

# Answering Complex Questions by Joining Multi-Document Evidence with Quasi Knowledge Graphs



Xiaolu Lu<sup>1</sup>, Soumajit Pramanik<sup>2</sup>, **Rishiraj Saha Roy**<sup>2</sup>, Abdalghani Abujabal<sup>3</sup>, Yafang Wang<sup>4</sup> and Gerhard Weikum<sup>2</sup>

<sup>1</sup>RMIT University, <sup>2</sup>**Max Planck Institute for Informatics**, <sup>3</sup>Amazon Alexa, <sup>4</sup>Ant Financial

# Question Answering: Vital for Search

- ★ Direct answers to questions
- ★ Saves time and effort
- ★ Natural in voice UI
- ★ Fact-centric, entity-seeking
- ★ Relies on knowledge graphs

# Question Answering: Vital for Search

- ★ Direct answers to questions
- ★ Saves time and effort
- ★ Natural in voice UI
- ★ Fact-centric, entity-seeking
- ★ Relies on knowledge graphs

What are some films directed by Nolan?

Christopher Nolan / Films directed



The Dark Knight  
2008



Interstellar  
2014

# Question Answering: Vital for Search

- ★ Direct answers to questions
- ★ Saves time and effort
- ★ Natural in voice UI
- ★ Fact-centric, entity-seeking
- ★ Relies on knowledge graphs

What are the Oscar nominations  
of Nolan?

Christopher Nolan Academy  
Awards Awards / Awards

**Best Picture**

2018 · Dunkirk

**Best Director**

2018 · Dunkirk

**Best Picture**

2011 · Inception

# Question Answering: Vital for Search

- ★ Direct answers to questions
- ★ Saves time and effort
- ★ Natural in voice UI
- ★ Fact-centric, entity-seeking
- ★ Relies on knowledge graphs

Which Nolan films won an Oscar?

This came back from a search



The “**Inception**” miss in particular was puzzling for a best picture nominee

# Question Answering: Vital for Search

- ★ Direct answers to questions
- ★ Saves time and effort
- ★ Natural in voice UI
- ★ Fact-centric, entity-seeking
- ★ Relies on knowledge graphs

Which Nolan films won an Oscar?

**Interstellar**

Academy Awards / Winners /...

# Question Answering: Vital for Search

- ★ Direct answers to questions
- ★ Saves time and effort
- ★ Natural in voice UI
- ★ Fact-centric, entity-seeking
- ★ Relies on knowledge graphs

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

Here you go. The first result is  
from Wikipedia

List of awards and  
nominations received by...  
<https://en.m.wikipedia.org...>

**Won.** 2010, Best Director of the  
Year ... Main article: Hochi **Film**  
Award .... (2002) Nominees **and...**

C  
-  
hi

C  
n  
B

# Direct Answers Today

- ★ Works for simple questions
- ★ Misses additional conditions
- ★ Depends upon KG completeness



# What's new?



- ★ **Entity** answers to **complex** questions directly from **text**
- ★ **Unsupervised method** using **Group Steiner Trees**
- ★ Handles **ungrammatical** constructs
- ★ **Robust** to high degrees of noise
- ★ Provides **explanatory** evidence

# Approach Outline



**QUEST: Question Answering with Steiner Trees**

# Multi-document Evidence

**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”**

# Semi-structuring Text with Open IE

<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>

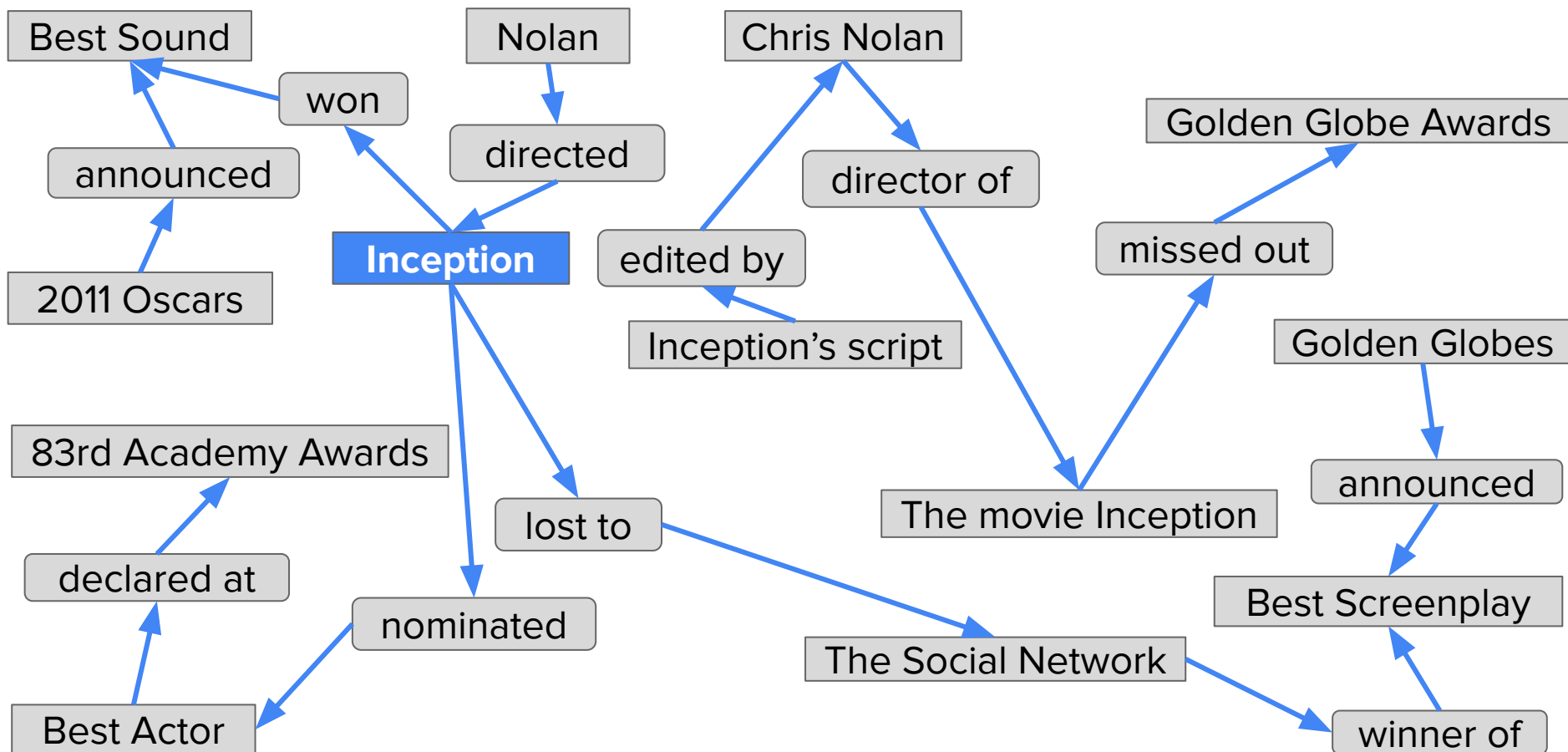
<Inception, lost to, The Social Network>

<Best Actor, declared at, 83rd Academy Awards>

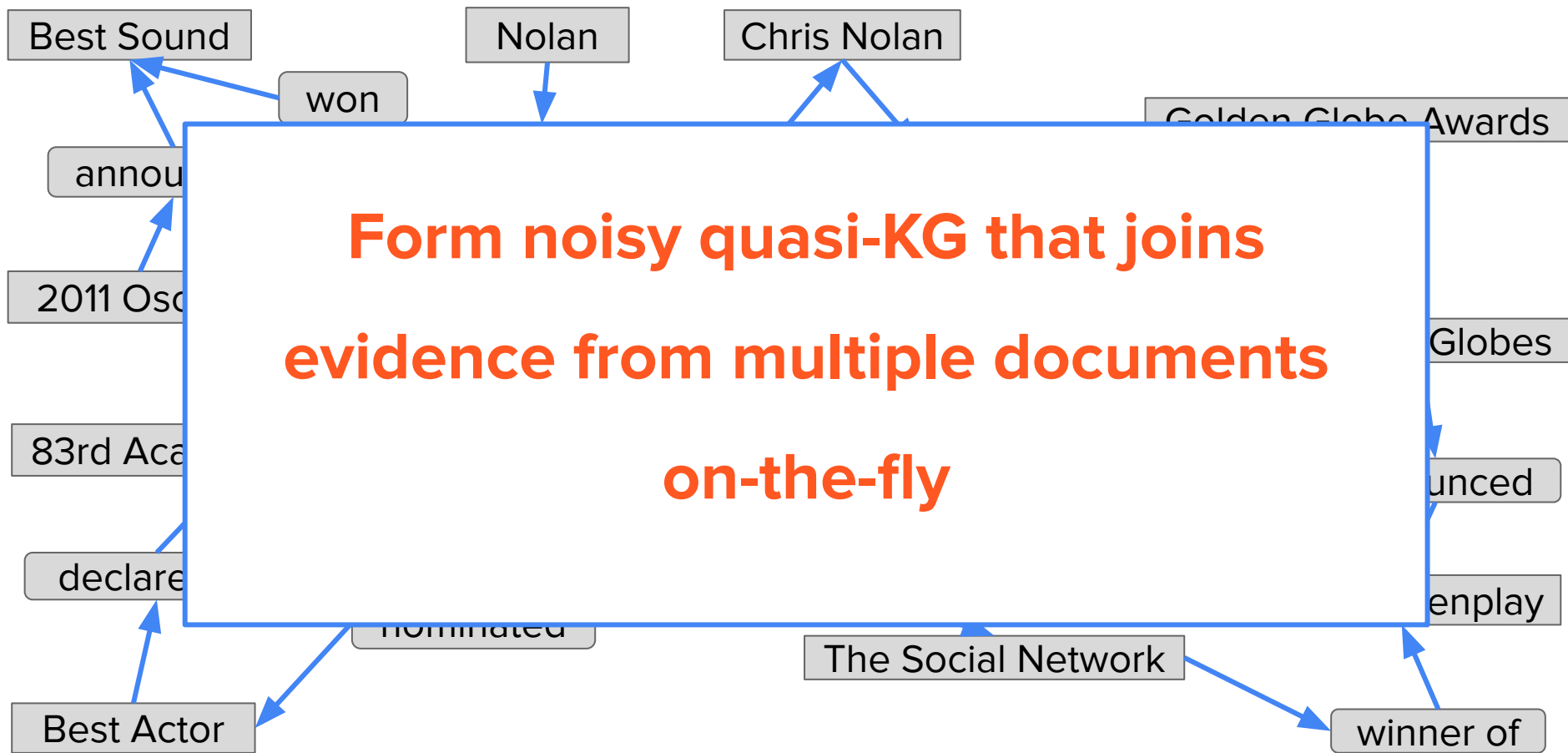
<The Social Network, winner of, Best Screenplay>

<Golden Globes, announced, Best Screenplay>

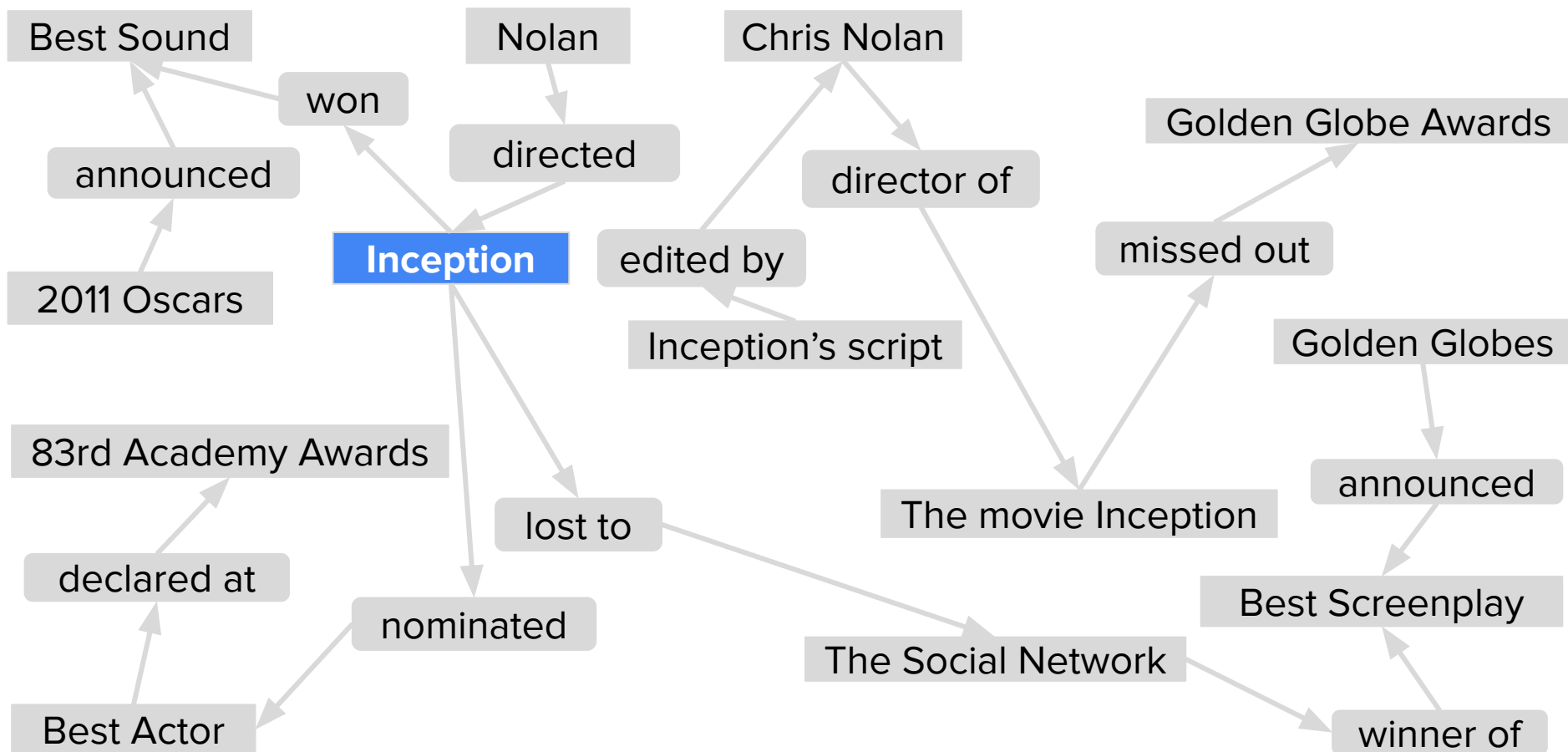
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



