Forcing Neural Link Predictors to Play by the Rules

- Sebastian Riedel

Collaborators

Johannes Welbl

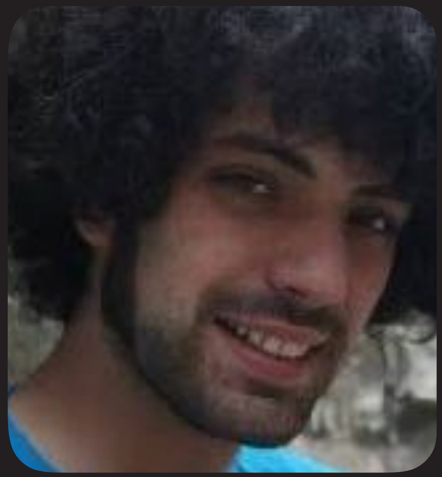
(UCL)

Pasquale Minervini

(UCL)

Thomas Demeester

(University of Ghent)



Theo Trouillon

(Xerox Research)

Tim Rocktäschel

(now at Oxford)

Guillaume Bouchard

(now at Bloomsbury AI)

Tim Dettmer

(University of Lugano)

Goal

If higher, more likely true

Calculate Truth Scores for **Statements**

learn from data and prior knowledge, use **neural networks**, **graphical models** etc.

in **natural language**, predicate **logic**, **atoms**, **rules** etc.

Running Example

$$s\big(\text{Enrique can speak Spanish}\big) = 0.9$$

Prior Knowledge (Text)

Enrique works in Mexico City

Enrique lives in Mexico

Training data (Text)

Luis speaks Spanish

Luis lives in Mexico

Training data (Structure)

livesIn(Ivan, Mexico)

speaks(Ivan, Spanish)

When someone lives in Mexico, they will likely speak Spanish

Compliance

$$s(\text{I can accept tickets to NBA court-side events}) = 0.1$$

You must not accept tickets to expensive sports events

NBA court-side tickets costs at least $1000

with Bloomsbury AI

Fact Checking

$$s(\text{ US Unemployment is 42\% }) = 0.1$$

The official unemployment rate is the US is 4.4%

There are 420 billion potential working hours

only 240 billion working hours were actually recorded

with FACTMATA

Deutsche Welle

BBC

European Commission

Scientific Text

$s($ DrugA interacts with DrugB $) = 0.9$

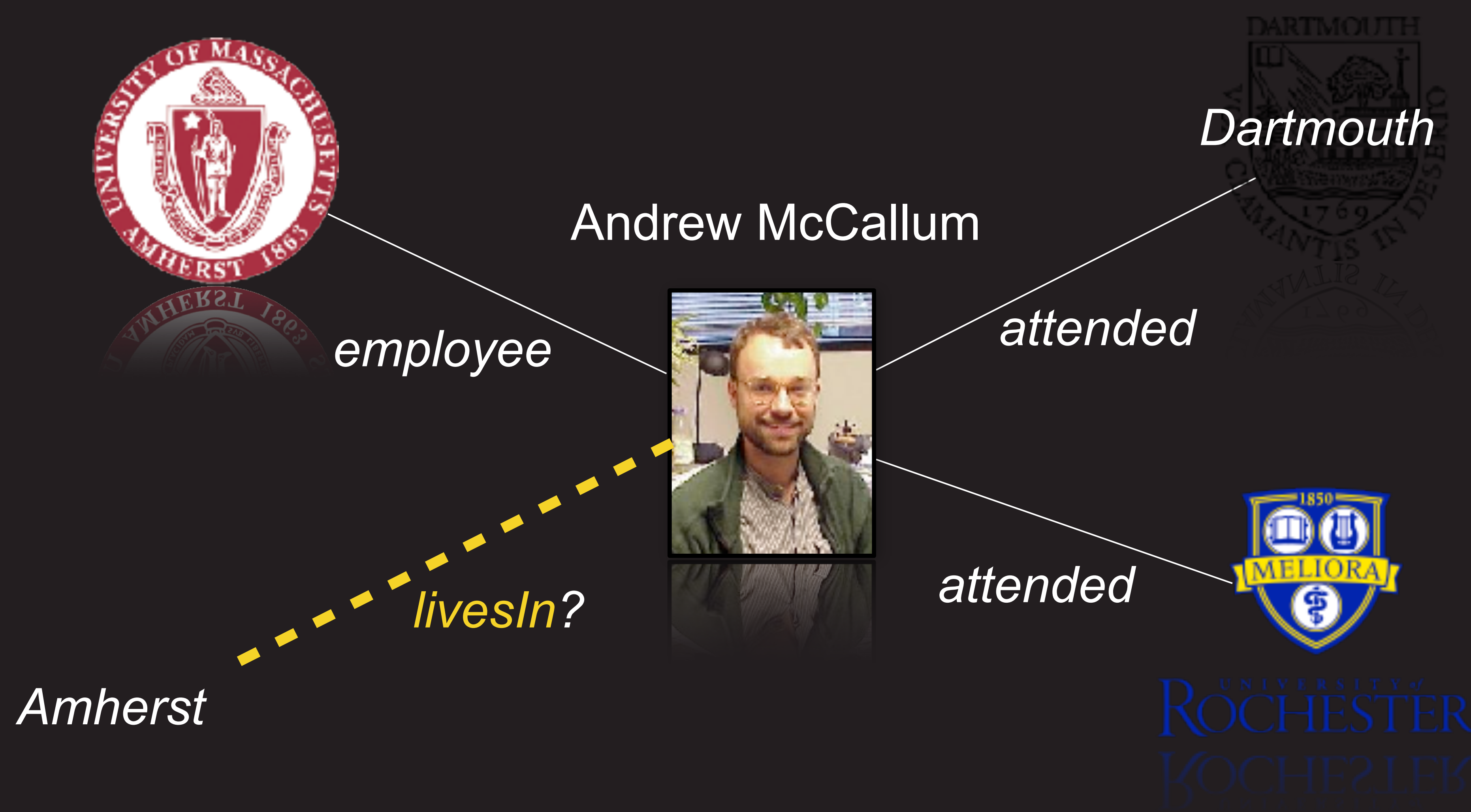
DrugA interacts with Protein1

Protein1 has been shown to stimulate Protein2

DrugB increases the rate of Protein2 …

Knowledge Graphs

$$S\left(\text{AndrewMcCallum livesIn Amherst}\right) = 0.9$$



Part 1: Learning To Score

$$s(\text{Enrique can speak Spanish}) = 0.9$$

Enrique works in Mexico City

Enrique lives in Mexico

Luis can speak Spanish

Luis lives in Mexico

livesIn(Ivan, Mexico)

speaks(Ivan, Spanish)

When someone lives in Mexico, they will likely speak Spanish

ICML 2016
JLMR 2017
AAAI 2018

Part 2: Injecting Prior Knowledge

s(Enrique can speak Spanish) = 0.9

Enrique works in Mexico City

Enrique lives in Mexico

Luis can speak Spanish

Luis lives in Mexico

livesIn(Ivan, Mexico)

speaks(Ivan, Spanish)

NAACL 2015
EMNLP 2016
UAI 2017

When someone lives in Mexico, they will likely speak Spanish

Part 1: Learning To Score

$$s(\text{Enrique can speak Spanish}) = 0.9$$

Enrique works in Mexico City

Enrique lives in Mexico

Luis can speak Spanish

Luis lives in Mexico

livesIn(Ivan, Mexico)

speaks(Ivan, Spanish)

When someone lives in Mexico, they will likely speak Spanish

Simplifying the Problem

$$s\left(\text{Enrique can speak Spanish}\right) = 0.9$$

Enrique works in Mexico City

Enrique lives in Mexico

Luis can speak Spanish

Luis lives in Mexico

livesIn(Ivan, Mexico)

speaks(Ivan, Spanish)

Simplifying the Problem: Link Prediction

$$s(\text{Enrique canSpeak Spanish}) = 0.9$$

Enrique worksIn MexicoCity

Enrique livesIn Mexico

Luis canSpeak Spanish

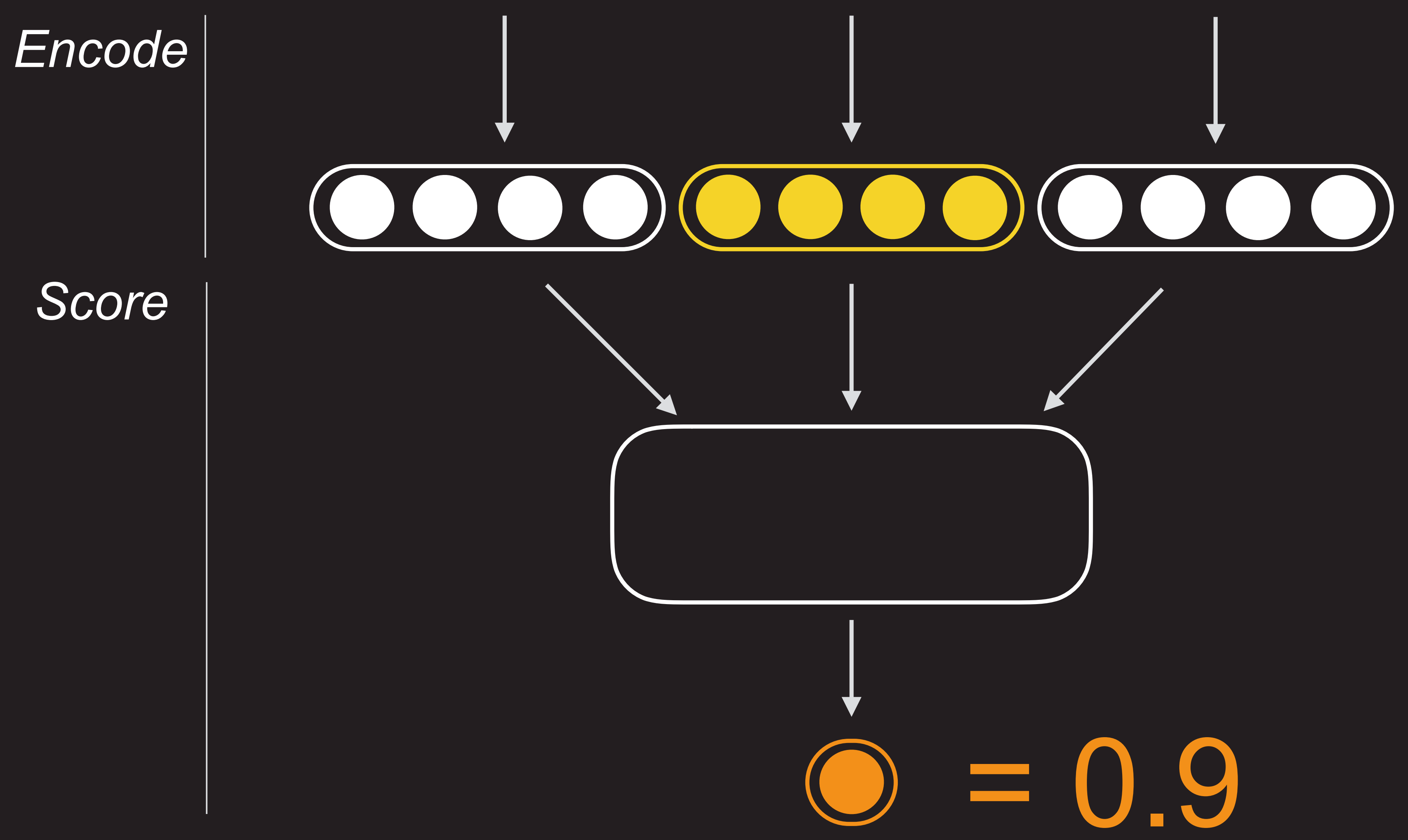
Luis livesIn Mexico

livesIn(Ivan, Mexico)

speaks(Ivan, Spanish)

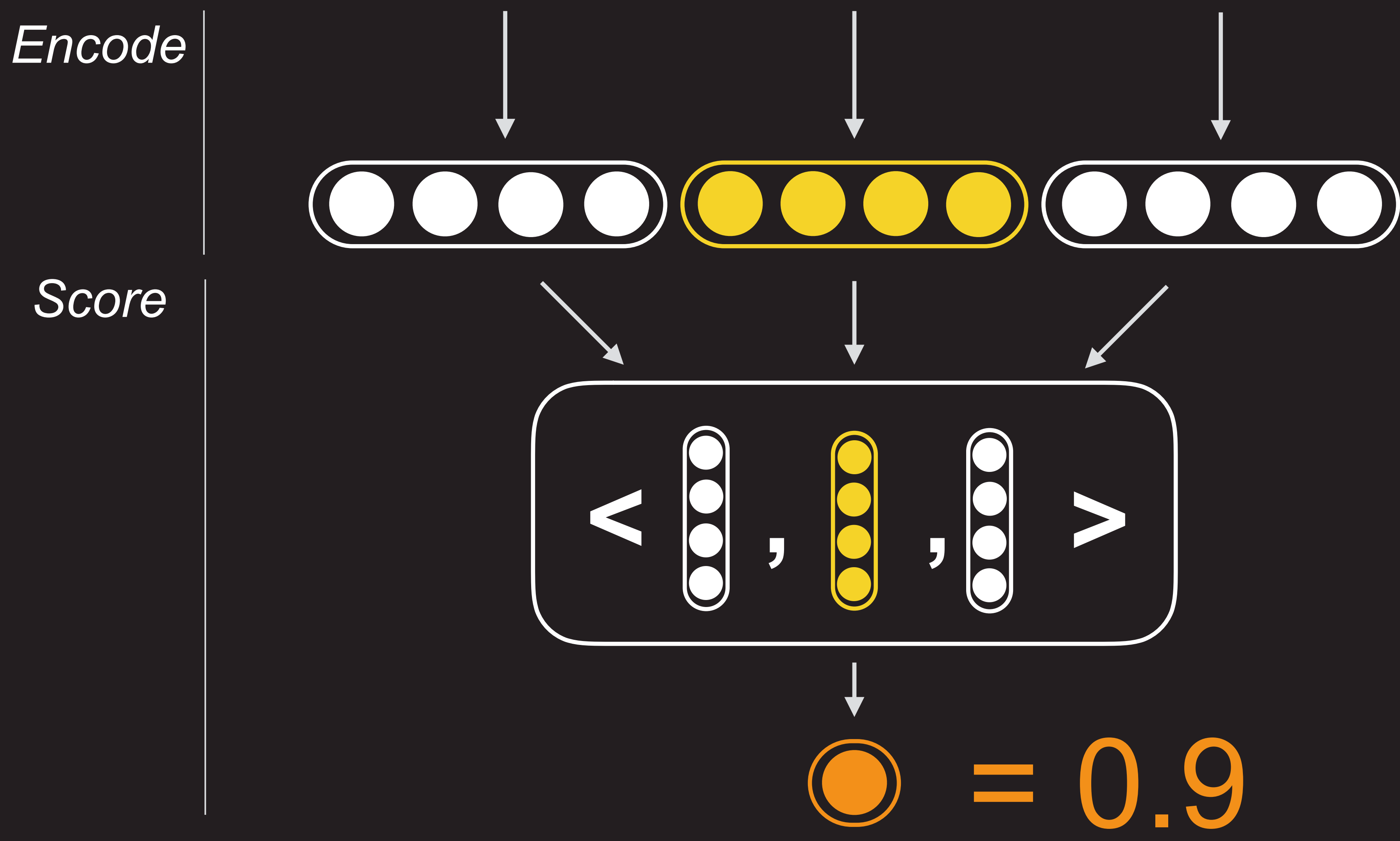
Simplifying the Problem: Triples

S(Enrique canSpeak Spanish) = 0.9



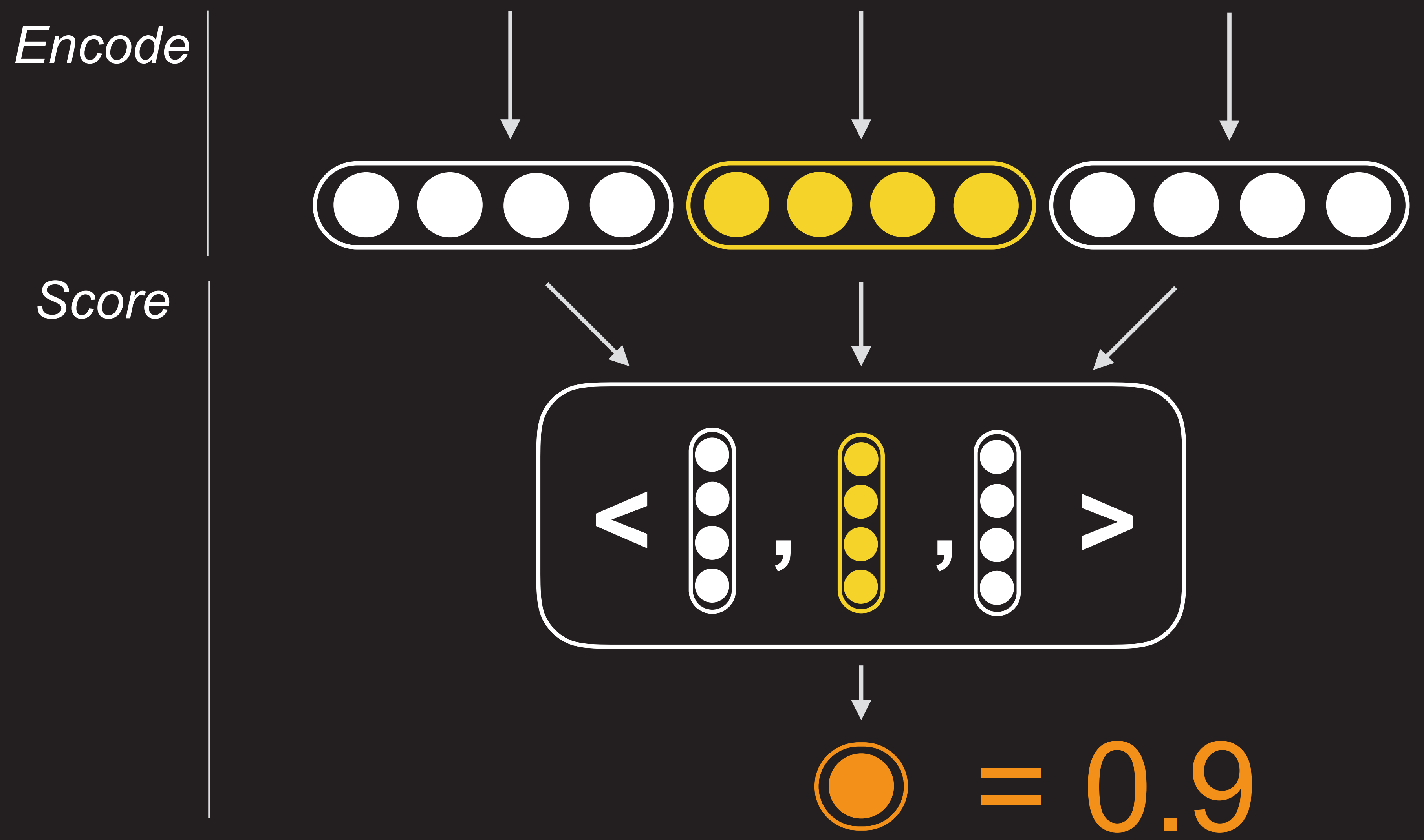
DistMult (Yang et. al, 2015)

