Graph Methods for Complex and Conversational Question Answering over Knowledge Bases

Rishiraj Saha Roy
Max Planck Institute for Informatics, Germany

Virtual Talk at CEDAR

26 November 2020

TEQUILA CIKM 18 S

QUEST SIGIR 19

CROWN SIGIR 20 D

PRINCE WSDM 20

FAIRY WSDM 19

Advancing QA over KBs

- Interpretability in question answering — **QUINT** EMNLP 17 D
- Improving efficiency in QA with answer types — **TIPI** IJCNLP 17 S
- A continuous learning framework for QA — **NEQA** WWW 18
- QA with temporal conditions — **TEQUILA** CIKM 18 S
- Answering complex questions from using QKGs — **QUEST** SIGIR 19
- A challenging benchmark for QA — **ComQA** NAACL 19
- Conversational QA over KGs and passages — **CROWN** SIGIR 20 D, **CONVEX** CIKM 19

- Parallel thread on transparency — **PRINCE** WSDM 20, **FAIRY** WSDM 19

Outline

- **Background:** Setup, benchmarks, metrics
- **Conversational QA:** Implicit context in multi-turn setup
- **Complex QA:** Multiple entities and predicates
- **Take-home:** Open problems and summary

What is question answering over knowledge graphs all about?

Question Answering: Vital for Search

What are some films directed by Nolan?

Question Answering: Vital for Search

What are some films directed by Nolan?

Google Assistant

Hey Siri

Hi. I'm Cortana.
Ask me a question!

amazon alexa

Christopher Nolan / Films directed

The Dark Knight
2008

Interstellar
2014

Question Answering: Vital for Search

What are some films directed by Nolan?

- Direct answers to questions
- Enabled by knowledge bases
- Saves time and effort
- Natural in voice UI

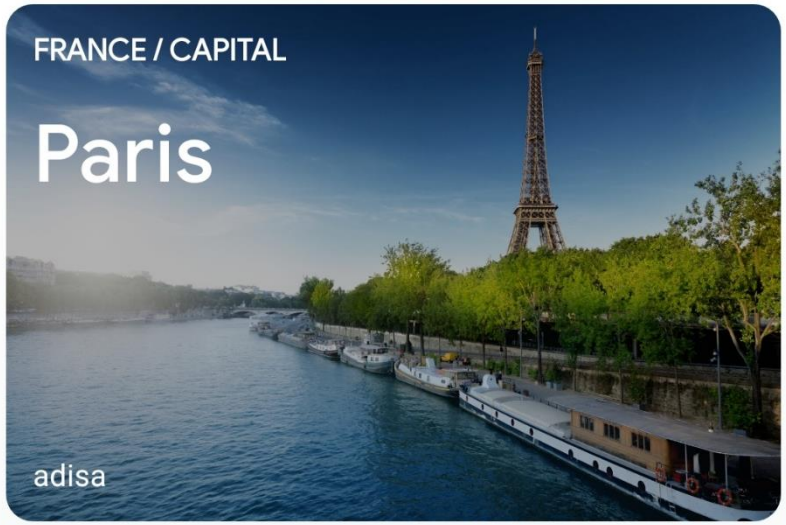
Christopher Nolan / Films directed

The Dark Knight
2008

Interstellar
2014

QA over KBs

what is the capital of France



what is the currency of India

where was Albert Einstein born

which club does Messi play for

LIONEL MESSI / CURRENT TEAM

FC Barce lona

what is the population of USA

United States / Population

328.2 million (2019)

Simple questions involving one entity and relation

which films star tom hanks and are directed by spielberg

All News Images Videos Shopping More Settings Tools

Steven Spielberg / Films directed / Tom Hanks / Movies

Bridge of Spies


The Terminal

Catch Me If You Can

The Post


Saving Private Ryan

what was spielberg's father's profession

All Images News Videos Shopping More Settings Tools

About 3.690.000 results (0,75 seconds)

Arnold Spielberg / Profession

Electrical engineer

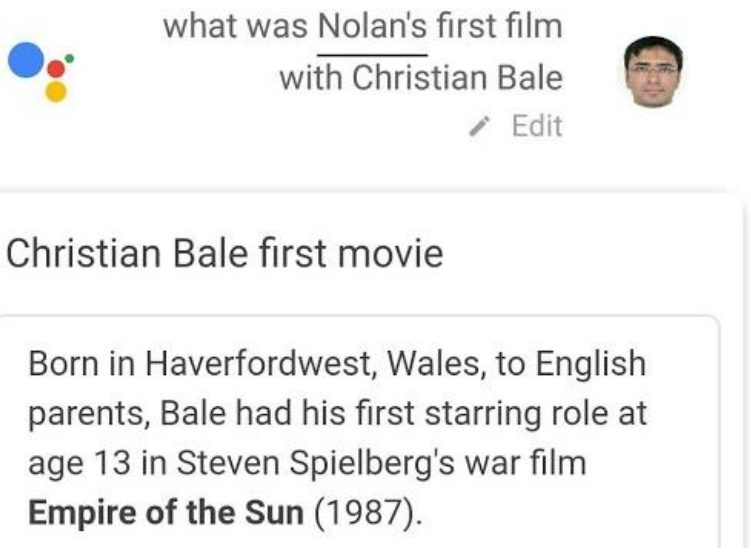
Complex questions involving multiple entities and relations

Conversational questions with implicit context

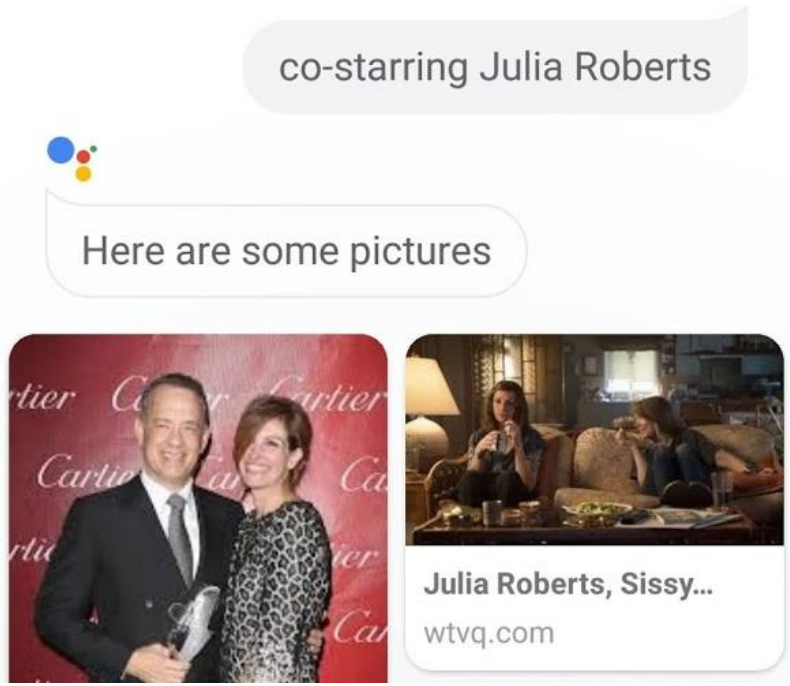
who played Batman in Dark Knight

and what about Alfred

Edit



Play with QA: Try out different formulations, entities, domains, complexities, assistants, sources, languages.... to expose brittleness of SoTA and take community forward!



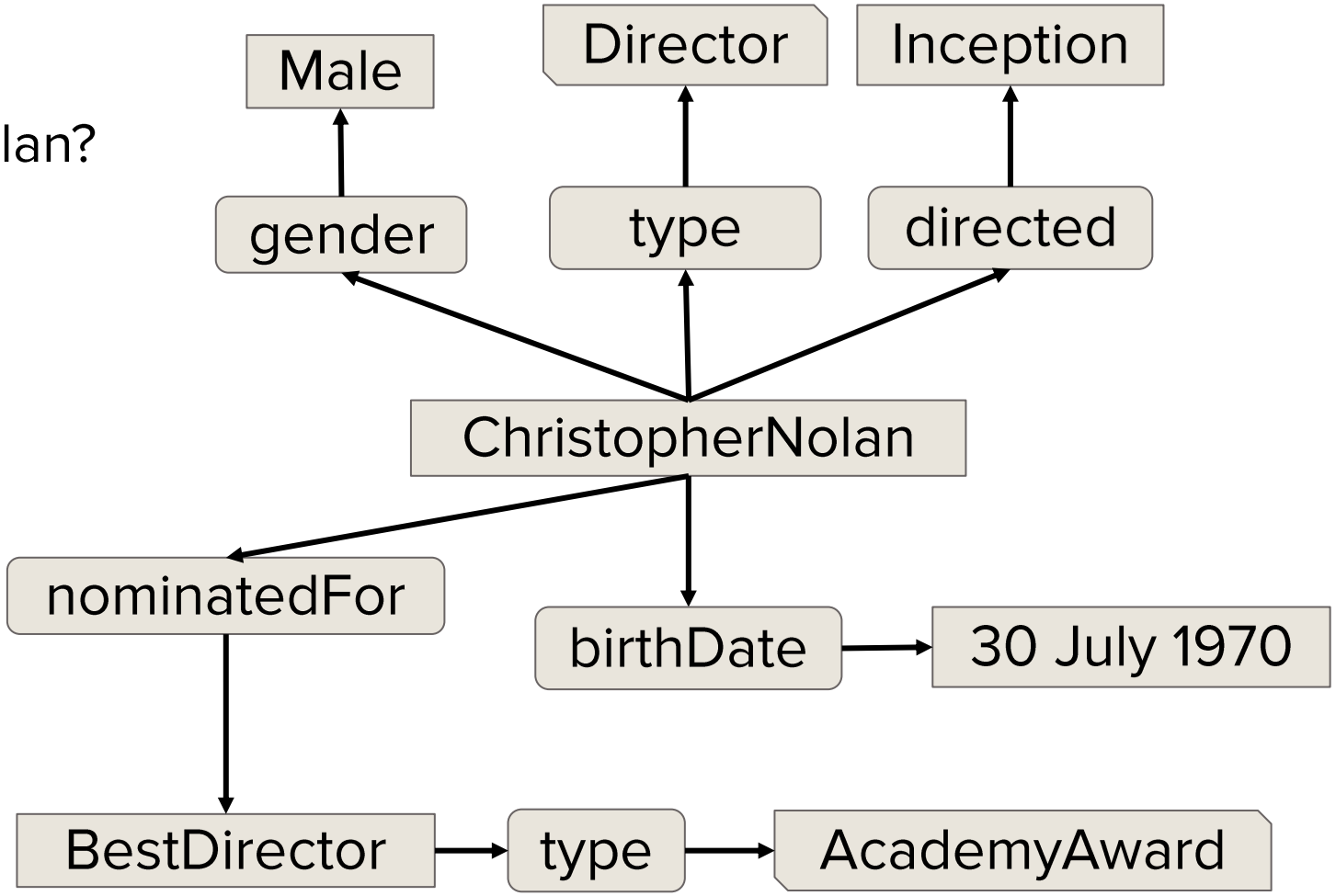
Significant progress has been made on **knowledge base contruction** over the last fifteen years or so; but for **question answering**, which is one of the most valuable applications of KBs, we are still at the **tip of iceberg**!

QA over knowledge bases (KB-QA)

What are the Oscar nominations of Nolan?

QA over knowledge bases (KB-QA)

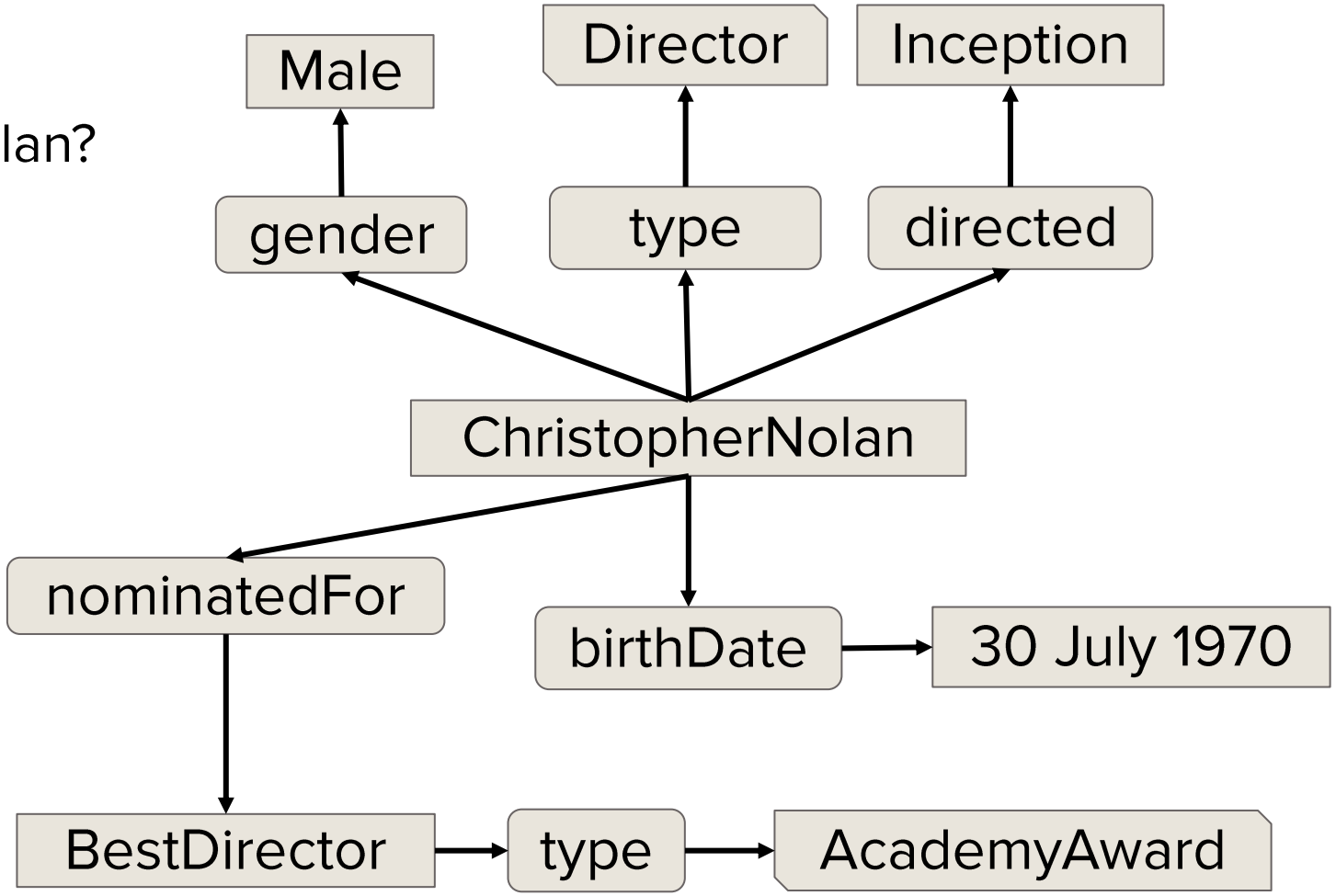
What are the Oscar nominations of Nolan?



QA over knowledge bases (KB-QA)

What are the Oscar nominations of Nolan?

- YAGO [Suchanek et al. 2007]
- DBpedia [Auer et al. 2007]
- Freebase [Bollacker et al. 2008]
- Wikidata [Vrandečić and Krötzsch 2014]



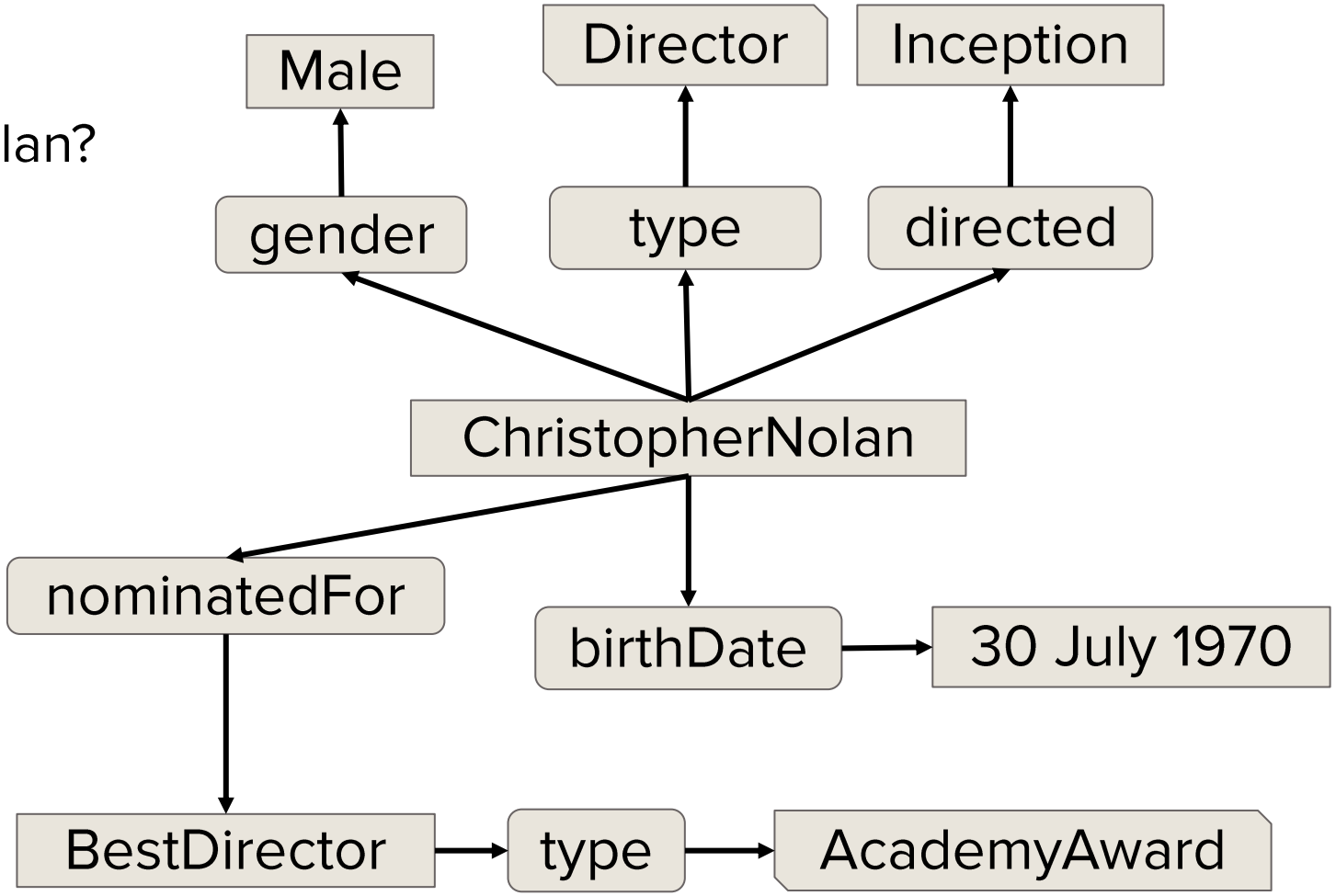
QA over knowledge bases (KB-QA)

What are the Oscar nominations of Nolan?

- YAGO [Suchanek et al. 2007]
- DBpedia [Auer et al. 2007]
- Freebase [Bollacker et al. 2008]
- Wikidata [Vrandečić and Krötzsch 2014]

Terminology varies across KGs

Here: Entities, predicates, types, literals

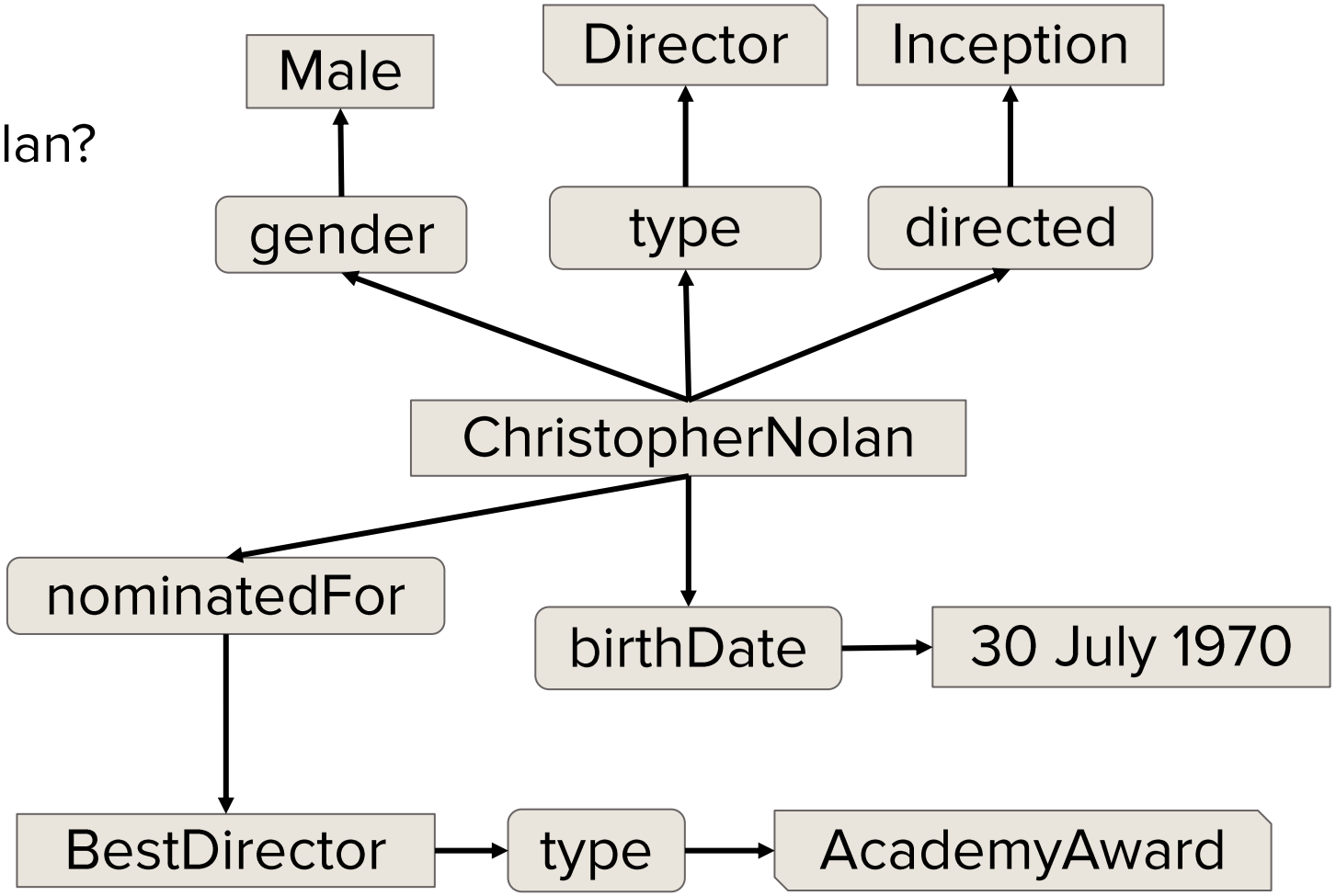


QA over knowledge bases (KB-QA)

What are the Oscar nominations of Nolan?

- YAGO [Suchanek et al. 2007]
- DBpedia [Auer et al. 2007]
- Freebase [Bollacker et al. 2008]
- Wikidata [Vrandečić and Krötzsch 2014]

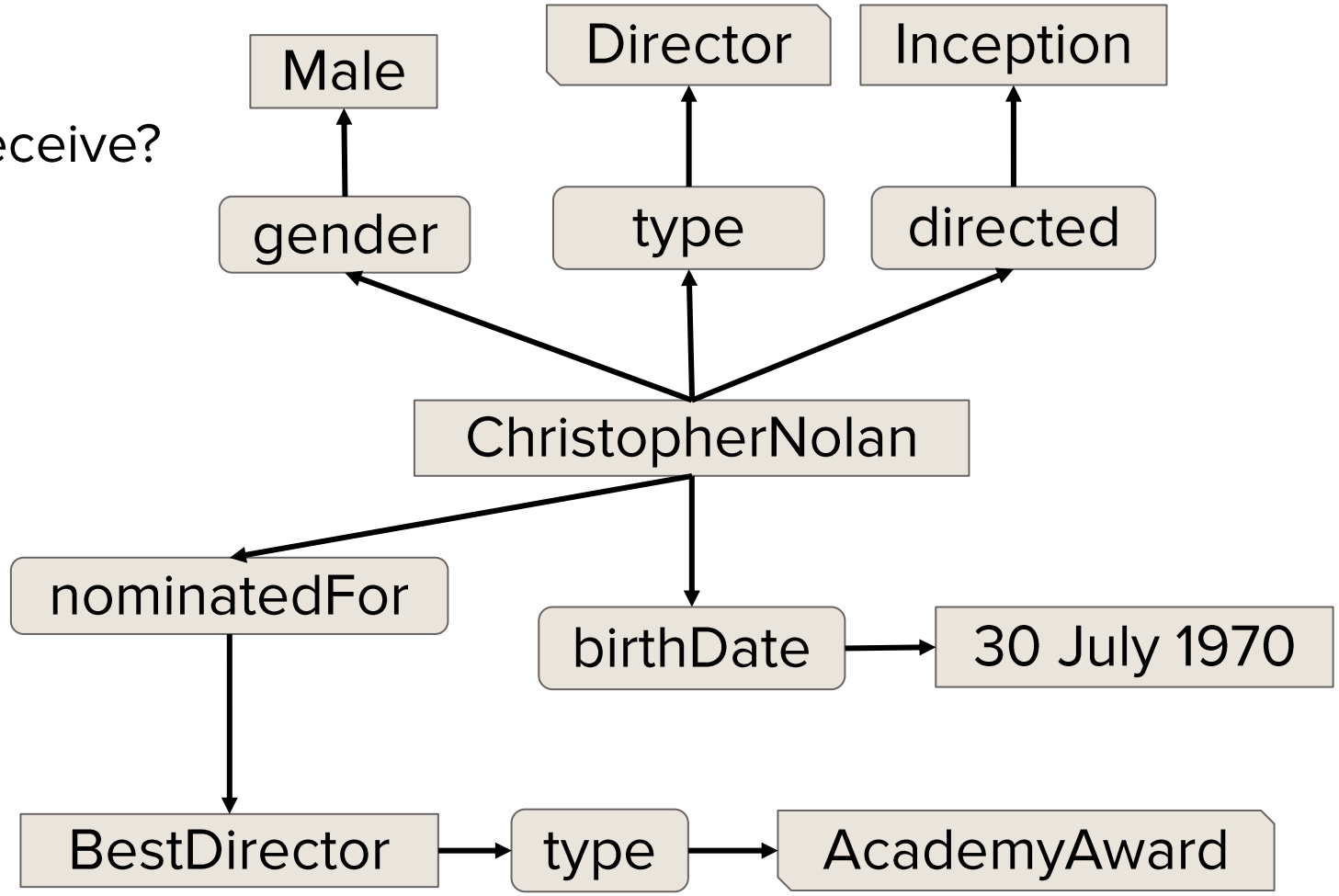
Wikidata: 12B facts, 84M entities, 7k predicates, 69k types



KGs and KBs are equivalent

Which Oscar nominations did Nolan receive?

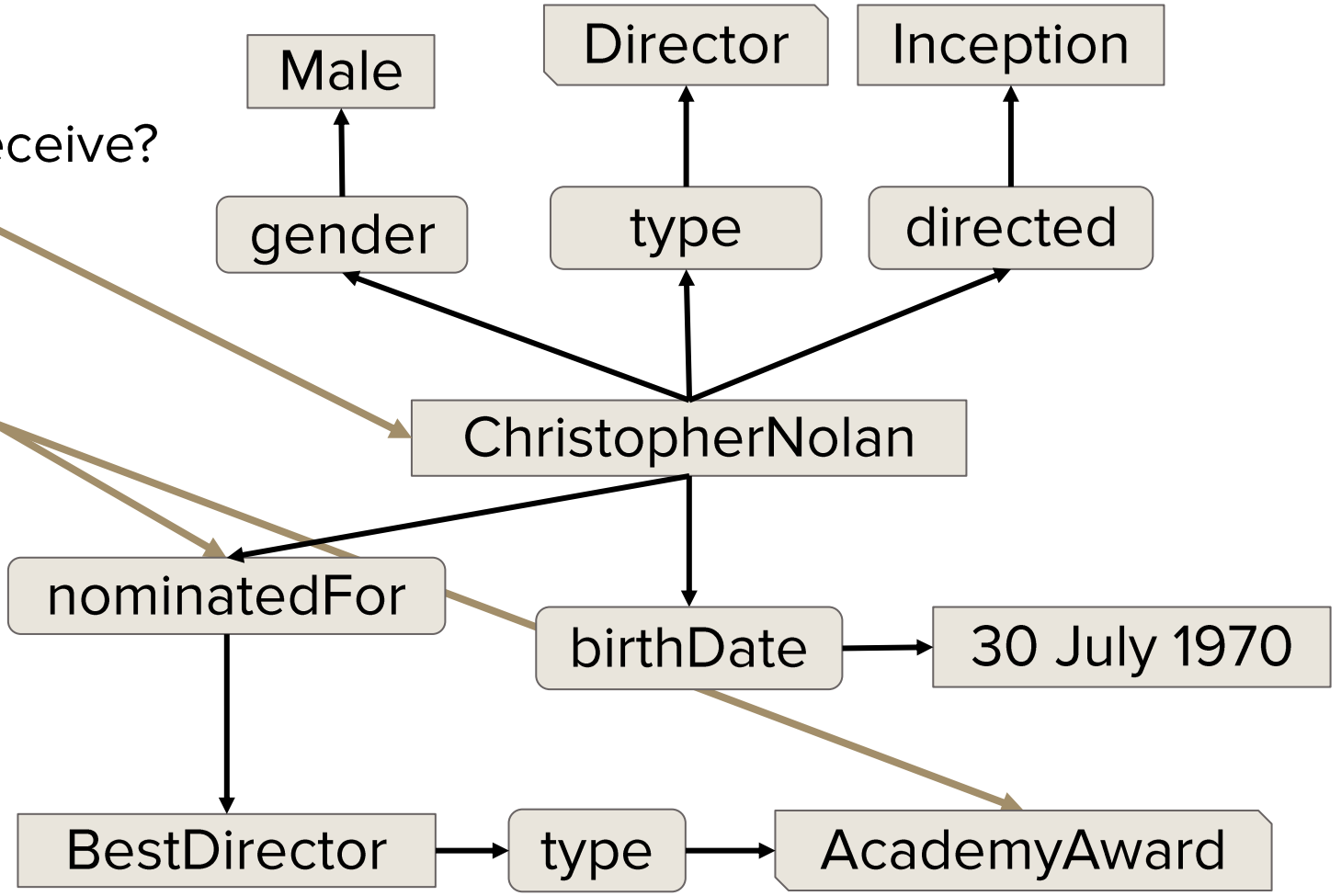
<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>



KG-QA Challenge 1: Bridge vocabulary gap

Which Oscar nominations did Nolan receive?

<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>



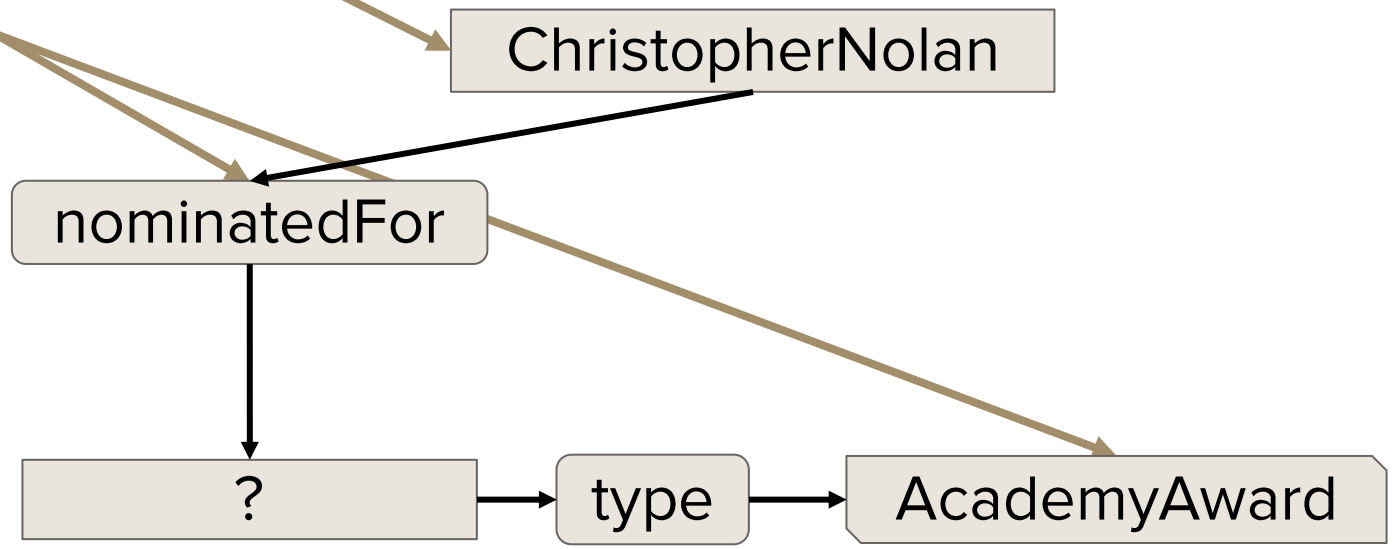
KG-QA Challenge 2: Query formulation

Which Oscar nominations did Nolan receive?

