

OxKBC

Outcome Explanation for Factorization Based Knowledge Base Completion

*Yatin Nandwani, Ankesh Gupta, Aman Agrawal, Mayank S. Chauhan, Parag Singla & Mausam
Department of Computer Science & Engineering
Indian Institute of Technology Delhi, INDIA*

Contribution

- Many accurate TF models for KBC, e.g. DistMult^{*}, ComplEx[#]

^{*}Yang B et al. "Embedding entities and relations for learning and inference in knowledge bases", ICLR 2015

[#]Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016.

Contribution

- Many accurate TF models for KBC, e.g. DistMult^{*}, ComplEx[#]
- Models are opaque
 - Reduces user trust in prediction

^{*}Yang B et al. "Embedding entities and relations for learning and inference in knowledge bases", ICLR 2015

[#]Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016.

Contribution

- Many accurate TF models for KBC, e.g. DistMult*, ComplEx#
- Models are opaque
 - Reduces user trust in prediction
- **OxKBC**: Generate explanations for model predictions

*Yang B et al. "Embedding entities and relations for learning and inference in knowledge bases", ICLR 2015

#Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016.

Contribution

- Many accurate TF models for KBC, e.g. DistMult*, ComplEx#
- Models are opaque
 - Reduces user trust in prediction
- **OxKBC**: Generate explanations for model predictions
 - Trains a *post-facto* explainer

*Yang B et al. "Embedding entities and relations for learning and inference in knowledge bases", ICLR 2015

#Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016.

Contribution

- Many accurate TF models for KBC, e.g. DistMult*, ComplEx#
- Models are opaque
 - Reduces user trust in prediction
- **OxKBC**: Generate explanations for model predictions
 - Trains a *post-facto* explainer
 - Faithful to the model

*Yang B et al. "Embedding entities and relations for learning and inference in knowledge bases", ICLR 2015

#Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016.

Contribution

- Many accurate TF models for KBC, e.g. DistMult*, ComplEx#
- Models are opaque
 - Reduces user trust in prediction
- **OxKBC**: Generate explanations for model predictions
 - Trains a *post-facto* explainer
 - Faithful to the model
 - Satisfies end user in a user study on MTurk

*Yang B et al. "Embedding entities and relations for learning and inference in knowledge bases", ICLR 2015

#Trouillon Théo et al. "Complex embeddings for simple link prediction." ICML, 2016.

What is a TF Model?

- **Knowledge Graph:**

- Nodes:

Entities e.g. Tory, Basketball

- Directed Edges:

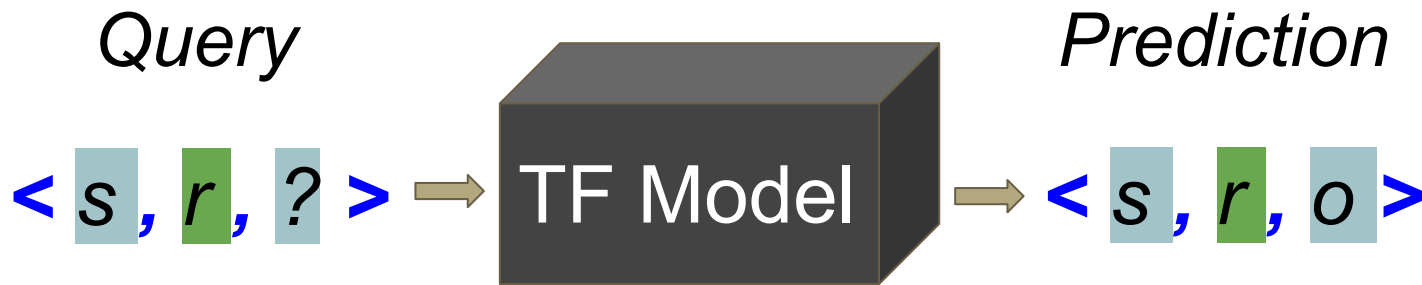
Relations b/w edges e.g. GoodAt

What is a TF Model?

