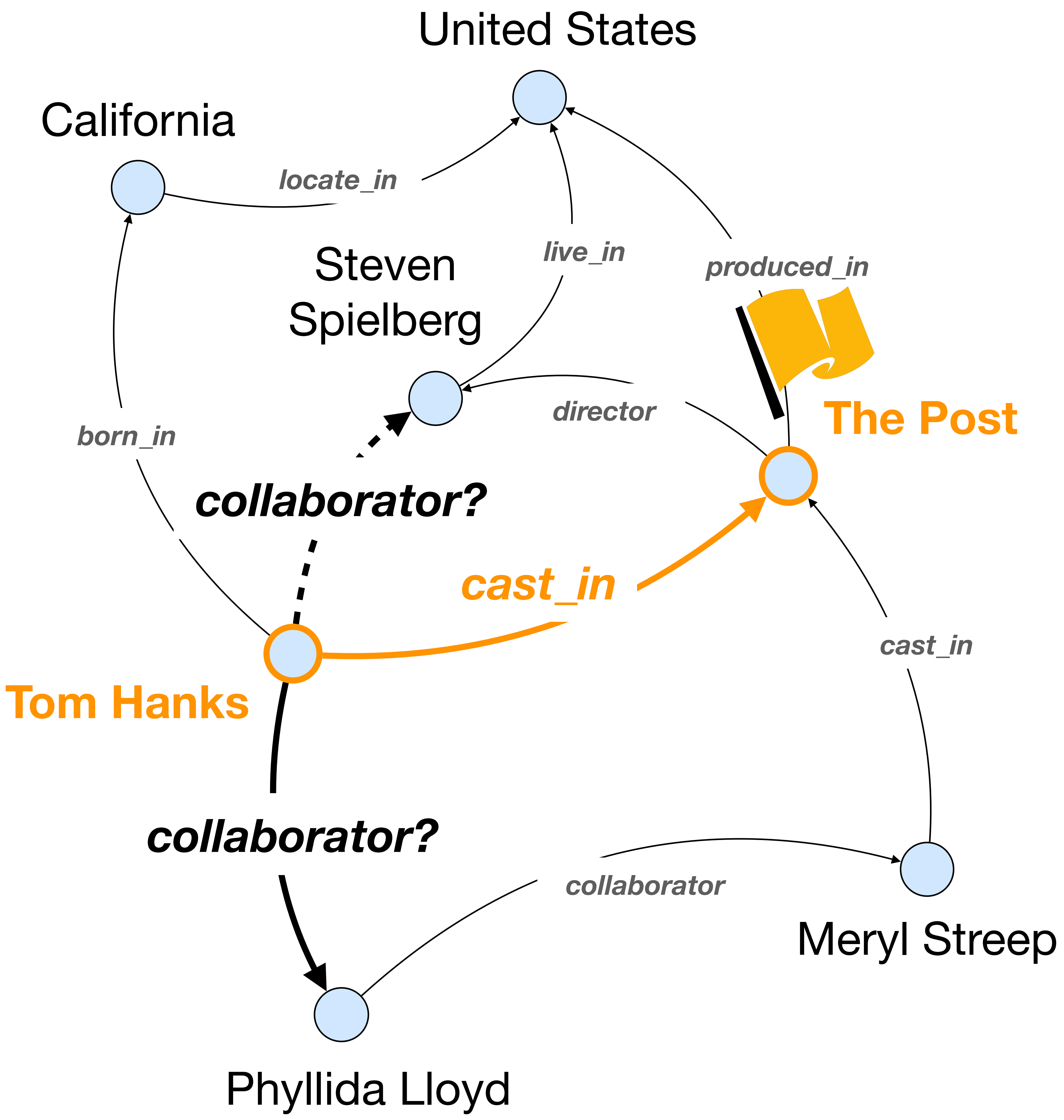
Topic entity

Tom Hanks



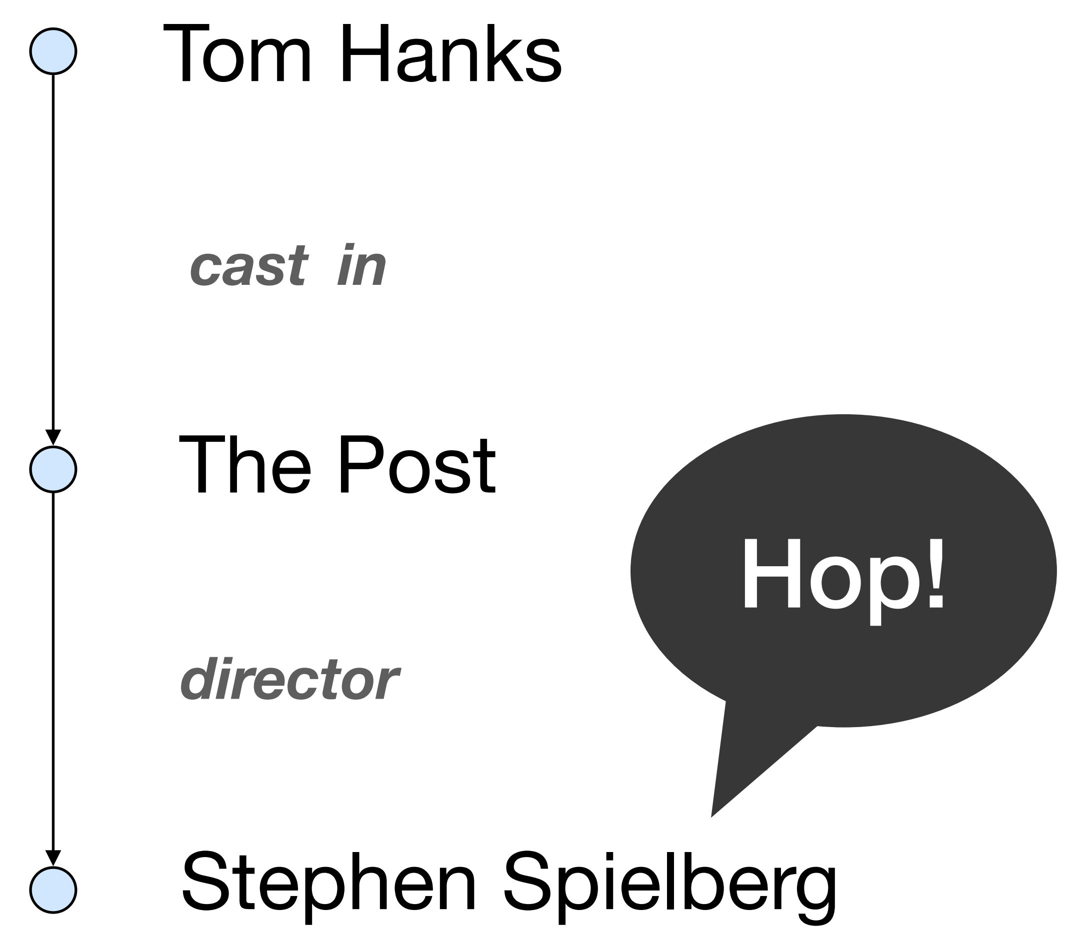
Which directors has Tom Hanks collaborated with?

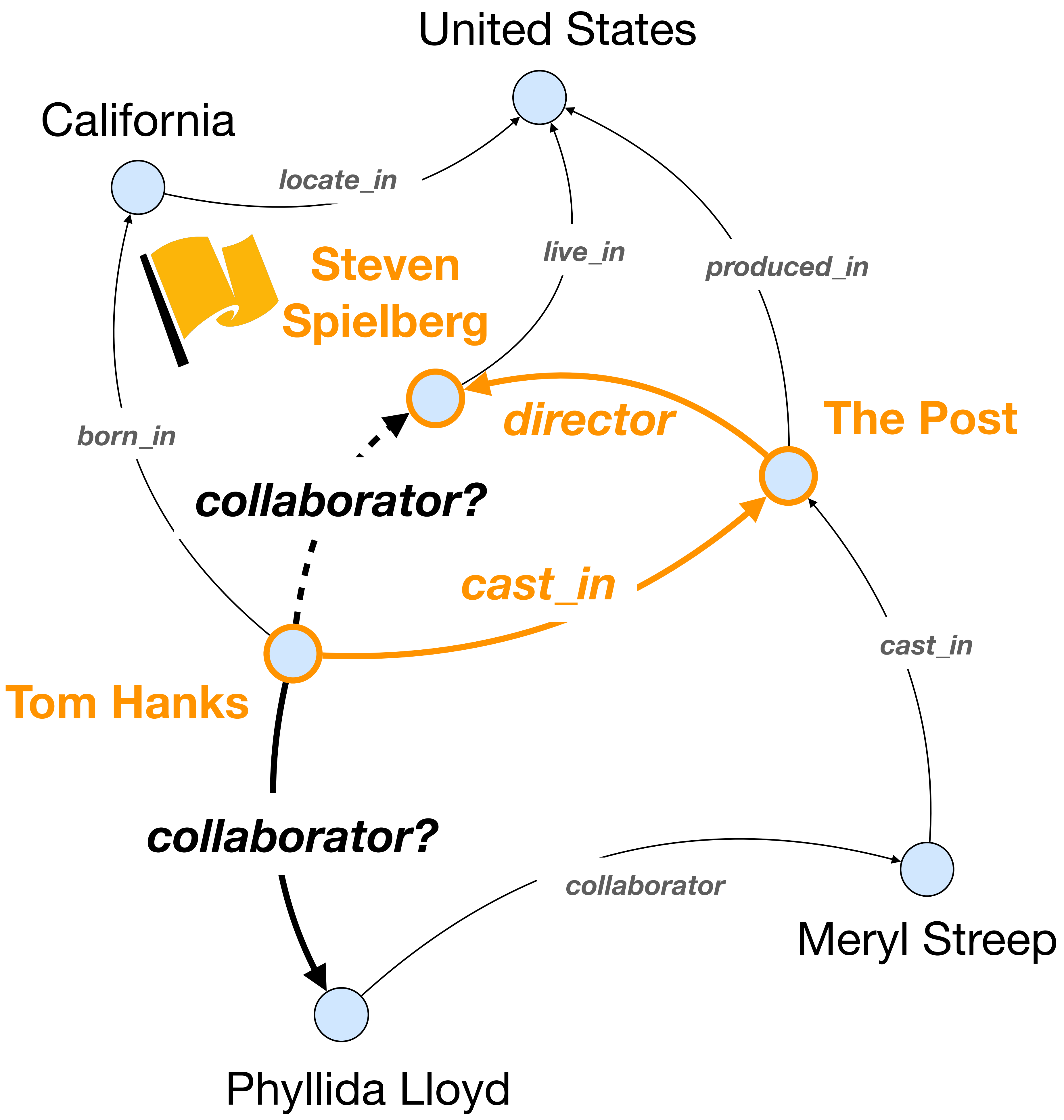
Sequential Multi-Hop Reasoning



Which directors has Tom Hanks collaborated with?

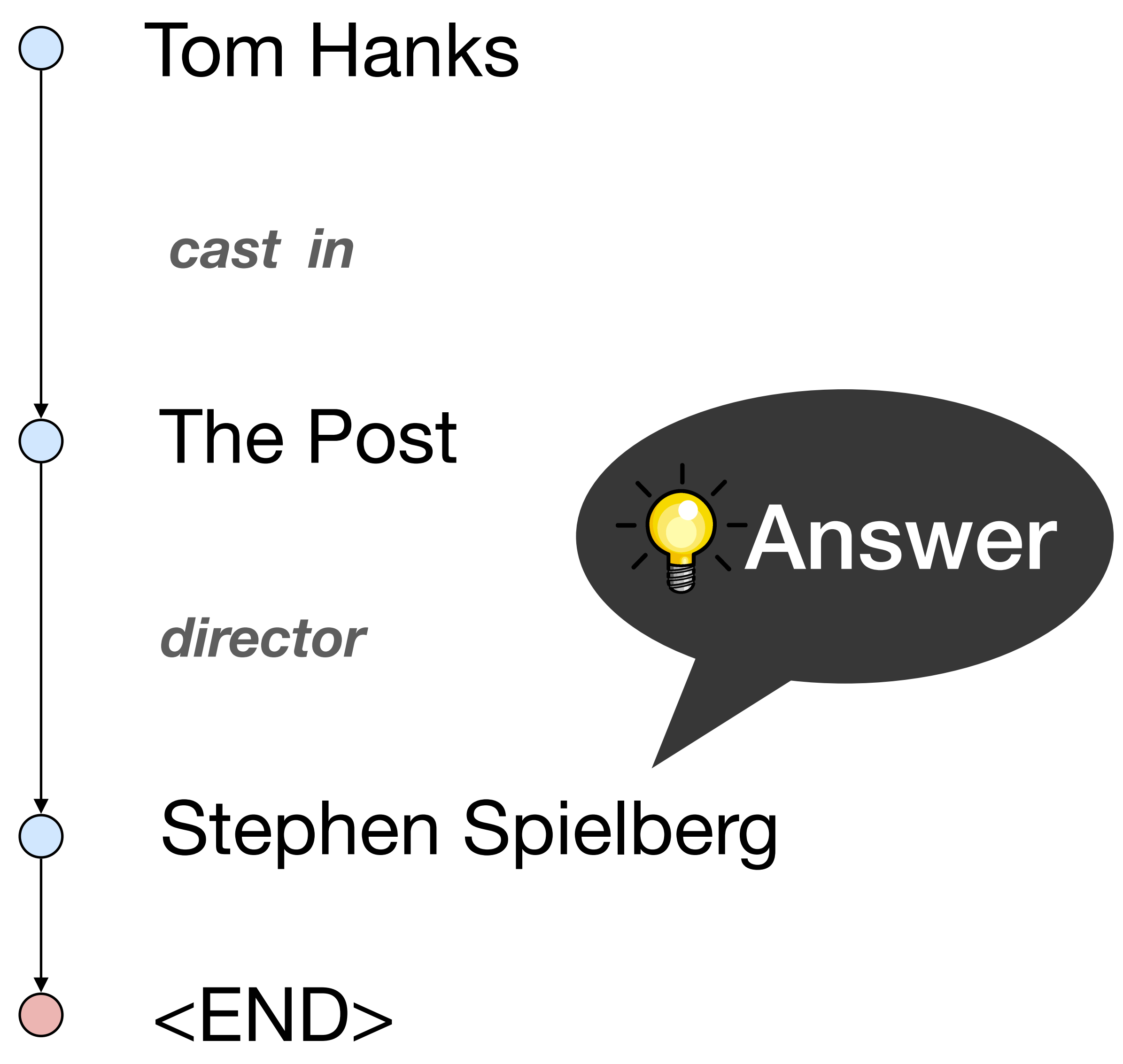
Sequential Multi-Hop Reasoning

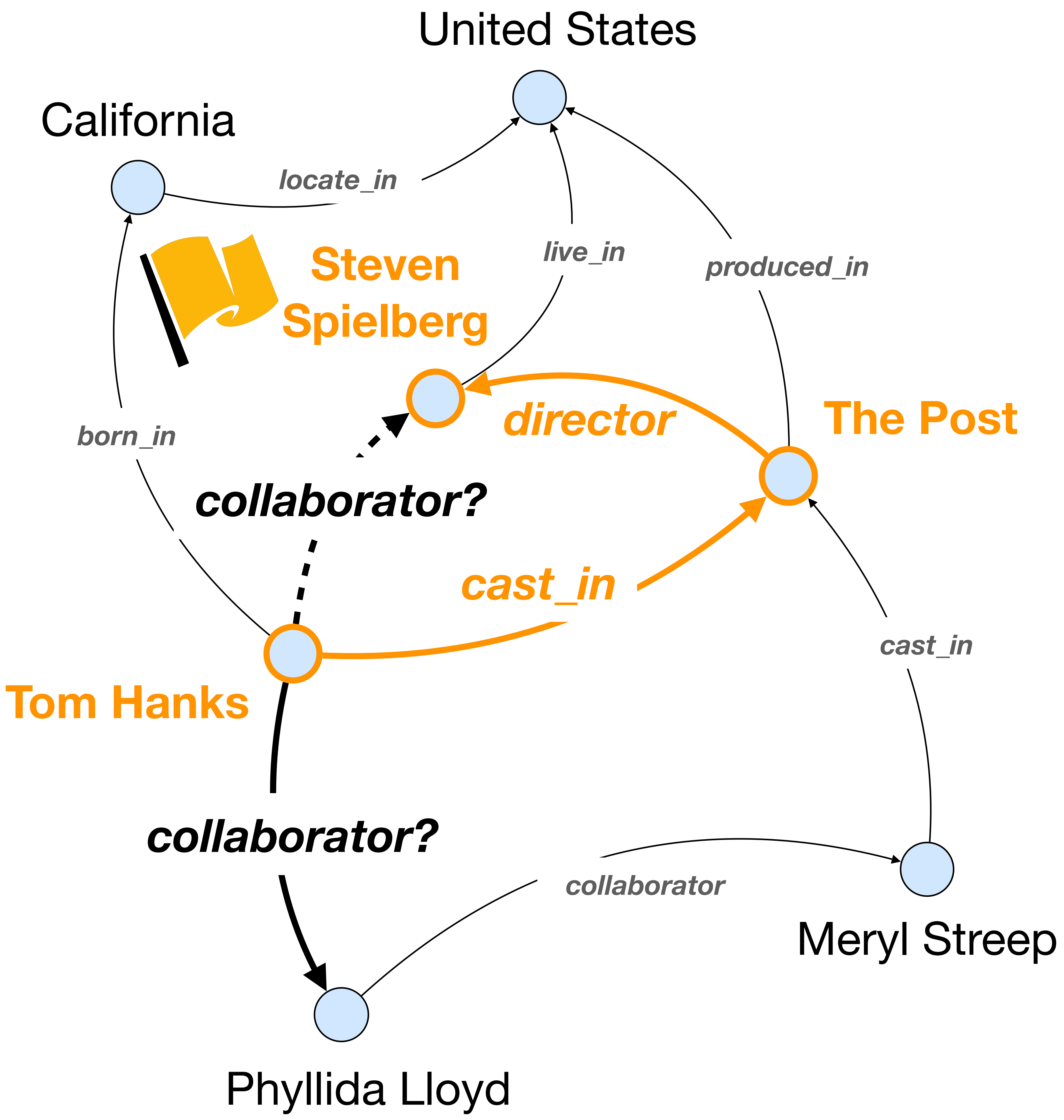




Which directors has Tom Hanks collaborated with?