<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>

SELECT ?ANS
WHERE {
ChristopherNolan nominatedFor ?ANS .
?ANS type AcademyAward }

SPARQL



KG-QA Challenge 2: Query formulation

Named Entity Recognition and Disambiguation (NERD) systems (aka Entity Detection and Linking): TagME, AIDA, Dandelion, Google NL API, MS Text Analytics, IBM NLU

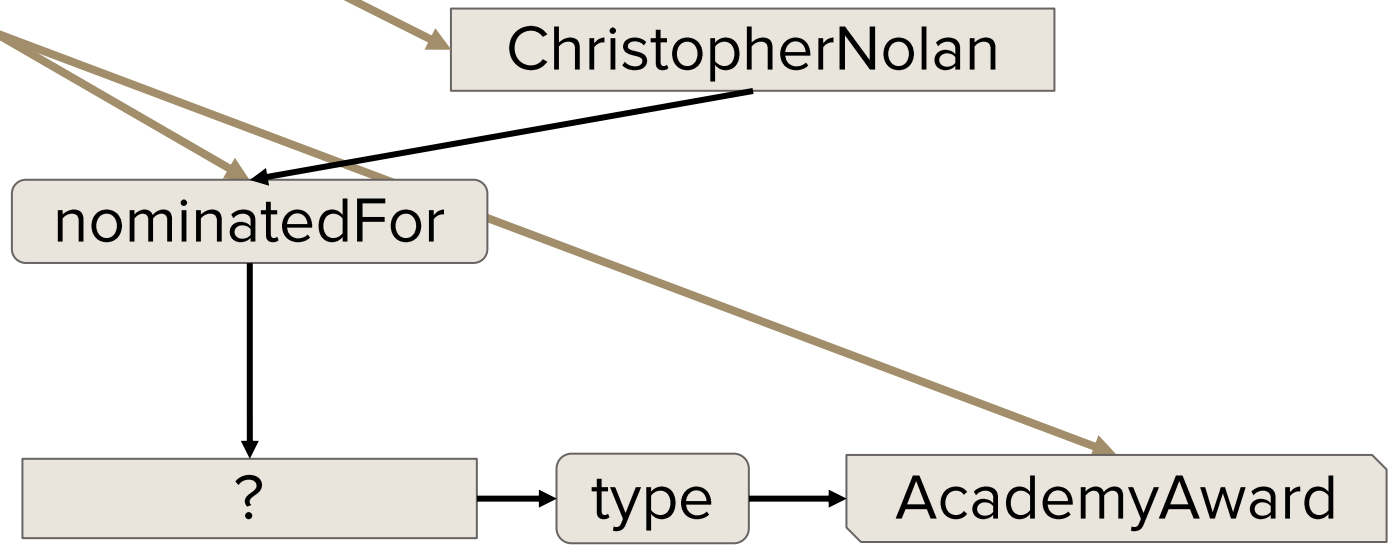
Named Entity Recognition (NER): Stanford NER, spaCy

Which Oscar nominations did Nolan receive?

<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>

SELECT ?ANS
WHERE {
ChristopherNolan nominatedFor ?ANS .
?ANS type AcademyAward }

SPARQL



Answering with query

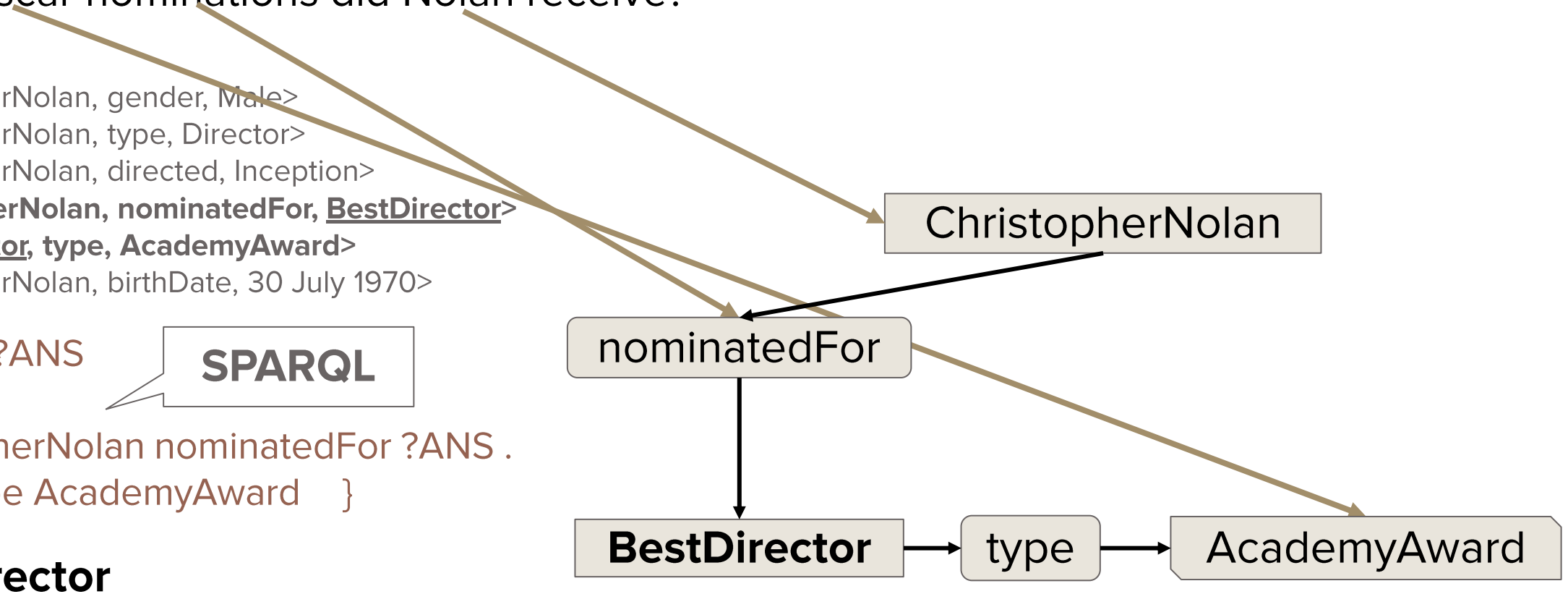
Which Oscar nominations did Nolan receive?

<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>

SELECT ?ANS
WHERE {
ChristopherNolan nominatedFor ?ANS .
?ANS type AcademyAward }

SPARQL

BestDirector



Structured queries and logical forms

Which Oscar nominations did Nolan receive?

Lambda-calculus $\lambda x.nominatedFor(ChristopherNolan, x) \wedge Type(x, AcademyAward)$

Lambda-DCS $nominatedFor.ChristopherNolan \sqcap type.AcademyAward$

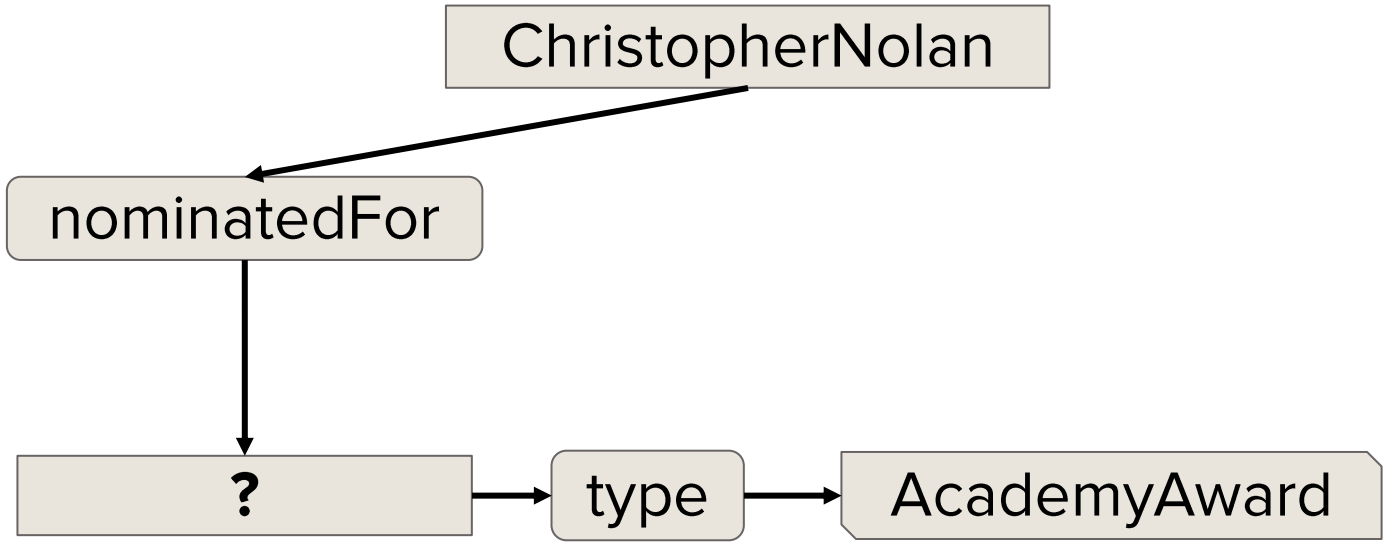
Neo4j CYPHER Graph QL

```
MATCH
(subj {name: 'ChristopherNolan' })-[:nominatedFor]-
>(obj:AcademyAward) RETURN obj.name
```

SPARQL BGP

```
SELECT ?ANS
WHERE {
ChristopherNolan nominatedFor ?ANS .
?ANS type AcademyAward     }
```

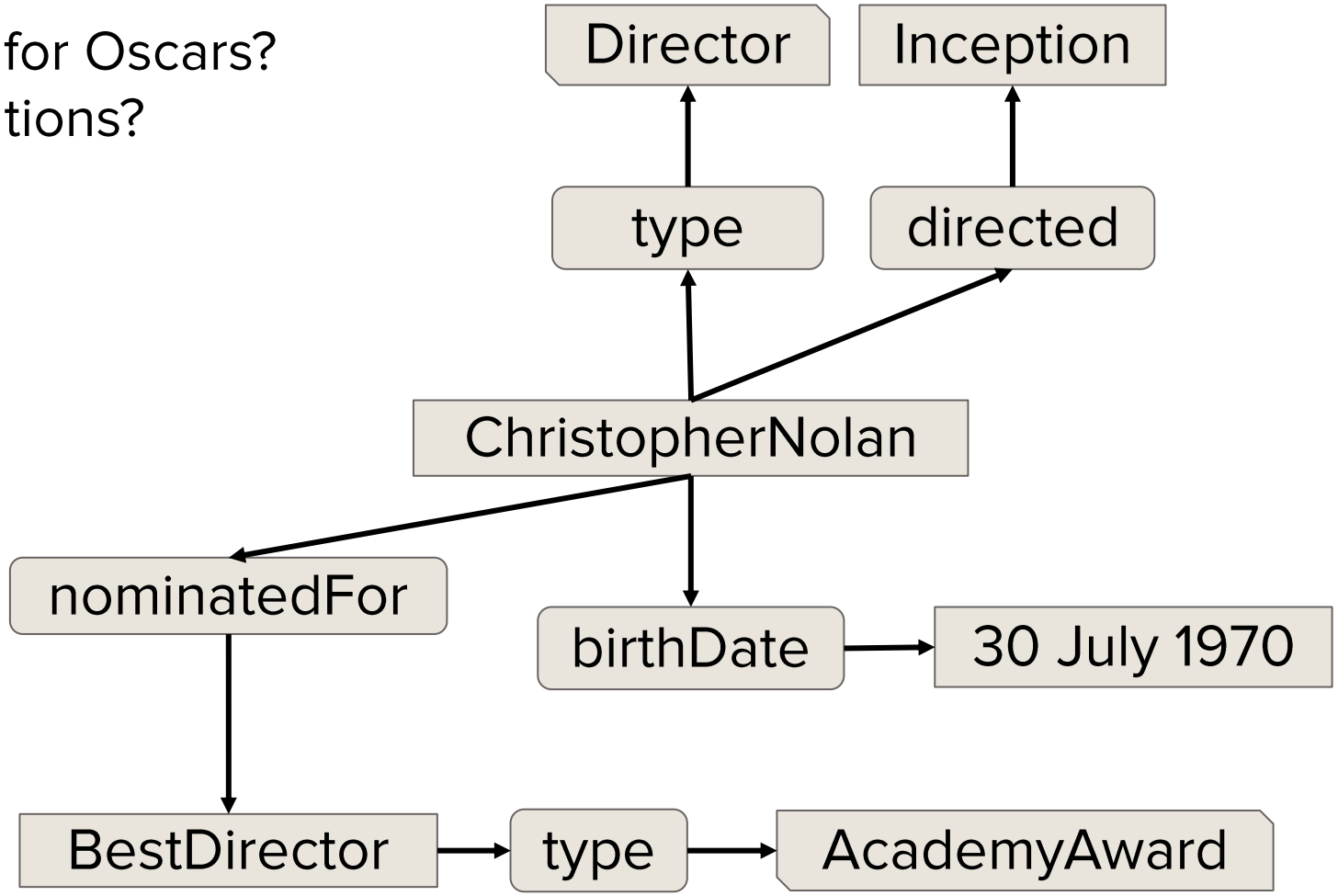
BestDirector



Reification: n-ary information in KGs

For which films was Nolan nominated for Oscars?
When did Nolan get his Oscar nominations?

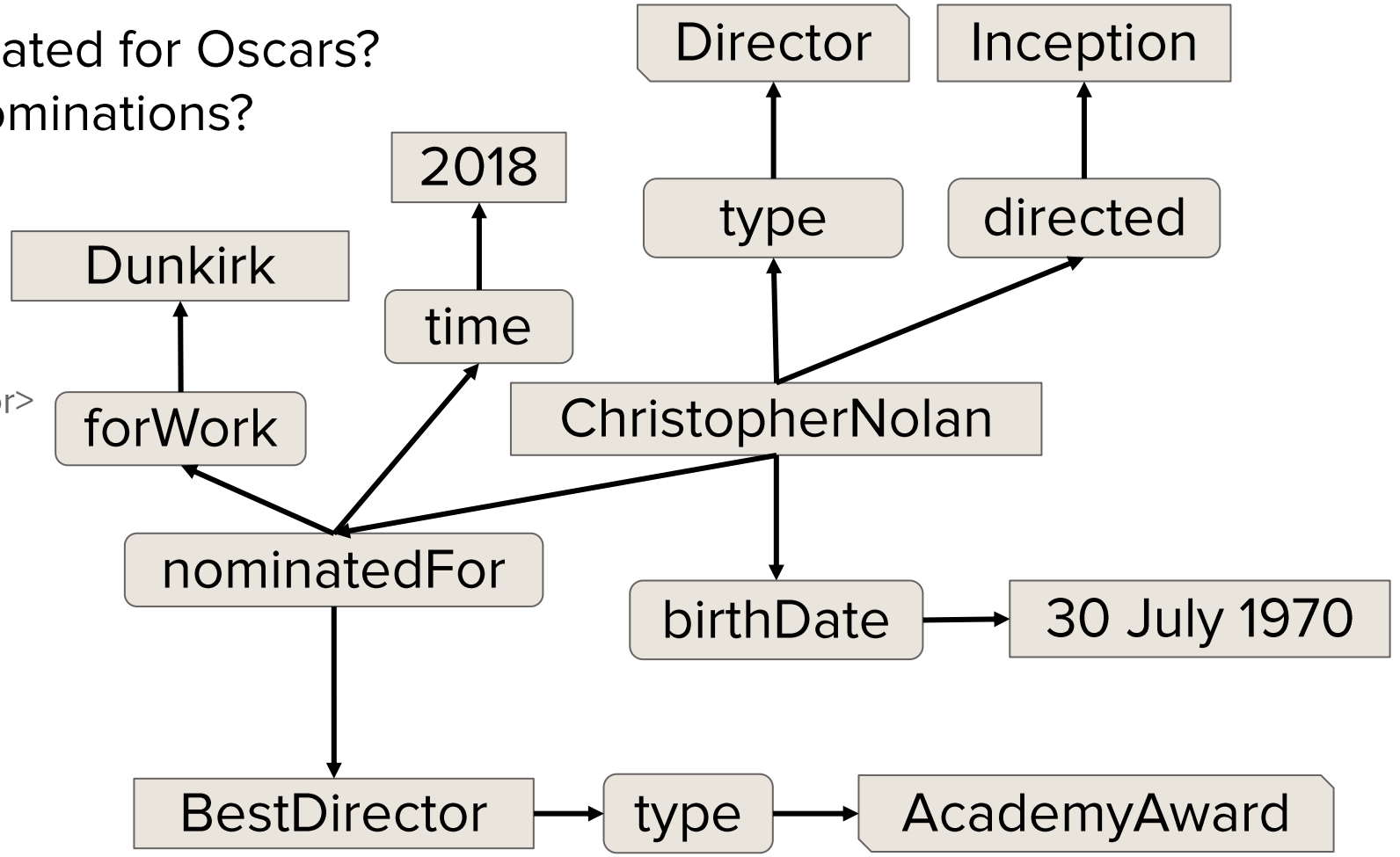
<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>



Reification: n-ary information in KGs

For which films was Nolan nominated for Oscars?
When did Nolan get his Oscar nominations?

<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>

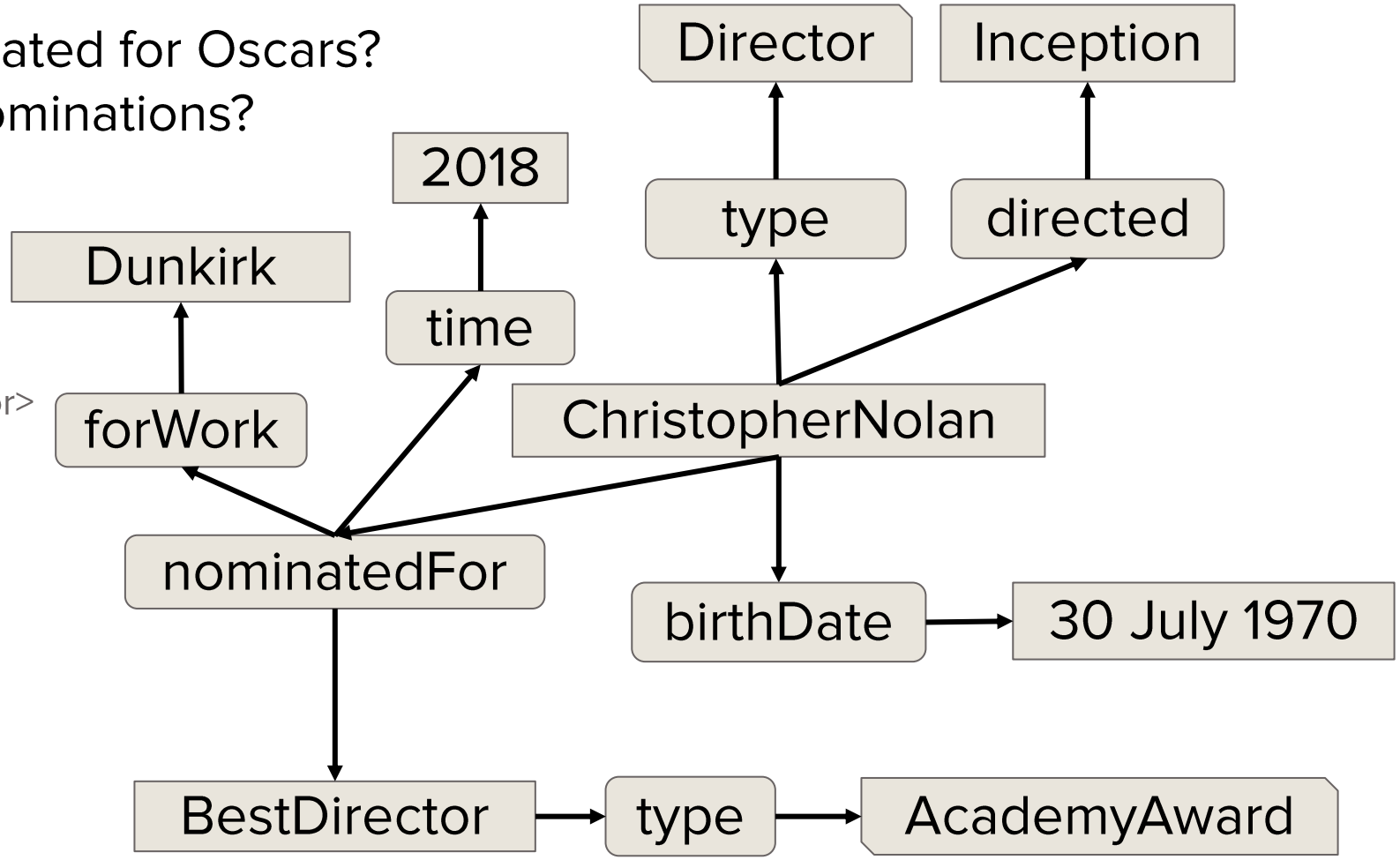


Reification: n-ary information in KGs

For which films was Nolan nominated for Oscars?
When did Nolan get his Oscar nominations?

<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>

<ChristopherNolan, nominatedFor, 123>
<123, nominatedFor, BestDirector>
<123, forWork, Dunkirk>
<123, year, 2018>



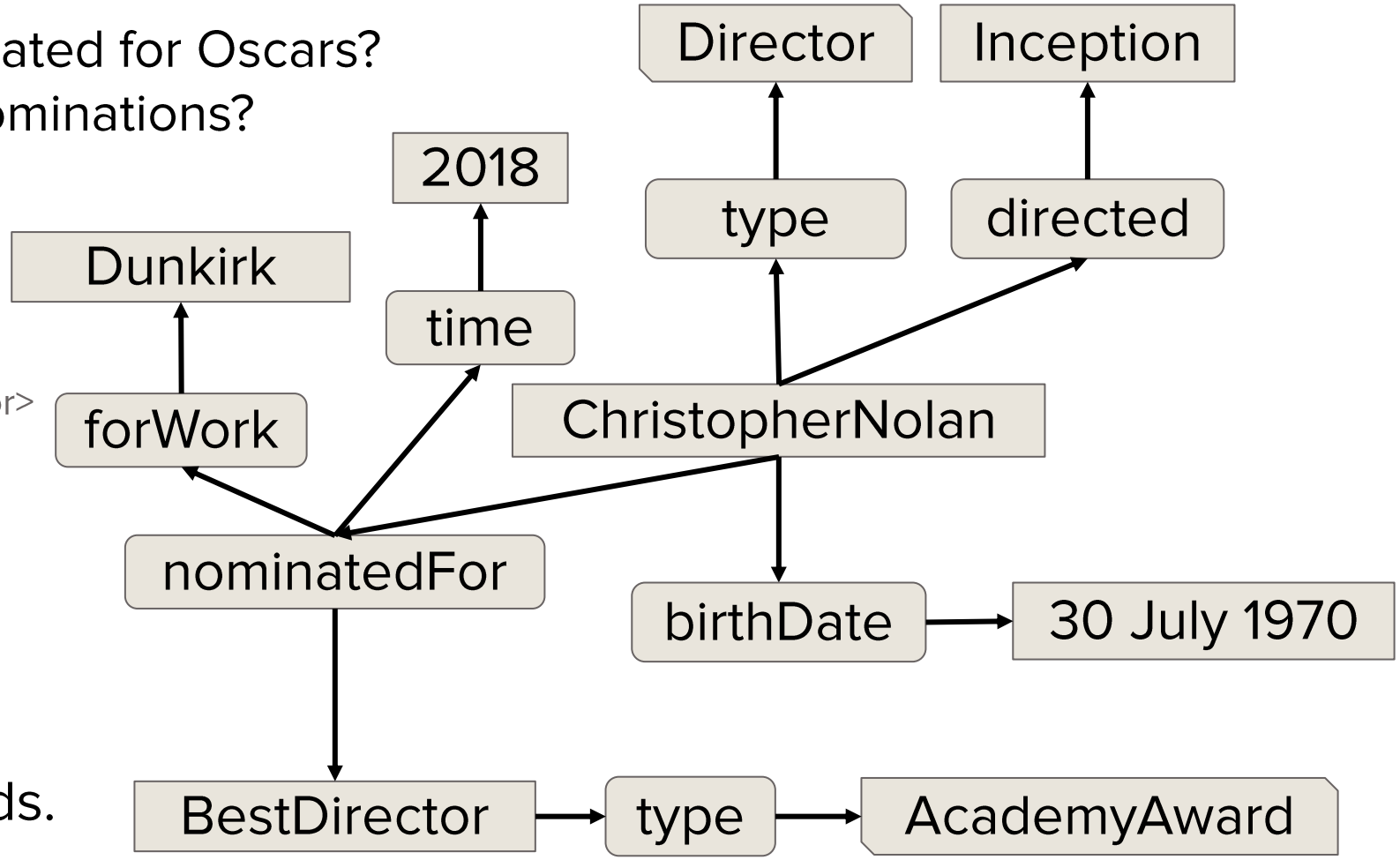
Qualifiers are a huge part of Wikidata

For which films was Nolan nominated for Oscars?
When did Nolan get his Oscar nominations?

<ChristopherNolan, gender, Male>
<ChristopherNolan, type, Director>
<ChristopherNolan, directed, Inception>
<ChristopherNolan, nominatedFor, BestDirector>
<BestDirector, type, AcademyAward>
<ChristopherNolan, birthDate, 30 July 1970>

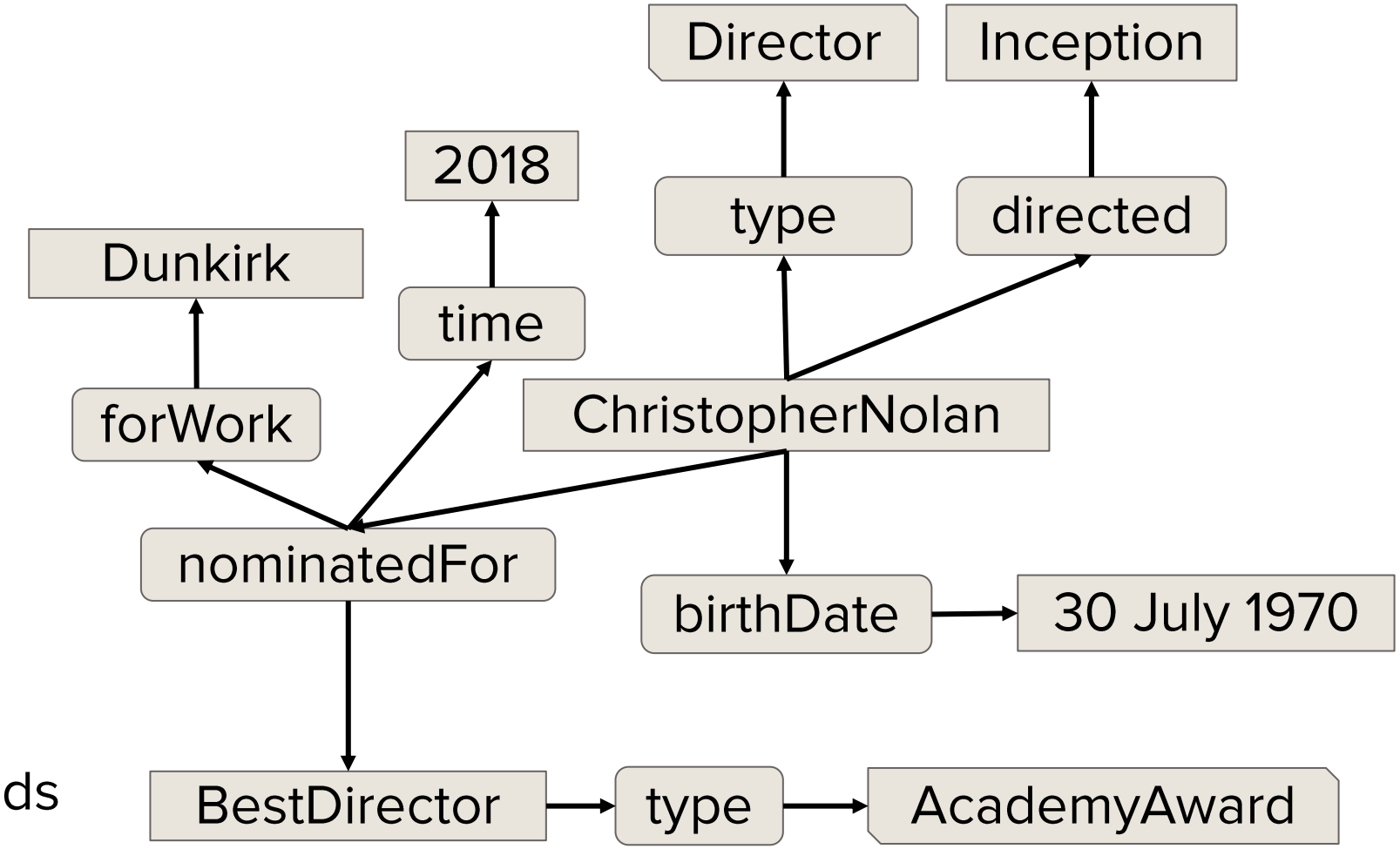
<ChristopherNolan, nominatedFor, 123>
<123, nominatedFor, BestDirector>
<123, forWork, Dunkirk>
<123, time, 2018>

Wikidata: Qualifiers, Statement-Ids.
6B triples part of reified facts!!



Questions that need reified triples

Who played Cobb in Inception?
Who did Leo play in Inception?
When did Neymar join PSG?
Who was Trump's first wife?
US president in 2016?
...



Wikidata: Qualifiers, Statement-Ids
6B triples part of reified facts!!

Explore Wikidata

Entity name / Subject

Entity id

The Dark Knight (Q163872)

Entity desc

2008 British-American superhero film directed by Christopher Nolan

TDK | Dark Knight

Entity aliases

Type

Type predicate

instance of	film

Predicate

genre	action film
director	Christopher Nolan

Object

nominated for	Academy Award for Best Supporting Actor	
	statement is subject of	81st Academy Awards
	nominee	Heath Ledger
	point in time	22 February 2009

part of the series	The Dark Knight Trilogy	
	follows	Batman Begins
	followed by	The Dark Knight Rises
	series ordinal	2
cast member	Christian Bale	
	character role	Bruce Wayne
	11 references	
	Michael Caine	
	character role	Alfred Pennyworth
	4 references	
	Heath Ledger	
	character role	Joker
	9 references	

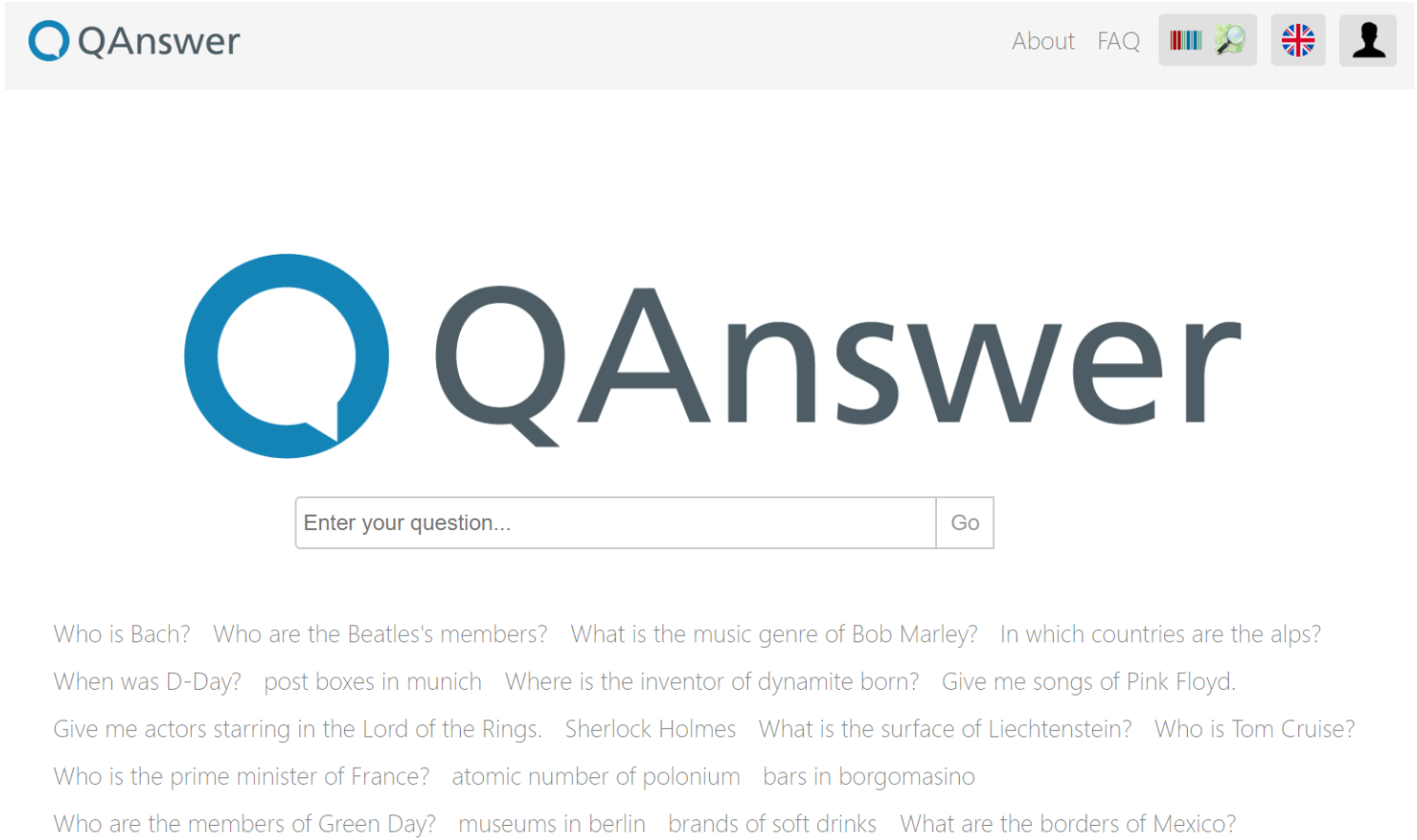
Qualifier predicate

Qualifier object

Explore Wikidata like a pro

- Wikidata: https://www.wikidata.org/wiki/Wikidata:Main_Page
- Wikidata data model: https://www.mediawiki.org/wiki/Wikibase/DataModel/Primer
- Wikidata dumps: https://www.wikidata.org/wiki/Wikidata:Database_download
- Download latest **n-triples** dump: https://dumps.wikimedia.org/wikidatawiki/entities/
- Wikidata SPARQL Endpoint: https://query.wikidata.org/
- Wikidata statistics: https://stats.wikimedia.org/#/wikidata.org
- More stats: https://www.wikidata.org/wiki/Wikidata:Statistics

Play with QA (over Wikidata)



Diefenbach et al., QAnswer: A Question answering prototype bridging the gap between a considerable part of the LOD cloud and end-users, WWW 2019. Available at https://qanswer-frontend.univ-st-etienne.fr/

Benchmarks

- **Simple questions**
  - WebQuestions (Berant et al. 2013) over Freebase
  - SimpleQuestions (Bordes et al. 2015) over Freebase
- **Complex questions**
  - LC-QuAD 2.0 (Dubey et al 2018) over Wikidata + DBpedia
  - MetaQA (Zhang et al. 2018) over Freebase
- **Conversational questions**
  - ConvQuestions (Christmann et al. 2019) over Wikidata
  - CSQA (Saha et al. 2018) over Wikidata

Benchmarks

- **Simple questions**
  -  WebQuestions* (Berant et al. 2013) over Freebase
    - SimpleQuestions (Bordes et al. 2015) over Freebase
- **Complex questions**
  -  LC-QuAD 2.0* (Dubey et al 2018) over Wikidata + DBpedia
    - MetaQA (Zhang et al. 2018) over Freebase
- **Conversational questions**
  -  ConvQuestions* (Christmann et al. 2019) over Wikidata
    - CSQA (Saha et al. 2018) over Wikidata

Recent benchmarks over Wikidata

More realistic benchmarks are smaller but harder

Much higher numbers on semi-synthetic benchmarks

“Vulnerable” to neural methods

* Need reified triples for answering

Benchmarks

- **Simple questions**
  -  WebQuestions* (Berant et al. 2013) over Freebase
    - SimpleQuestions (Bordes et al. 2015) over Freebase
- **Complex questions**
  -  LC-QuAD 2.0* (Dubey et al 2018) over Wikidata + DBpedia
    - MetaQA (Zhang et al. 2018) over Freebase
- **Conversational questions**
  -  ConvQuestions* (Christmann et al. 2019) over Wikidata
    - CSQA (Saha et al. 2018) over Wikidata

Many, many more:

LC-QuAD
(Trivedi et al. 2017)

ComQA
(Abujabal et al. 2019)

GraphQuestions
(Su et al. 2016)

QALD
(Usbeck et al. 2018)

TempQuestions
(Jia et al. 2018)

ComplexWebQuestions
(Talmor and Berant 2018)

WikiMovies
(Miller et al. 2016)

ComplexQuestions
(Bao et al. 2016)

Benchmarks: WebQuestions

- Real questions: Collected using the Google Suggest API
- Mostly simple questions using one fact or reified triple
- 3778 train, 2032 test questions
- Available at: https://nlp.stanford.edu/software/sempre/

who was richard nixon married to?

what high school did harper lee go to?

what was the capital city of the east roman empire?

who plays ken barlow in coronation street?

where is the fukushima daiichi nuclear plant located?