- **Knowledge Graph:**
 - Nodes: Entities e.g. Tory, Basketball
 - Directed Edges: Relations b/w edges e.g. GoodAt
- **TF Model:** Learn embeddings for each entity and relation

What is a TF Model?

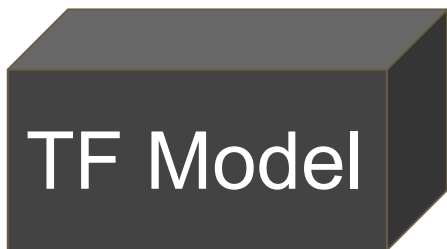
- **Knowledge Graph:**
 - Nodes: Entities e.g. Tory, Basketball
 - Directed Edges: Relations b/w edges e.g. GoodAt
- **TF Model:** Learn embeddings for each entity and relation



OxKBC: Outcome explanation for KBC

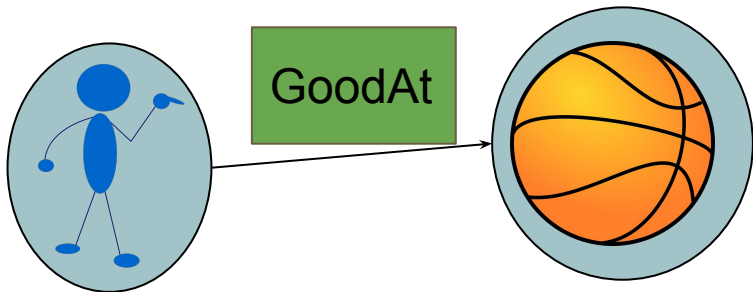
Query

$\langle s, r, ? \rangle$



Prediction

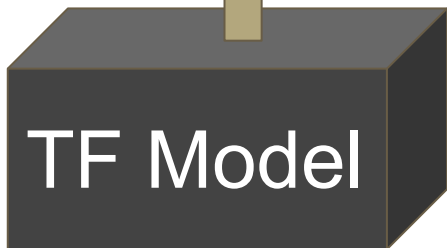
$\langle s, r, o \rangle$



OxKBC: Outcome explanation for KBC

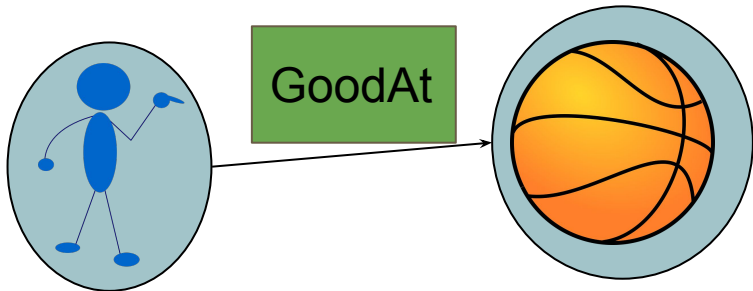
Query

$\langle s, r, ? \rangle$

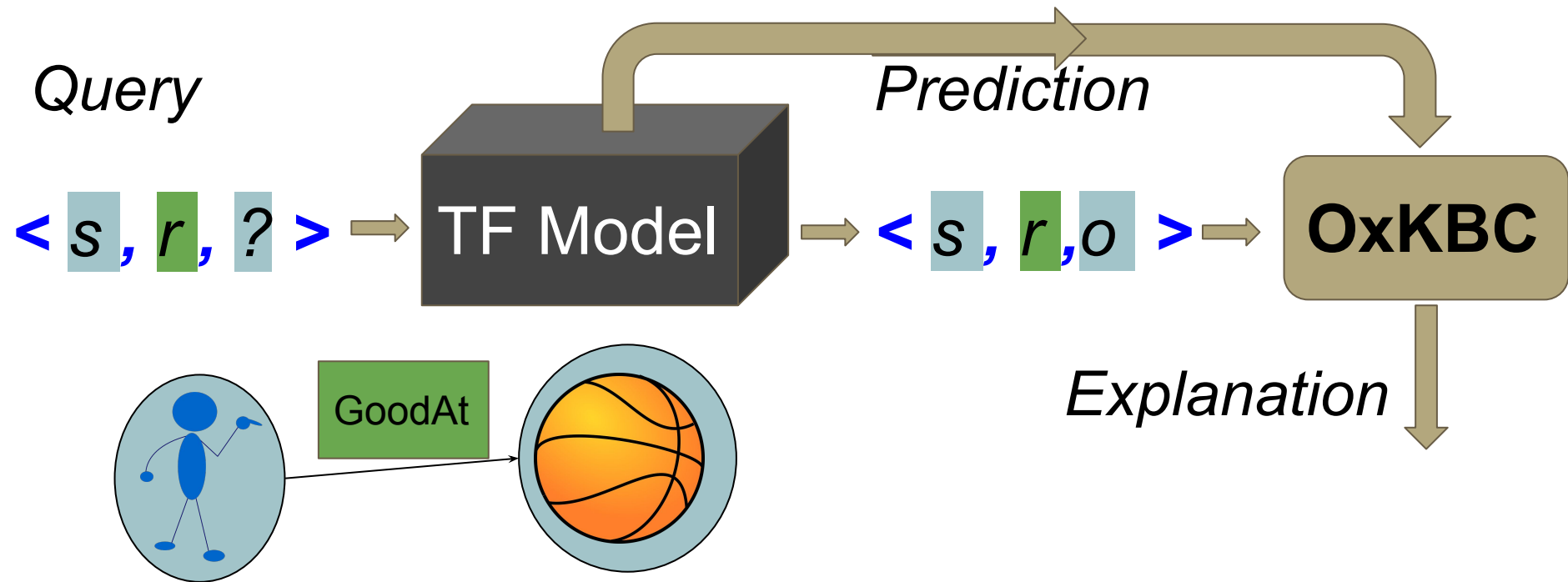


Prediction

$\langle s, r, o \rangle$



OxKBC: Outcome explanation for KBC



OxKBC: Outcome explanation for KBC

Query