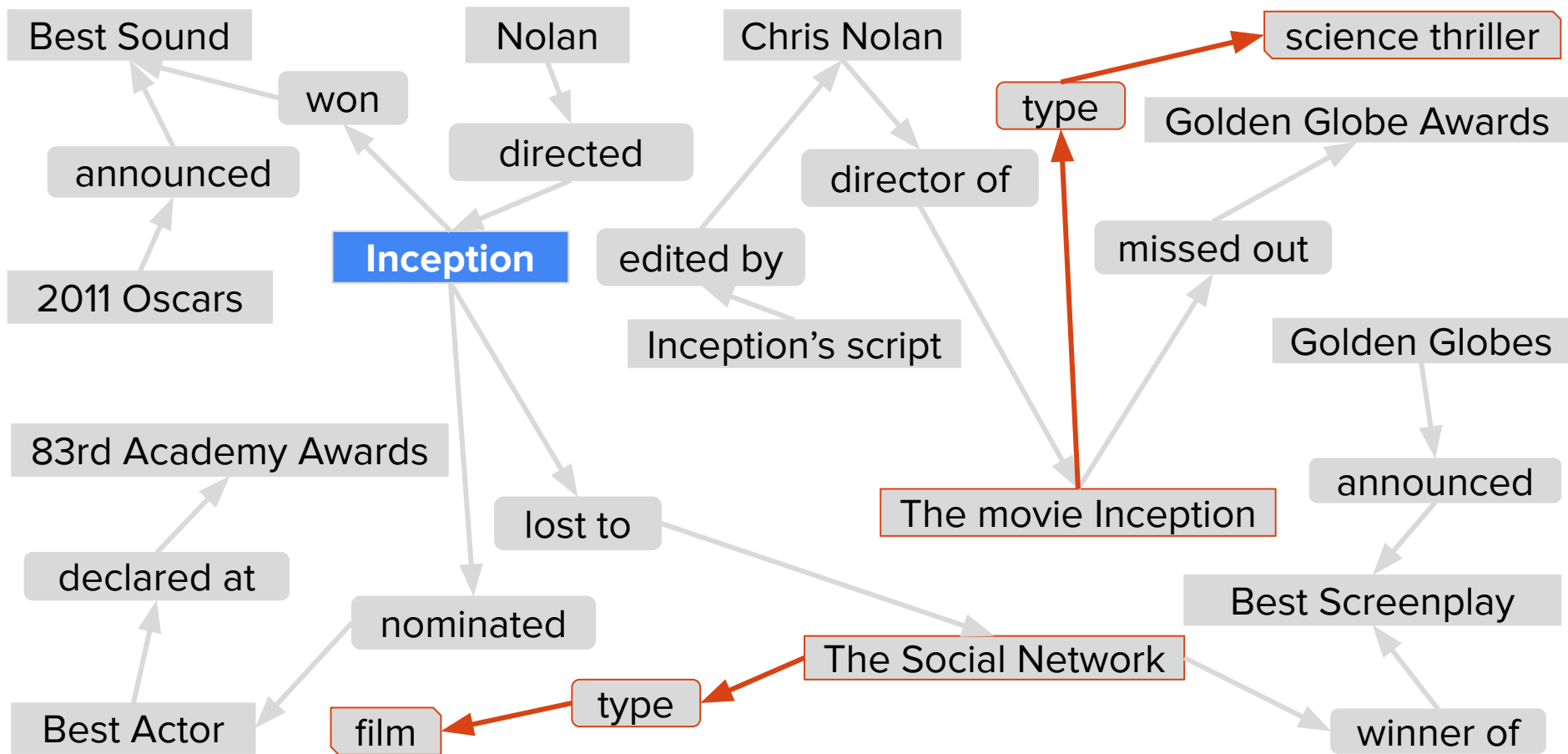
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



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

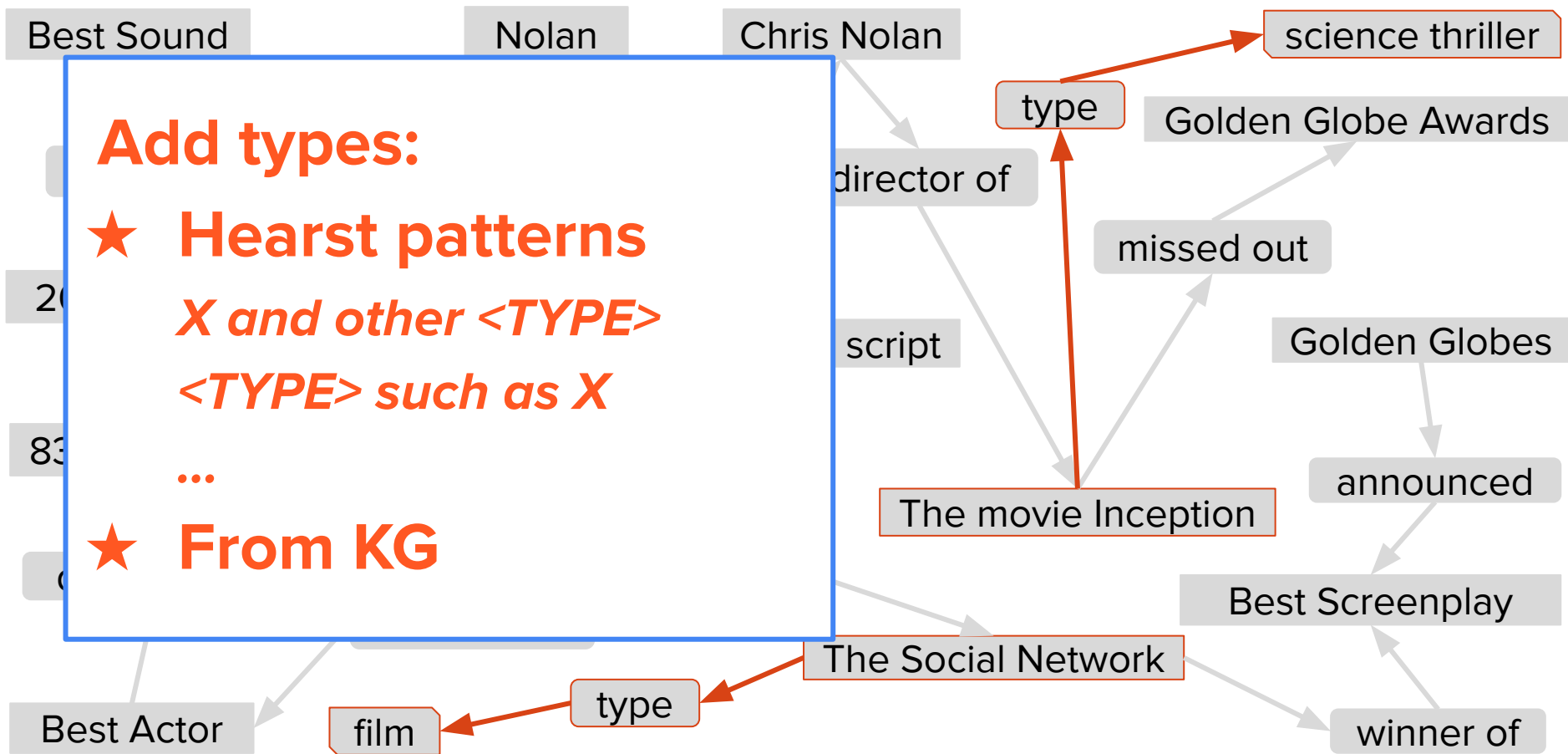


**Question:** Which Nolan films won an Oscar but missed a Golden Globe?

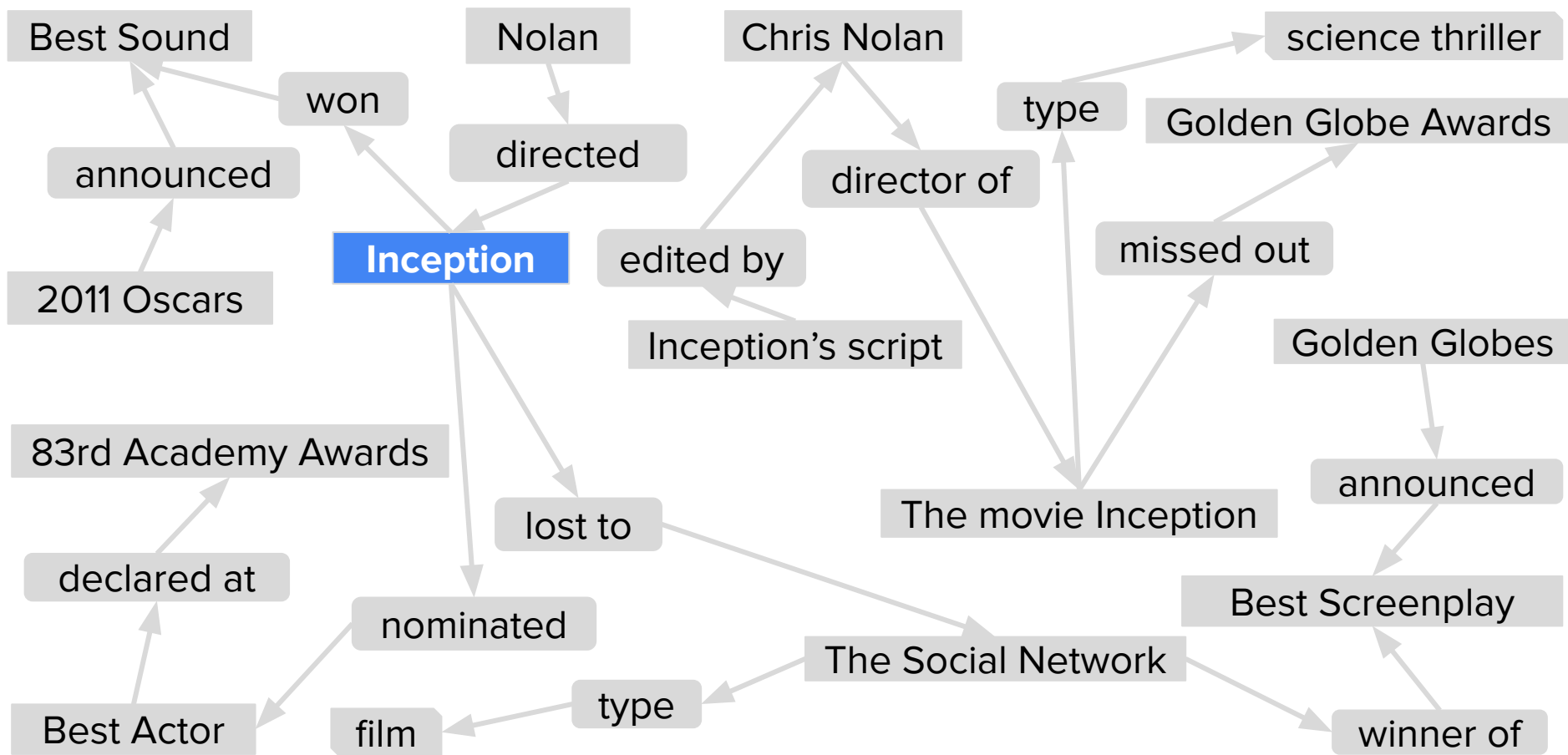




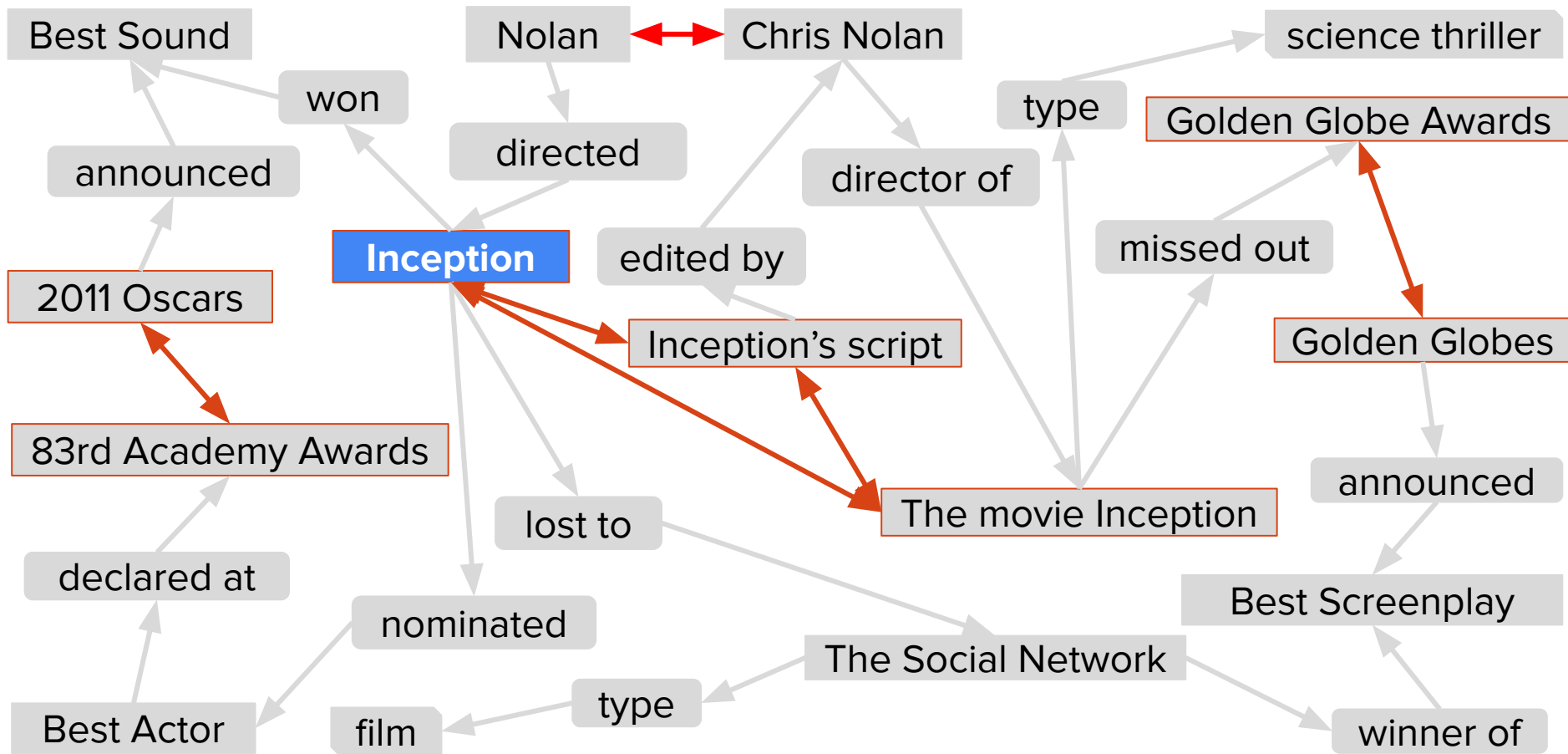
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