Benchmarks: LC-QuAD 2.0

- Sampled SPARQL queries via templates, verbalized by crowdworkers
- Complex (and simple) questions involving multiple entities and relations
- 23954 train, 6046 test questions
- Available at: http://lc-quad.sda.tech/

What city is the twin city of Oslo and also the setting for “A Tree Grows in Brooklyn”?

What Empire used to have Istanbul as its capital?

How long was Shirley Temple the United States Ambassador to Ghana?

Were Dutch and Hungarian the official languages of the Holy Roman Empire?

Who replaced Albus Dumbledore as headmaster of Hogwarts?

Benchmarks: ConvQuestions

- Natural conversations by crowdworkers after choosing topic
- Both simple and complex
- Five domains
- 6720 train, 2240 dev, 2240 test conversations
- Available at:

https://convex.mpi-inf.mpg.de/

Books	Movies	Soccer	Music	TV series
When was the first book of the book series The Dwarves published ?	Who played the joker in The Dark Knight?	Which European team did Diego Costa represent in the year 2018?	Led Zeppelin had how many band members?	Who is the actor of James Gordon in Gotham?
2003	Heath Ledger	Atlético Madrid	4	Ben McKenzie
What is the **name of the second book?**	When did he die?	Did they win the Super Cup the previous year?	Which was released first: Houses of the Holy or Physical Graffiti?	**What about Bullock?**
The War of the Dwarves	22 January 2008	No	Houses of the Holy	Donal Logue
Who is the author ?	**Batman actor?**	**Which club was the winner?**	**Is the rain song and immigrant song there?**	Creator?
Markus Heitz	Christian Bale	Real Madrid C.F.	No	Bruno Heller
In which city was he born ?	Director?	Which English club did Costa play for before returning to Atlético Madrid?	Who wrote those songs?	**Married to in 2017?**
Homburg	Christopher Nolan	Chelsea F.C.	Jimmy Page	Miranda Phillips Cowley
When was he born ?	**Sequel name?**	**Which stadium is this club's home ground?**	**Name of his previous band?**	**Wedding date first wife?**
10 October 1971	The Dark Knight Rises	Stamford Bridge Stadium	The Yardbirds	19 June 1993

Metrics

- Answers as sets (for systems using explicit structured queries)
  - Precision, Recall, F1-Score
- Answers as ranked lists (systems w/o explicit queries: approx. graph search)
  - Precision@1, MRR, MAP
  - Hit@5
- Single answer
  - Accuracy

Outline

- **Background:** Setup, benchmarks, metrics
- **Conversational QA:** Implicit context in multi-turn setup
- **Complex QA:** Multiple entities and predicates
- **Take-home:** Open problems and summary

How can we deal with information needs spread over multi-turn conversations?

Conversational KG-QA

Which actor voiced the character Unicorn in The Last Unicorn?

Mia Farrow

Which role was voiced by Alan Arkin in the Last Unicorn?

Schmendrick

Who performed the songs in the movie The Last Unicorn?

America

What is the genre of the band that performed the songs in The Last Unicorn?

Folk rock

Who was the director of the movie The Last Unicorn?

Jules Bass

Conversational KG-QA

Which actor voiced the character Unicorn in The Last Unicorn?

Mia Farrow

And Alan Arkin was behind ...?

Schmendrick

The songs were by...?

America

Genre of this band?

Folk rock

By the way, who directed the movie?

Jules Bass

Conversational KG-QA

- Information needs **rarely one-off**
- Sequence of **follow-up questions** on a topic
- Analogous to **search sessions** and **interactive retrieval**
- Users want to simulate **natural experience** with assistant
- **Leave context unspecified** in follow-ups

Conversational KG-QA

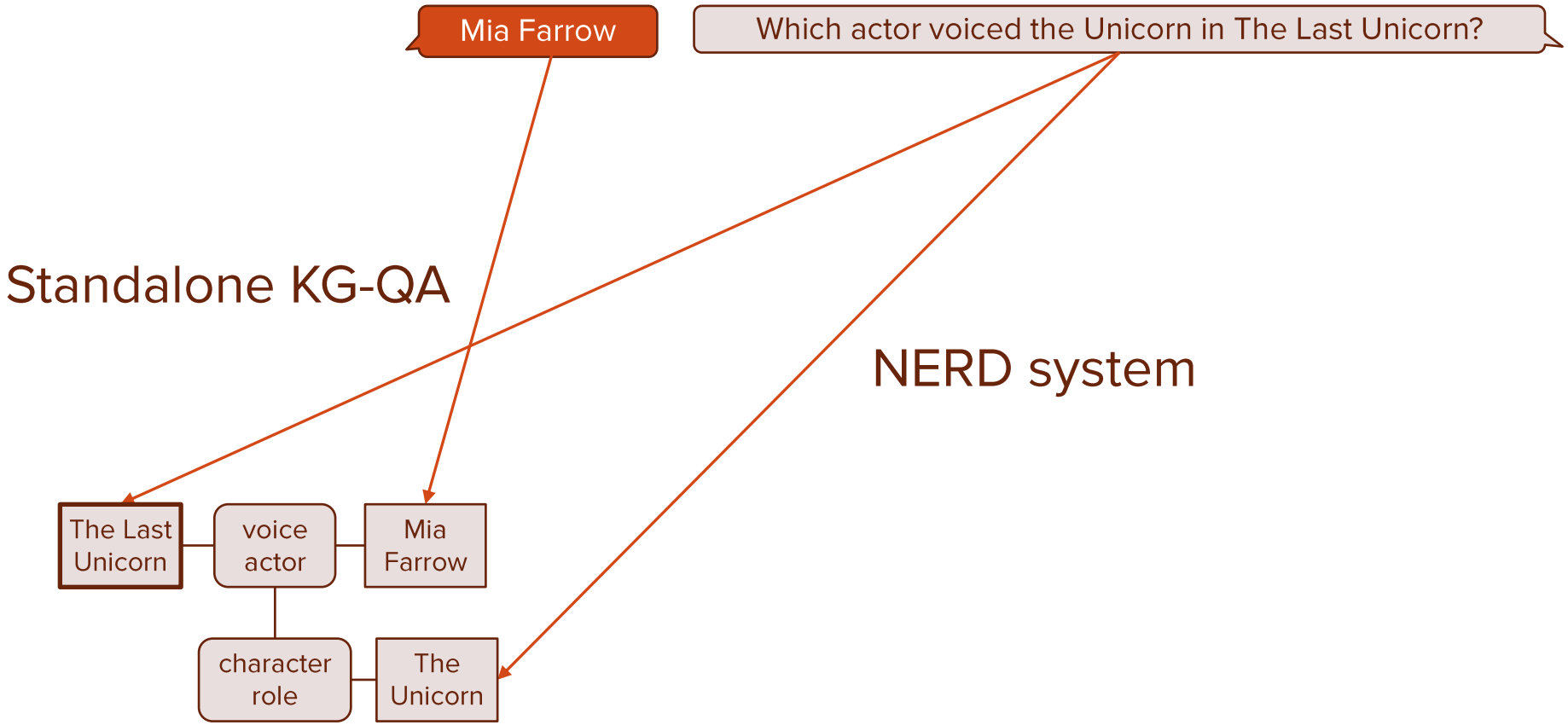
- **Key challenges** in conversational (KG-)QA
  - Infer implicit context
  - Handle ad hoc formulations
- Initially explored over **small tables** as sequential QA (Iyyer et al. 2017)
- Key direction for KG-QA now (Saha et al. 2018, Guo et al. 2018, **Christmann et al. 2019**, Shen et al. 2019)

Conversational QA: Graph traversal

- The **CONVEX** system (Christmann et al. 2019)
- Large topic jumps in conversations are rare: establish **localized KG context**
- Harness **KG-connectivity:** No need to complete/rewrite questions
- **Expand context judiciously** with relevant entities and predicates in neighborhood
- **Unsupervised** iterative graph traversal (c.f. supervised graph traversal in PullNet)
- CONVEX works on top of **any KG-QA system** to handle conversations

Christmann et al., Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion, CIKM 2019.

Initial context

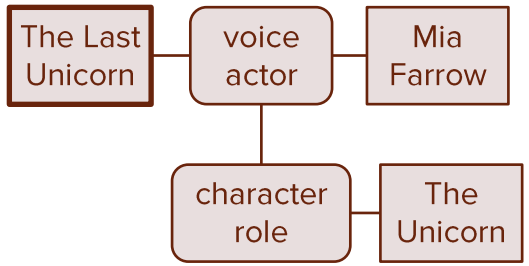


Initial context

Mia Farrow

Which actor voiced the Unicorn in The Last Unicorn?

And Alan Arkin was behind . . .?



How to expand the context?

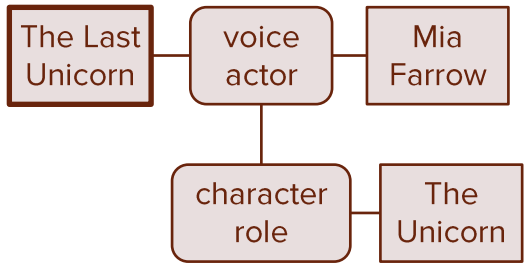

Neighborhood of
Mia Farrow
Neighborhood of
The Last Uni
borhood of
Unicorn

Judicious context expansion

Mia Farrow

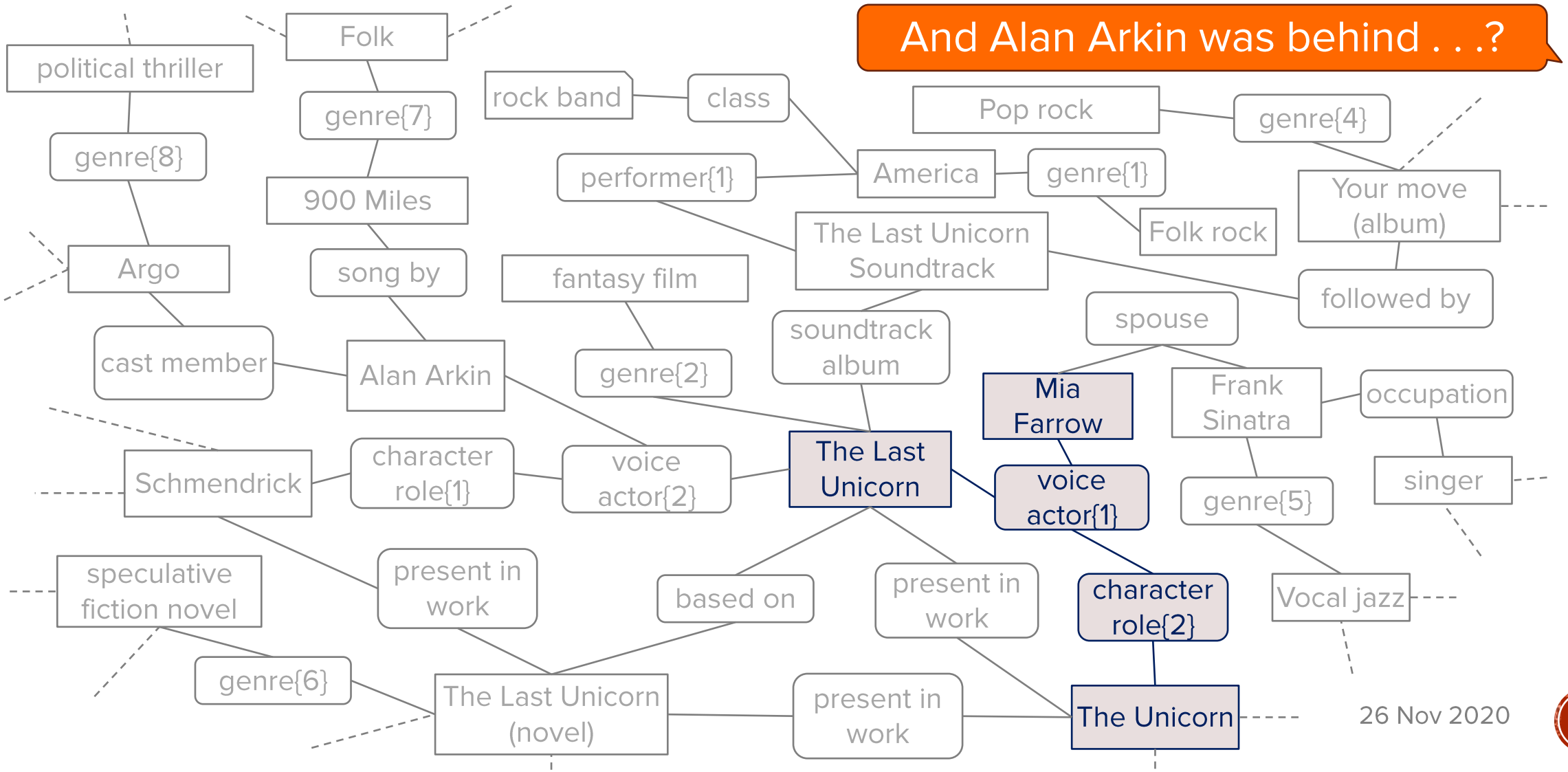
Which actor voiced the Unicorn in The Last Unicorn?

And Alan Arkin was behind . . .?

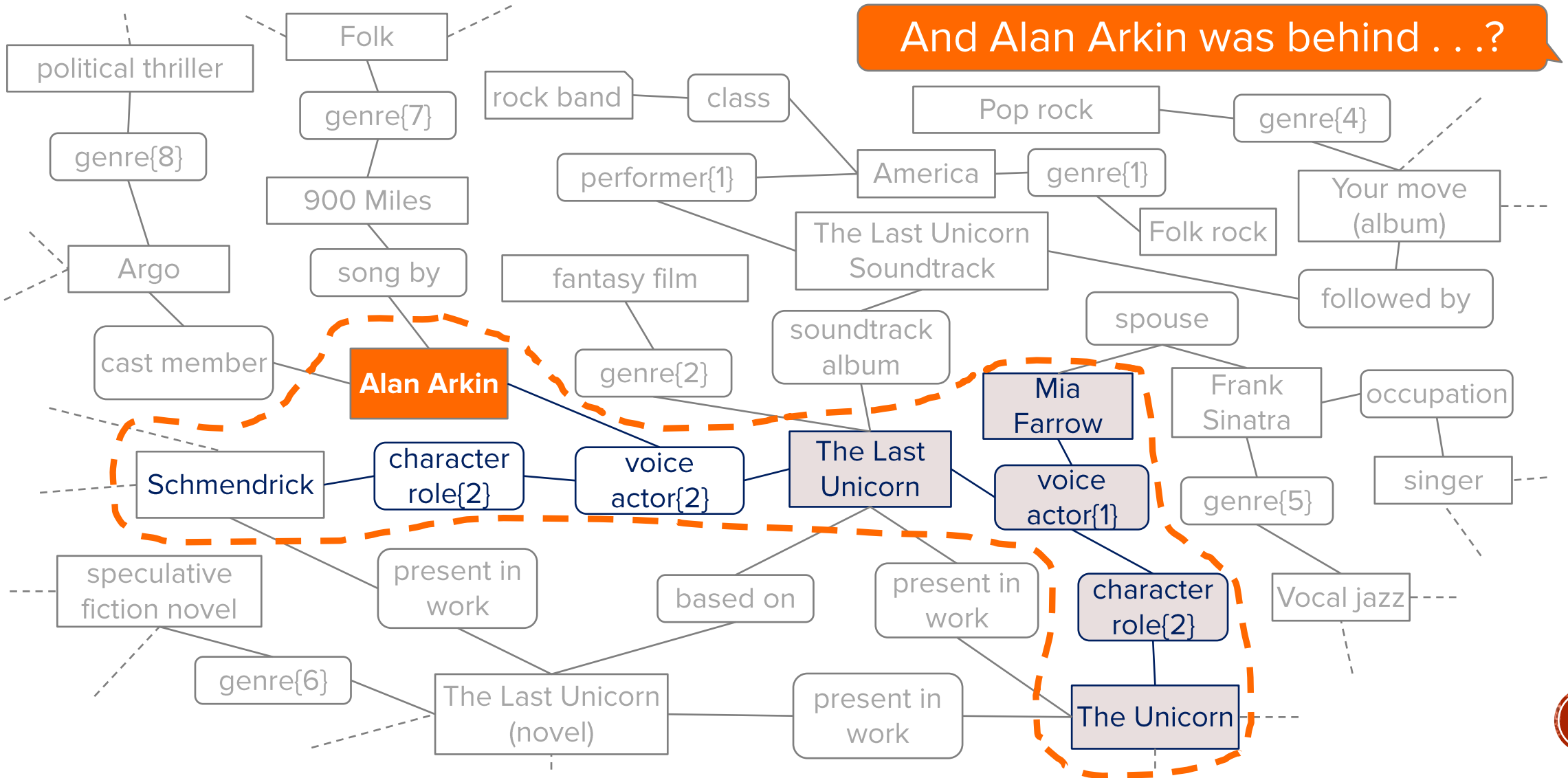


Do not expand with the complete neighborhood!

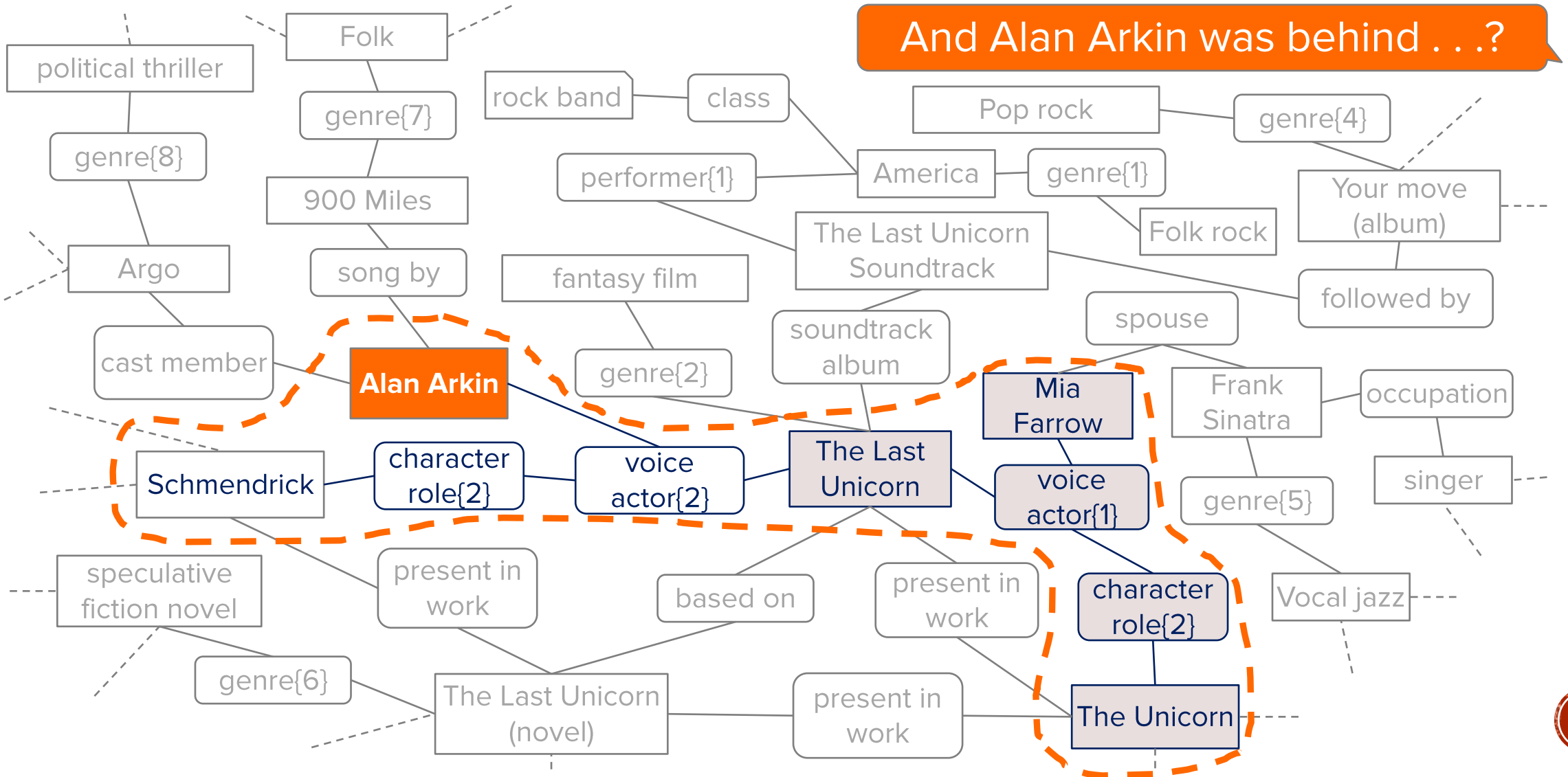
Exploring context neighborhood



Exploring context neighborhood



Find **frontier nodes** to define expansion border

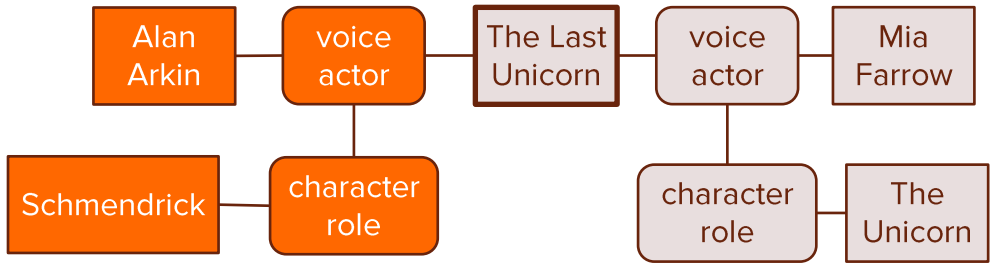


Context graph

Mia Farrow

Which actor voiced the Unicorn in The Last Unicorn?

And Alan Arkin was behind . . .?



Graph expanded with relevant facts only

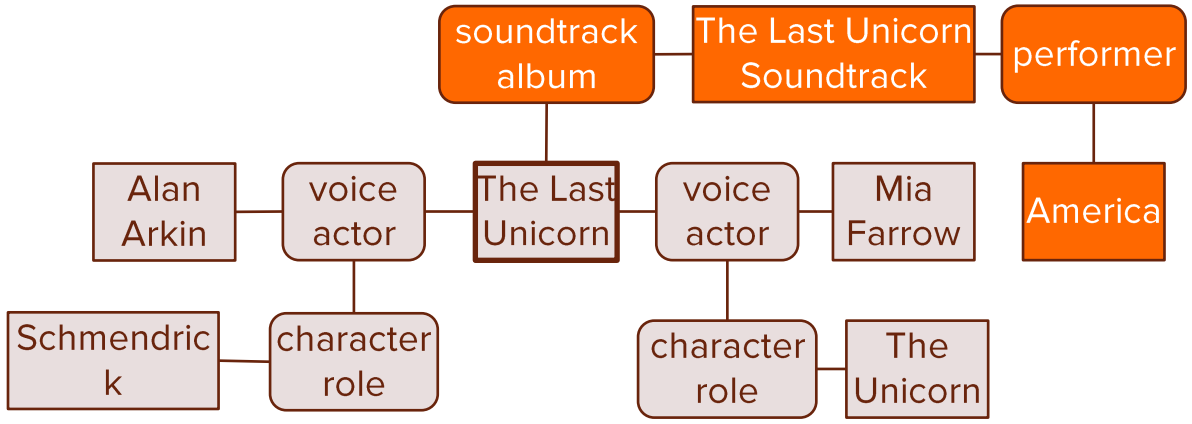
Context graph

Mia Farrow

Which actor voiced the Unicorn in The Last Unicorn?

And Alan Arkin was behind . . .?

So who performed the songs?



Graph expanded with relevant facts only

Context graph

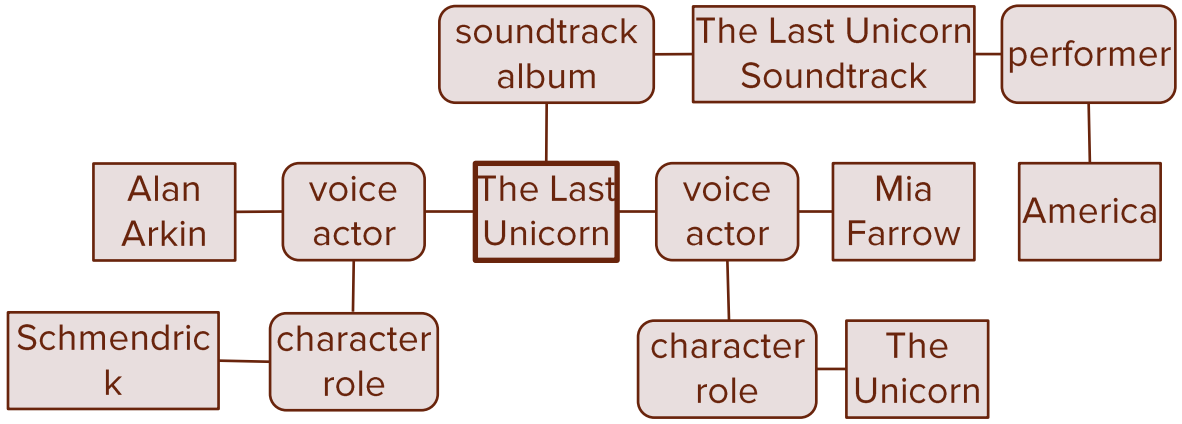
Mia Farrow

Which actor voiced the Unicorn in The Last Unicorn?

And Alan Arkin was behind . . .?

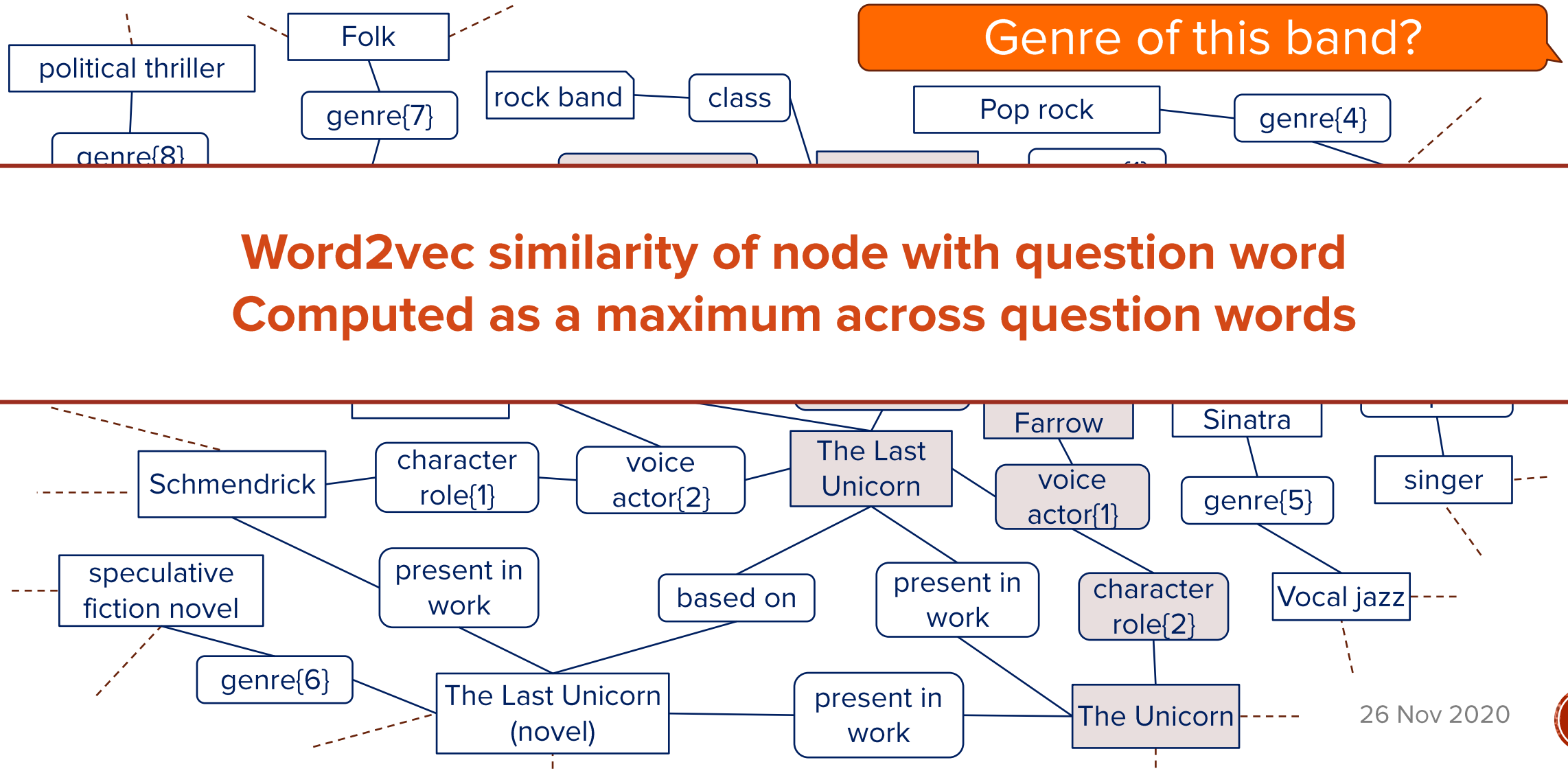
So who performed the songs?

Genre of this band?

