



Extended Question Answering on Freebase

Donnerstag, 15. September 2016
Abschlusskolloquium

Matthias Urban

Lehrstuhl für Algorithmen und Datenstrukturen

Universität Freiburg



Einleitung

- Fortsetzung von „More Accurate Question Answering on Freebase“ von Hannah Bast und Elmar Haussmann
- Kann nur strukturell einfache Fragen beantworten
- Zu komplexe Fragen: Zwei Fragen in einer, Temporal-Fragen, Superlativ-Fragen
- Answer Type Matching



Inhalt

- Bisheriges System
- Erweiterungen
- Answer Type Matching
- Evaluation
- Fazit

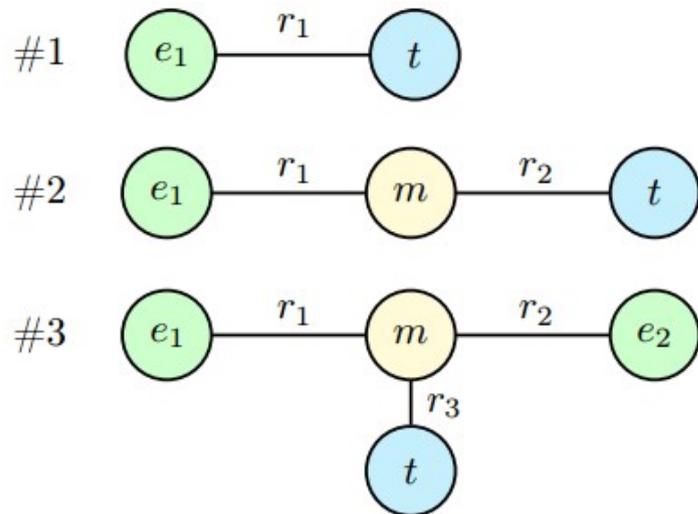


Bisheriges System

- Identifizierung der Entitäten
- Template Matching
- Relation Matching
- Answer Type Matching
- Features und Ranking

Bisheriges System - Template Matching

Template



Example Candidate

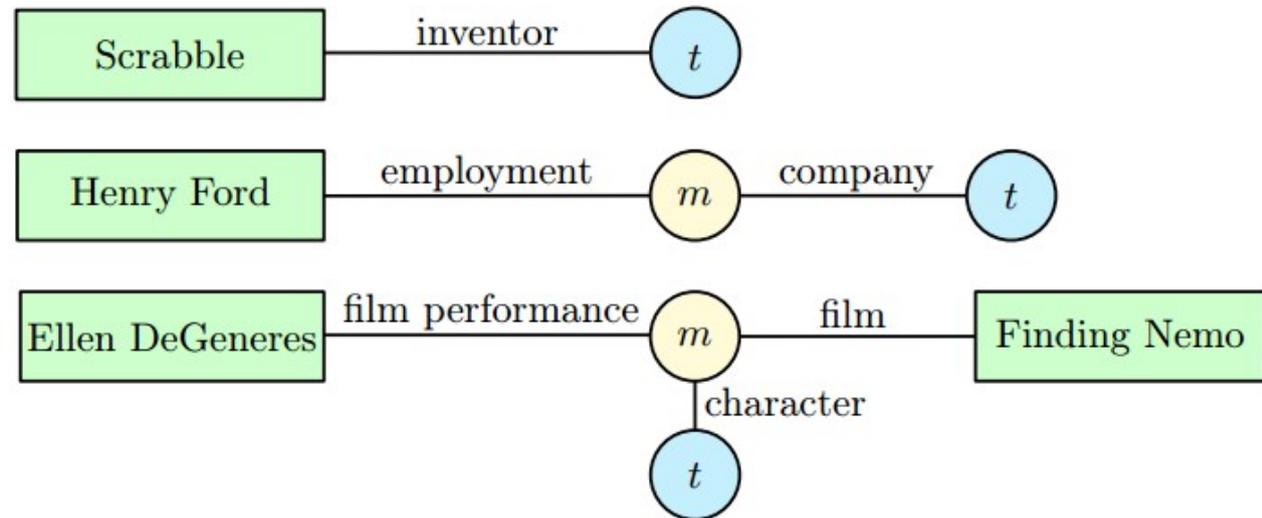


Bild aus „More Accurate Question
Answering on Freebase“



Bisheriges System

- Identifizierung der Entitäten
- Template Matching
- Relation Matching
- Answer Type Matching
- Features und Ranking