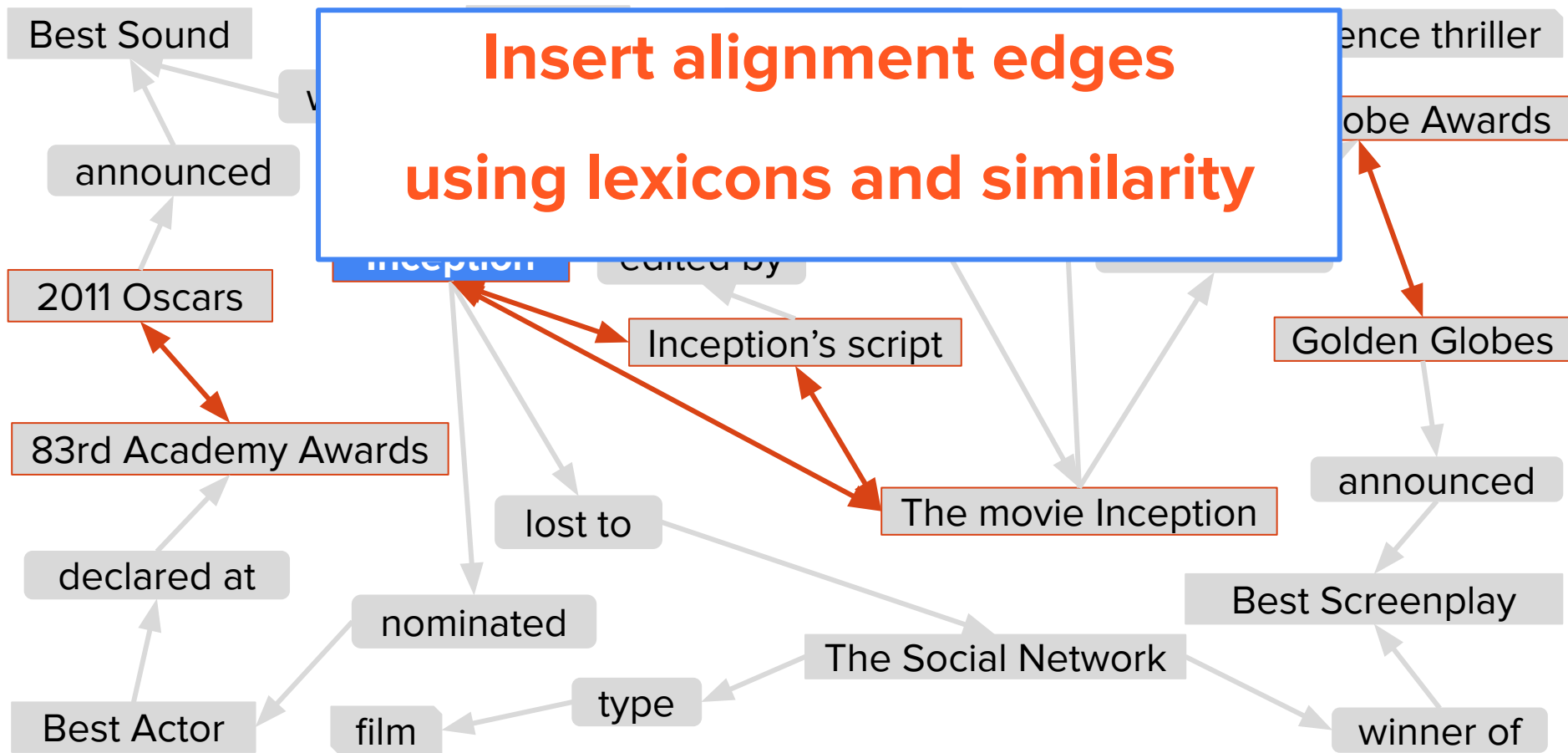
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



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

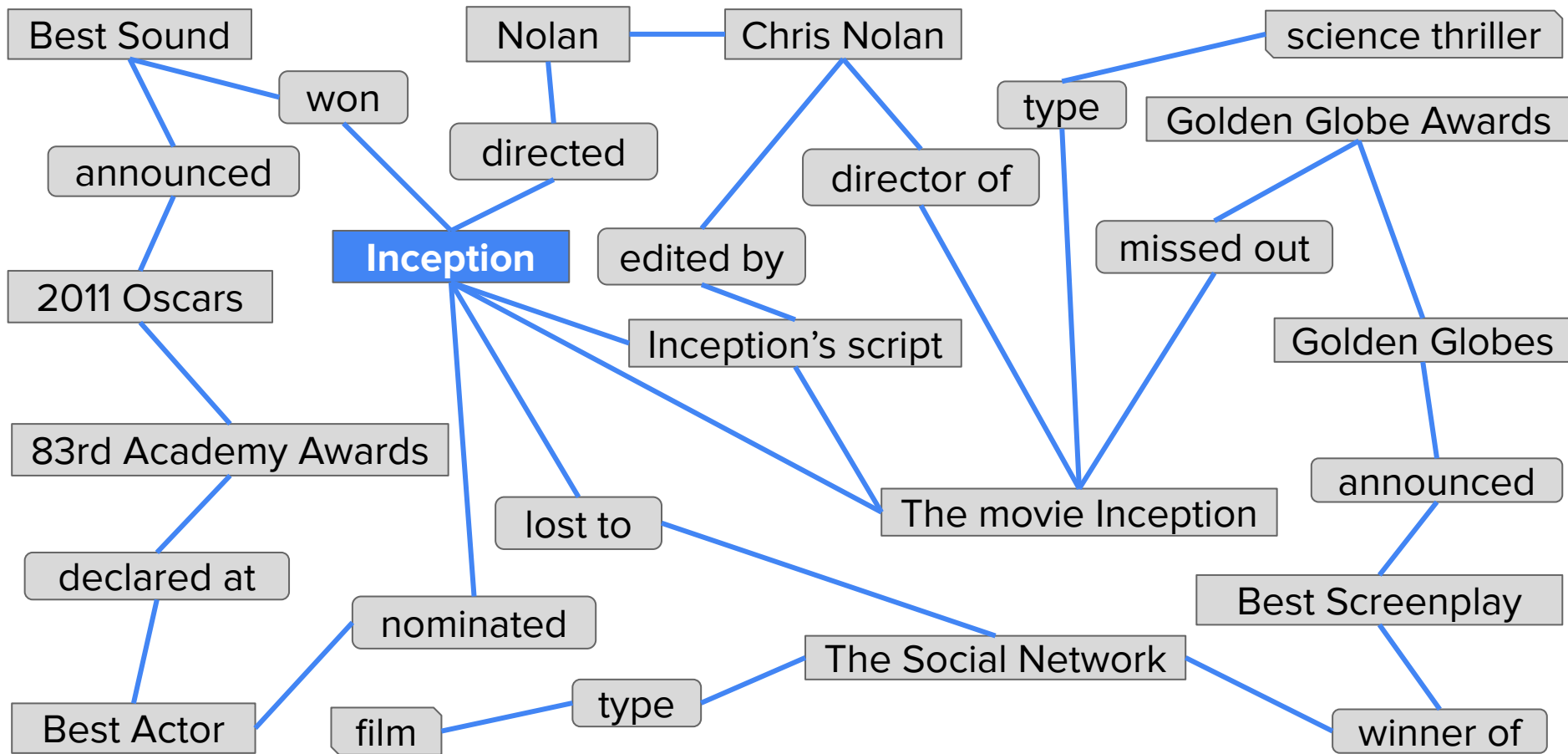


**Question:** Which Nolan films won an Oscar but missed a Golden Globe?

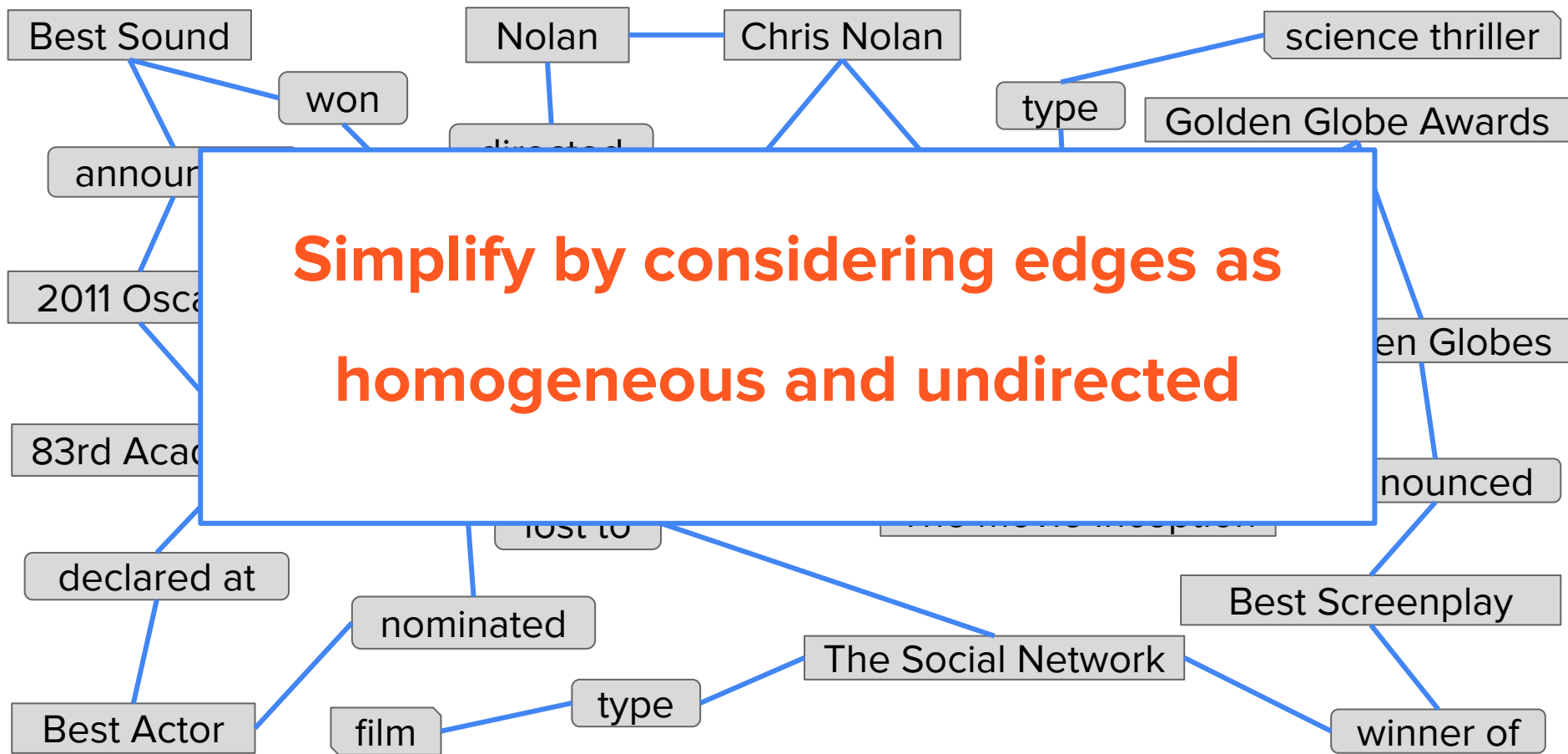




**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



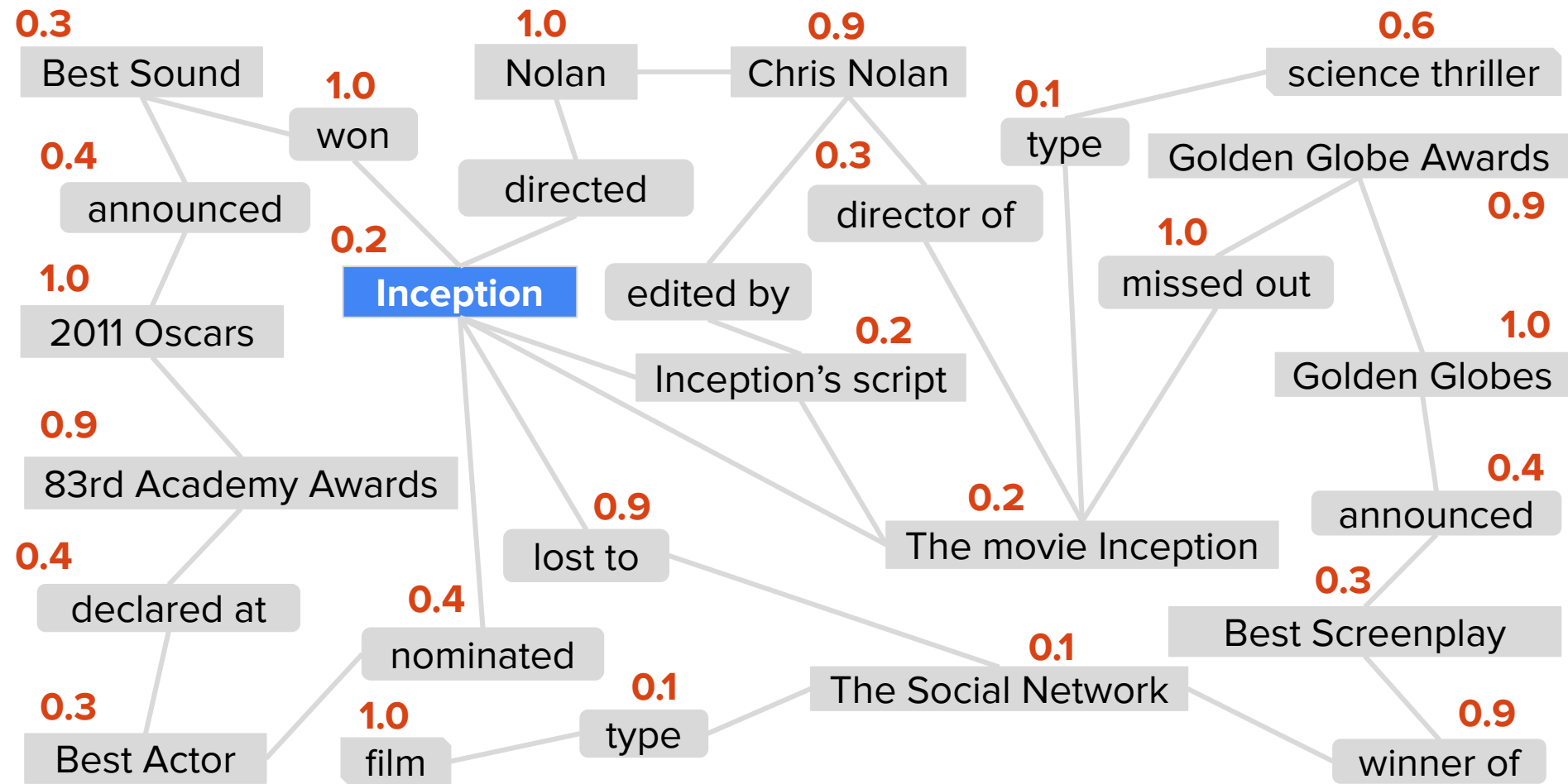
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