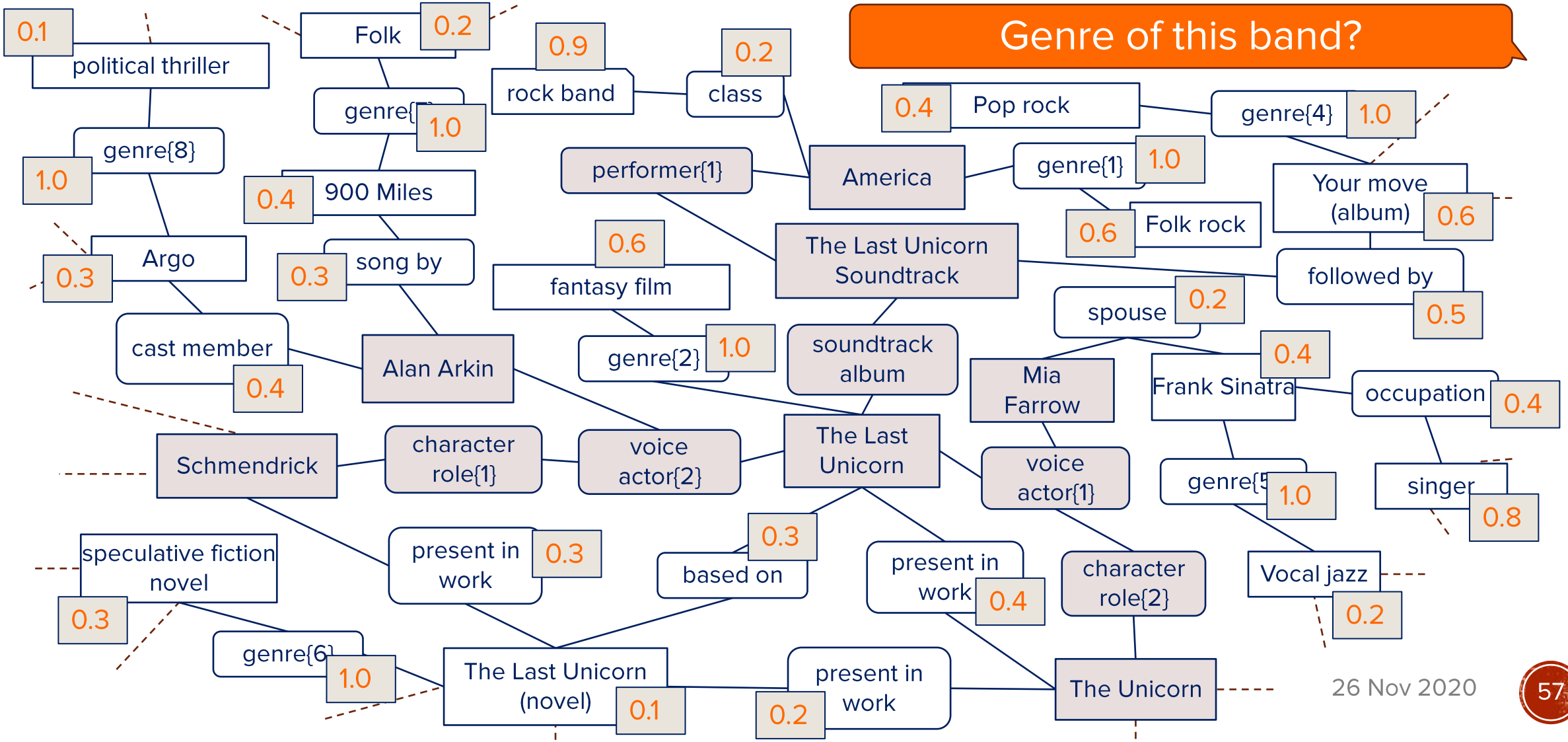


How to determine Frontier nodes?

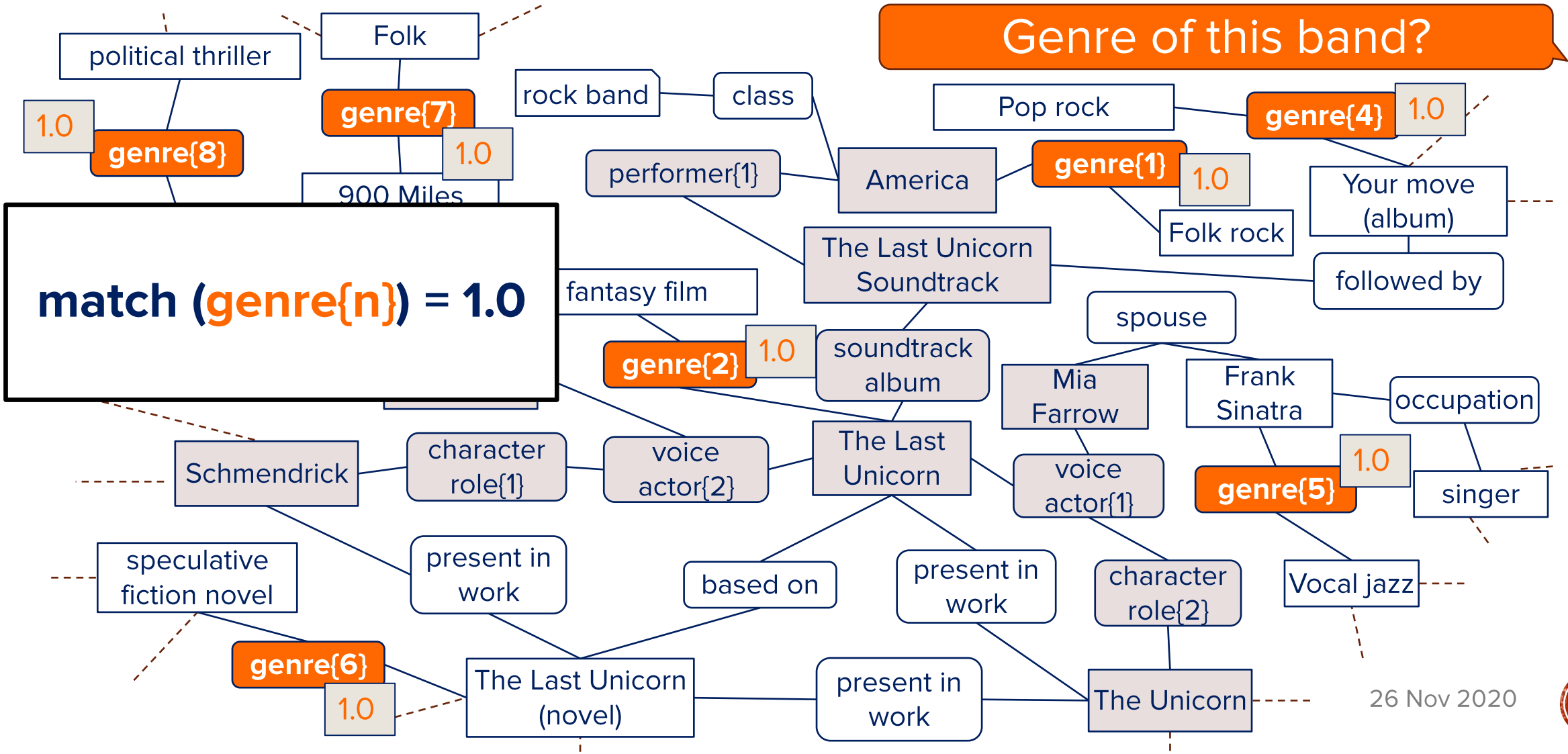
Relevance to the question



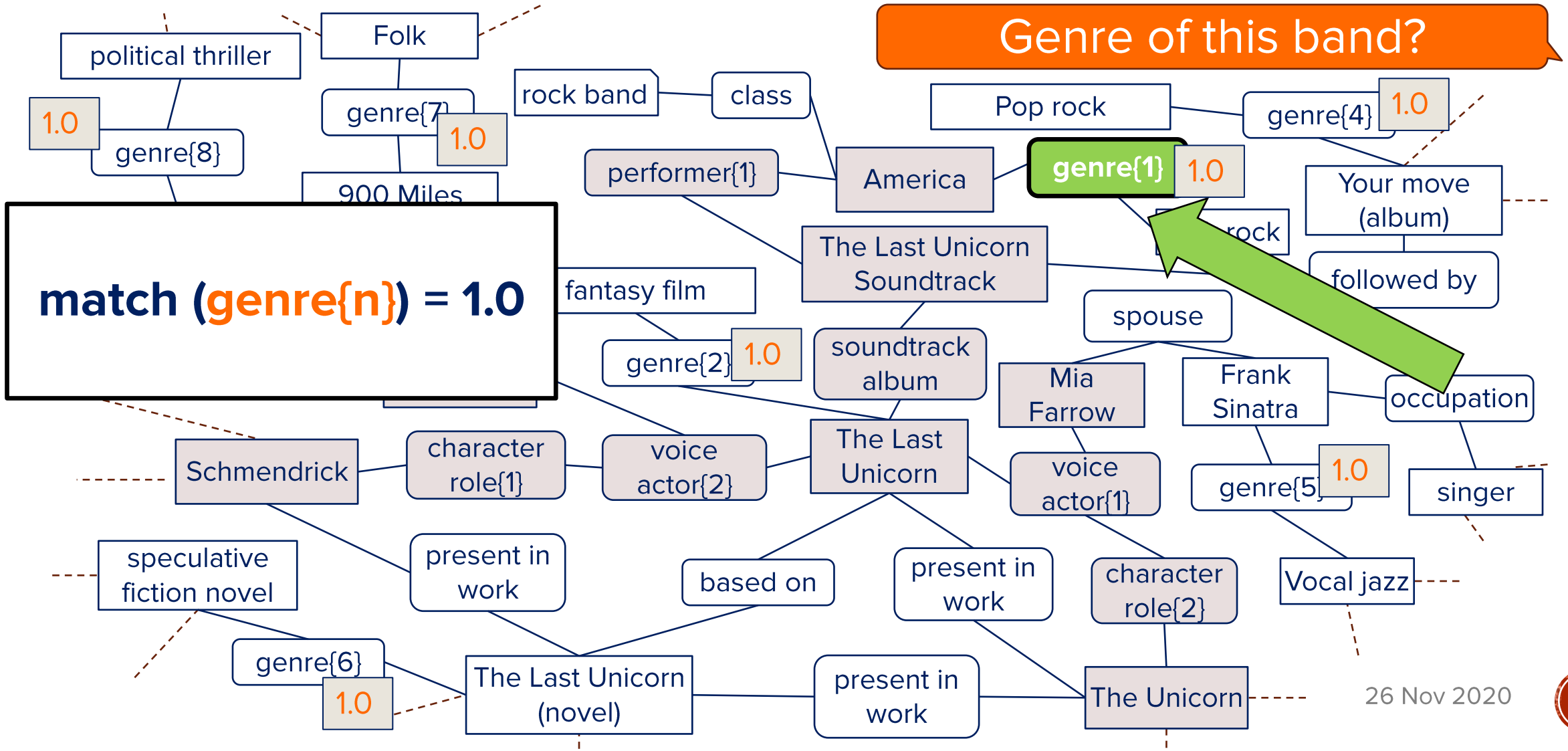
Relevance to the question



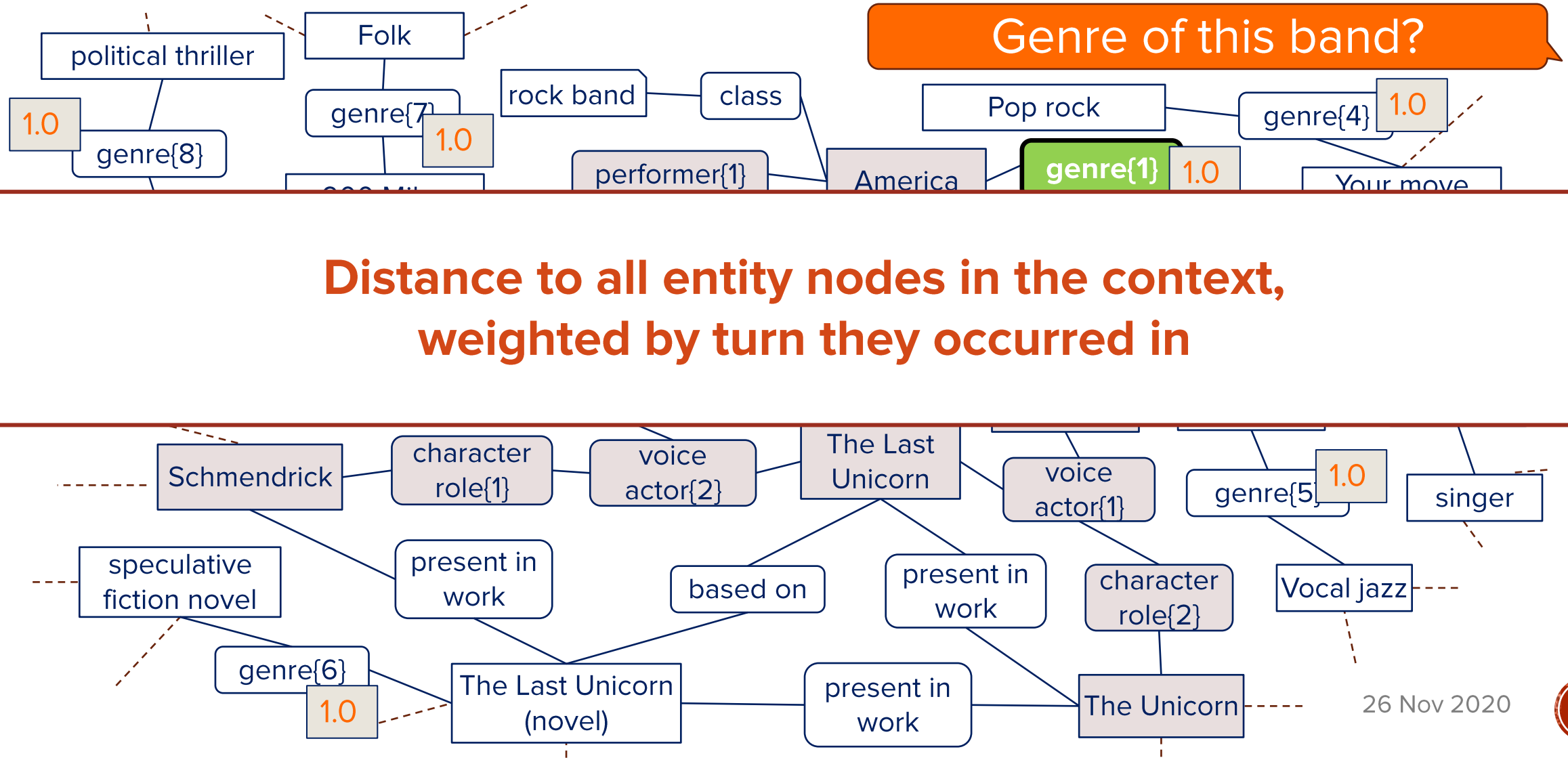
The great disambiguation



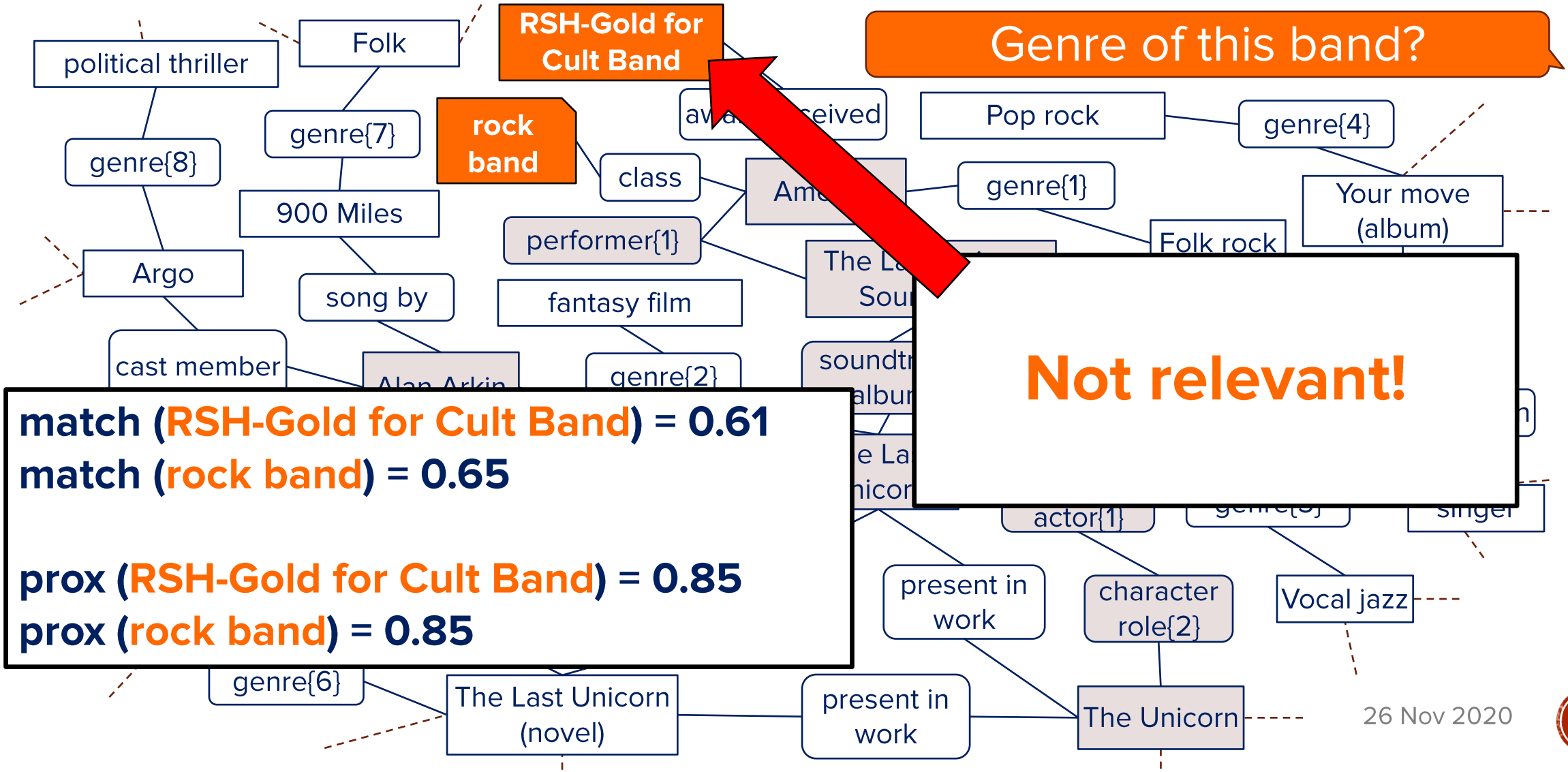
Relevance to the context



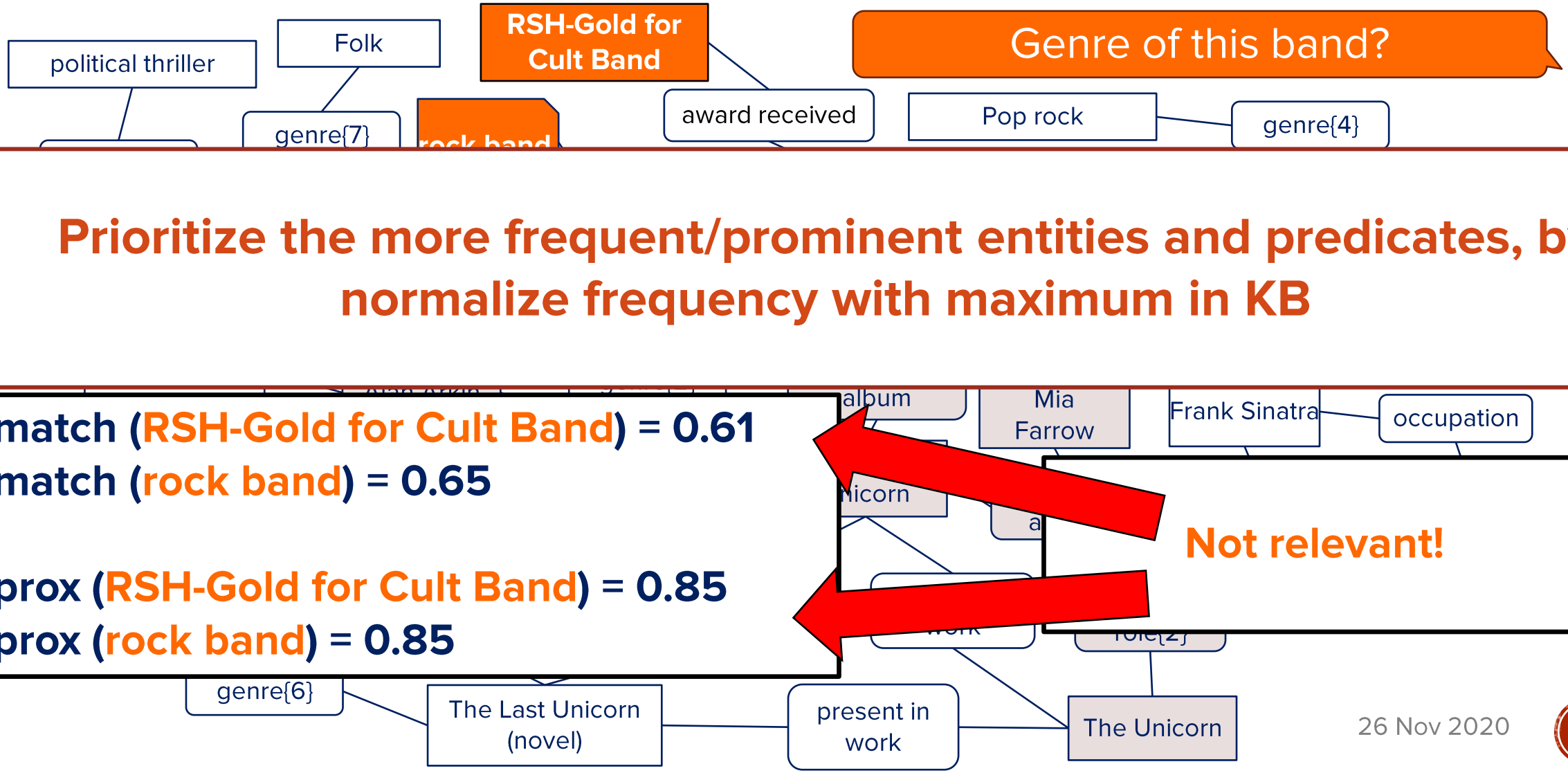
Relevance to the context



KG priors


RSH-Gold for Cult Band
Genre of this band?
rock band
political thriller
Folk
genre{8}
genre{7}
900 Miles
Argo
song by
cast member
class
performer{1}
fantasy film
genre{2}
Pop rock
genre{4}
genre{1}
Your move (album)
Folk rock
Not relevant!
match (RSH-Gold for Cult Band) = 0.61
match (rock band) = 0.65
prox (RSH-Gold for Cult Band) = 0.85
prox (rock band) = 0.85
actor{1}
present in work
character role{2}
Vocal jazz
genre{6}
The Last Unicorn (novel)
present in work
The Unicorn

KG priors


political thriller
Folk
RSH-Gold for Cult Band
Genre of this band?
award received
Pop rock
genre{4}
genre{7}
rock band
Prioritize the more frequent/prominent entities and predicates, by normalize frequency with maximum in KB
match (RSH-Gold for Cult Band) = 0.61
match (rock band) = 0.65
prox (RSH-Gold for Cult Band) = 0.85
prox (rock band) = 0.85
Not relevant!
album
Mia Farrow
Frank Sinatra
occupation
genre{6}
The Last Unicorn (novel)
present in work
The Unicorn

Frontier score

Matching similarity	**Context relevance**	**KG priors**
match (candidate c)	*prox (candidate c)*	*prior (candidate c)*

$$frontier\_score(candidate\ c) = h_1 \cdot match(c) + h_2 \cdot prox(c) + h_3 \cdot prior(c)$$

With hyperparameters h_1, h_2, h_3

Frontier nodes

Matching similarity

Candidate	*Match*
genre{1}	1.00
genre{2}	1.00
…	…
folk rock band	0.89
RSH-Gold for Cult Band	0.87
fantasy film	0.36
…	…

Context relevance

Candidate	*Prox*
genre{1}	0.91
folk rock band	0.86
RSH-Gold for Cult Band	0.86
…	…
genre{2}	0.34
fantasy film	0.36
…	…

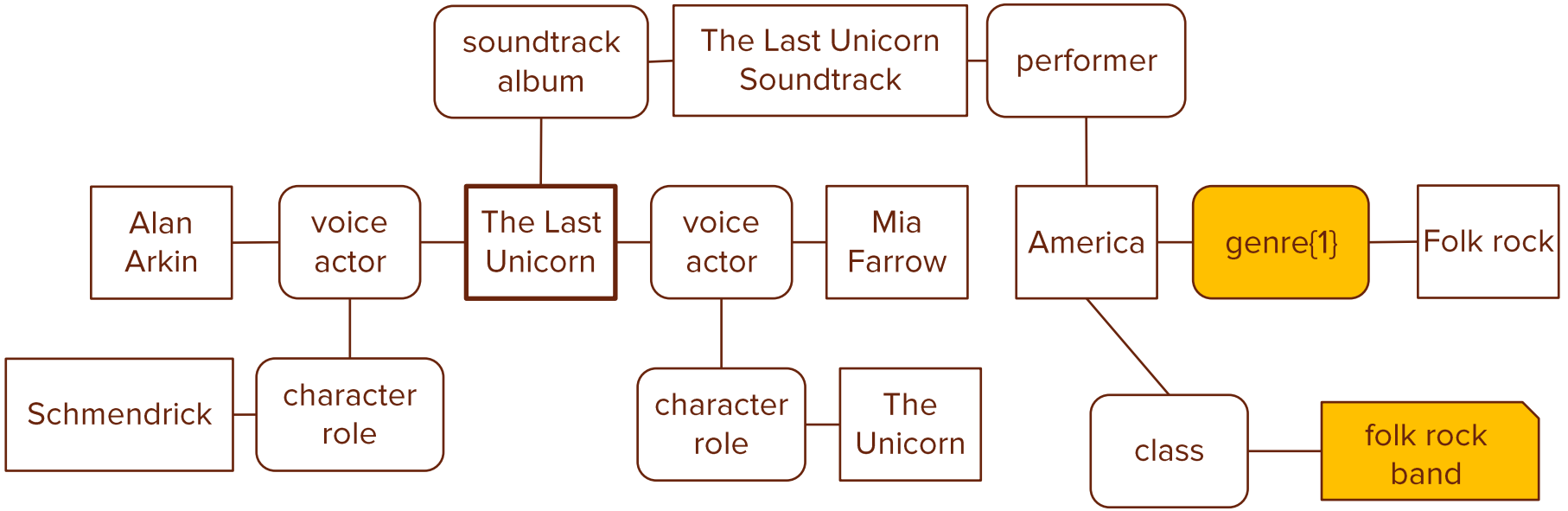
KG priors

Candidate	*KG priors*
…	…
genre{1}	0.56
genre{2}	0.56
…	…
folk rock band	0.34
…	…
RSH-Gold for Cult Band	0.01

Fagin's Threshold Algorithm (FTA) to retrieve top-*k* ranked nodes according to frontier score

Frontier nodes

Genre of this band?



Frontier nodes

Answer to the question

Genre of this band?

- Distance to Frontier nodes
  - Weighted by the frontier score
  - distance_F

=> Explicit part

- Distance to all nodes in context graph *X*
  - Weighted by the turn they occurred in
  - distance_X

=> Implicit part

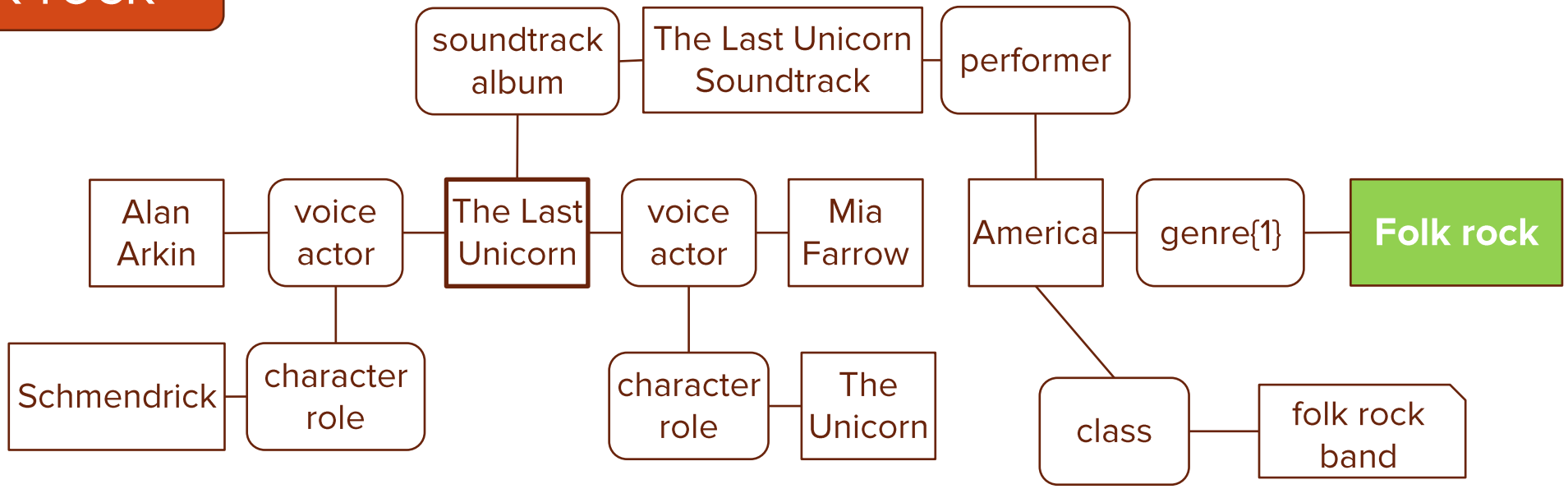
$$answer\_score(candidate\ c) = h_4 \cdot distance\_F + h_5 \cdot distance\_X$$

Christmann et al., Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion, CIKM 2019.

Answer detection

Genre of this band?