S(Enrique canSpeak Spanish) = 0.9



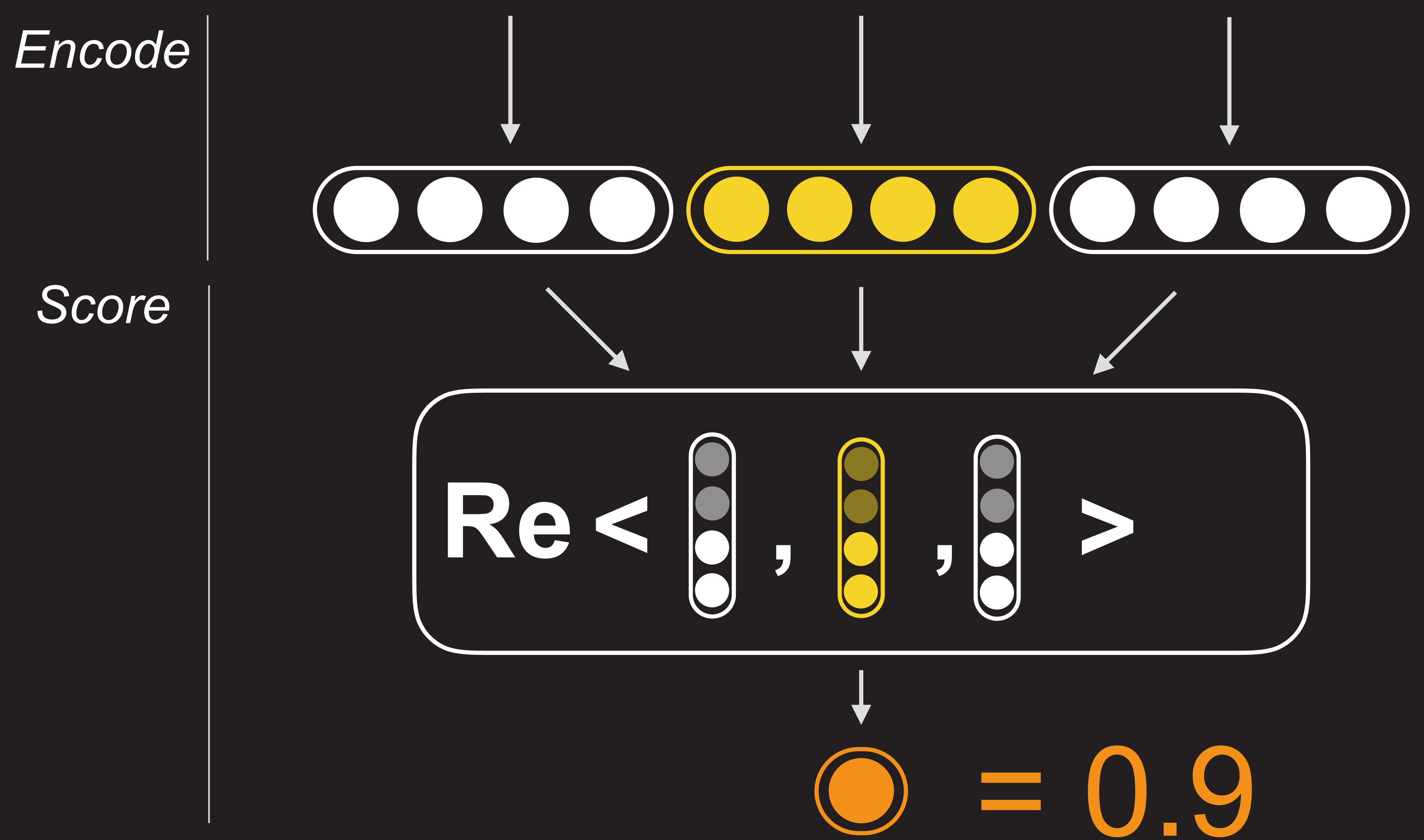
Symmetric

$$s(\text{Spanish canSpeak Enrique}) = 0.9$$



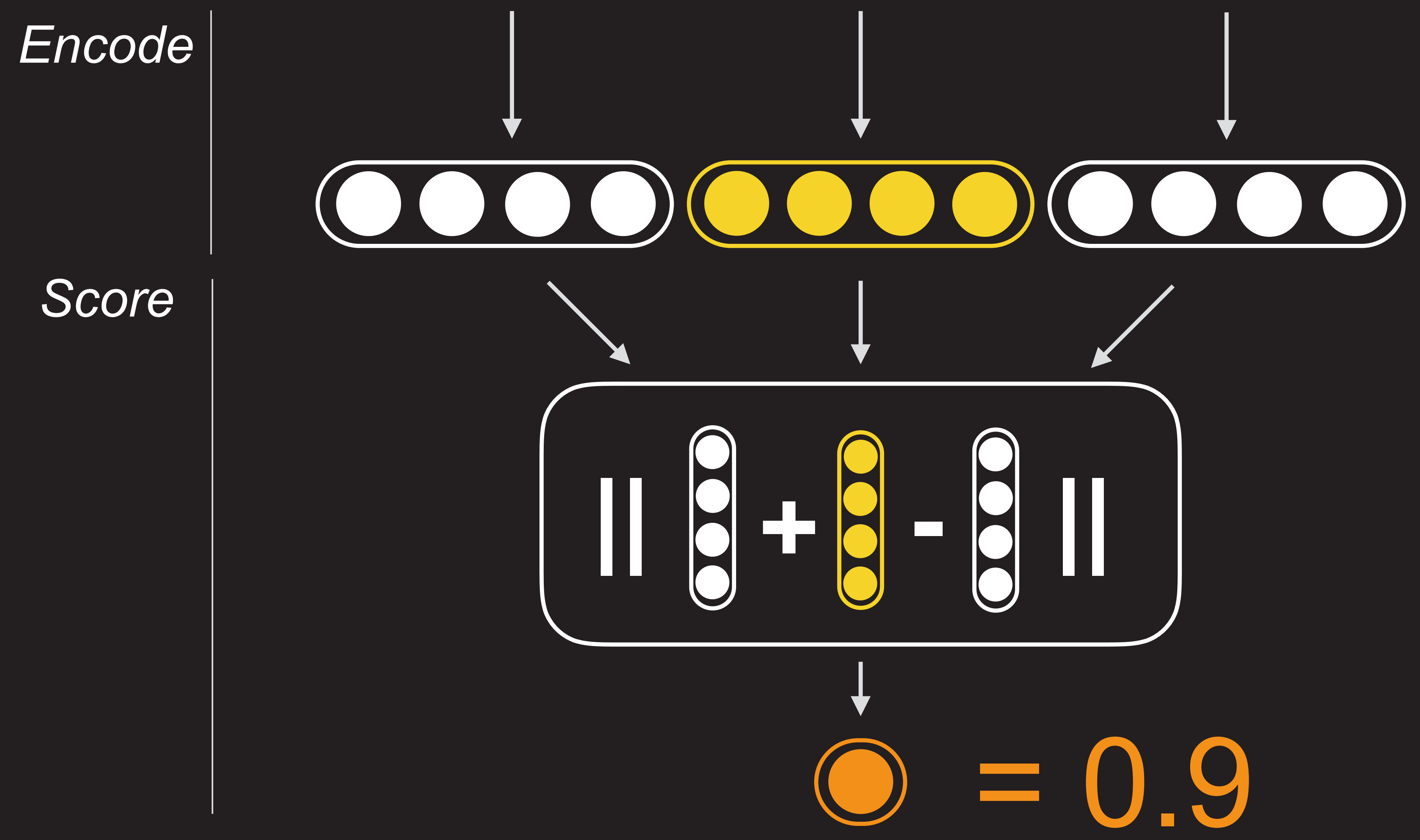
Complex (Trouillon et al, 2016)

S(Enrique canSpeak Spanish) = 0.9

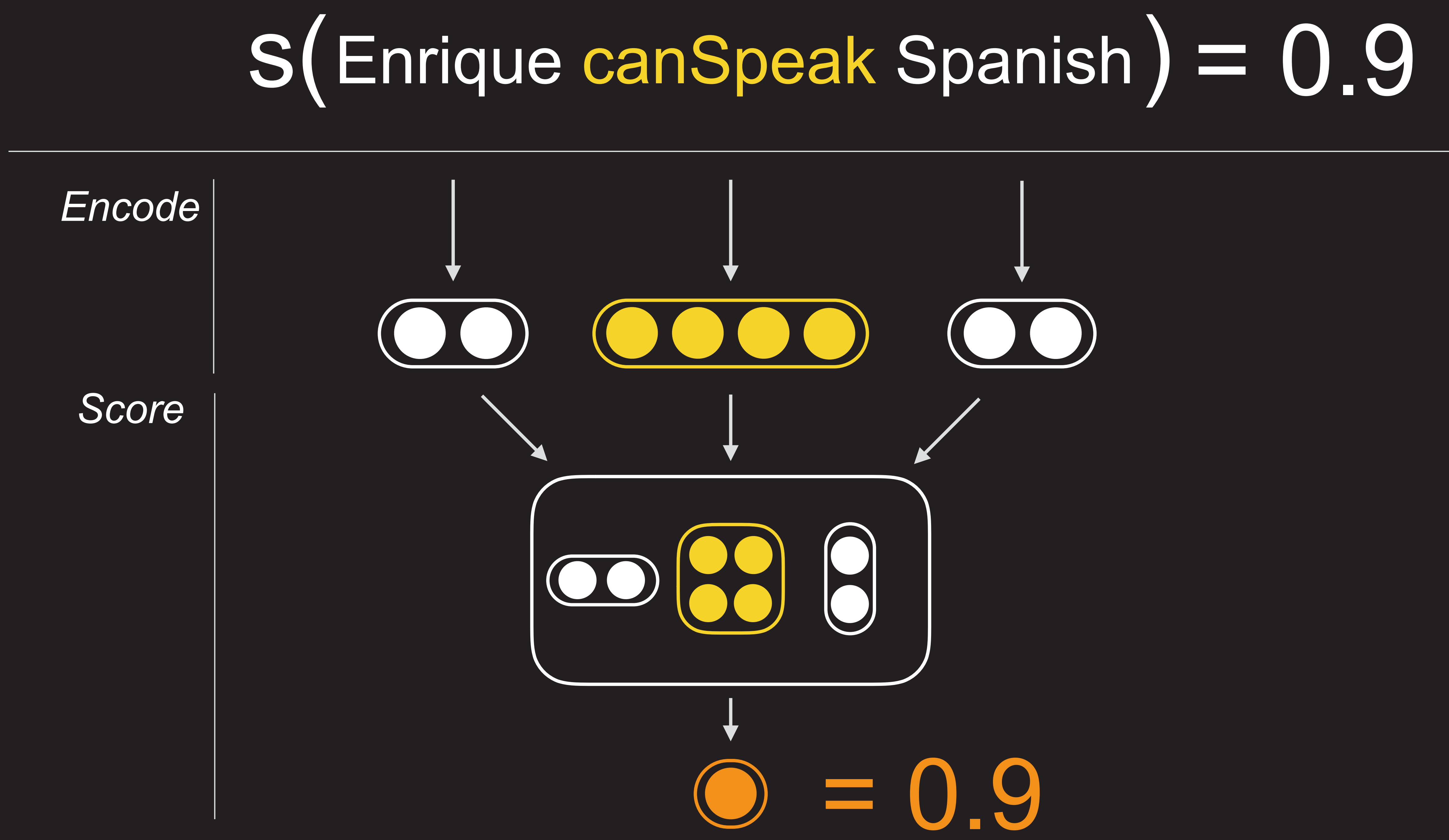


TransE (Bordes et al. 2013)

S(Enrique canSpeak Spanish) = 0.9

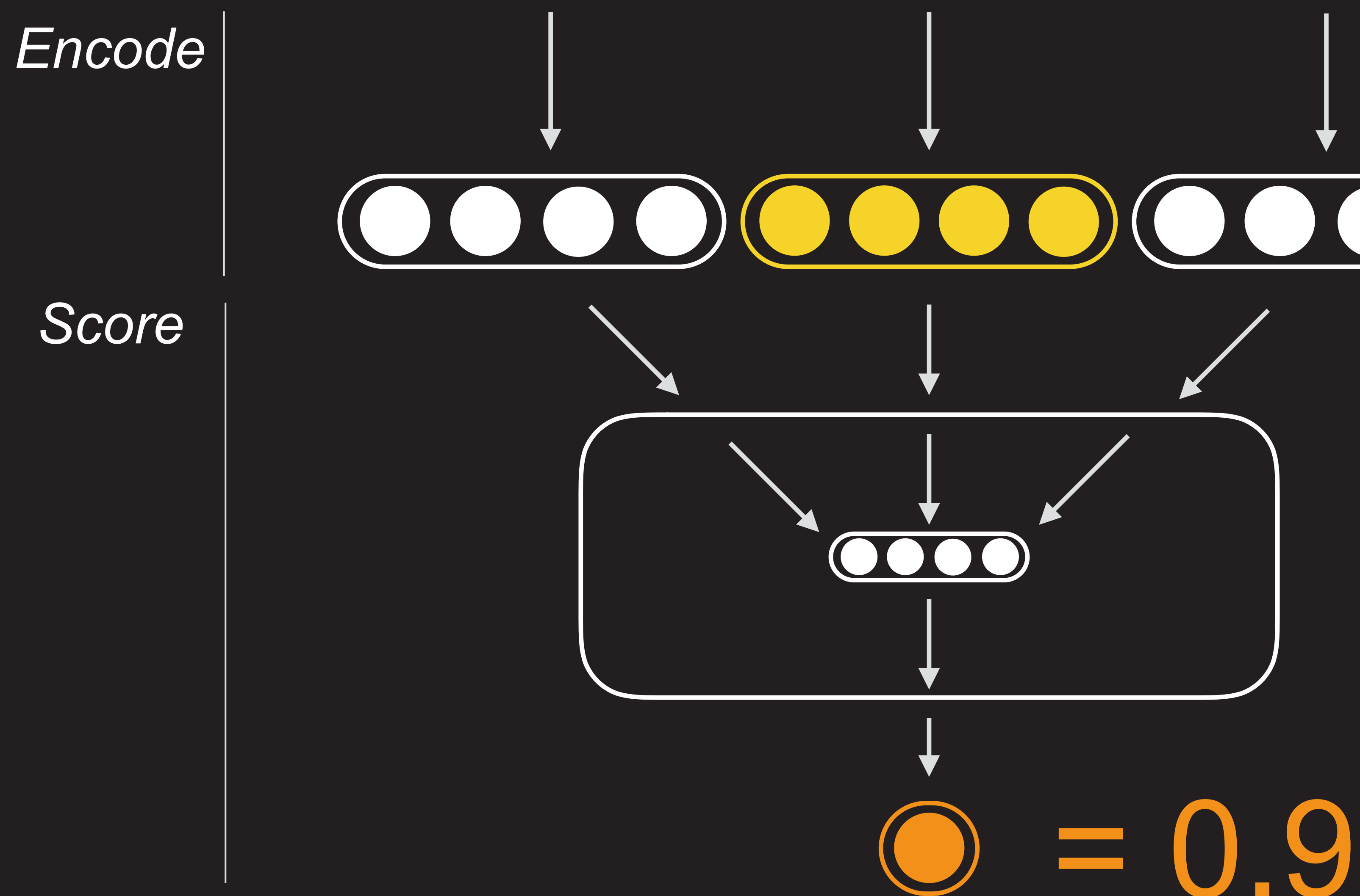


Rescal (Nickel et. al 2013)



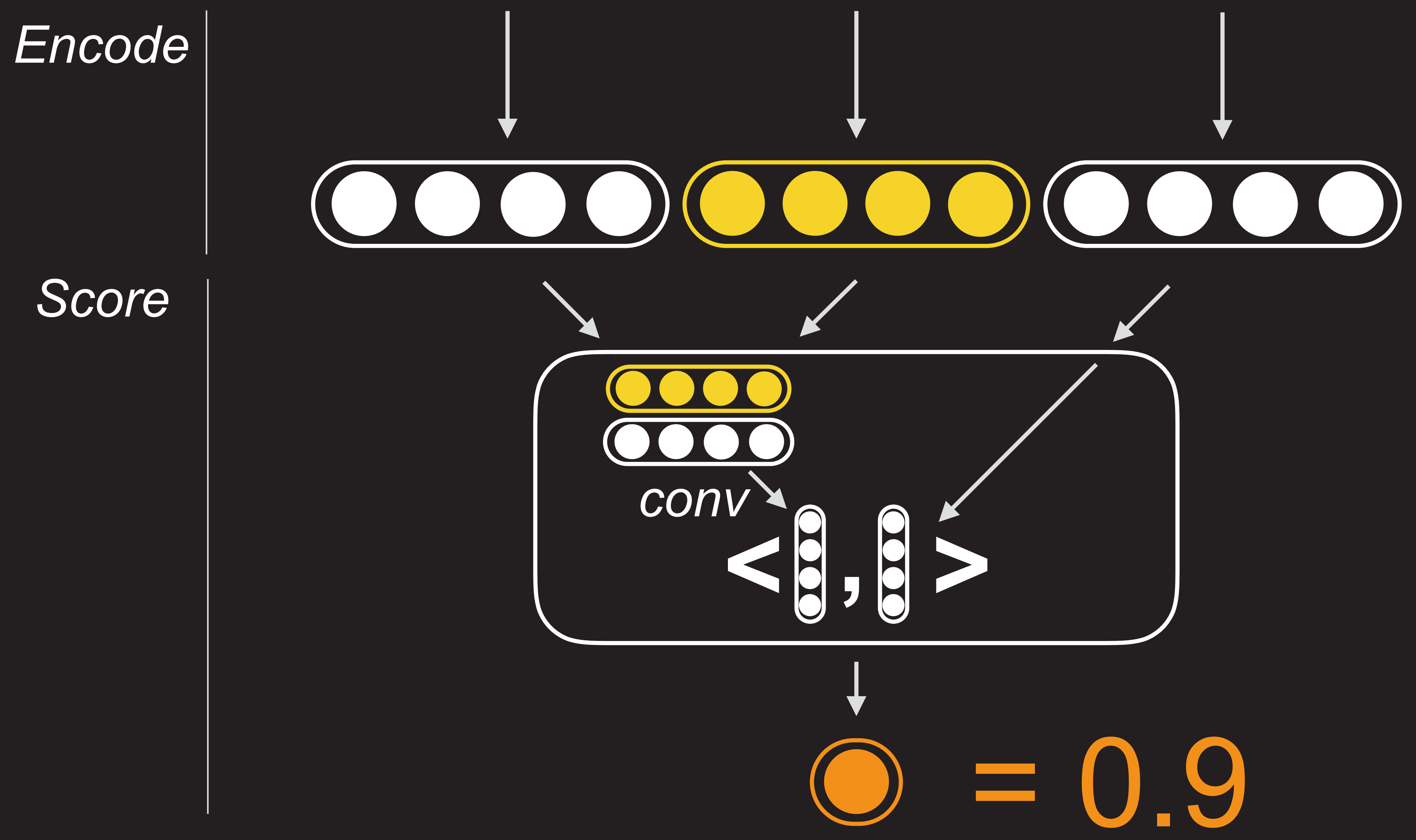
ERMLP (Dong et al. 2014)

