SIC Saarland Informatics Campus

UNIVERSITÄT DES SAARLANDES

max planck institut informatik

Question Answering Systems

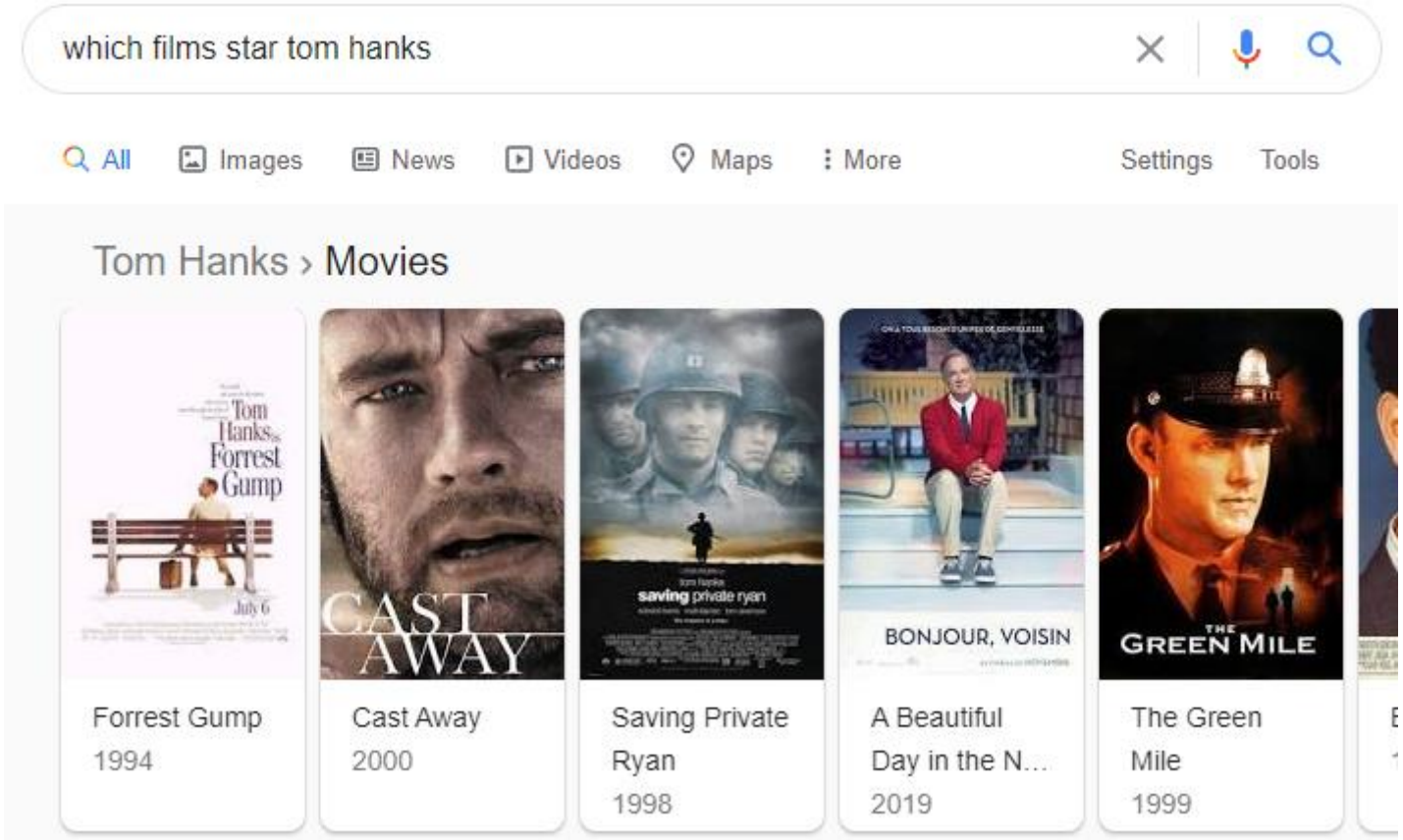
Open KGs: Templates, paraphrases and graphs

Rishiraj Saha Roy
Max Planck Institute for Informatics, Germany

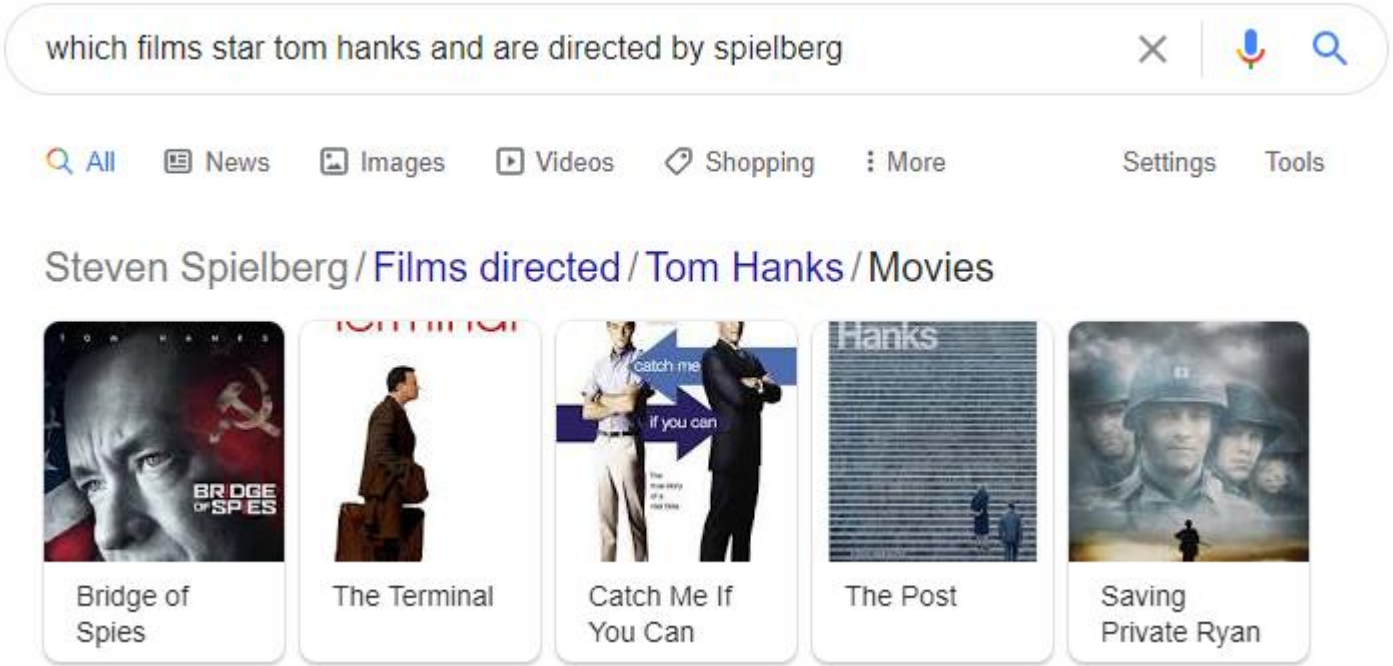
3

Saarland University, Summer Semester 2020

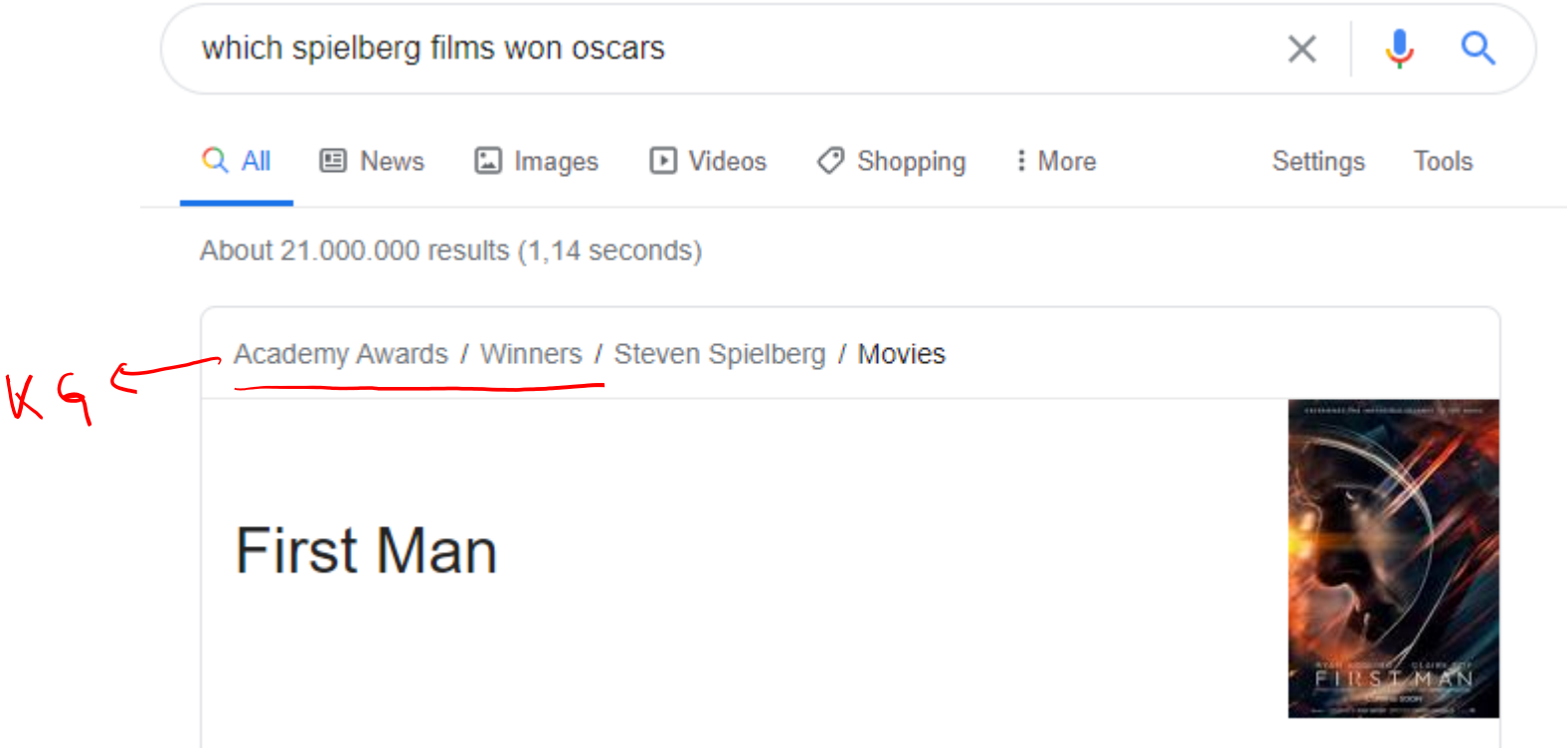
Play with QA


which films star tom hanks
All
Images
News
Videos
Maps
More
Settings
Tools
Tom Hanks › Movies
Forrest Gump
1994
Cast Away
2000
Saving Private Ryan
1998
A Beautiful Day in the N…
2019
The Green Mile
1999

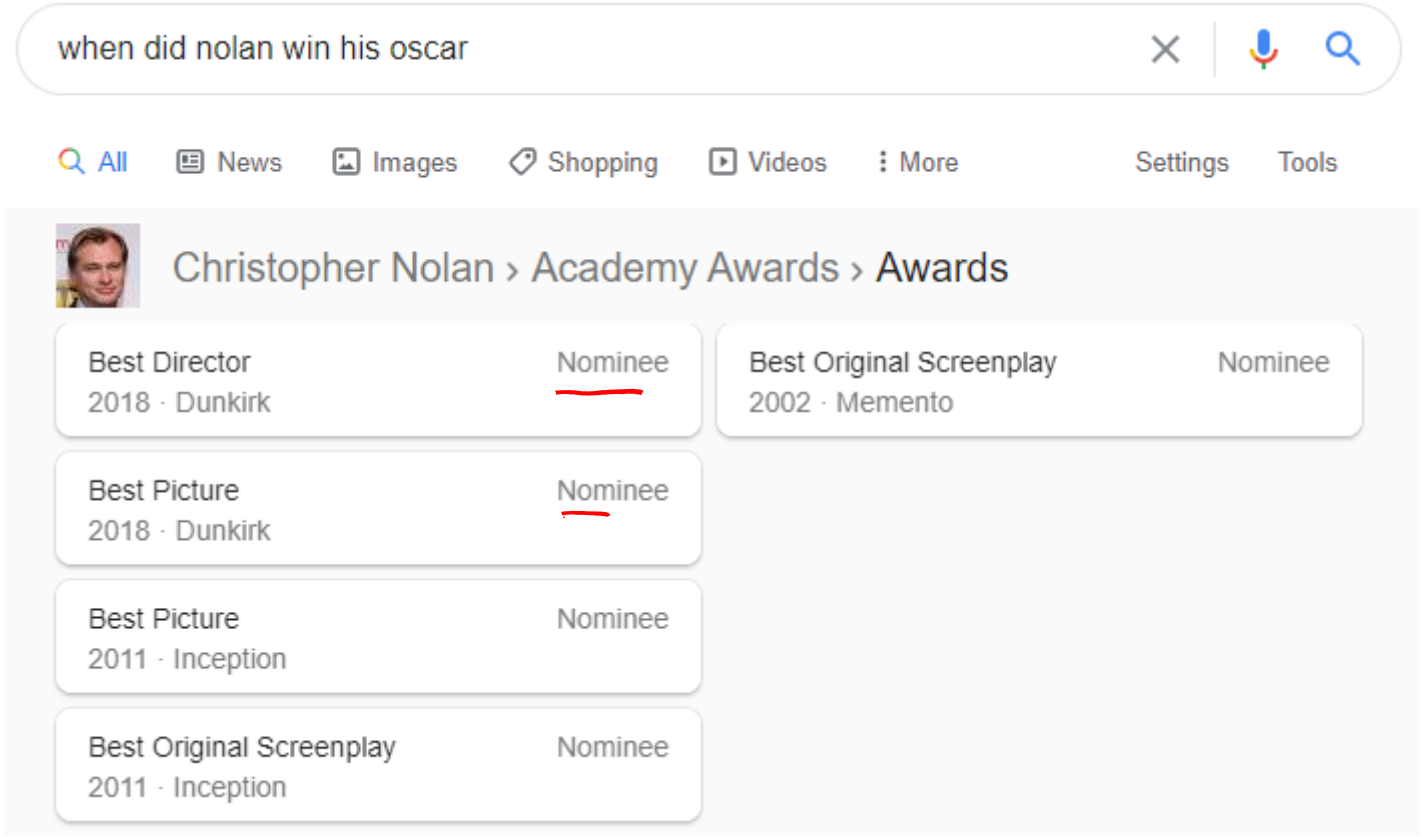
Play with QA


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

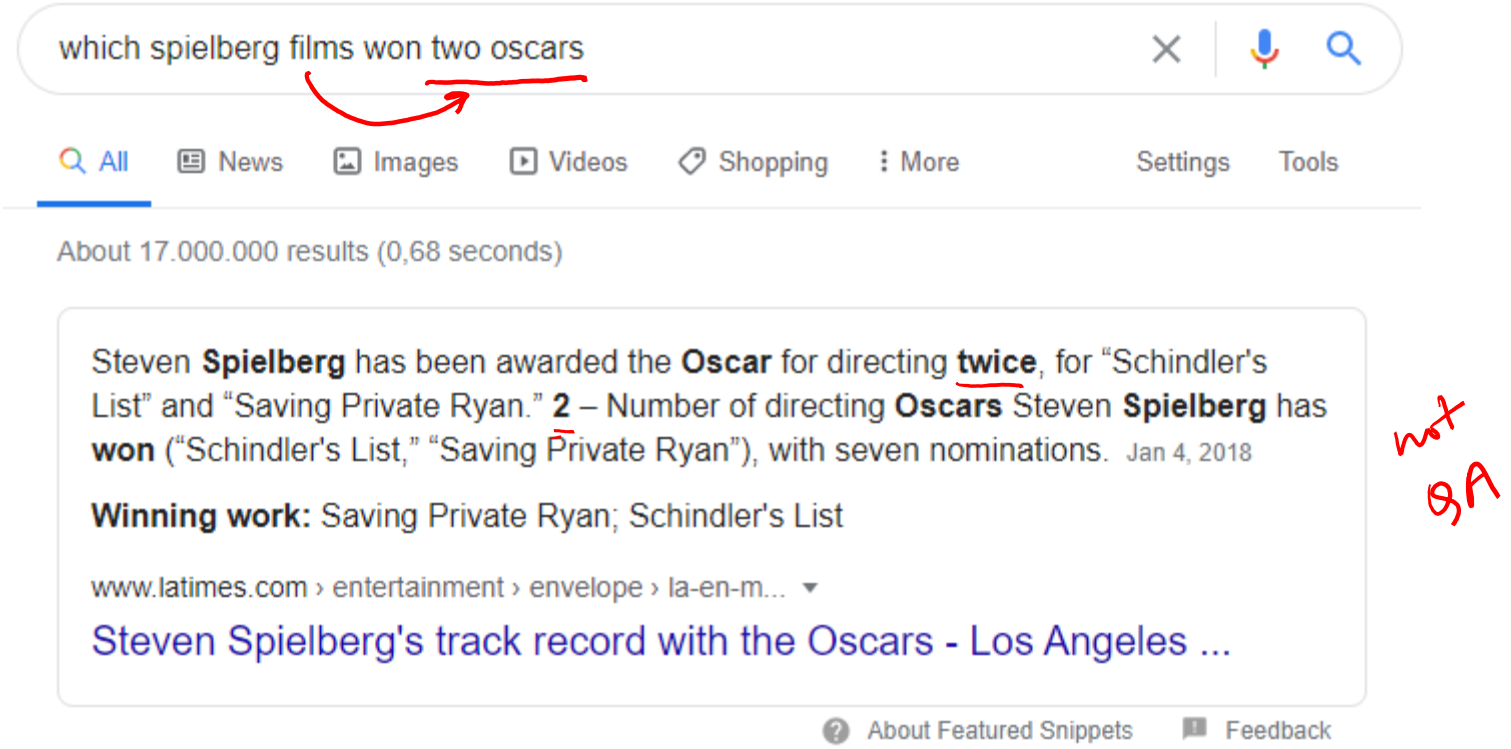
Play with QA


which spielberg films won oscars
All
News
Images
Videos
Shopping
More
Settings
Tools
About 21.000.000 results (1,14 seconds)
KG
Academy Awards / Winners / Steven Spielberg / Movies
First Man
FIRST MAN

Play with QA


when did nolan win his oscar
All
News
Images
Shopping
Videos
More
Settings
Tools
Christopher Nolan › Academy Awards › Awards
Best Director
Nominee
2018 · Dunkirk
Best Original Screenplay
Nominee
2002 · Memento
Best Picture
Nominee
2018 · Dunkirk
Best Picture
Nominee
2011 · Inception
Best Original Screenplay
Nominee
2011 · Inception

Play with QA


which spielberg films won two oscars
All
News
Images
Videos
Shopping
More
Settings
Tools
About 17.000.000 results (0,68 seconds)
Steven Spielberg has been awarded the Oscar for directing twice, for "Schindler's List" and "Saving Private Ryan." 2 – Number of directing Oscars Steven Spielberg has won ("Schindler's List," "Saving Private Ryan"), with seven nominations. Jan 4, 2018
Winning work: Saving Private Ryan; Schindler's List
www.latimes.com › entertainment › envelope › la-en-m...
Steven Spielberg's track record with the Oscars - Los Angeles ...
About Featured Snippets
Feedback

Explore variations!!