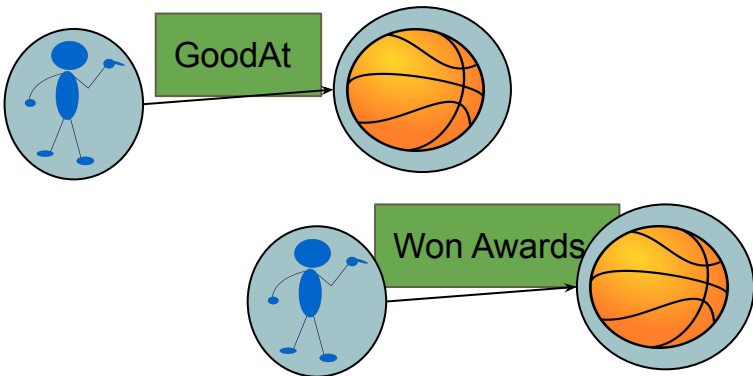
$\langle s, r, ? \rangle$

TF Model

Prediction

$\langle s, r, o \rangle$

OxKBC



Explanation

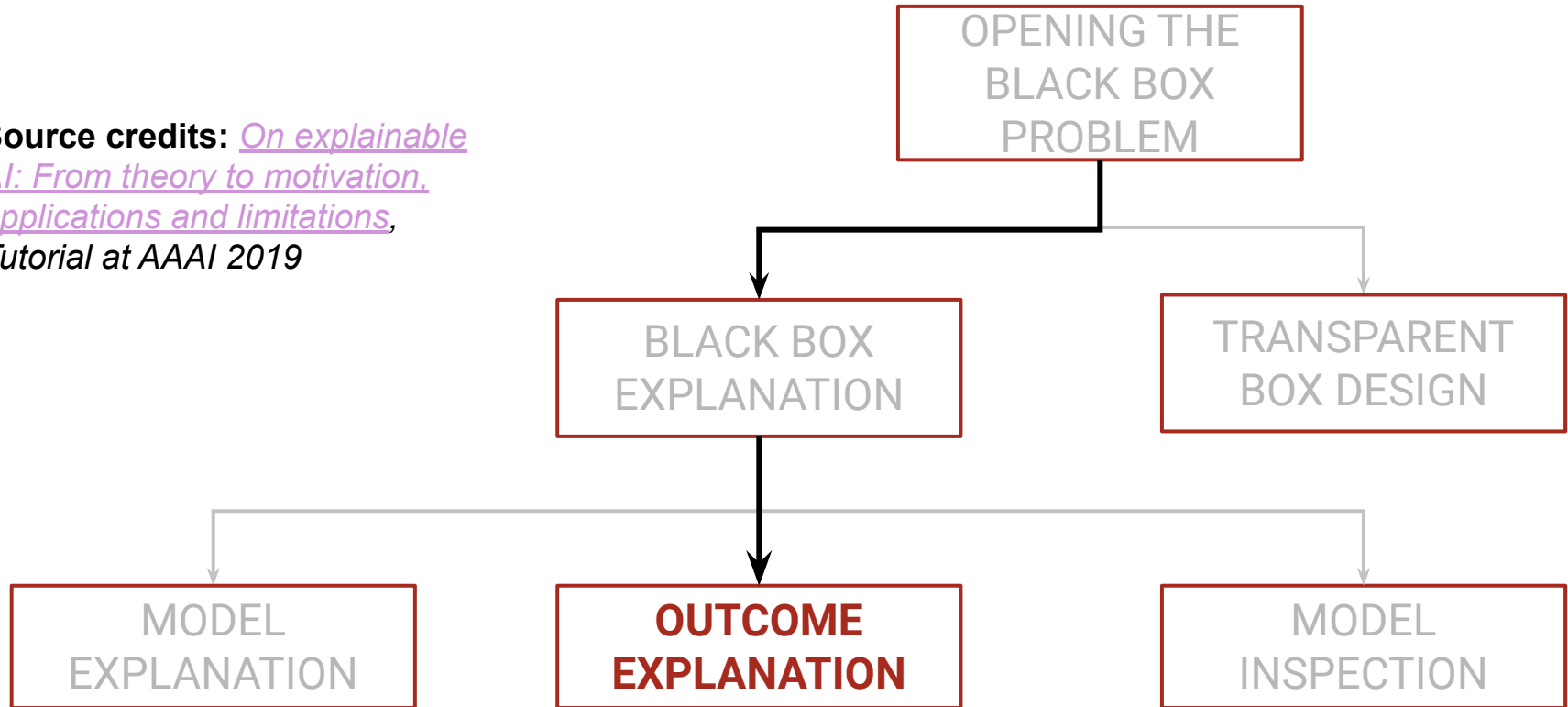
$\langle s, r, ? \rangle$

because

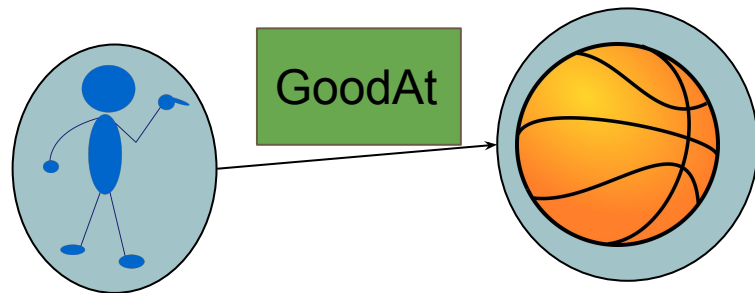
$\langle s, r', ? \rangle \& r' \sim r$

Taxonomy of Explanation Engines

Source credits: [On explainable AI: From theory to motivation, applications and limitations](#),
Tutorial at AAAI 2019