S(Enrique canSpeak Spanish) = 0.9



ConvE (Dettmers et al. 2017)

s(Enrique canSpeak Spanish) = 0.9



Training Objective

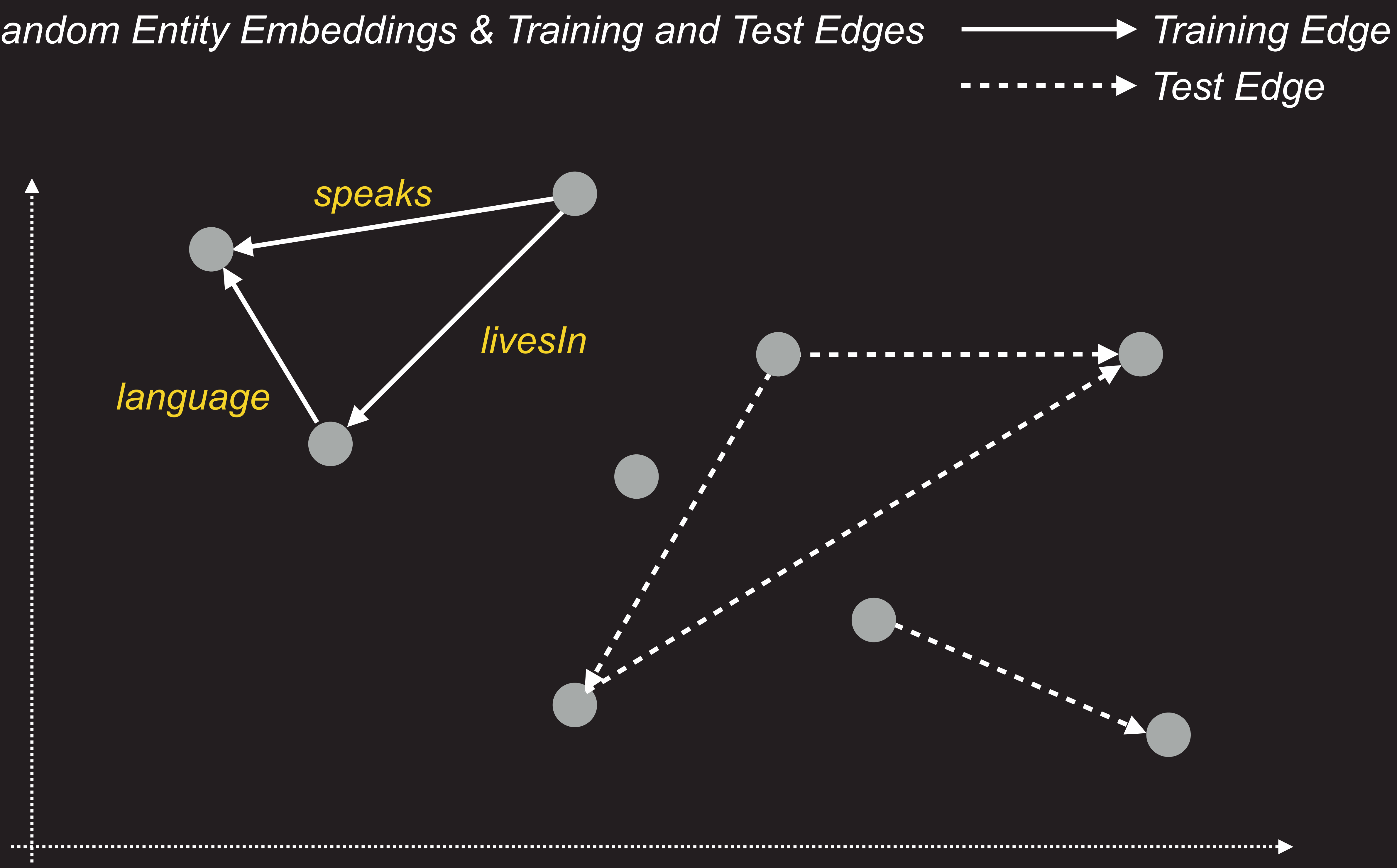
- Terms in **training objective for observed facts**

$$\sum_{\substack{\text{subject, predicate, object}\\ \text{in training data}}} \overset{\text{high}}{\downarrow}\!\!\!\!\overline{\text{loss}}\;\Big(\, \underset{\substack{\uparrow\\ \text{low}}}{\underline{s\,(\,\text{subject, predicate, object}\,)}}\,\Big)$$

- Plus some term with **negative sampling**
- Optimised by Stochastic Gradient Descent

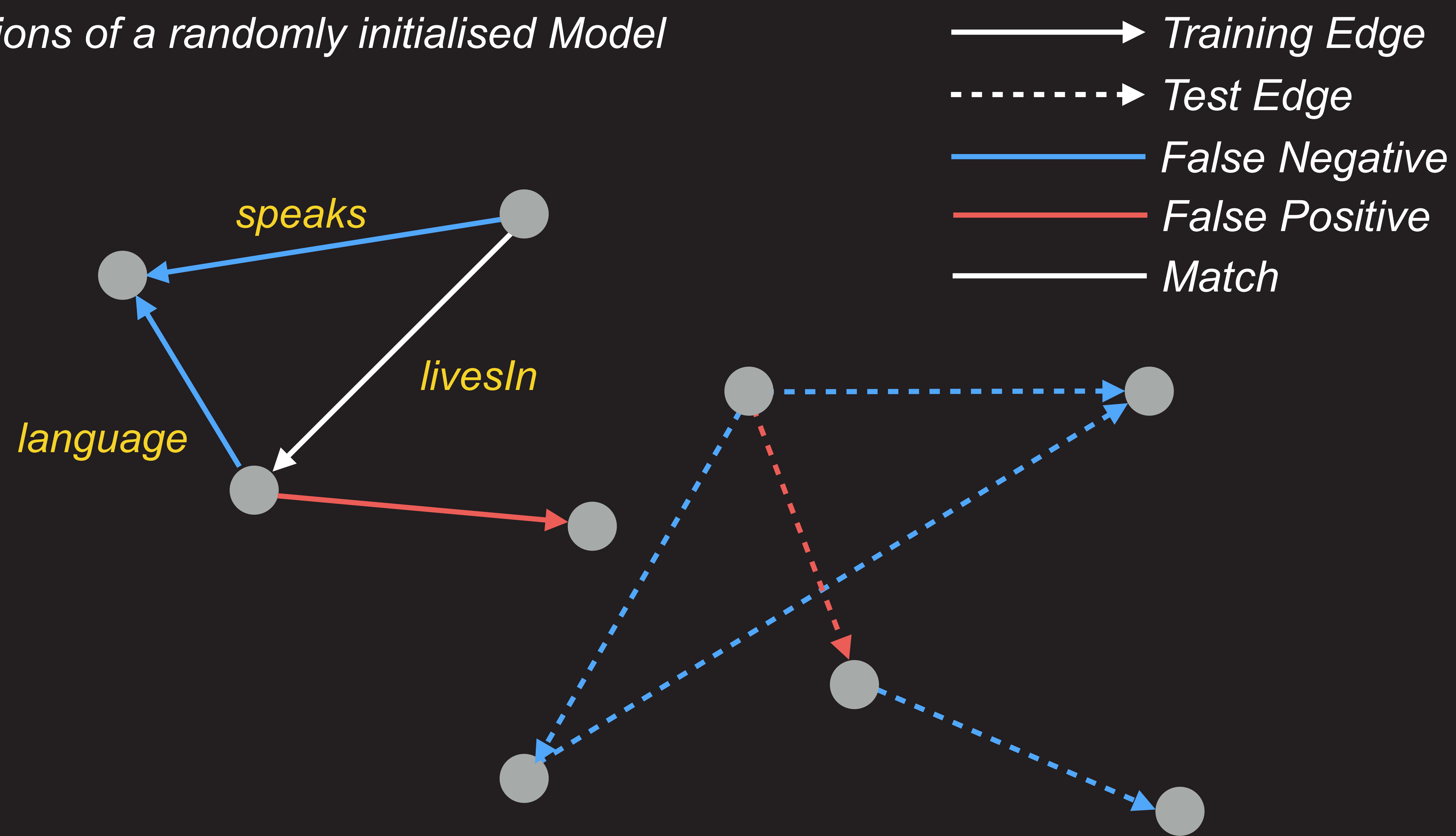
Learning Visualised

Random Entity Embeddings & Training and Test Edges



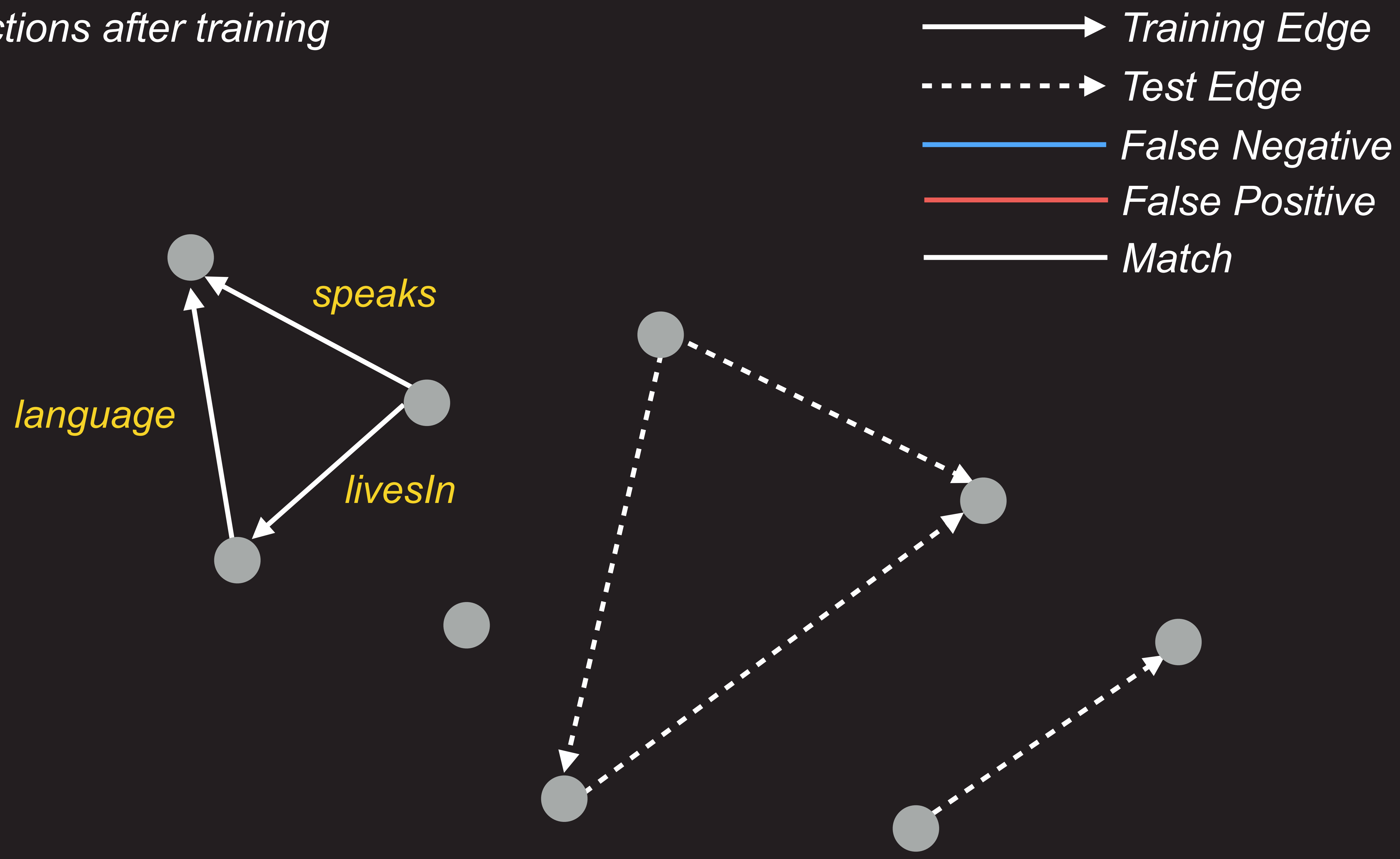
Learning Visualised

Predictions of a randomly initialised Model



Learning Visualised

Predictions after training



Part 2: Injecting Prior Knowledge

$$s(\text{Enrique can speak Spanish}) = 0.9$$

Enrique works in Mexico City

Enrique lives in Mexico

Luis can speak Spanish

Luis lives in Mexico

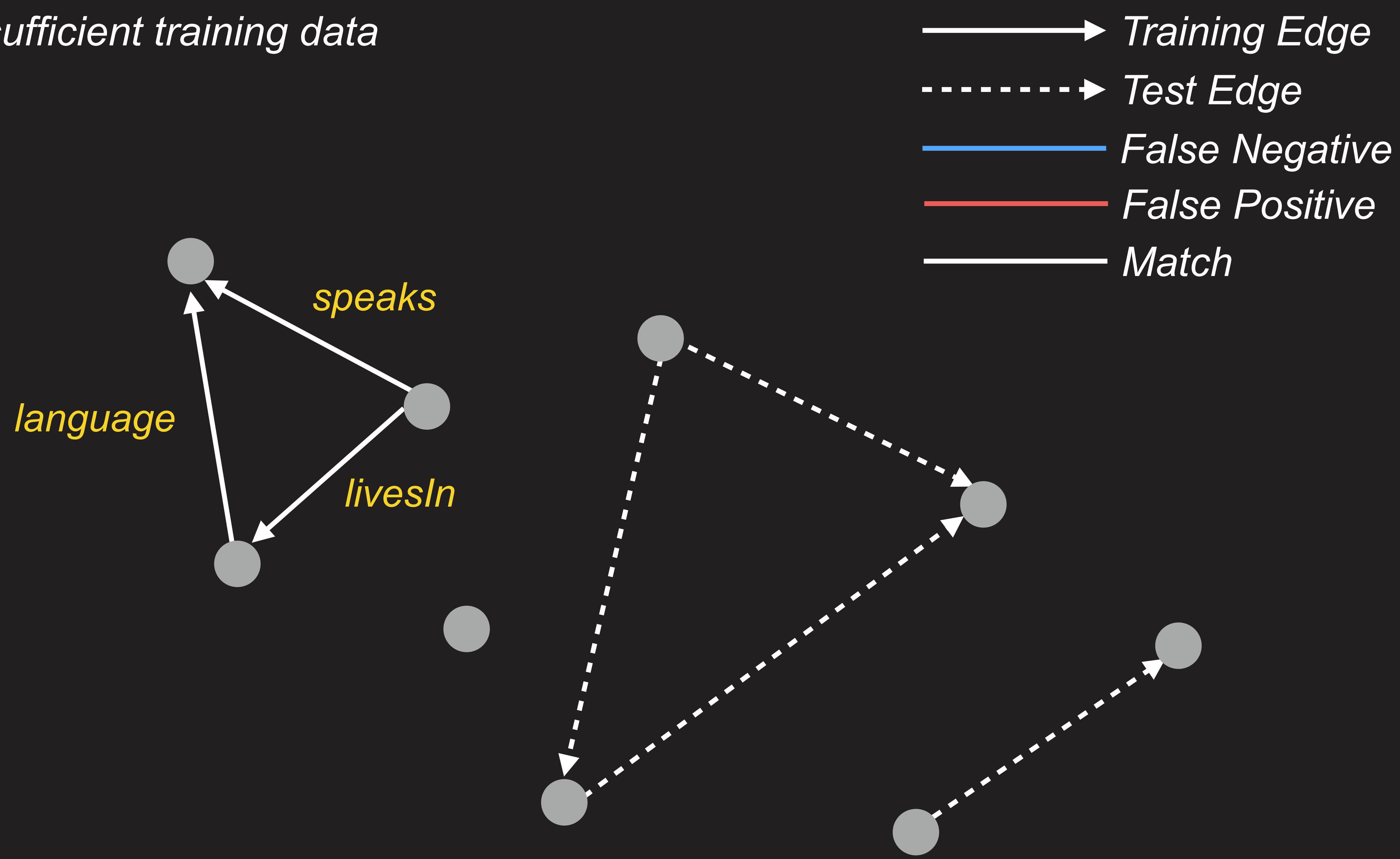
livesIn(Ivan, Mexico)

speaks(Ivan, Spanish)