- Types of questions : comparatives, aggregations, temporal, ...
- Web search and conversational assistants
- Google, Cortana, Alexa, Siri → Amazon device
- What changes across languages? English, German
- What changes across persons? Individualized
- Domains? Complex questions? → Soccer, Music, ...
- From text? KG? ... / ... ent pred ent

18 April bold + passage / sentence

Research prototype: QAnswer

- https://qanswer-frontend.univ-st-etienne.fr/

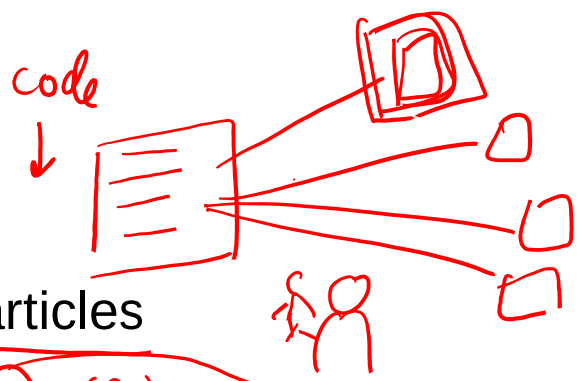
reliable
parallel
fast
quality ✓

Lecture overview

- Writing a good review
- Using templates and paraphrases
- Graph-based answering for complex questions

The publication pipeline

- Do new research ☺
- Write an article
- Include references to useful articles
- Communicate to publisher (conference/journal)
- Article is peer-reviewed (typically three)
- Article is accepted or rejected for publication
- If accepted, the references contribute to citation counts of respective papers
- Review process is critical to advancement of scientific community!!



Writing a good review: Dos and donts

- Write in a structured manner (review form)
- Summary: Unbiased
  - Problem/motivation
  - Method
  - Evaluation
  - Do not copy from abstract – paraphrase!
  - Not we… but "The authors …"
  - Show understanding of article
- Positives and negatives
  - Concise
  - Reduce redundancy
  - Pinpoint to section, quote numbers
  - Illustrate with examples
  - Attention to detail
  - Passive, 3rd person, polite
  - Some but not too many first persons
  - *Constructive*

Positives and negatives

- Focus areas
  1. Motivation
  2. Method
  3. Evaluation
- Stay objective *(I hate templates!)*
- Try to position paper properly
- Look at the bigger picture: main advantages and disadvantages
- No nitpicking
- Do not find flaws in (or praise) future work!
- ~~Do no~~t point out too many grammar and spelling issues
- Cl~~arity, p~~resentation, reproducibility very important
  - But don't harp on ~~missing~~ details
- *Be diplomatic but take a stand*

Question of the day

How can we answer questions on open knowledge graphs?

You'll find this covered in

- Paraphrase-Driven Learning for Open Question Answering
  - Fader et al.
  - ACL 2013
  - https://www.aclweb.org/anthology/P13-1158.pdf
- Answering Complex Questions by Joining Multi-Document Evidence with Quasi Knowledge Graphs
  - Lu et al.
  - SIGIR 2019
  - https://arxiv.org/pdf/1908.00469.pdf

Research paper 1

Paraphrase-Driven Learning for Open Question Answering

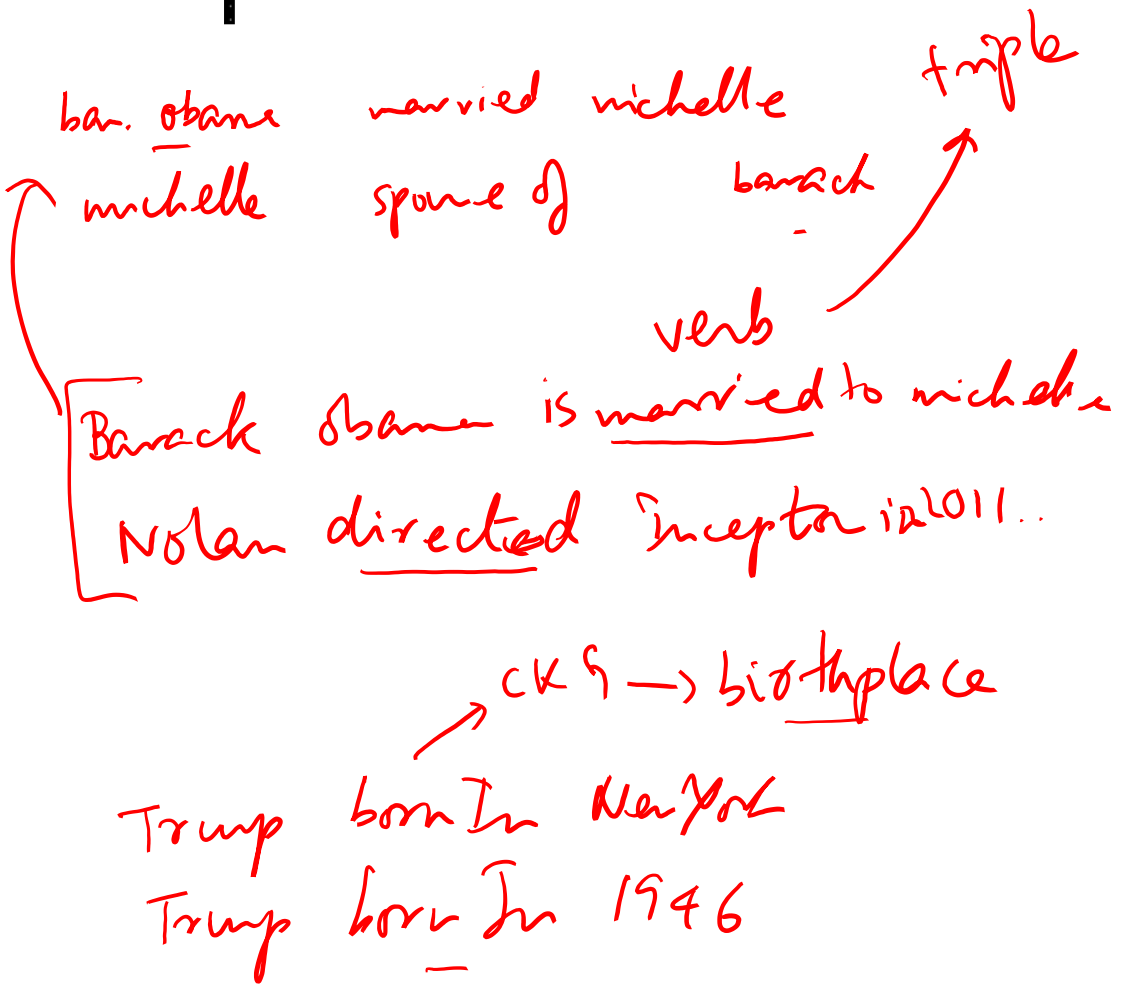
Open question answering over curated and extracted knowledge bases A Fader, L Zettlemoyer, O Etzioni Proceedings of the 20th ACM SIGKDD international conference on Knowledge …	301	2014
Paraphrase-driven learning for open question answering A Fader, L Zettlemoyer, O Etzioni Proceedings of the 51st Annual Meeting of the Association for Computational …	267	2013

Curated Knowledge Graphs

- Wikidata, YAGO, Freebase, Dbpedia
- Predicate and classes moderated
- Crowdsourced, or from Wikipedia infoboxes
- Contains selected but clean information
- Entities, predicates, classes, literals canonicalized
- Moderated schema (predicate signatures, class ontology)
- Intricate construction process

Open Knowledge Graphs

- Open vocabulary knowledge graphs
  - NELL: http://rtw.ml.cmu.edu/rtw/
- Automatically extracted from text (fast)
  - Reverb, Stanford OpenIE
  - IE course?
- Higher coverage
- No canonicalization, noisy
- No moderated schema (bornIn city/date)



The PARALEX System

- Leverages paraphrases from WikiAnswers
- Runs on OKGs
- ~~End-to-end~~ system
- Factoid questions
- Open-domain QA
- Uses distant supervision
  - No direct manual annotation!
- Single ~~relation~~ queries

Who wrote the Winnie the Pooh books? Who is the author of winnie the pooh? What was the name of the authur of winnie the pooh? Who wrote the series of books for Winnie the poo? Who wrote the children's storybook 'Winnie the Pooh'? Who is poohs creator?	C1
What relieves a hangover? What is the best cure for a hangover? The best way to recover from a hangover? Best remedy for a hangover? What takes away a hangover? How do you lose a hangover? What helps hangover symptoms?	C2
What are social networking sites used for? Why do people use social networking sites worldwide? Advantages of using social network sites? Why do people use social networks a lot? Why do people communicate on social networking sites? What are the pros and cons of social networking sites?	C3
How do you say Santa Claus in Sweden? Say santa clause in sweden? How do you say santa clause in swedish? How do they say santa in Sweden? In Sweden what is santa called? Who is sweden santa?	C4

OKG for PARALEX

- ReVerb extractions from ClueWeb09
  - http://reverb.cs.washington.edu/
    - https://www.lemurproject.org/clueweb09.php/
- 6B triples
- Subset used: 15M
- 600k relations, 2M entities

PARALEX: Terminology

- Familiar terms!
- Tuple store
- Patterns, templates, lexical structures
- Lexicons → dictionary
- Derivations
- Lexical equivalences (synonyms)
- Executable queries ~ logical form / SPARQL
- Database concepts → KG items (E, P)

RDF triple store
triple, n-tuple, 3-tuple, 4-tuple, ..., quadruple
KG / DB

tuple
authored(milne, winnie-the-pooh)
treat(bloody-mary, hangover-symptoms)
≡
milne authored winnie-the-pooh
bloody treats hangover
r(?, e) ≡ ?x r e
?x p o
r(e, ?) s p ?x
P ≐ R

PARALEX: Overview

- Problem
  - Learn a question-query mapping function
- Model
  1. Lexicon
  2. Linear ranking function
- Learning
  - Lexicon induction
  - Parameter learning
- Evaluation

Entry Type	NL Pattern	DB Concept
Entity	*nyc*	`new-york`
Relation	*big*	`population`
Question (1-Arg.)	*how big is* `e`	`population(?, e)`
Question (2-Arg.)	*how* `r` *is* `e`	`r(?, e)`

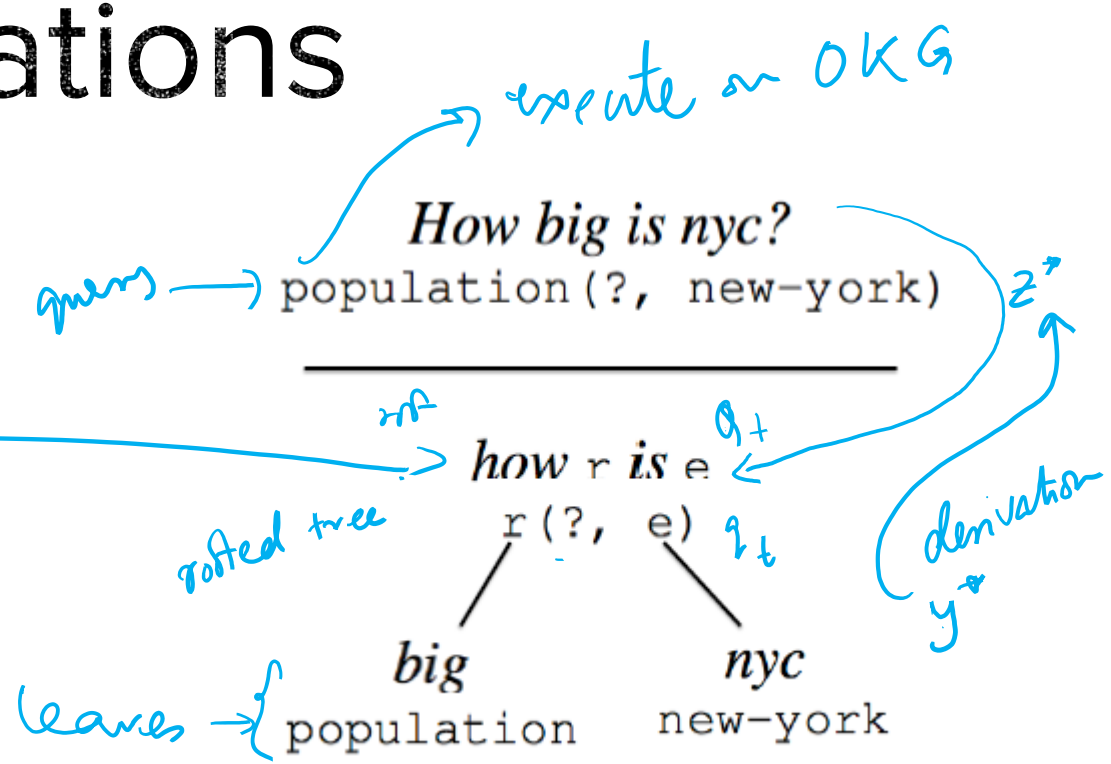
PARALEX: QA model

(1) ▪ Lexicon and derivations

(2) ▪ Linear ranking function

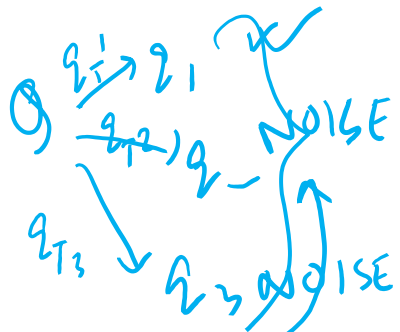
Lexicon and derivations

- Entity patterns
- Relation patterns
- Question patterns
- Derivation in two steps
  1. Match question to template
  2. Map surface forms to DB entries



Linear ranking function

- Multiple queries from single question
- Lots of noise
- Score queries with linear function
- Create feature vectors for question query pairs ϕ
- Learn parameter vector (weights) θ
- Prune by N-best if necessary



$\theta \cdot \phi(x, y)$

$\vec{w} \cdot \vec{z}$

$GEN(x; L)$

$y^*(x) = \arg\max_{y \in GEN(x;L)} \theta \cdot \phi(x, y)$

$z^*(x)$

Ranked list of answers

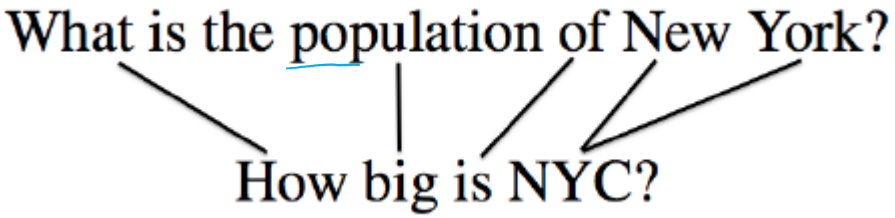
PARALEX: Learning

- Lexical learning
- Parameter learning

Lexical learning

- For generalizing from initial lexicon
- Given:
  - Paraphrase pair
  - Derivation of one question
  - Word alignment
- Produces
  - New lexicon entries

what is the r *of* e = r(?, e)
population = population
new york = new-york



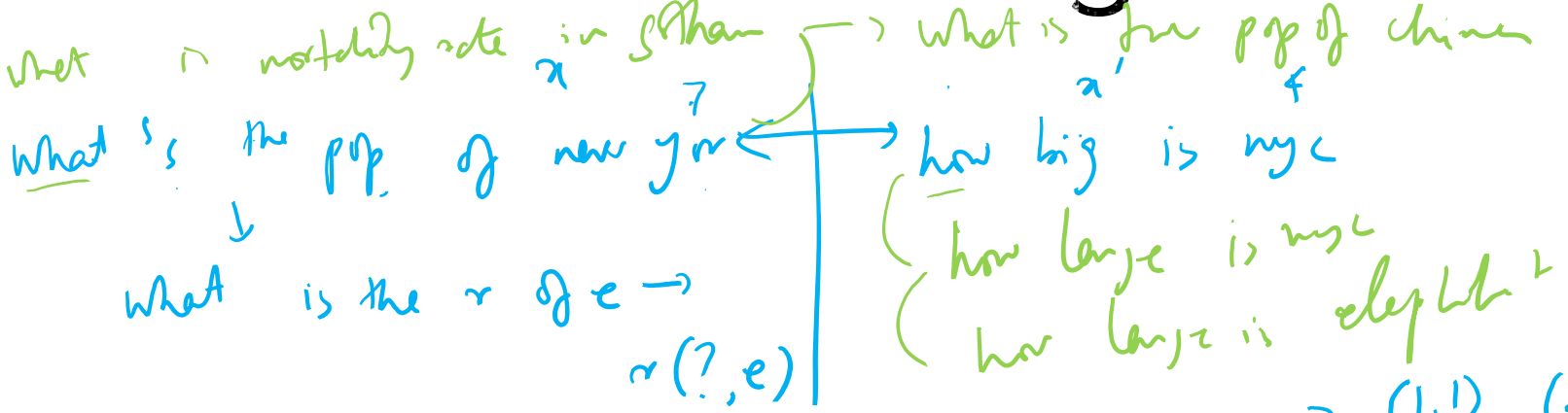
how r *is* e = r(?, e)
big = population
nyc = new-york

Lexical learning: Initial patterns

Question Pattern	Database Query
who r e	r(?, e)
what r e	r(?, e)
who does e r	r(e, ?)
what does e r	r(e, ?)
what is the r *of* e	r(?, e)
who is the r *of* e	r(?, e)
what is r *by* e	r(e, ?)
who is e*'s* r	r(?, e)
what is e*'s* r	r(?, e)
who is r *by* e	r(e, ?)
when did e r	r-in(e, ?)
when did e r	r-on(e, ?)
when was e r	r-in(e, ?)
when was e r	r-on(e, ?)
where was e r	r-in(e, ?)
where did e r	r-in(e, ?)

+ e/r entries

Lexical learning: Walkthrough



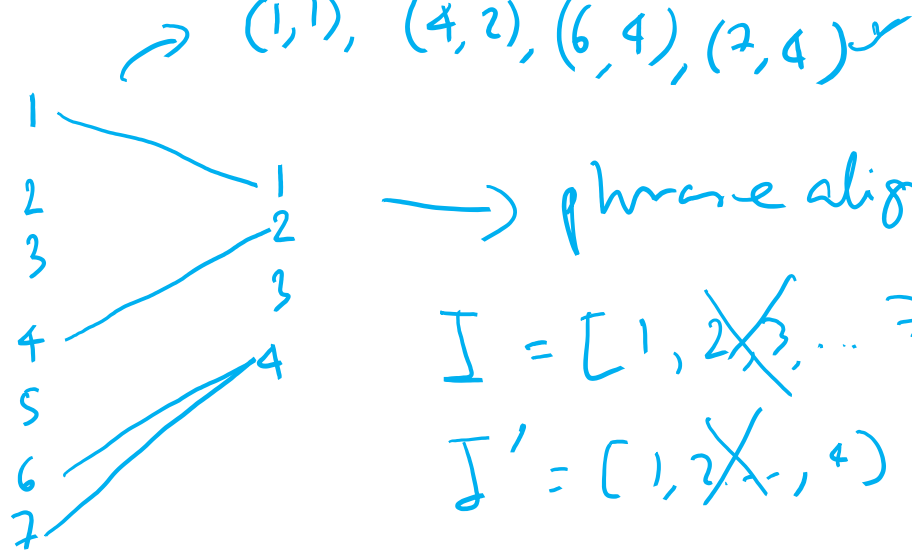
population → pop (KG)

new york → new-york (KG)

word $\underline{A} \in [n] \times [n']$

$n \to 7$

$n' \to 4$



→ phrase alignment

$I = [1, 2, 3, \ldots 7) = [1, 4, 6, 7]$

$I' = [1, 2, \ldots, 4) = [1, 2, 4]$

Lexicon growth

Returns all triples

1. p'_q, p'_r, p'_e are a partition of the words in x'.
2. The phrase pairs $(p_q, p'_q), (p_r, p'_r), (p_e, p'_e)$ are consistent with the word alignment A.
3. The p'_r and p'_e are contiguous spans of words in x'.

function LEARNLEXICON

Inputs:

- A corpus $\mathcal{C}$ of paraphrases (x, x'). (Table 1)
- An initial lexicon L_0 of (pattern, concept) pairs.
- A word alignment function $\text{WordAlign}(x, x')$. (Section 6)
- Initial parameters θ_0.
- A function $\text{GEN}(x; L)$ that derives queries from a question x using lexicon L. (Section 4)
- A function $\text{InduceLex}(x, x', y, A)$ that induces new lexical items from the paraphrases (x, x') using their word alignment A and a derivation y of x. (Section 5.1)

Output: A learned lexicon L.