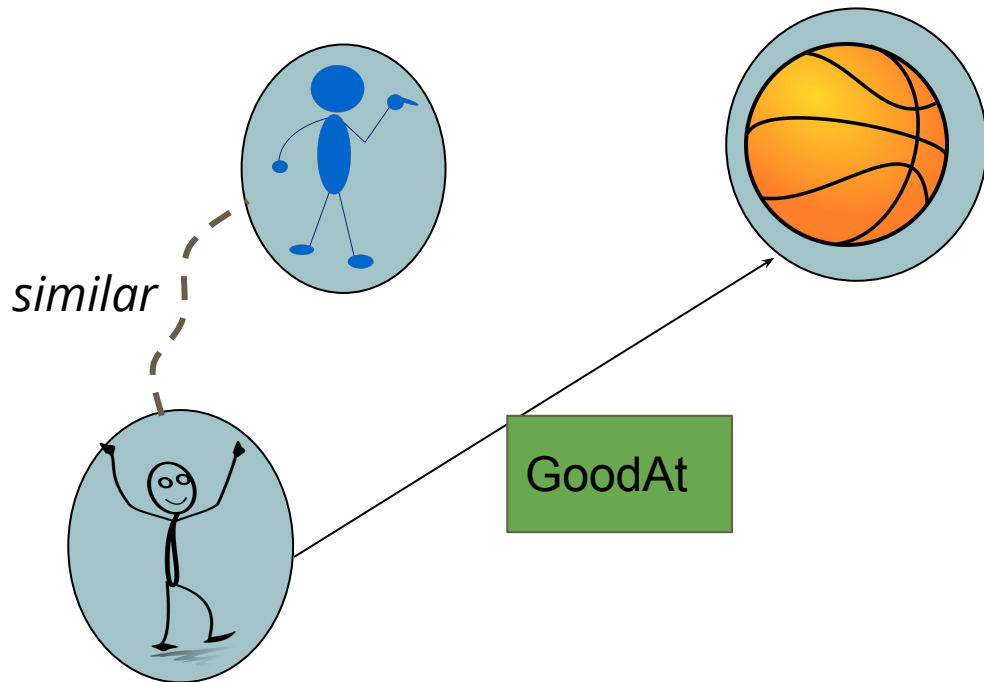


OxKBC: Different Templates of Explanations



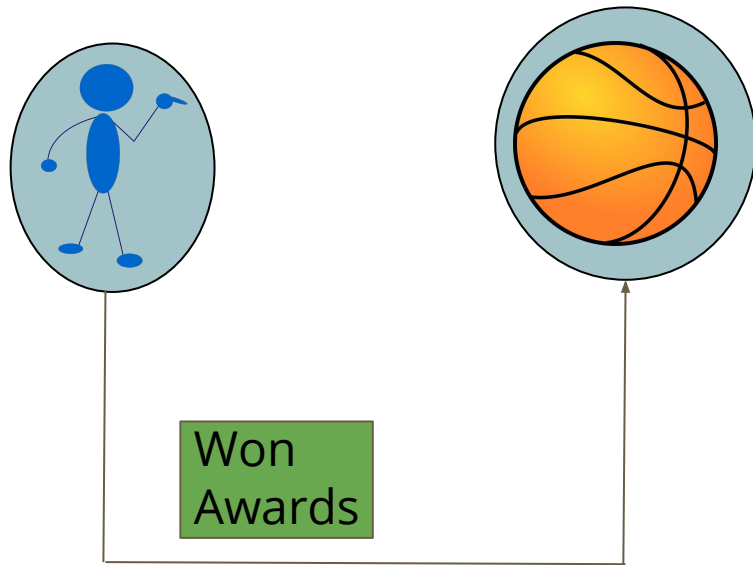
OxKBC: Different Templates of Explanations

- Entity Similarity (T1)