Folk rock

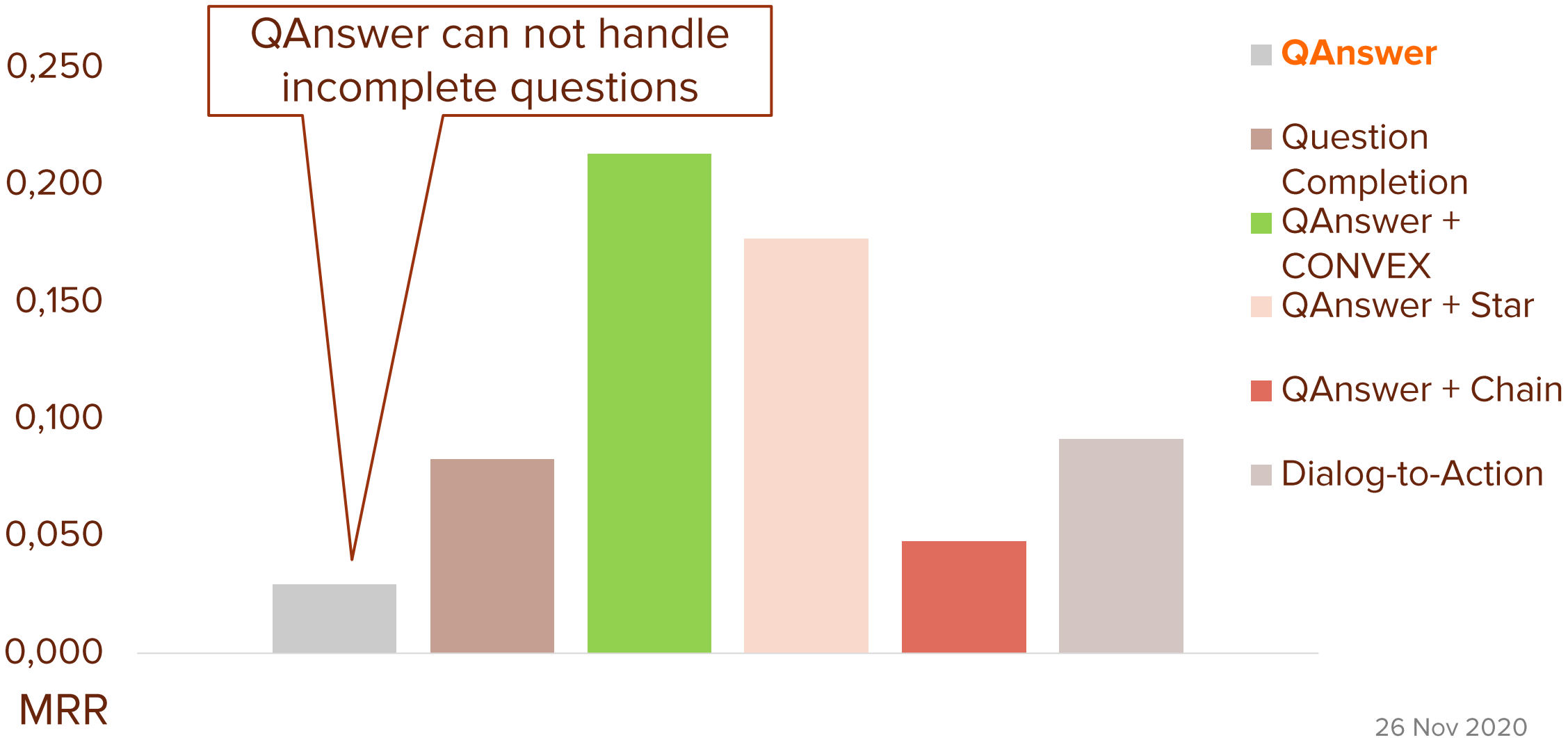


Top-ranked node according to *answer_score*

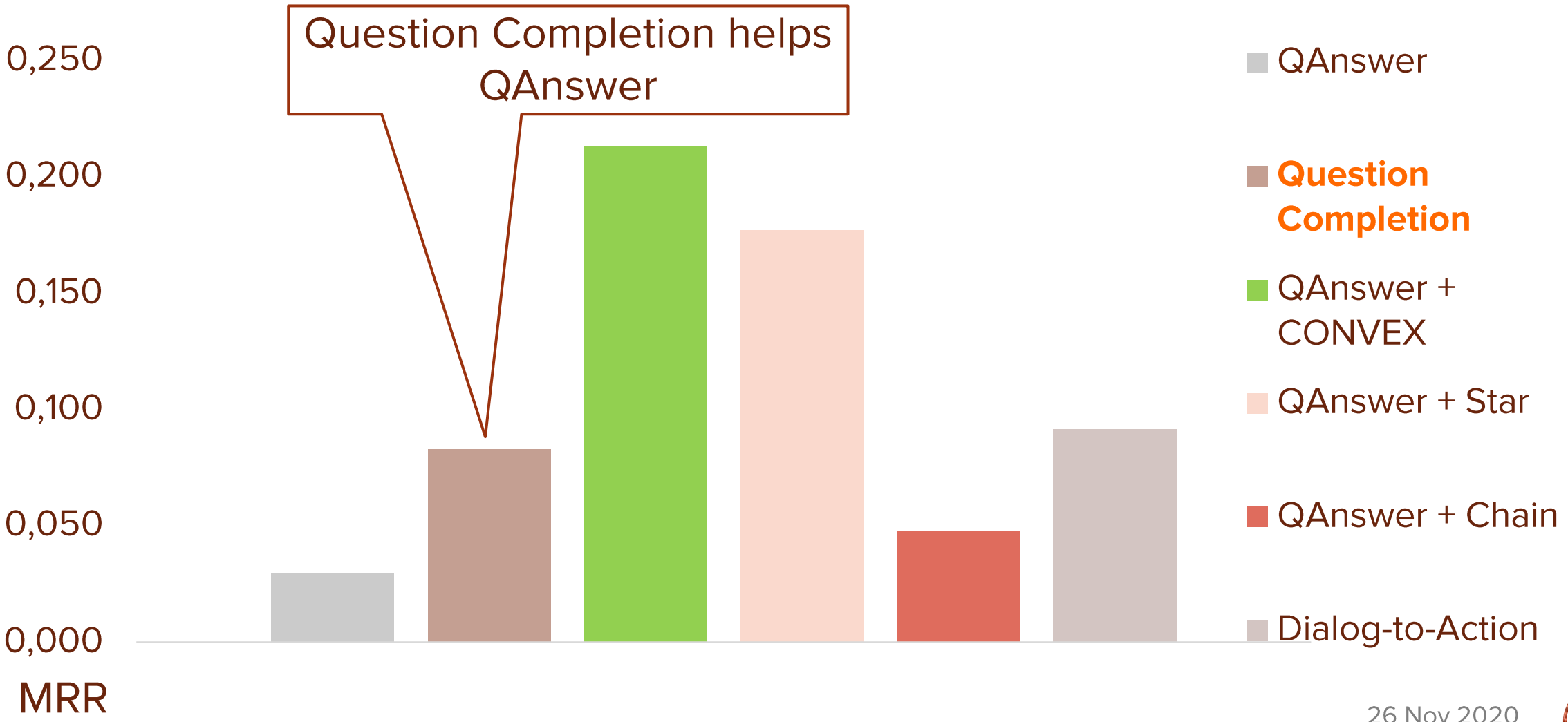
Evaluation

- Benchmark: ConvQuestions
- Metrics: **MRR**, P@1, Hit@5
- Baselines
  - Qanswer
  - Question completion
  - Star and chain models
  - Dialog-to-Action (Guo et al. 2018)

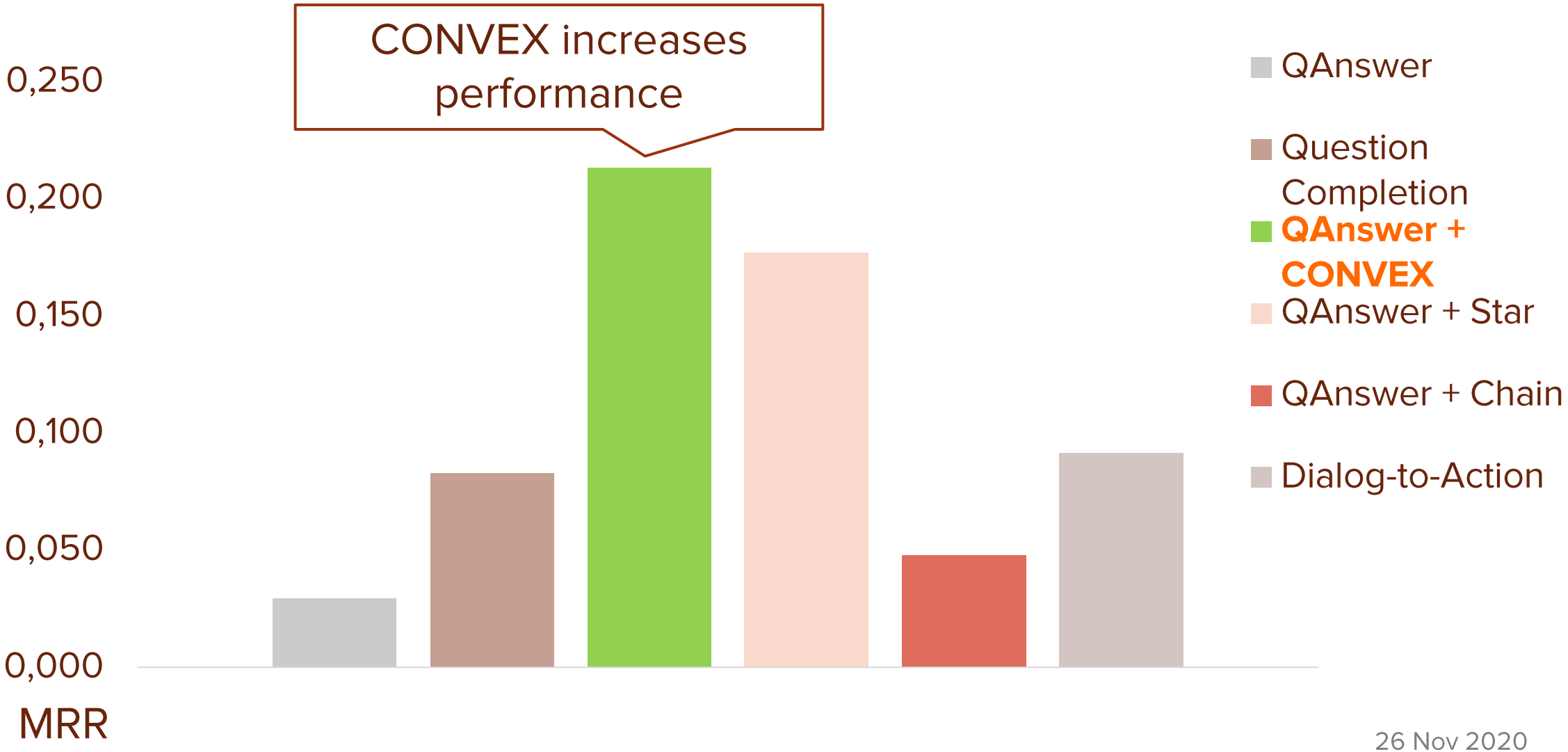
Experimental results



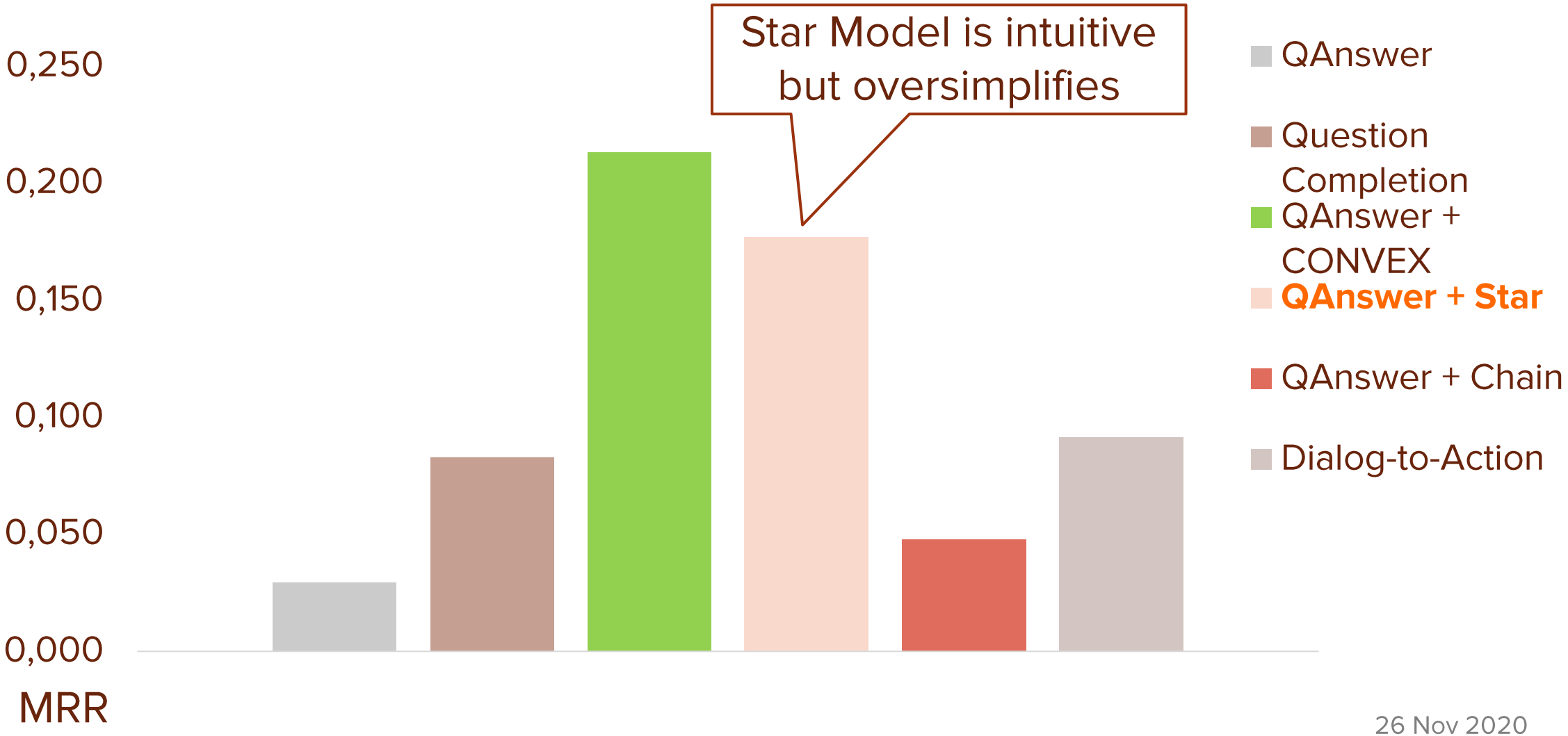
Experimental results



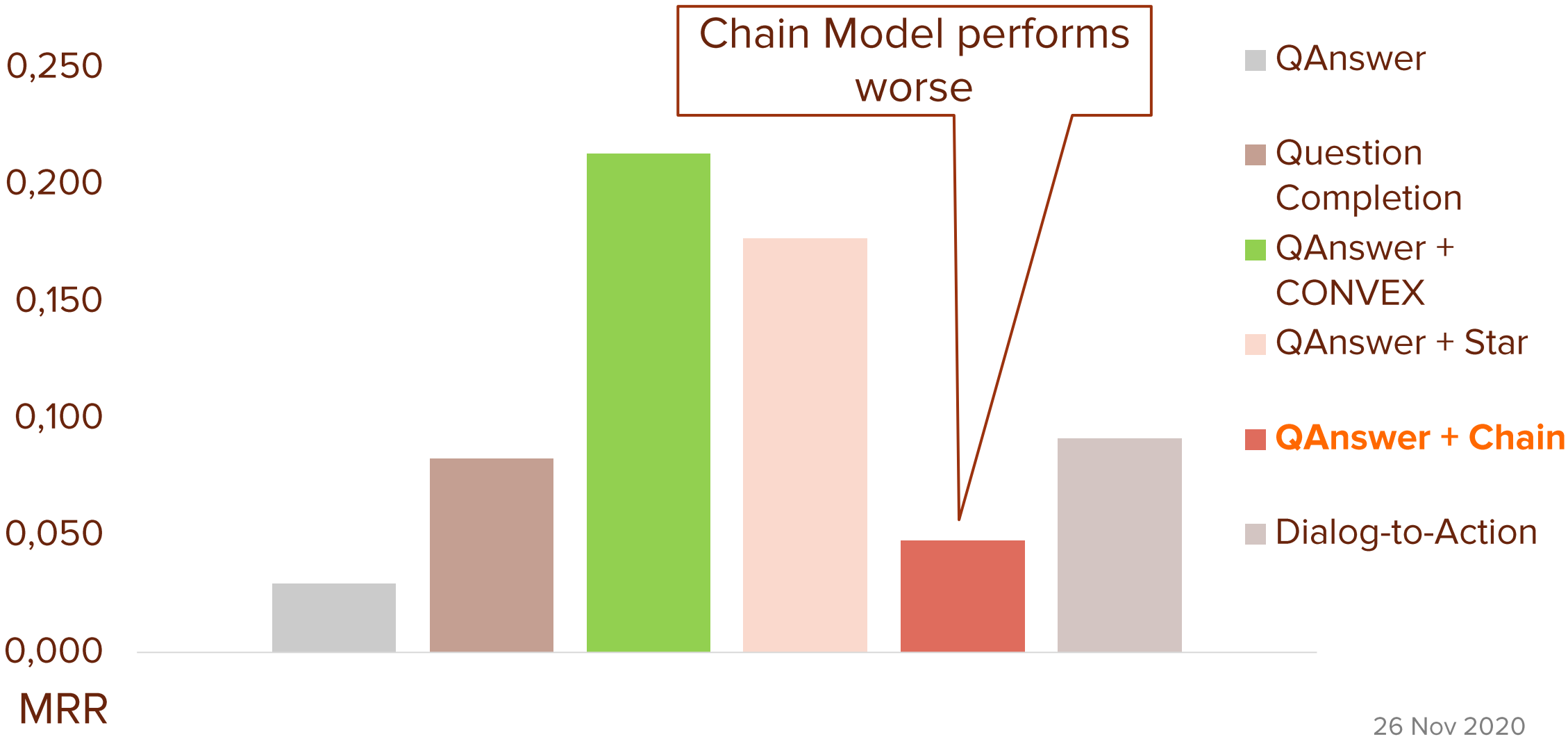
Experimental results



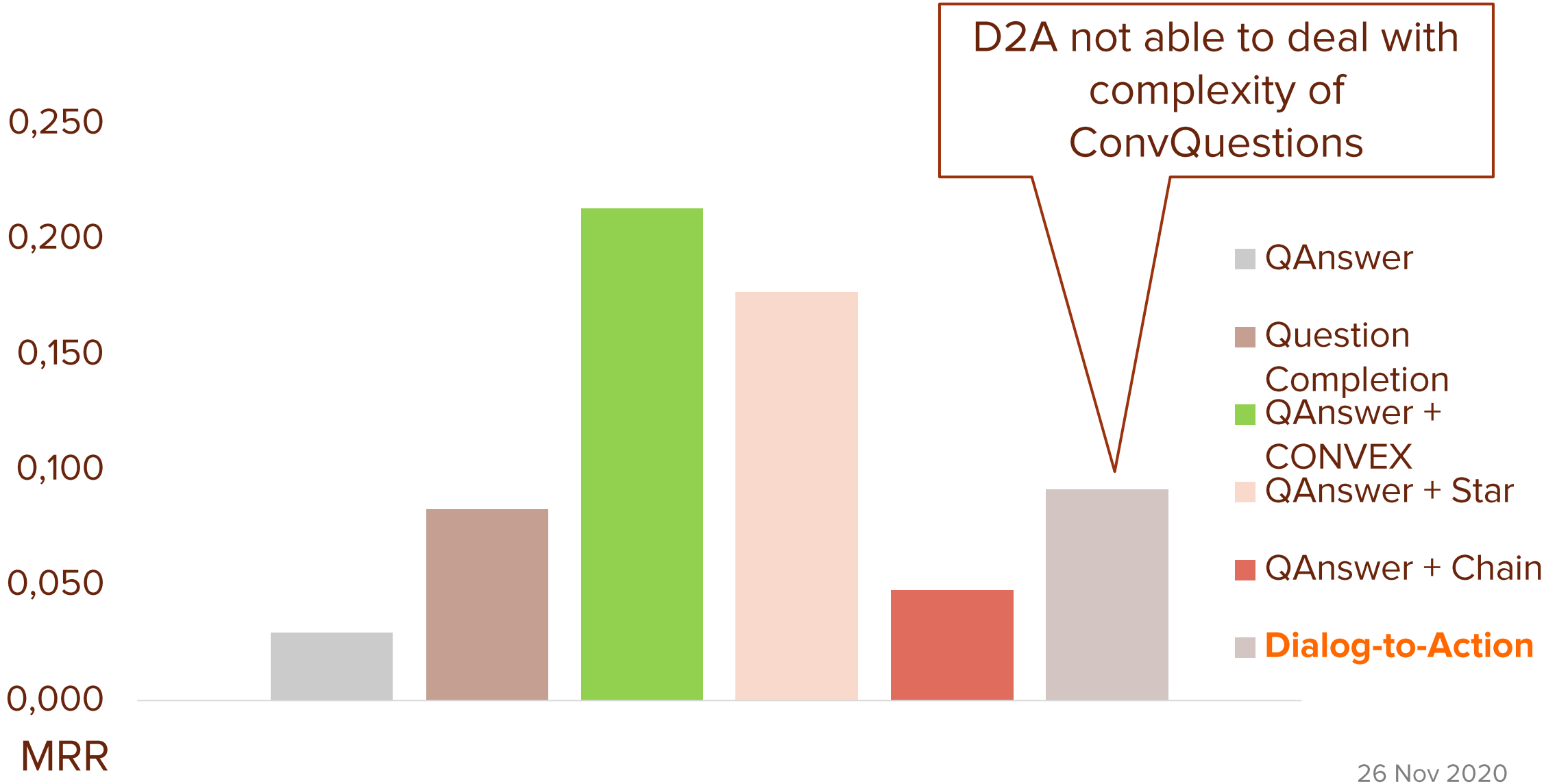
Experimental results



Experimental results



Experimental results



Contributions

- CONVEX
  - First unsupervised system for ConvKG-QA
  - Enables any standalone KG-QA with conversational support
  - Judicious context expansion enables efficient answering
- ConvQuestions
  - First realistic benchmark on Conversational KG-QA
  - 11,200 conversations from 5 domains

Outline

- **Background:** Setup, benchmarks, metrics
- **Conversational QA:** Implicit context in multi-turn setup
- **Complex QA:** Multiple entities and predicates
- **Take-home:** Open problems and summary

How can we answer more complex questions with multiple entities and predicates?

Complex questions

- Two basic types
  - Star joins
    - Who played for Barcelona and Real Madrid?
  - Chain joins
    - What is the profession of Messi's father?

```
SELECT ?x WHERE
?x playedFor Barcelona .
?x playedFor RealMadrid .
```

Single variable

```
SELECT ?y WHERE
?x fatherOf Messi .
?x profession ?y .
```

Two or more variables

Complex questions

- **Much more:** Aggregations, comparatives, superlatives, reasoning, existential, temporal, ...
- Focus on **substructures** in questions and queries (Bhutani et al. 2019, Ding et al. 2019, Sun et al. 2020)
- Often rely on question **decomposition** (Bao et al. 2016, Talmor and Berant 2018, Sun et al. 2020)
- **Joint disambiguation** of question concepts (Yahya et al. 2012, Lu et al. 2019)

Which female **actor played in Casablanca and is married to** a writer who was born in Rome?

Where is the **founder of Tesla born**?

Who was the **second wife** of Tom Cruise?

Which **Portuguese speaking countries** import **fish from Brazil**?

Who wrote **more books:** Enid Blyton or Agatha Christie?

Which is the **third highest** mountain in Asia?

How many **movies have the same director** as The Shawshank Redemption?

How many movies were directed by the graduate of Burbank High School?

Did any cosmonauts die in the same place they were born in?

Complex questions

- Early efforts in Yahya et al. (2012)
- Further explorations in Bao et al. (2016), Abujabal et al. (2017) and Cui et al. (2017)
- Dedicated methods for complex questions in Ding et al. (2019), Hu et al. (2018), Luo et al. (2018), Bhutani et al. (2019), **Lu et al. (2019)**, Vakulenko et al. (2019), …

Which female **actor played in Casablanca and is married to** a writer who was born in Rome?

Where is the **founder of Tesla born**?

Who was the **second wife** of Tom Cruise?

Which **Portuguese speaking countries** import **fish from Brazil**?

Who wrote **more books:** Enid Blyton or Agatha Christie?

Which is the **third highest** mountain in Asia?

How many **movies have the same director** as The Shawshank Redemption?

How many movies were directed by the graduate of Burbank High School?

Did any cosmonauts die in the same place they were born in?

Complex QA: Computing compact subgraphs

- The **QUEST** system (Lu et al. 2019)
- Works over open vocabulary quasi KGs (Bhutani et al. 2019b, Yin et al. 2015, Fader et al. 2013, 2014)
- Augment quasi KGs with alignments and types
- Spot question cornerstones in quasi KG
- **Unsupervised compact subgraph computation:** Compute Group Steiner Tree (GST) with cornerstones as terminals for joint disambiguation of question concepts

Lu et al., Answering Complex Questions by Joining Multi-Document Evidence with Quasi Knowledge Graphs, SIGIR 2019.

Complexity in information needs

Question: Which Nolan films won an Oscar but missed a Golden Globe?

Answers: Inception, Interstellar

- **Join evidence** from multiple documents on-the-fly

Join evidence from multiple docs on-the-fly

Question: Which Nolan films won an Oscar but missed a Golden Globe?

The 2011 **Oscar** award just announced that Inception is the **winner** of the Best Sound Editing award. Other **winners** of the day…

Nolan directed the **movie** Inception and other **science thrillers**…

Inception narrowly **lost to** The Social Network for Best Screenplay at the 68th **Golden Globe** Awards, which were declared in the afternoon..

This illustrates "joining evidence"

Creating a quasi KG

<Nolan, directed, Inception>
<Inception, won, Best Sound>
<2011 Oscars, announced, Best Sound>
<Inception, nominated, Best Actor>
<The movie Inception, missed out, Golden Globe Awards>
<Chris Nolan, director of, The movie Inception>
<Inception's script, edited by, Chris Nolan>
<Inception, lost to, The Social Network>
<Best Actor, declared at, 83rd Academy Awards>
<The Social Network, winner of, Best Screenplay>
<Golden Globes, announced, Best Screenplay>

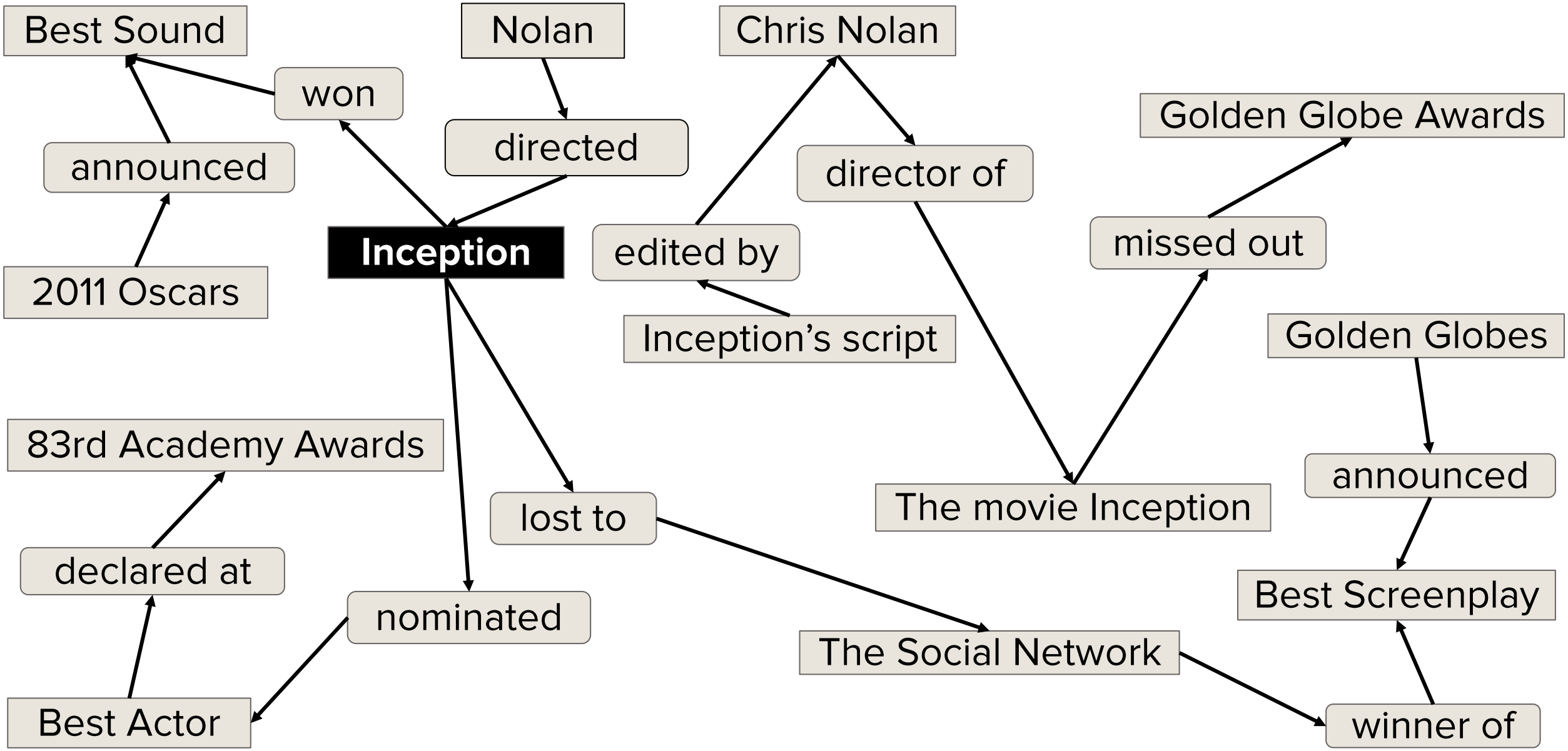
Compile an open-vocabulary triple store

Triples can ideally come from text (via Open IE), KG, or both

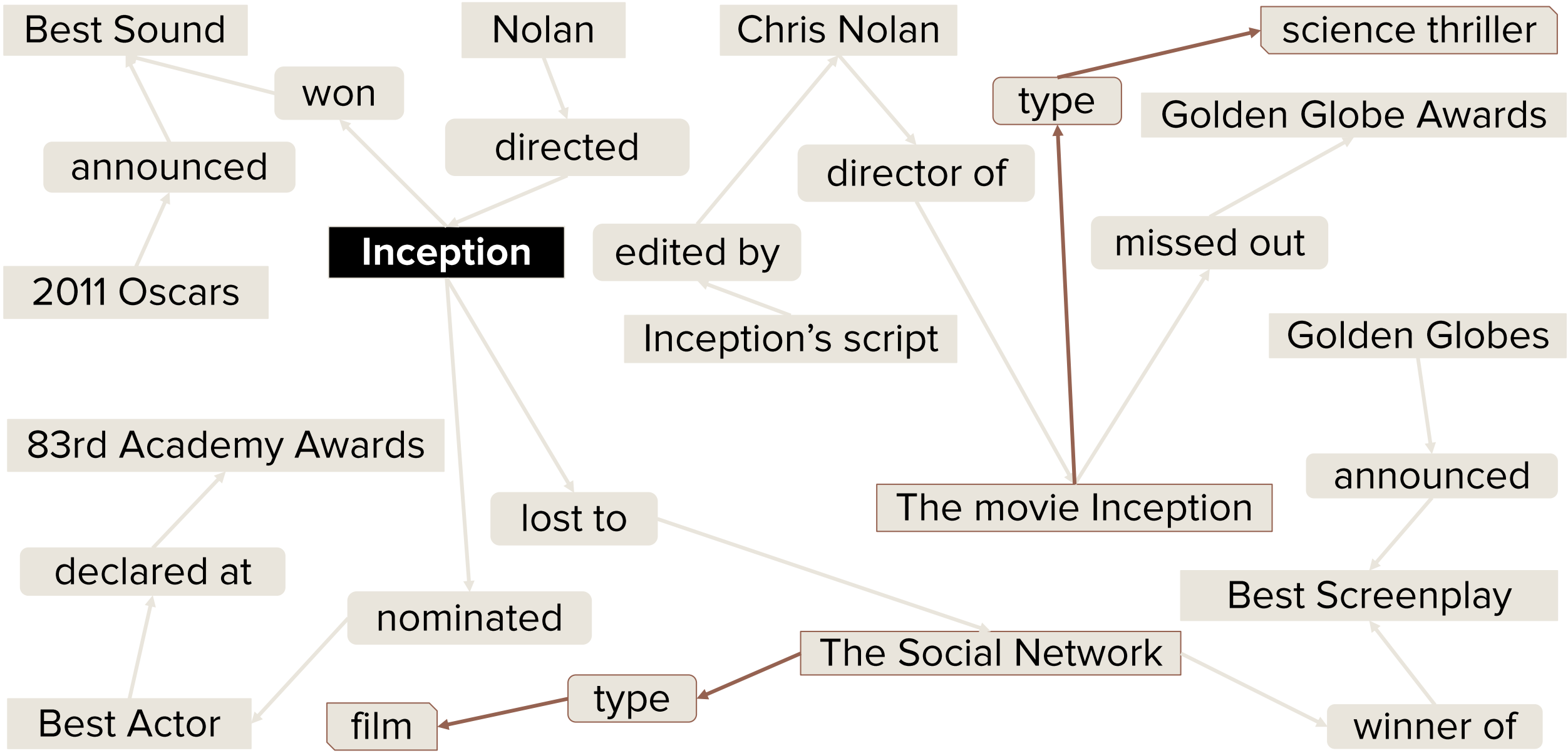
Open IE extracts KG-style triples by running pattern extraction over raw text: Stanford Open IE, ClausIE, OpenIE 5.0, …

Lu et al., Answering Complex Questions by Joining Multi-Document Evidence with Quasi Knowledge Graphs, SIGIR 2019.

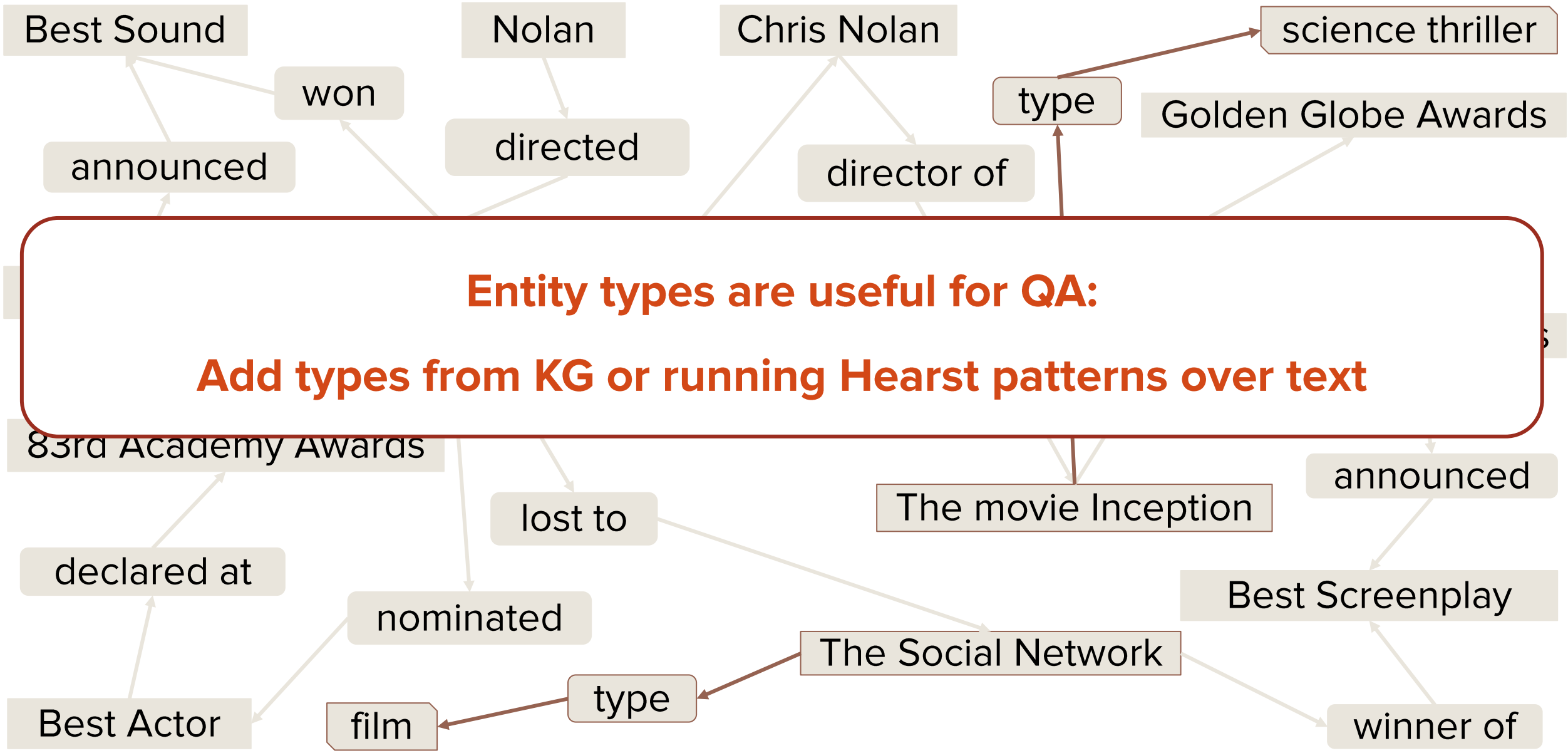
Question: Which Nolan films won an Oscar but missed a Golden Globe?