$L = \{\}$
for all $x, x' \in \mathcal{C}$ **do**
  if $\text{GEN}(x; L_0)$ is not empty **then**
    $A \leftarrow \text{WordAlign}(x, x')$
    $y^* \leftarrow \arg\max_{y \in \text{GEN}(x; L_0)} \theta_0 \cdot \phi(x, y)$
    $L \leftarrow L \cup \text{InduceLex}(x, x', y^*, A)$

return L

Lexicon growth

String	Learned Database Relations for String
get rid of	treatment-for, cause, get-rid-of, cure-for, easiest-way-to-get-rid-of
word	word-for, slang-term-for, definition-of, meaning-of, synonym-of
speak	speak-language-in, language-speak-in, principal-language-of, dialect-of
useful	main-use-of, purpose-of, importance-of, property-of, usefulness-of

String	Learned Database Entities for String
smoking	smoking, tobacco-smoking, cigarette, smoking-cigar, smoke, quit-smoking
radiation	radiation, electromagnetic-radiation, nuclear-radiation
vancouver	vancouver, vancouver-city, vancouver-island, vancouver-british-columbia
protein	protein, protein-synthesis, plasma-protein, monomer, dna

Parameter learning

- Filter out noise
- Create feature vectors
- Perceptron algorithm

  https://en.wikipedia.org/wiki/Perceptron#Learning_algorithm
- Get top-*k* queries and execute!
- Leads to answer ranking!

function LEARNPARAMETERS

Inputs:

- A corpus $\mathcal{C}$ of paraphrases (x, x'). (Table 1)
- An initial lexicon L_0 of (pattern, db concept) pairs.
- A learned lexicon L of (pattern, db concept) pairs.
- Initial parameters θ_0.
- Number of perceptron epochs T.
- Number of training-data shards K.
- A function $\text{GEN}(x; L)$ that derives queries from a question x using lexicon L. (Section 4)
- A function $\text{PerceptronEpoch}(\mathcal{T}, \theta, L)$ that runs a single epoch of the hidden-variable structured perceptron algorithm on training set $\mathcal{T}$ with initial parameters θ, returning a new parameter vector θ'. (Section 5.2)

Output: A learned parameter vector θ.

// Step 1: Generate Training Examples $\mathcal{T}$
$\mathcal{T} = \{\}$
for all $x, x' \in \mathcal{C}$ **do**
  if $\text{GEN}(x; L_0)$ is not empty **then**
    $y^* \leftarrow \arg\max_{y \in \text{GEN}(x; L_0)} \theta_0 \cdot \phi(x, y)$
    $z^* \leftarrow$ query of y^*
    Add (x', z^*) to $\mathcal{T}$

// Step 2: Learn Parameters from $\mathcal{T}$
Randomly partition $\mathcal{T}$ into shards $\mathcal{T}_1, \ldots, \mathcal{T}_K$
for $t = 1 \ldots T$ **do**
  // Executed on k processors
  $\theta_{k,t} = \text{PerceptronEpoch}(\mathcal{T}_k, \theta_{t-1}, L)$
  // Average the weights
  $\theta_t = \frac{1}{K} \sum_k \theta_{k,t}$

return θ_T

Research paper 2

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

Question Answering Systems: Complexity?

Which Nolan films won an Oscar?

Question Answering Systems: Complexity?

Which Nolan films won an Oscar?

This came back from a search not QA

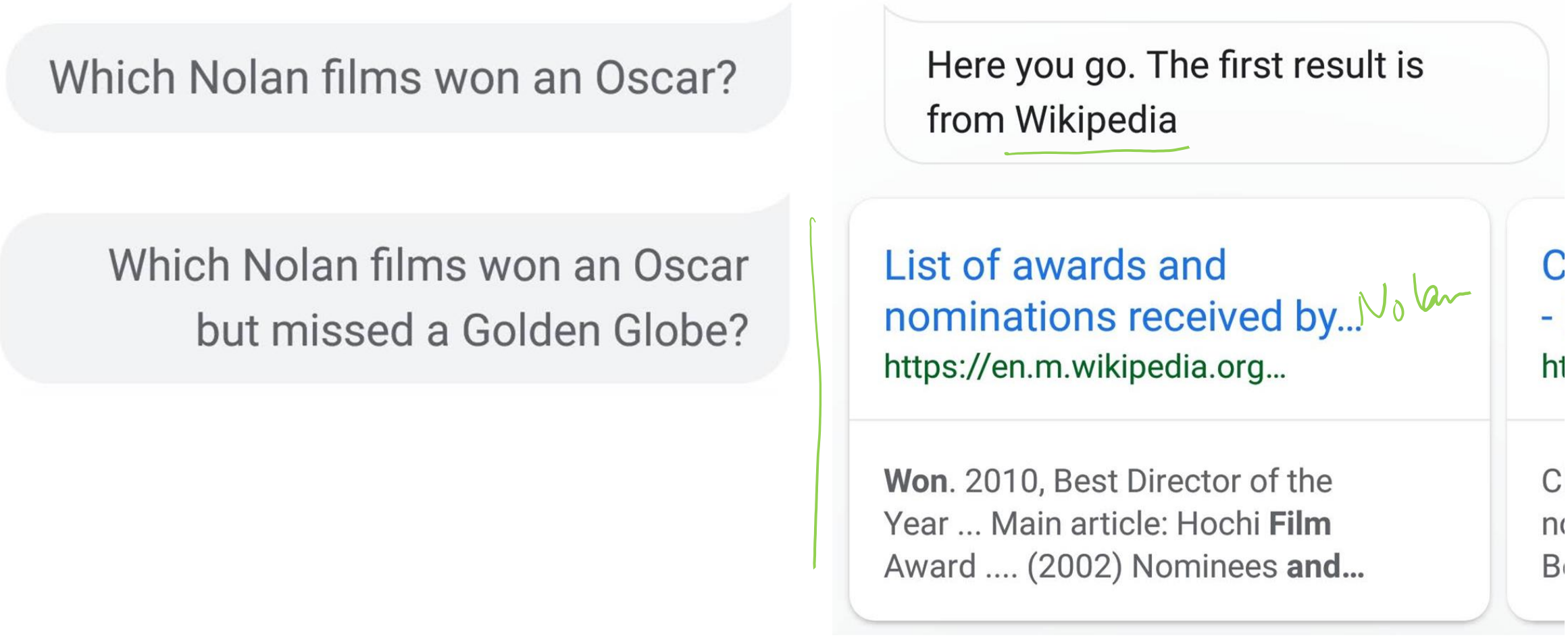
The “**Inception**” miss in particular was puzzling for a best picture nominee that ended up being the year's most awarded film at the Oscars that year. (Nolan did, however, receive screenplay Oscar nominations for “**Memento**” and “**Inception**,” as well as a best picture nomination for the latter.)

Question Answering Systems: Complexity?

Which Nolan films won an Oscar?

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

Question Answering Systems: Complexity?



Challenges in QA systems

- Complexity in information needs
  - Works only for simple questions
  - Misses additional conditions

Complexity in information needs

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

Answers: Inception, Interstellar

- KGs are inherently incomplete
- Answer from text (+ KG)
- Join evidence from multiple documents on-the-fly
- Hardly any training data: develop unsupervised approach

Tackling complexity: Our 5-point agenda

1. Construct noisy quasi-KG on-the-fly
2. Augment with alignments and types
3. Spot question cornerstones
4. Compute Group Steiner Tree
5. Rank answer candidates

Graph construction

Graph algorithm

QUEST: Question Answering with Steiner Trees

Reviving Text-QA: Passage Retrieval

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..

Reviving Text-QA: Passage Retrieval

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...

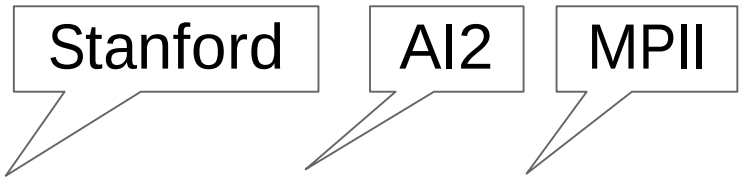
Inception narrowly lost to The Social Network for Best Screenplay at the 68th Golden Globe Awards, which were declared in the afternoon..

THIS ILLUSTRATES "JOINING EVIDENCE"

Semi-structuring raw text

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

- Use Open Information Extraction
- Existing approaches (Stanford Open IE, ReVerb, ClausIE) not suited
- Use own noisy extractor:
  - … *X* … *verb(+preposition)* … *Y* … <Nolan, directed, the movie Inception>
  - … *X* … *noun+preposition* … *Y* … <Inception, winner of, Best Sound>

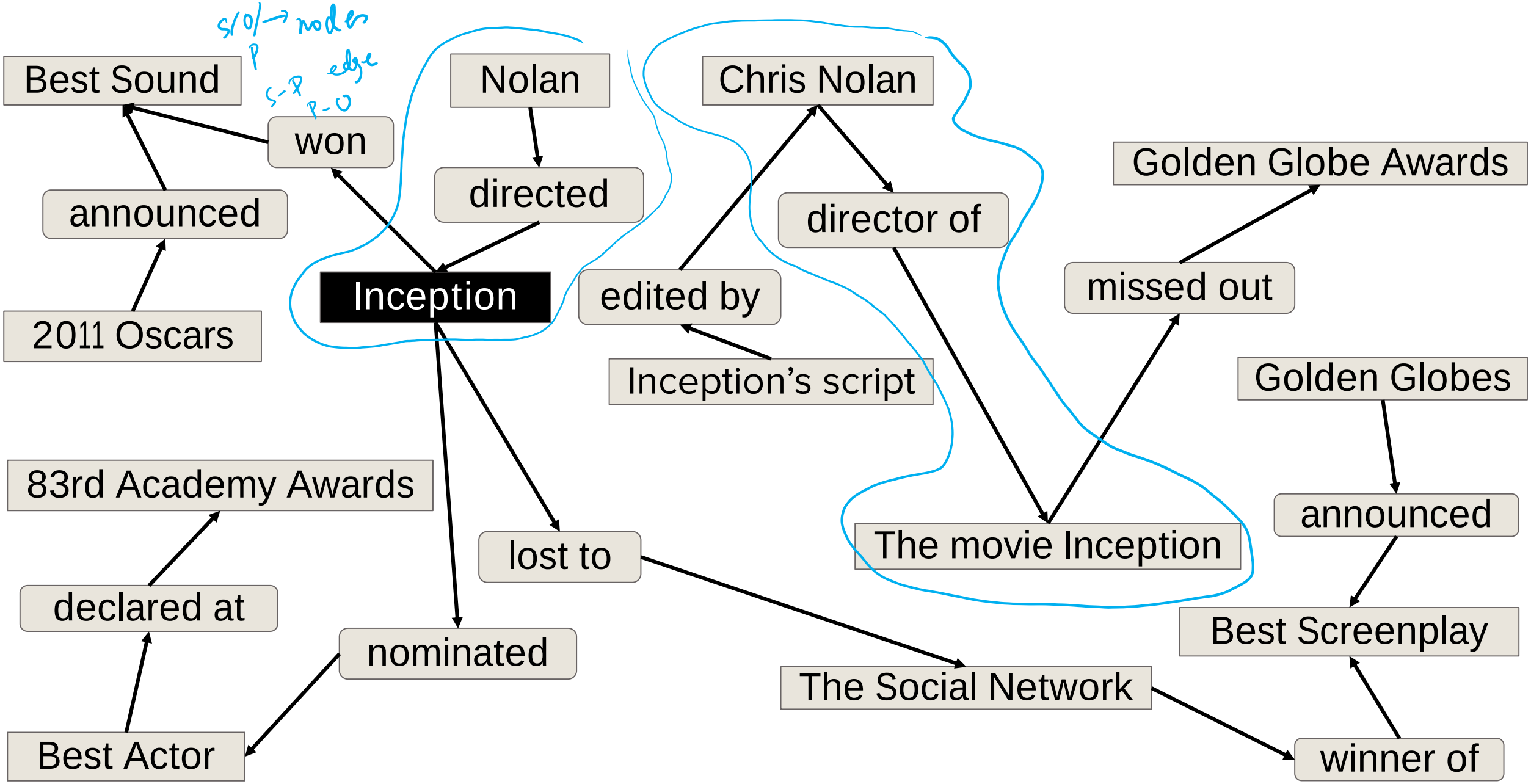


Semi-structuring raw text

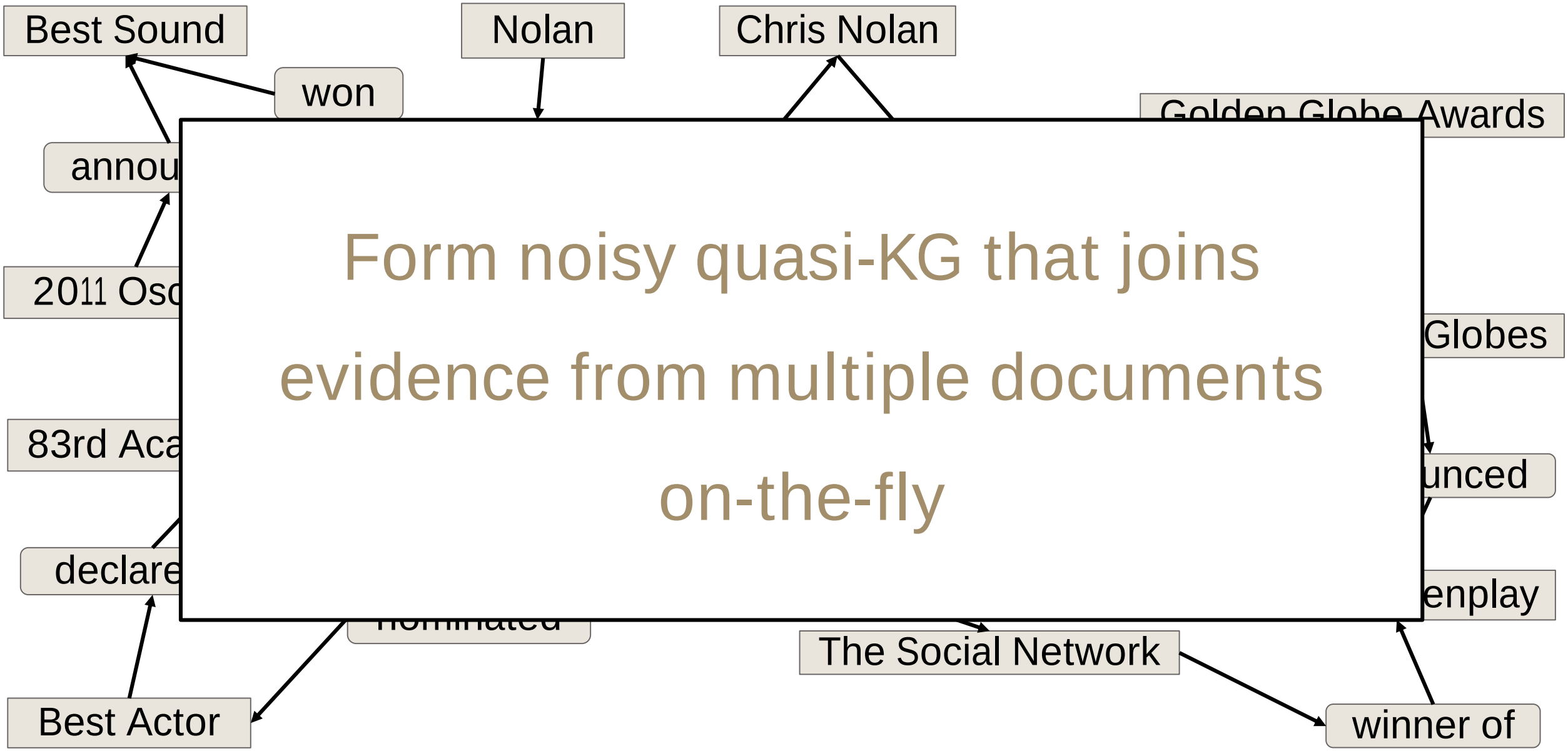
<Nolan, directed, Inception>
<Inception, won, Best Sound>
<2011 Oscars, announced, Best Sound>
<Inception, nominated, Best Actor>
<The movie Inception, missed out, Golden Globe Awards>
<Chris Nolan, director of, The movie Inception>
<Inception's script, edited by, Chris Nolan>
<Inception, lost to, The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

FROM HERE ON, THE APPROACH IS OUTLINED BY EXAMPLE

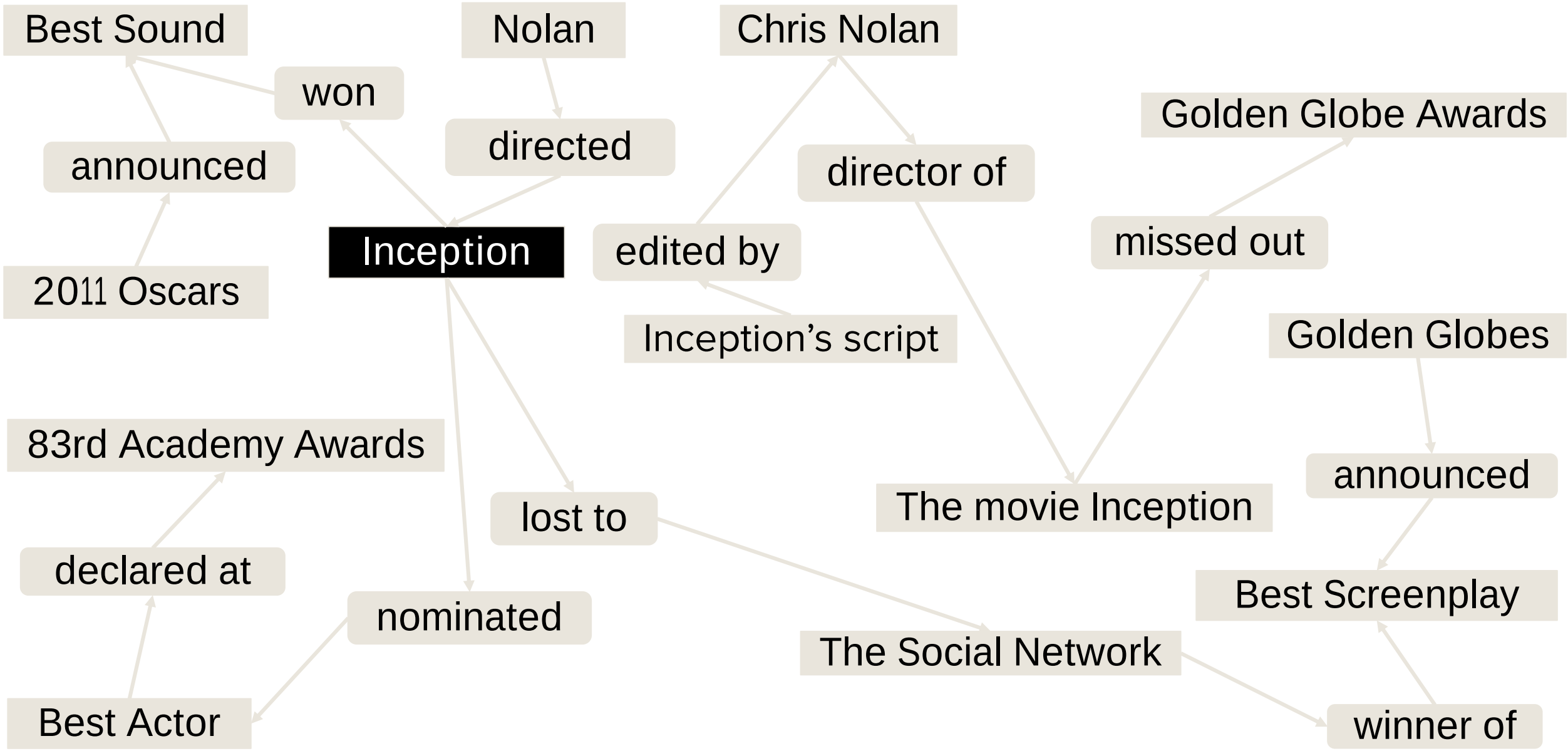
Question: Which Nolan films won an Oscar but missed a Golden Globe?