Erweiterungen - Idee

Question

who was a wife of a president of the united states

Translation Options

Ranker WQ_Ranker Uses relation oracle Uses entity oracle

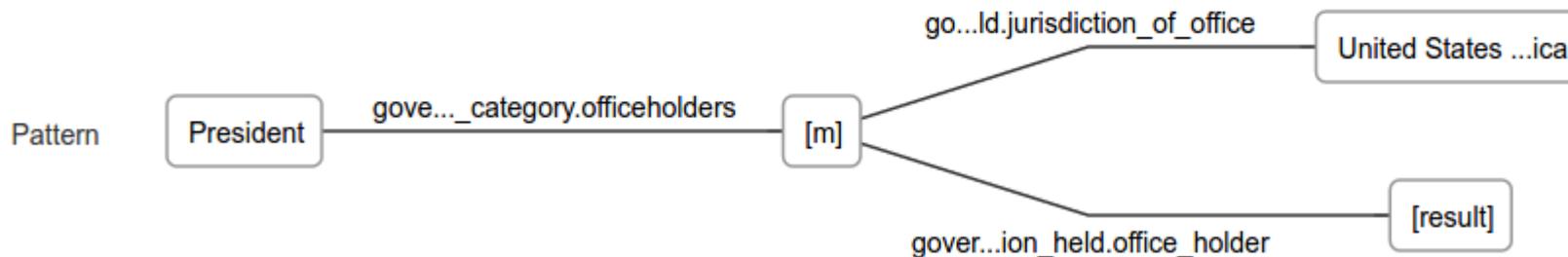
Translate and execute

Identify entites

Total translation time : 9305.81 ms

Candidate 1 of 60 ▾

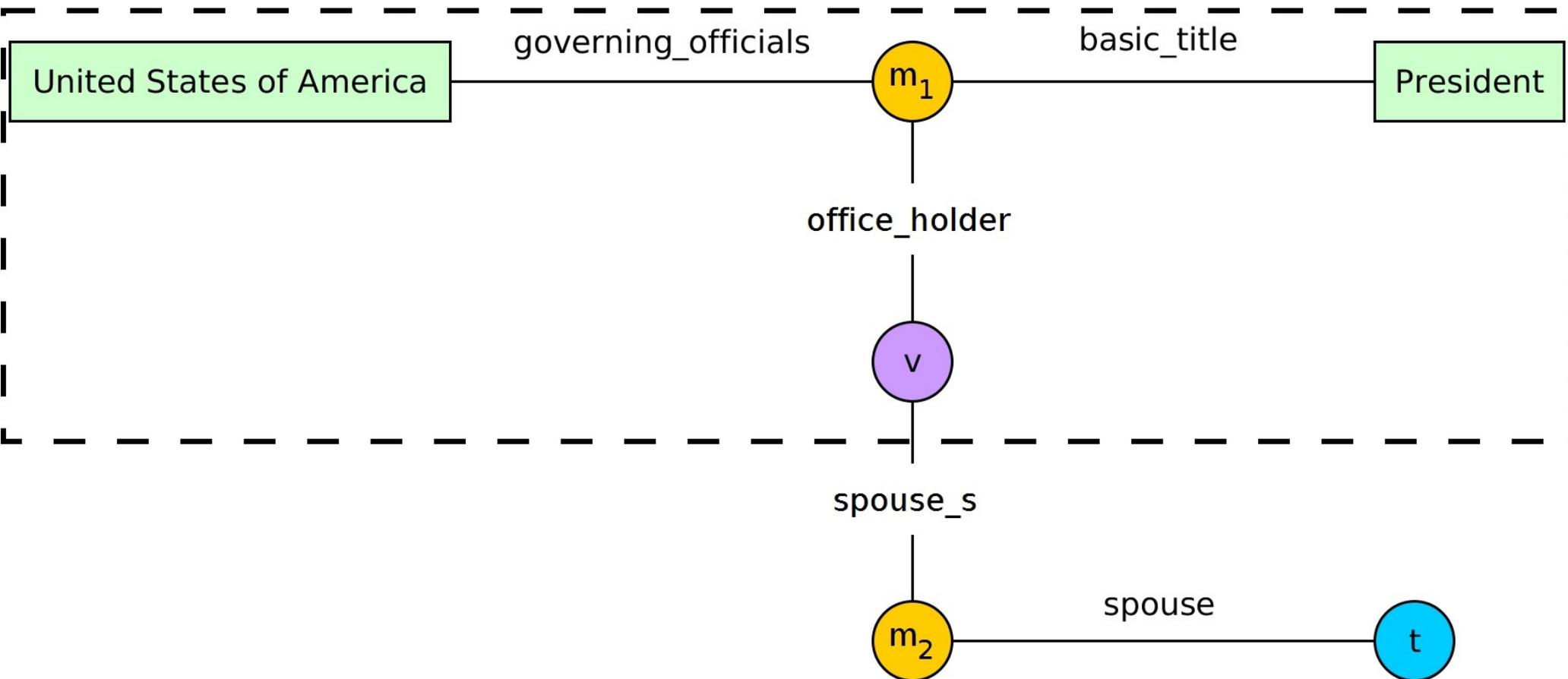
Question who was a wife of a **president** of the **united states**
0.22 | 12.74 0.96 | 18.82