When someone lives in Mexico, they will likely speak Spanish

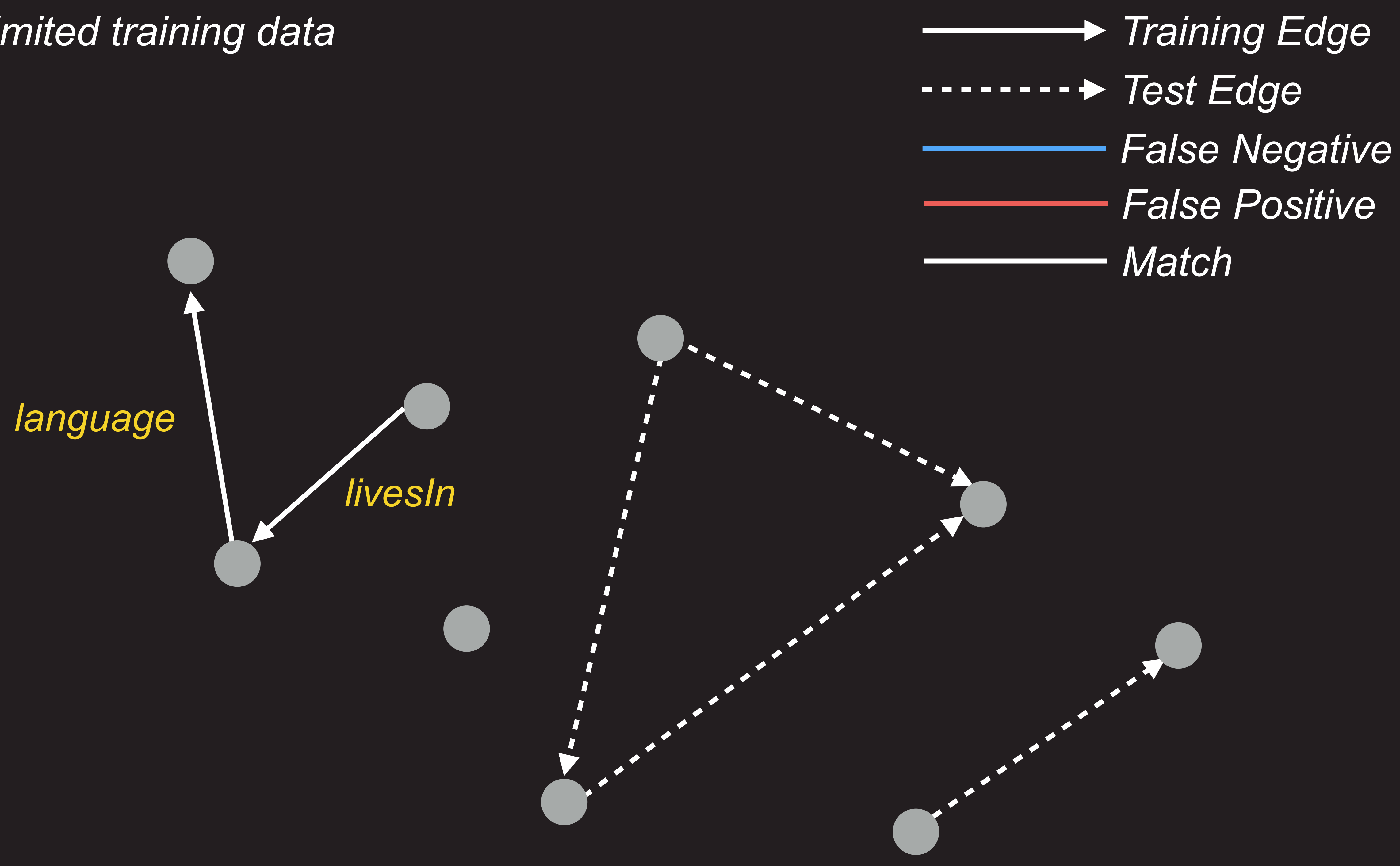
Learning Visualised

With sufficient training data



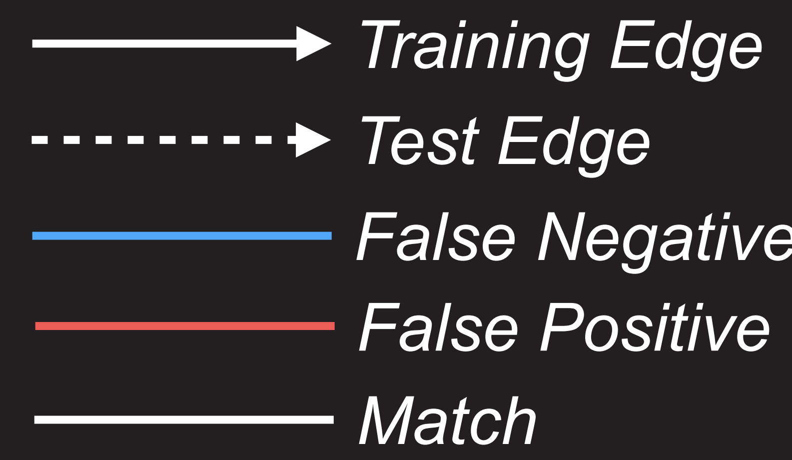
Learning Visualised

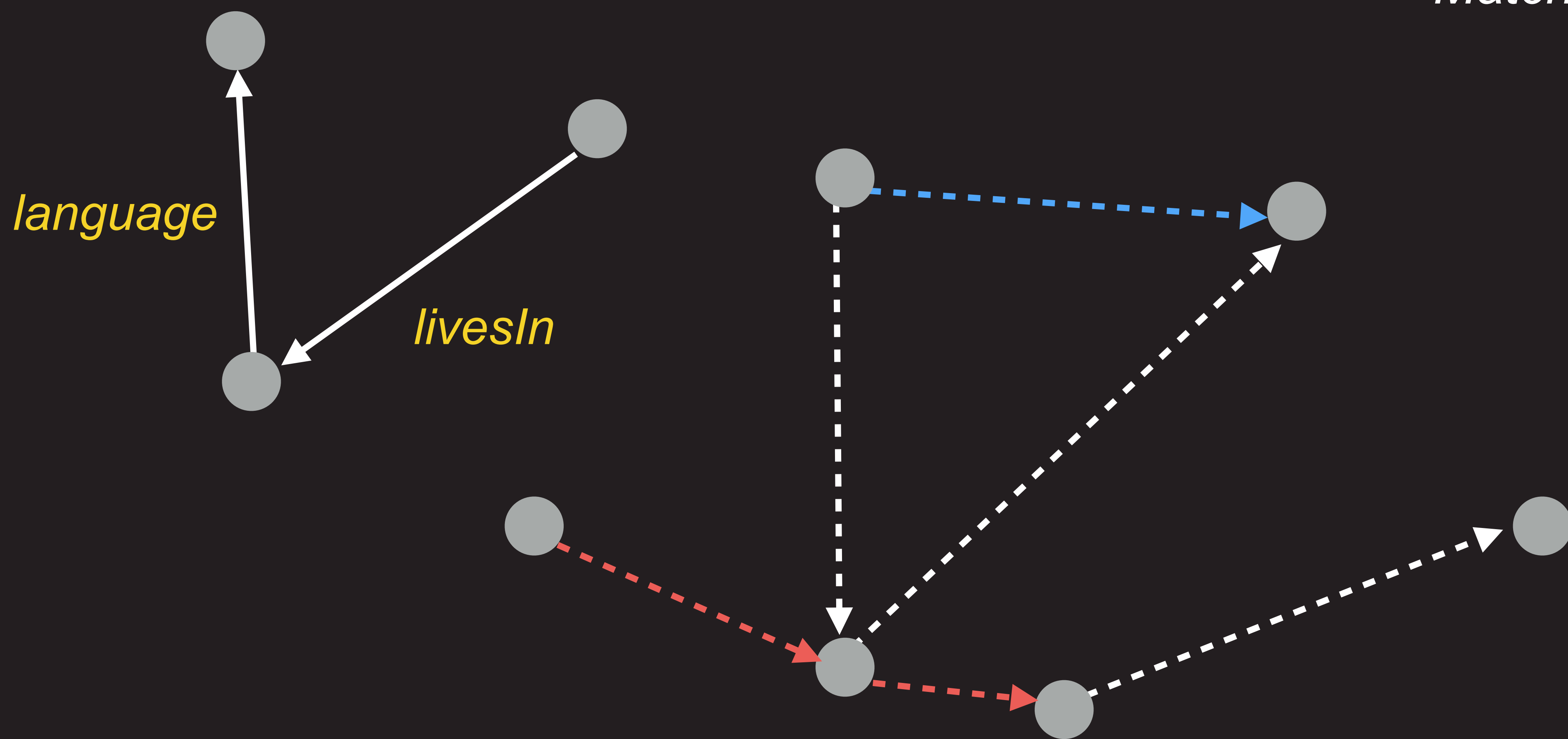
With limited training data



Limited training Data

(Obviously) Leads to errors

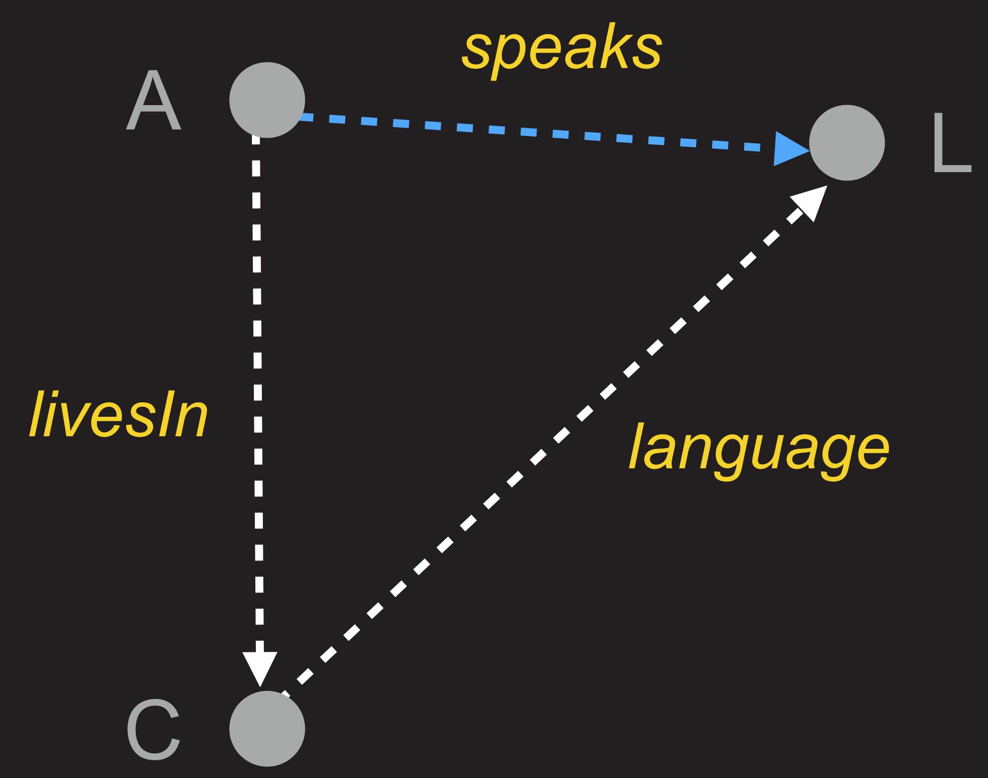




Errors

May violate our ***Prior Knowledge***

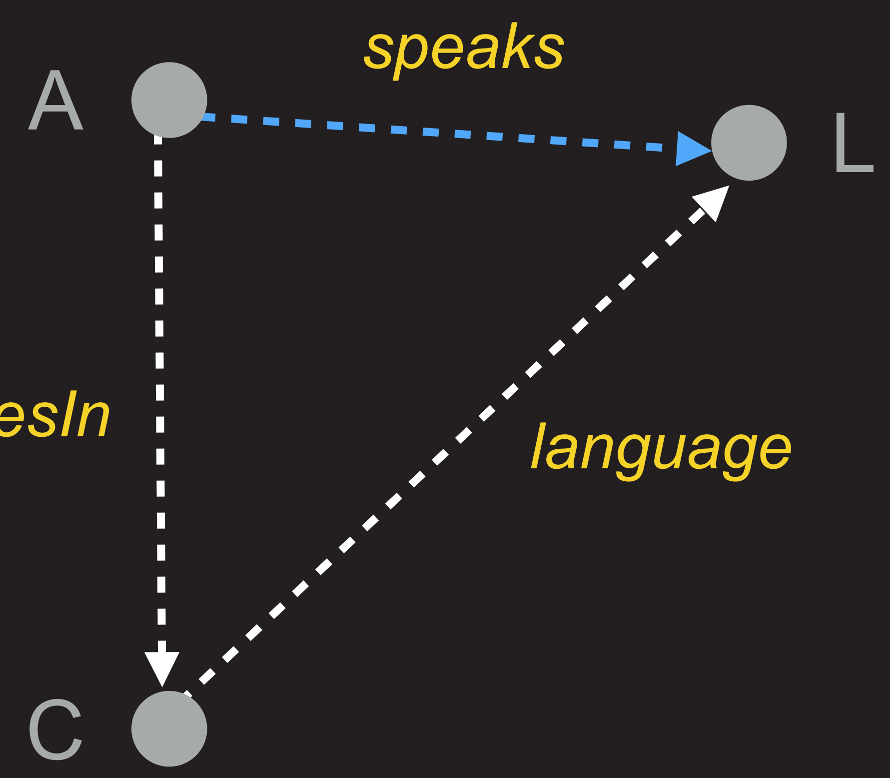
When A lives in C and C's language is L then A likely speaks L



Solution

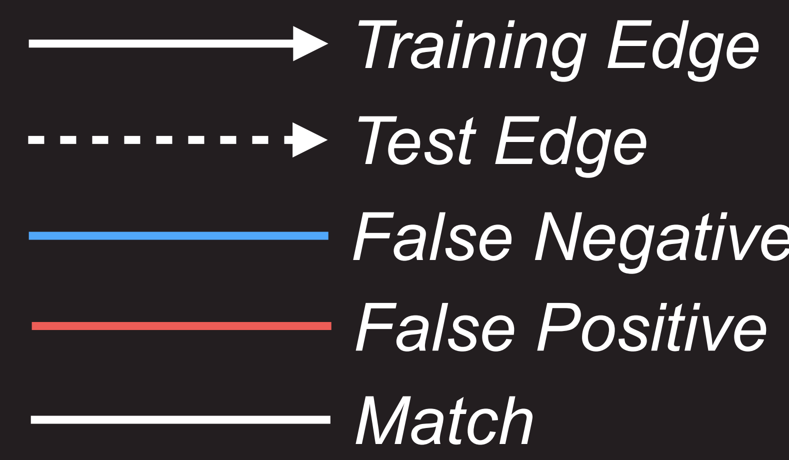
Add a loss term that punishes this specific violation

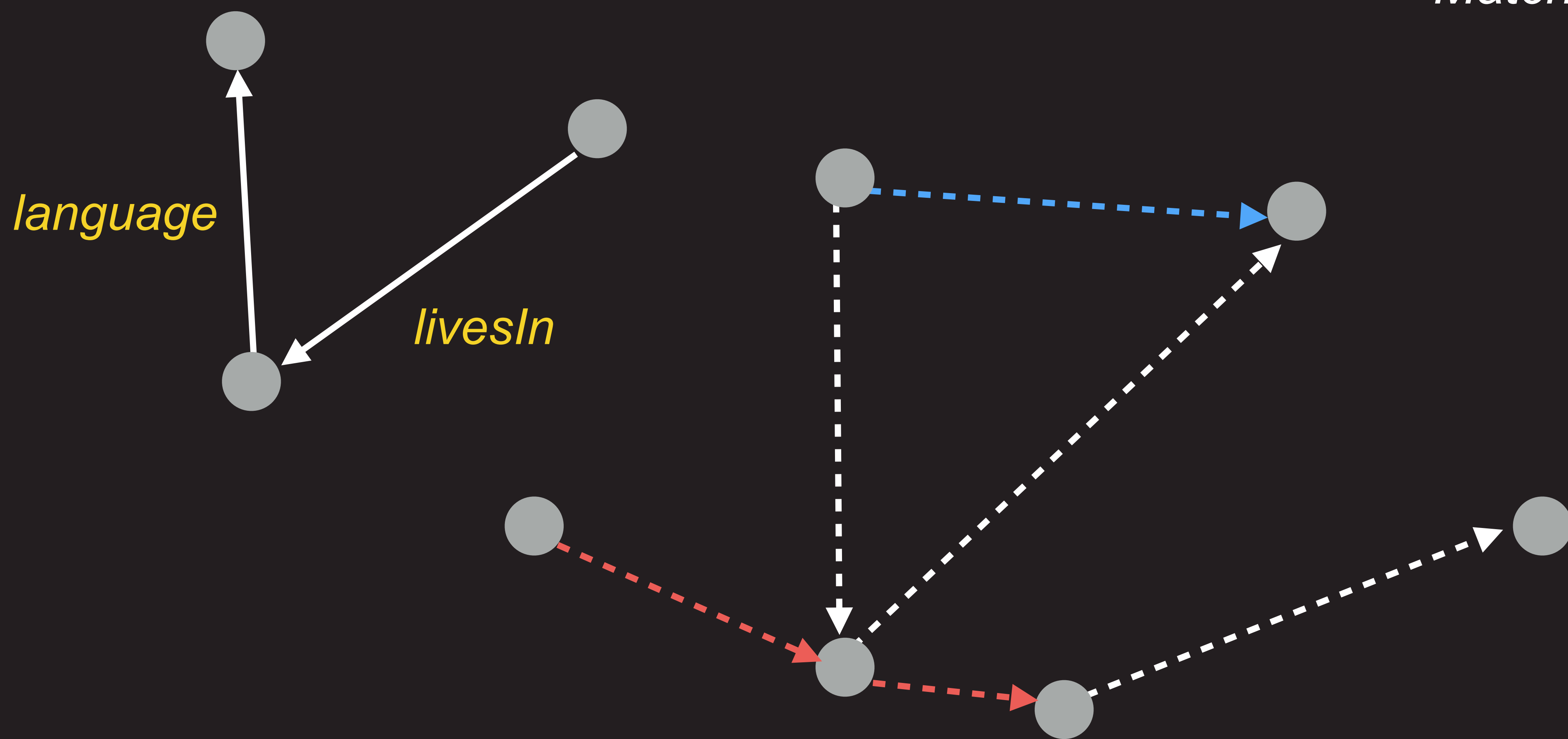
$$\ldots + \text{loss}(v_A, v_B, v_C, v_{lives}, v_{speaks}, v_{lang})$$