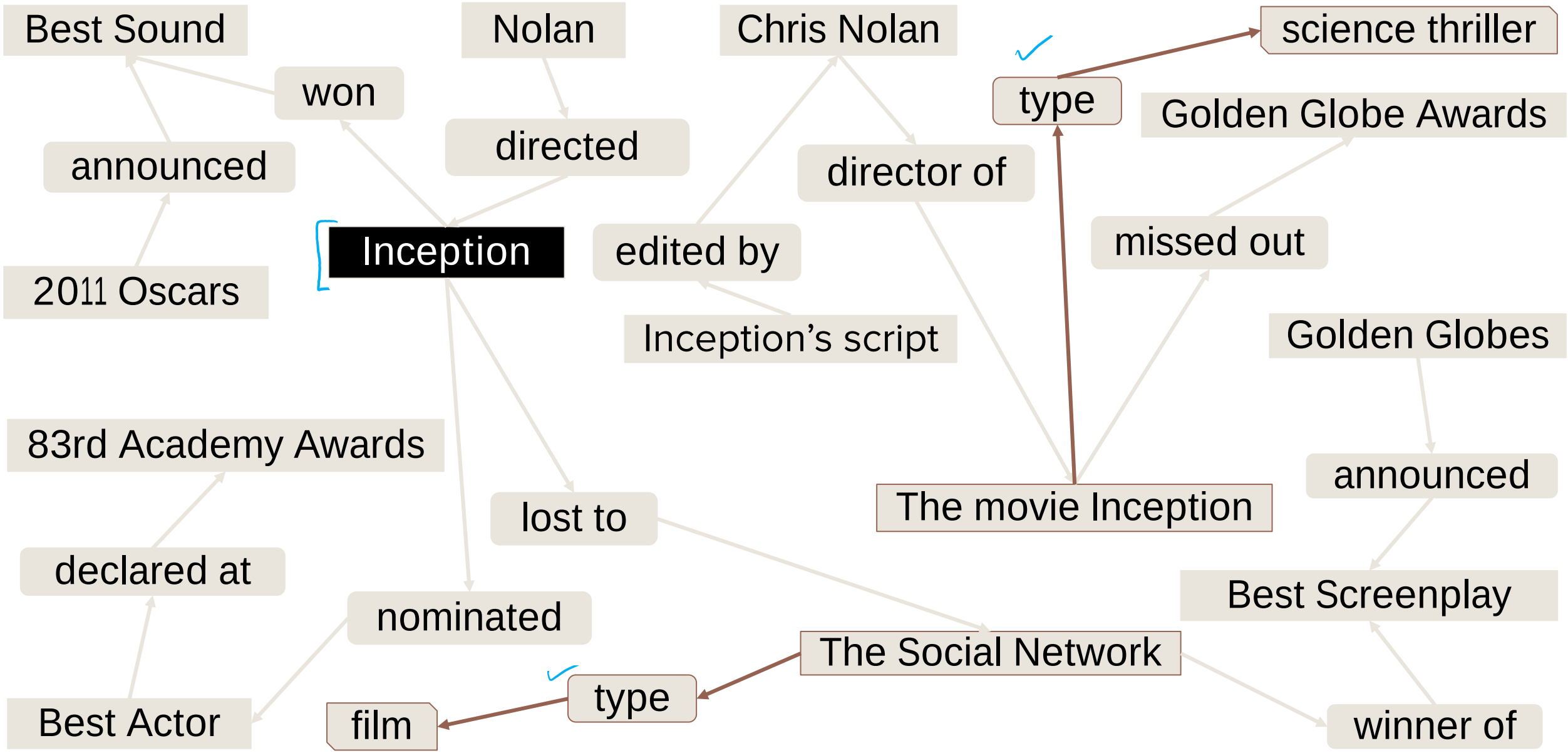
Question: Which Nolan films won an Oscar but missed a Golden Globe?



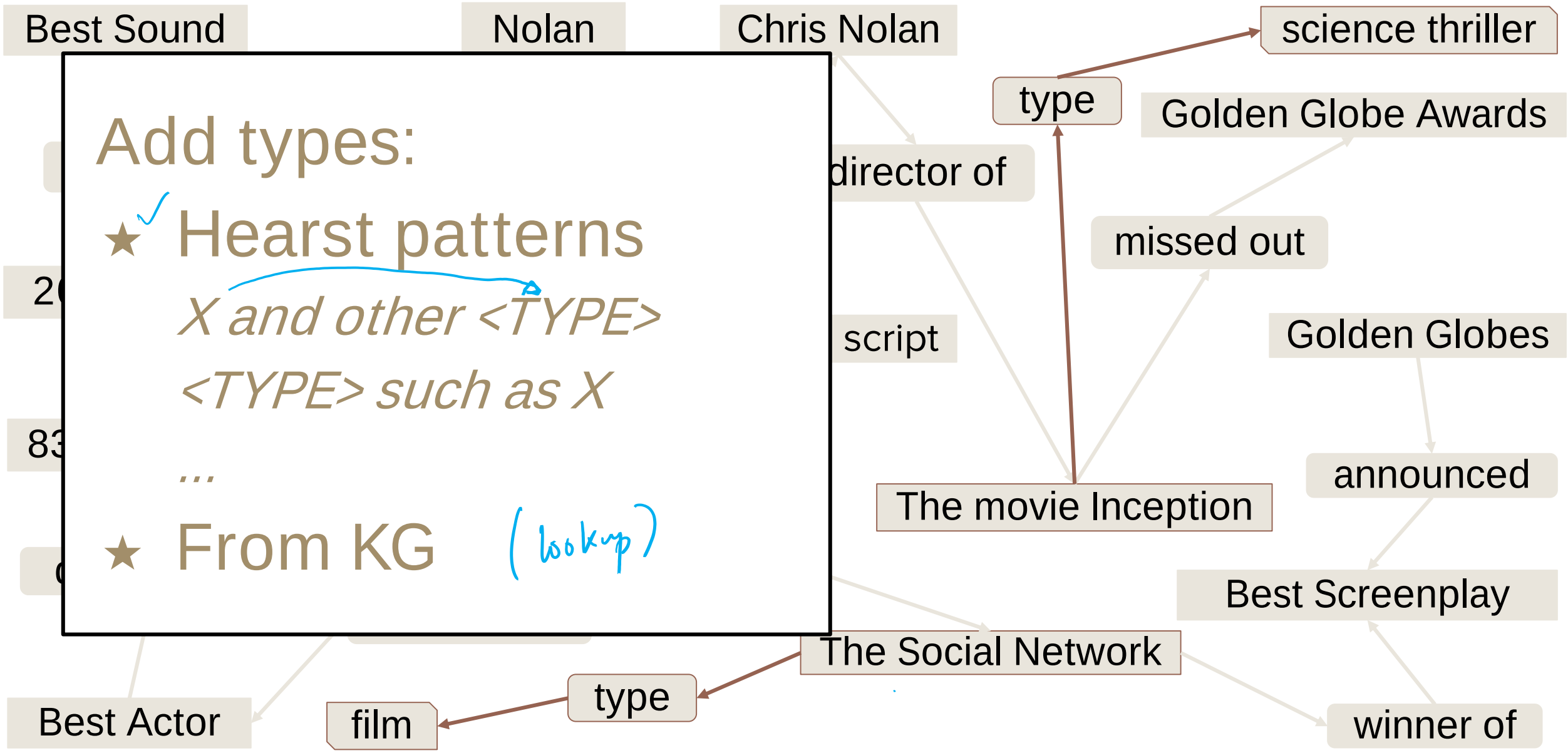
Question: Which Nolan films won an Oscar but missed a Golden Globe?



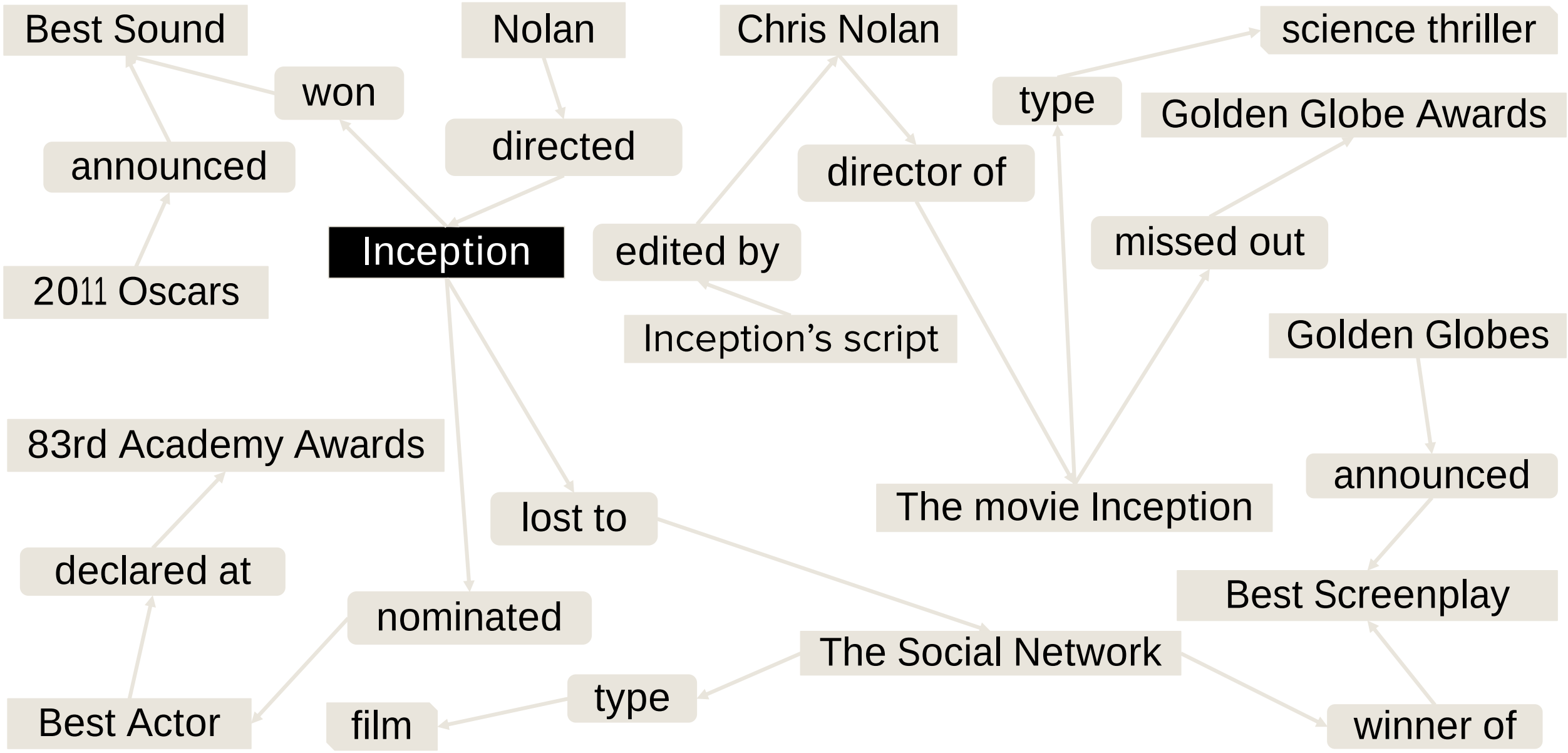
Question: Which Nolan films won an Oscar but missed a Golden Globe?



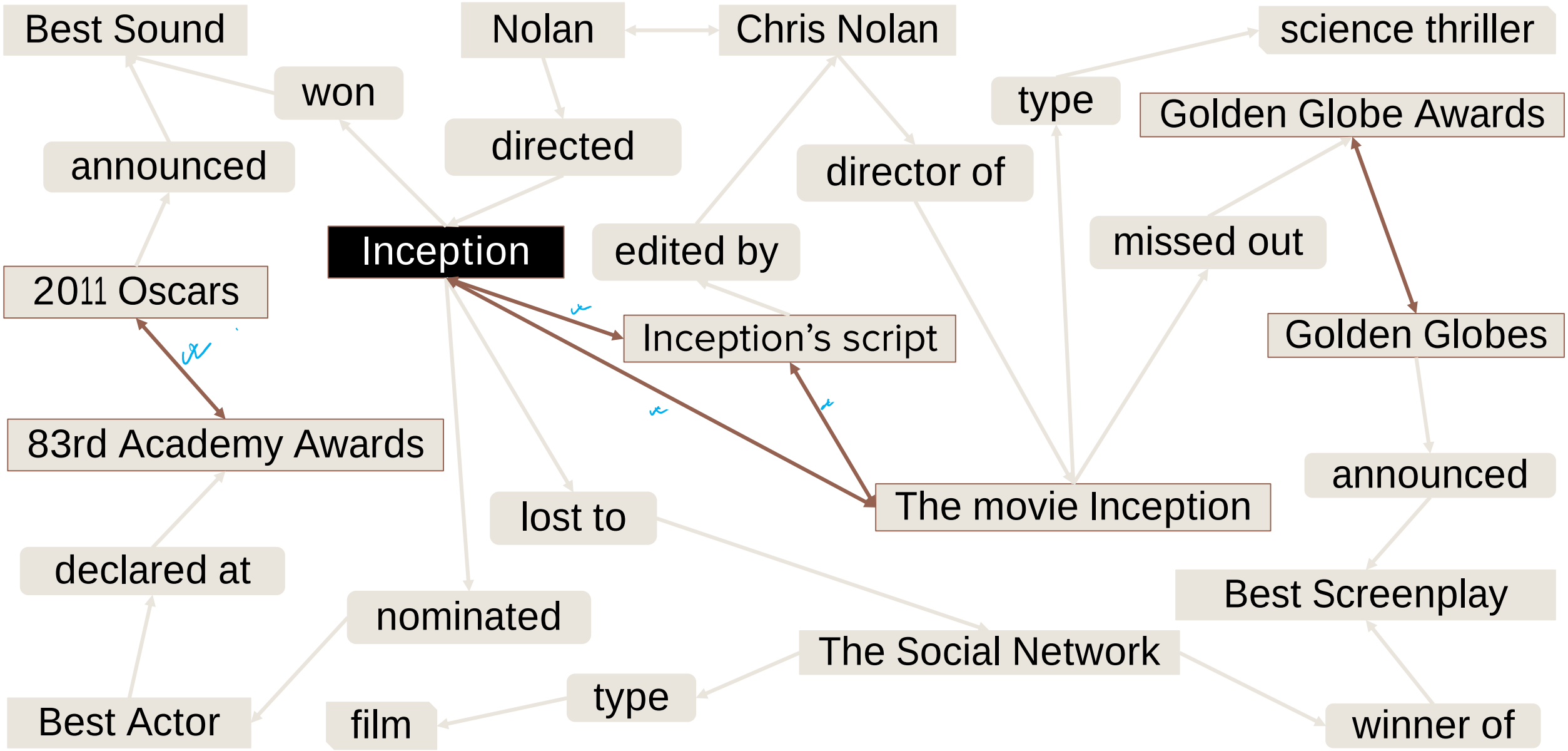
Question: Which Nolan films won an Oscar but missed a Golden Globe?



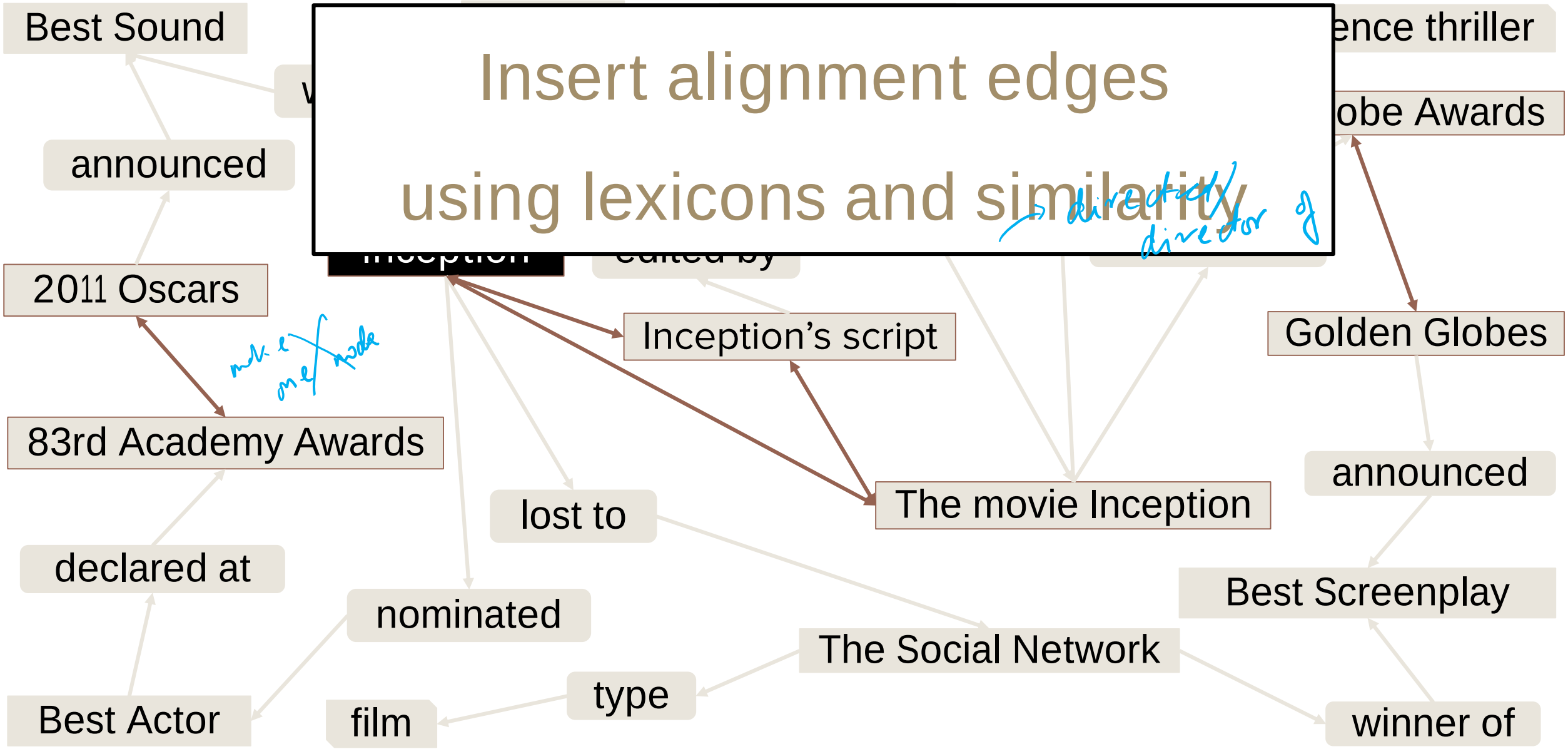
Question: Which Nolan films won an Oscar but missed a Golden Globe?