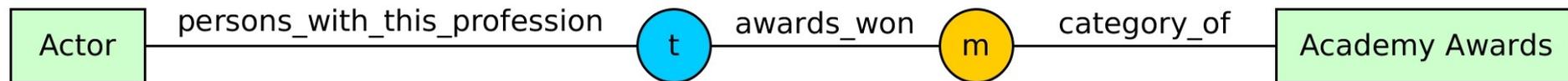
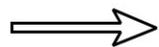
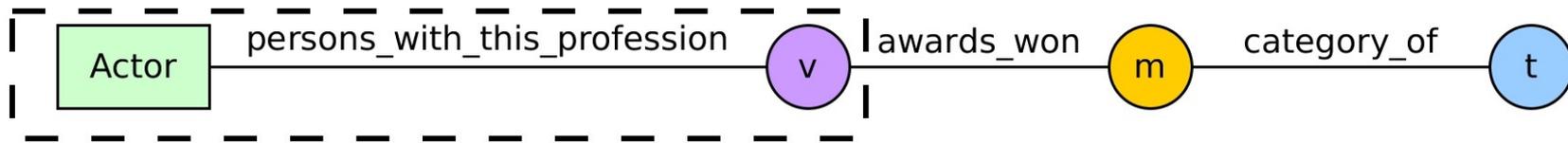
Result (43) Benjamin Harrison James Monroe William Henry Ha...n James A. Garfield Franklin D. Roos...t Geor...