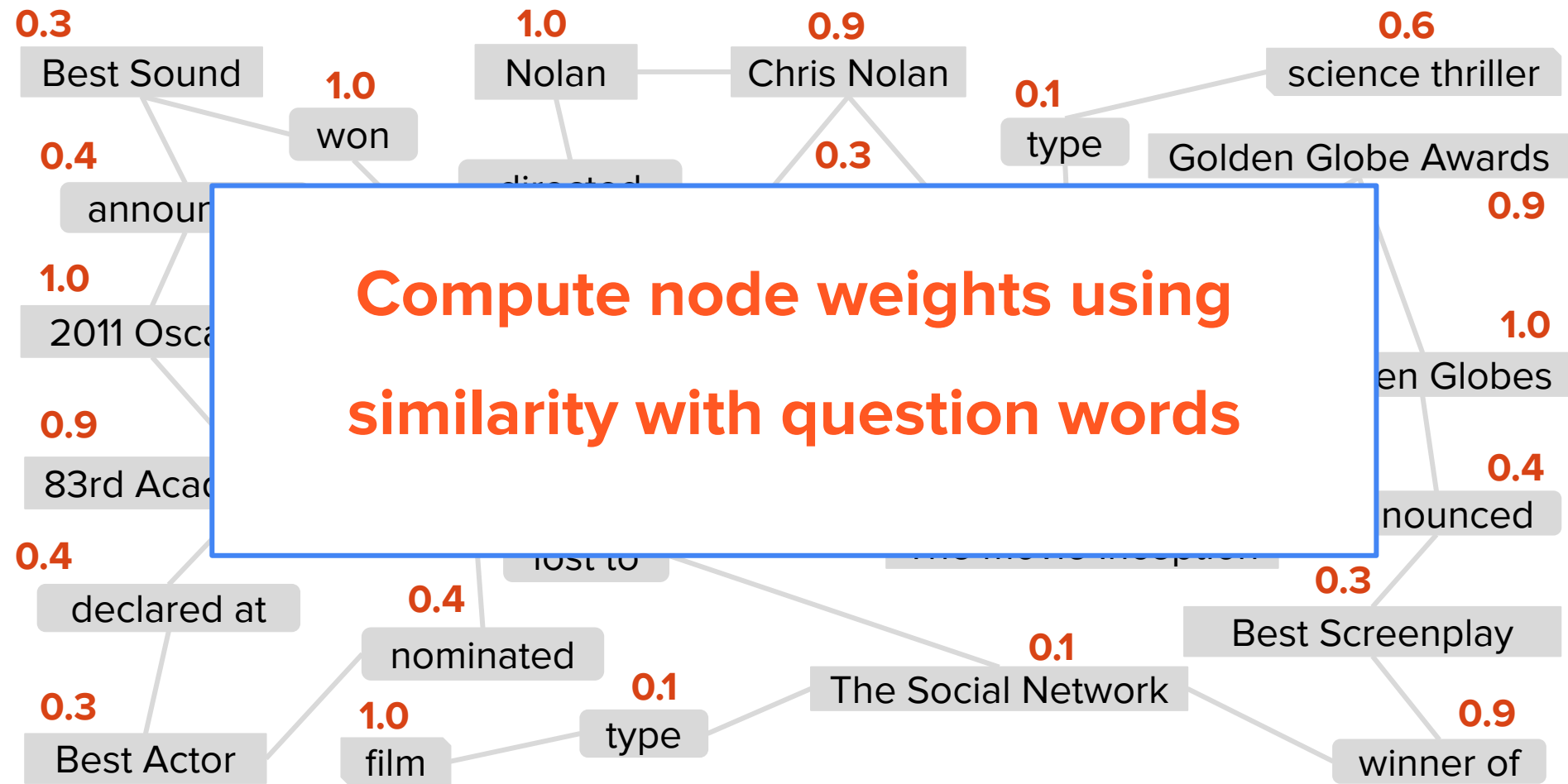




**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



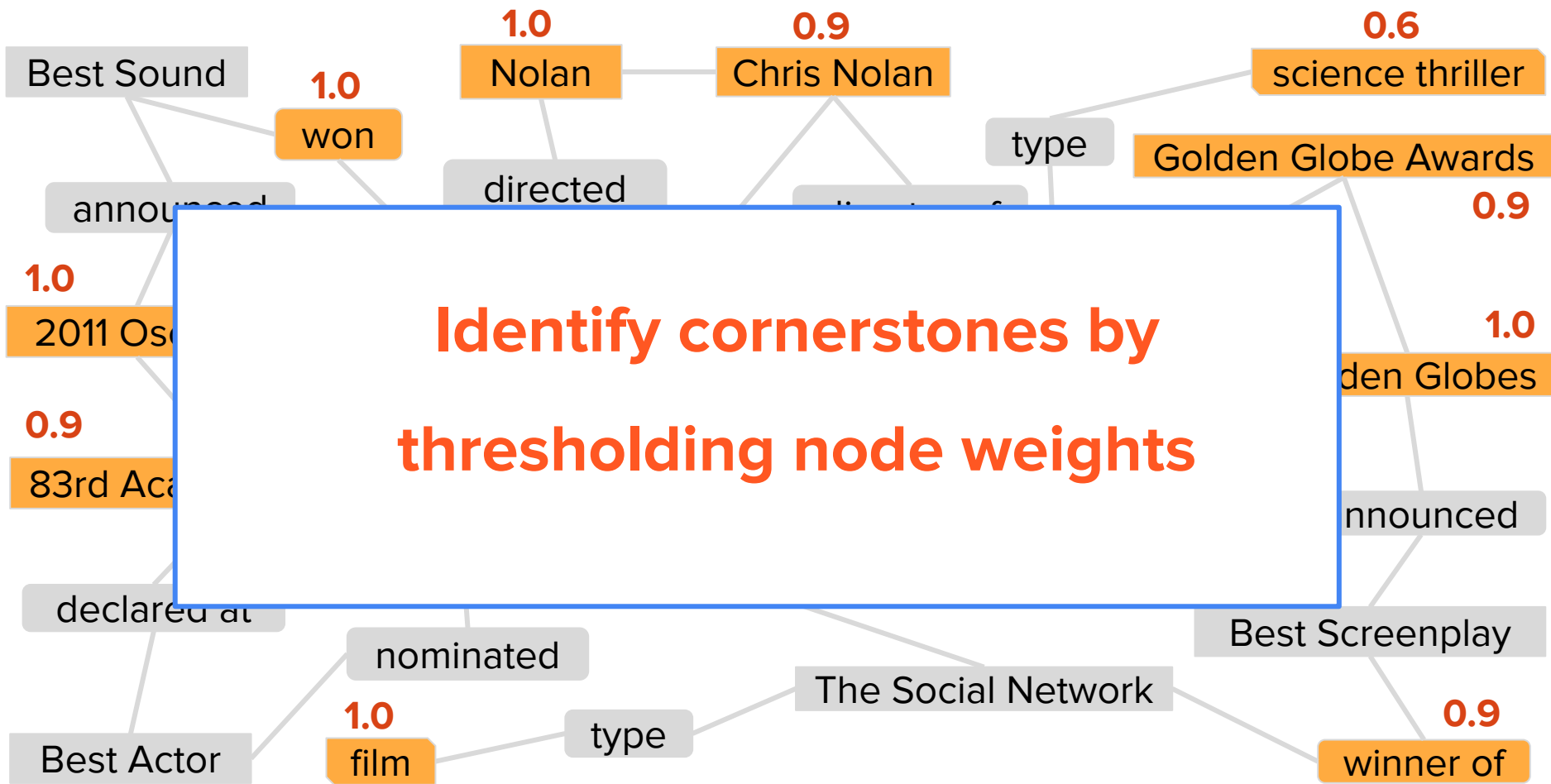
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?