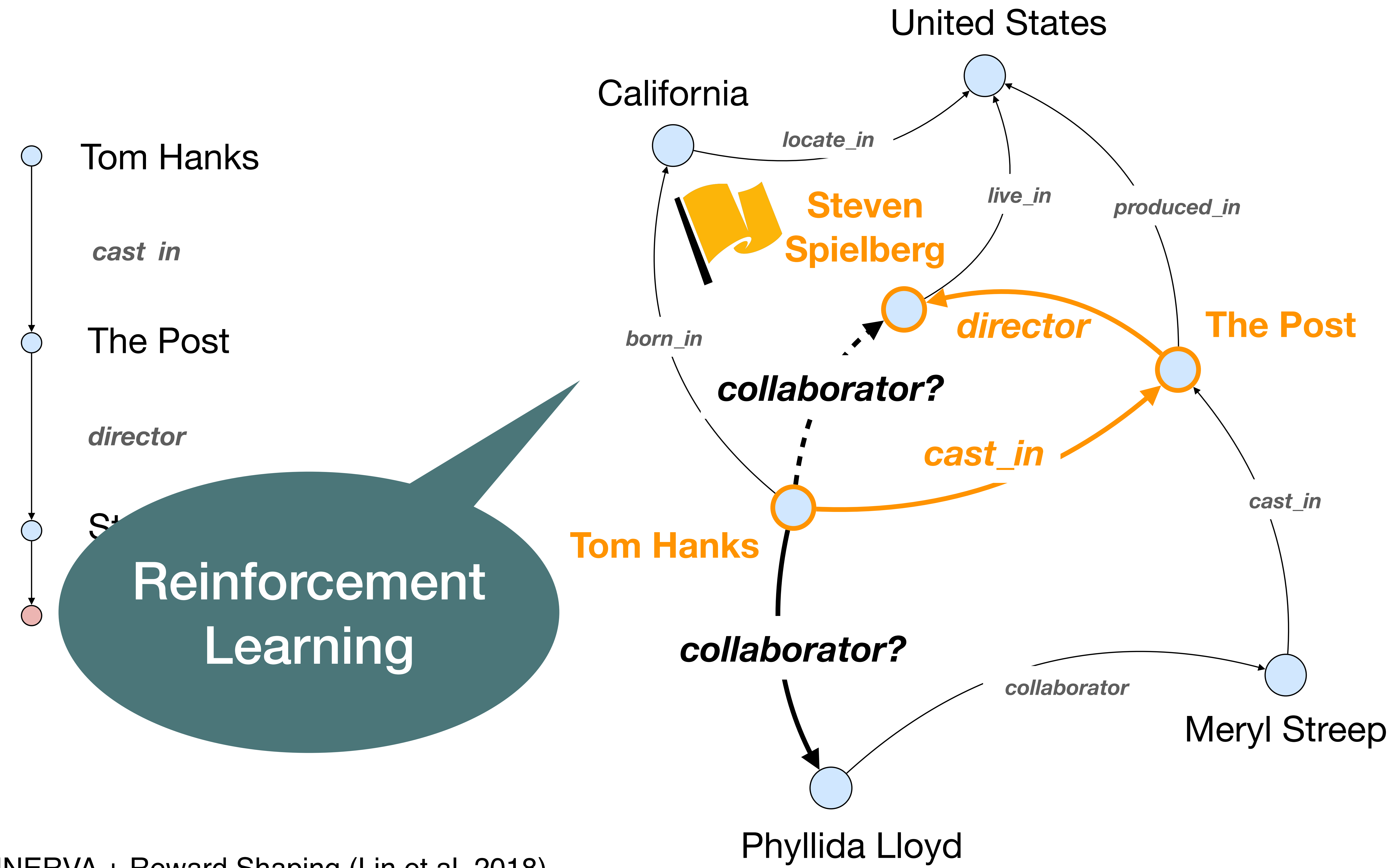
Sequential Multi-Hop Reasoning





Which directors has Tom Hanks collaborated with?

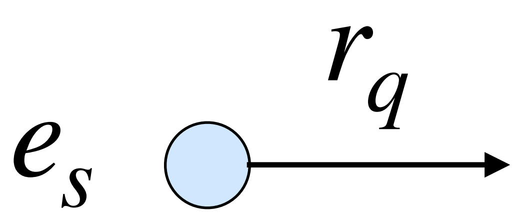
Sequential Multi-Hop Reasoning



MINERVA (Das et al. 2018); MINERVA + Reward Shaping (Lin et al. 2018)

Reinforcement Learning Framework

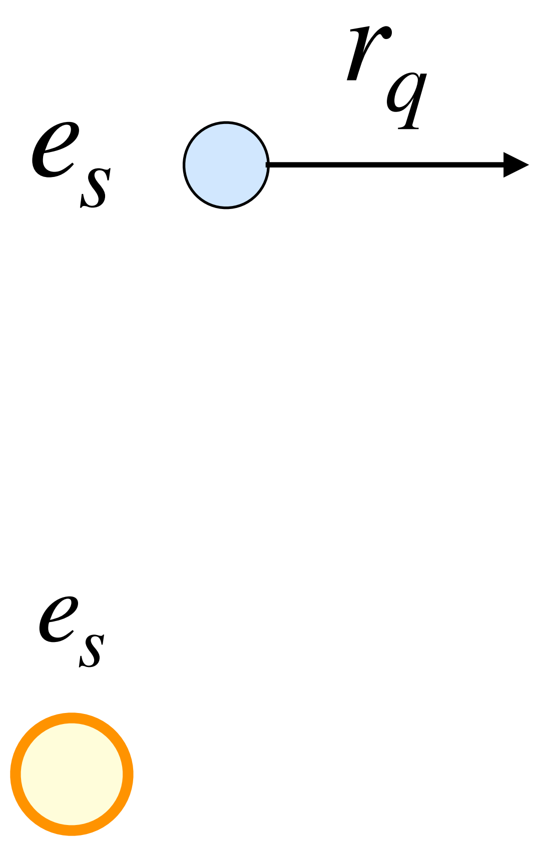
Environment State Action Transition Reward



Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning
(Das et al. 2018)

Reinforcement Learning Framework

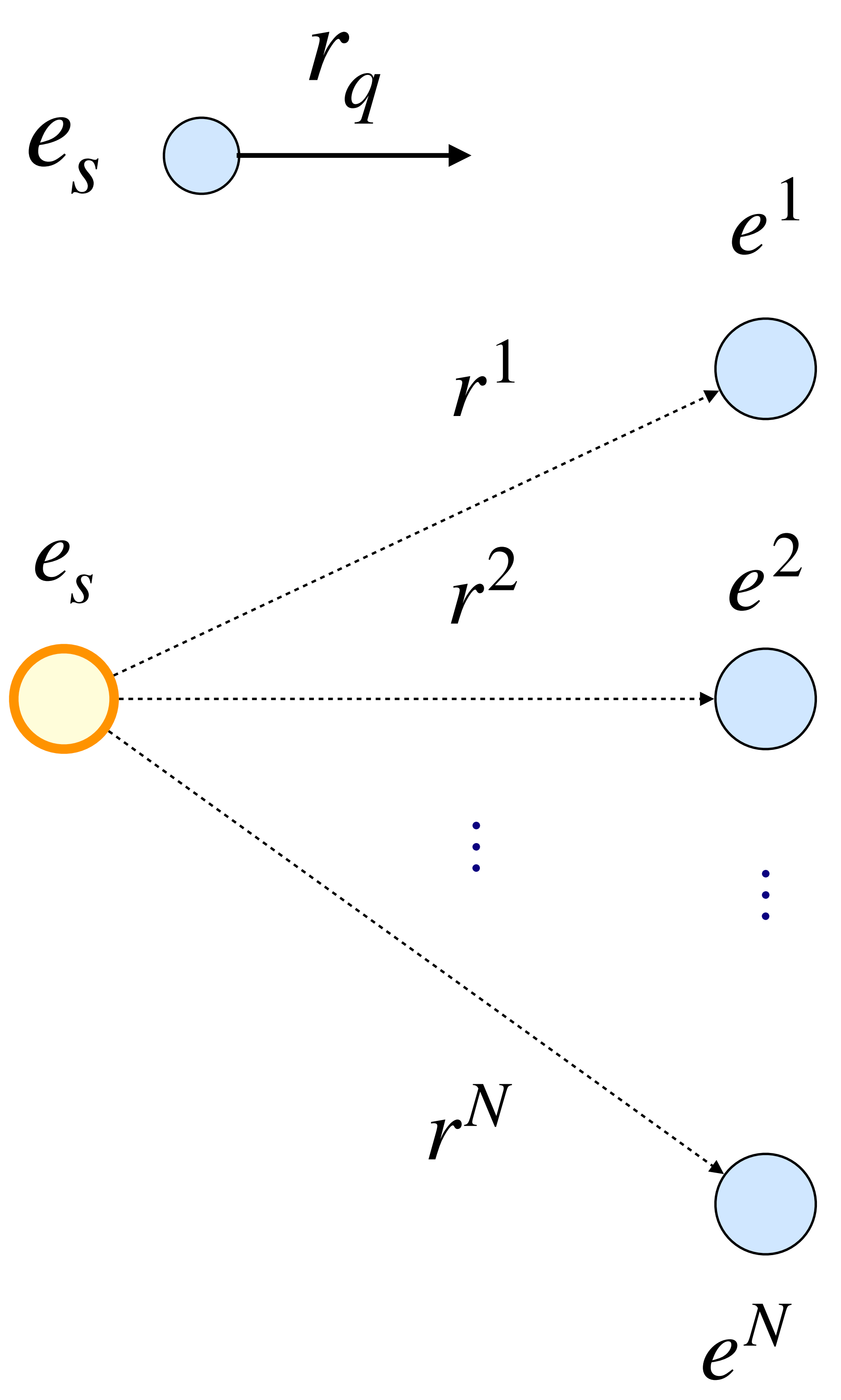
Environment | **State** | Action | Transition | Reward



Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning
(Das et al. 2018)

Reinforcement Learning Framework

Environment | State | **Action** | Transition | Reward

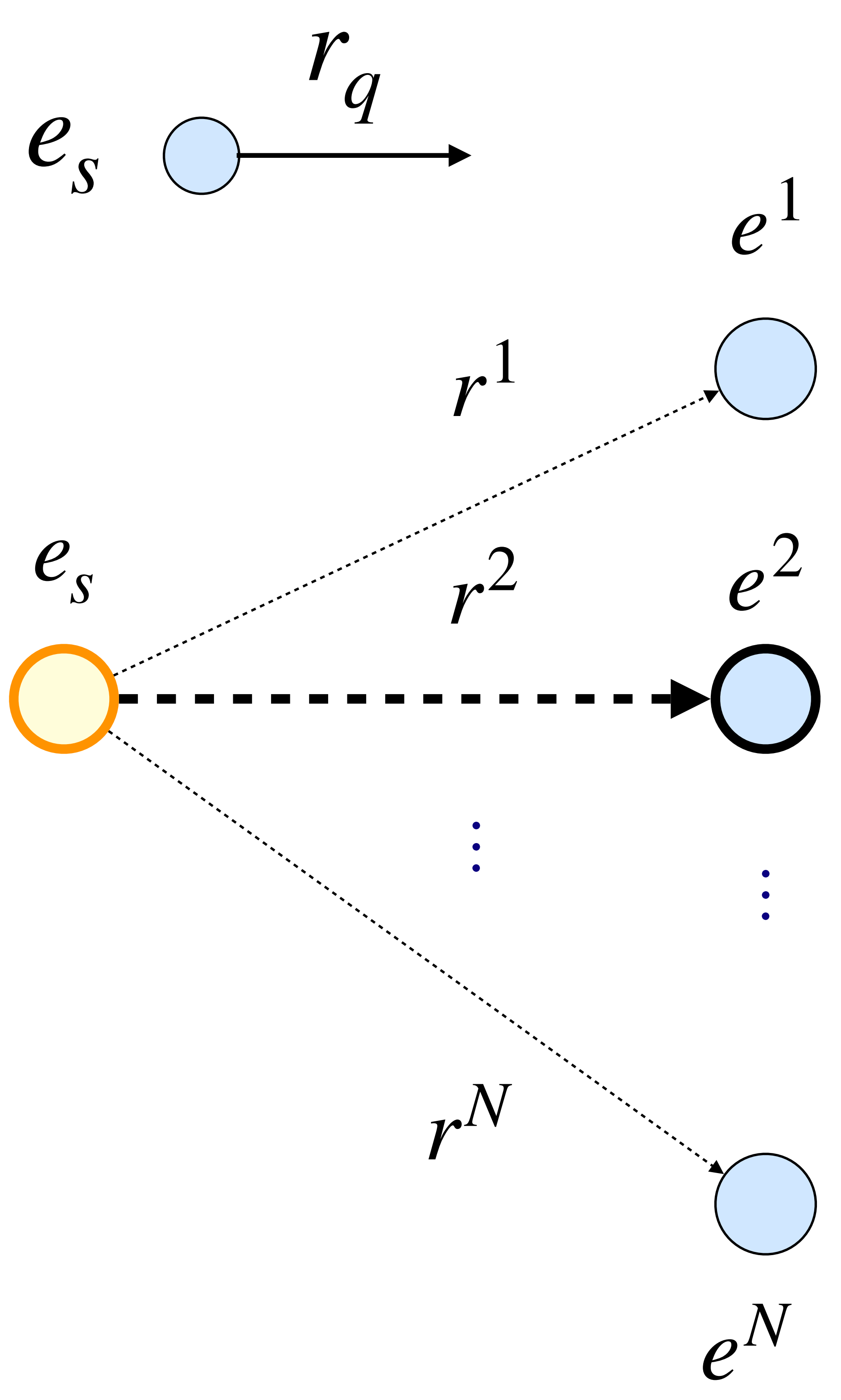


A_t

Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning
(Das et al. 2018)

Reinforcement Learning Framework

Environment | State | **Action** | Transition | Reward



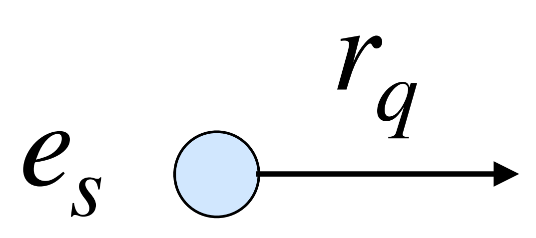
A_t

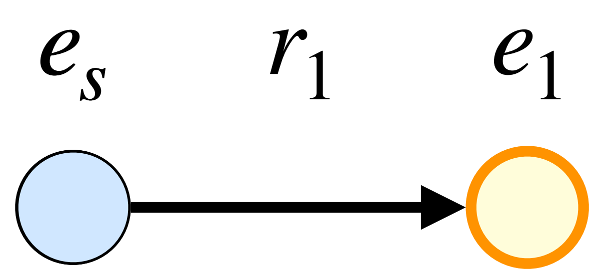
Action space: All nodes with incoming edges from e_s in the graph

Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

Reinforcement Learning Framework

Environment | State | Action | **Transition** | Reward

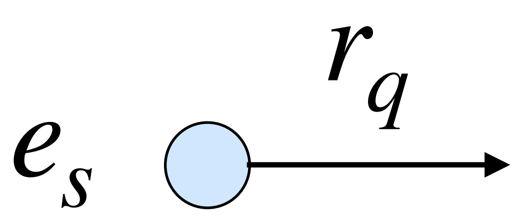


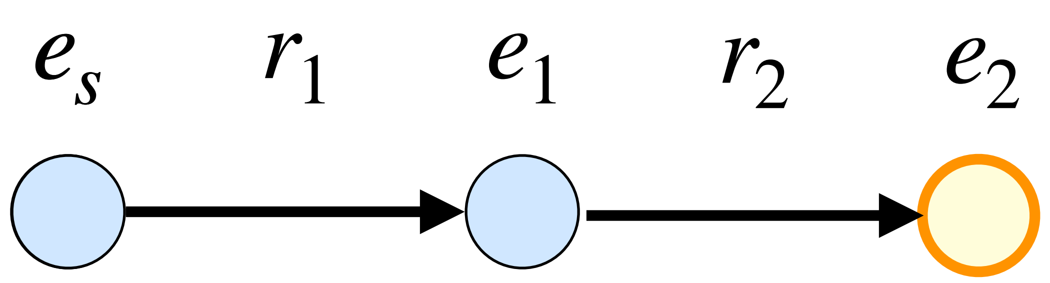


Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

Reinforcement Learning Framework

Environment | State | Action | **Transition** | Reward

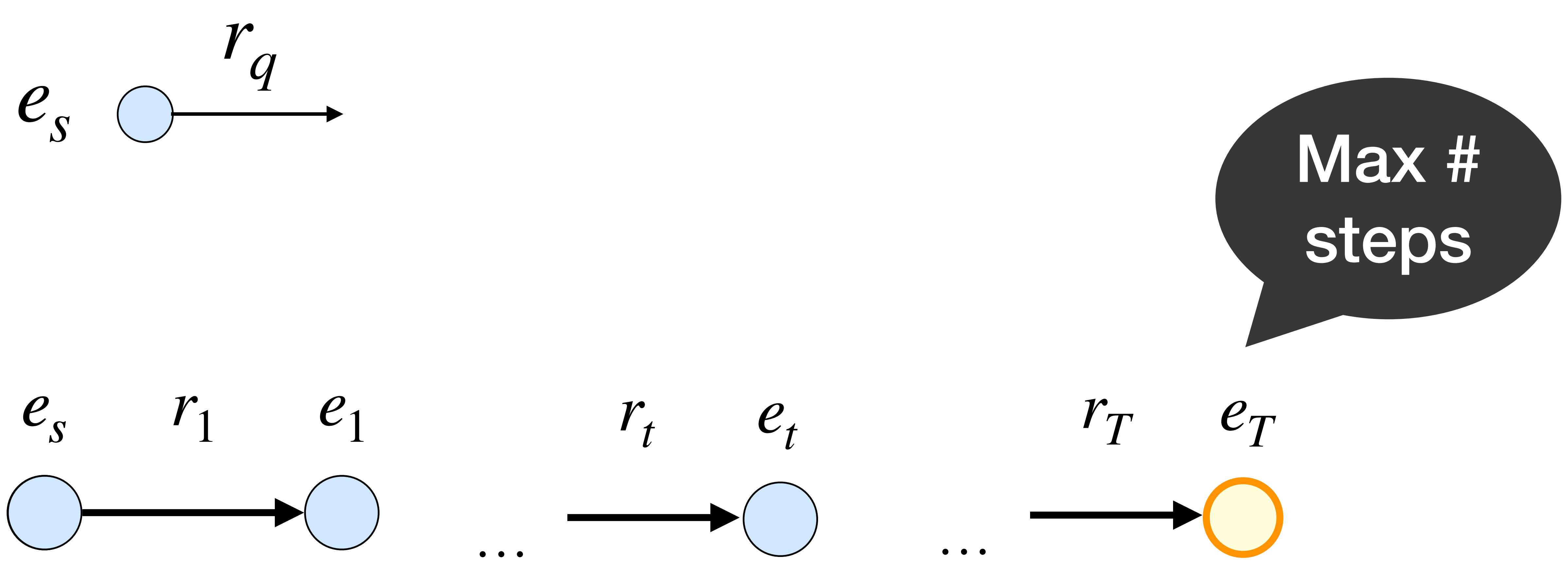




Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

Reinforcement Learning Framework

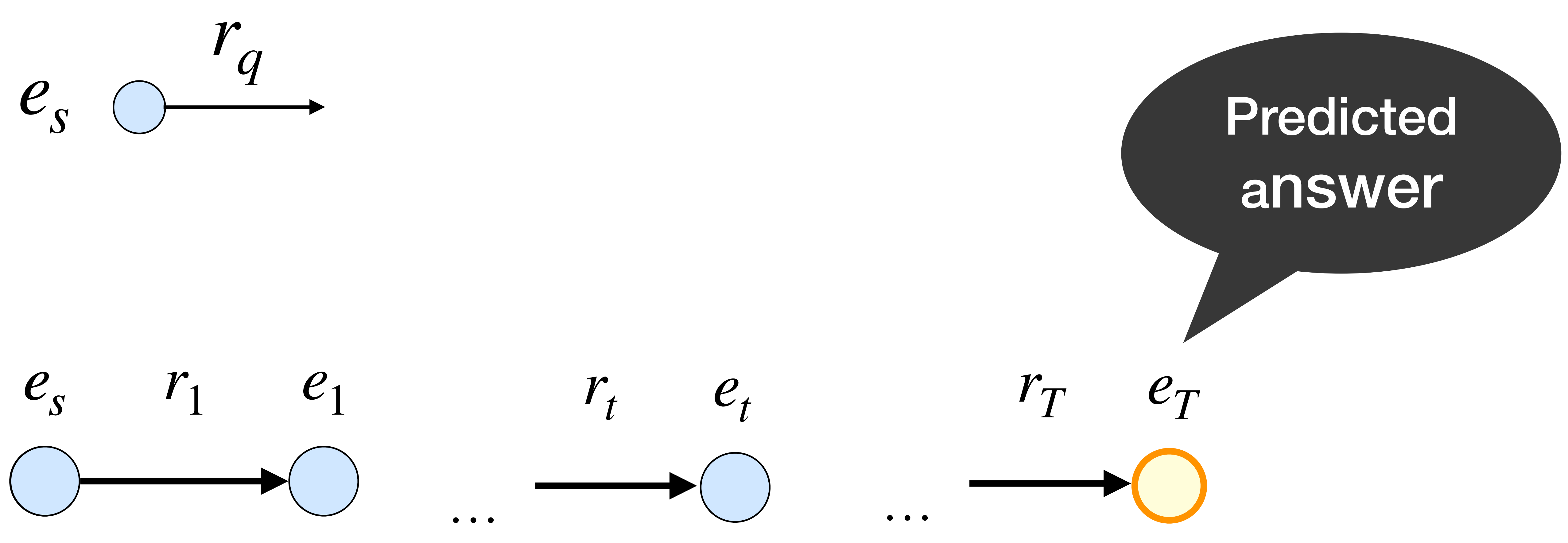
Environment | State | Action | **Transition** | Reward



Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

Reinforcement Learning Framework

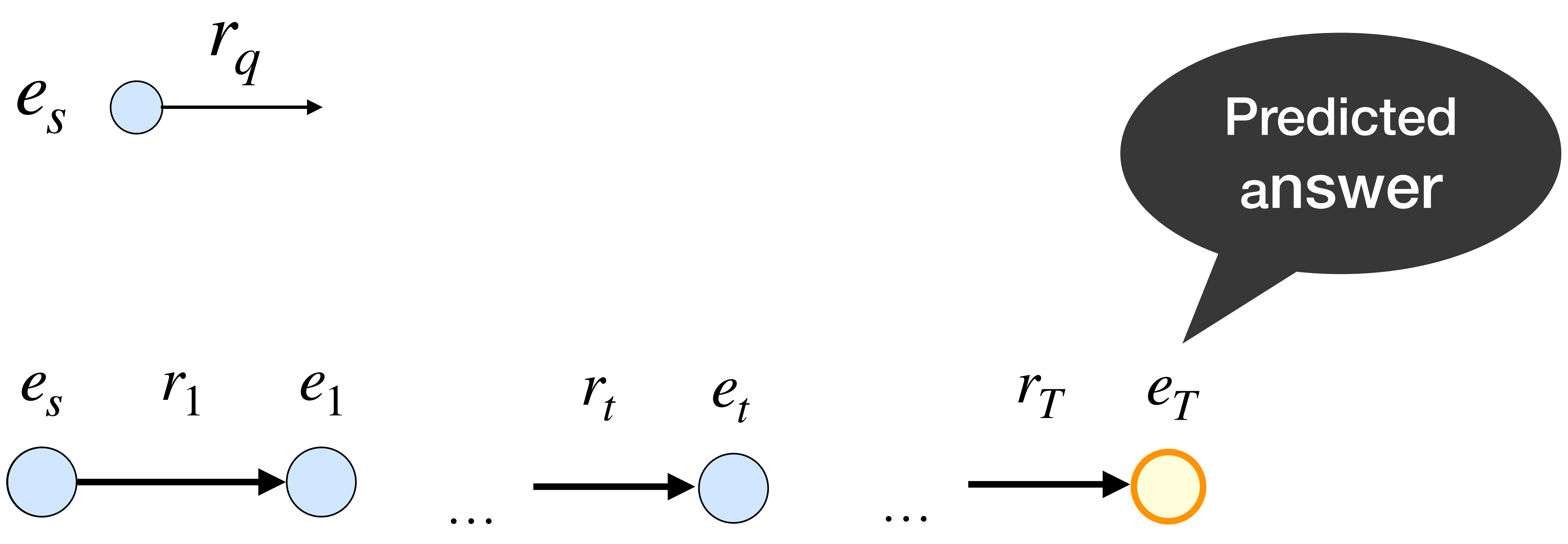
Environment · State · Action · **Transition** · Reward



Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

Reinforcement Learning Framework

Environment | State | Action | Transition | **Reward**

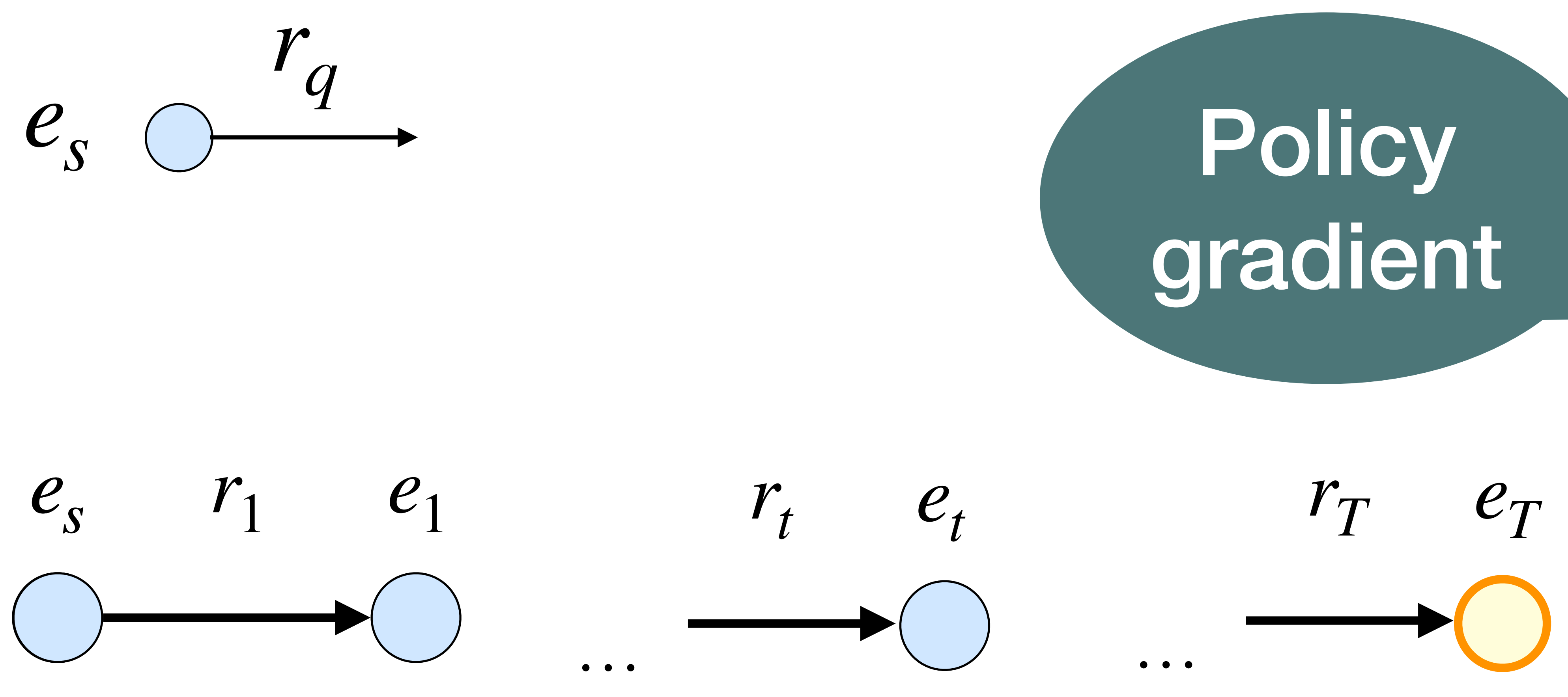


$$R_b(s_T) = \mathbf{1}\{(e_s, r_q, e_T) \in G\}$$

Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

Reinforcement Learning Framework

Environment | State | Action | Transition | **Reward**



Policy gradient

Learn ***which action to choose*** given a state

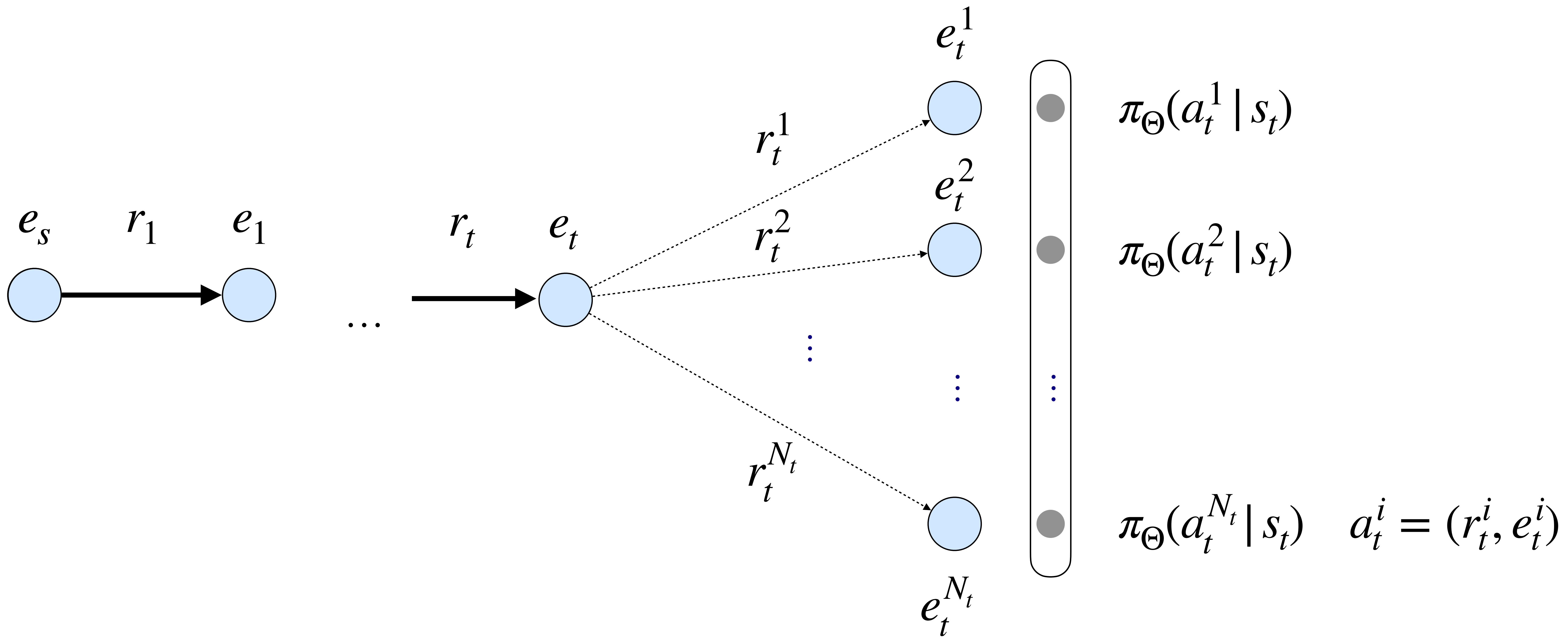
$$R_b(s_T) = \mathbf{1}\{(e_s, r_q, e_T) \in G\}$$

Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

Policy Gradient

Policy function $\pi_\Theta(a_t \mid s_t)$

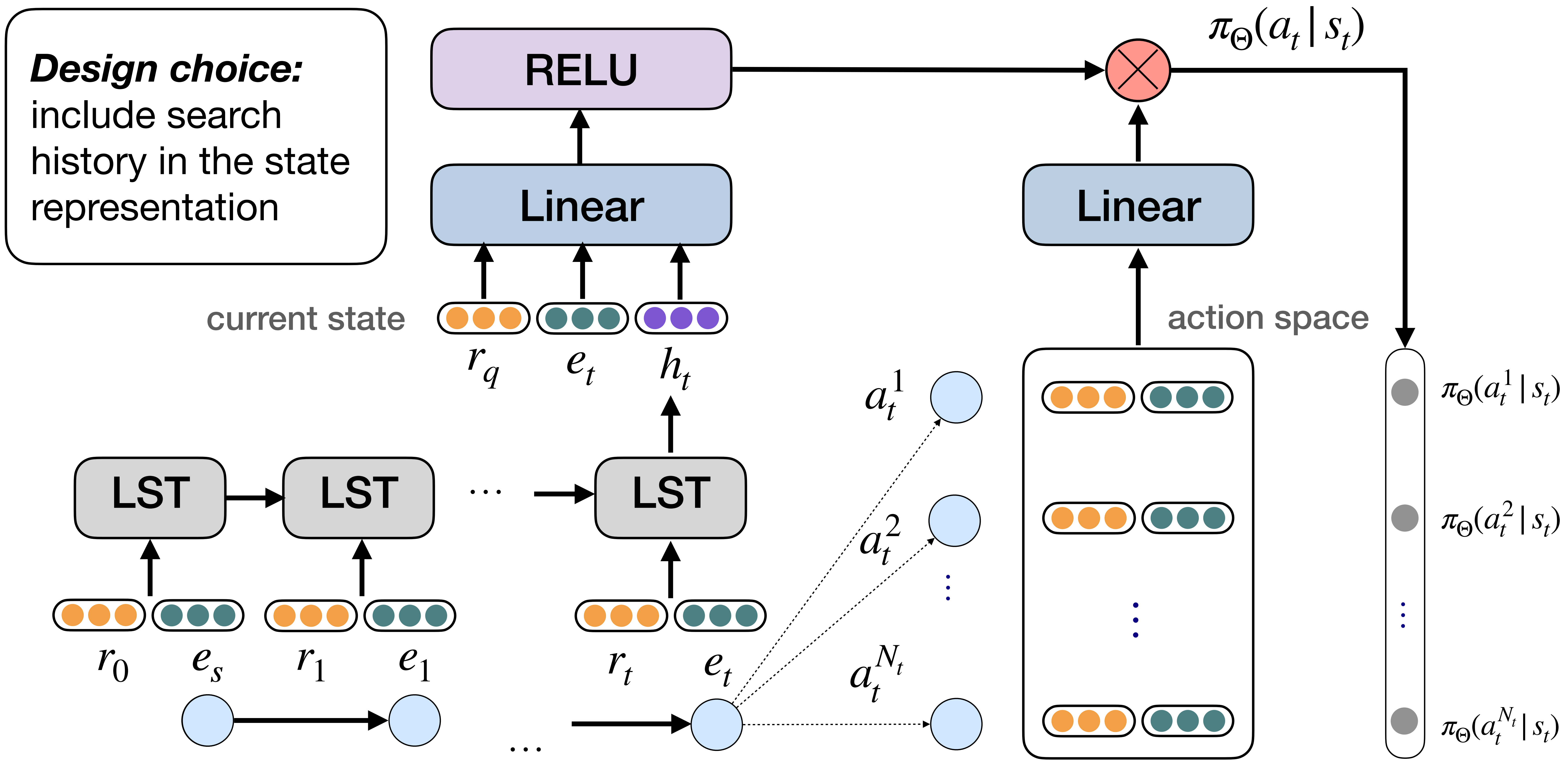
Probability of choosing an action given the current state



Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

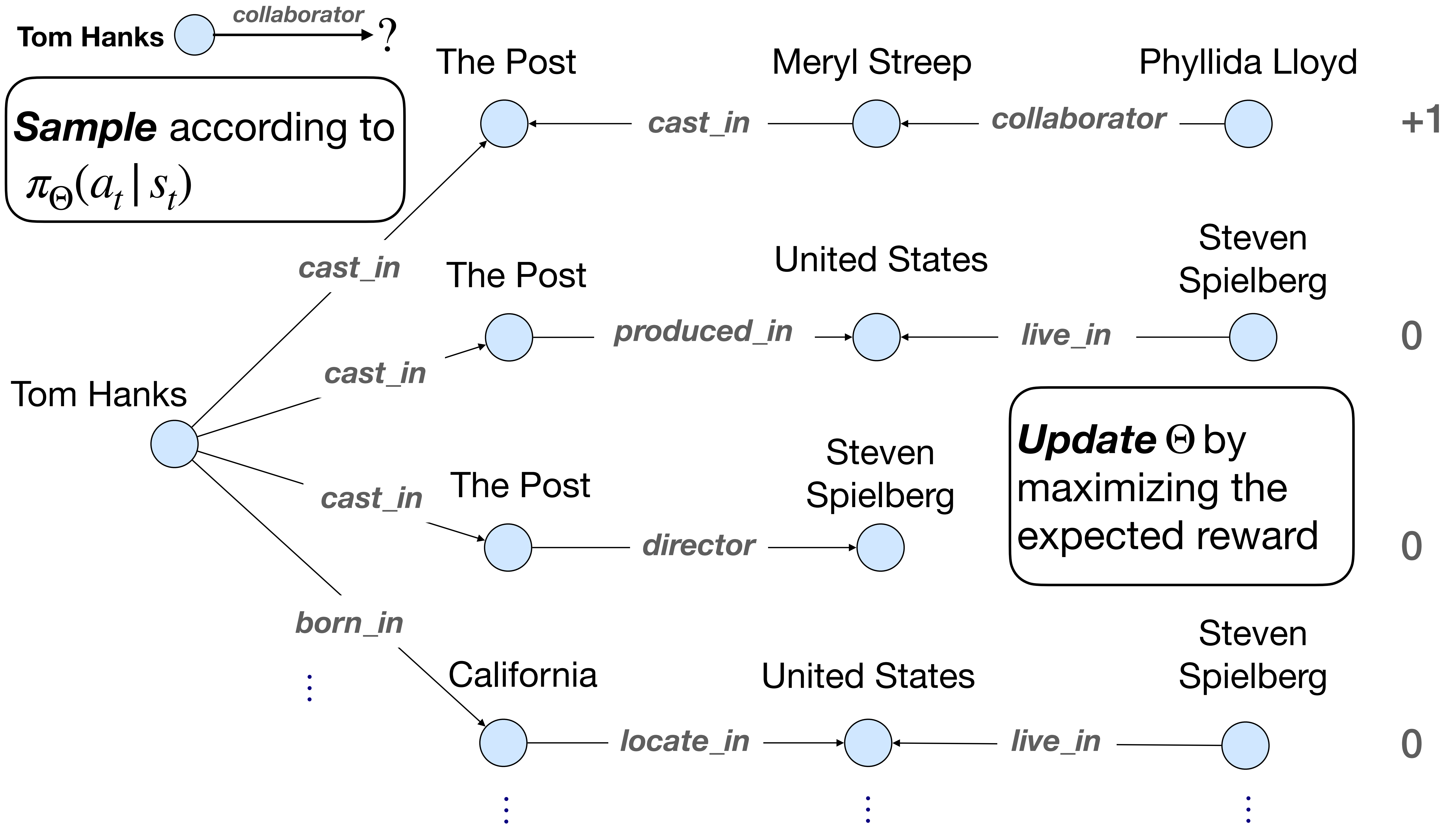
Policy Gradient

Policy function $\pi_\Theta(a_t | s_t)$



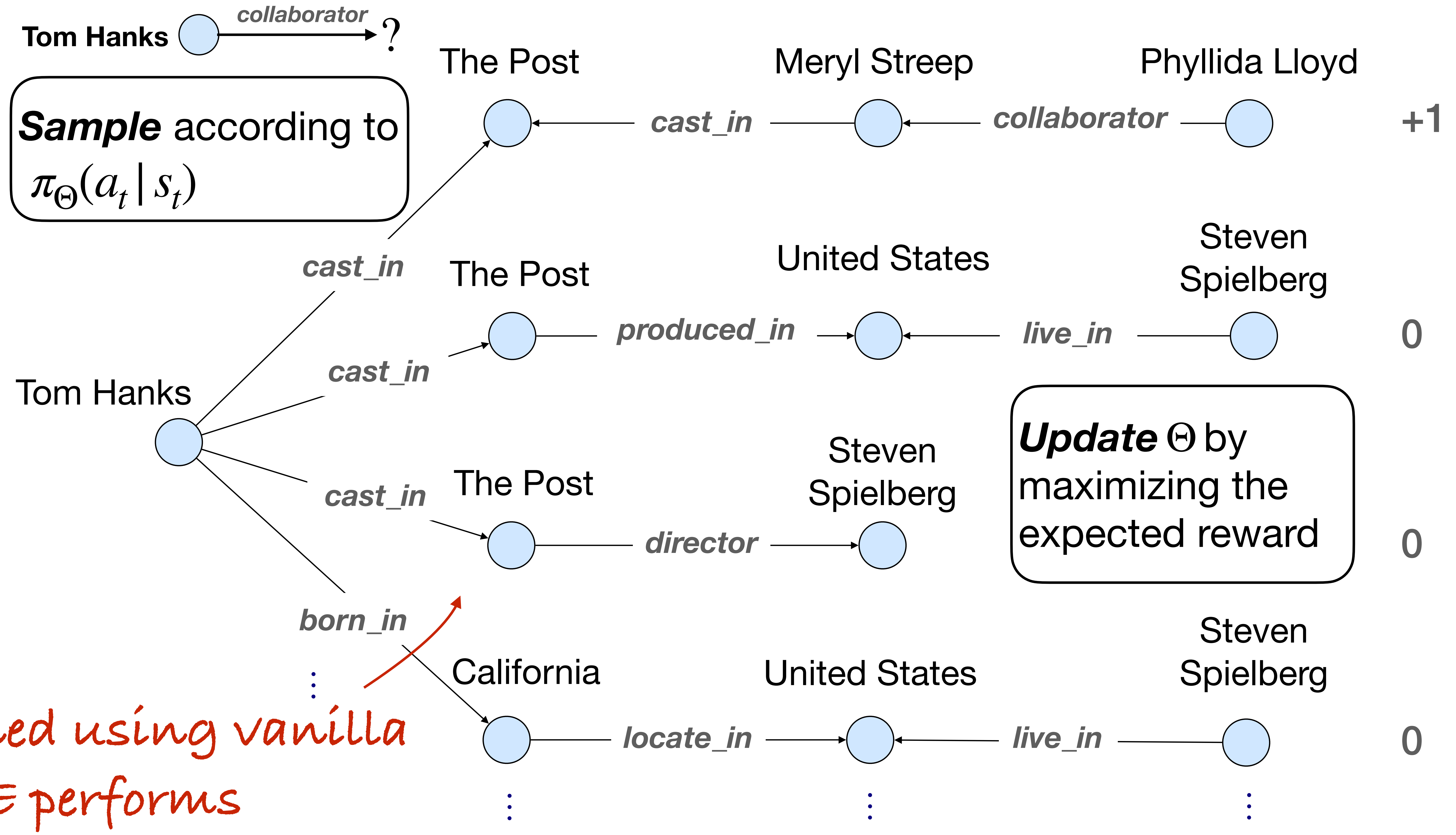
Go for A Walk And Arrive at the Answer: Reasoning Over Paths in Knowledge Bases using Reinforcement Learning (Das et al. 2018)

REINFORCE Training



REINFORCE (Williams, 1992)

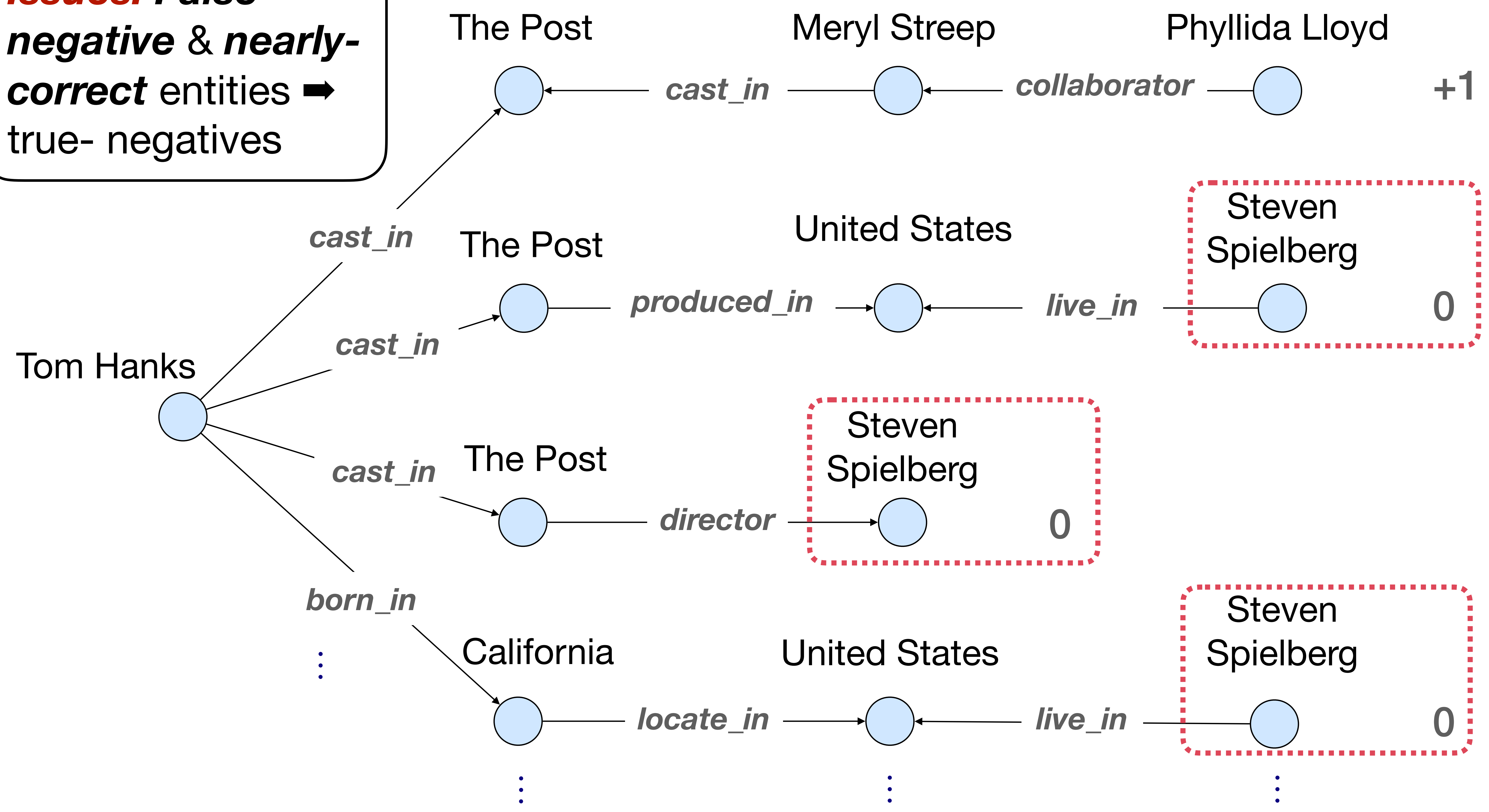
REINFORCE Training



Models trained using vanilla REINFORCE performs significantly worse compared to KG embedding baselines