Re-training

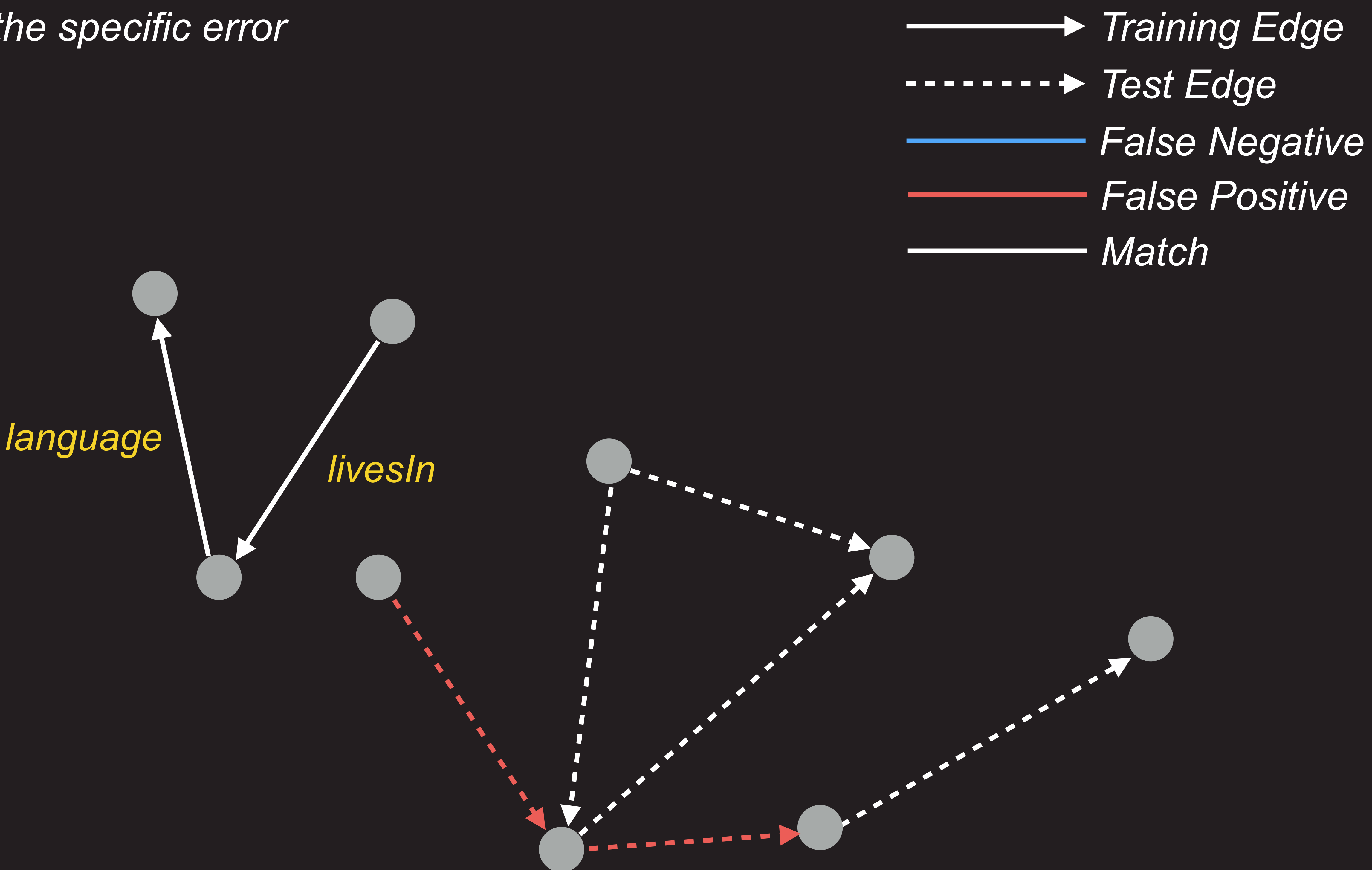
with the specific loss term added





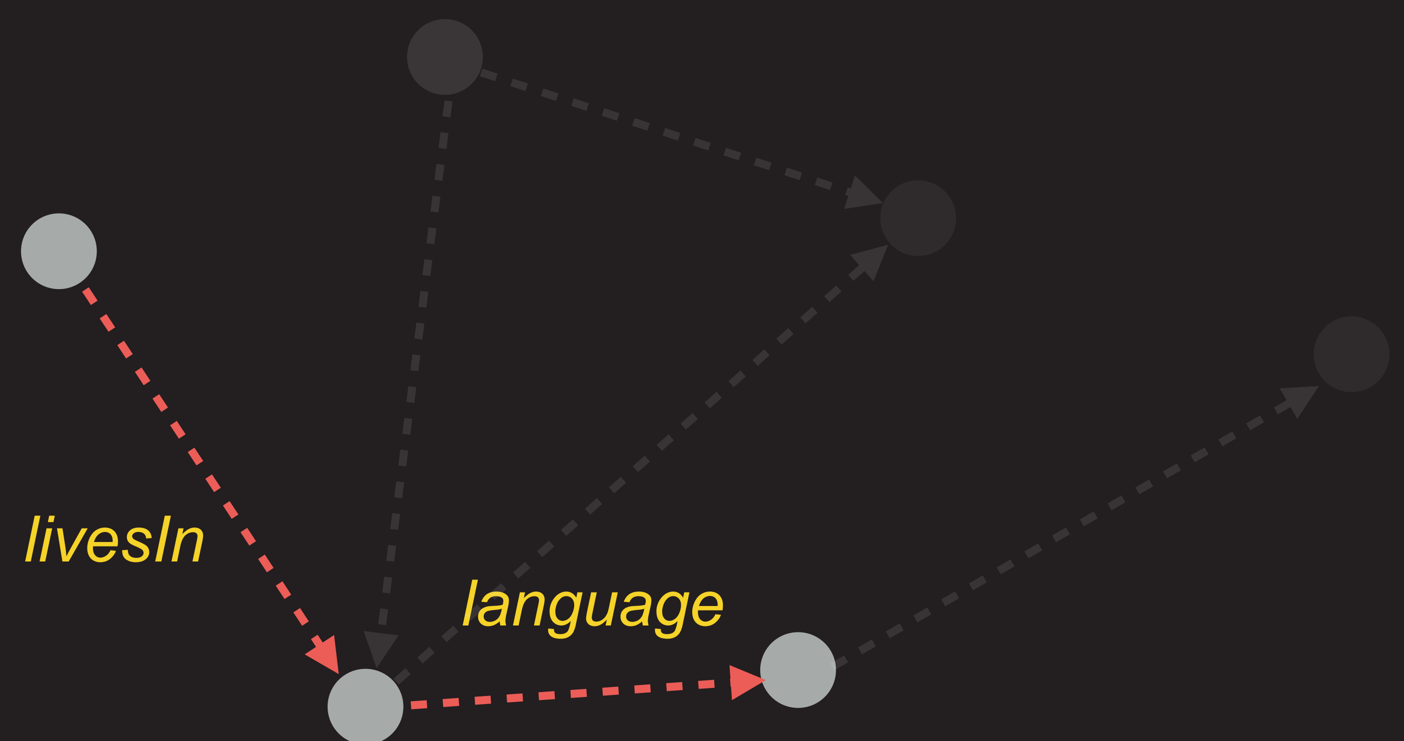
Re-training

fixes the specific error



Re-training

but not all

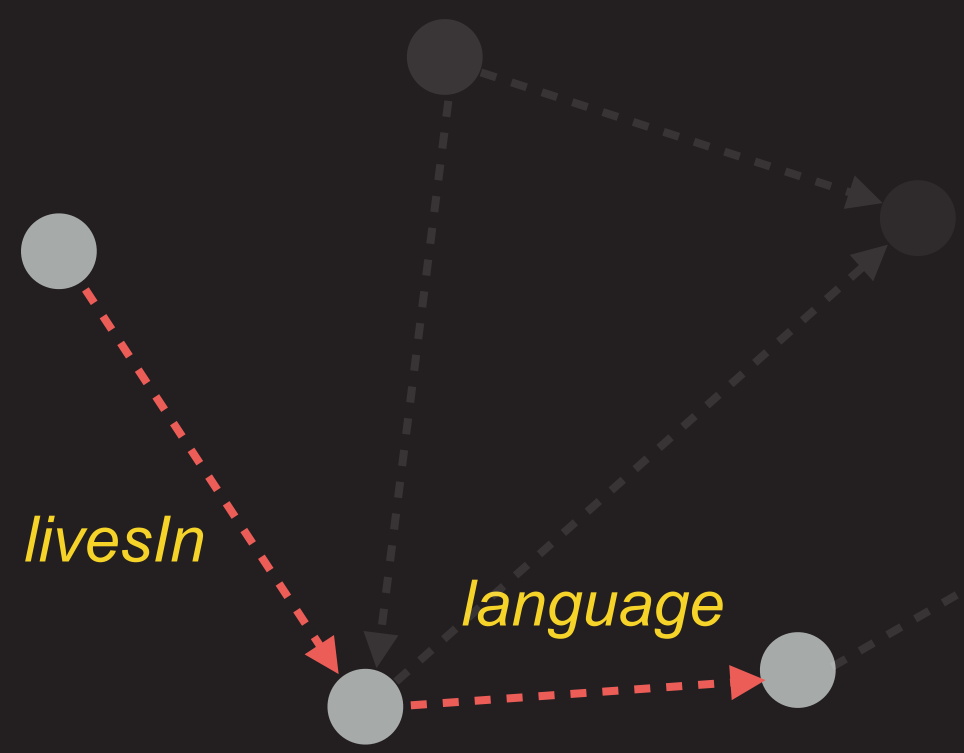


Loss

*So minimise **the loss of the worst case***

$$\ldots + \boxed{\underset{A,B,C}{\mathrm{Max}}}\ \mathrm{loss}(v_A, v_B, v_C, v_{lives}, v_{speaks}, v_{lang})$$

expensive

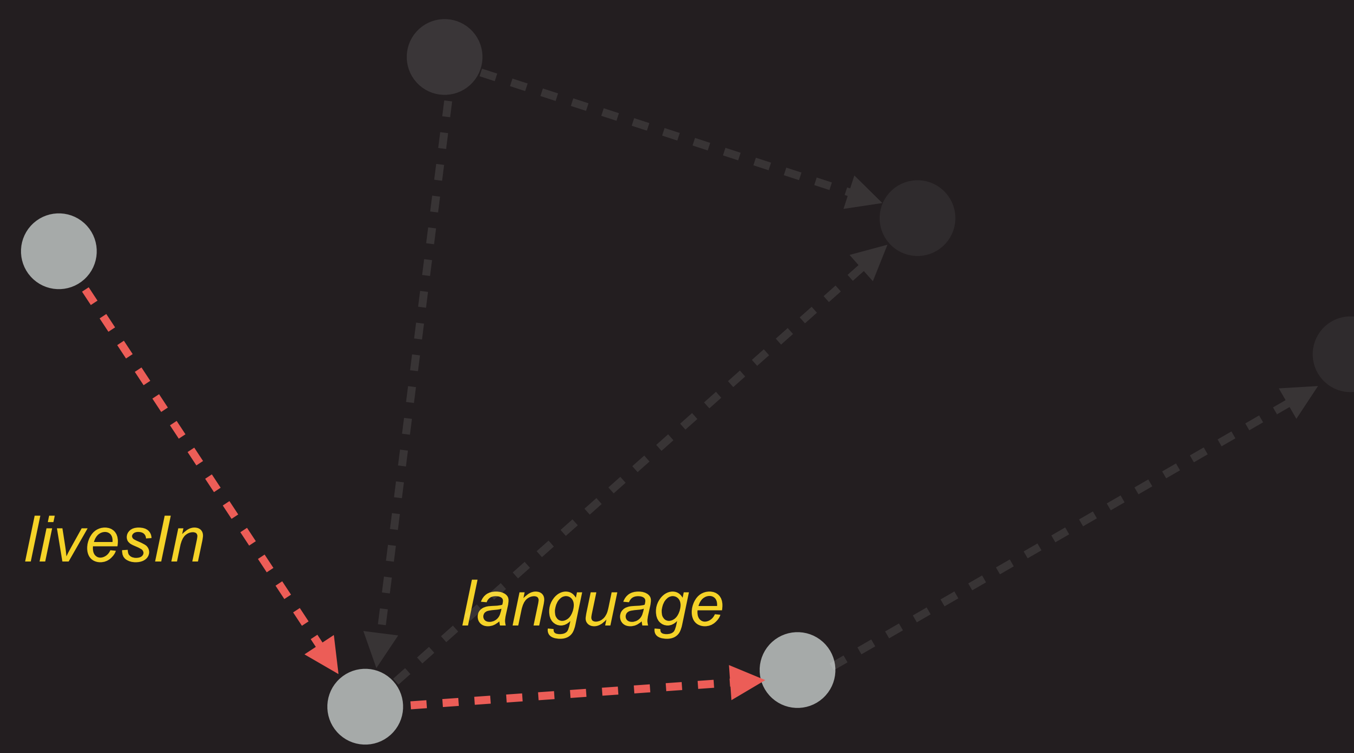


Loss

*Avoid combinatorial optimisation by maximising over **embeddings***

$$\ldots + \boxed{\max_{v_A, v_B, v_C}} \text{loss}(v_A, v_B, v_C, v_{lives}, v_{speaks}, v_{lang})$$

cheaper



Full Loss

- Terms in **training objective for observed facts**

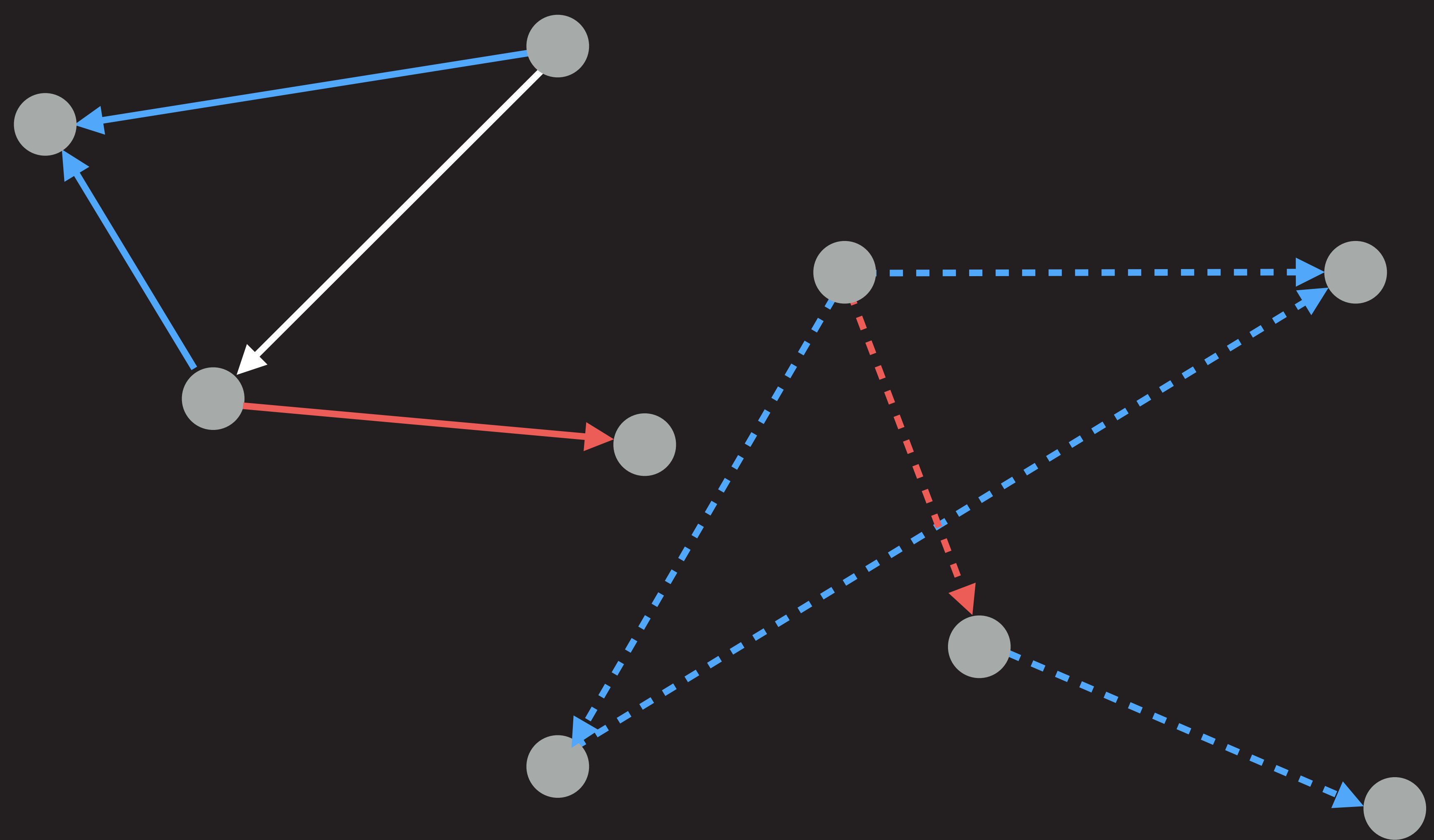
$$\sum_{\substack{\text{subject, predicate, object} \\ \text{in training data}}} \text{loss}\,(\,s\,(\,\text{subject, predicate, object}\,)\,)$$

$$+ \max_{v_A,\, v_B,\, v_C} \text{loss}(v_A,\ v_B,\ v_C,\ v_{lives},\ v_{speaks},\ v_{lang}\ )$$