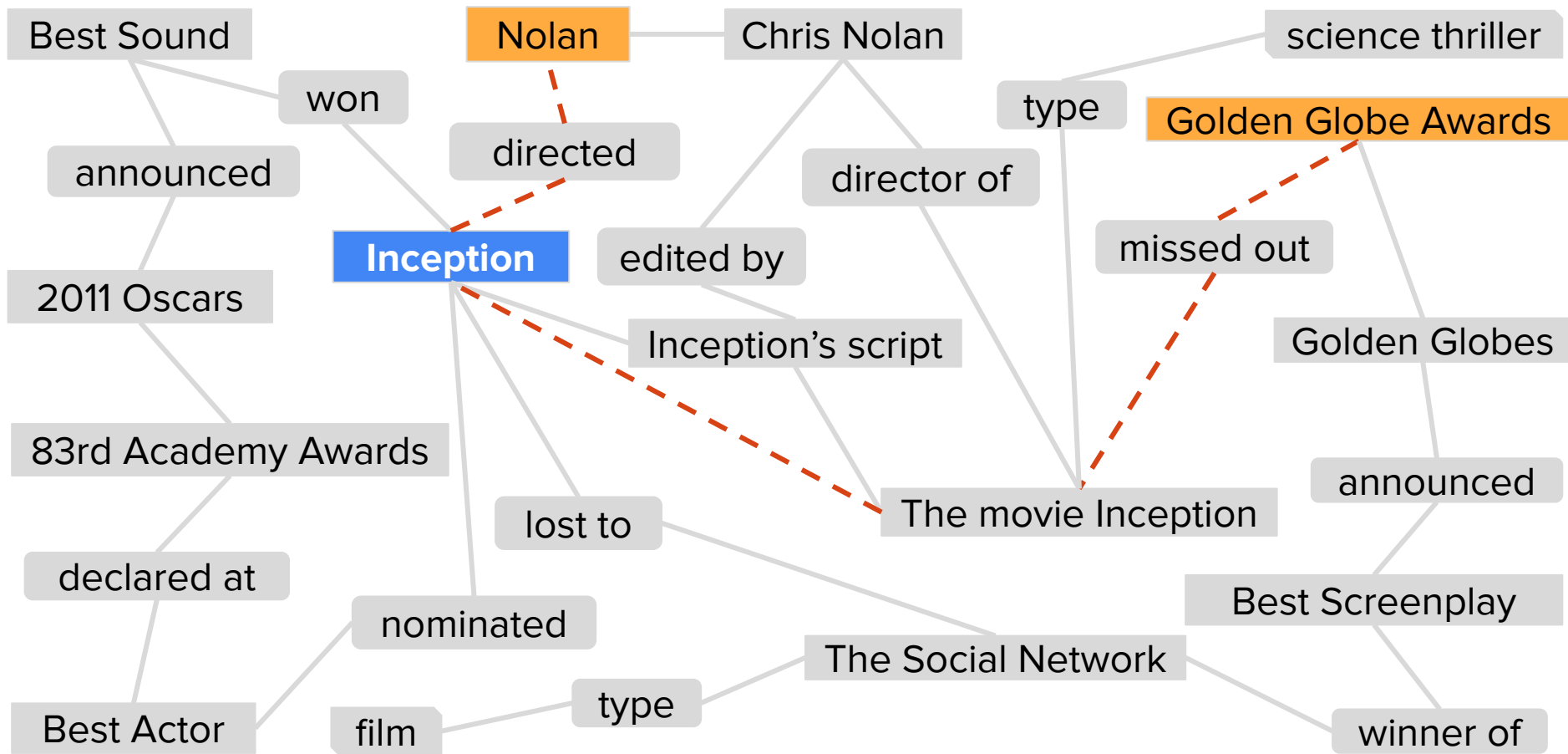


**Question:** Which Nolan films won an Oscar but missed a Golden Globe?

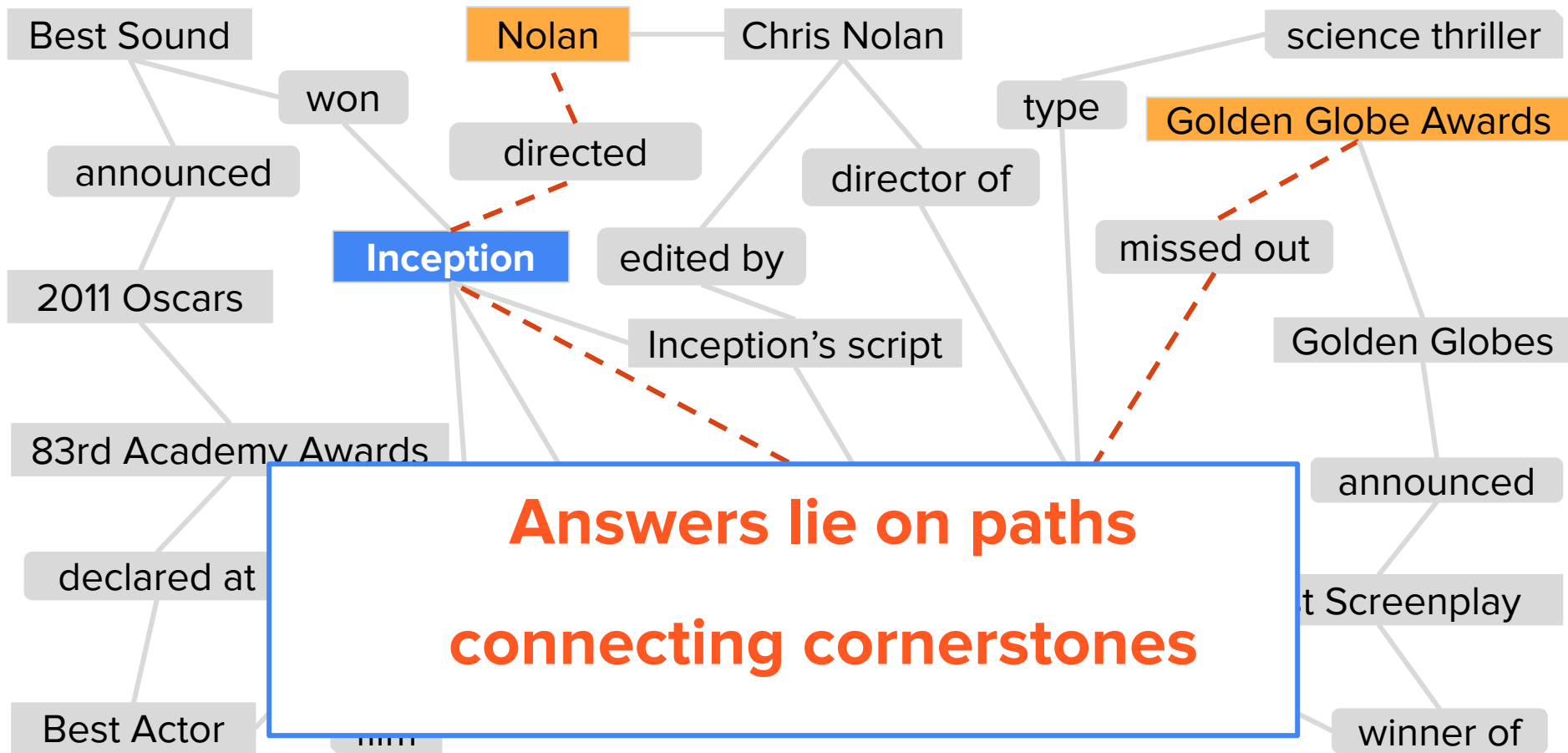




**Question:** Which Nolan films won an Oscar but missed a Golden Globe?

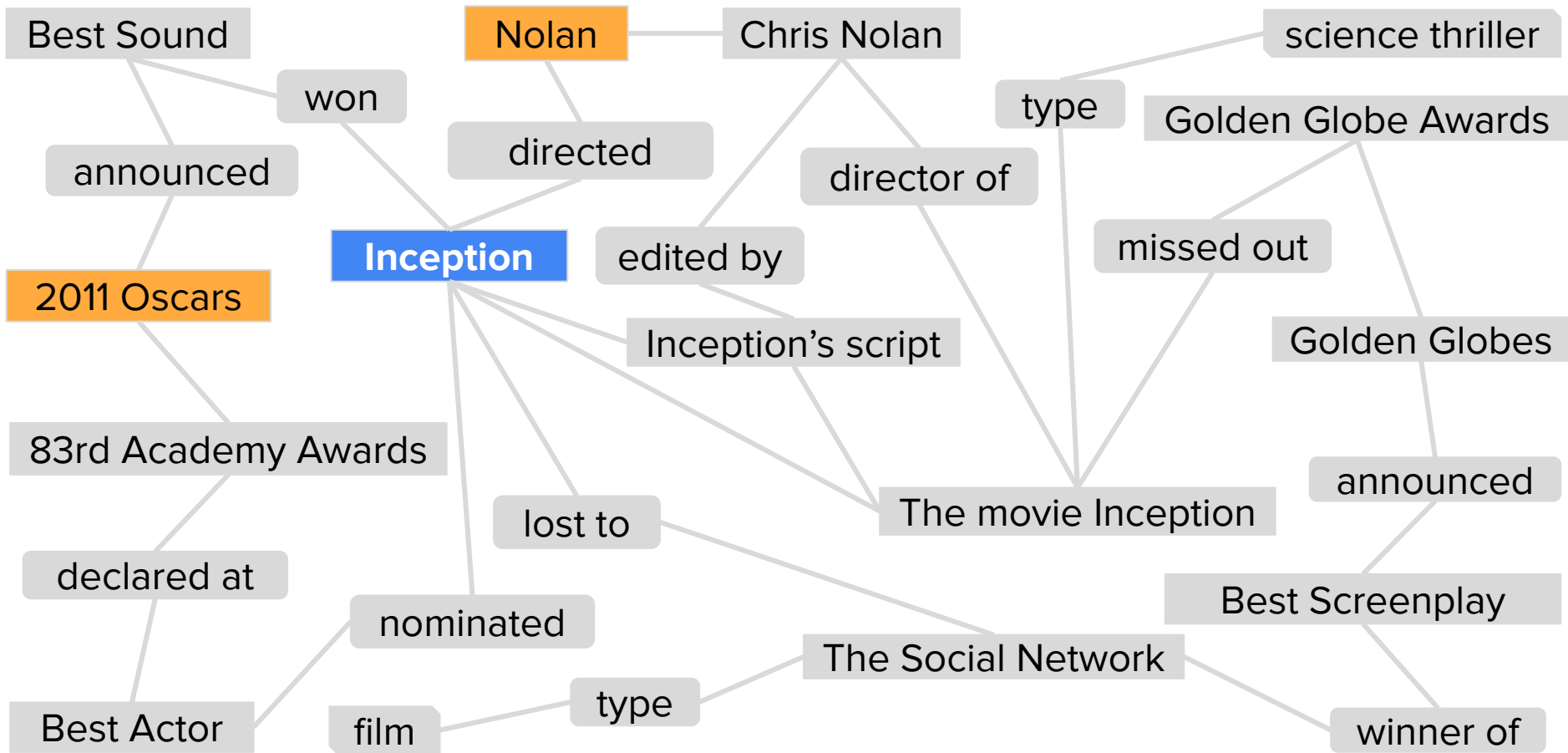


**Question:** Which Nolan films won an Oscar but missed a Golden Globe?

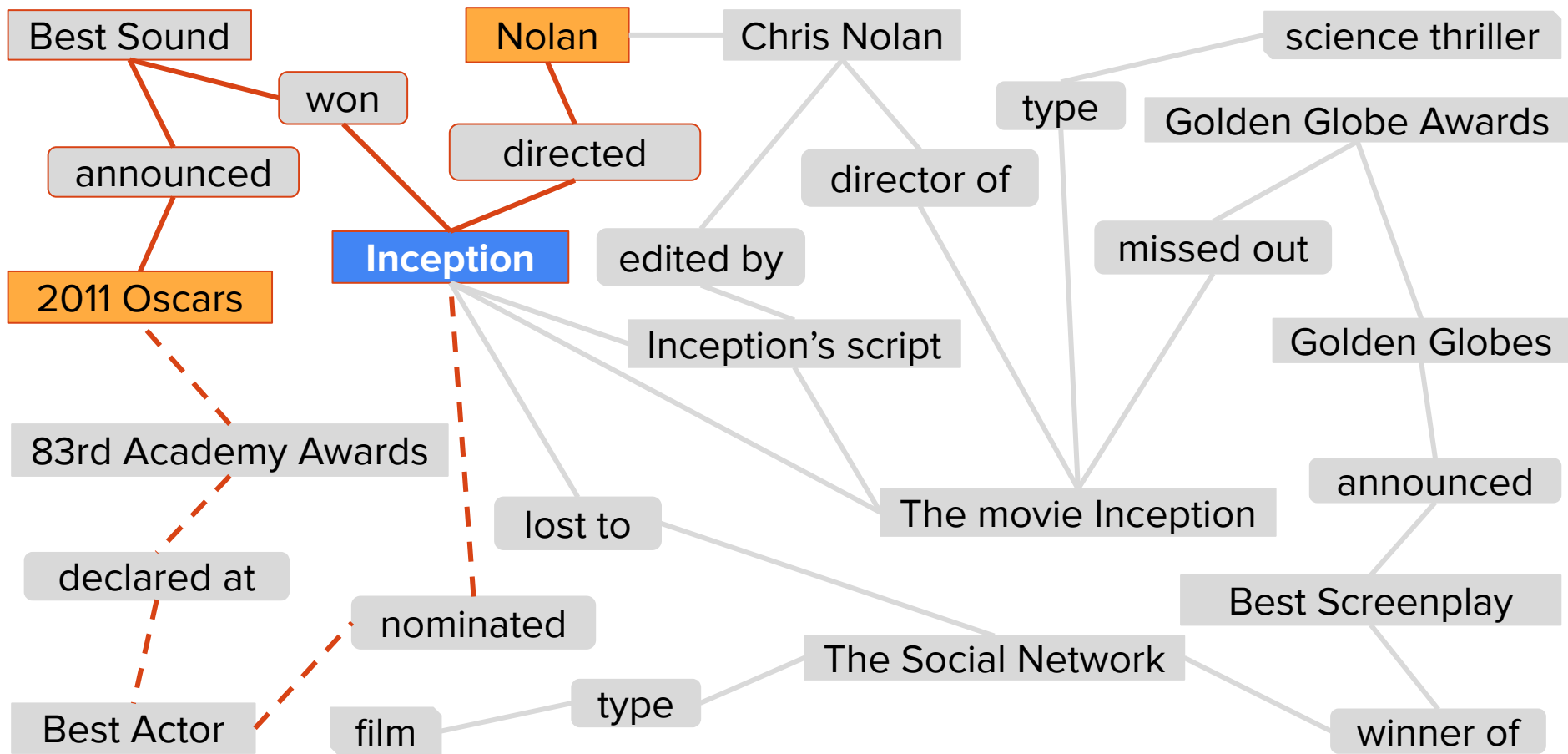




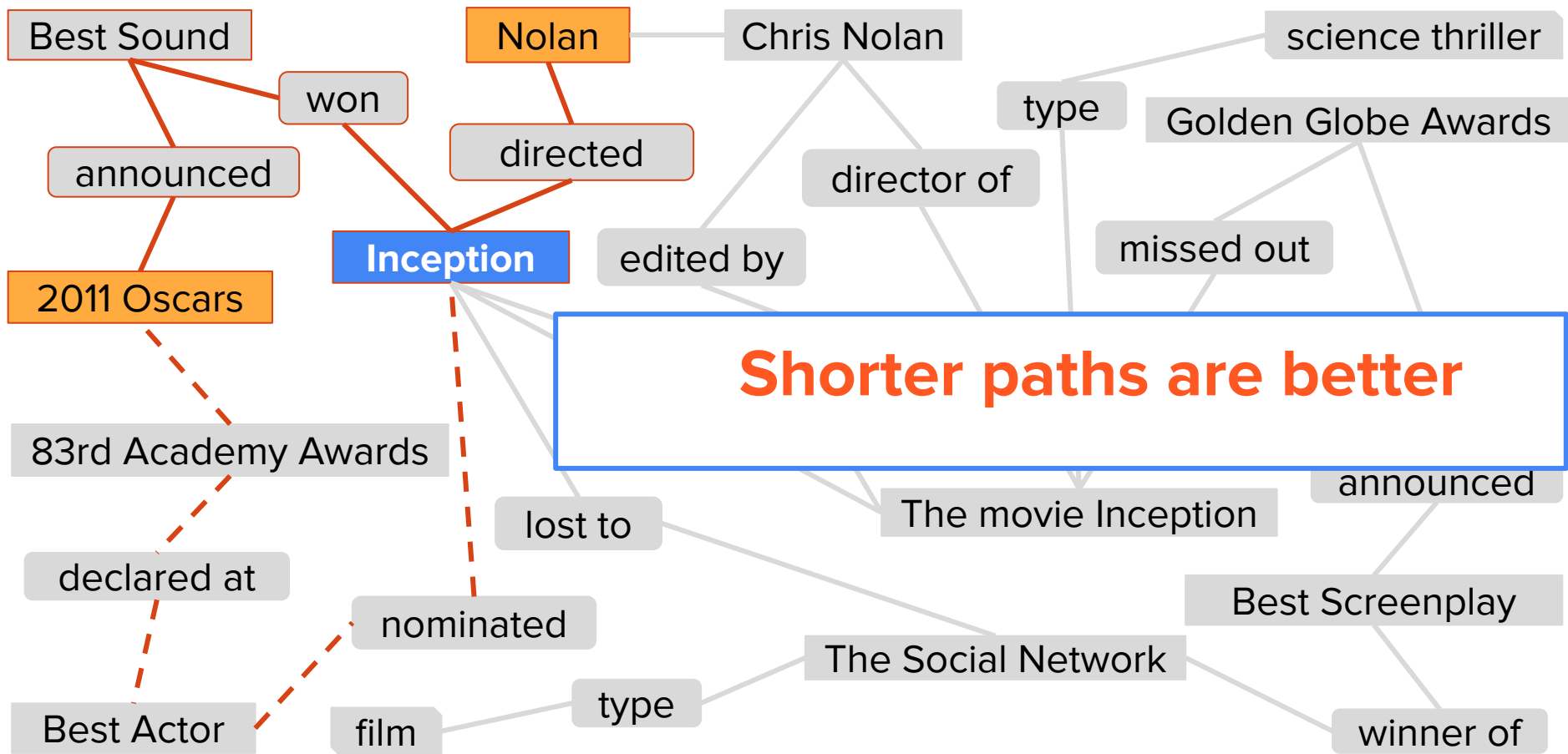
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



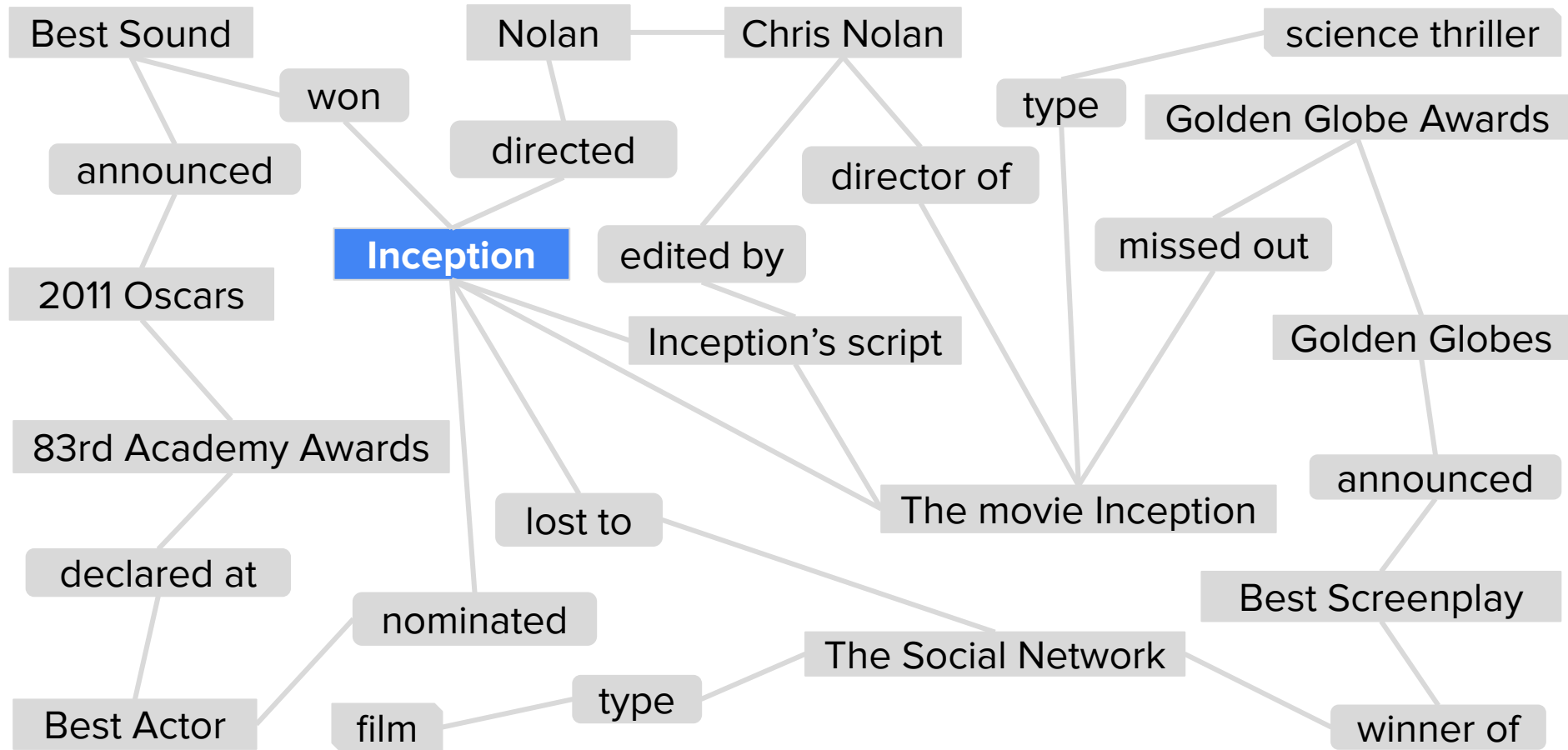
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