Details ▶

Erweiterungen - Zwei Fragen in einer

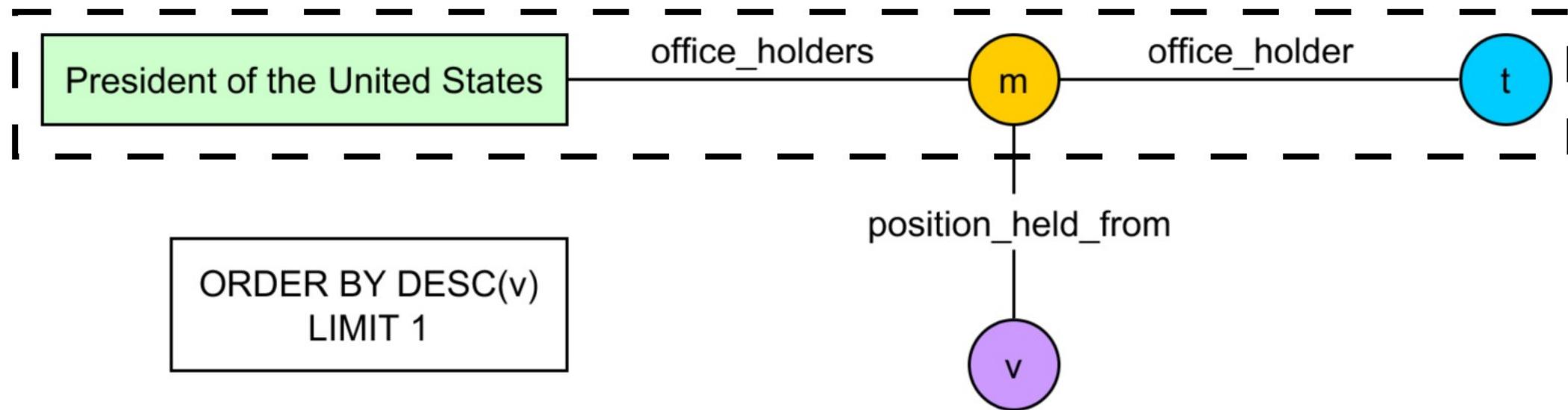


Erweiterungen - Zwei Fragen in einer

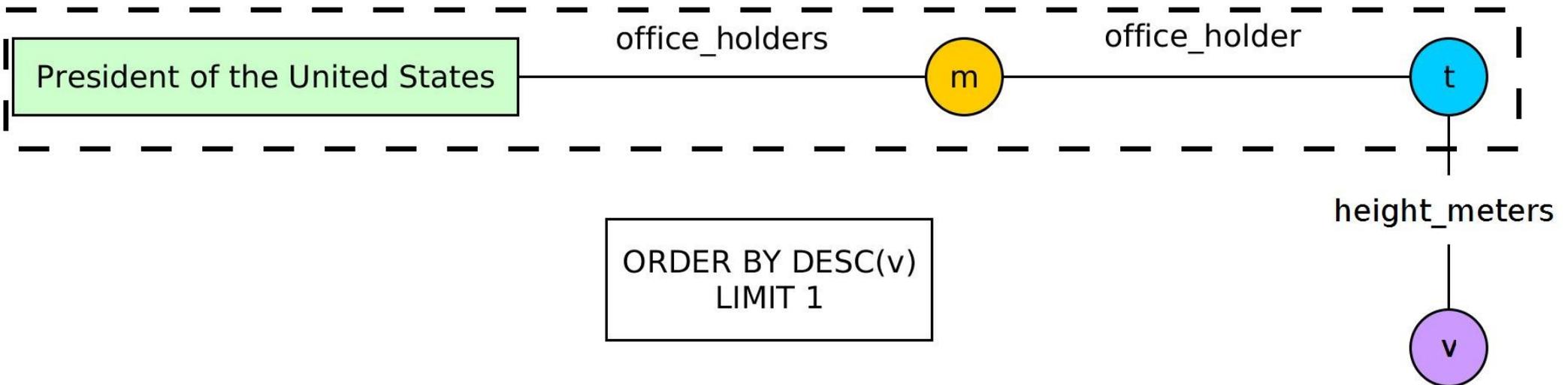


Erweiterungen - Temporal-Fragen

- Signalwörter: last, current, currently, at the moment, final, now, newest, latest, first, oldest, earliest



Erweiterungen - Superlativ-Fragen





Answer Type Matching - Idee

- Ein auf Machine Learning basierender Klassifizierer
- Syntaktische und semantische Features
- Typen von Freebase

Answer Type Matching - Features

1-grams	Who.WP, is.VBZ, the.DT, wife.NN, of.IN, the.DT, president.NN, of.IN, ...
---------	--------------------------------------------------------------------------

2-grams	Who.WP is.VBZ, is.VBZ the.DT, the.DT wife.NN, wife.NN of.IN, ...
---------	------------------------------------------------------------------

NP-chunks	the.DT wife.NN, the.DT president.NN, the.DT united.NNP states.NNPS
-----------	--------------------------------------------------------------------

VP-chunks	is.VBZ
-----------	--------

Head NP-chunks	the.DT wife.NN
----------------	----------------

Head VP-chunks	is.VBZ
----------------	--------

Head Word	wife.NN
-----------	---------

Answer Type Matching - Features

Entitäten	President of the United States, The Wife, United States of America, ...
Entitätstypen	government_office, book, location, ...
Entitäten im Head NP-chunk	The Wife, ...
Entitätstypen im Head NP-chunk	Book, ...

Sense IDs	president_of_the_united_states.n.01, president.n.04, ...
Hyperonym IDs	head_of_state.n.01, spouse.n.01, ...
Hyponym IDs	run_into.v.01, fall.v.16, diverge.v.02, reach.v.06

Evaluation - Erweiterungen

	Accuracy
Insgesamt	50 %
Zwei Fragen in einer – Typ 1	70 %
Zwei Fragen in einer – Typ 2	70 %
Einfache Temporal-Fragen	50 %
Superlativ Fragen	10 %

Evaluation - Erweiterungen

	Mit Erweiterungen trainiert. Mit Erweiterungen evaluiert.	Ohne Erweiterungen trainiert. Mit Erweiterungen evaluiert.	Ohne Erweiterungen
Accuracy	57,2%	40,2%	65,9%

Evaluation - Answer Type Matching

Datensatz	WebQuestions	Free917
Accuracy	0,63	0.54
Average Precision	0,398	0.25
Average Recall	0,18	0.35
Anzahl Antworttypen, die nie korrekt bestimmt wurden	78 (insgesamt 116 Antworttypen im Testset)	28 (insgesamt 48 Antworttypen im Testset)



Fazit



**Vielen Dank für Ihre
Aufmerksamkeit**