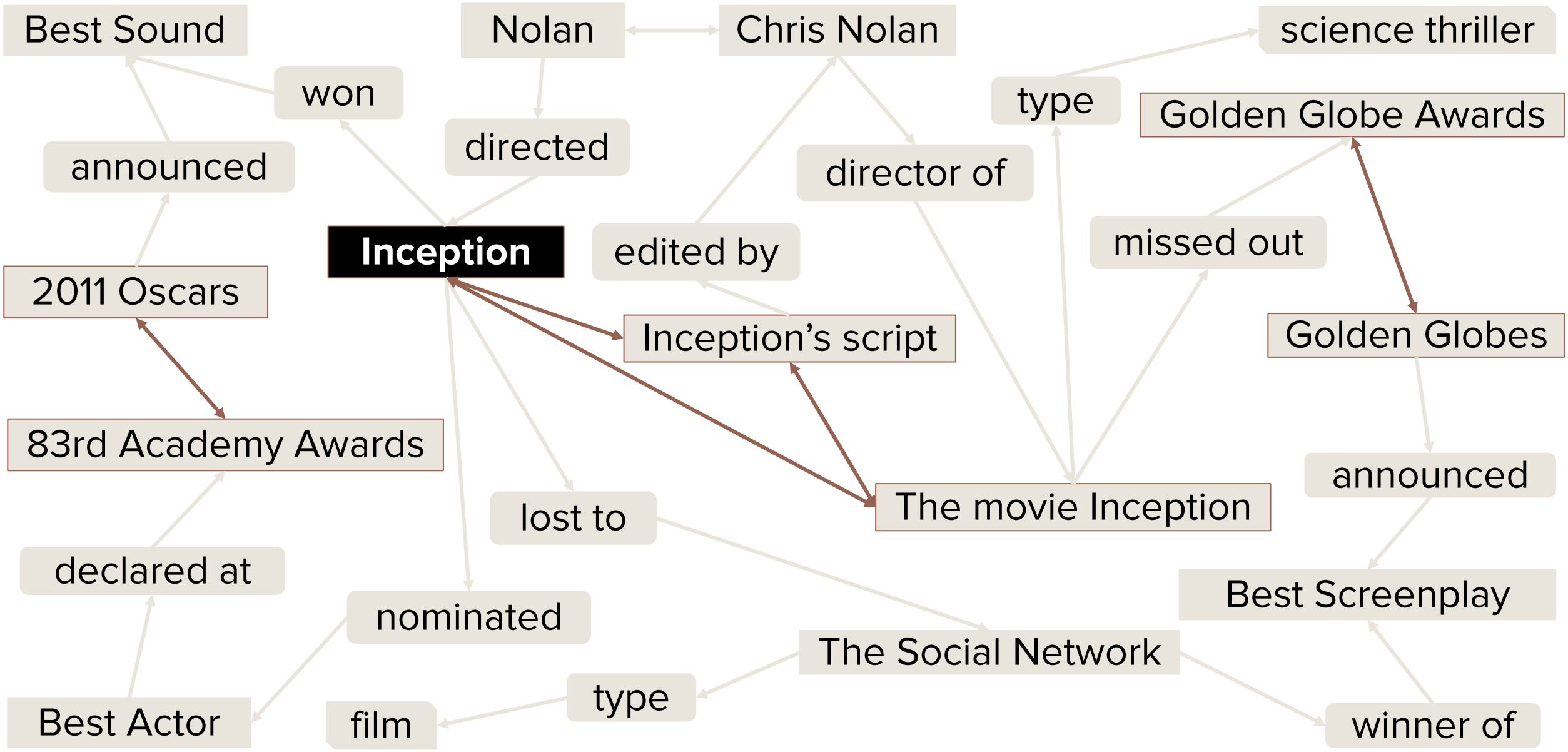
Question: Which Nolan films won an Oscar but missed a Golden Globe?



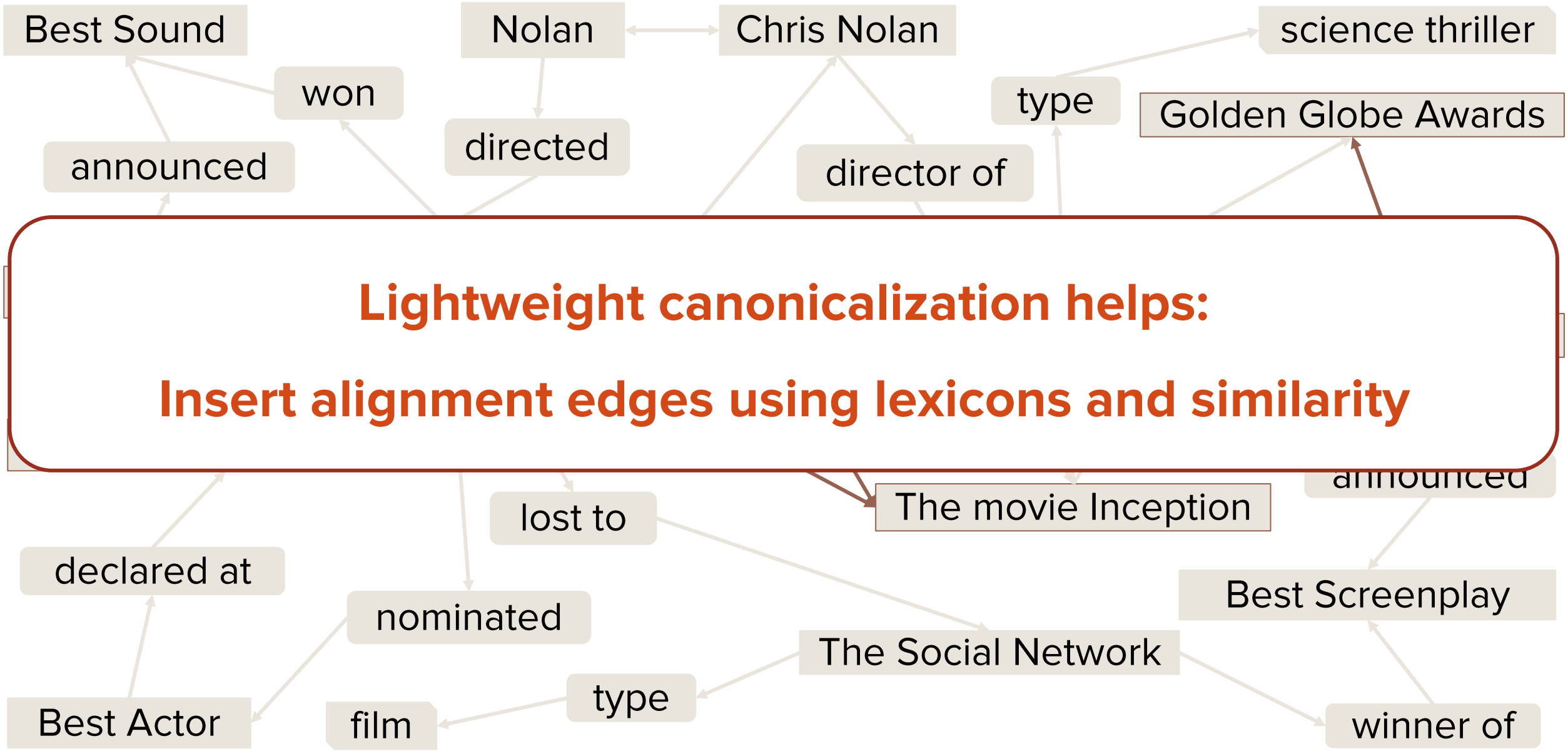
Question: Which Nolan films won an Oscar but missed a Golden Globe?



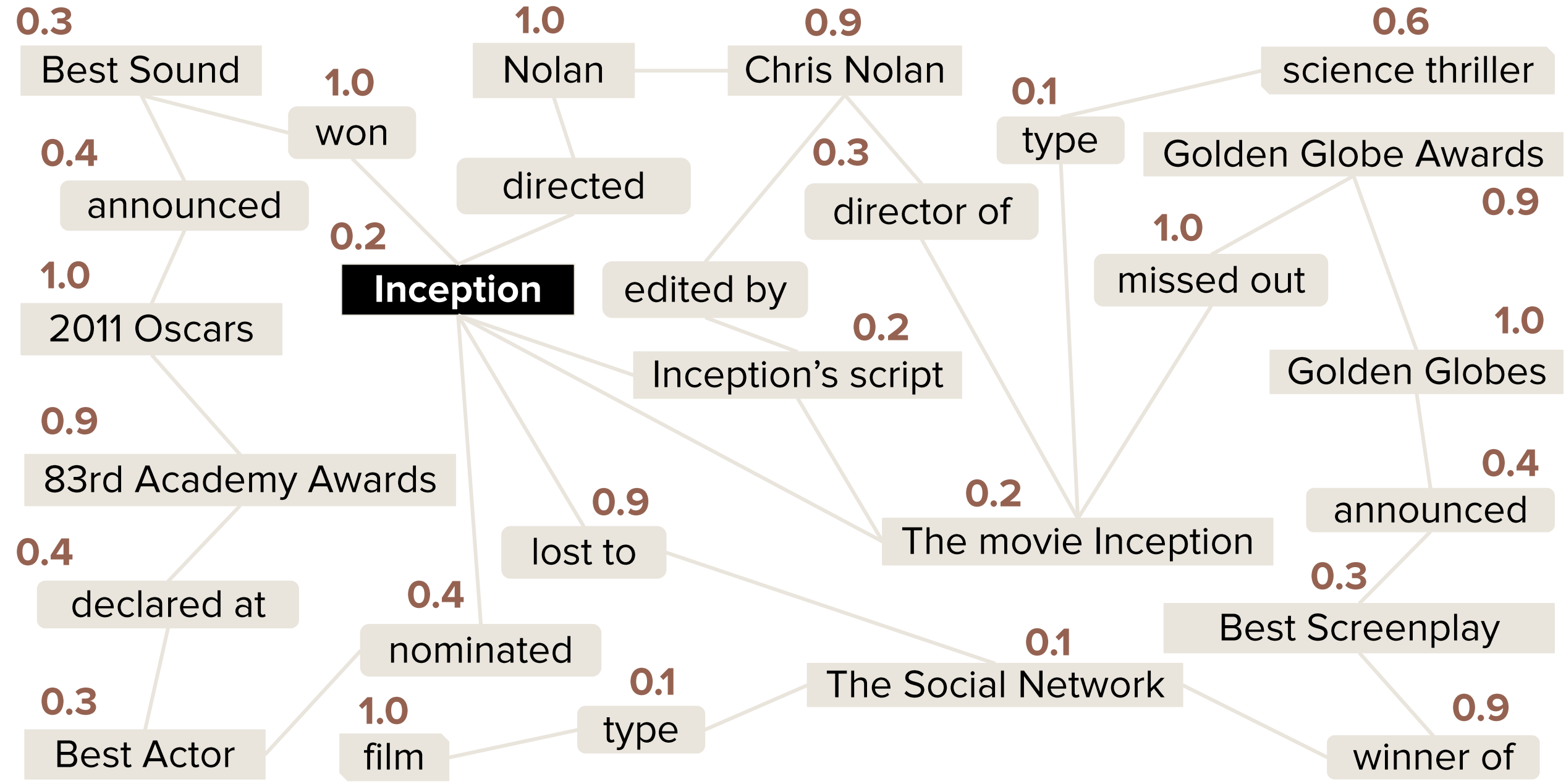
Question: Which Nolan films won an Oscar but missed a Golden Globe?



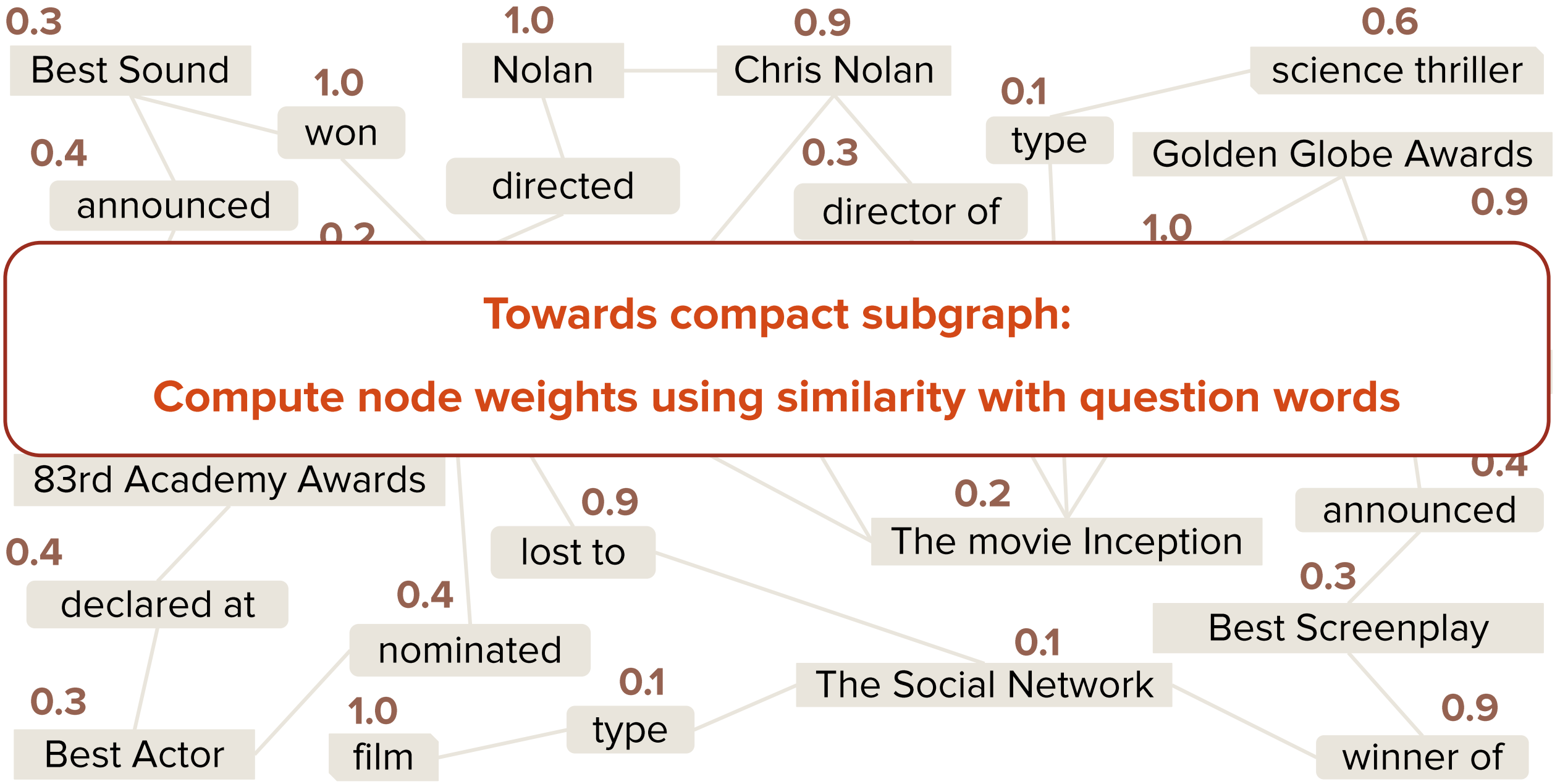
Question: Which Nolan films won an Oscar but missed a Golden Globe?



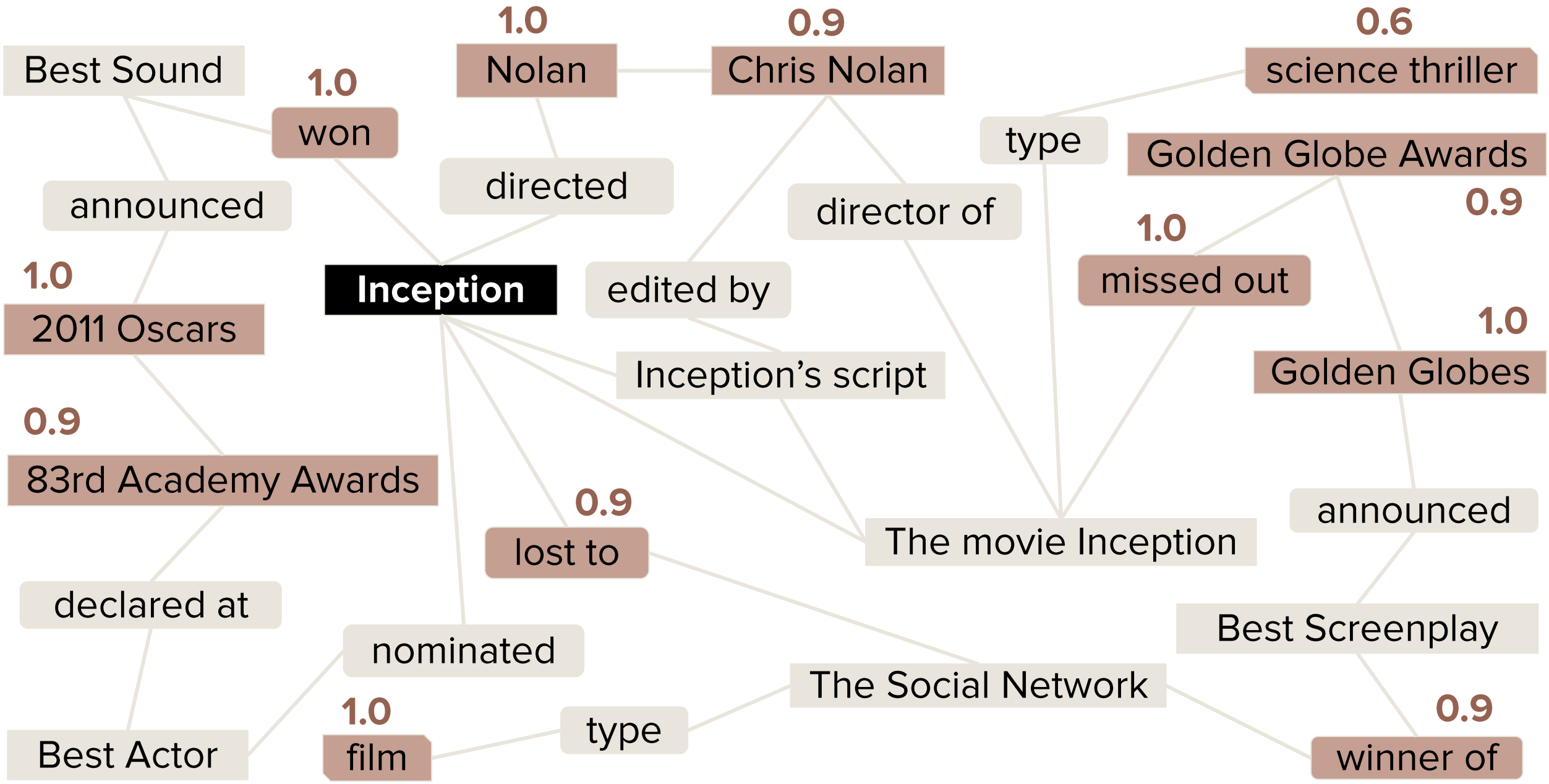
Question: Which Nolan films won an Oscar but missed a Golden Globe?



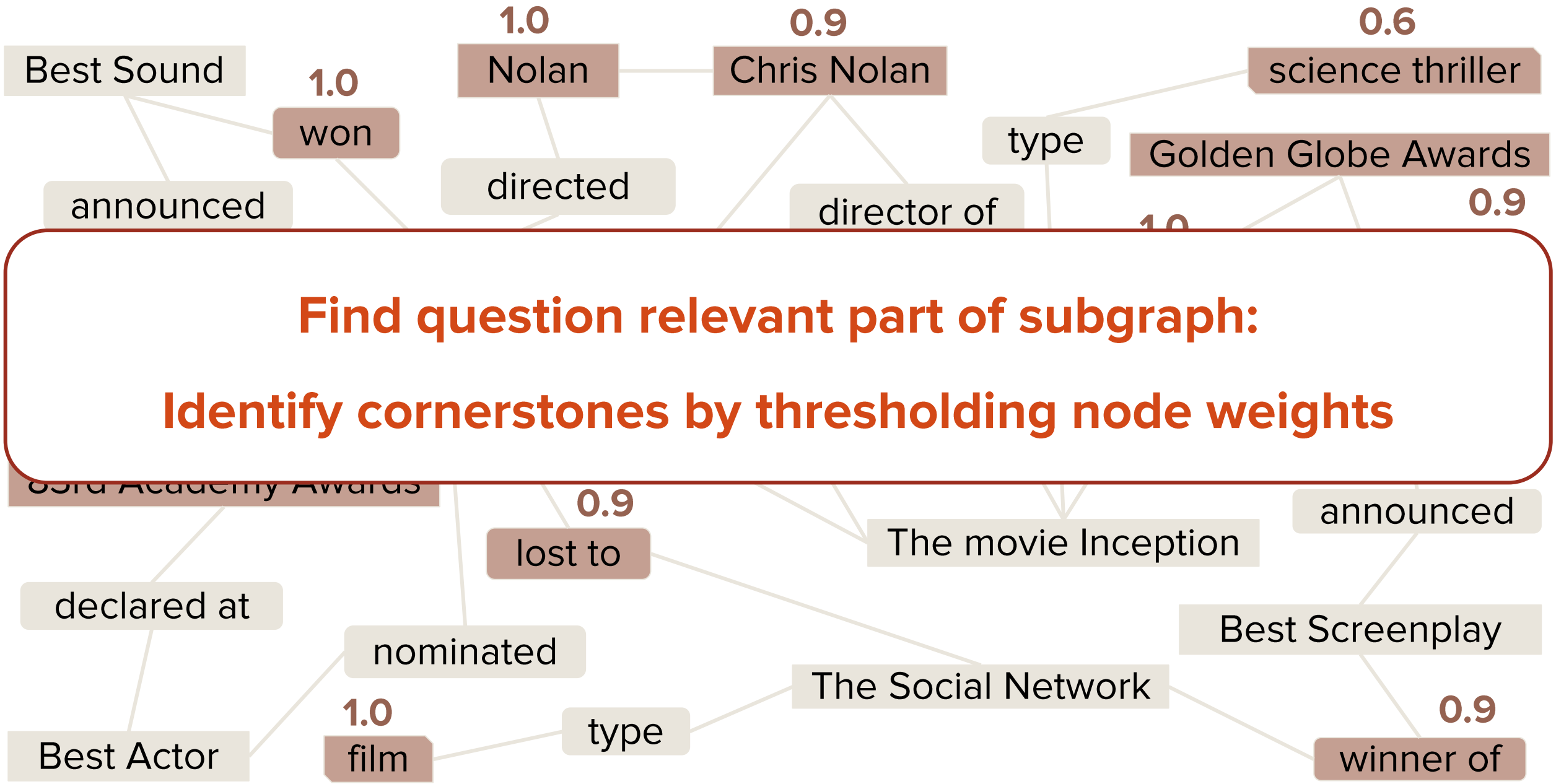
Question: Which Nolan films won an Oscar but missed a Golden Globe?



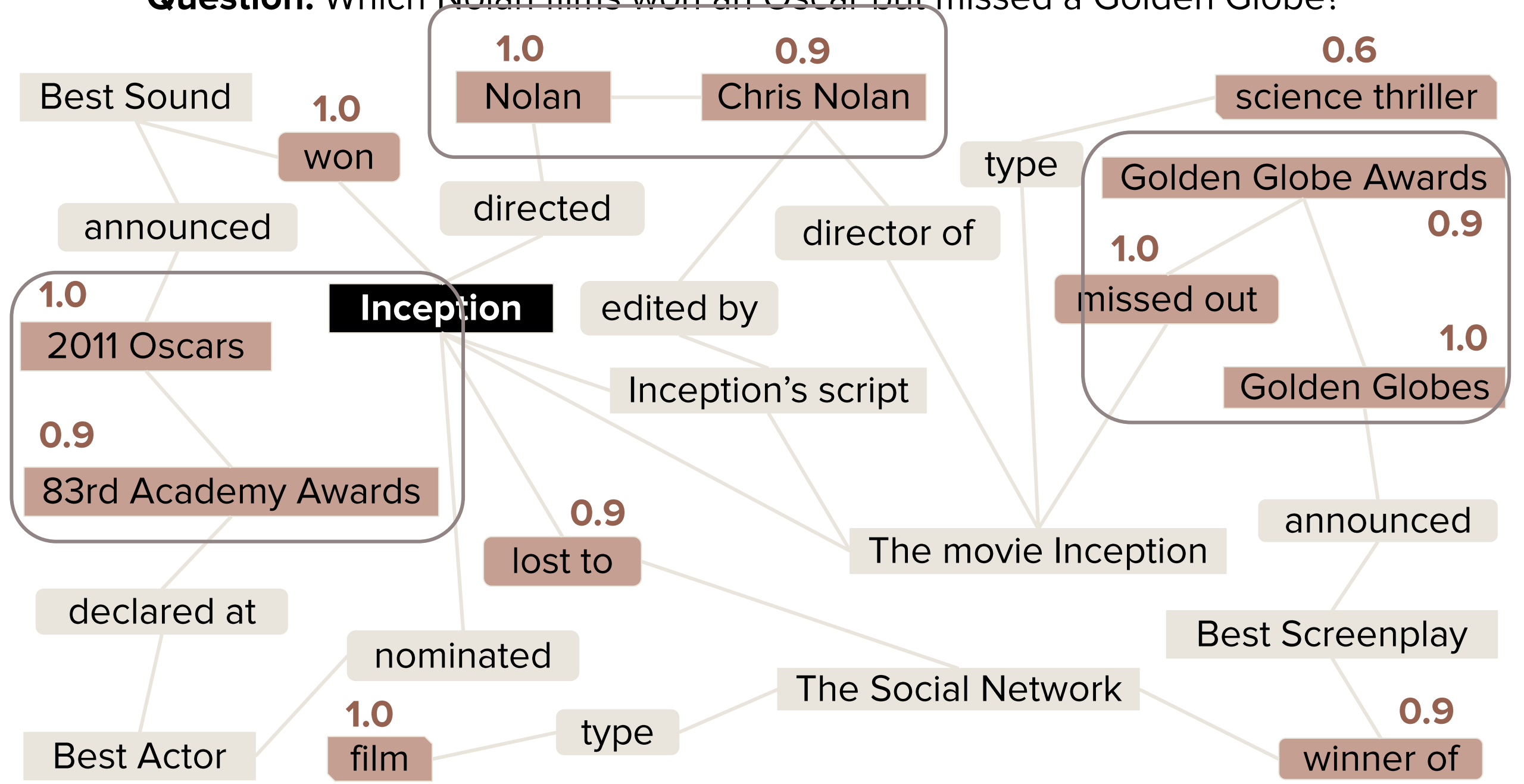
Question: Which Nolan films won an Oscar but missed a Golden Globe?



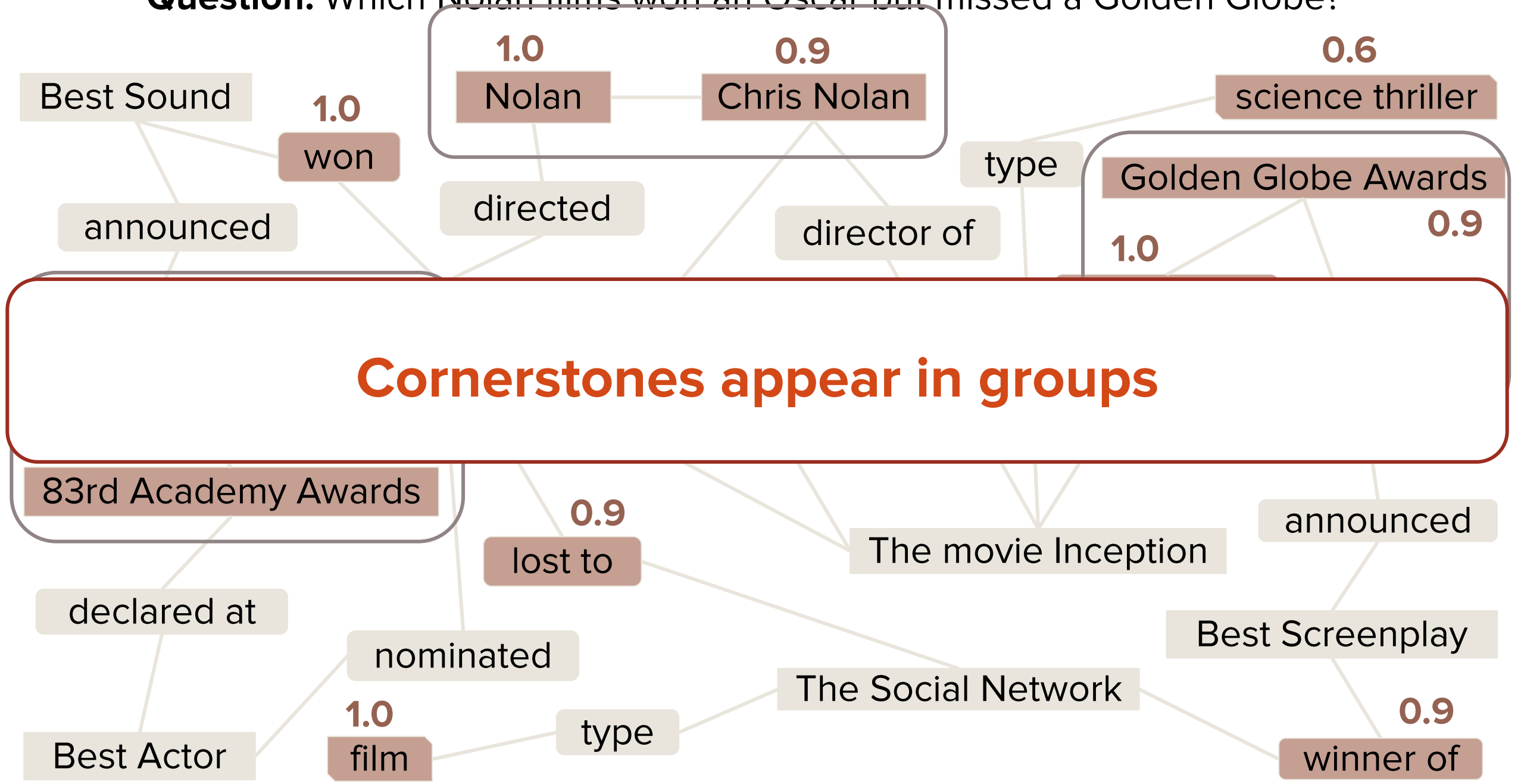
Question: Which Nolan films won an Oscar but missed a Golden Globe?



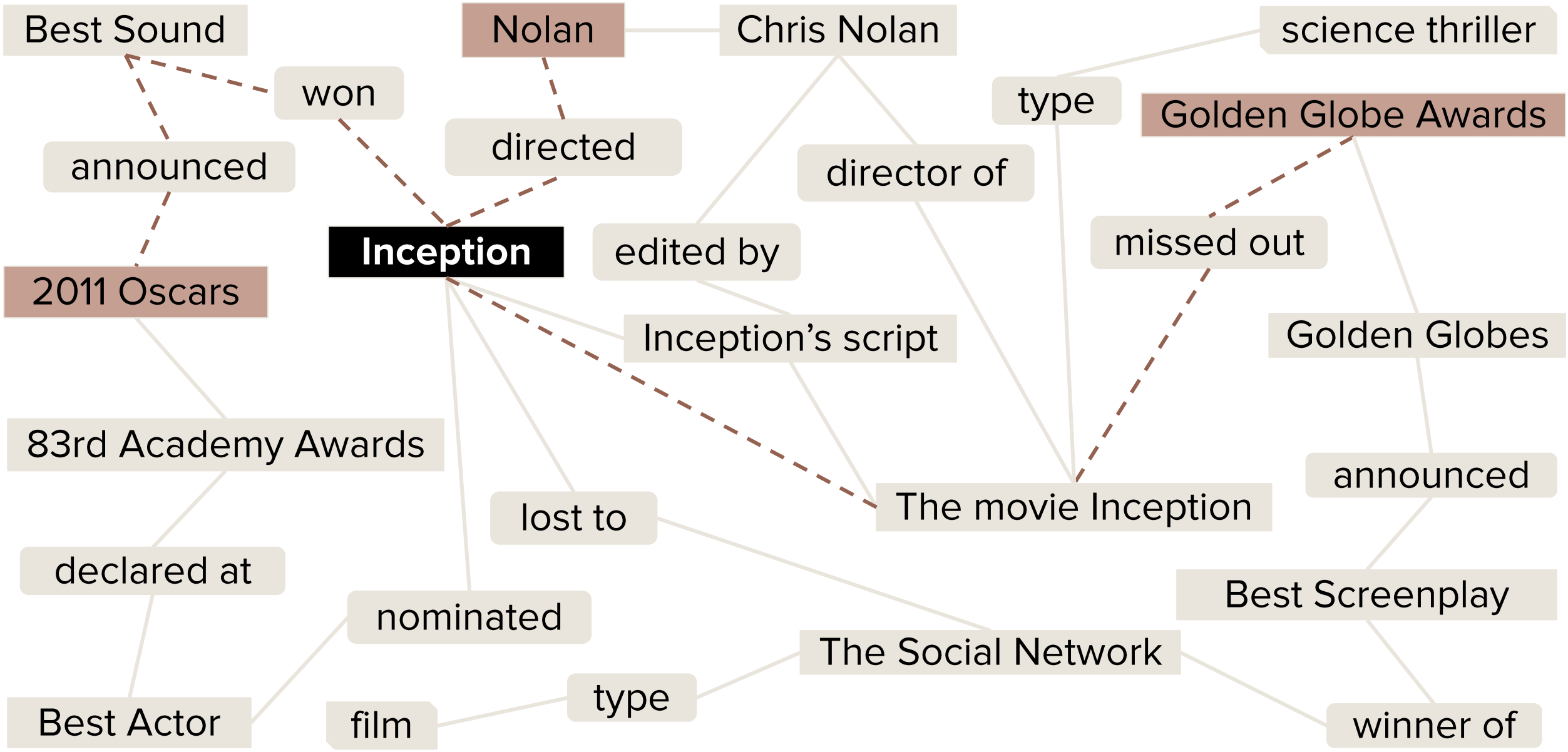
Question: Which Nolan films won an Oscar but missed a Golden Globe?