- Minimised in **two player game**

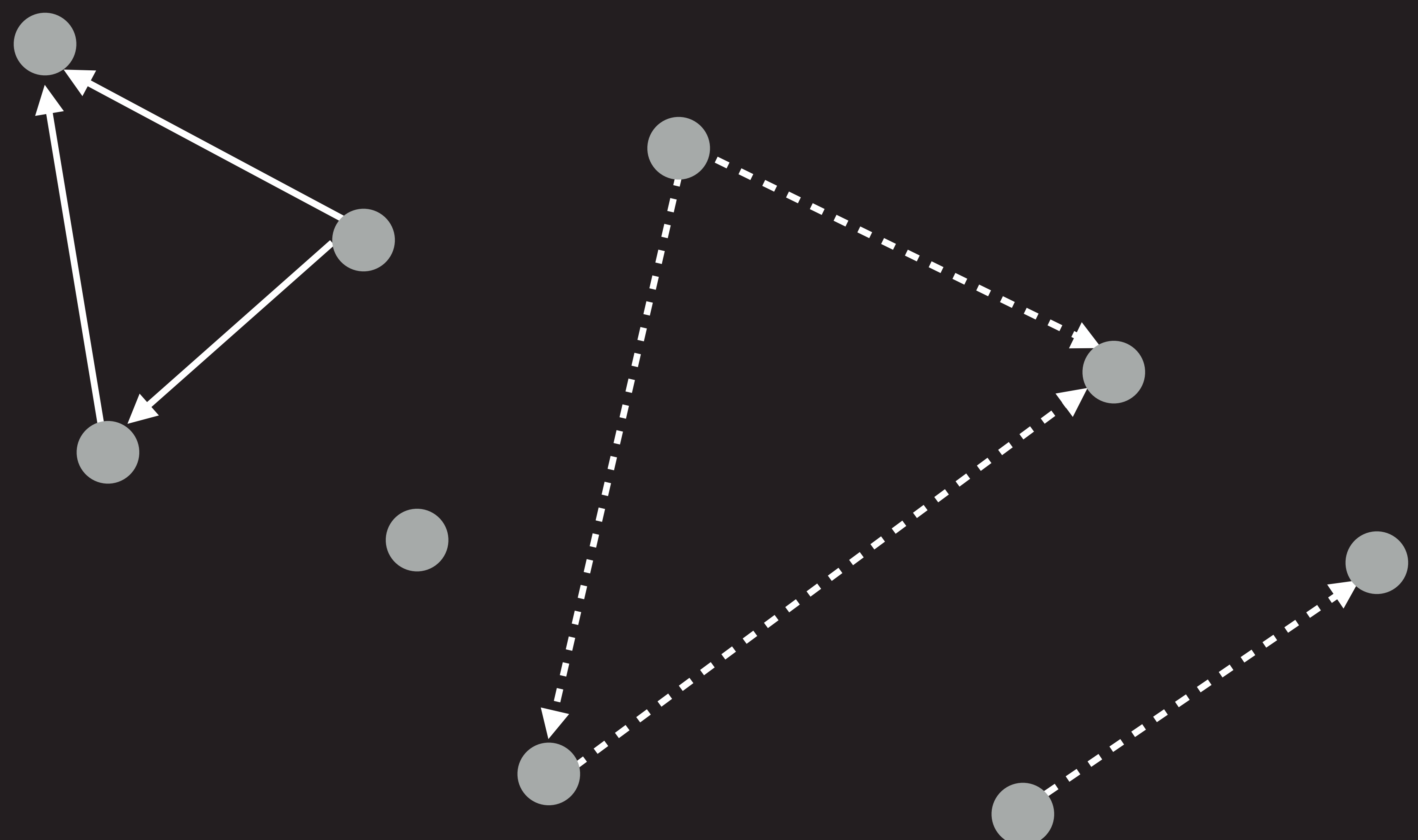
Player 1: Link Predictor

*Estimates entity and relation embeddings to **fix the graph** (gradient **descent**)*



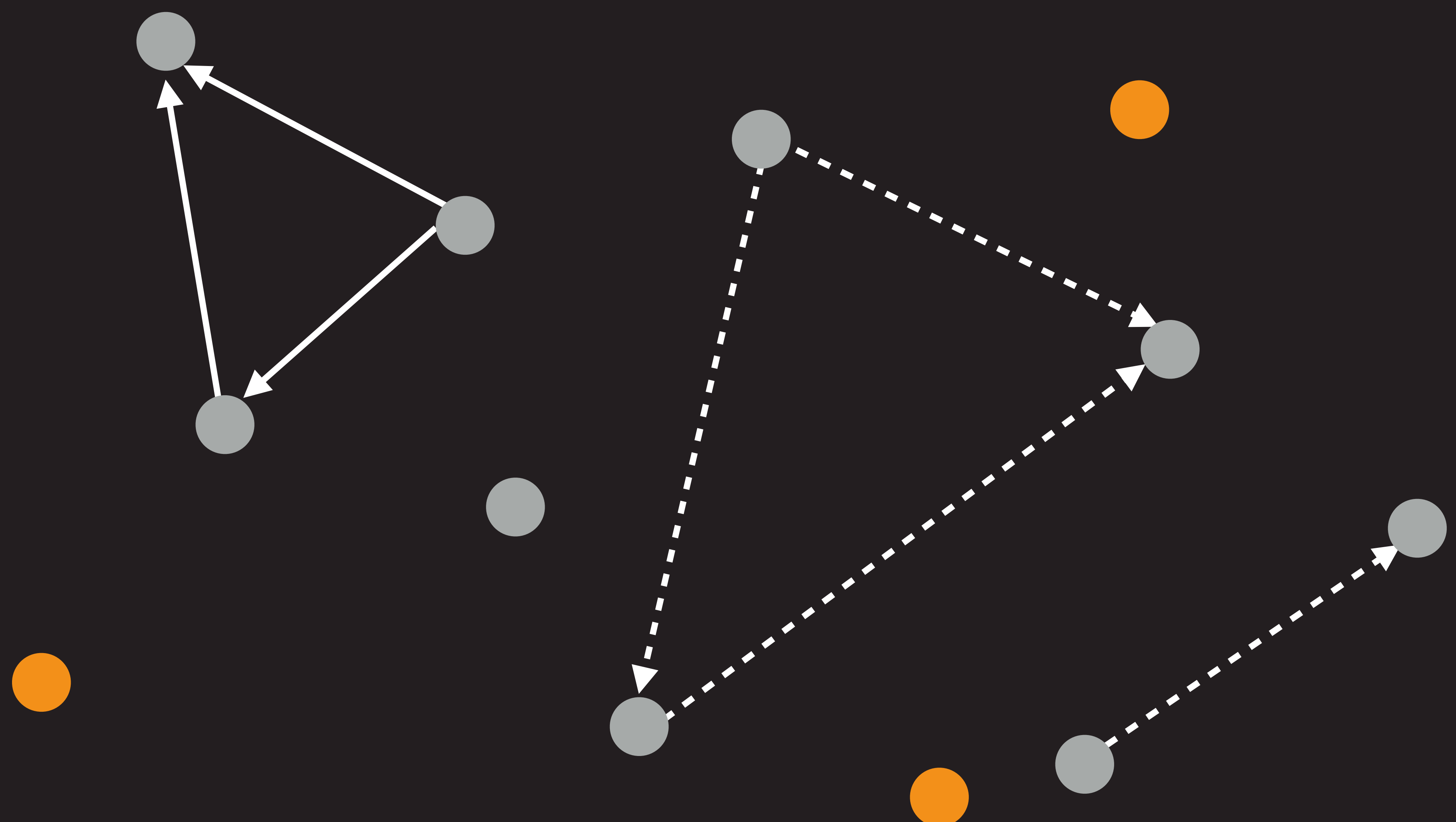
Player 1

Predictions after training



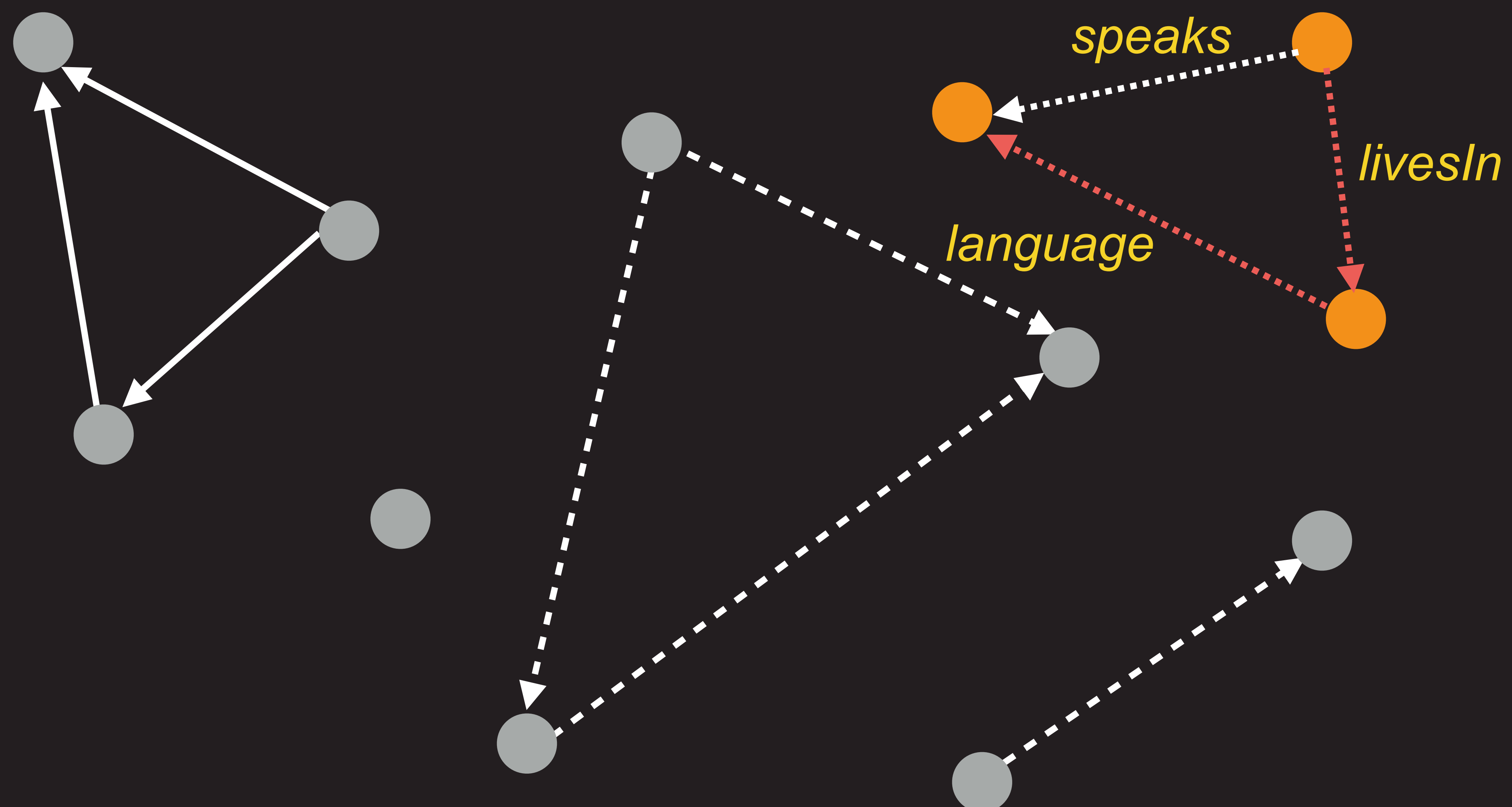
Player 2: Adversary

Synthesises entity embeddings that ***break the rules*** *(gradient* ***ascent****)*