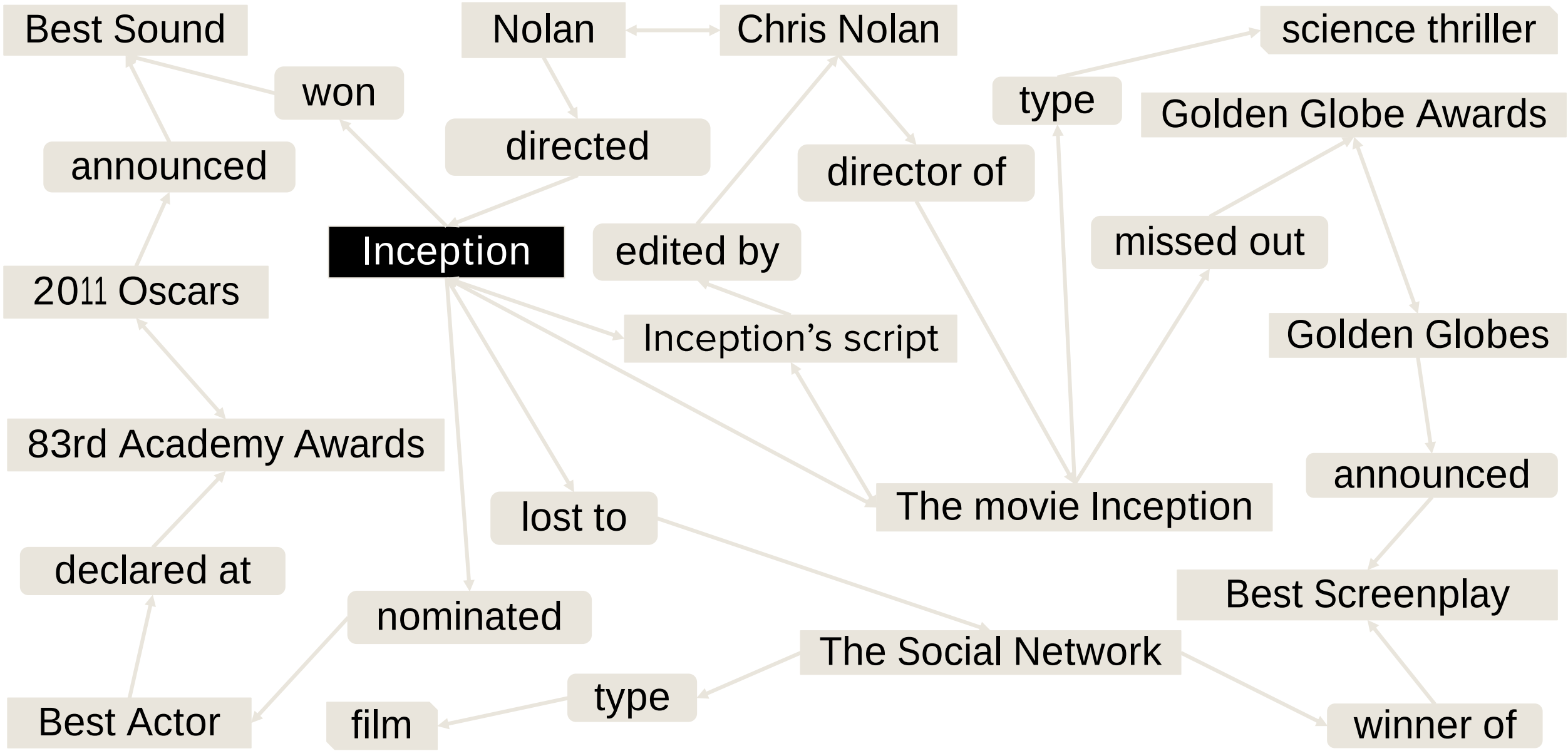
Question: Which Nolan films won an Oscar but missed a Golden Globe?



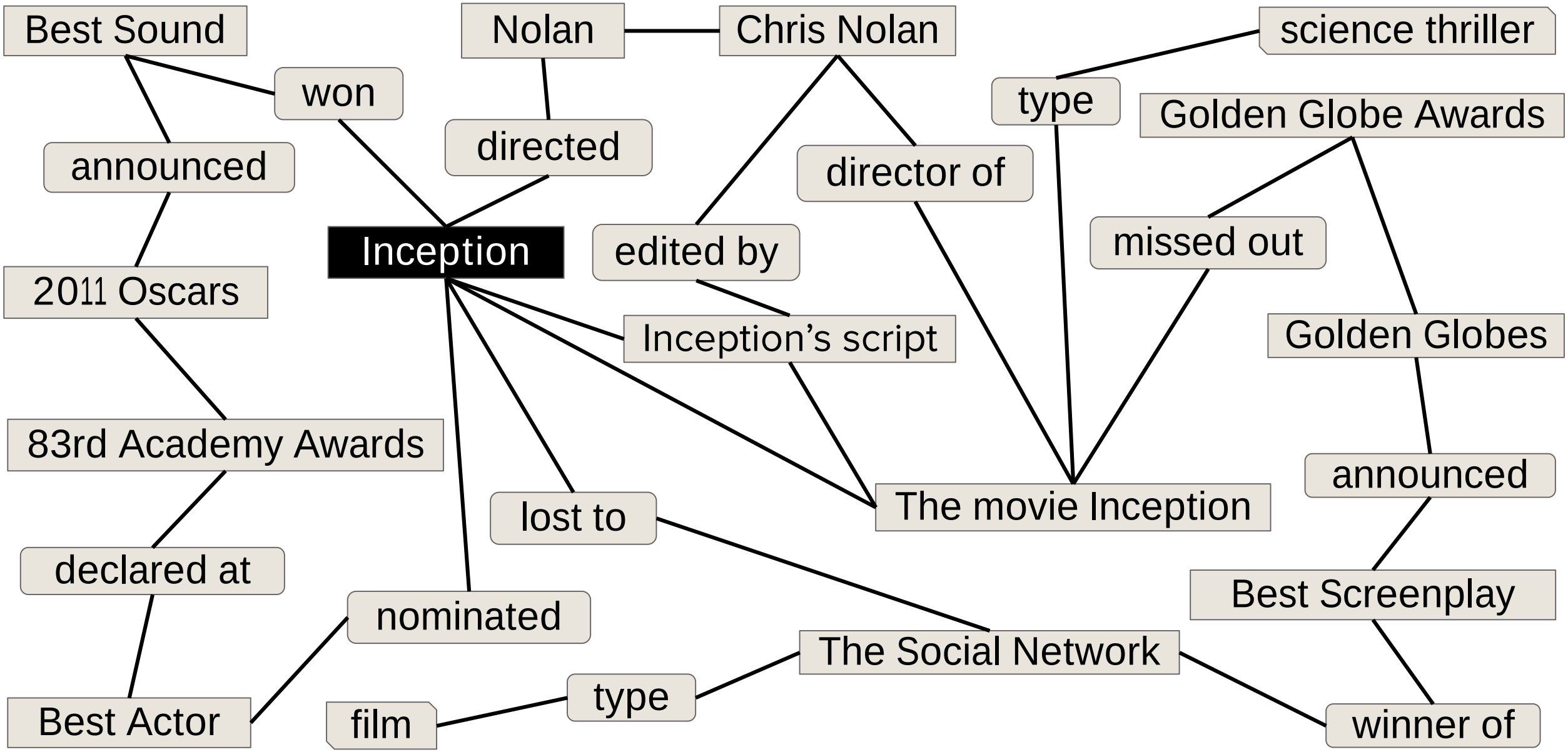
Question: Which Nolan films won an Oscar but missed a Golden Globe?



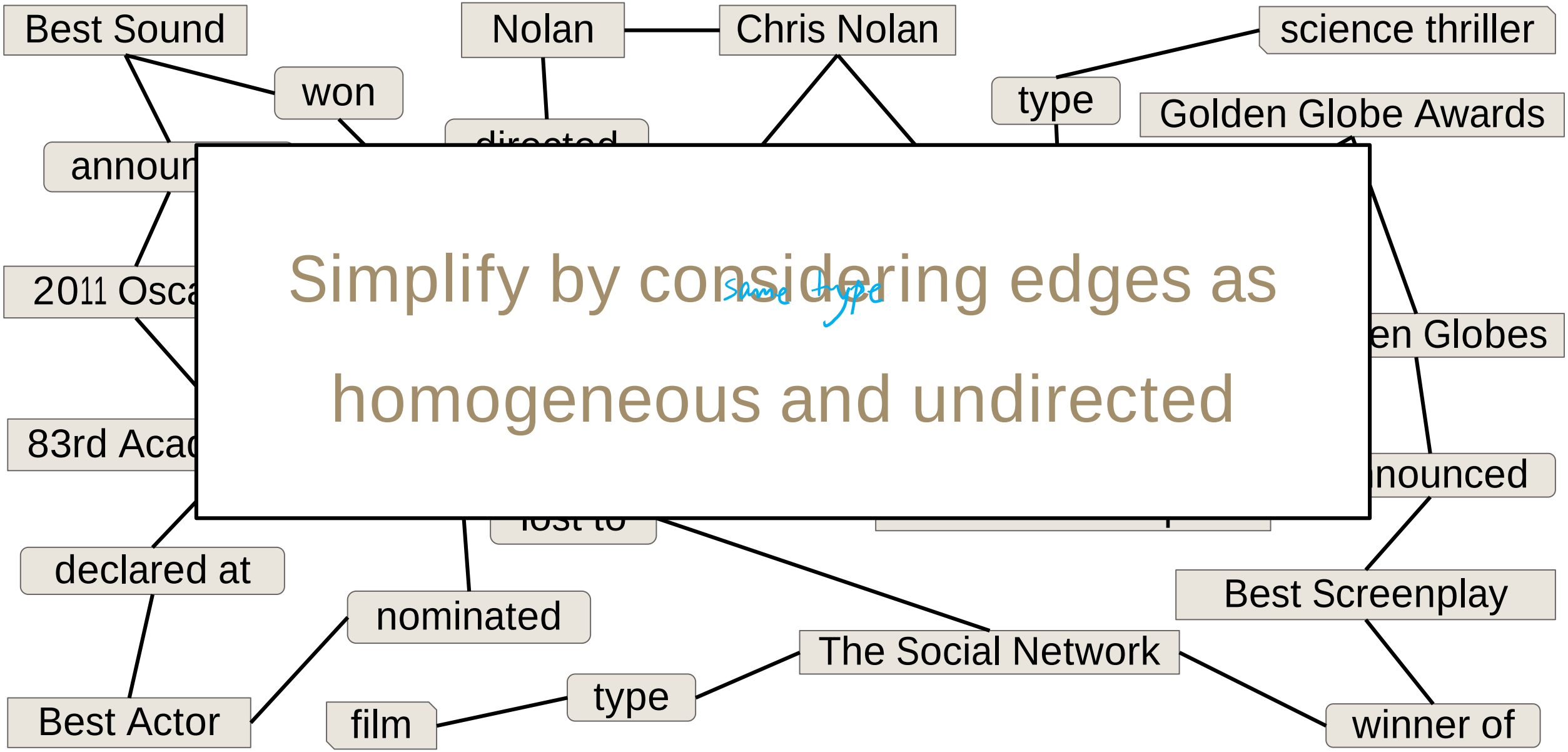
Question: Which Nolan films won an Oscar but missed a Golden Globe?



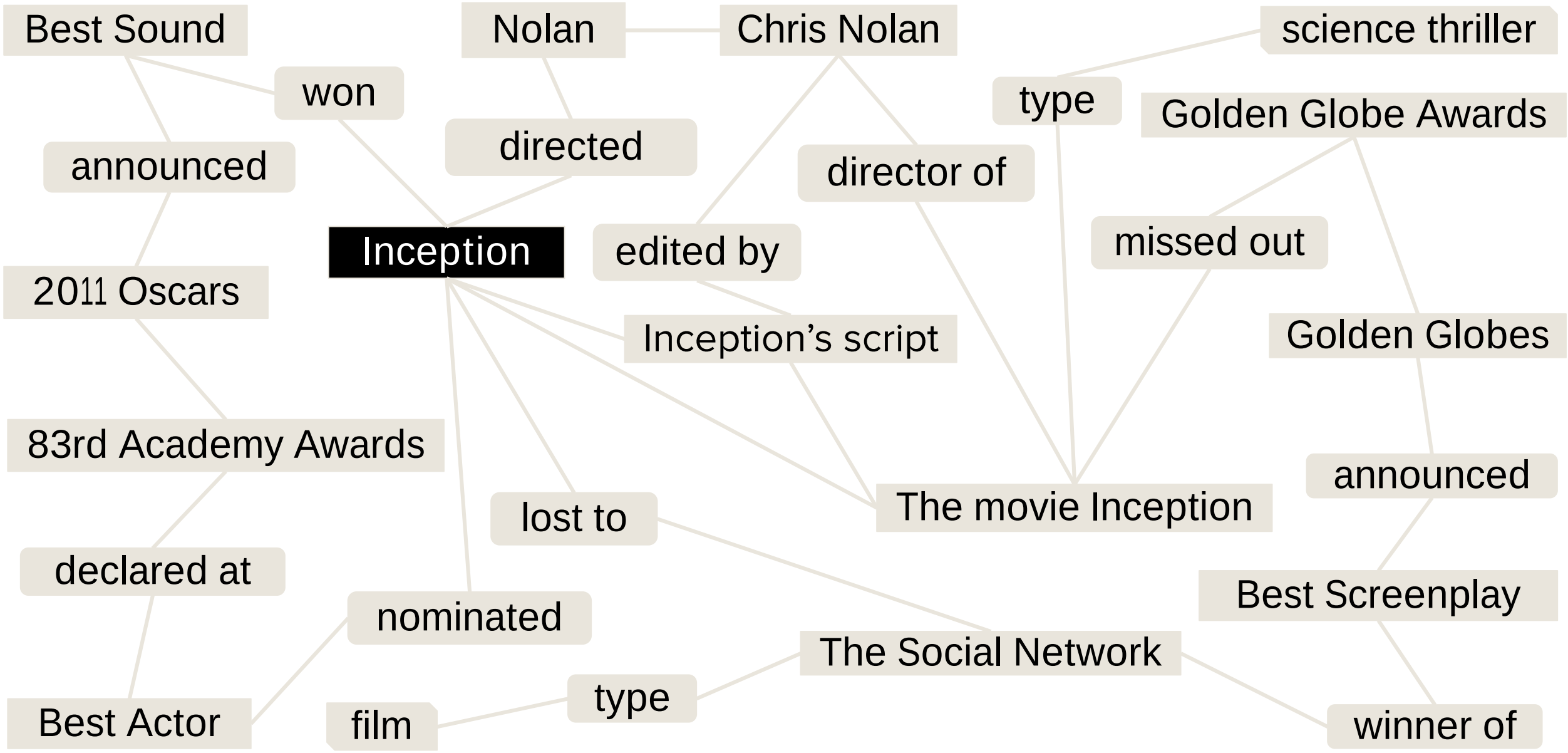
Question: Which Nolan films won an Oscar but missed a Golden Globe?



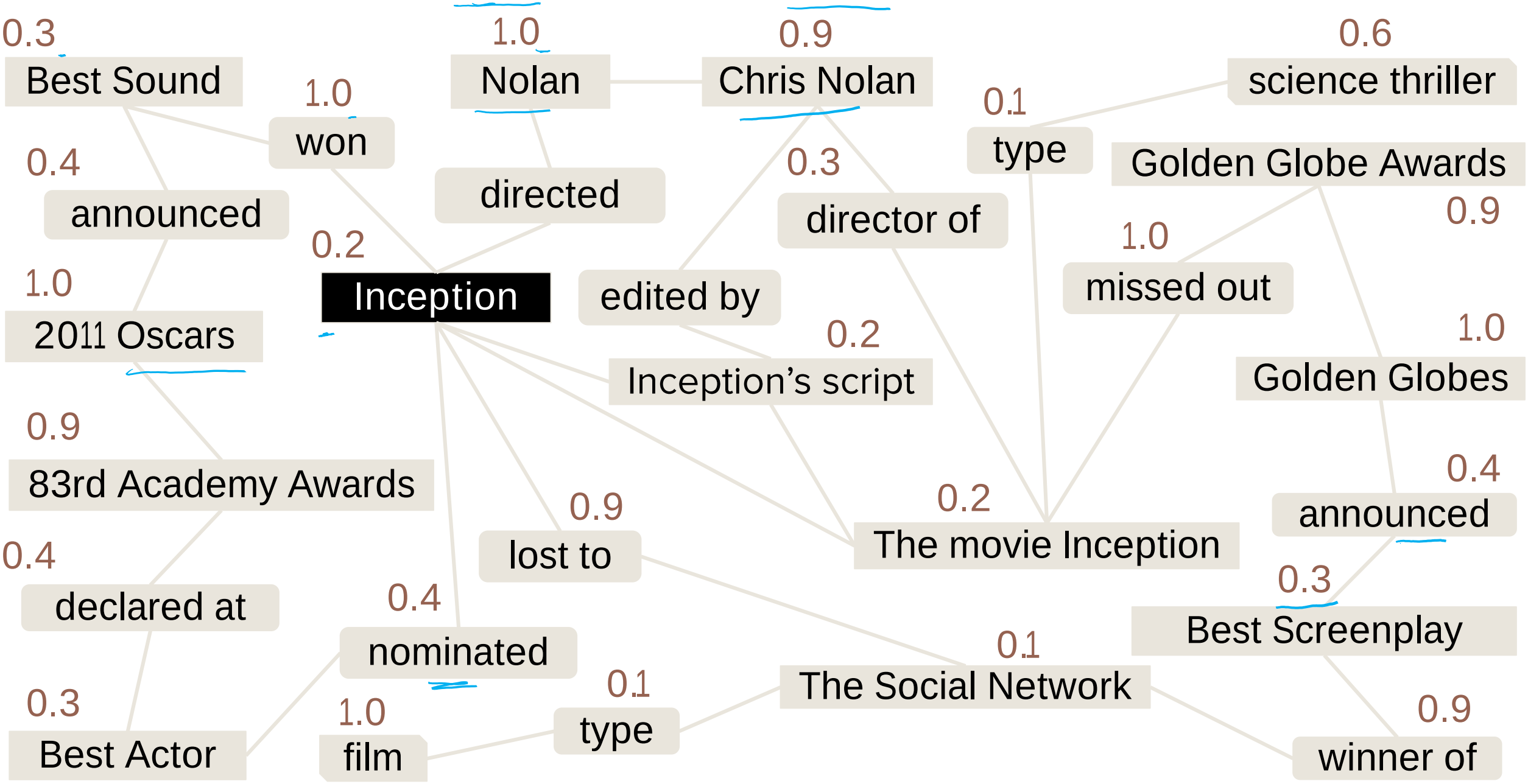
Question: Which Nolan films won an Oscar but missed a Golden Globe?