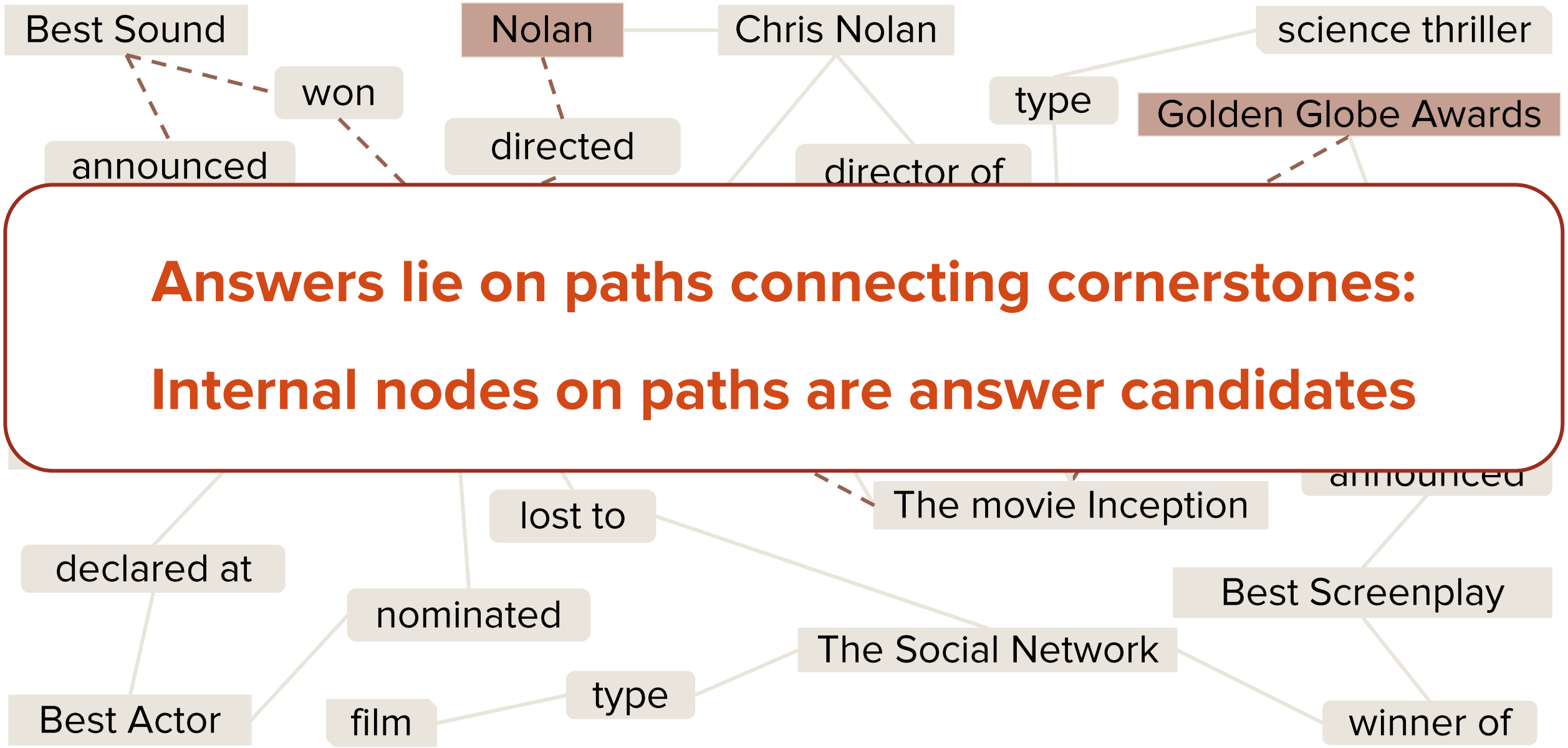
Question: Which Nolan films won an Oscar but missed a Golden Globe?



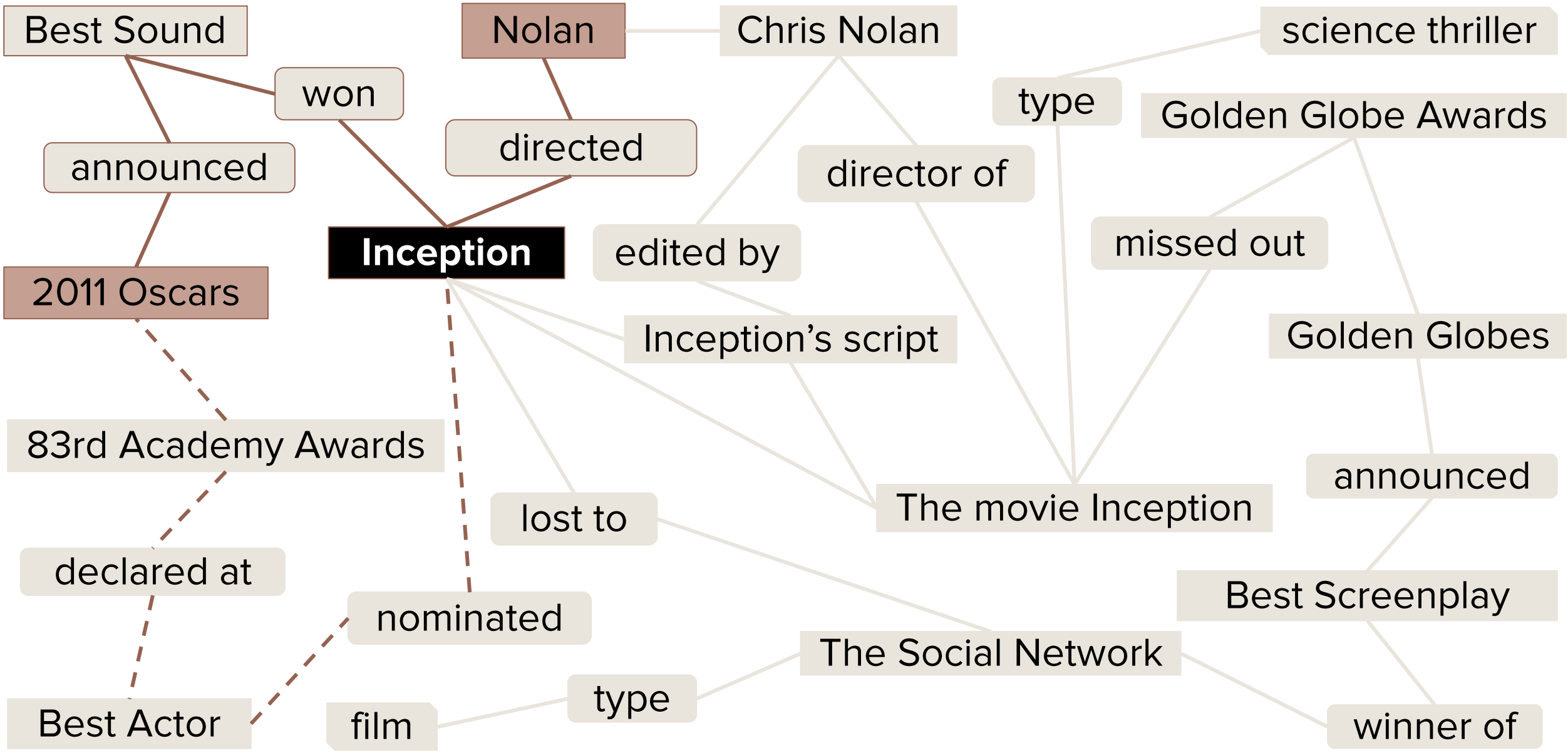
Question: Which Nolan films won an Oscar but missed a Golden Globe?



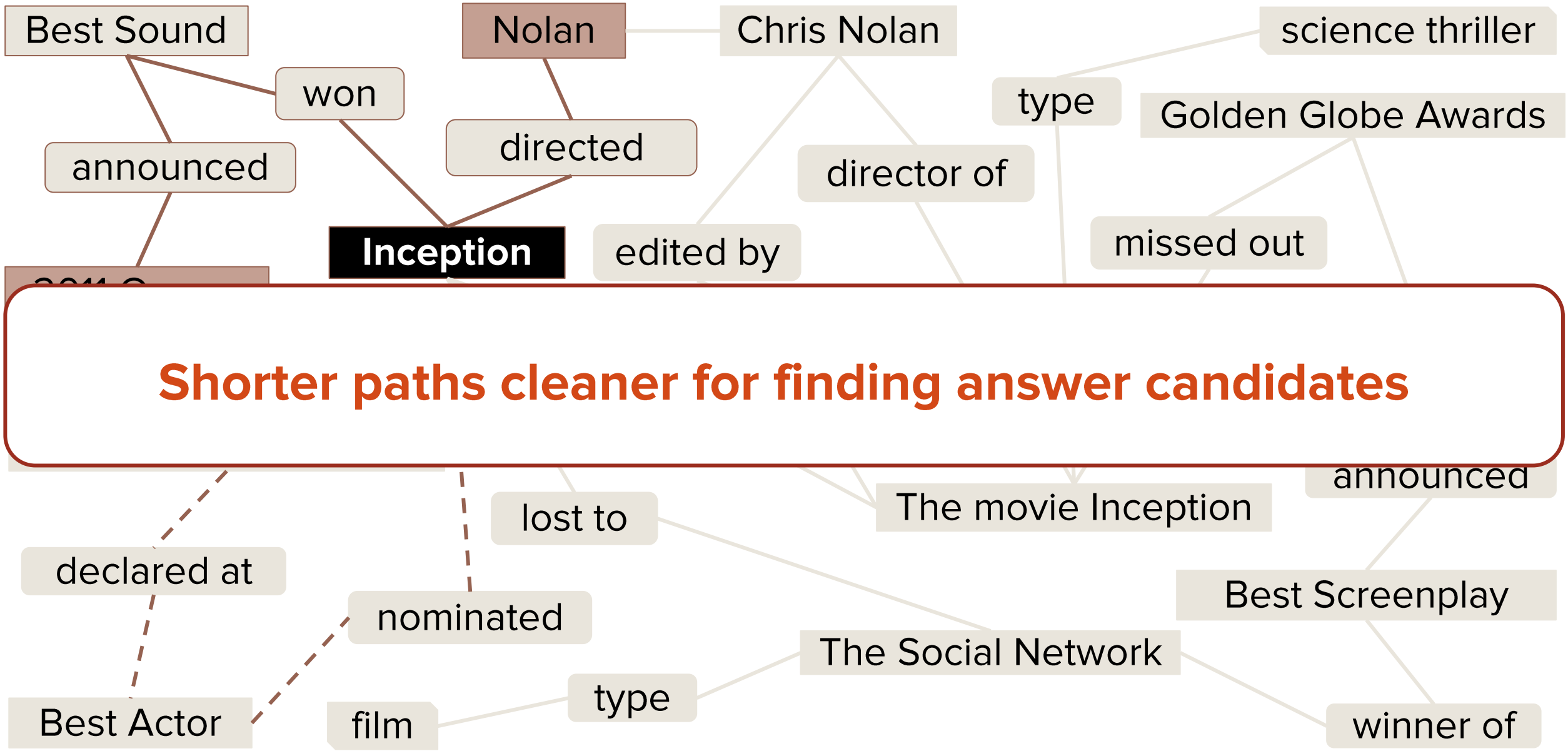
Question: Which Nolan films won an Oscar but missed a Golden Globe?



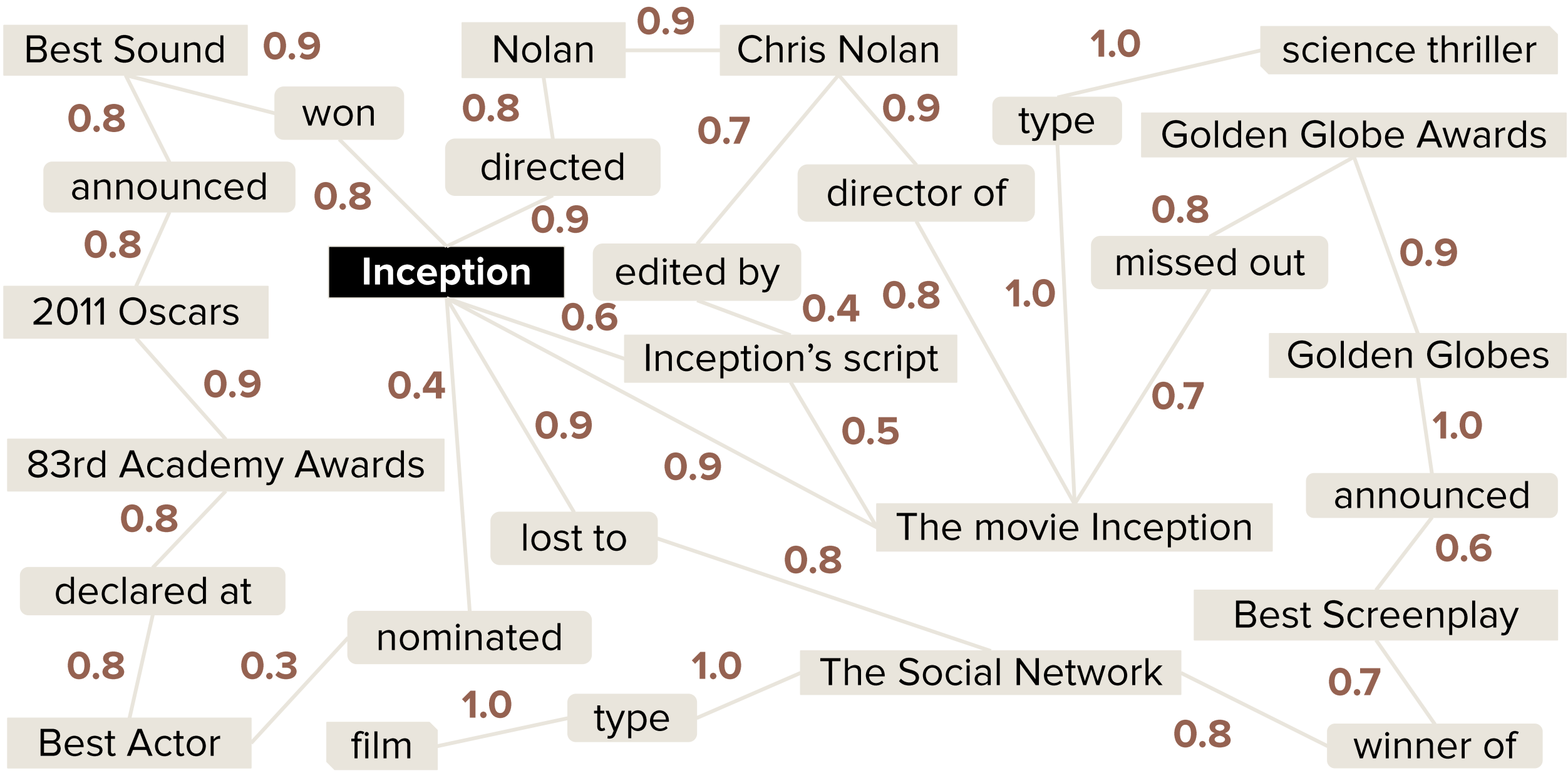
Question: Which Nolan films won an Oscar but missed a Golden Globe?



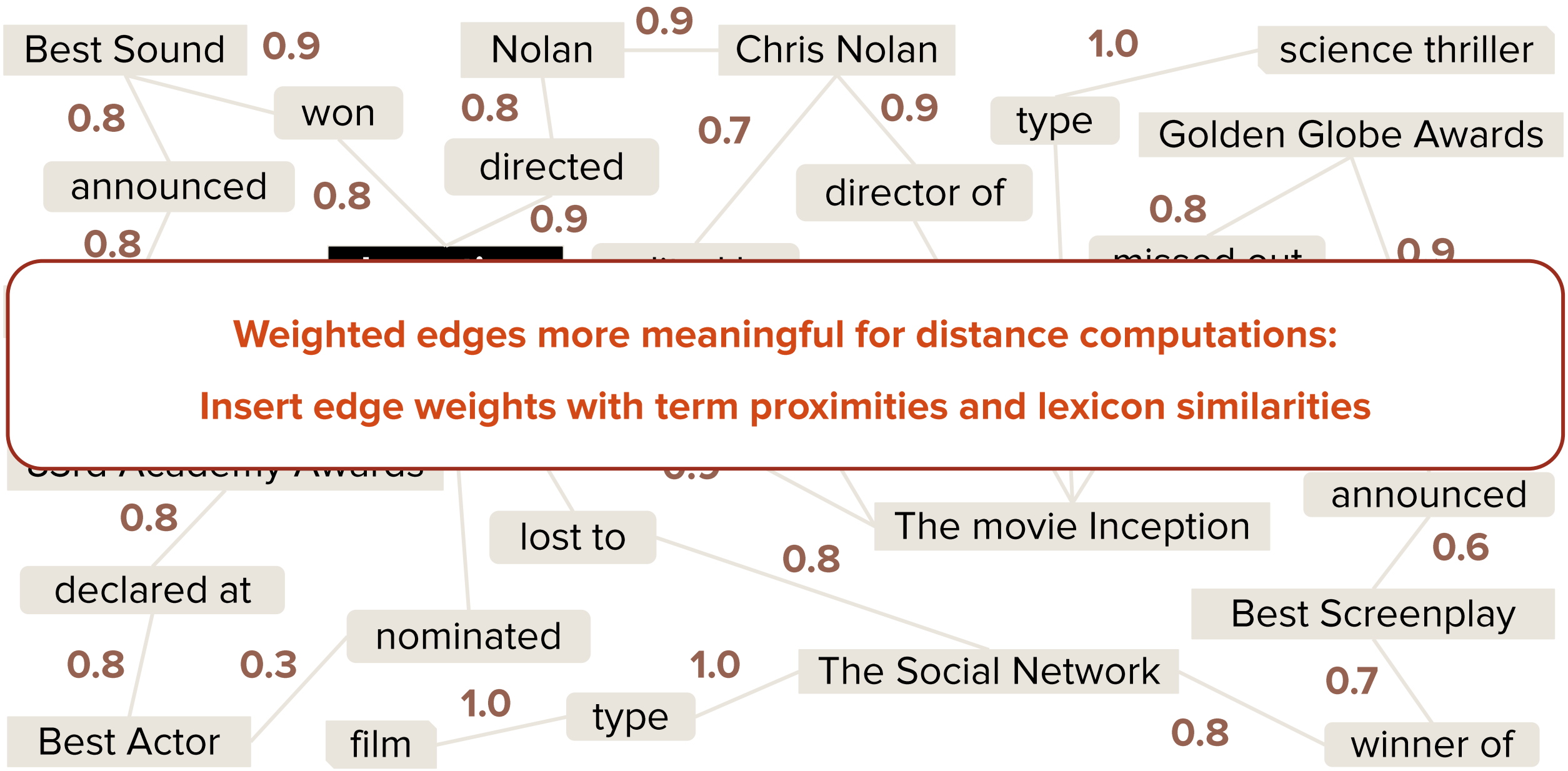
Question: Which Nolan films won an Oscar but missed a Golden Globe?



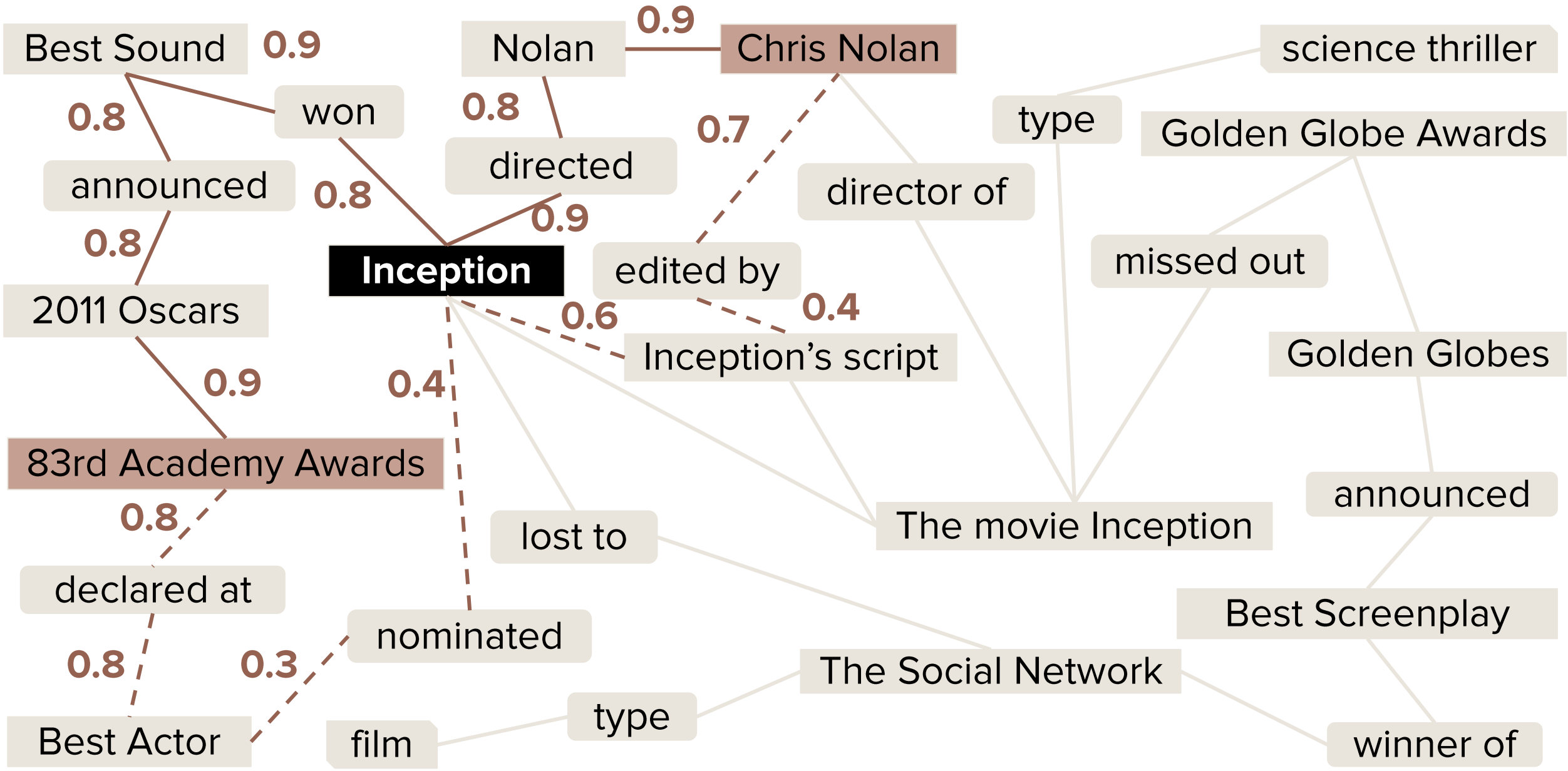
Question: Which Nolan films won an Oscar but missed a Golden Globe?



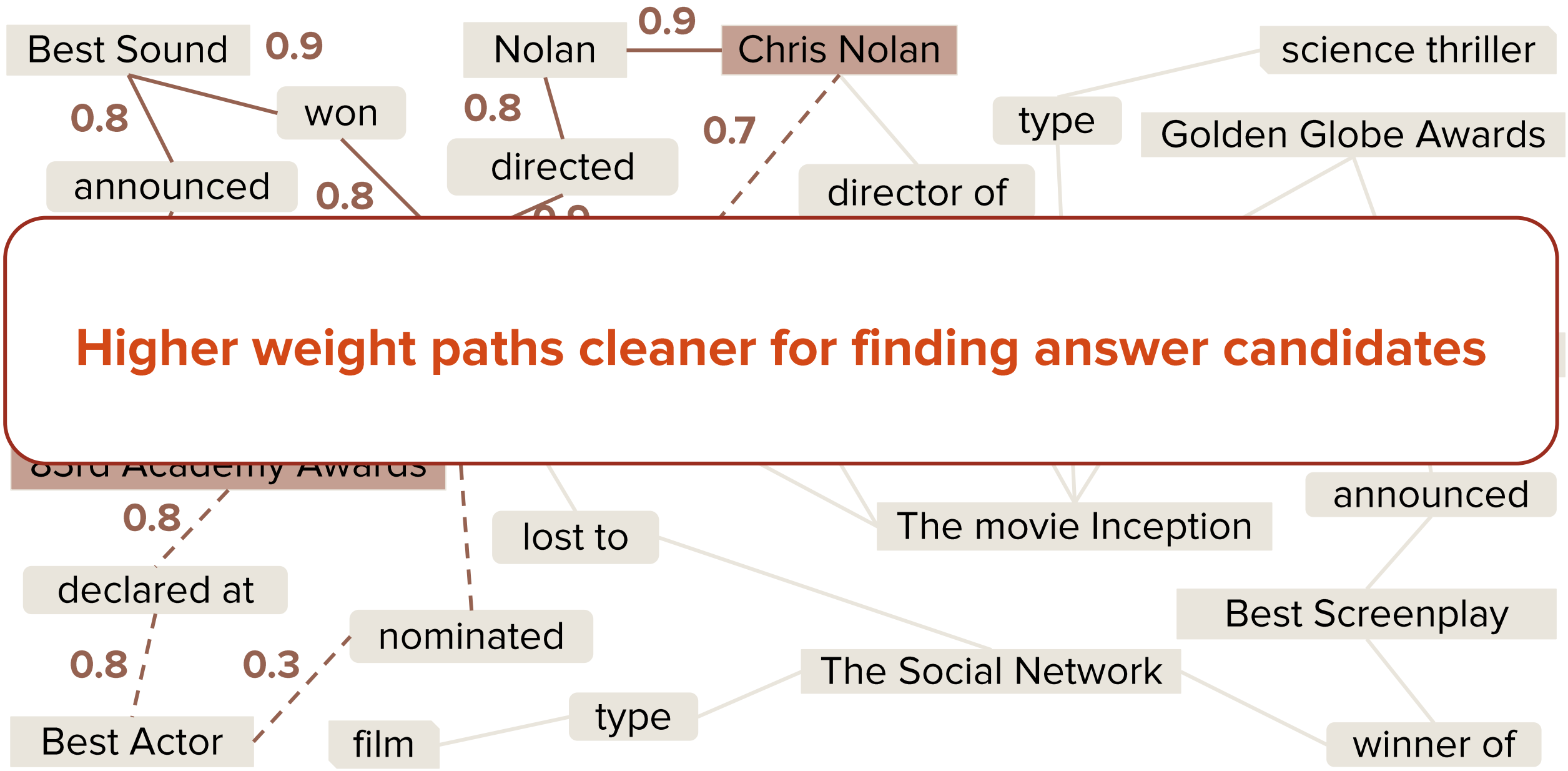
Question: Which Nolan films won an Oscar but missed a Golden Globe?



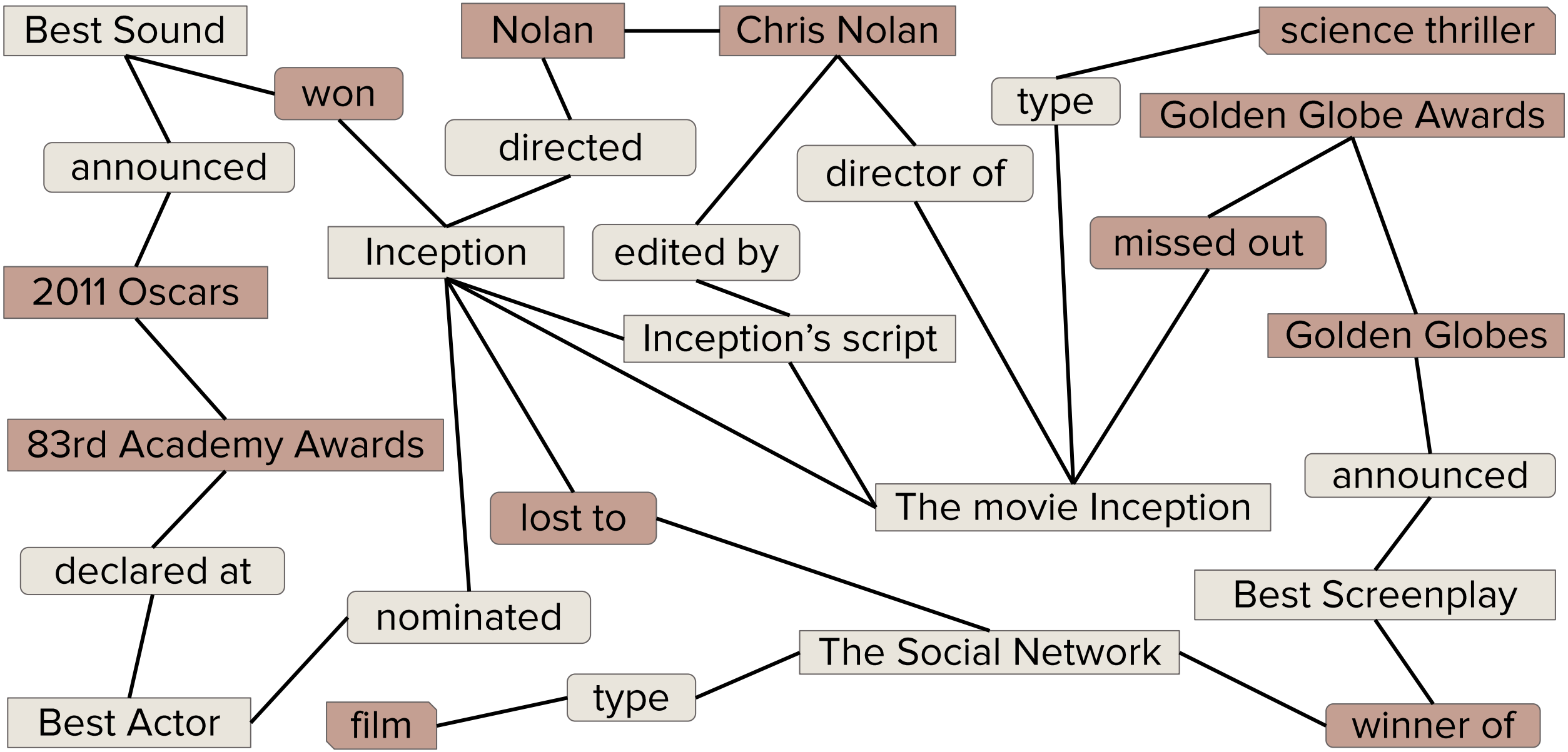
Question: Which Nolan films won an Oscar but missed a Golden Globe?



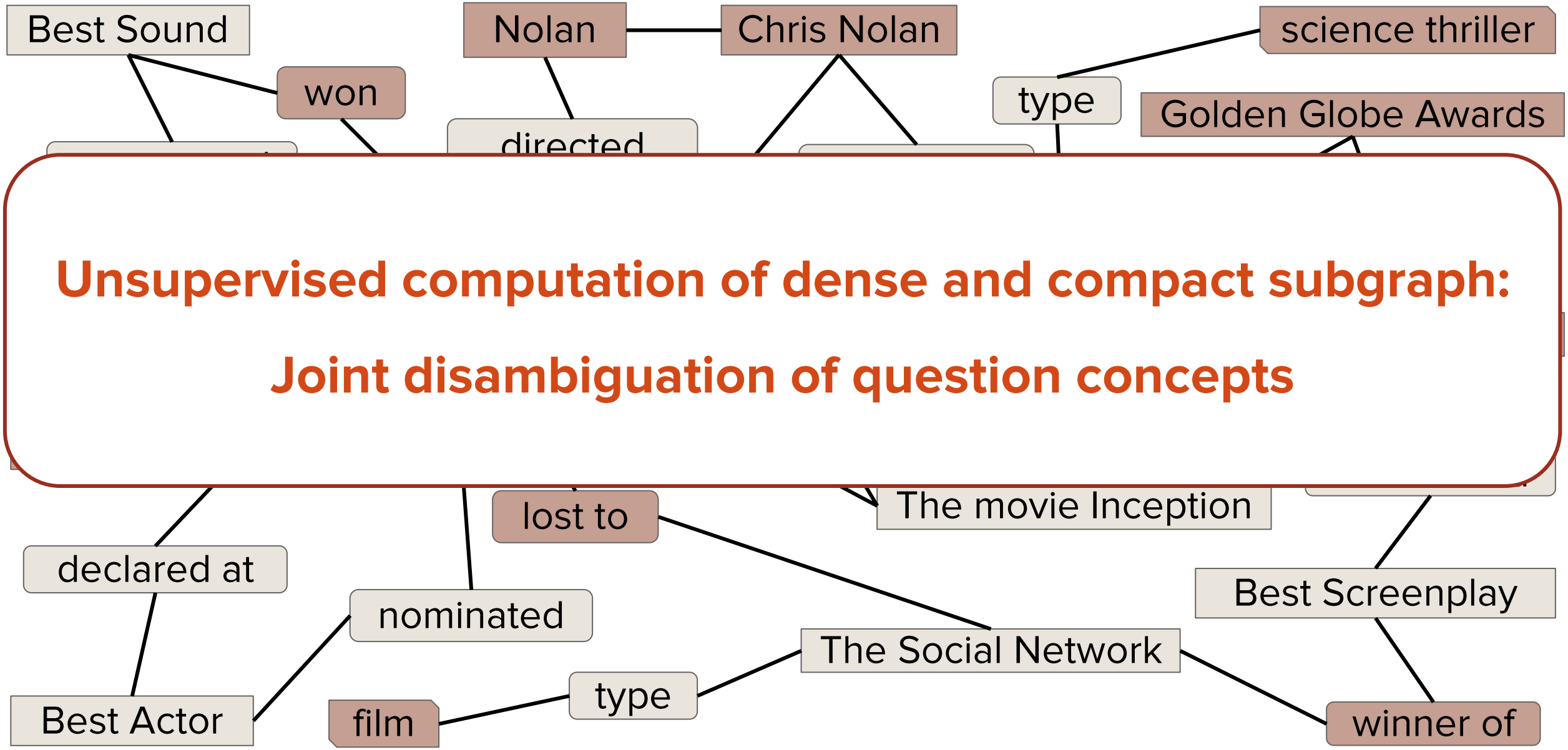
Question: Which Nolan films won an Oscar but missed a Golden Globe?