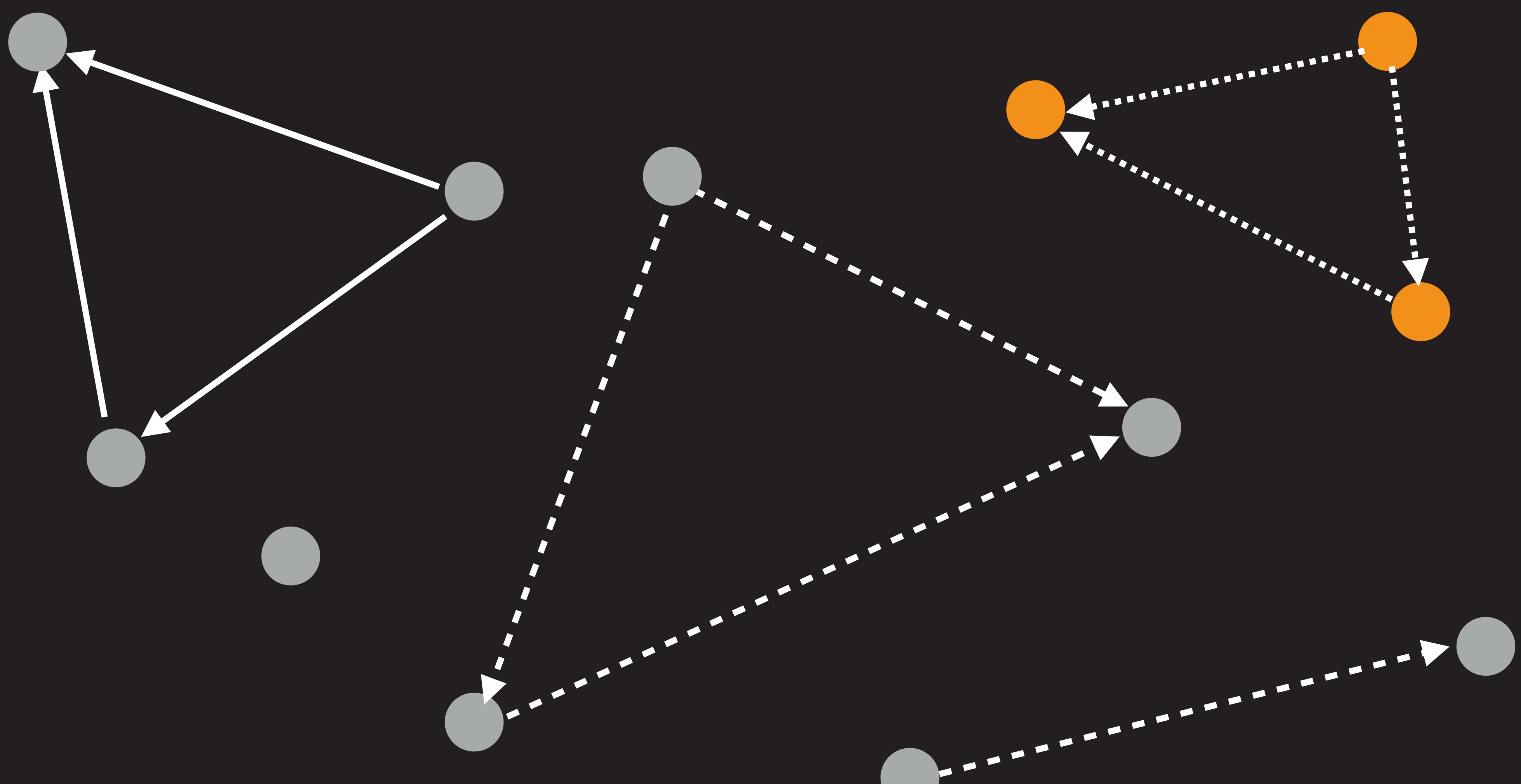
Player 2: Adversary

*Synthesises entity embeddings that **break the rules** (gradient **ascent**)*



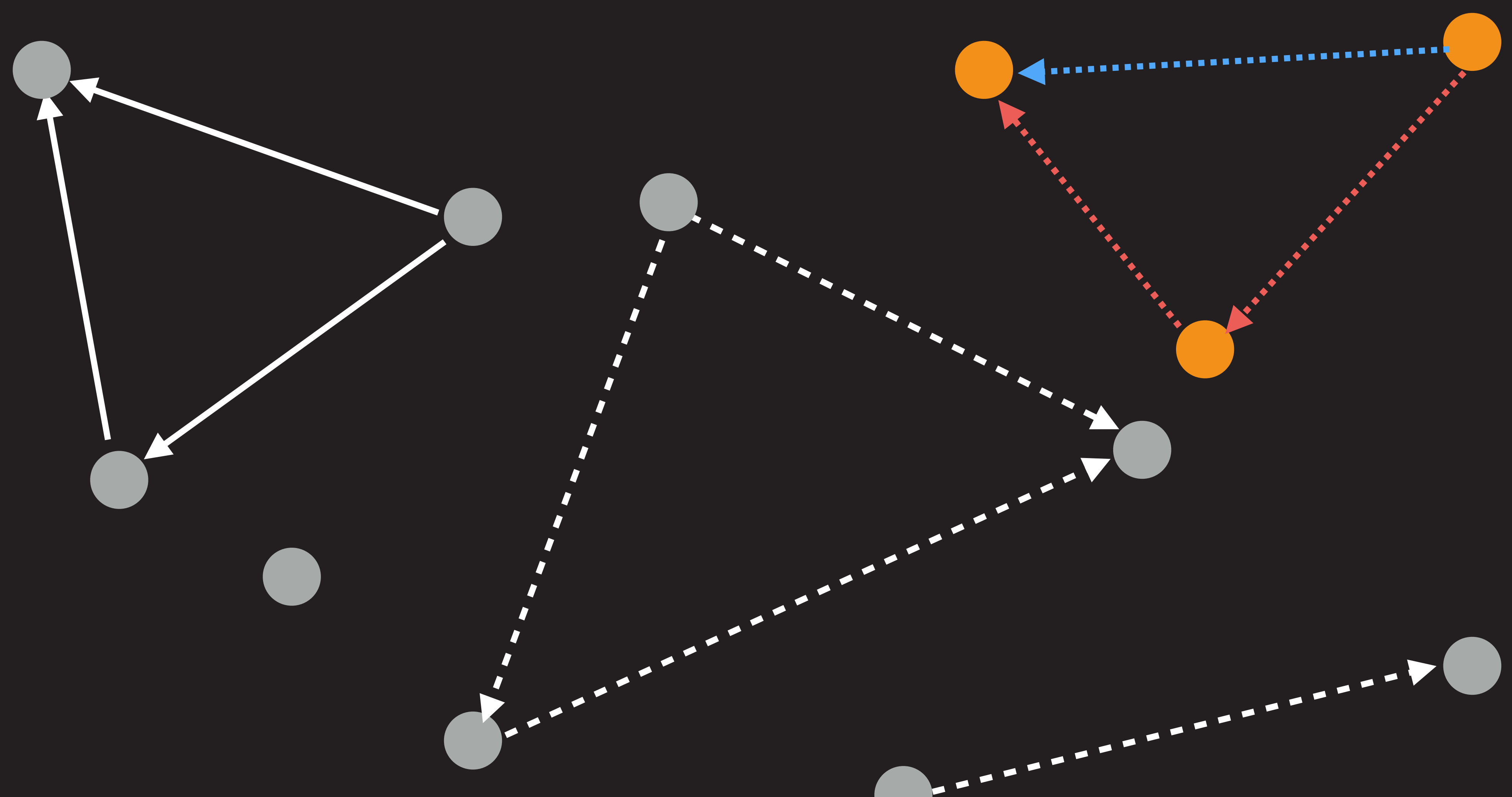
Player 1: Link Predictor

Estimates entity and relation embeddings to ***fix the graph and violations***



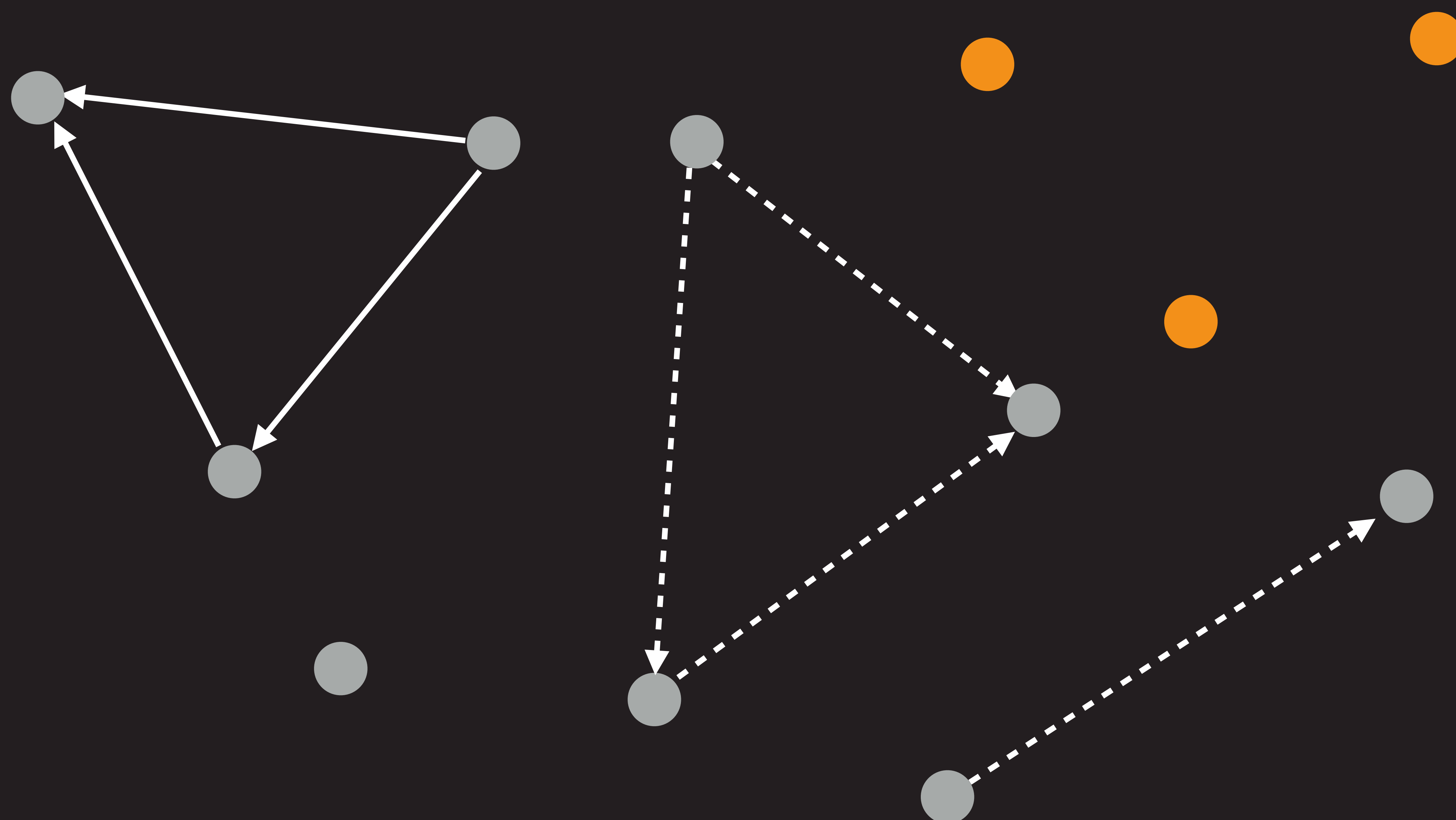
Player 2: Adversary

Estimates entity and relation embeddings to ***fix the graph and violations***



Player 1: Link Predictor

Estimates entity and relation embeddings to ***fix the graph and violations***



Closed Form Solutions

For some rules and models the max expression has a closed form solution

For DistMult, Complex, TransE:

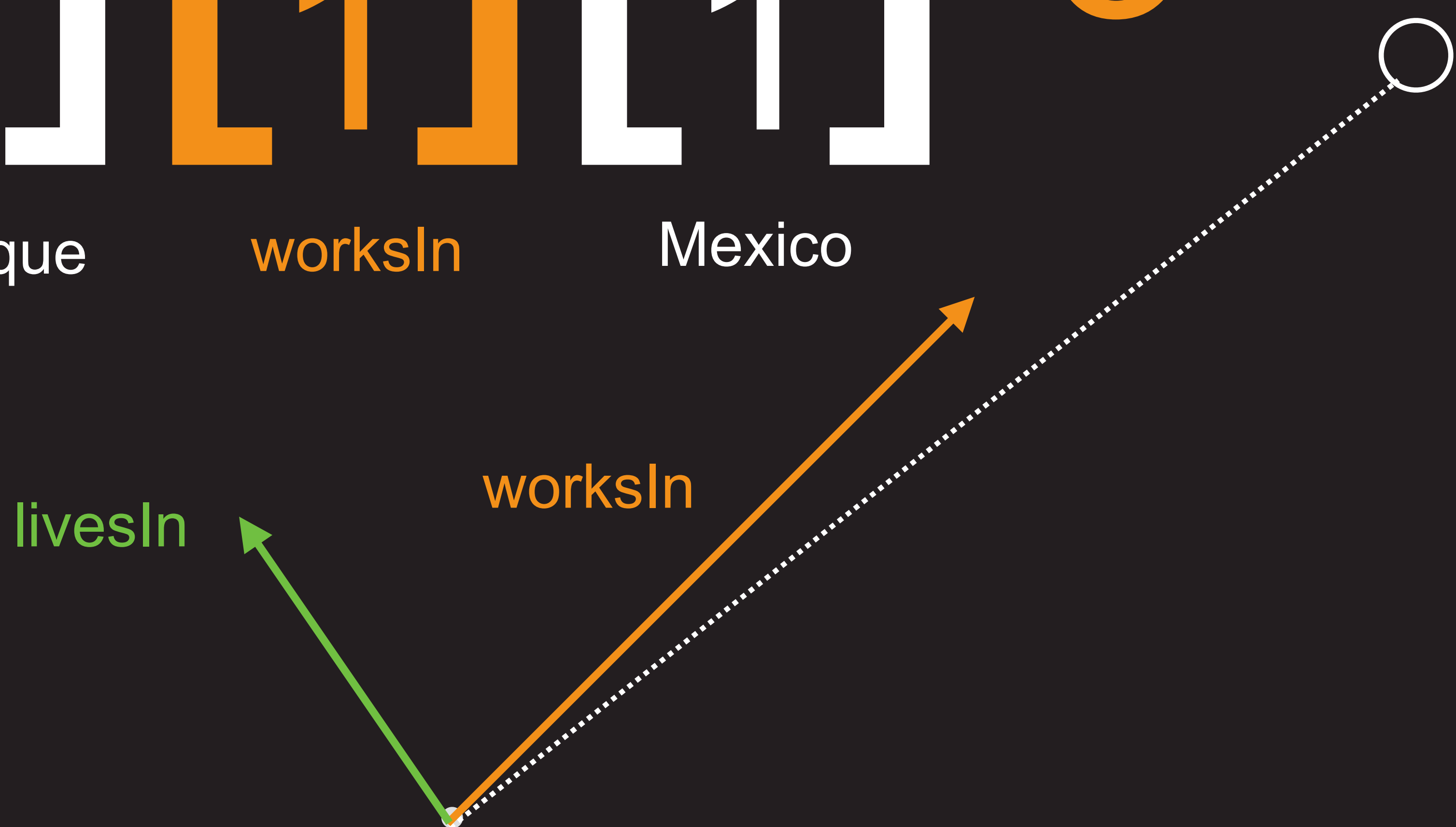
When A relation1 B then A relation2 B

$$\max_{v_A, v_B} \text{loss}(v_A, v_B, v_{rel1}, v_{rel2})$$

$$= \text{loss}_{\text{closed-form}}(v_{rel1}, v_{rel2})$$

Entailment in Vector Space

$$\underbrace{\begin{bmatrix}1\\1\end{bmatrix}}_{\text{Enrique}} * \underbrace{\begin{bmatrix}1\\1\end{bmatrix}}_{\text{worksIn}} * \underbrace{\begin{bmatrix}2\\1\end{bmatrix}}_{\text{Mexico}} = 3$$

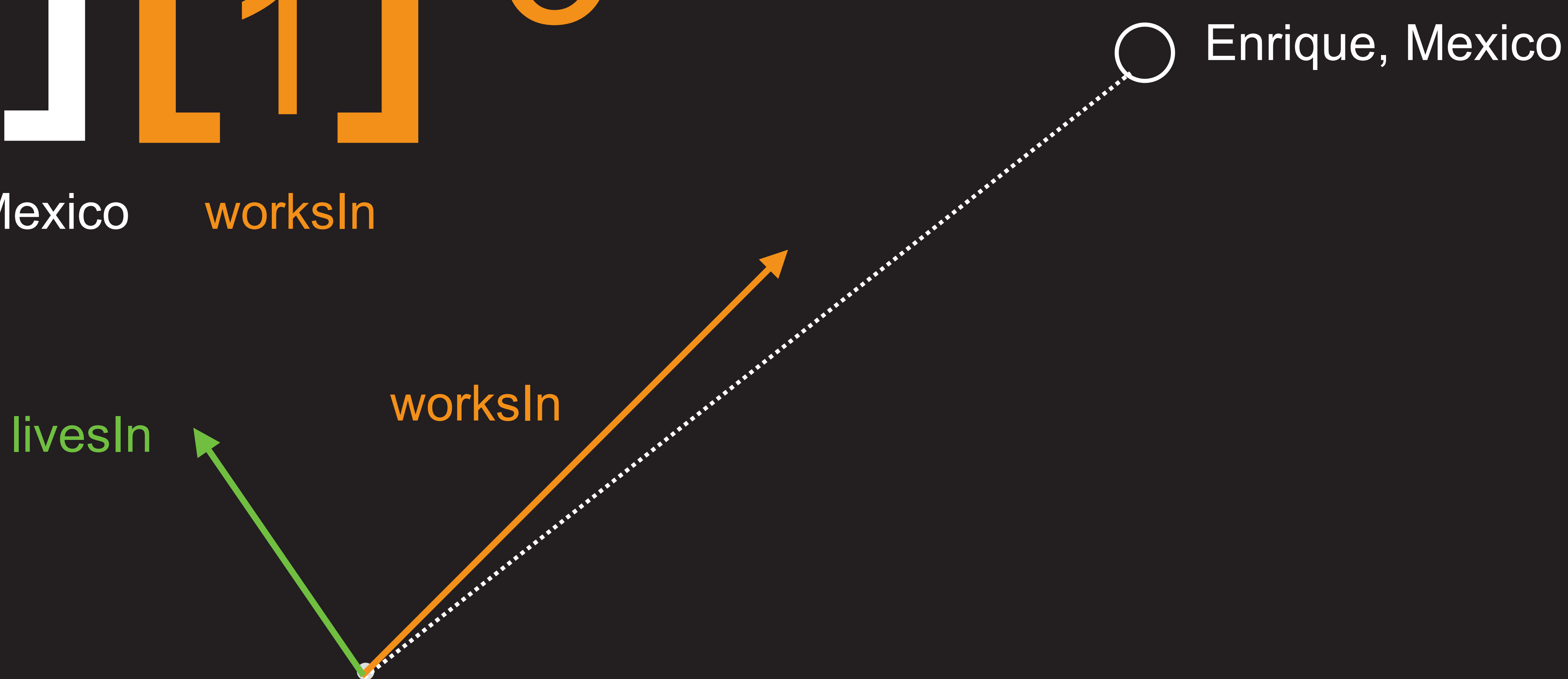


- Where to put livesIn vector to be implied by worksIn

Entailment in Vector Space

$$\begin{bmatrix}2\\1\end{bmatrix} * \begin{bmatrix}1\\1\end{bmatrix} = 3$$

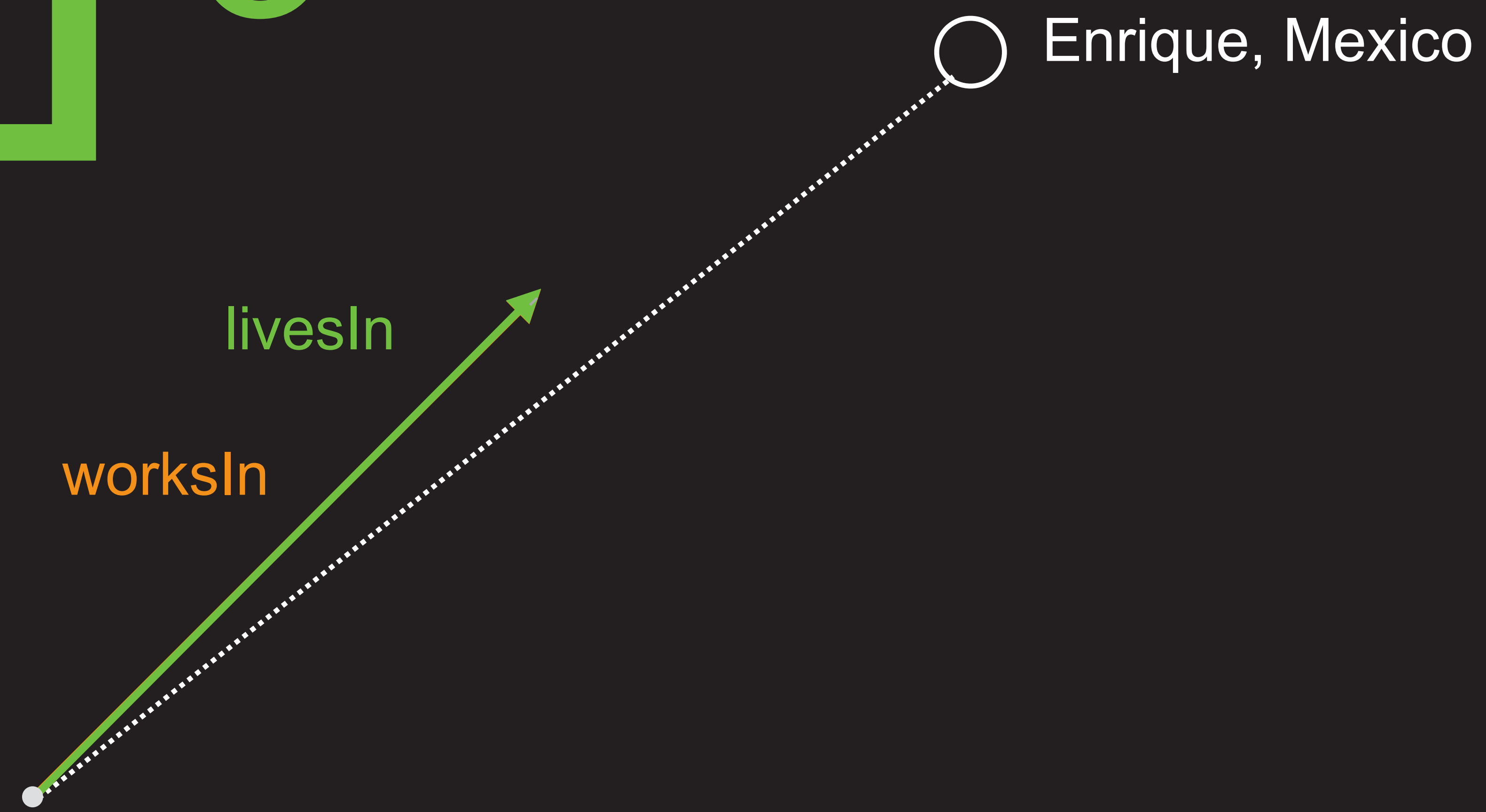
Enrique, Mexico worksIn



- Entailment: livesIn score is at least as high as worksIn

Entailment in Vector Space

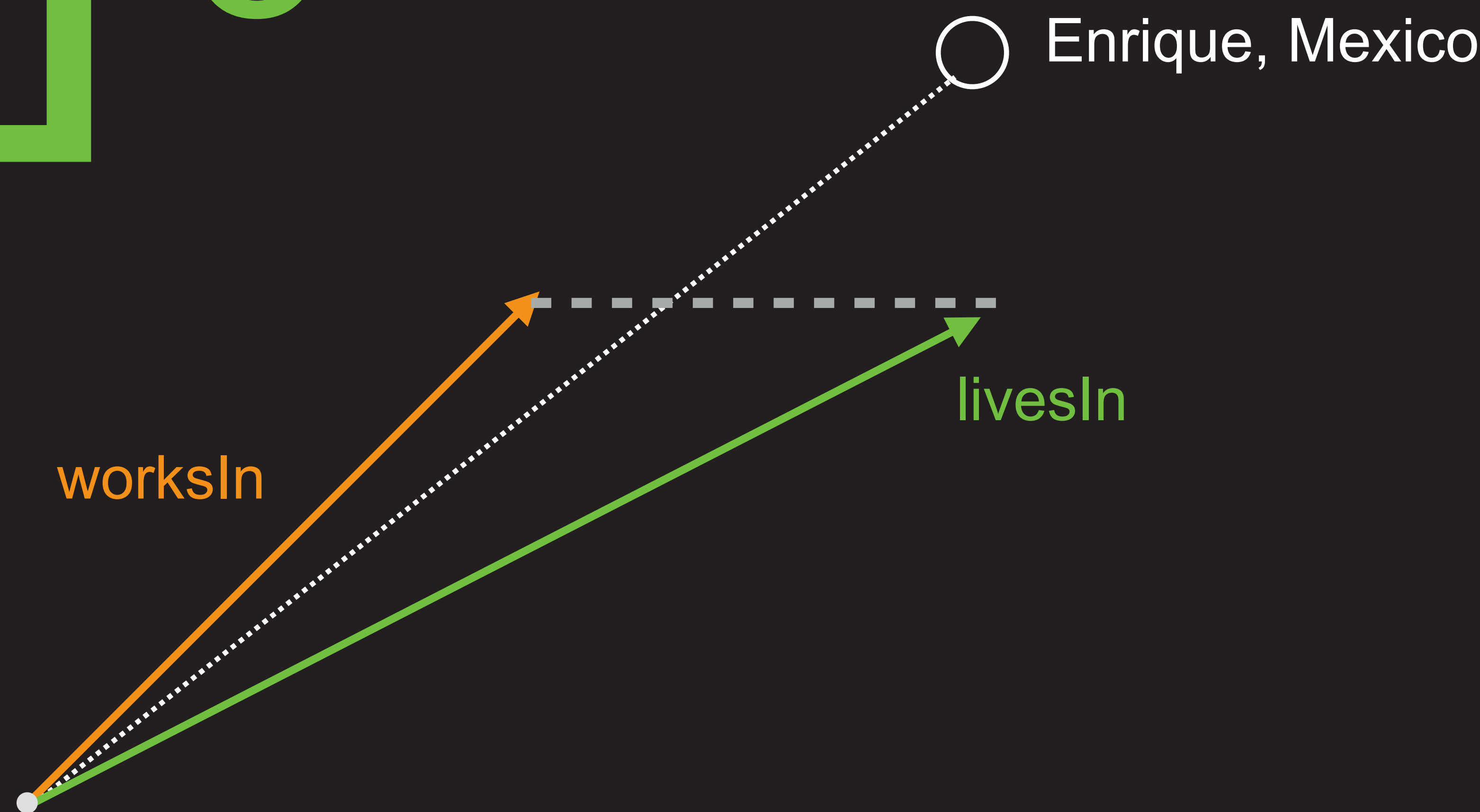
$$\begin{bmatrix}2\\1\end{bmatrix} * \begin{bmatrix}1\\1\end{bmatrix} = \begin{matrix}3\\3\end{matrix}$$