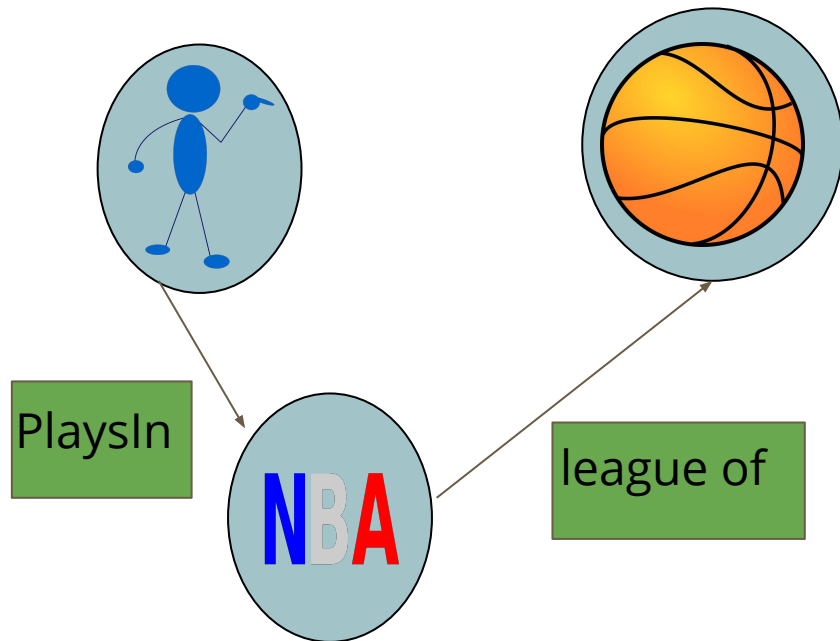
OxKBC: Different Templates of Explanations

- Entity Similarity (T1)
- **Relation Similarity (T2)**



OxKBC: Different Templates of Explanations

- Entity Similarity (T1)
- Relation Similarity (T2)
- **Two Length Paths (T3)**



OxKBC: Selection Module

- How to pick the best explanation?

OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.

OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
- **Selection Module:**

OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
- **Selection Module:**
 - Novel features for each template

OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
- **Selection Module:**
 - Novel features for each template
 - Novel unsupervised loss function

OxKBC: Selection Module

- How to pick the best explanation?
- No annotation over explanations.
- **Selection Module:**
 - Novel features for each template
 - Novel unsupervised loss function
 - Handful of annotation reduces variance

Turk Experiments

- How good are OxKBC's explanations?
- Do end users prefer OxKBC vs Rule Mining*?

Turk Experiments

Question: Turun Palloseura football team has a play position _____ ?

Answer: Midfielder

Explanation A: Turun Palloseura football team has a play position Defender and Defender is a position in the football team Lierse S.K. and Lierse S.K. has a position of Midfielder

Explanation B: In our Knowledge Base, *many football teams (598 of 745)* have players at Midfielder

(Australia national soccer team , Olympique de Marseille and *596 more...*) football team has a play position Midfielder

- A is better than B
- B is better than A
- Both A and B are equally good
- Both A and B are bad

Turk Experiments

	OxKBC Better	Rules Better	Tie	Total
	145	49	18	212

Faithfulness of Explanations

- What if explaining facts are removed from KB?

Faithfulness of Explanations

- What if explaining facts are removed from KB?
- How will it impact model's prediction?

Faithfulness of Explanations

- What if explaining facts are removed from KB?
- How will it impact model's prediction?
- Measure through reduction in MRR:
 - More decrease => More faithful

Faithfulness of Explanations

- What if explaining facts are removed from KB?
- How will it impact model's prediction?
- Measure through reduction in MRR:
 - More decrease => More faithful

	MRR		
Steps	T1	T2	T3
1	1.00	1.00	1.00
1	0.50	0.60	0.40
2	0.30	0.34	0.24
5	0.20	0.18	0.13

Conclusion

- OxKBC provides post-hoc explanations
 - for **any** factorization based KBC models.
- Faithful to the underlying model
- Satisfies end user in a user study on MTurk
- Increases trust in the underlying model

All code and data at: <https://github.com/dair-iitd/OxKBC>