Sparse Reward

Issues:* False-negative** & ***nearly-correct entities ➡ true- negatives

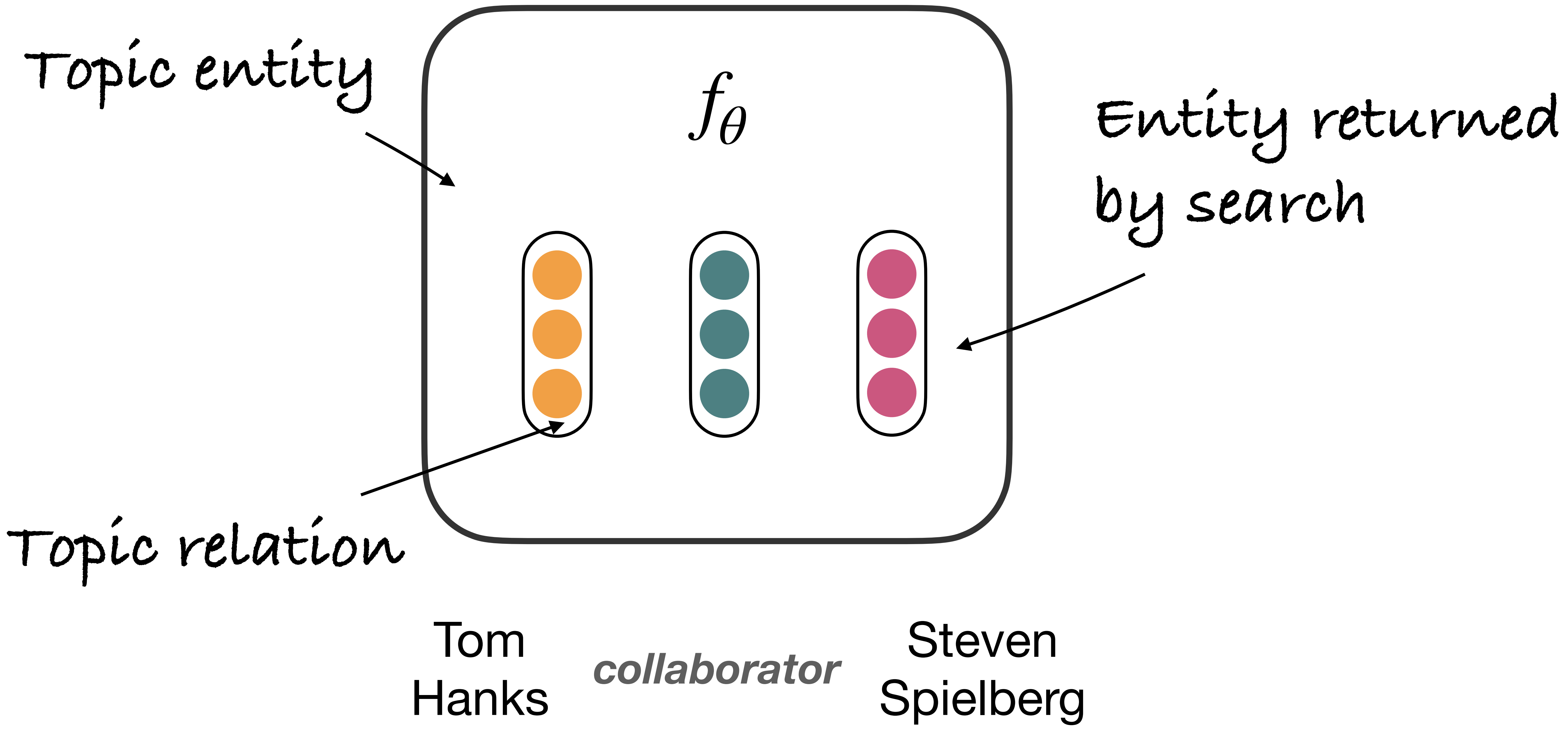


Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Reward Shaping

Unobserved facts

Soft correctness

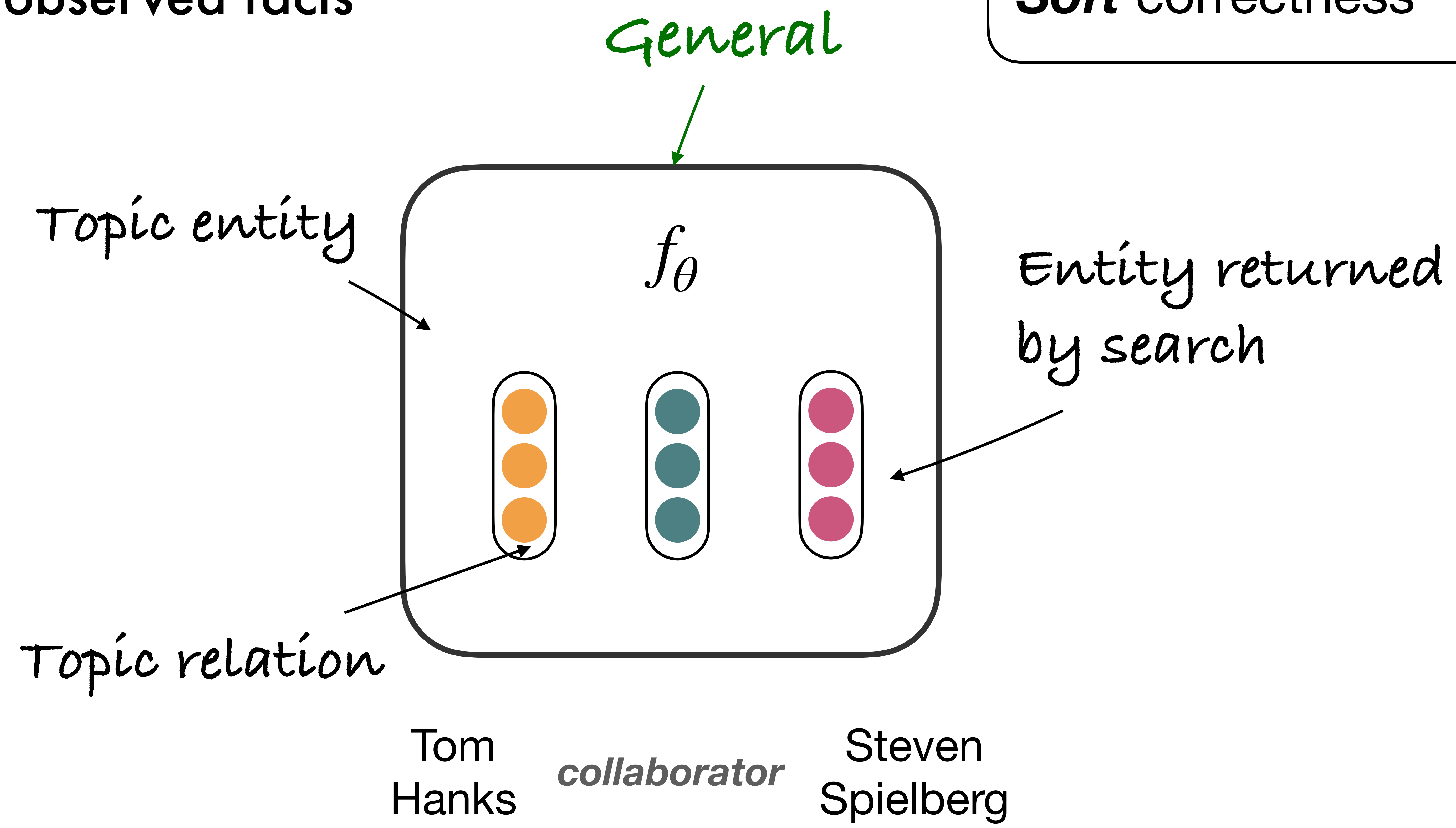


Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Reward Shaping

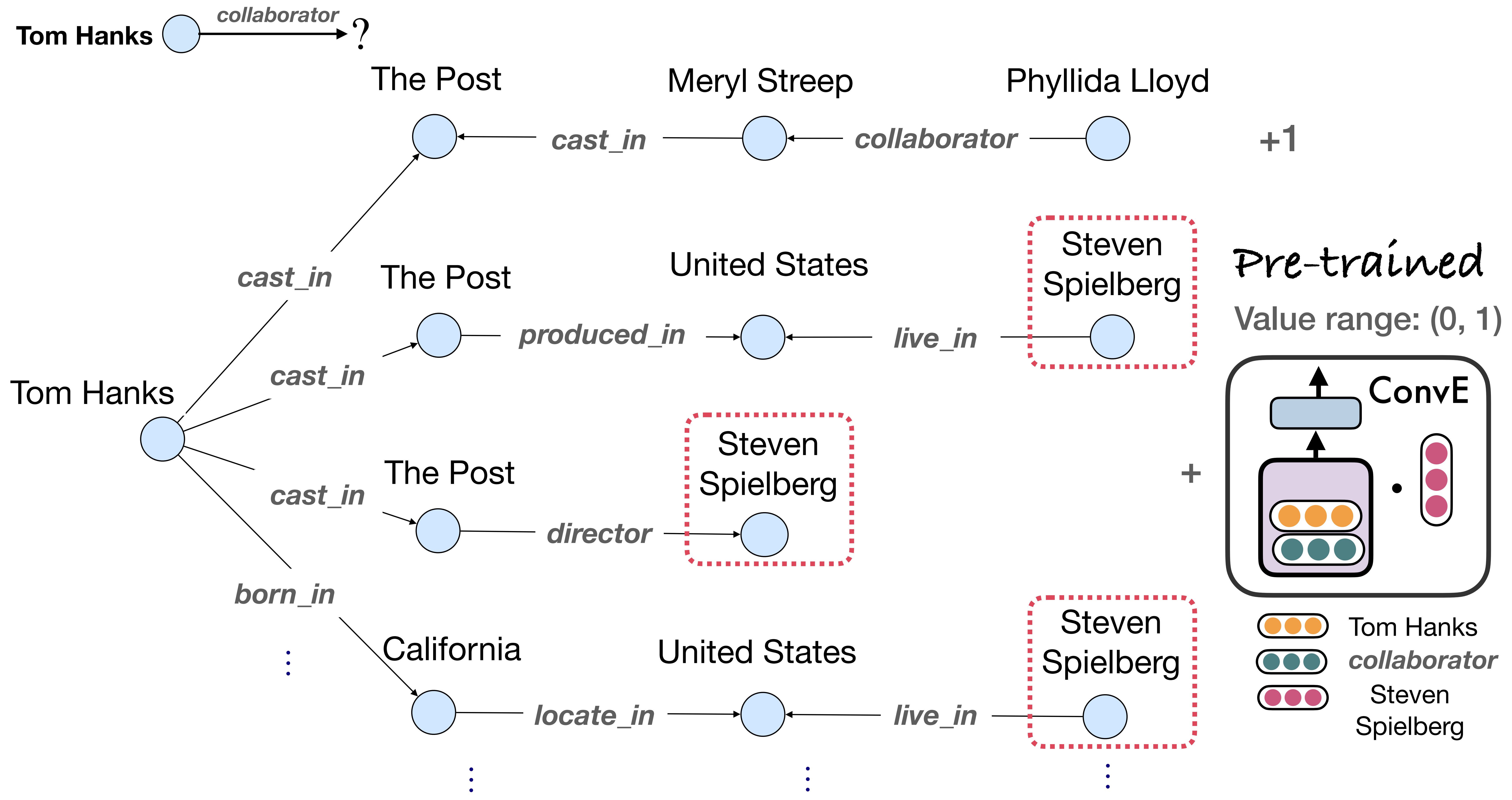
Unobserved facts

Soft correctness



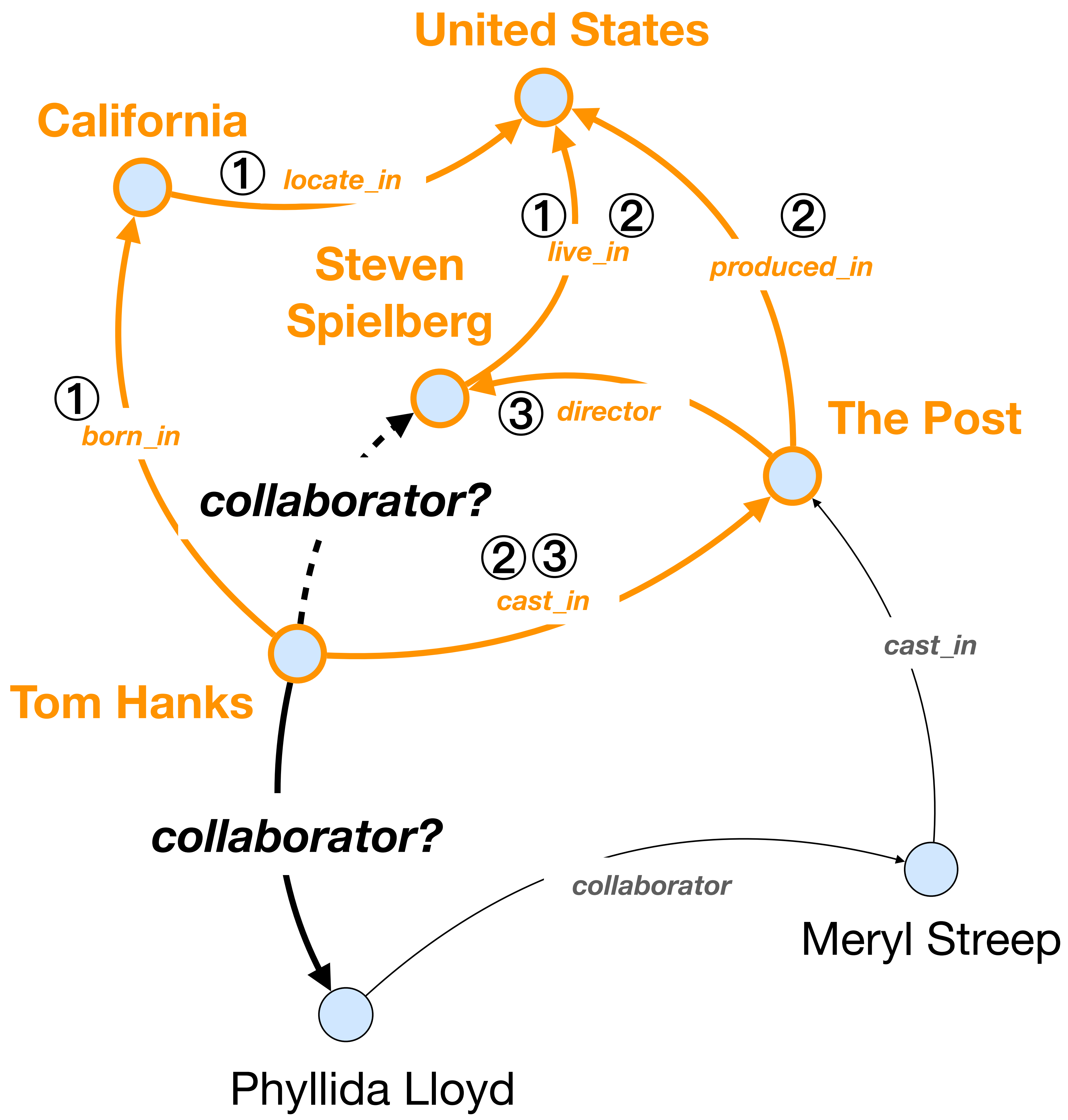
Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Reward Shaping



Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Spurious Path

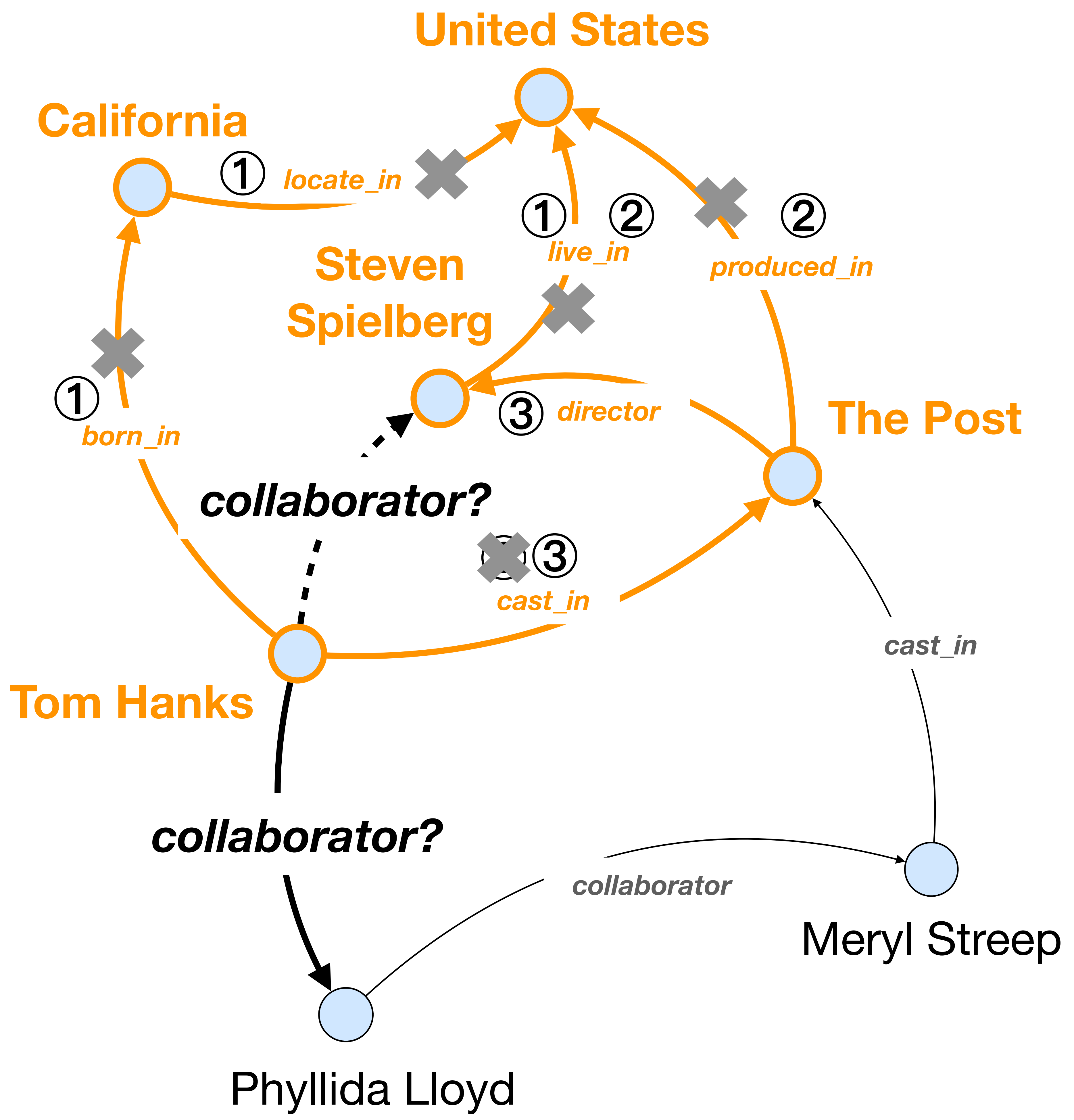


Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Spurious Path

False positive (spurious) paths ① ②

Overfit to the spurious paths



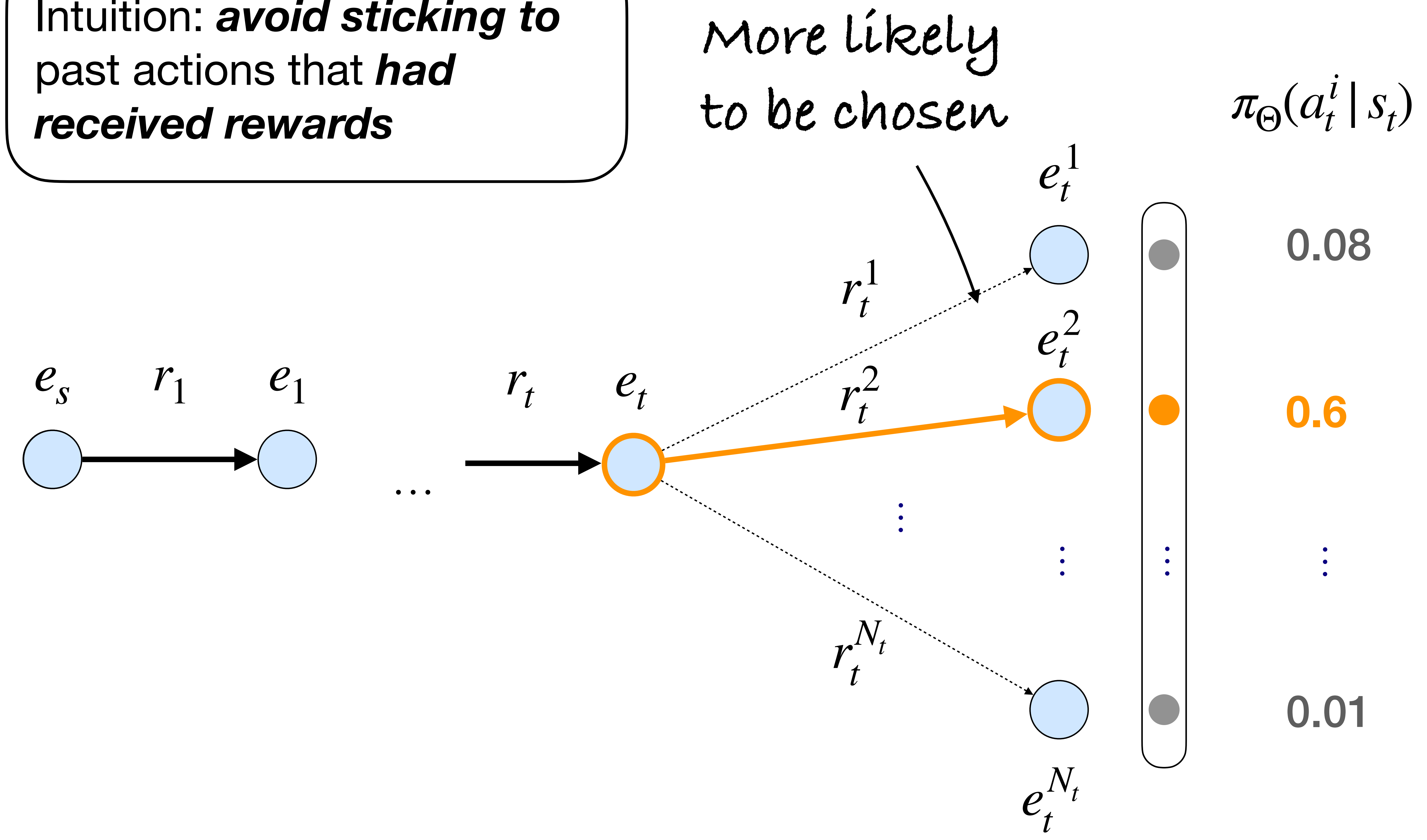
Action Dropout

Intuition: ***avoid sticking to*** past actions that ***had received rewards***

Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Action Dropout

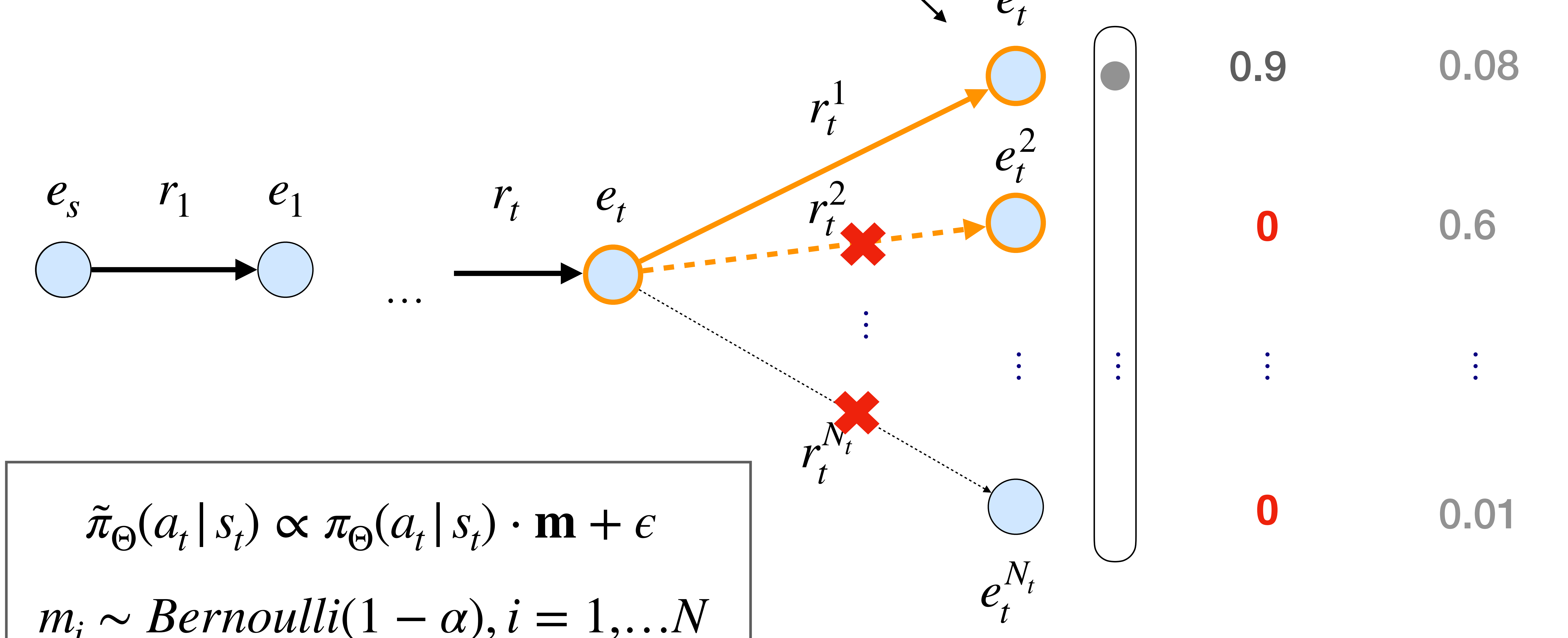
Intuition: ***avoid sticking to*** past actions that ***had received rewards***



Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Action Dropout

Randomly offset the ***sampling probabilities*** w/ rate α and renormalize



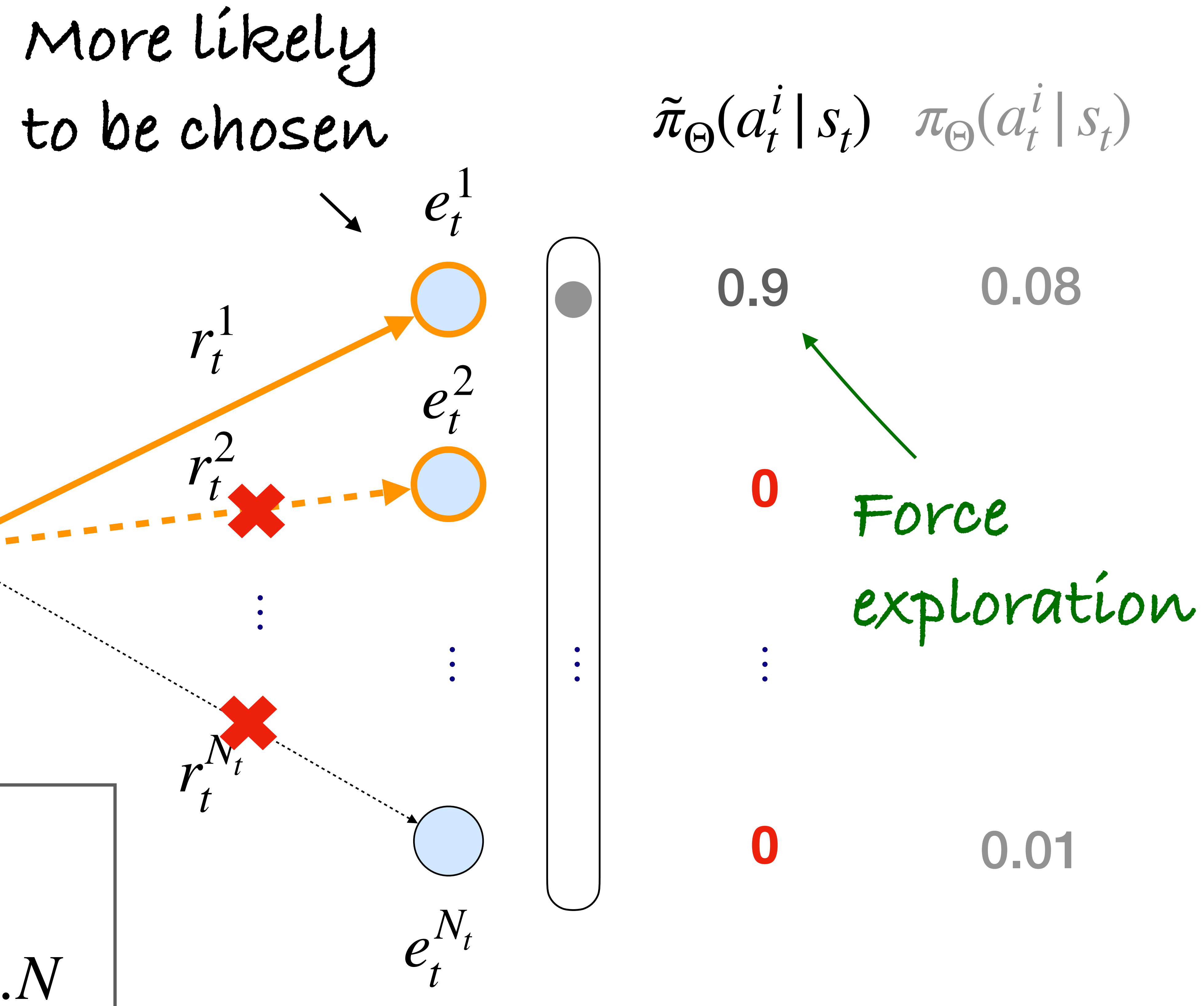
$$\tilde{\pi}_\Theta(a_t \mid s_t) \propto \pi_\Theta(a_t \mid s_t) \cdot \mathbf{m} + \epsilon$$

$$m_i \sim Bernoulli(1-\alpha), i = 1, \ldots N$$

Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Action Dropout

Randomly offset the ***sampling probabilities*** w/ rate α and renormalize



$$\tilde{\pi}_\Theta(a_t \mid s_t) \propto \pi_\Theta(a_t \mid s_t) \cdot \mathbf{m} + \epsilon$$

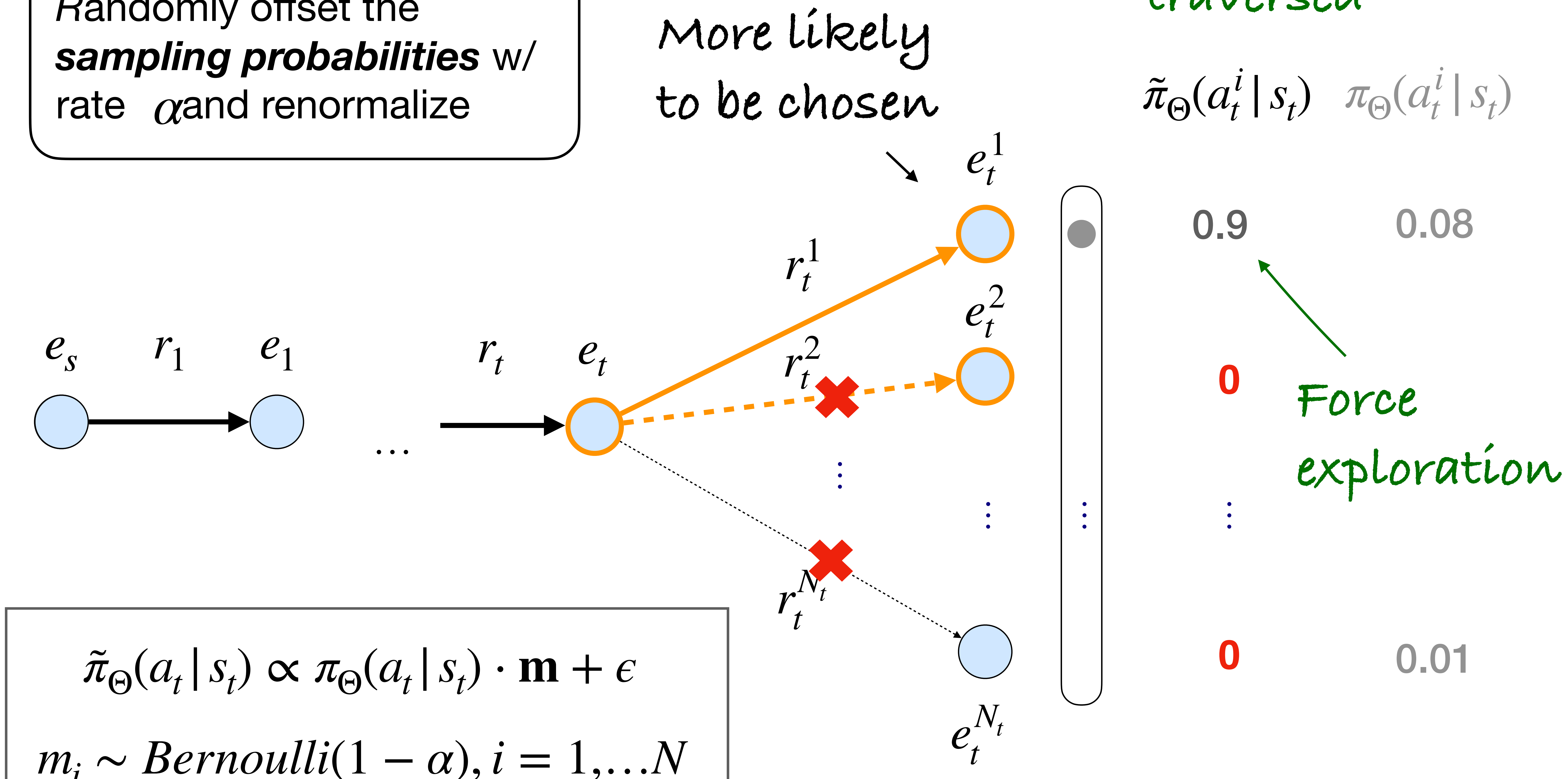
$$m_i \sim Bernoulli(1-\alpha), i = 1, \ldots N$$

Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Action Dropout

Randomly offset the ***sampling probabilities*** w/ rate α and renormalize

Up to ×8 # path traversed



$$\tilde{\pi}_\Theta(a_t | s_t) \propto \pi_\Theta(a_t | s_t) \cdot \mathbf{m} + \epsilon$$

$$m_i \sim Bernoulli(1-\alpha), i = 1, \ldots N$$

Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Experiment Setup

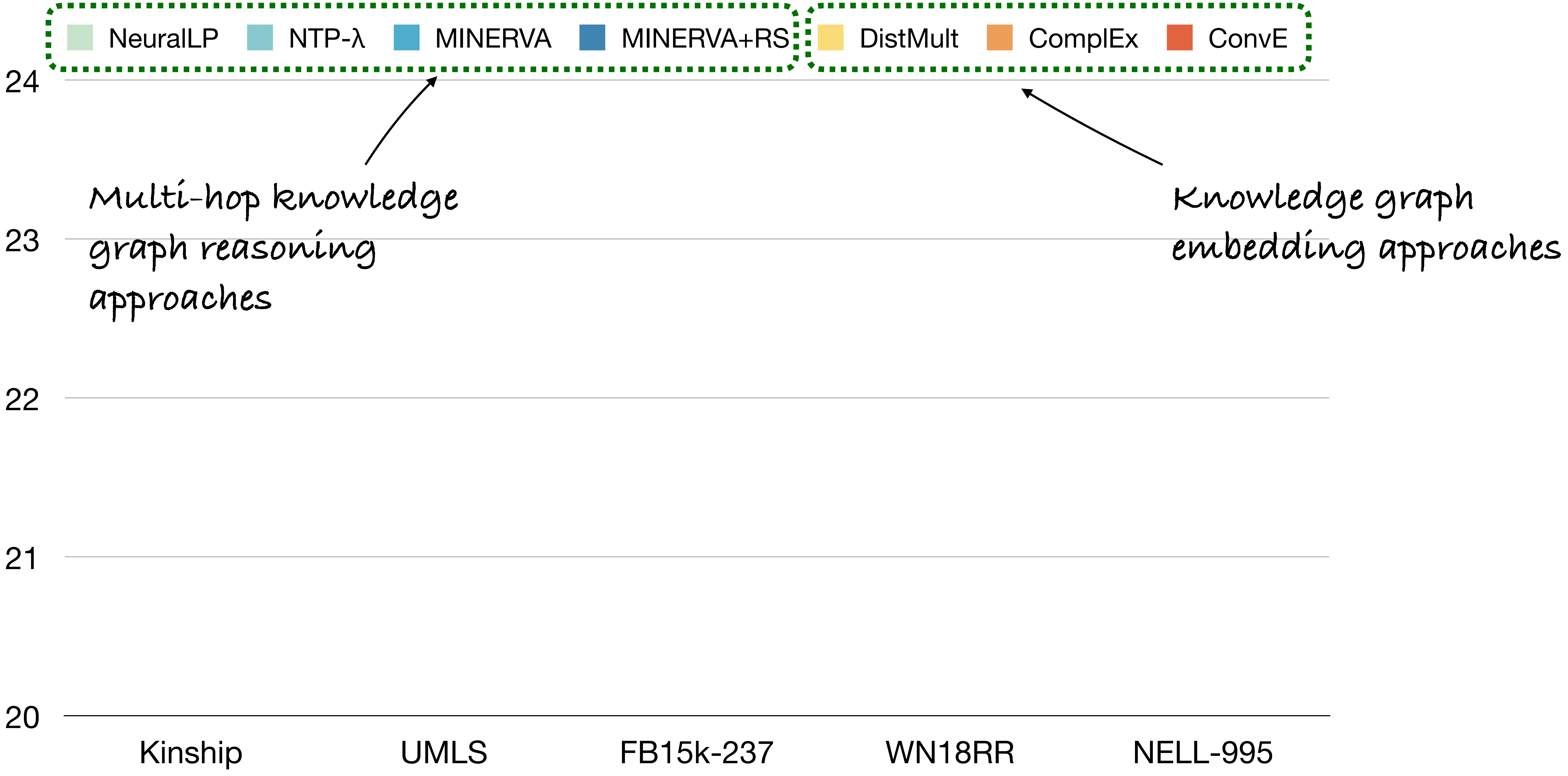
KG Benchmarks

Name	# Ent.	# Rel.	# Fact	# Degree Avg	# Degree Median
Kinship	104	25	8,544	85.15	82
UMLS	135	46	5,216	38.63	28
FB15k-237	14,505	237	272,115	19.74	14
WN18RR	40,945	11	86,835	2.19	2
NELL-995	75,492	200	154,213	4.07	1

Decreasing connectivity

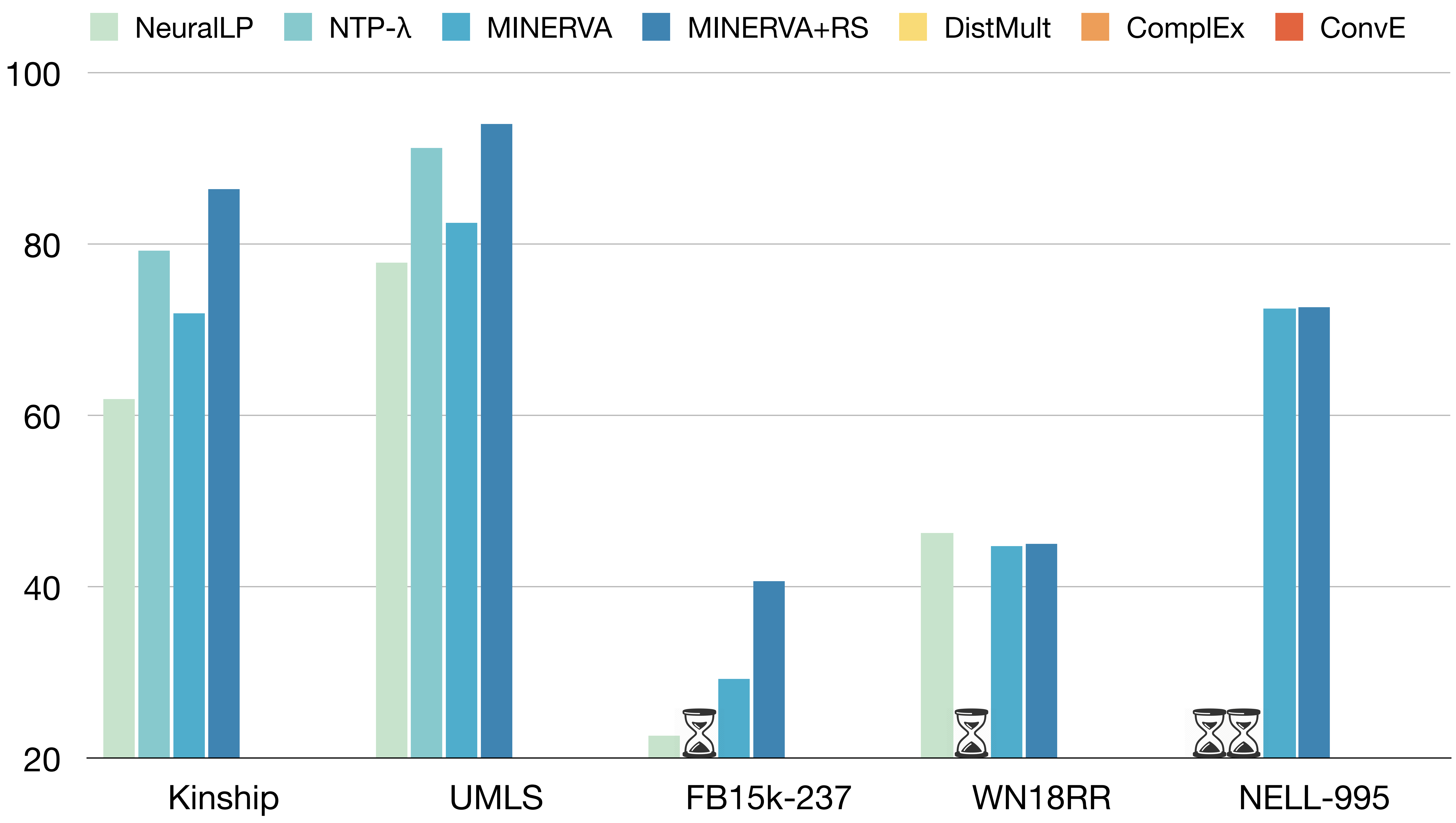
Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Main Results



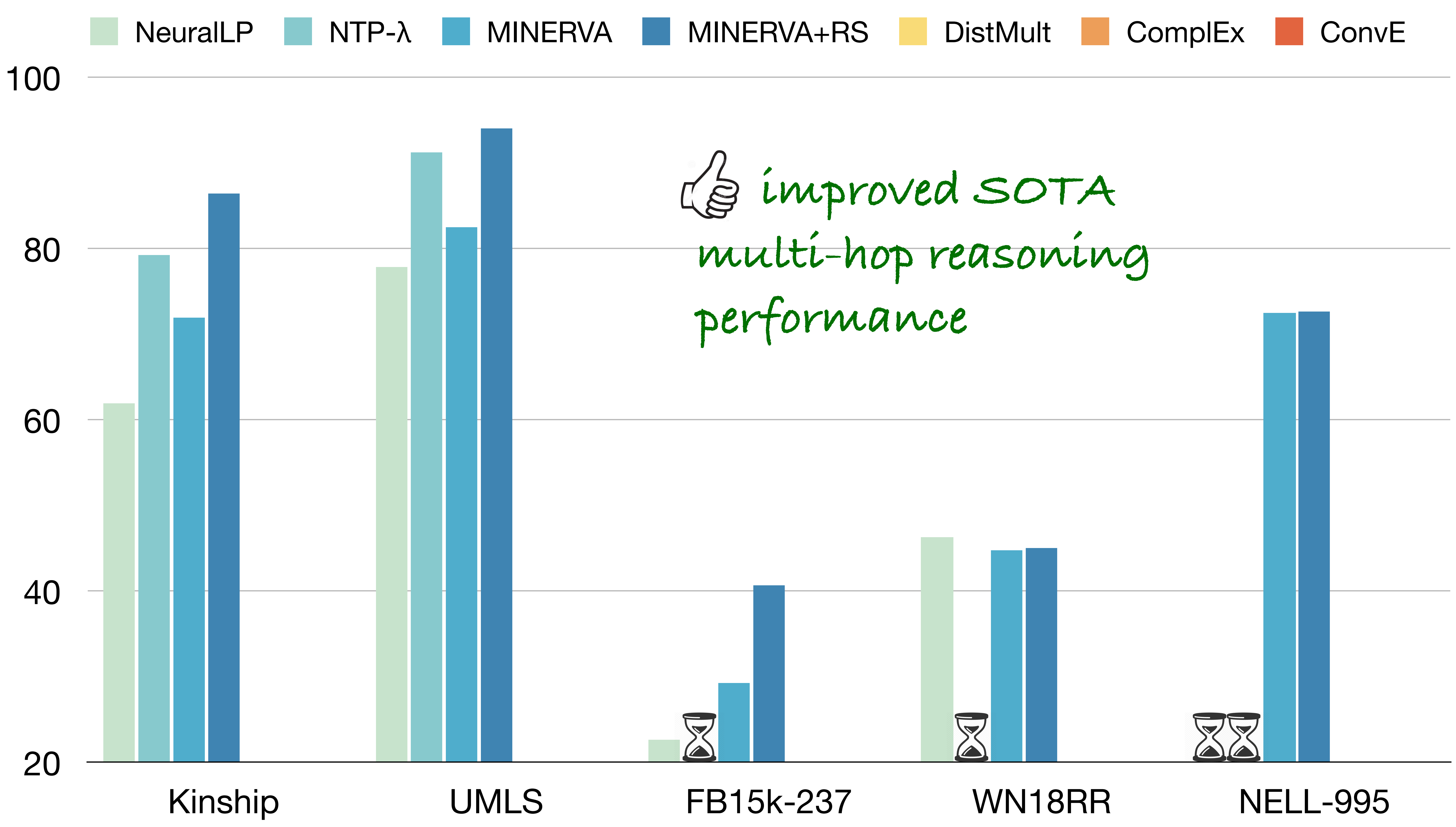
Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Main Results



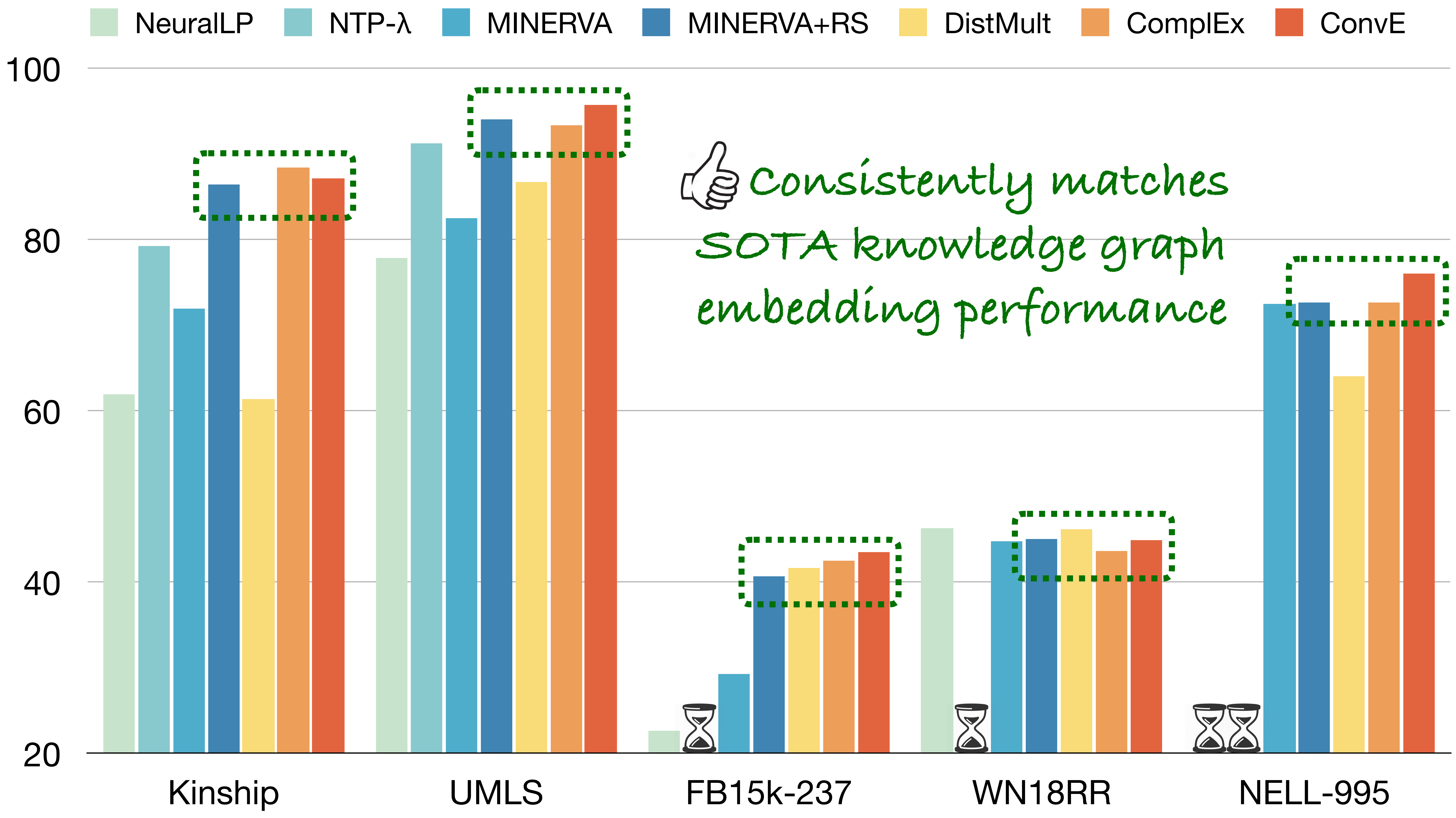
Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Main Results