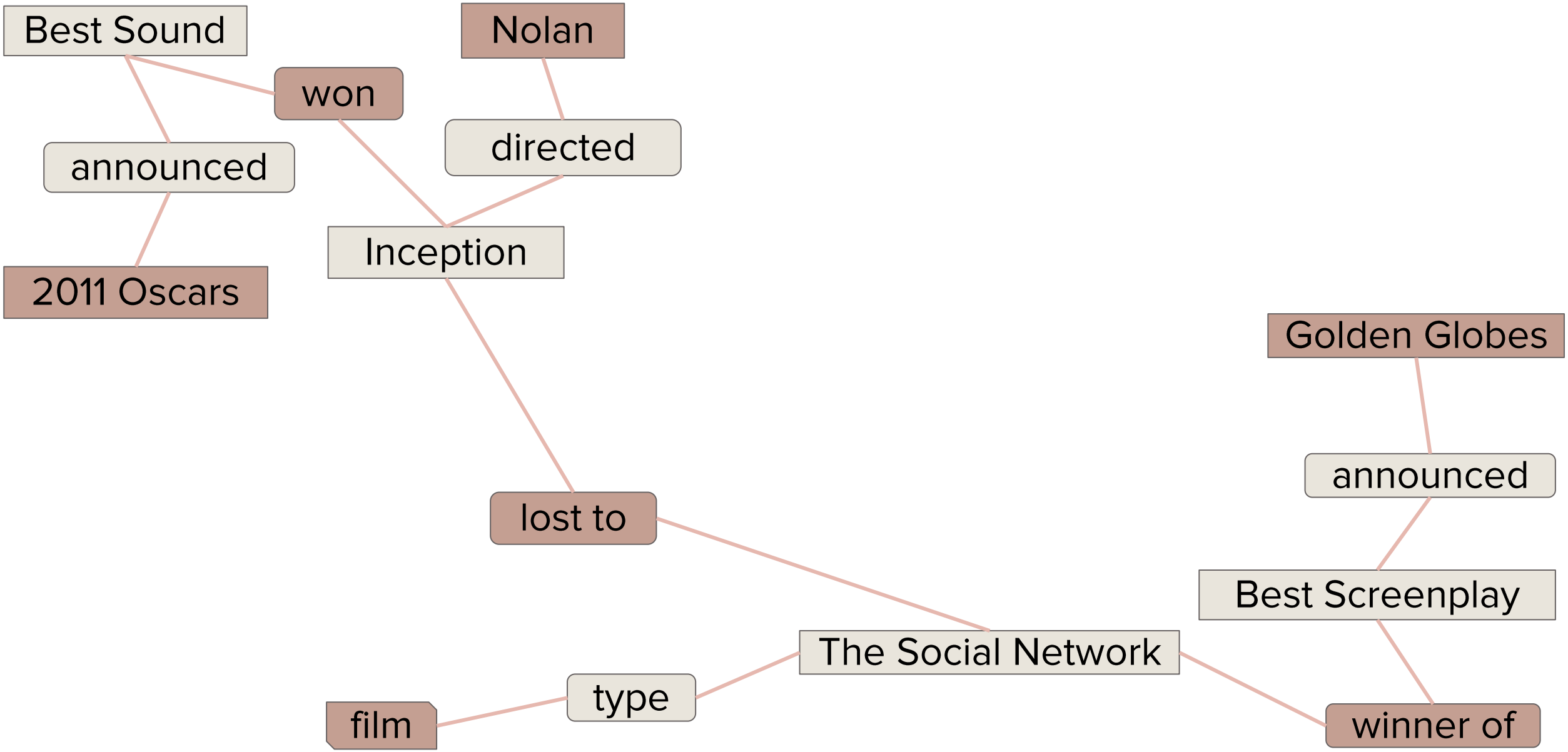
Question: Which Nolan films won an Oscar but missed a Golden Globe?



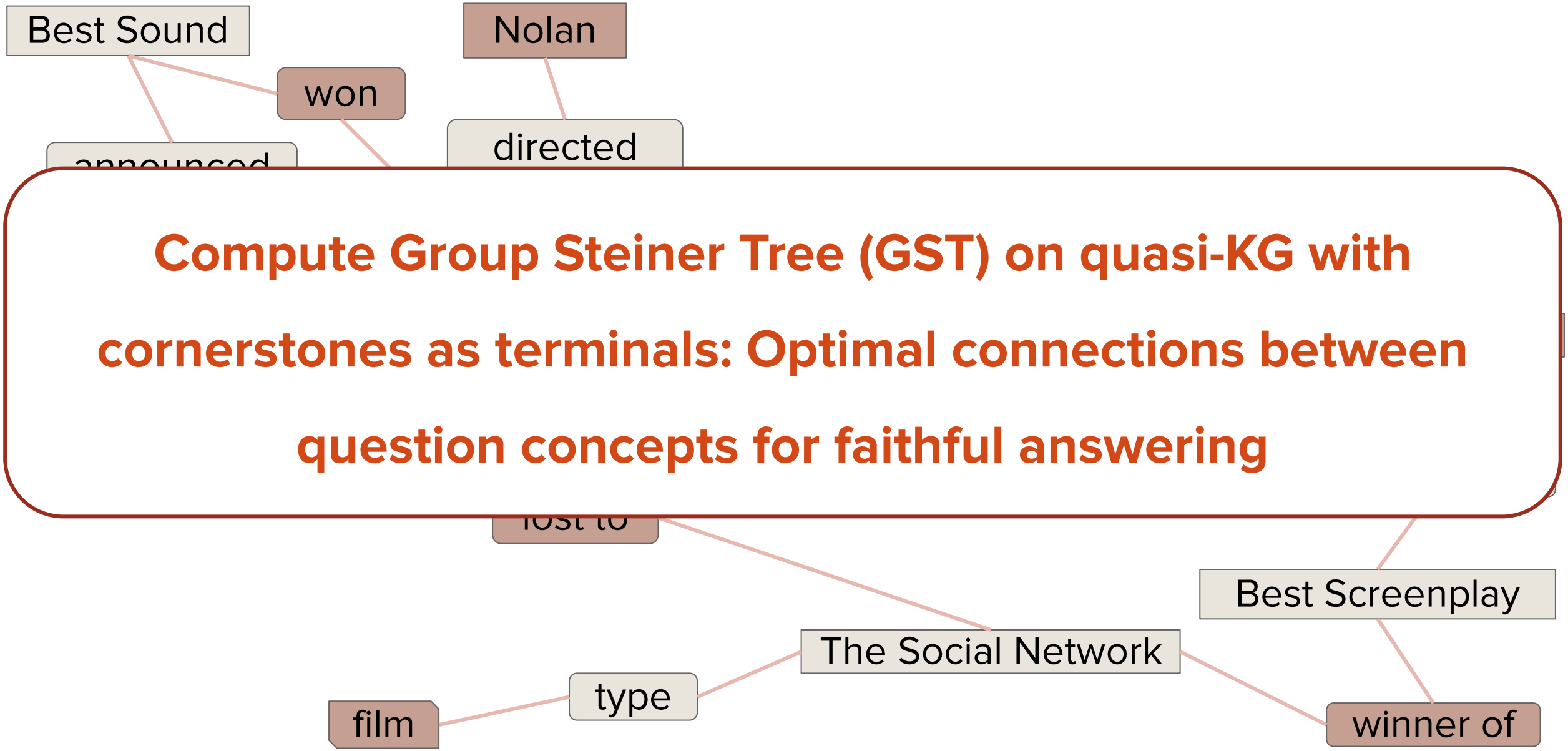
Question: Which Nolan films won an Oscar but missed a Golden Globe?



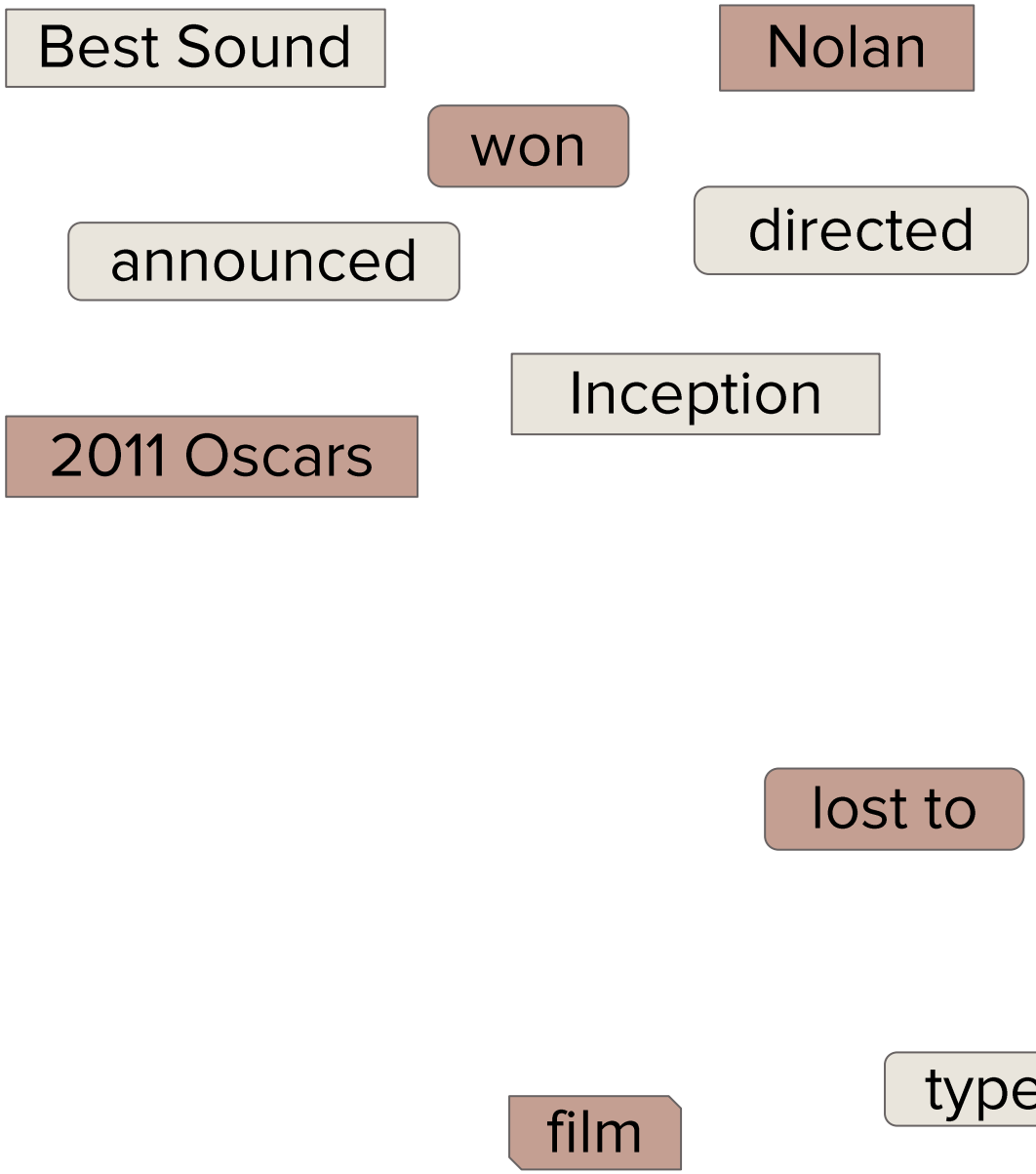
Question: Which Nolan films won an Oscar but missed a Golden Globe?



Question: Which Nolan films won an Oscar but missed a Golden Globe?

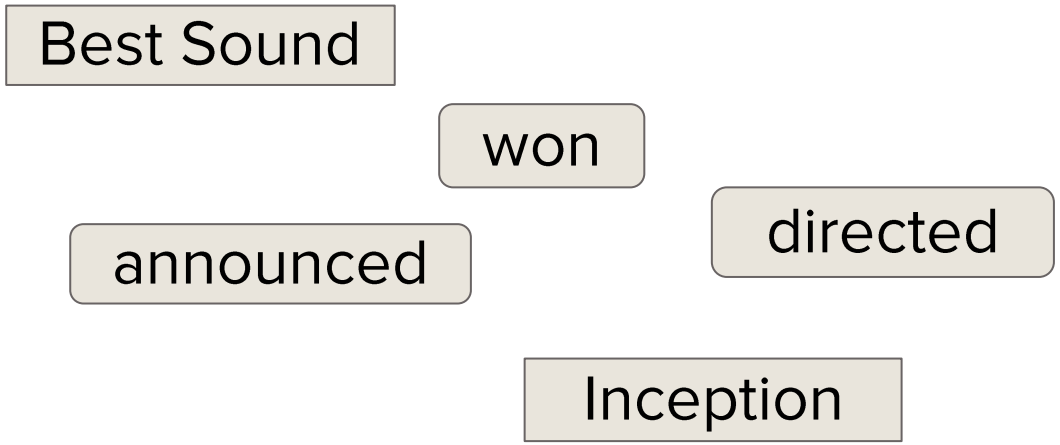


Question: Which Nolan films won an Oscar but missed a Golden Globe?



- **Answers are nodes**
- Cornerstones are not answers
- Only entities
- Must respect type constraints

Question: Which Nolan films won an Oscar but missed a Golden Globe?



- Answers are nodes
- **Cornerstones are not answers**
- Only entities
- Must respect type constraints

Question: Which Nolan films won an Oscar but missed a Golden Globe?

Best Sound

Inception

Best Screenplay

The Social Network

- Answers are nodes
- Cornerstones are not answers
- **Only entities**
- Must respect type constraints

Question: Which Nolan films won an Oscar but missed a Golden Globe?

- Answers are nodes
- Cornerstones are not answers
- Only entities
- **Must respect type constraints**

Inception

The Social Network

Question: Which Nolan films won an Oscar but missed a Golden Globe?

- **Answers ranked by aggregation**
- Best answer chosen

Inception

Number of GSTs	**5**

Number of GSTs	2

The Social Network

Question: Which Nolan films won an Oscar but missed a Golden Globe?

- Answers ranked by multiple criteria
- **Best answer chosen**

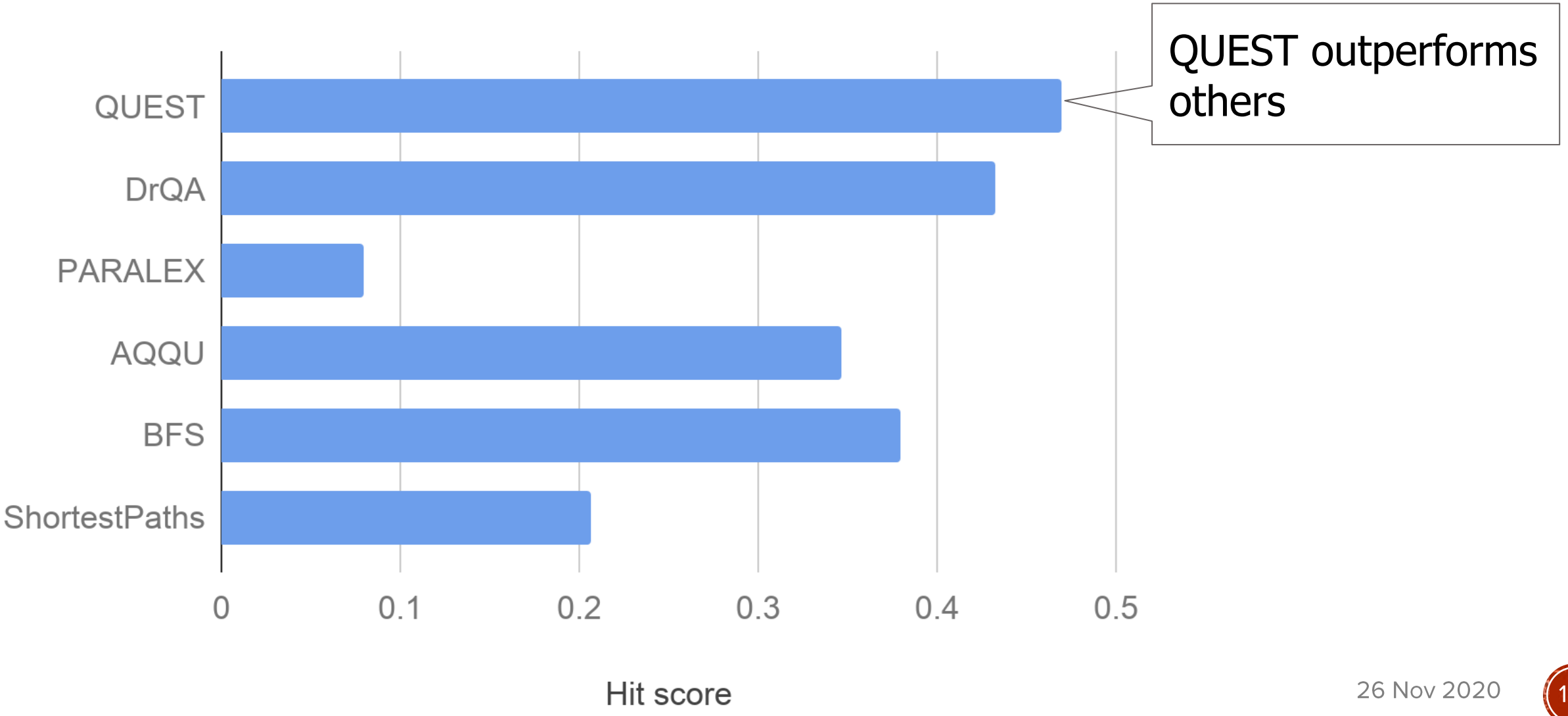
Inception

Lu et al., Answering Complex Questions by Joining Multi-Document Evidence with Quasi Knowledge Graphs, SIGIR 2019.

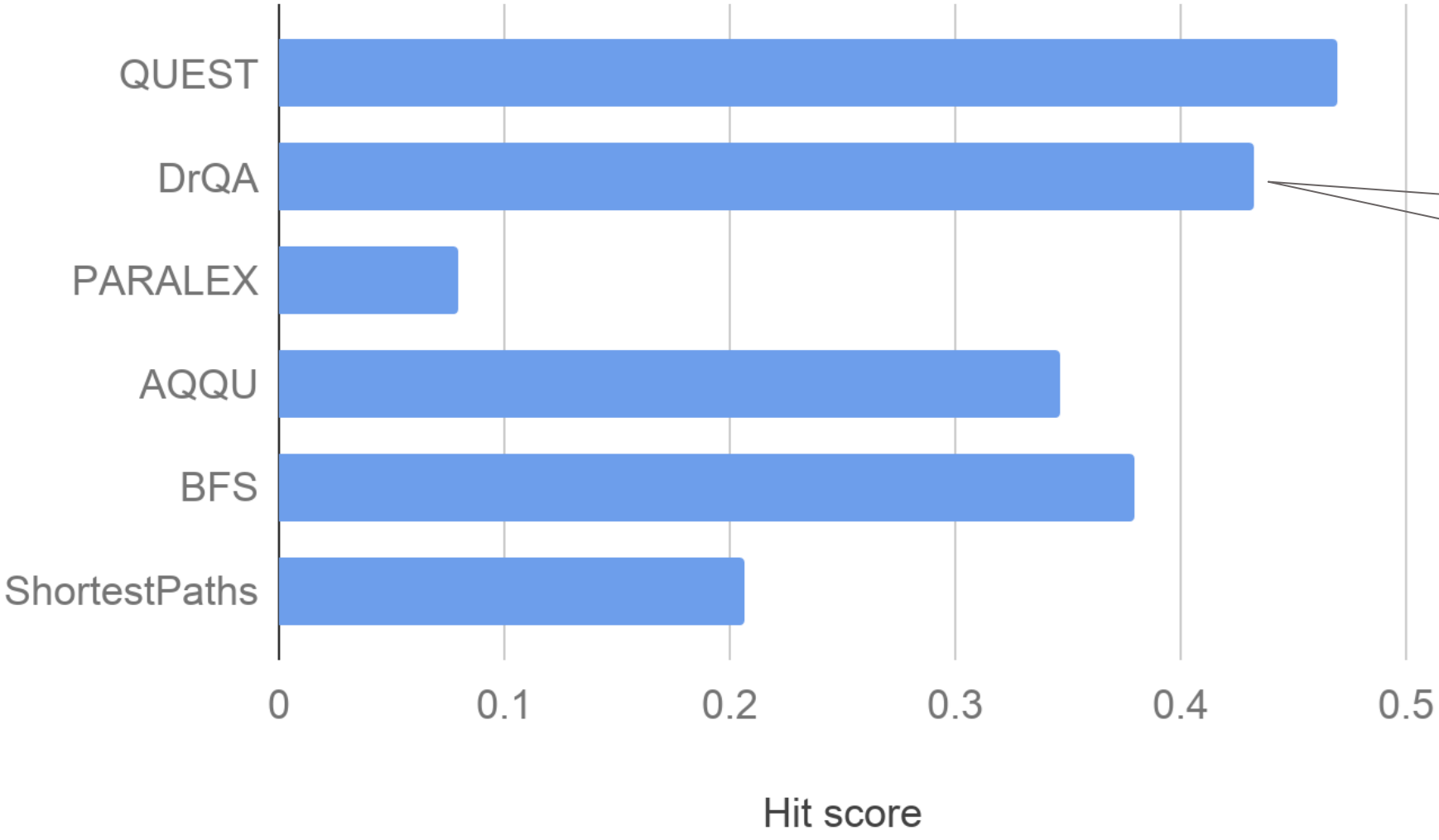
Evaluation

- Benchmark: 300 complex questions
- Metrics: MRR, P@1, **Hit@5**
- Baselines: QA algorithms and graph methods

Experimental results

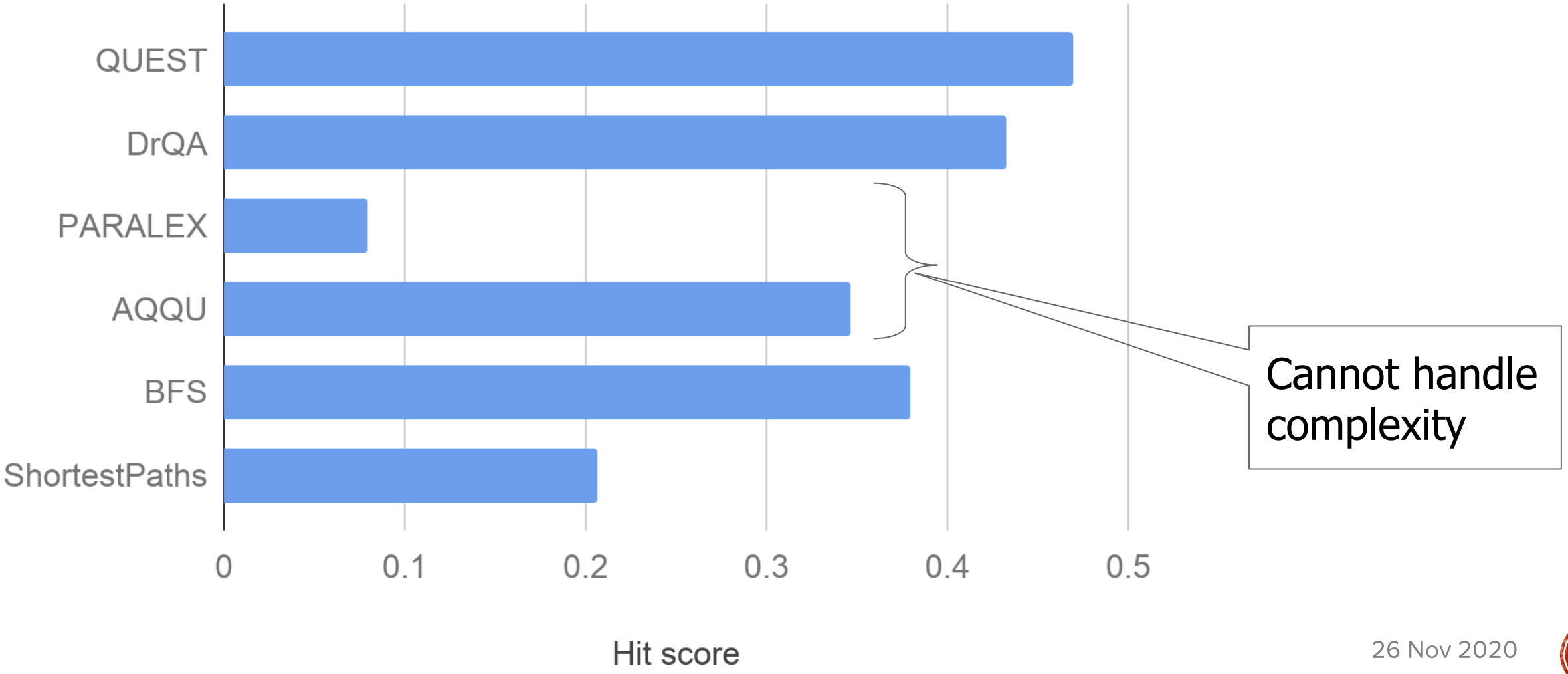


Experimental results

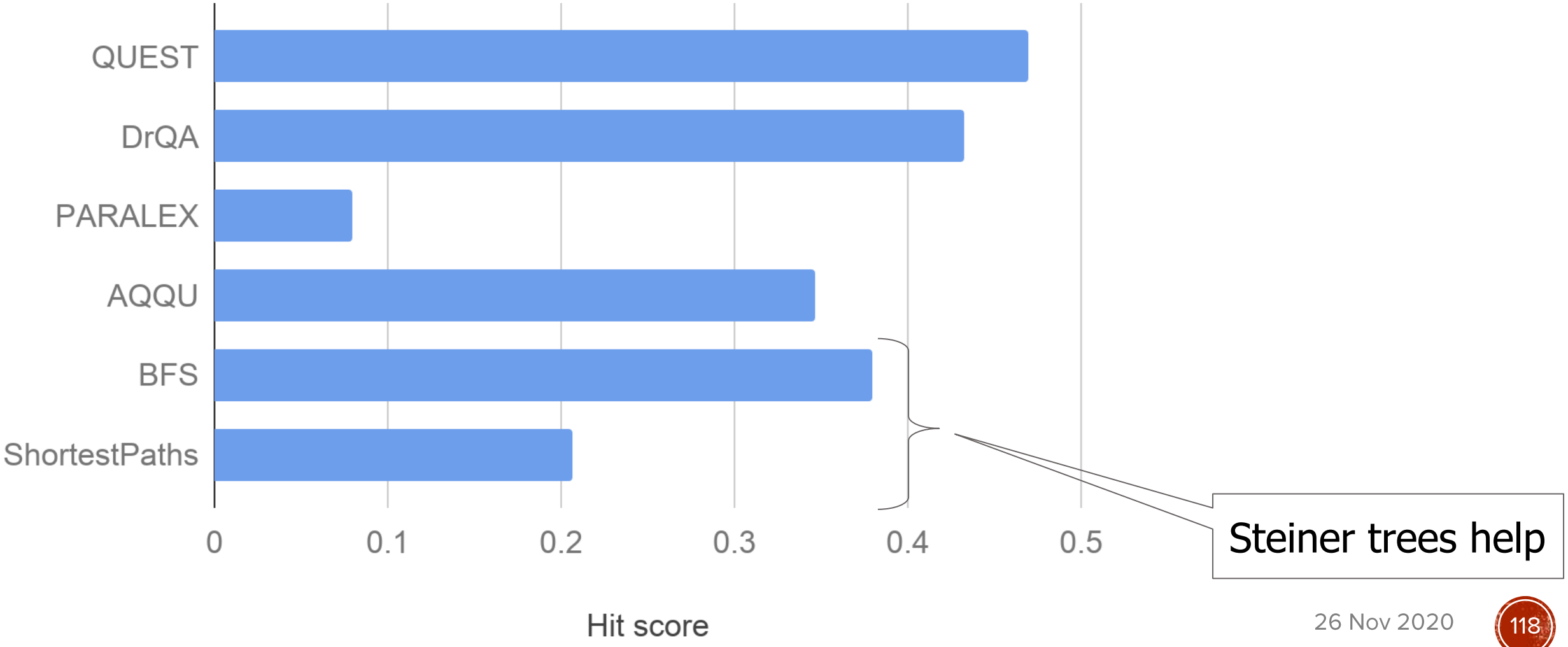


Cannot join multi-doc evidence

Experimental results



Experimental results



Anecdotal examples

- Which Japanese baseball player was contracted for Los Angeles Angels who also played for Hokkaido Nippon-ham fighters?
- Which aspiring model split with Chloe Moretz and is dating Lexi Wood?
- Where did Sylvie Vartan meet her future husband Johnny Hallyday?

Contributions in QUEST

- Answers complex questions directly over text using quasi KGs
- On-the-fly joining of multi-document evidence
- Robust to high degrees of noise
- Robust to ungrammatical constructs
- No need for question decomposition

Outline

- **Background:** Setup, benchmarks, metrics
- **Conversational QA:** Implicit context in multi-turn setup
- **Complex QA:** Multiple entities and predicates
- **Take-home:** Open problems and summary

Methodology: Pros and cons

Aspect	With explicit structured query (SPARQL-like)	Without explicit structured query (approx. graph search)
Simple questions	☺	☺
Single answer	☺	☺
List answer	☺	☹?
Efficiency	☺	☹?
Complex questions	☹?	☺
Conversational questions	☹?	☺
Heterogenous sources	☹?	☺
Handling reified triples	☹?	☺

☺ Preferable ☹? Less preferable but scope for improvement

Open problems

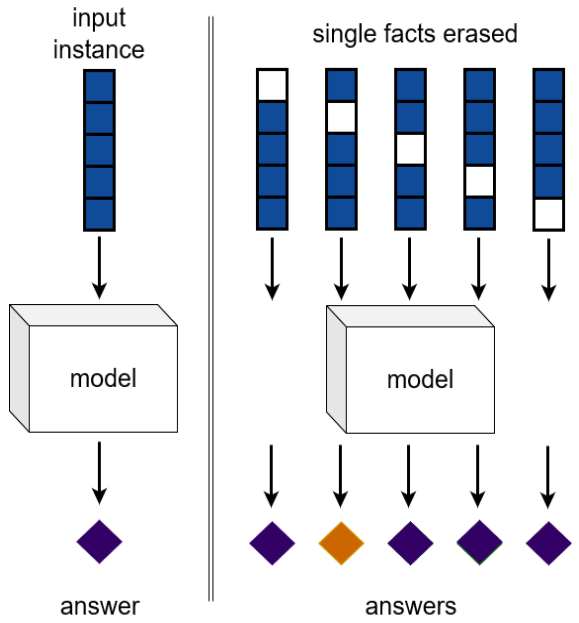
- Unanswerability
- Interpretability
- Interactivity
- Efficiency
- Robustness

Open problems: Unanswerability

- Learn when to stay quiet and prevent embarrassment ☺
  - *Where was Messi's father born?*
  - *Who was the first man on Mars?*
- Knowing when answer is:
  - Not confident
  - Not in KG
  - Null
- Open and closed world assumptions
- Learn when to consult text

Open problems: Interpretability

- Are your system's answers explainable? To the developer? What about the end user?
- Interpretability increases trust and guides user in case of mistakes
- Template- and graph-based methods construct interpretable evidence for answers - an unsolved concern for neural methods
- Sydorova et al. (2019) provide insights with **input perturbation** and **evaluation** of interpretability
- But very much an open problem!



Open problems: Interactivity

- Towards mixed initiative systems (Radlinski and Craswell 2017)
- Can your system absorb feedback?
- Positive and negative feedback?
- What kinds of feedback?
- Can your system ask clarifications?

Open problems: Efficiency

- Critical component of QA systems
- Largely unexplored
- Identify bottlenecks
- Measure trade-offs

Open problems: Robustness

- Think out of the ~~box~~ benchmark
- What is **open-domain** question answering?
- What happens for entities not seen during training?
- What about unseen predicates and vocabulary?

Summary

- Basic framework of KB-QA
- Graph methods for conversational and complex QA
  - Judicious context expansion helps deal with multi-turn scenario
  - Group Steiner Trees help in joint disambiguation for complex questions
- Several open problems in the key areas of focus

QA@MPII-D5: Visit qa.mpi-inf.mpg.de

- **Course** on QA systems: https://www.mpi-inf.mpg.de/question-answering-systems/
- **CONVEX:** Conversational QA over KGs [CIKM 2019]: https://convex.mpi-inf.mpg.de/
- **CROWN:** Conversational QA over passages [SIGIR 2020]: https://crown.mpi-inf.mpg.de/
- **QUEST:** Complex question answering [SIGIR 2019]: https://quest.mpi-inf.mpg.de/
- **ComQA:** QA benchmark with paraphrase [NAACL 2019]: http://qa.mpi-inf.mpg.de/comqa/
- **TEQUILA:** Temporal question answering [CIKM 2018]: https://tequila.mpi-inf.mpg.de/
- **QUINT:** Template-based question answering [EMNLP 2017]: https://quint.mpi-inf.mpg.de/
- Send an email to rishiraj@mpii.de in case of any issues!

Tutorial on QA

- Visit our SIGIR 2020 tutorial at

  https://www.avishekanand.com/talk/sigir20-tute/

Contents

1. **Question Answering over Knowledge Graphs** by Rishiraj Saha Roy. [Full Video]
   - Background.
   - Simple QA.
   - Complex QA.
   - Heterogeneous QA.
   - Conversational QA.
   - Summary.

2. **Question Answering over Textual Sources** by Avishek Anand
   - Background
   - Machine Reading Comprehension
   - Open-domain QA
   - Feedback and Interpretability
   - Conversational QA
   - Summary