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



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

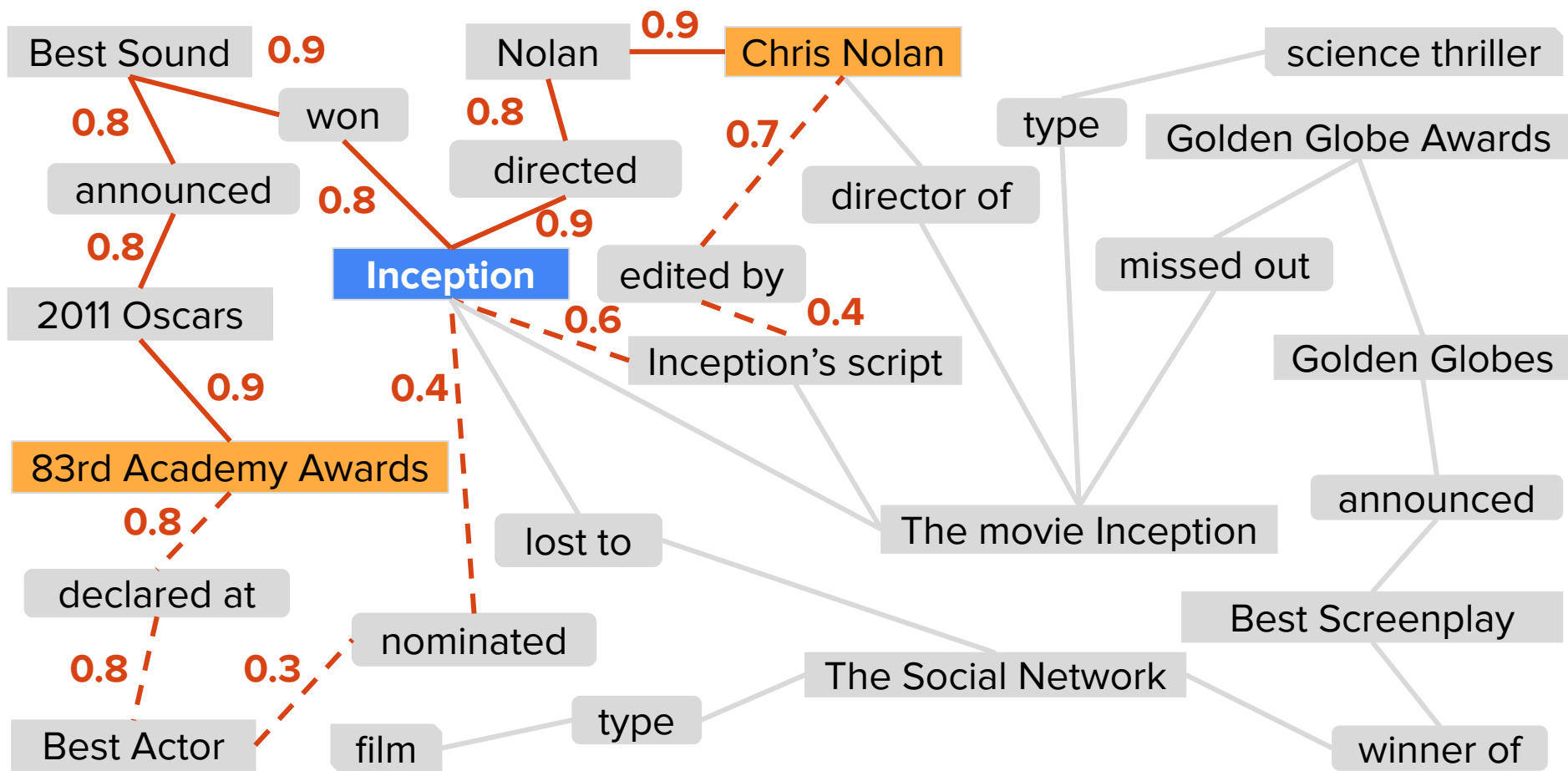






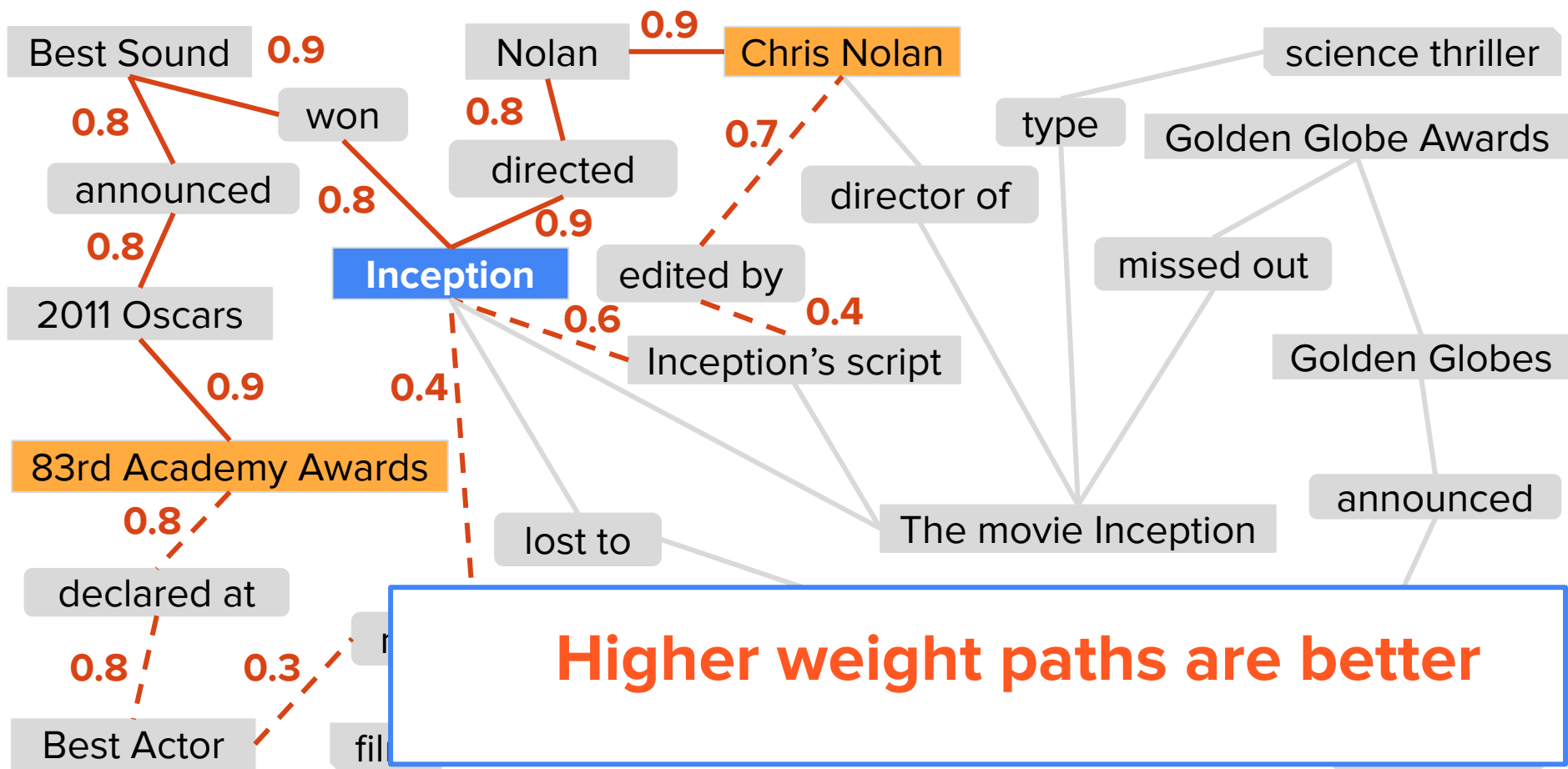


**Question:** Which Nolan films won an Oscar but missed a Golden Globe?





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



# Formalizing Intuitions: What we Have

★ Quasi-KG:  $G = (\underbrace{N, E}_{\text{Nodes and edges}}, \underbrace{M_N^T, M_E^T}_{\text{Mapping nodes and edges to types}}, \underbrace{M_N^W, M_E^W}_{\text{Mapping nodes and edges to weights}})$

Nodes and edges

Mapping nodes and  
edges to types

Mapping nodes and  
edges to weights

- ★ Answers on paths connecting cornerstones
- ★ Higher weight paths are preferred
- ★ Shorter paths are preferred