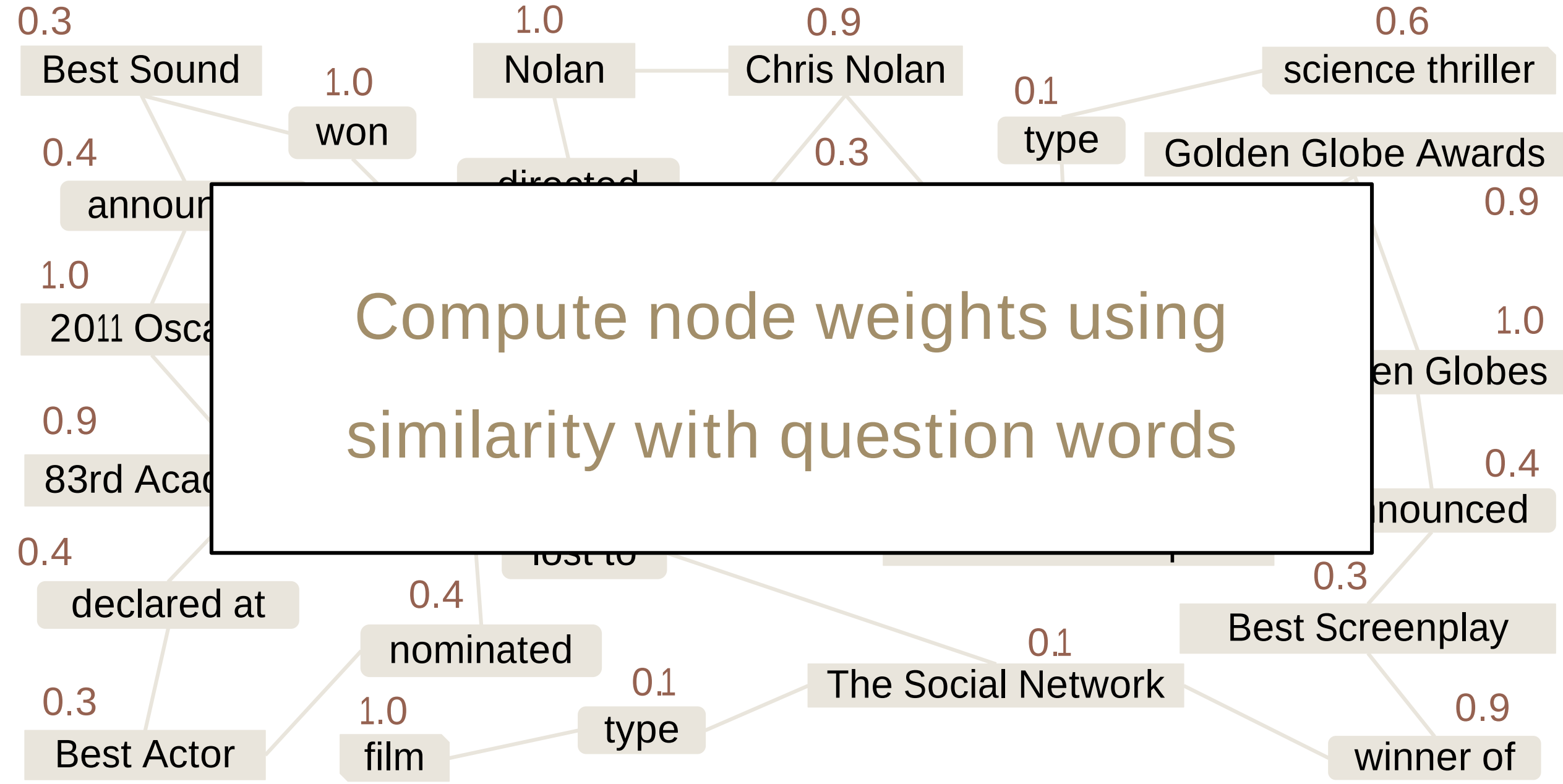
Question: Which Nolan films won an Oscar but missed a Golden Globe?



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

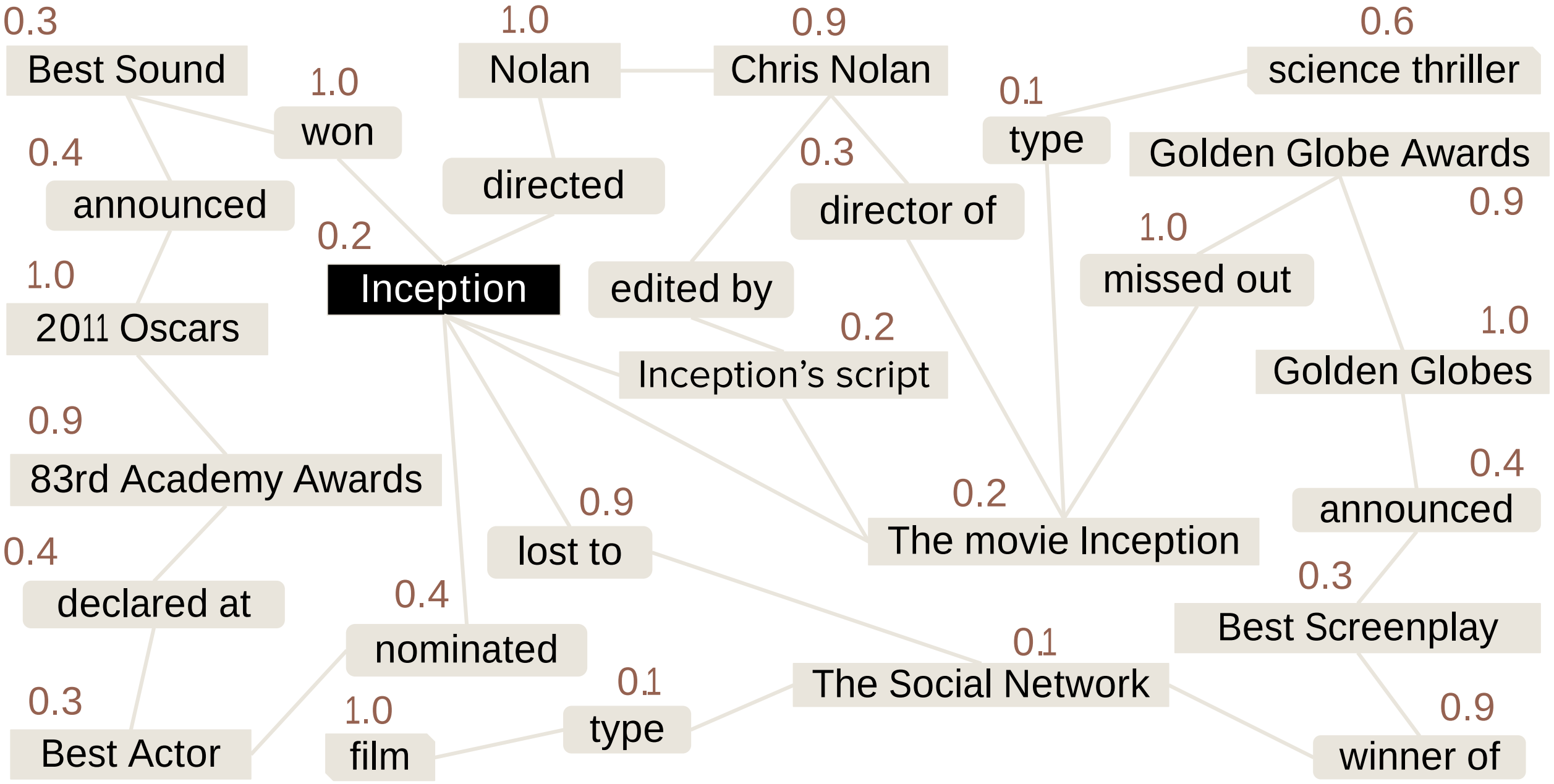


Question: Which Nolan films won an Oscar but missed a Golden Globe?

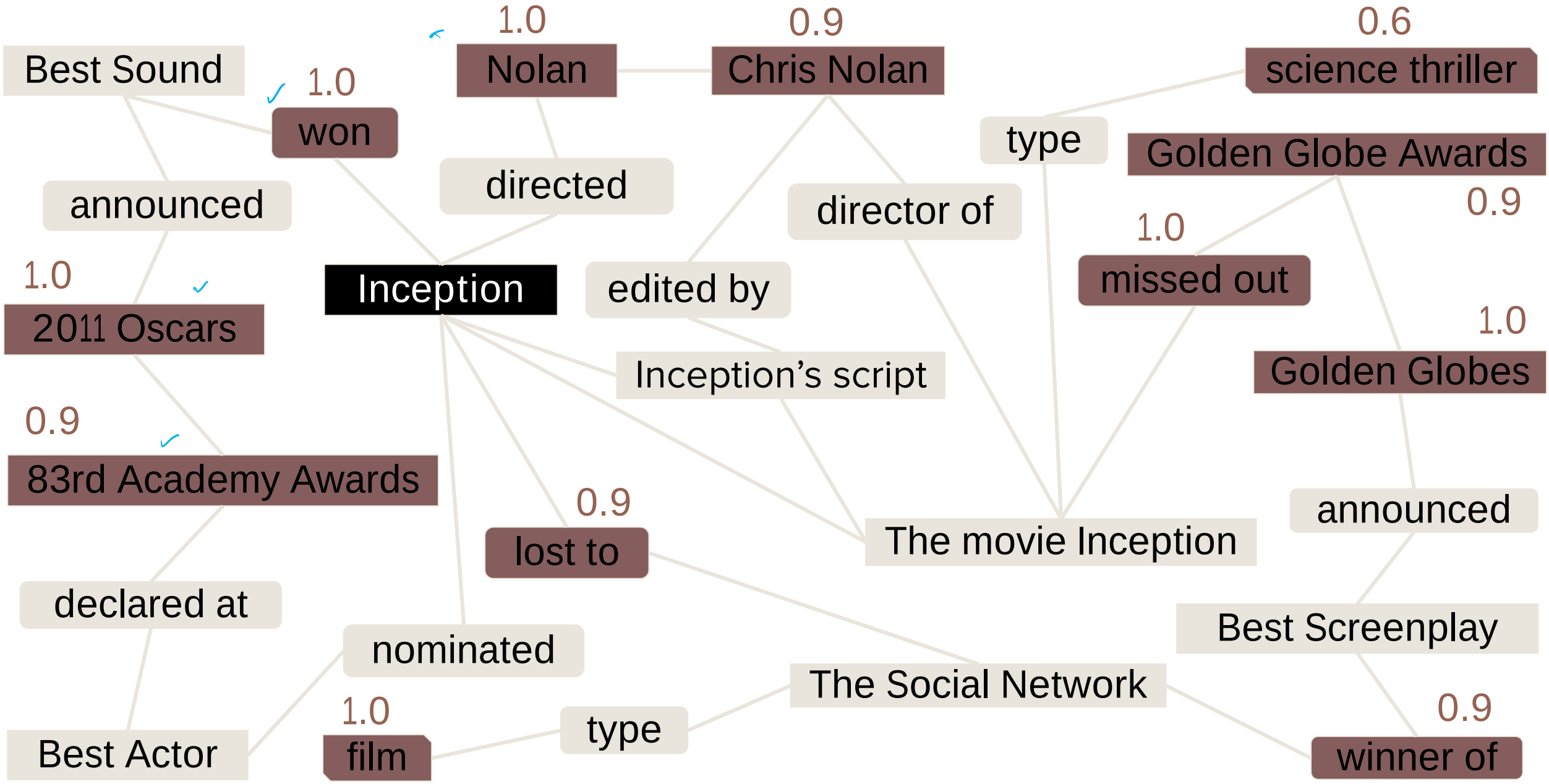


Compute node weights using similarity with question words

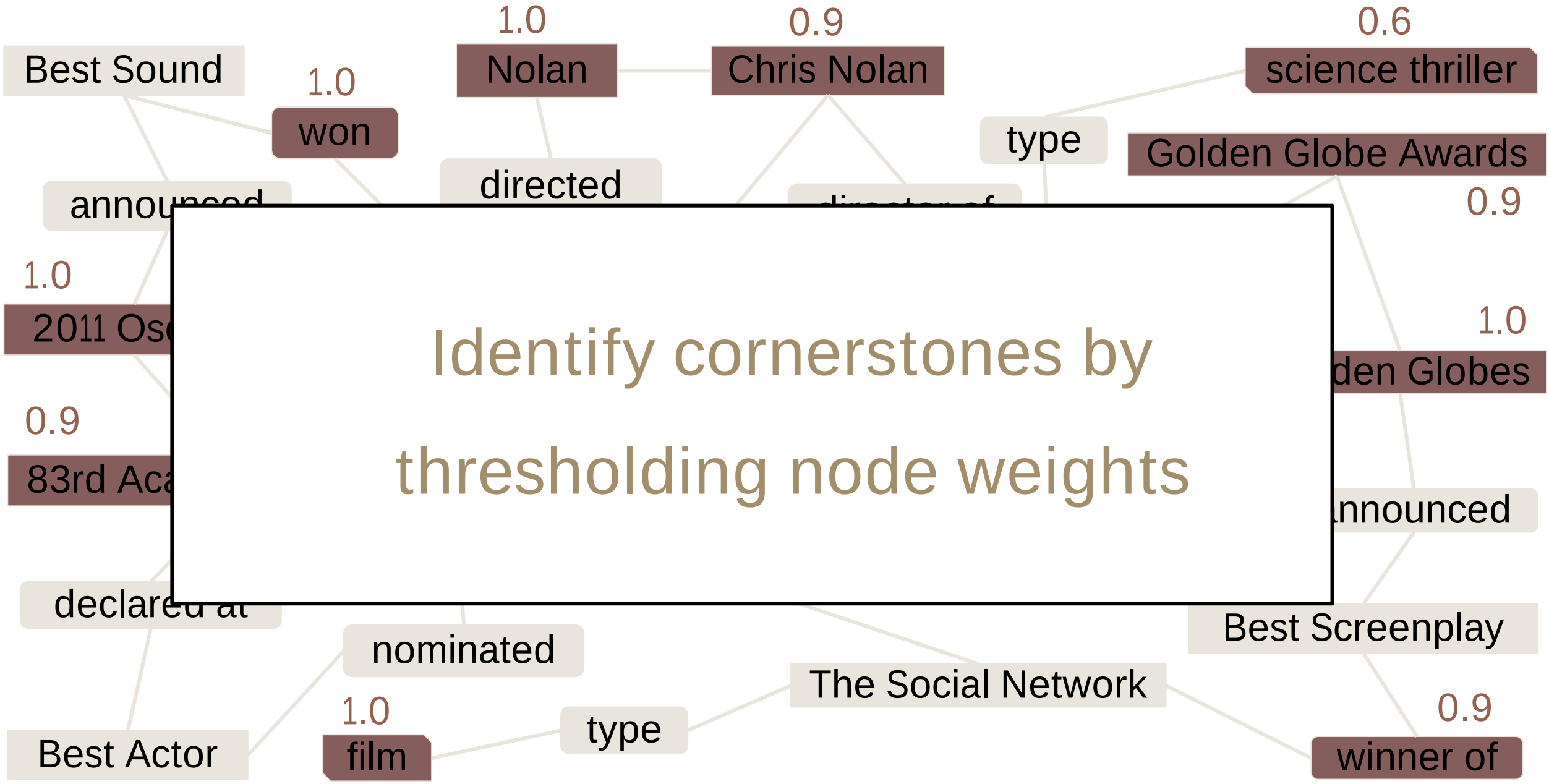
Question: Which Nolan films won an Oscar but missed a Golden Globe?