- When it's the same vector

Entailment in Vector Space

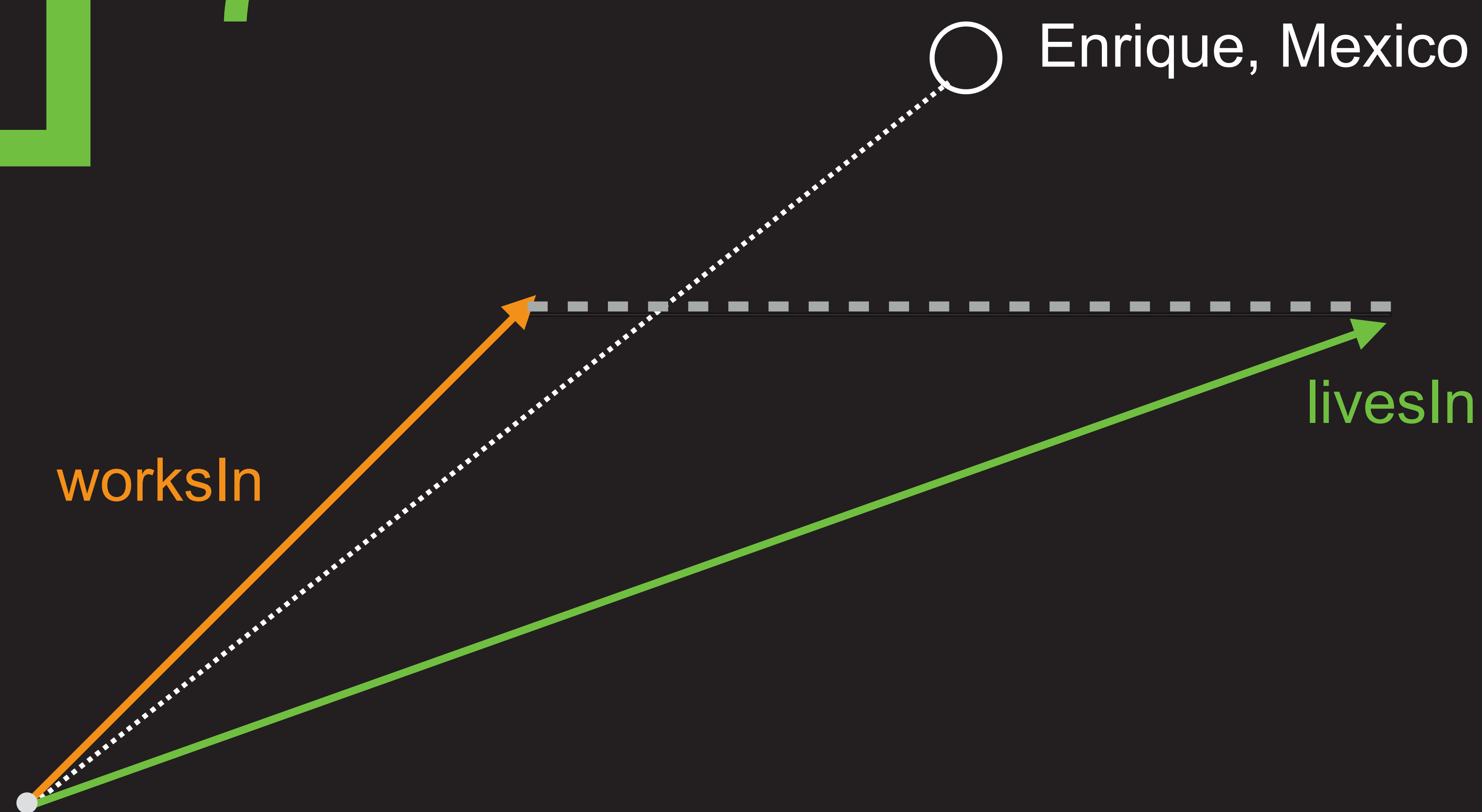
$$\begin{bmatrix}2\\1\end{bmatrix} * \begin{bmatrix}2\\1\end{bmatrix} = \begin{matrix}3\\5\end{matrix}$$



- When the **first component is larger**

Entailment in Vector Space

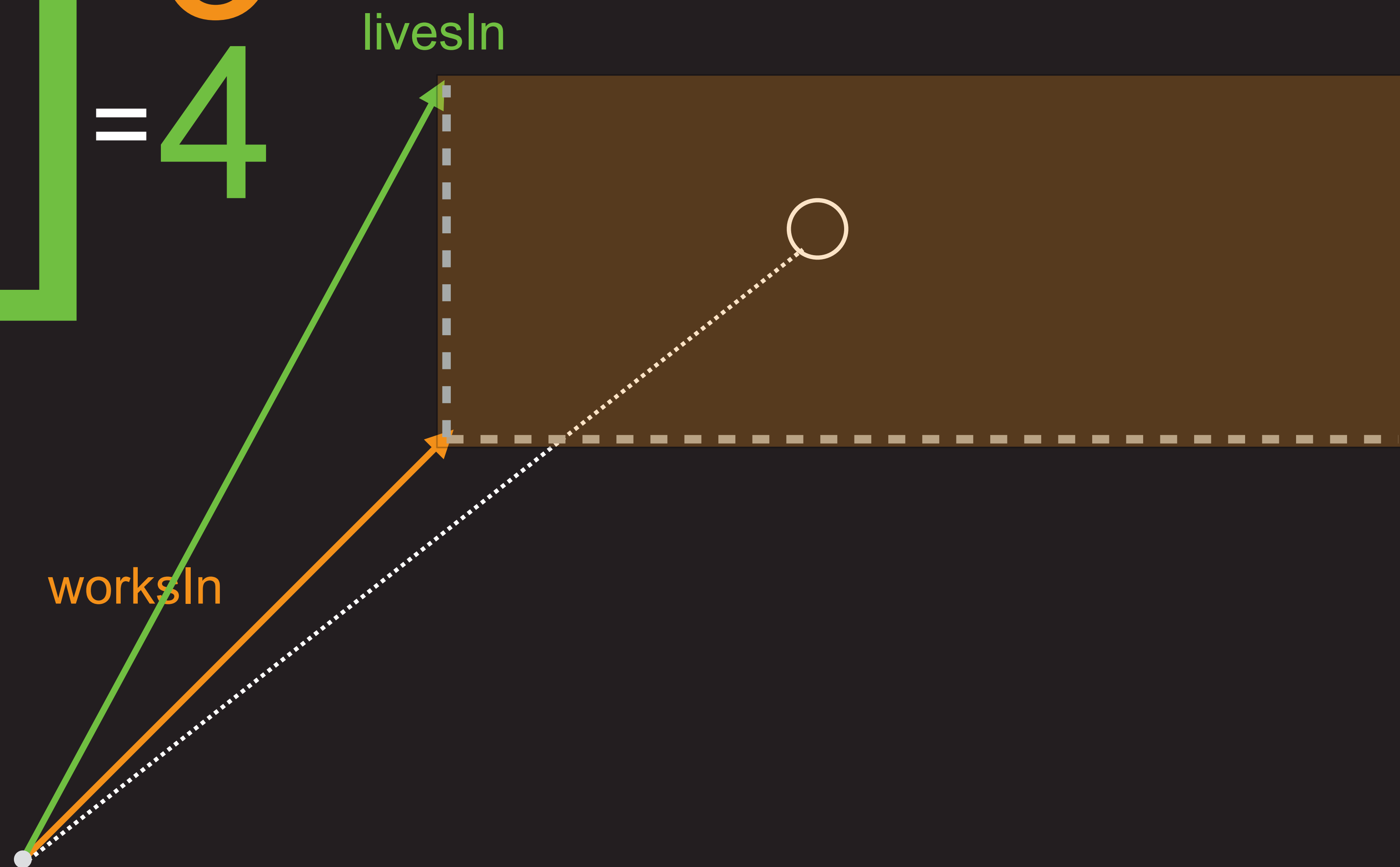
$$\begin{bmatrix}2\\1\end{bmatrix} * \begin{bmatrix}3\\1\end{bmatrix} = \begin{matrix}3\\7\end{matrix}$$



- When the **first component is larger**

Entailment in Vector Space

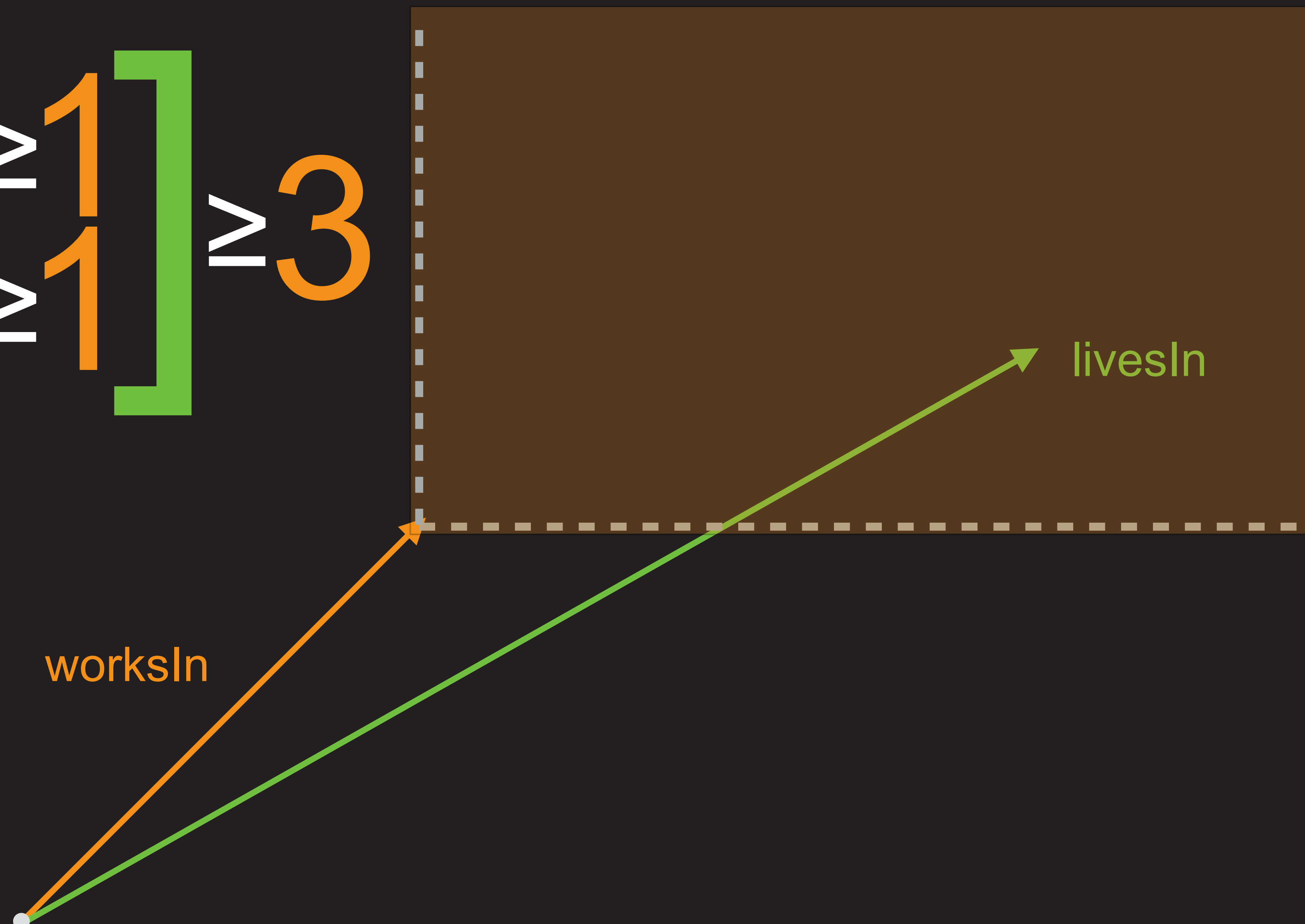
$$\begin{bmatrix}2\\1\end{bmatrix} * \begin{bmatrix}1\\2\end{bmatrix} = \overset{3}{4}$$



- When the **second component is larger**

Entailment in Vector Space

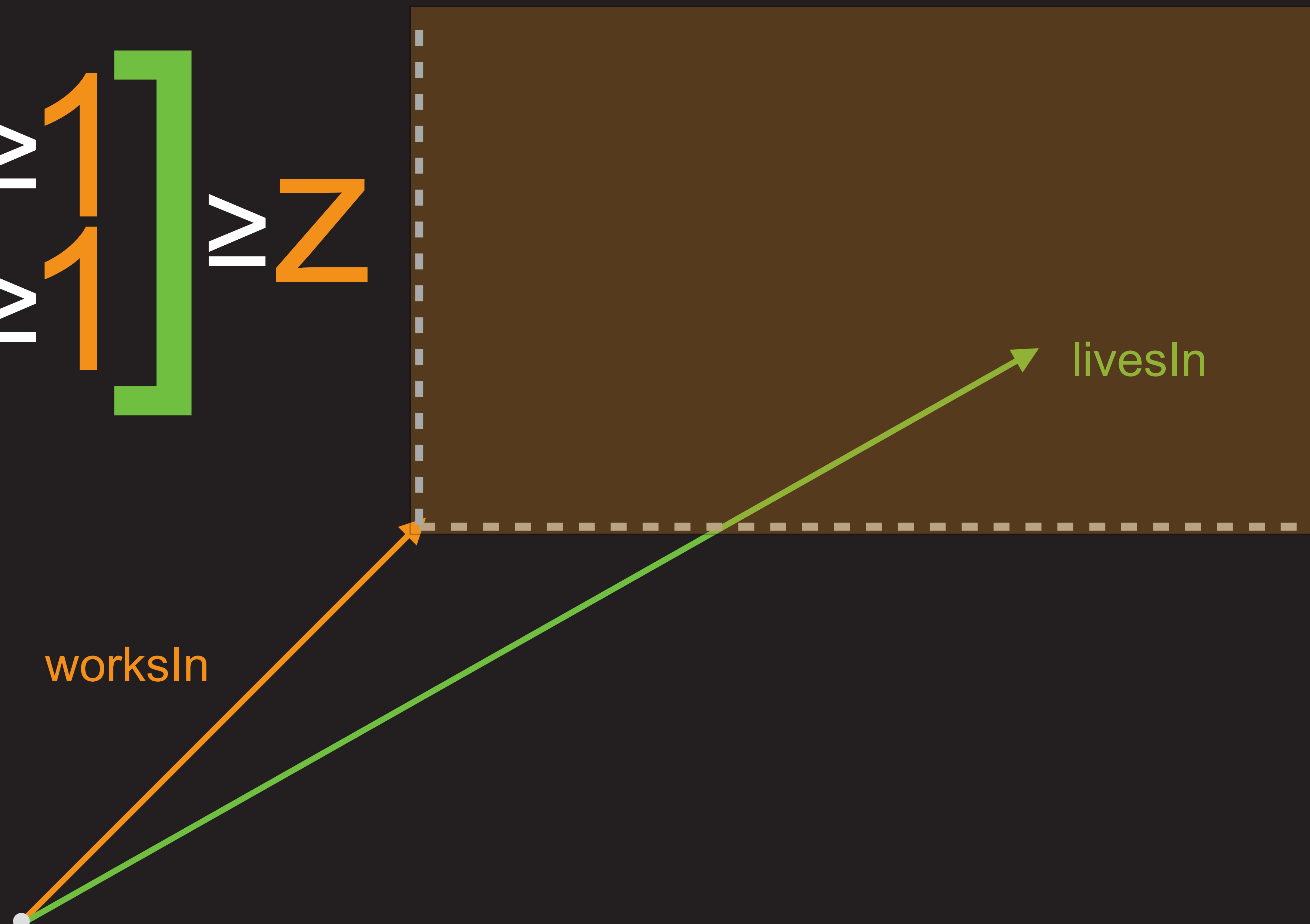
$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} * \begin{bmatrix} x \geq 1 \\ y \geq 1 \end{bmatrix} \geq 3$$



- For any **linear combination**

Entailment in Vector Space

$$\underset{\geq 0}{\begin{bmatrix}a\\b\end{bmatrix}} * \begin{bmatrix}x \geq 1\\ y \geq 1\end{bmatrix} \geq z$$



- And **any (non-negative) input entity pair**

*Compare **Order Embeddings** (Vendrov et al. 2016)*

Results

	FB122 Hits@3
TransE	59%
Kale-TransE (Guo 16)	62%
DistMult	67%
…with rules	71%
Complex	67%
…with rules	**72%**

- Applied to general link prediction problems
- Integrated rules such as
  `nationalityOf(P,N),officialLang(N,L)=>speaks(P,L)`

Conclusion

- Train Neural “Sentence Scorer” with limited data
- and prior knowledge
- Implemented via 2-player game
  - Player 1 learns to predict training facts, and follow rules
  - Player 2 creates tuples (in embedding space) that violate the rules
- Runtime “independent of domain size”
- Future work:
  - inject prior knowledge in natural language
  - extract explanations (see Sameer Singh’s talk)

Papers presented in this work

- **Complex Embeddings for Simple Link Prediction**, Trouillon, Théo, Welbl, Johannes, Bouchard, Guillaume, Riedel, Enriqueastian and Gaussier, Eric, International Conference on Machine Learning 2016
- **Injecting Logical Background Knowledge into Embeddings for Relation Extraction**, Rocktaschel, Tim, Singh, Sameer and Riedel, Enriqueastian, Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2015
- **Lifted Rule Injection for Relation Embeddings**, Demeester, Thomas, Rocktaschel, Tim and Riedel, Enriqueastian, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP) 2016
- **Adversarial Sets for Regularising Neural Link Predictors**, Minervini, Pasquale, Demeester, Thomas, Rocktaschel, Tim, Riedel, Enriqueastian, Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI) 2017

Entailed Predicates need to live in the boxes of their premises