# Steiner Trees

## ★ Given:

- Undirected and weighted graph  $G$
- Subset of vertices = Terminals

## ★ Find:

- Minimum cost tree containing all terminals

## ★ Cornerstones = Terminals!

## ★ Costs = 1 - weights

$$G = (N, E)$$

$$w_{ij} \geq 0$$

$$T \subseteq N$$

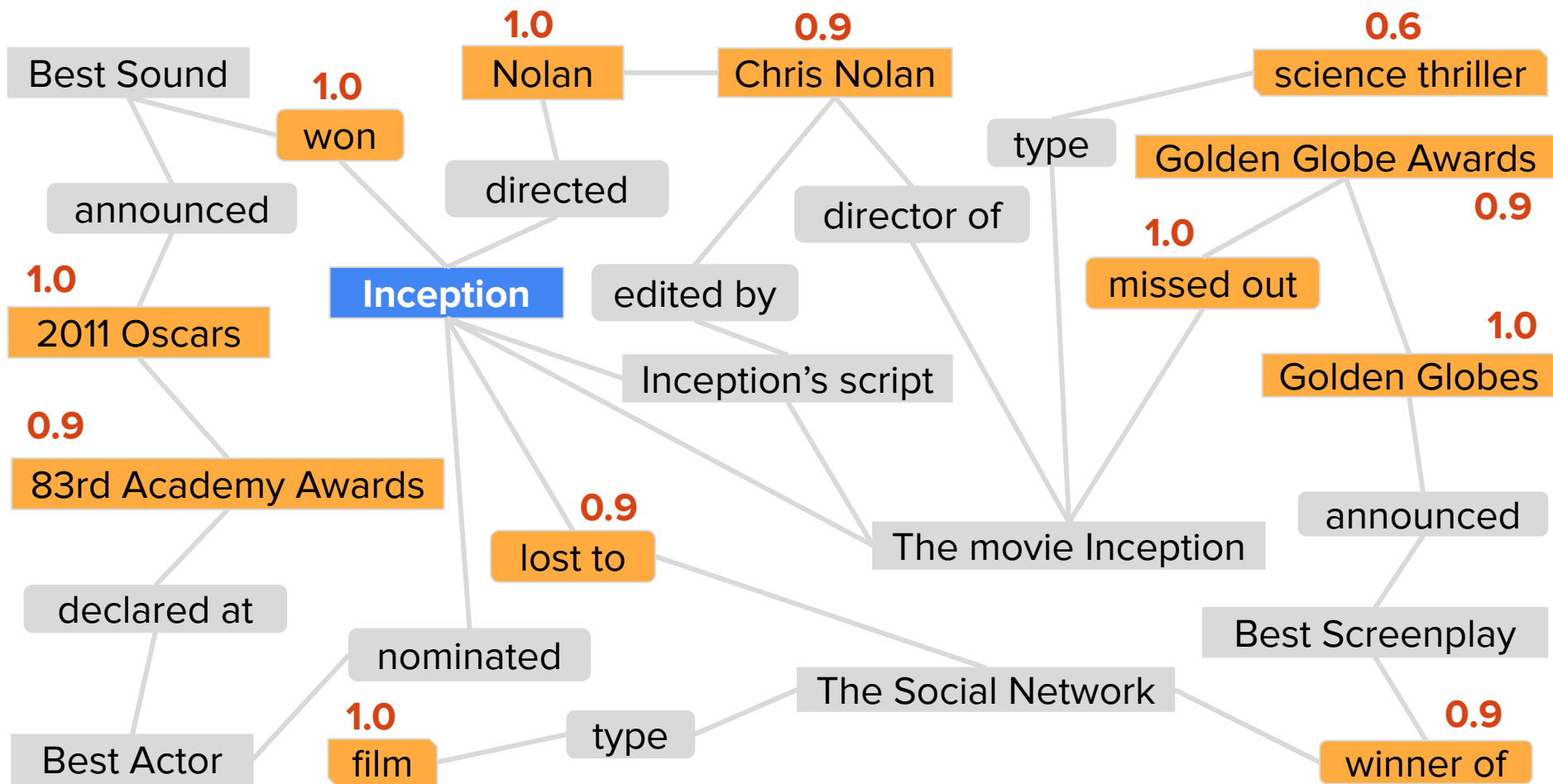
$$\mathcal{T} = (N^*, E^*)$$

$$N^* \subseteq N, E^* \subseteq E$$

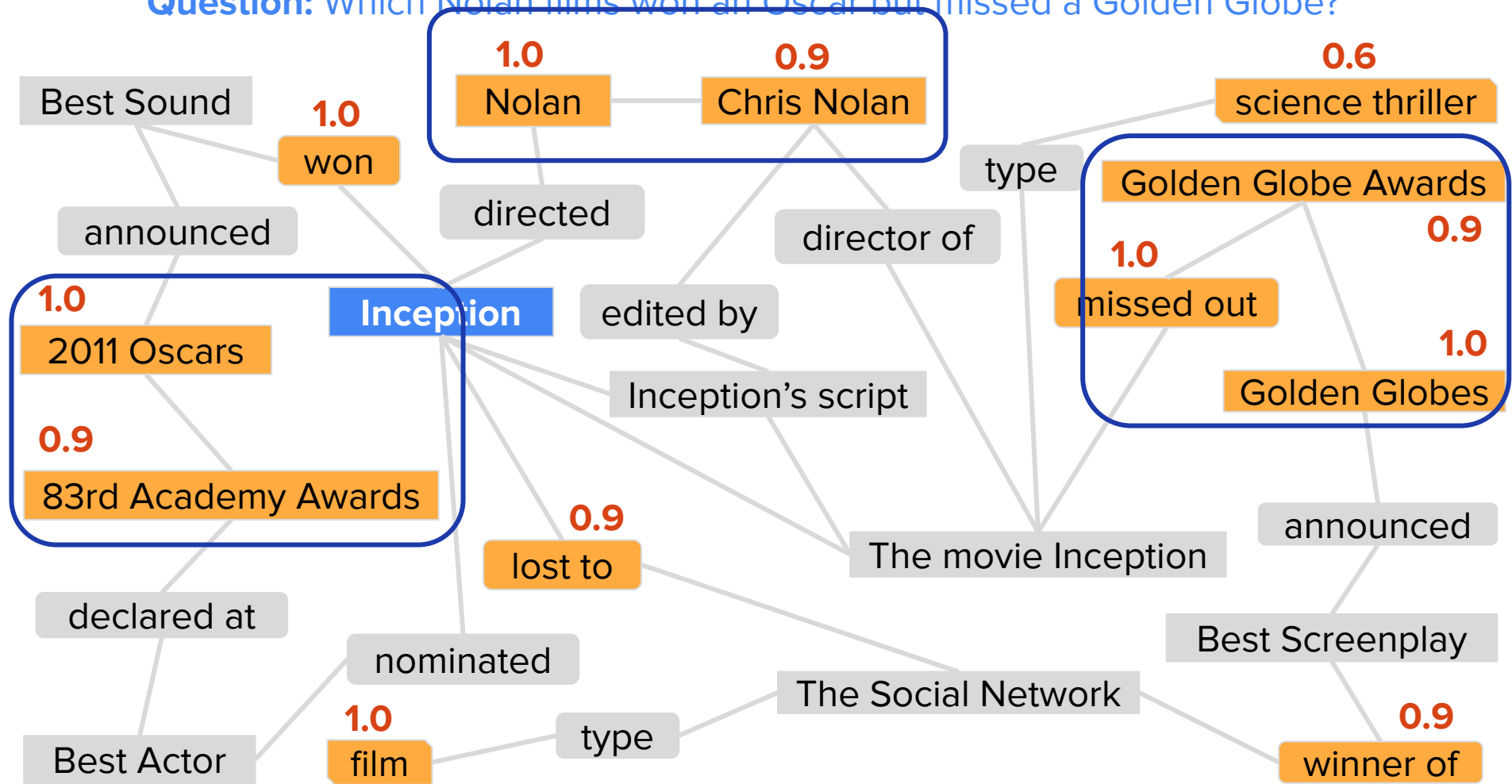
$$\min \sum_{ij \in E^*} w_{ij}$$

$$T \subseteq N^*$$

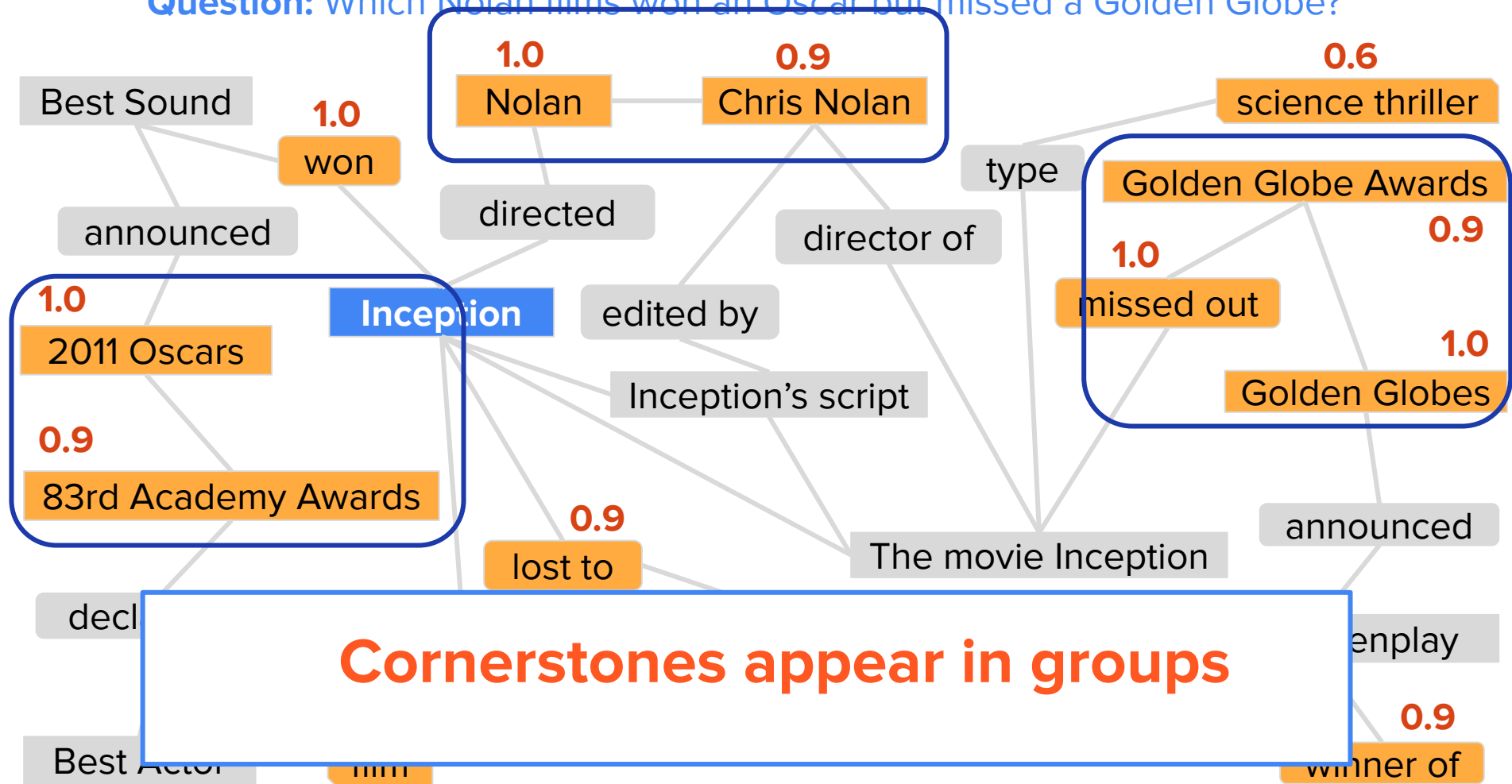
**Question:** Which Nolan films won an Oscar but missed a Golden Globe?



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



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



# Group Steiner Trees (GST)

- ★ Contains **at least one terminal** from each group
- ★ **Compute Group Steiner Trees on quasi-KG**
- ★ Method from Ding et al. [ICDE 2007]
- ★ Dynamic programming exponential in #keywords
- ★  $O(n \log n)$  in graph size

$$G = (N, E)$$

$$w_{ij} \geq 0$$

$$T \subseteq N$$

$$\mathcal{T} = (N^*, E^*)$$

$$N^* \subseteq N, E^* \subseteq E$$

$$\min \sum_{ij \in E^*} w_{ij}$$

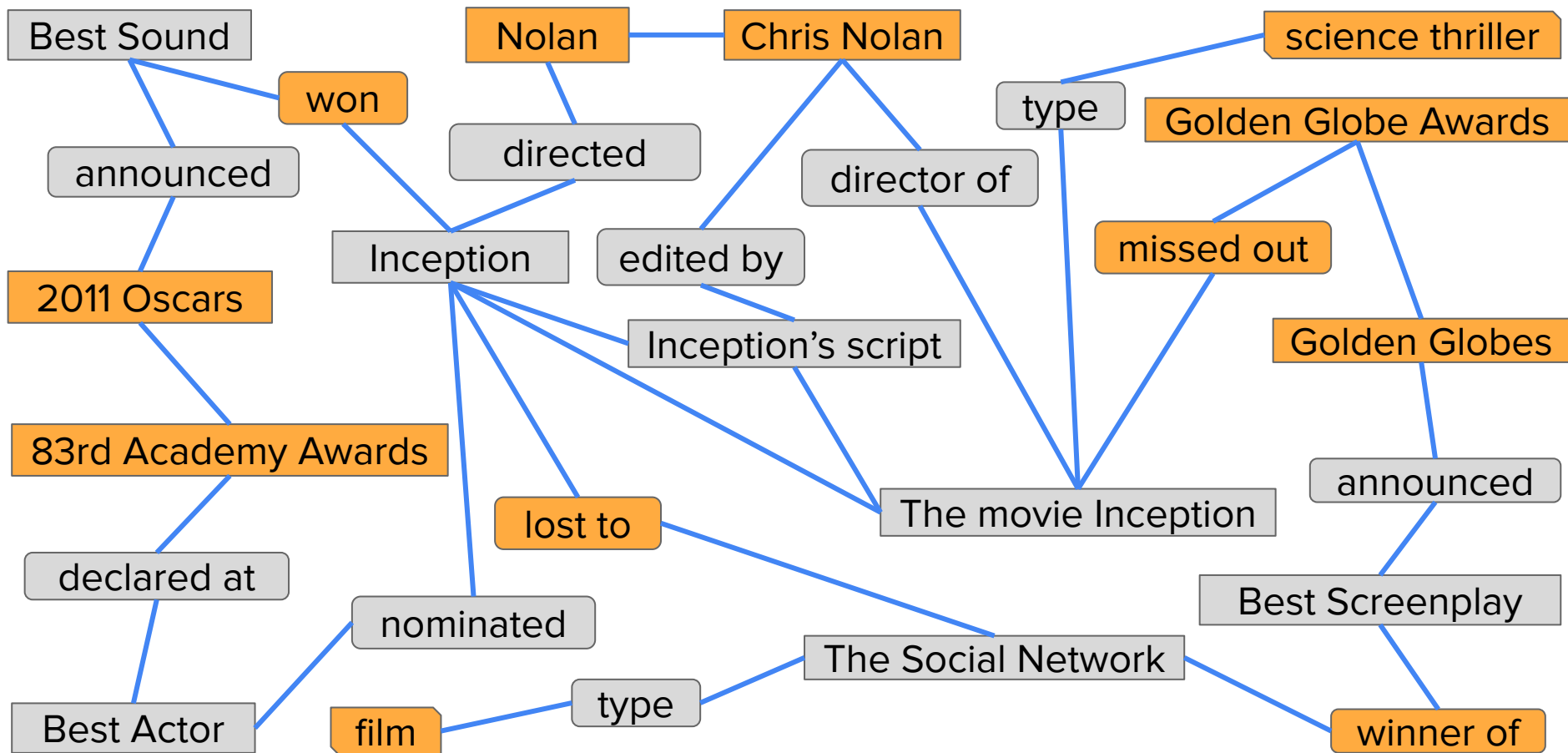
$$T \subseteq N^*$$

---

$$T = \{T_n\}, T_n \subseteq N$$

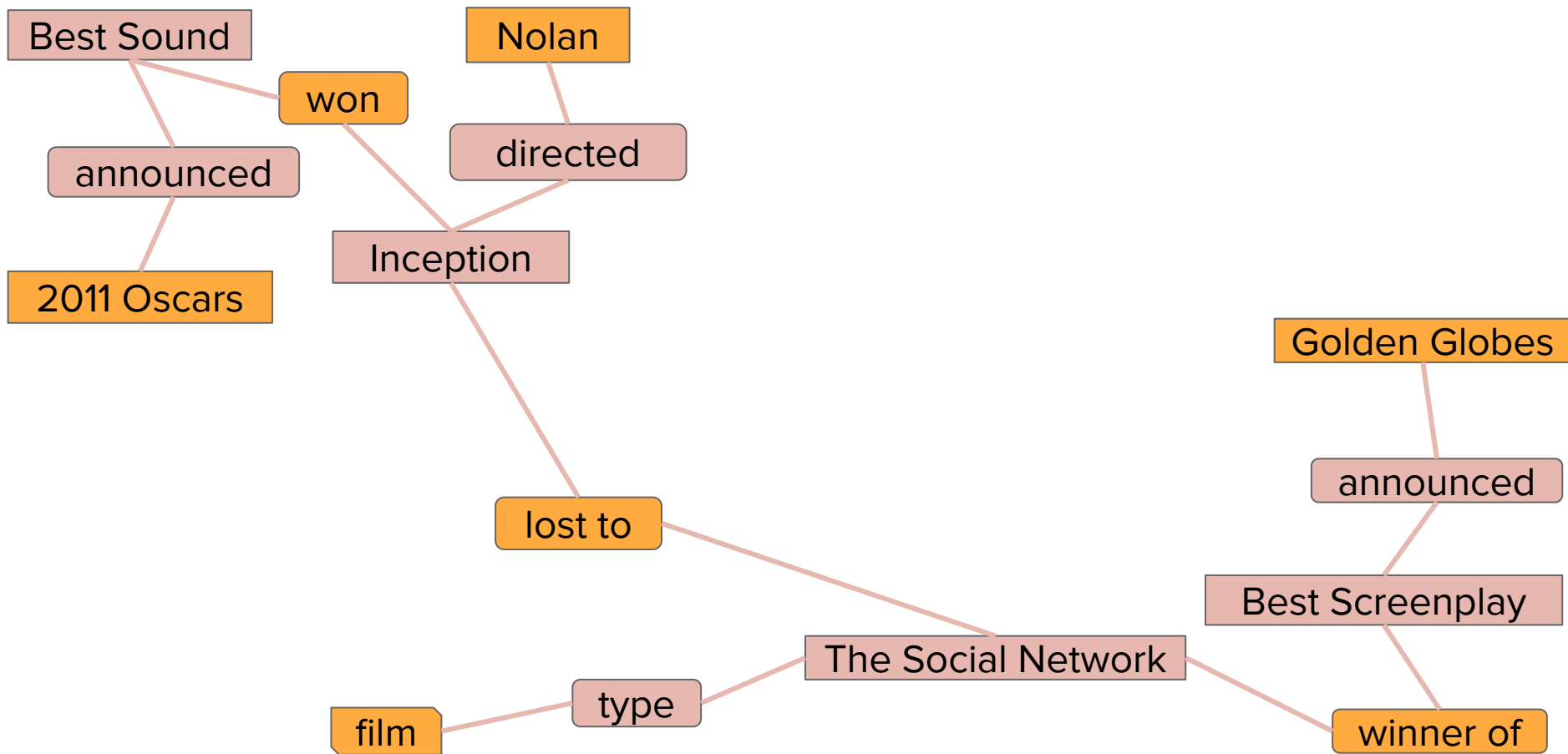
$$T_n \cap N^* = \phi, \forall T_n$$

**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?

Best Sound

Nolan

won

announced

directed

Inception

2011 Oscars

Golden Globes

announced

lost to

Best Screenplay

The Social Network

film

type

winner of

★ **Answers are nodes**

★ Cornerstones are not answers

★ Answers are entities

★ Answers respect type constraints

★ Answers ranked by intuitive criteria

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

Best Sound

won

announced

directed

Inception

lost to

type

- ★ Answers are nodes
- ★ **Cornerstones are not answers**
- ★ Answers are entities
- ★ Answers respect type constraints
- ★ Answers ranked by intuitive criteria

announced

Best Screenplay

The Social Network

winner of

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

Best Sound

Inception

- ★ Answers are nodes
- ★ Cornerstones are not answers
- ★ **Answers are entities**
- ★ Answers respect type constraints
- ★ Answers ranked by intuitive criteria