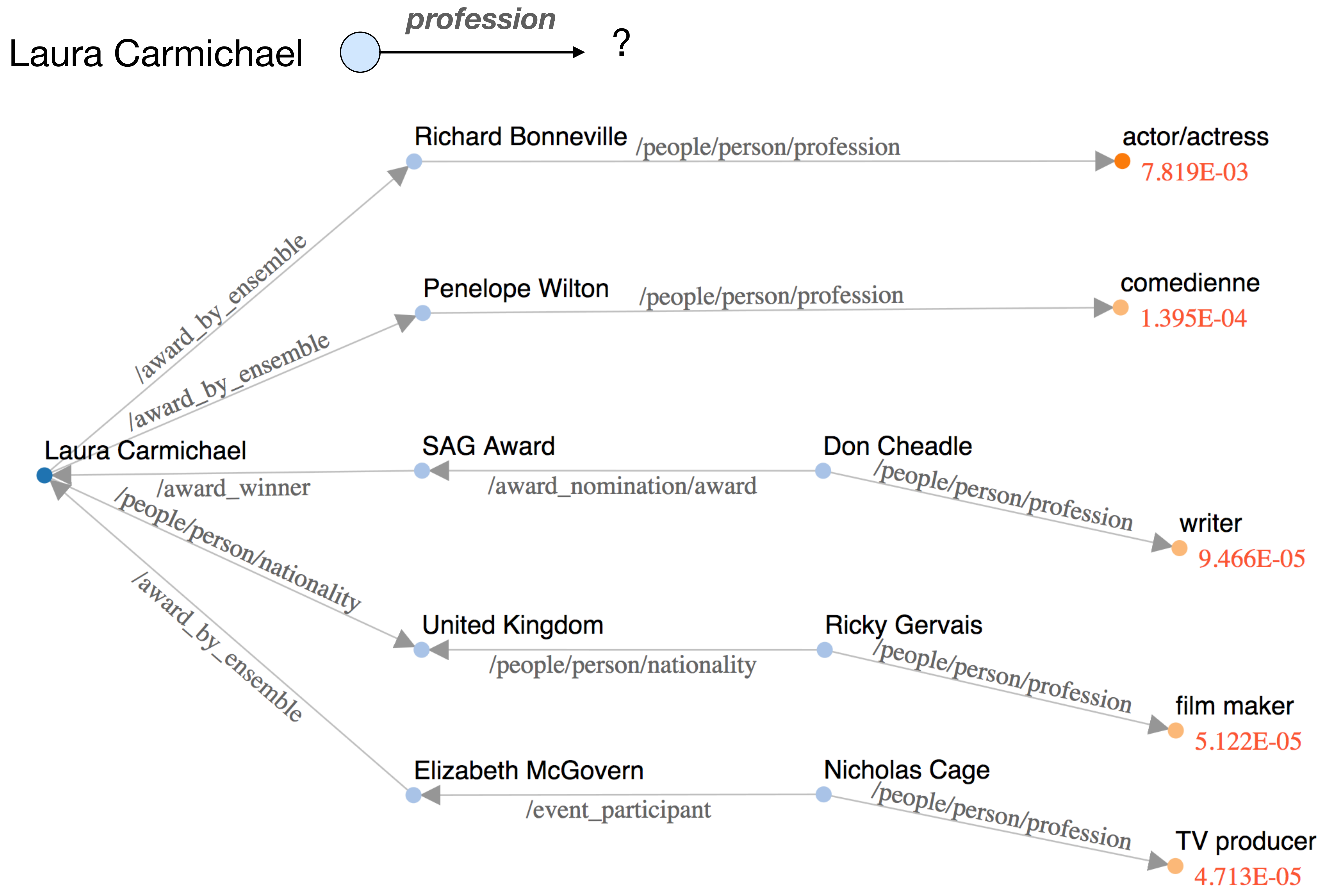
Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Main Results



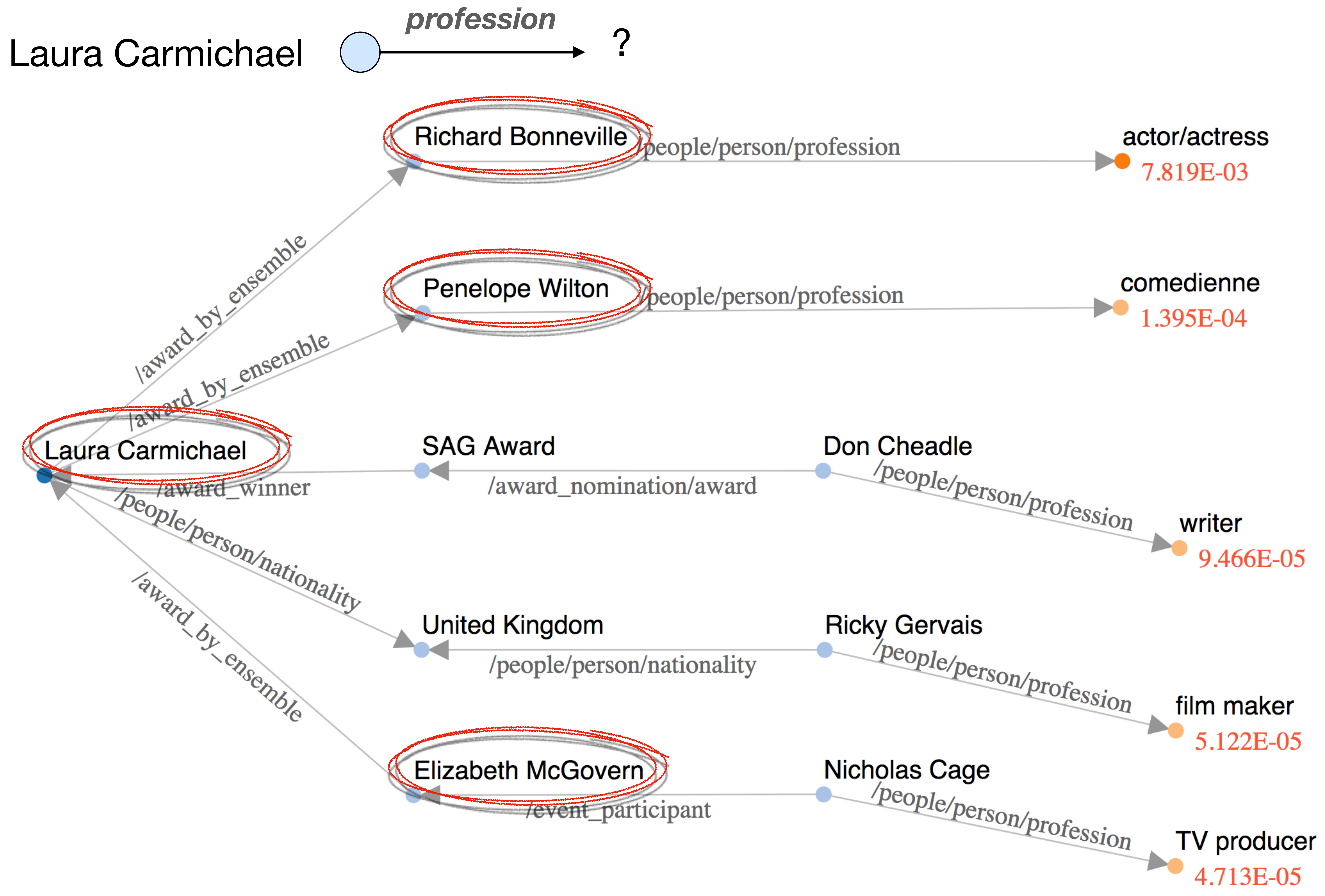
Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Interpretable Results



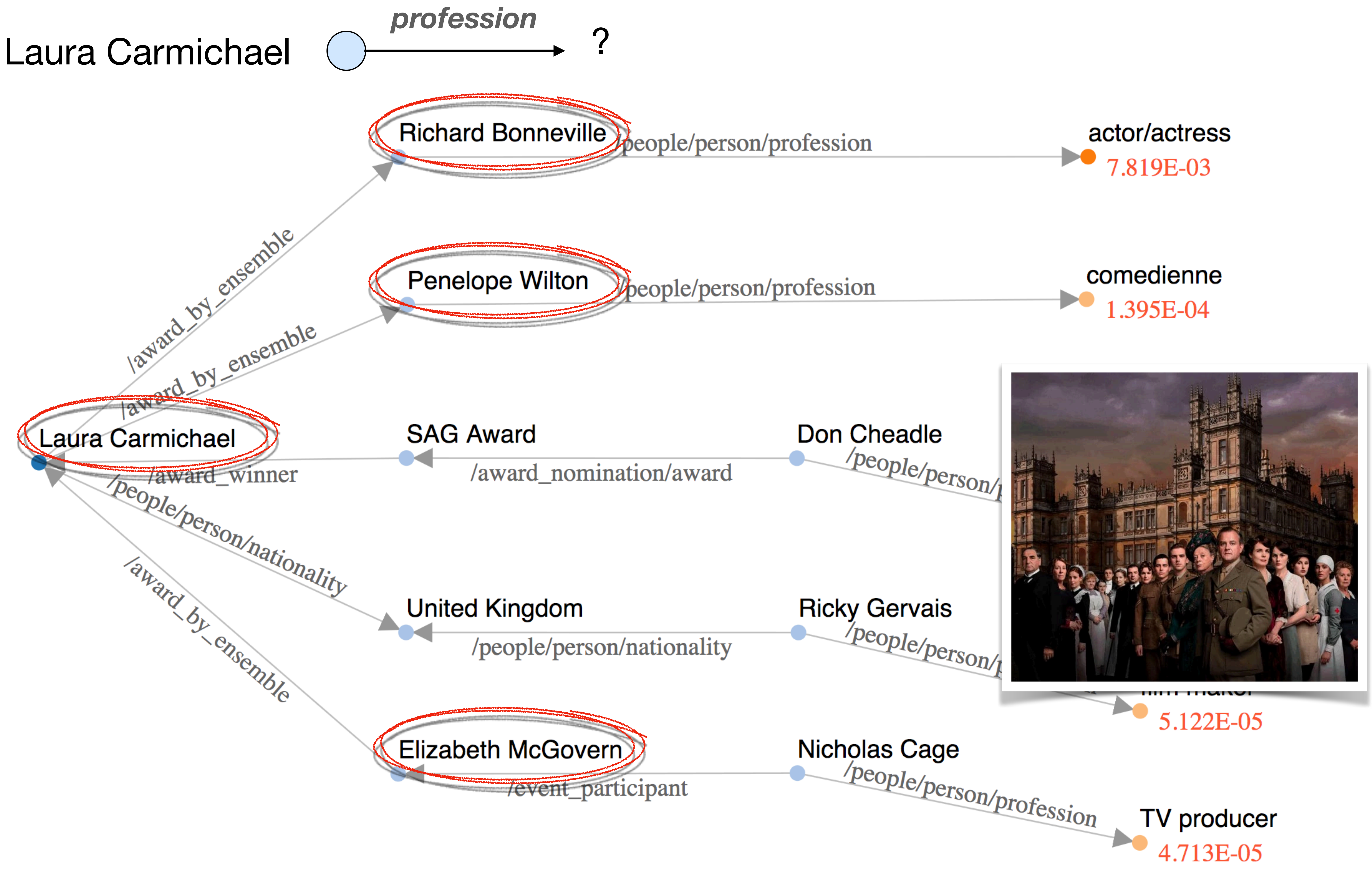
Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Interpretable Results



Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

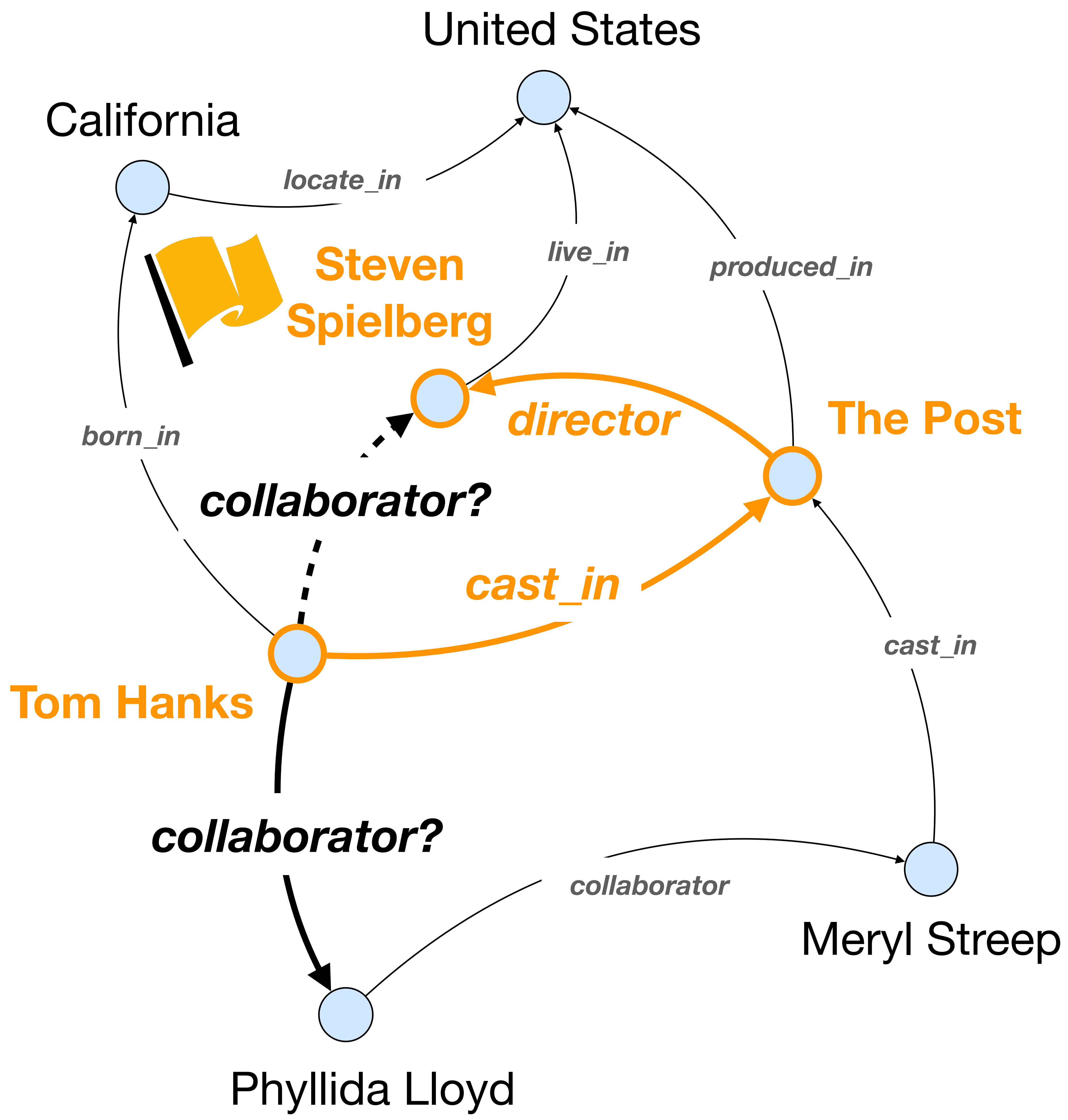
Interpretable Results



Multi-Hop Knowledge Graph Reasoning with Reward Shaping (Lin et. al. 2018)

Sequential Multi-Hop Reasoning

- Efficient for $e_s \xrightarrow{r_q} ?$ queries
- Return multiple predicted target entities with best-first search
- Can resolve $e_s \xrightarrow{r_q?} e_t$ queries



Open Source: https://github.com/salesforce/MultiHopKG

- Knowledge graph reasoning is critical for KG-based applications as KGs are intrinsically incomplete
- **(Deep) reinforcement learning** provides a strong family of algorithms for learning **informative** reasoning paths while being **time and space efficient**
- Our work combines policy network with KG embedding based reward shaping is **the first sequential multi-hop reasoning approach that matches the performance of KG embedding based approaches** on multiple benchmarks

  Multi-Hop Knowledge Graph Reasoning with Reward Shaping. Xi Victoria Lin, Richard Socher and Caiming Xiong. EMNLP 2018.

- Future work could learn more from core RL research to resolve generic (e.g. sparse reward) and KG-specific learning challenges

  M-walk: Learning to Walk over Graphs Using Monte Carlo Tree Search (Shen et. al. 2018)
  Reinforcement Knowledge Graph Reasoning for Explainable Recommendation (Xian et. al. 2019)
  Collaborative Policy Leaning for Open Knowledge Graph Reasoning (Fu et. al. 2019)
  Path Reasoning over Knowledge Graph: A Multi-Agent and Reinforcement Learning Based Method (Li et. al. 2019)
  Reinforcement Learning Based Meta-Path Discovery in Large-Scale Heterogeneous Information Networks (Wan et. al. 2020)
  Reasoning Like Human: Hierarchical Reinforcement Learning for Knowledge Graph Reasoning (Wan et. al. 2020)


thank you
BLAZE YOUR TRAIL
salesforce