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

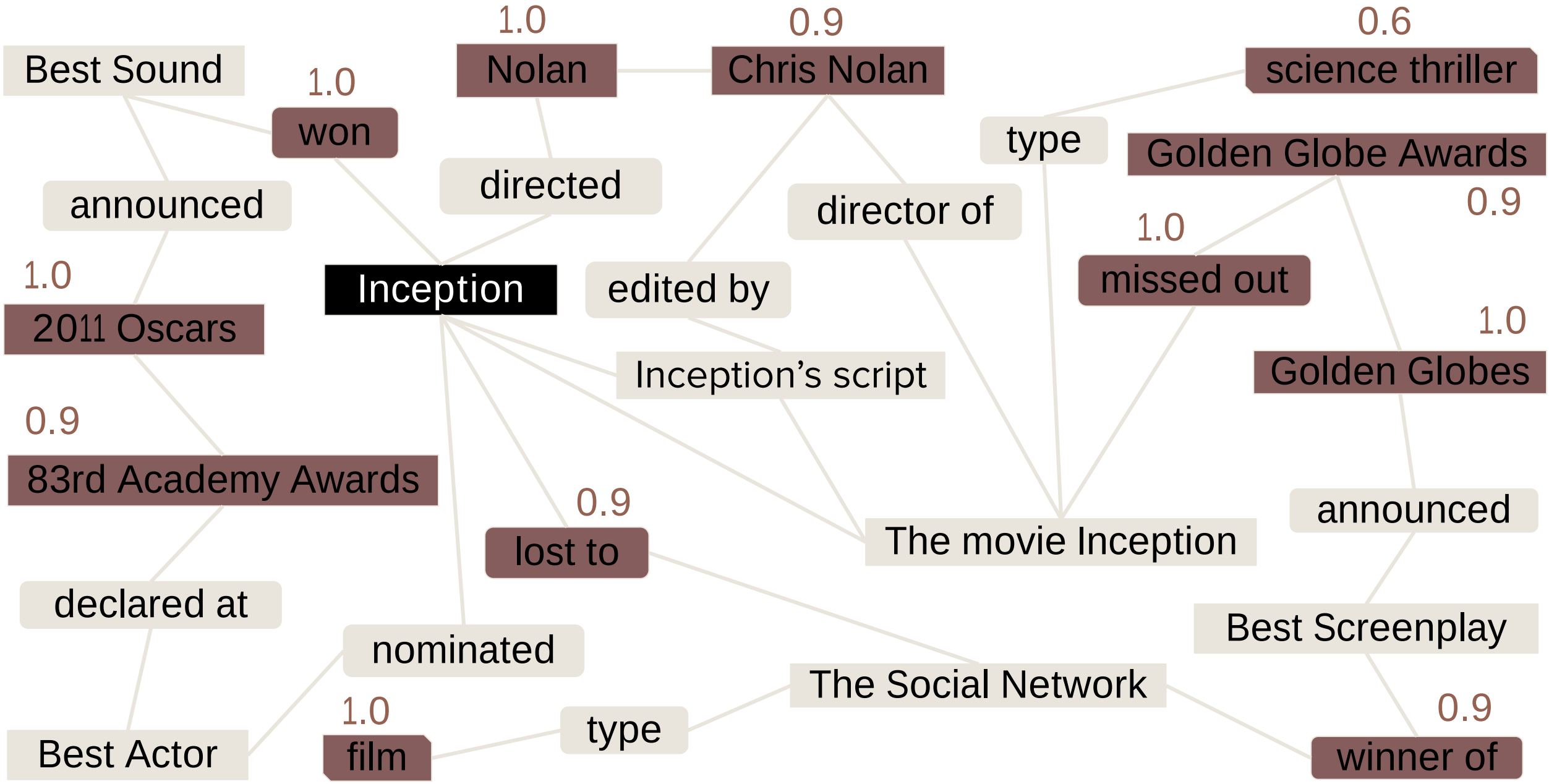


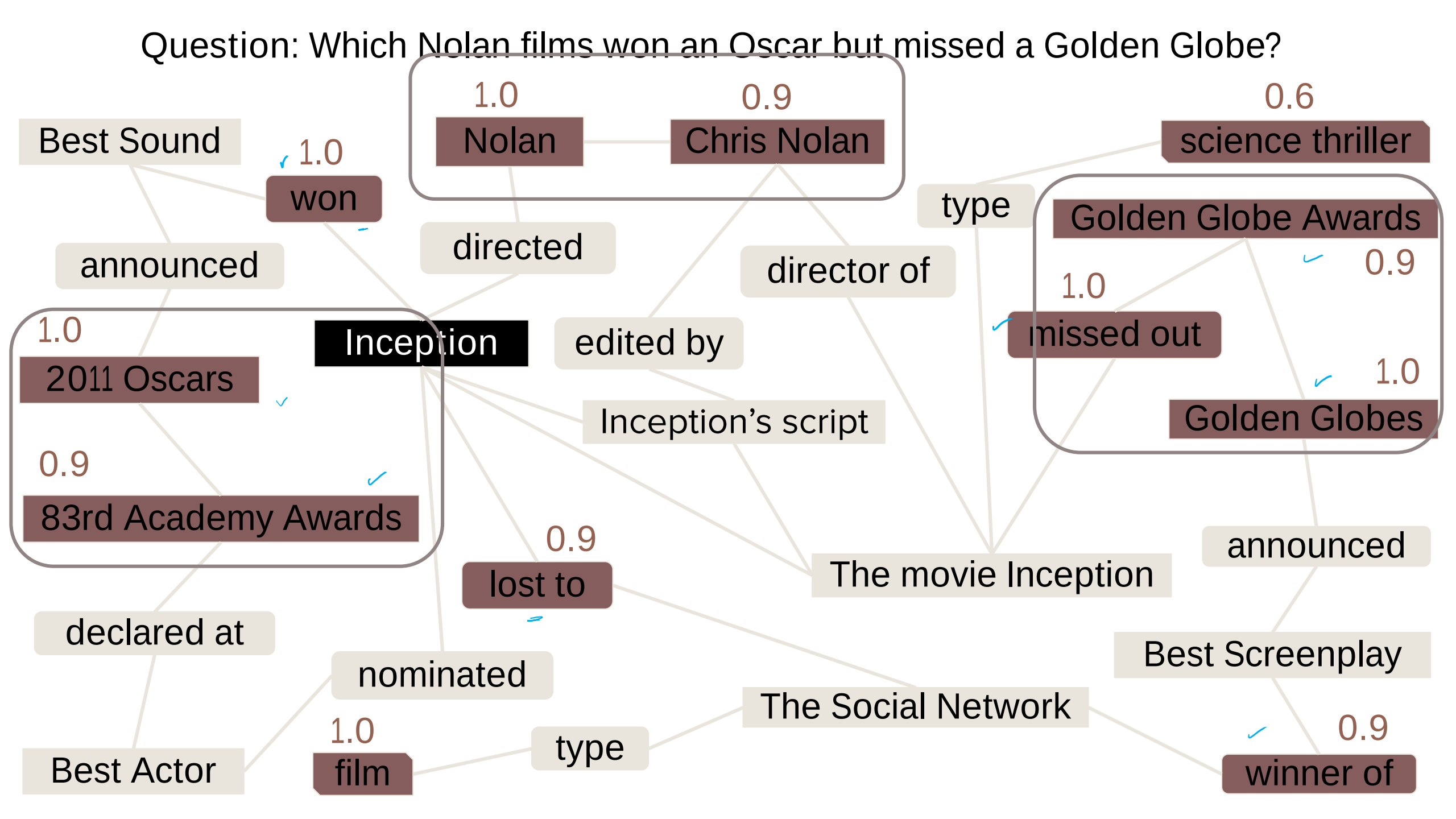
Question: Which Nolan films won an Oscar but missed a Golden Globe?

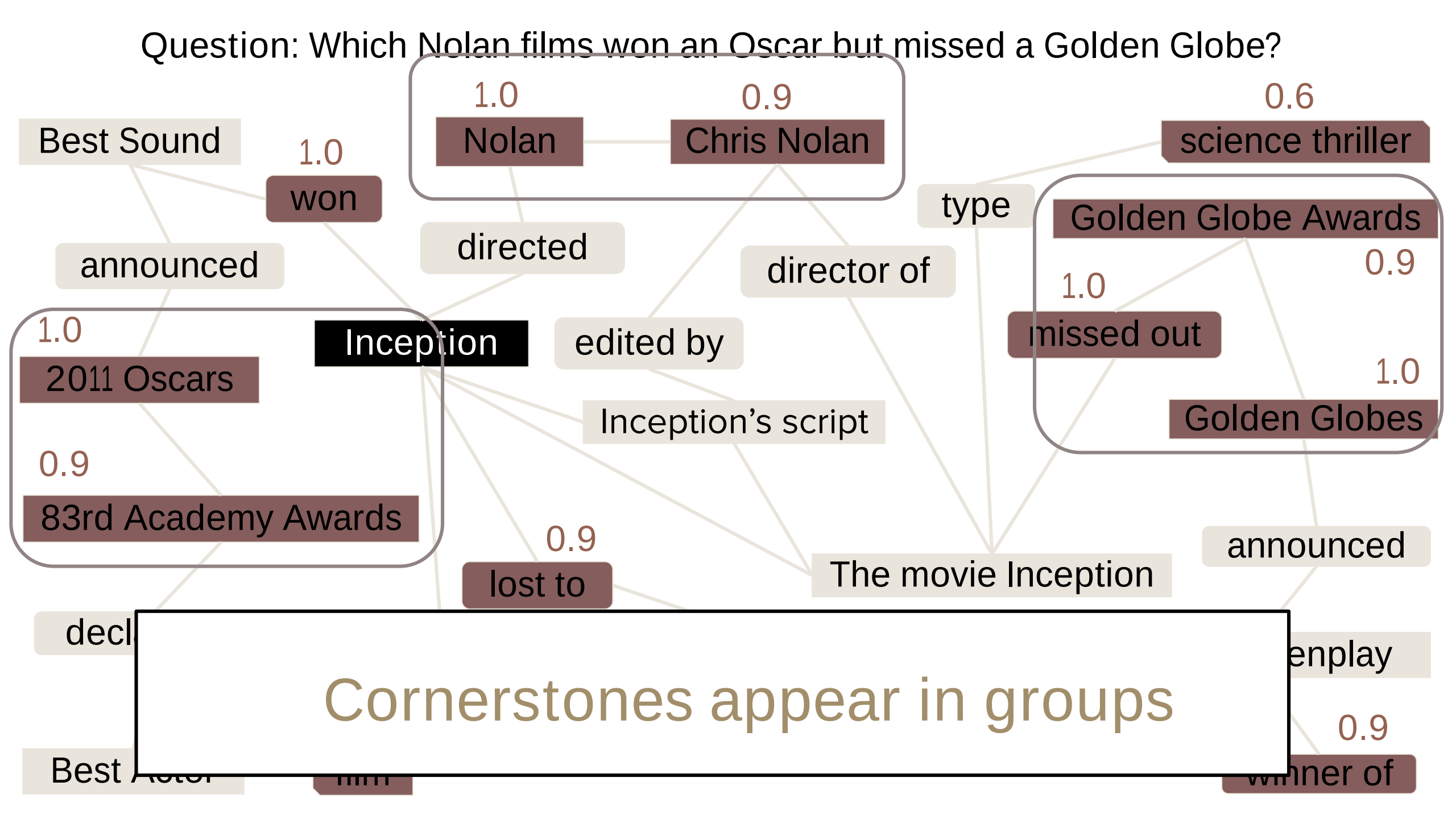


Identify cornerstones by thresholding node weights

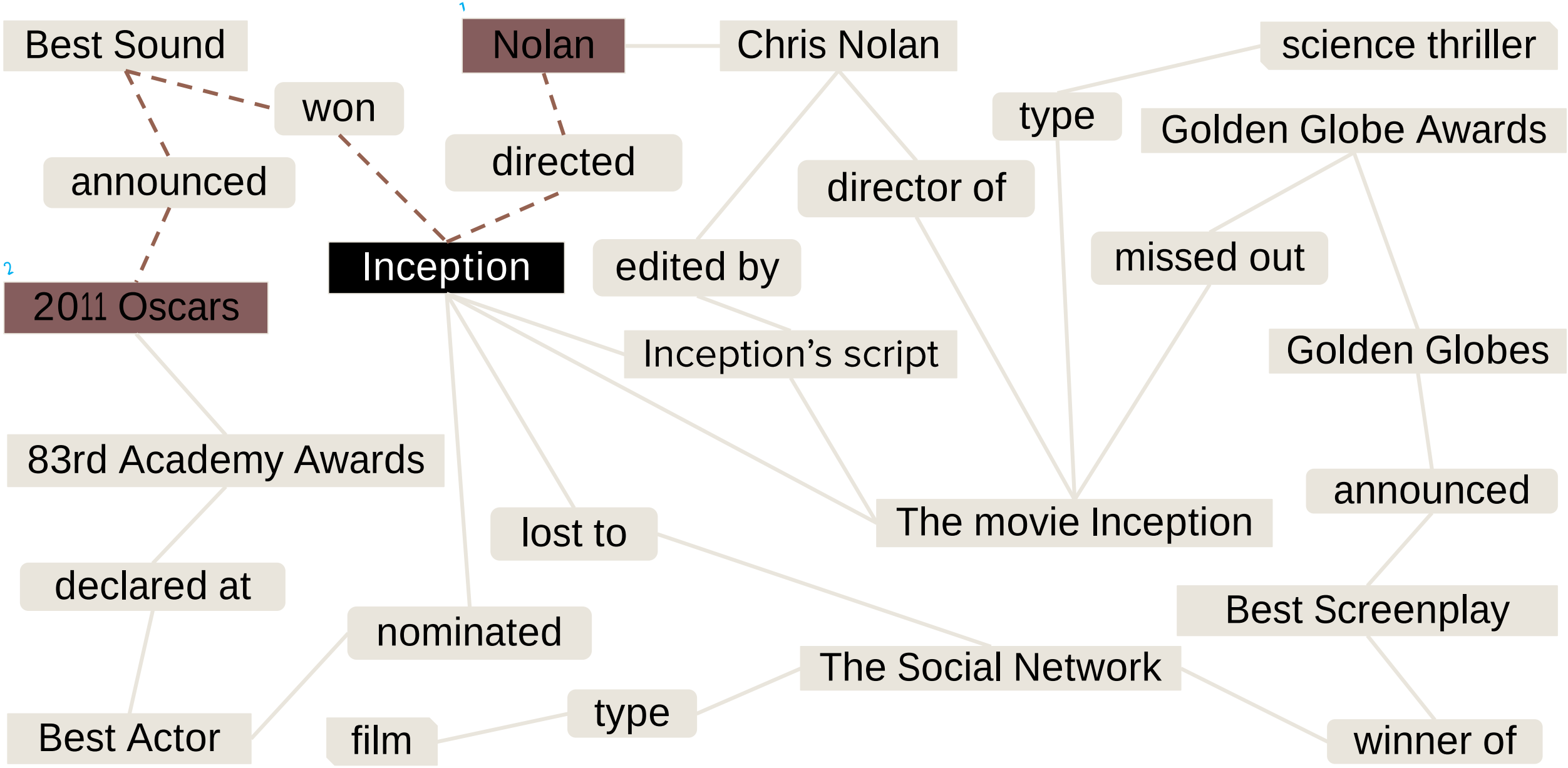
Question: Which Nolan films won an Oscar but missed a Golden Globe?



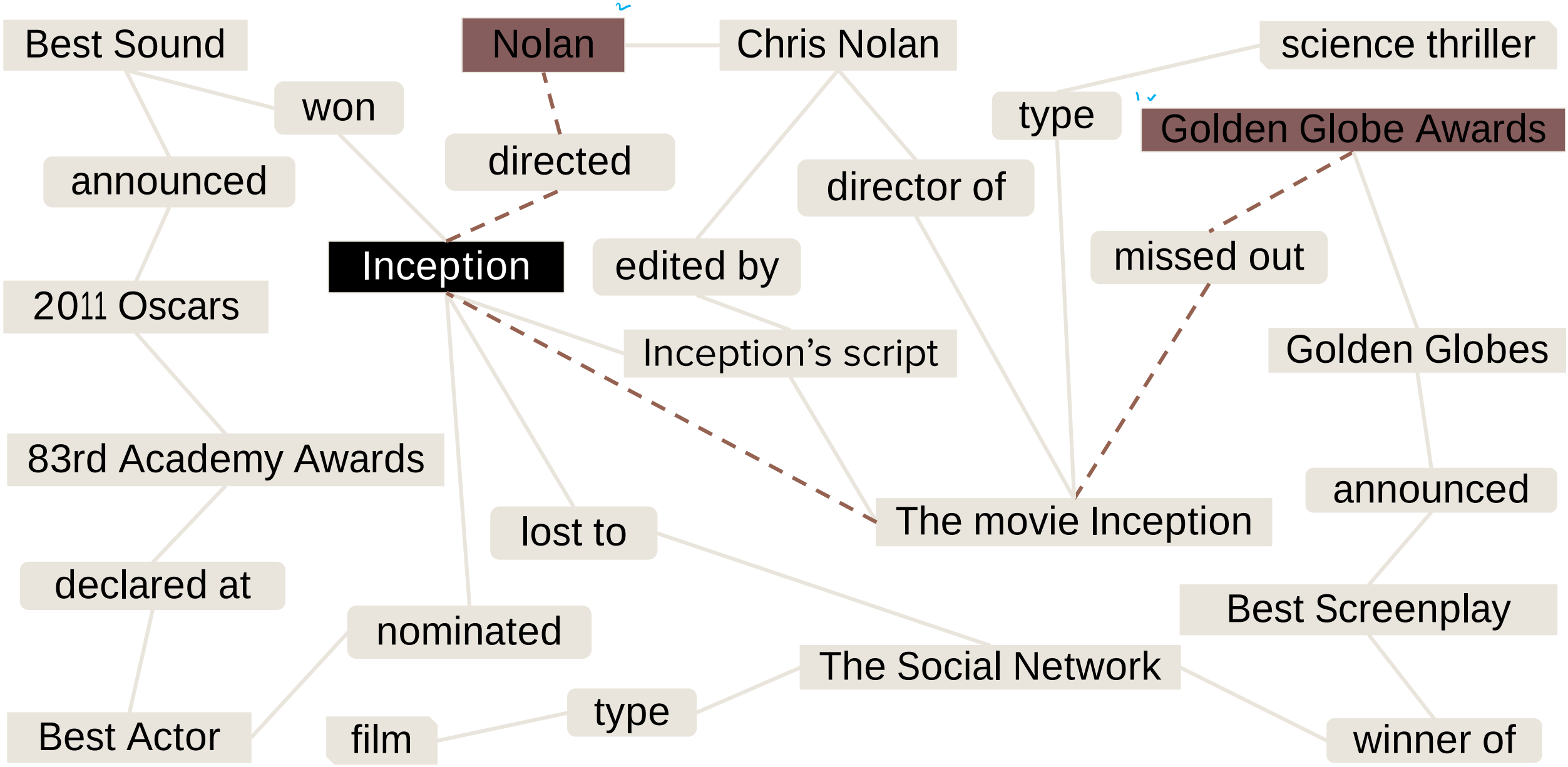

Question: Which Nolan films won an Oscar but missed a Golden Globe?
1.0
Nolan
0.9
Chris Nolan
Best Sound
1.0
won
0.6
science thriller
type
Golden Globe Awards
0.9
1.0
missed out
1.0
Golden Globes
directed
director of
announced
1.0
2011 Oscars
Inception
edited by
Inception's script
0.9
83rd Academy Awards
0.9
lost to
The movie Inception
announced
declared at
Best Screenplay
nominated
The Social Network
1.0
film
type
0.9
winner of
Best Actor


Question: Which Nolan films won an Oscar but missed a Golden Globe?
1.0
Nolan
0.9
Chris Nolan
Best Sound
1.0
won
0.6
science thriller
type
directed
Golden Globe Awards
announced
director of
0.9
1.0
missed out
1.0
Inception
edited by
2011 Oscars
1.0
Golden Globes
Inception's script
0.9
83rd Academy Awards
0.9
lost to
announced
The movie Inception
Cornerstones appear in groups
0.9

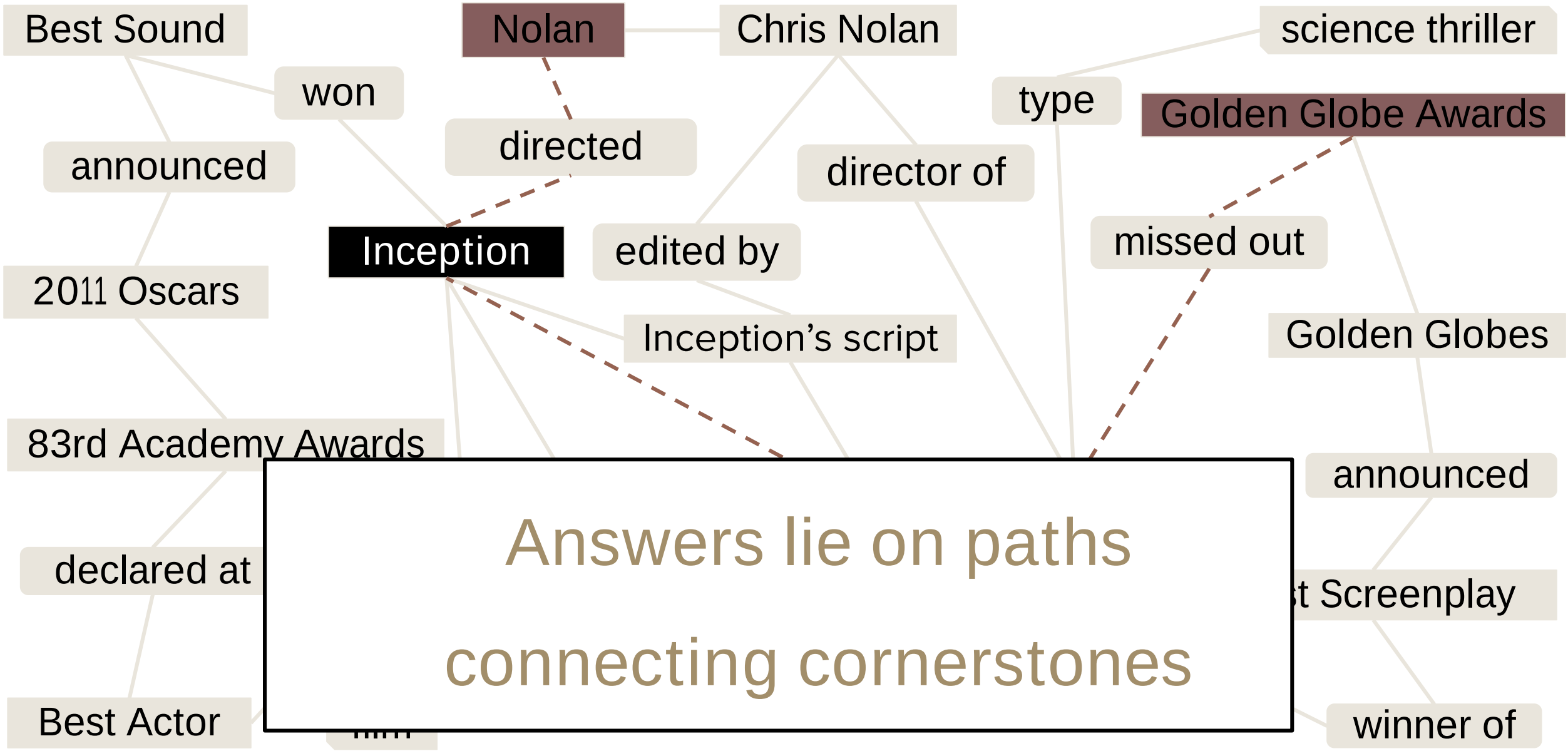
Question: Which Nolan films won an Oscar but missed a Golden Globe?