Best Screenplay

The Social Network

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

Inception

- ★ Answers are nodes
- ★ Cornerstones are not answers
- ★ Answers are entities
- ★ **Answers respect type constraints**
- ★ Answers ranked by intuitive criteria

The Social Network

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

### Inception

Sum of node weights	<b>9.75</b>
Number of GSTs	<b>5</b>
Sum of tree costs	<b>6.78</b>
Distance to cornerstones	<b>19</b>

- ★ Answers are nodes
- ★ Cornerstones are not answers
- ★ Answers are entities
- ★ Answers respect type constraints
- ★ **Answers ranked by intuitive criteria**

Sum of node weights	8.56
Number of GSTs	2
Sum of tree costs	7.98
Distance to cornerstones	27

### The Social Network

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

### Inception

Sum of node weights	<b>9.75</b>
Number of GSTs	<b>5</b>
Sum of tree costs	<b>6.78</b>
Distance to cornerstones	<b>19</b>

- ★ Answers are nodes
- ★ Cornerstones are not answers
- ★ Answers are entities
- ★ Answers respect type constraints
- ★ **Answers ranked by intuitive criteria**

# Experimental Results: Setup

## ★ Benchmarks

- 150 Complex questions from WikiAnswers [Abujabal et al. 2017]
- 150 Complex questions from Google Trends
- **Corpus:** Google Top-10 / Stratified sampling from Web / News / Wikipedia

## ★ Metric

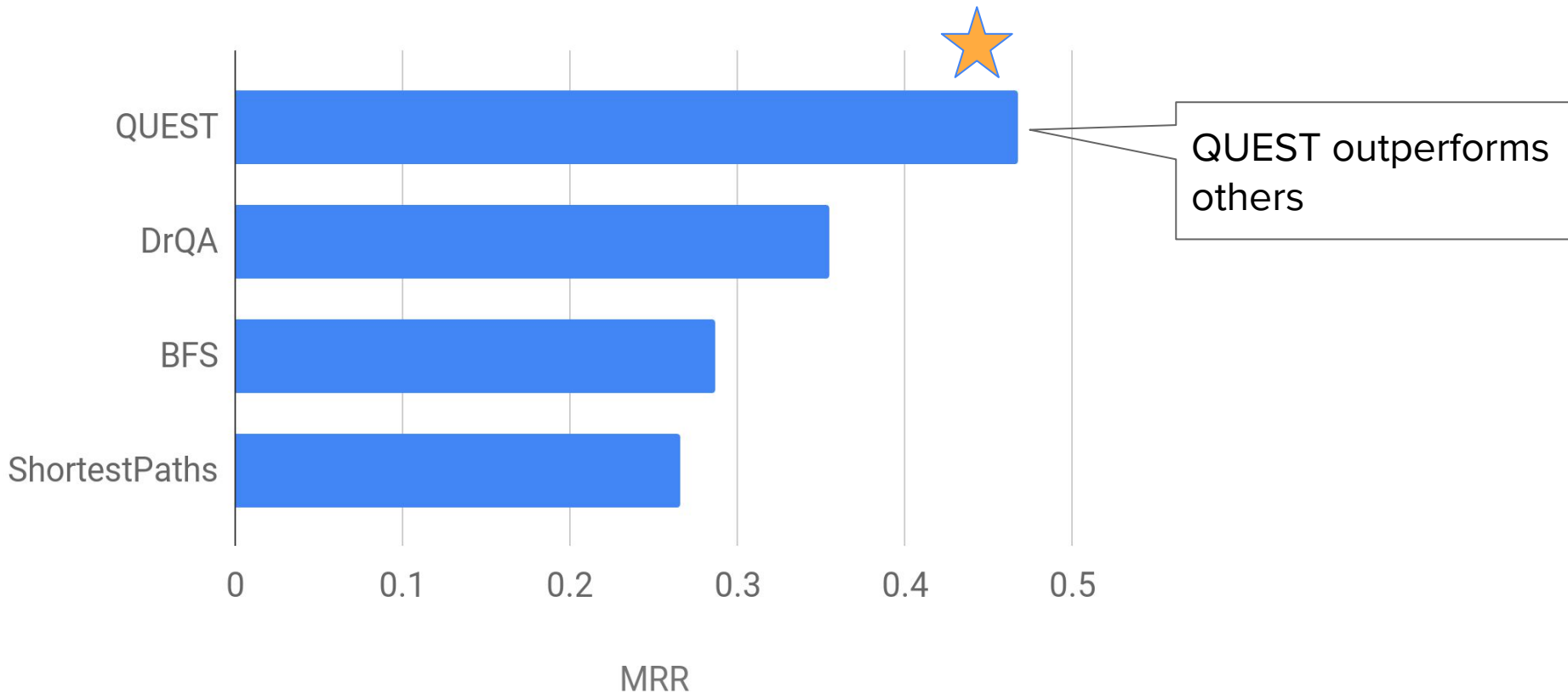
- Mean Reciprocal Rank
- Precision@1
- Hit@5

## ★ Baselines

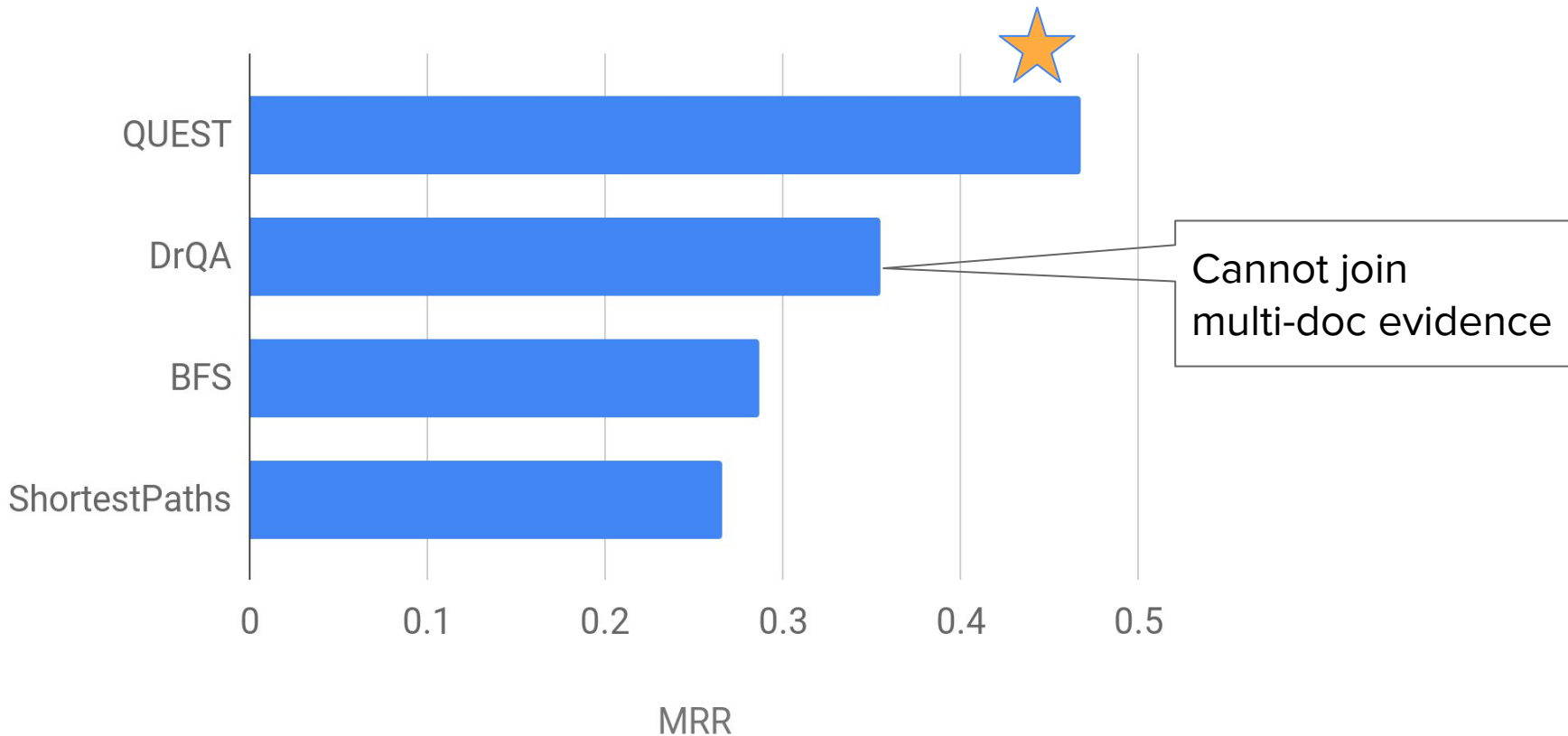
- Neural QA
- Graph methods



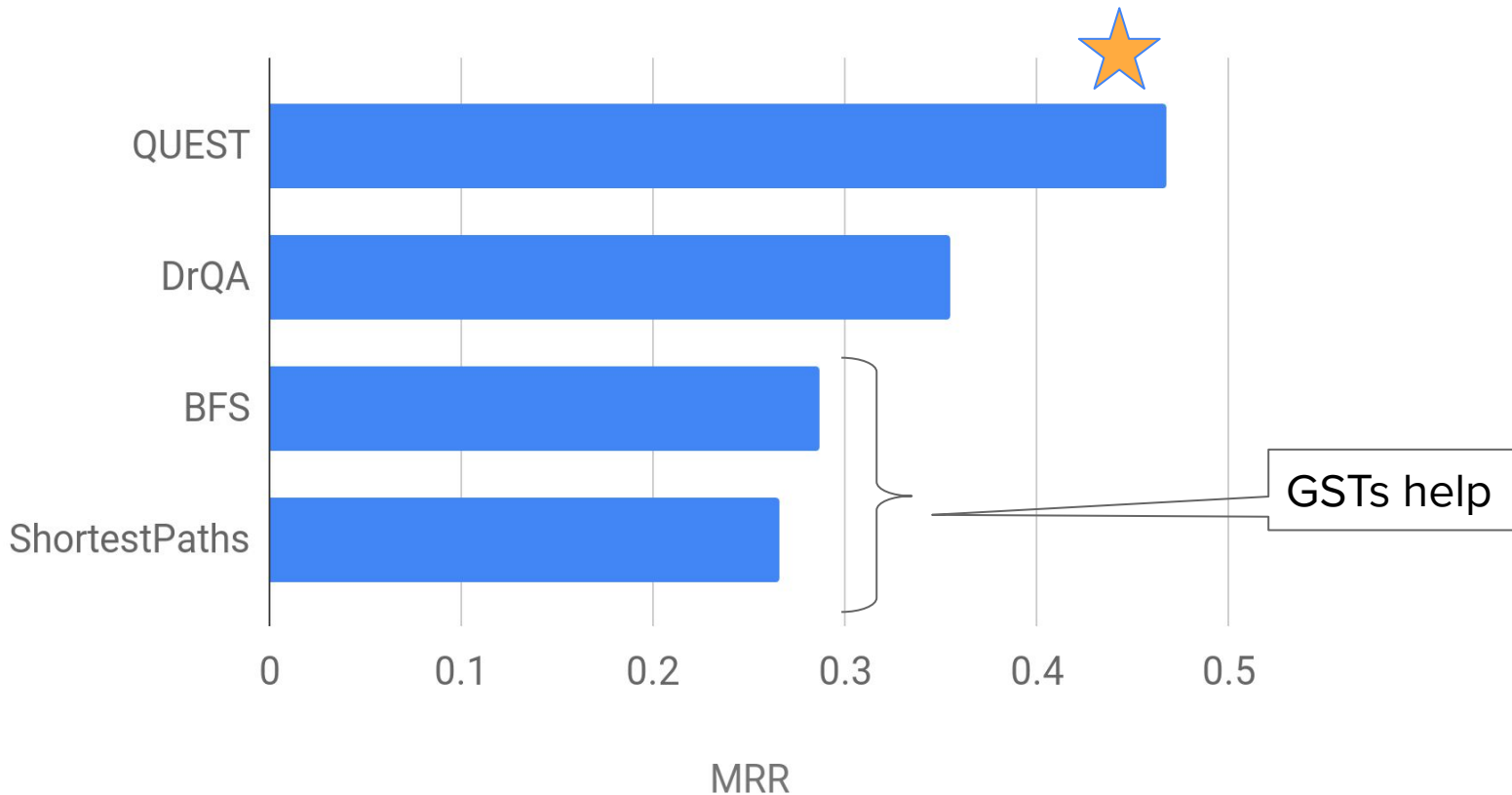
# Experimental Results: Performance



# Experimental Results: Performance



# Experimental Results: Performance



# Experimental Results: Anecdotes

- ★ 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

- ★ **Entity** answers to **complex** questions from **multi-document quasi-KG**
- ★ Design rationale: **Allow noise** and cope later
- ★ **Unsupervised** method using **Group Steiner Trees**
- ★ **Robust** to noise and ungrammatical constructs
- ★ **No decomposition** of question necessary
- ★ Provides **explanatory** evidence

Demo, code and data:

<http://qa.mpi-inf.mpg.de/quest/>

***Thank you!***