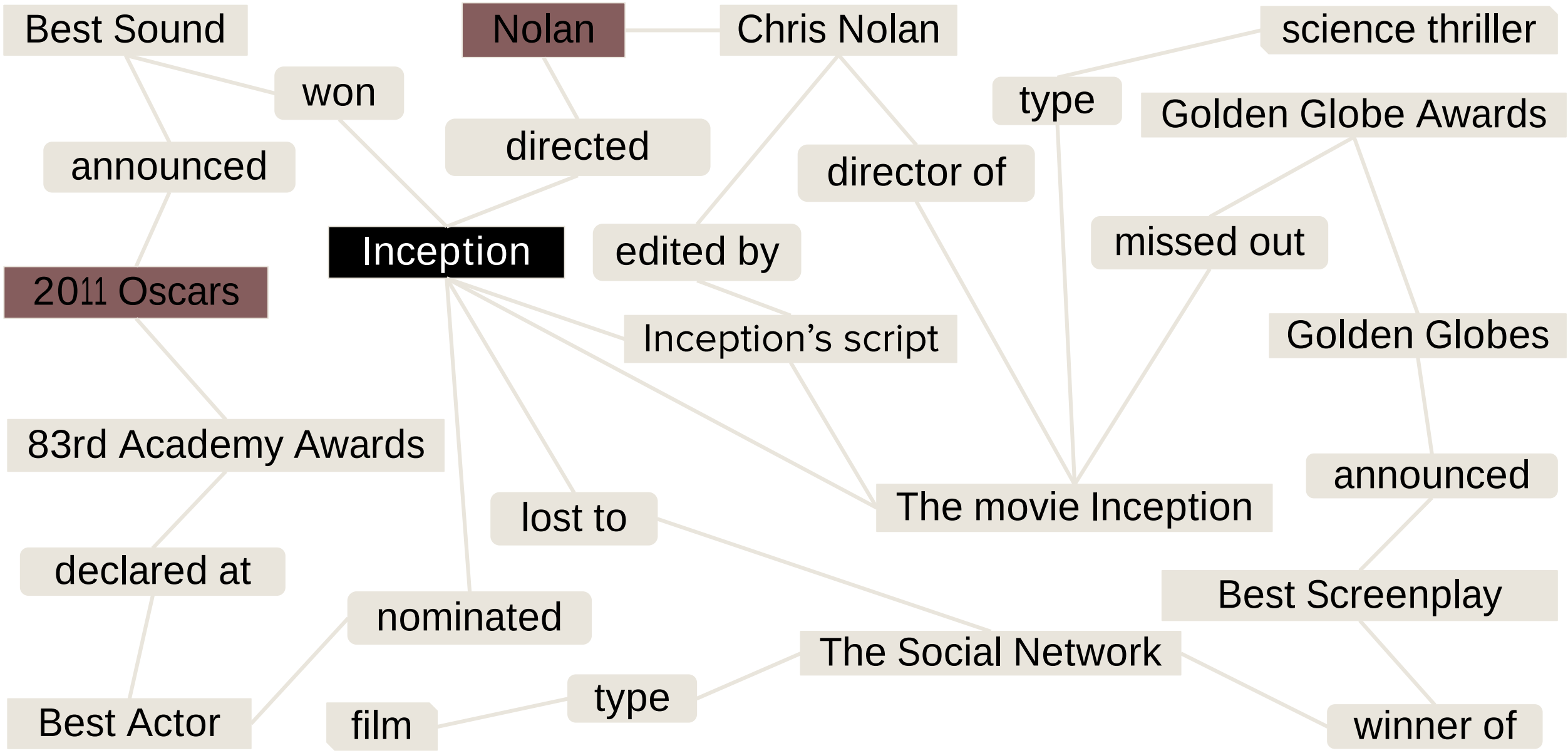
Question: Which Nolan films won an Oscar but missed a Golden Globe?



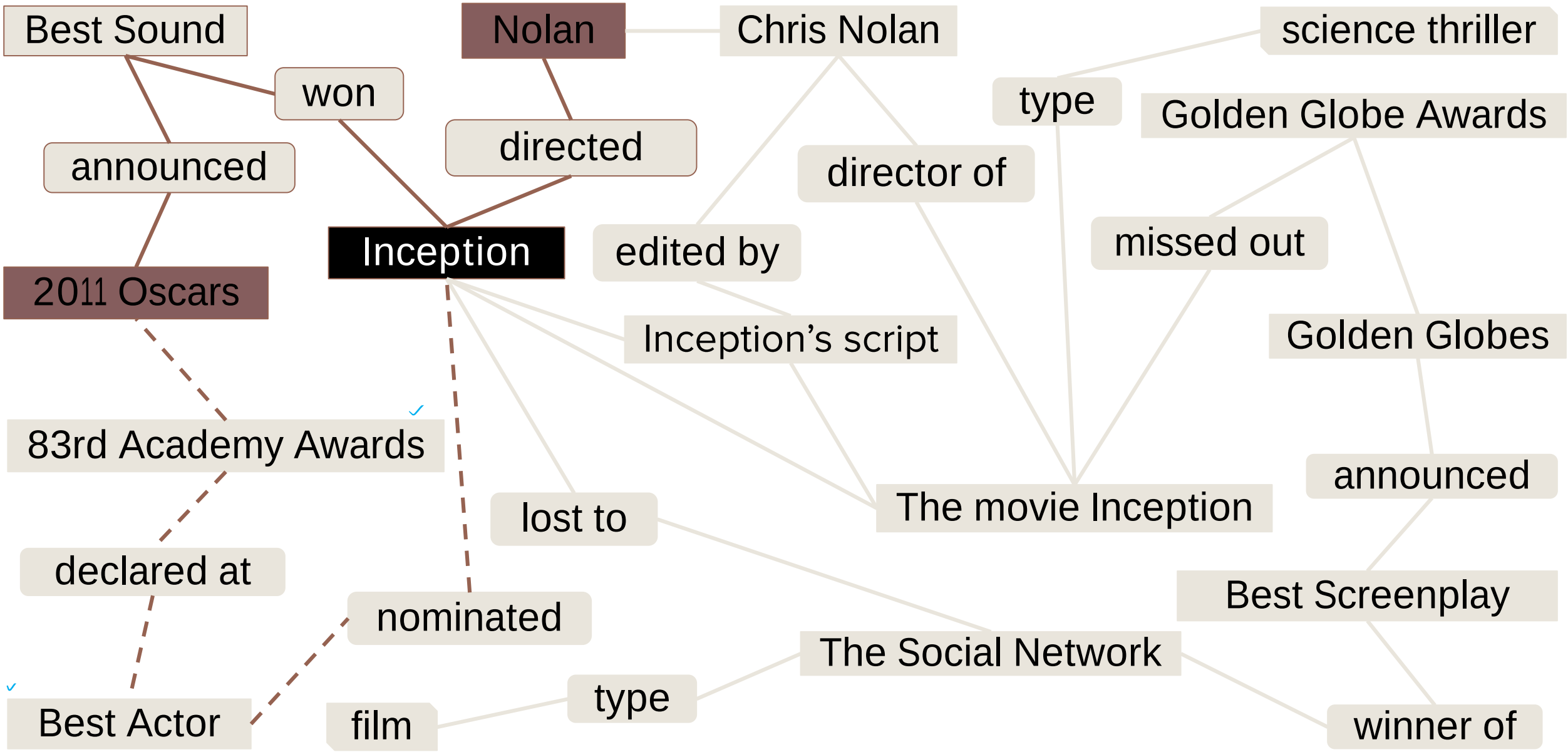
Question: Which Nolan films won an Oscar but missed a Golden Globe?



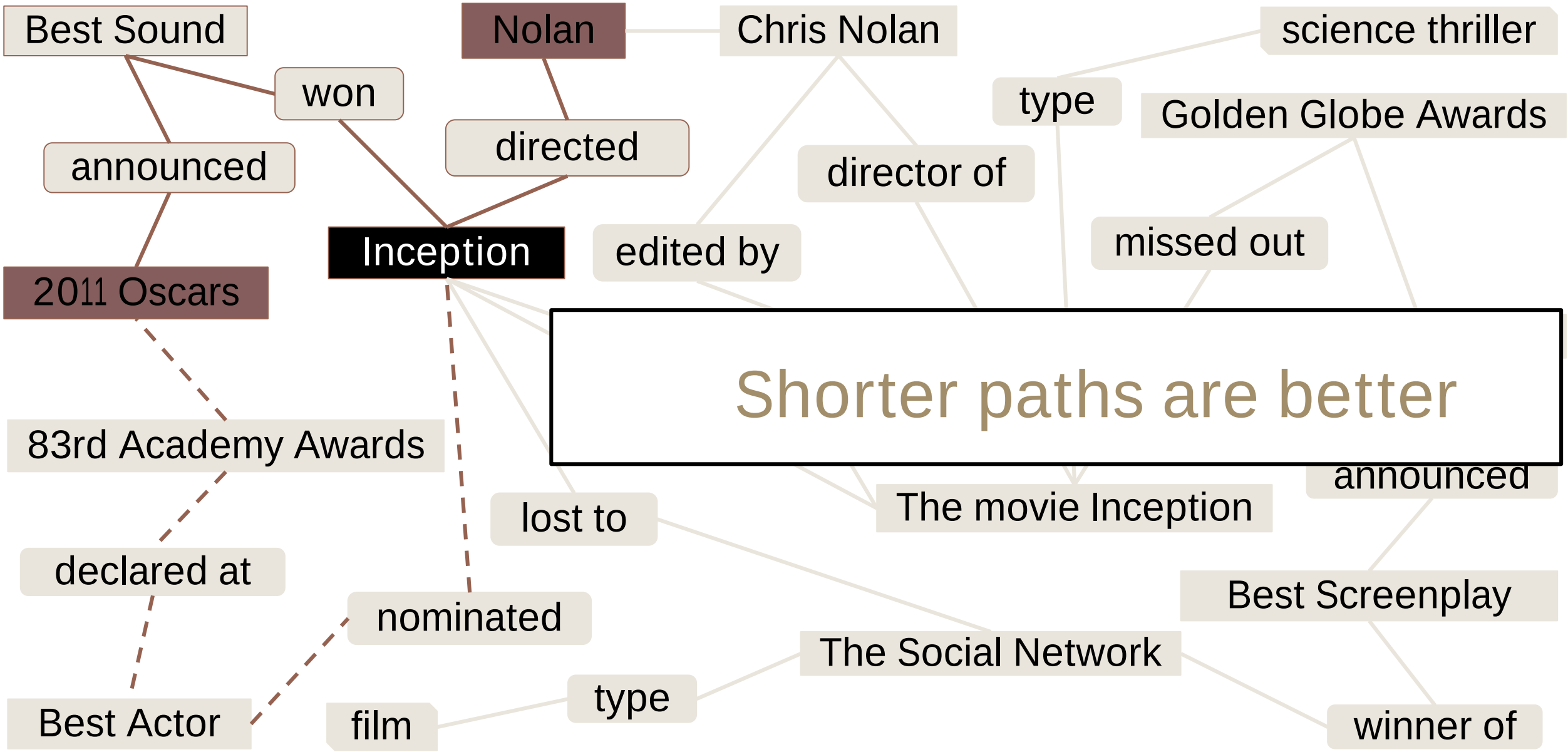
Question: Which Nolan films won an Oscar but missed a Golden Globe?



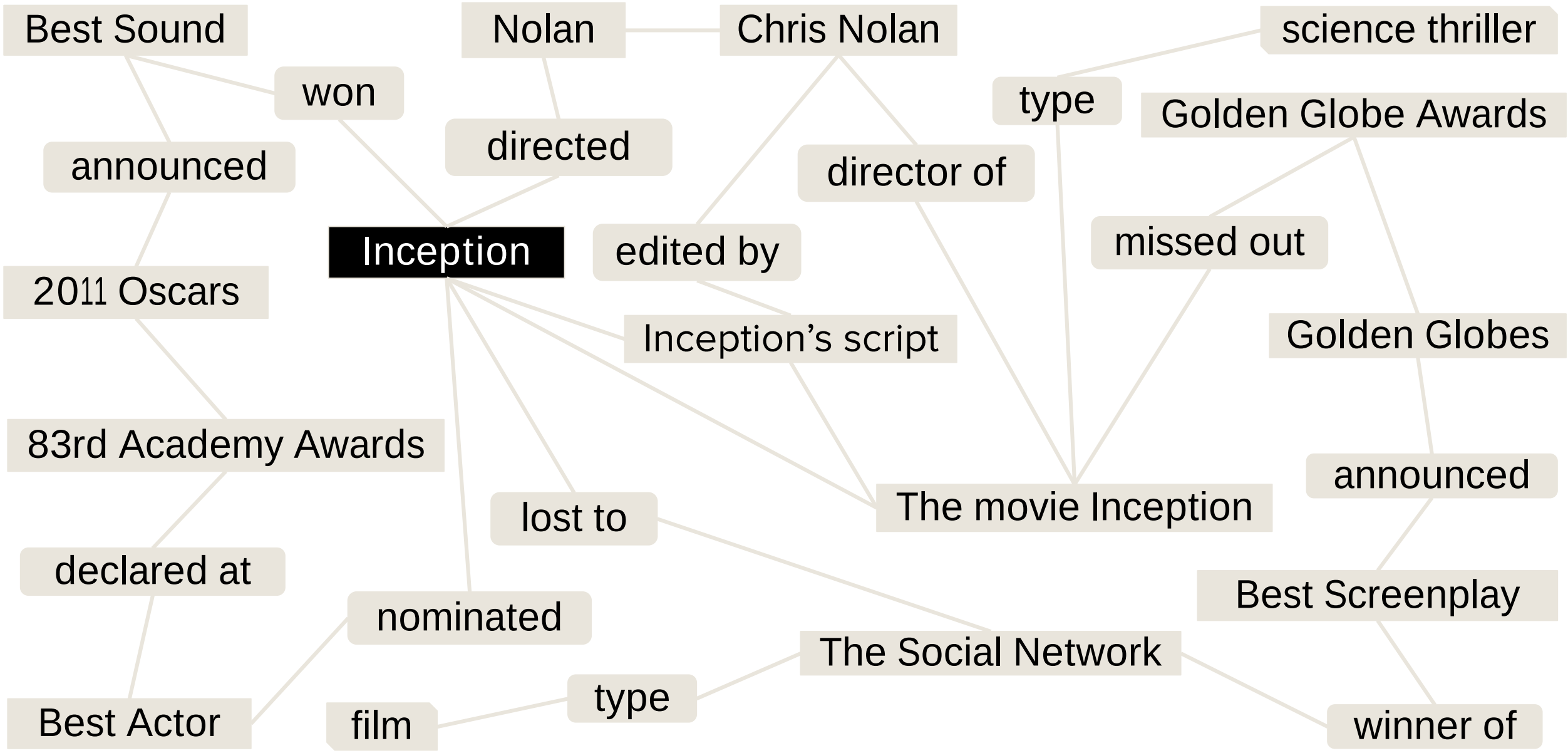
Question: Which Nolan films won an Oscar but missed a Golden Globe?



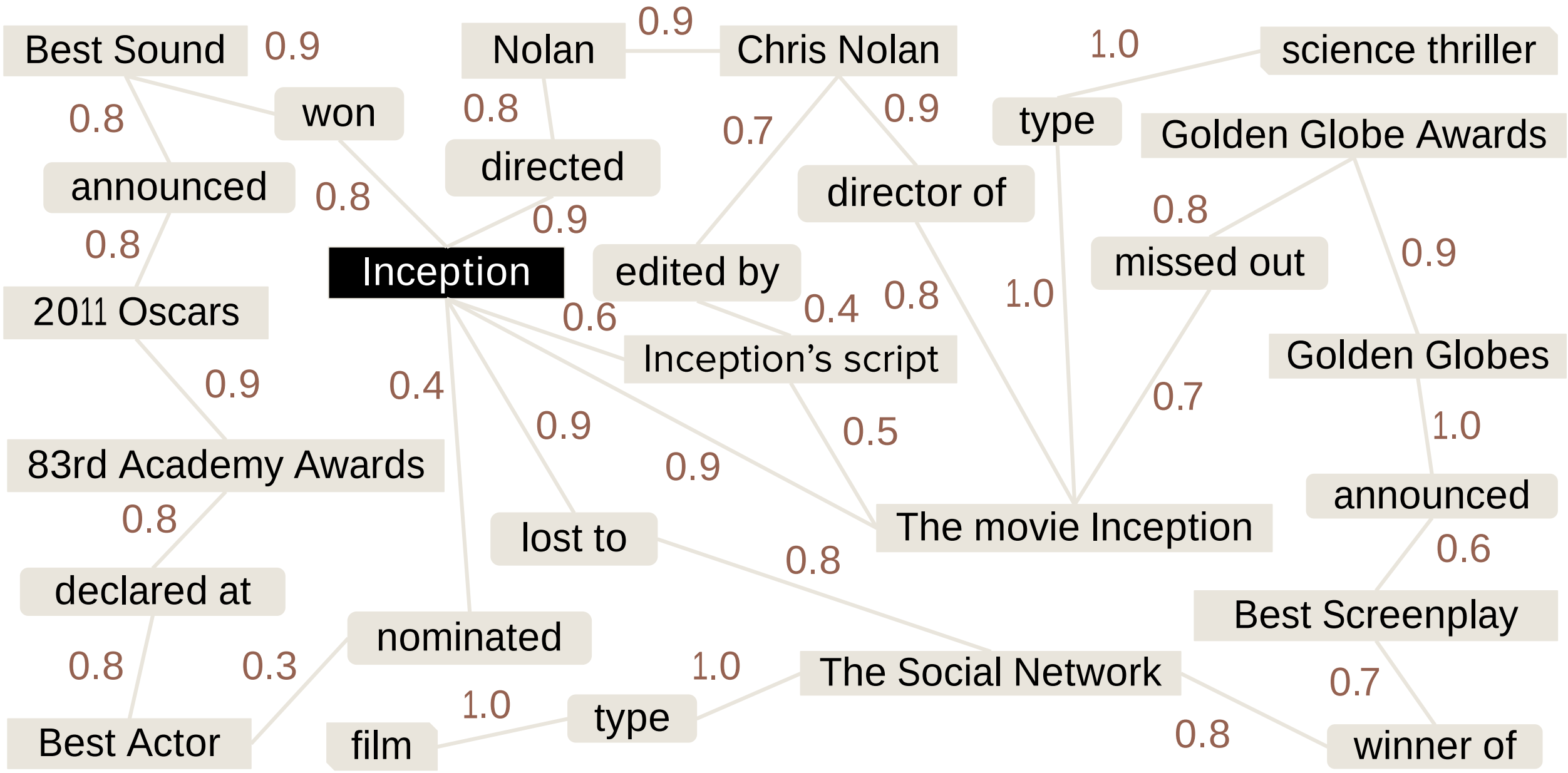
Question: Which Nolan films won an Oscar but missed a Golden Globe?



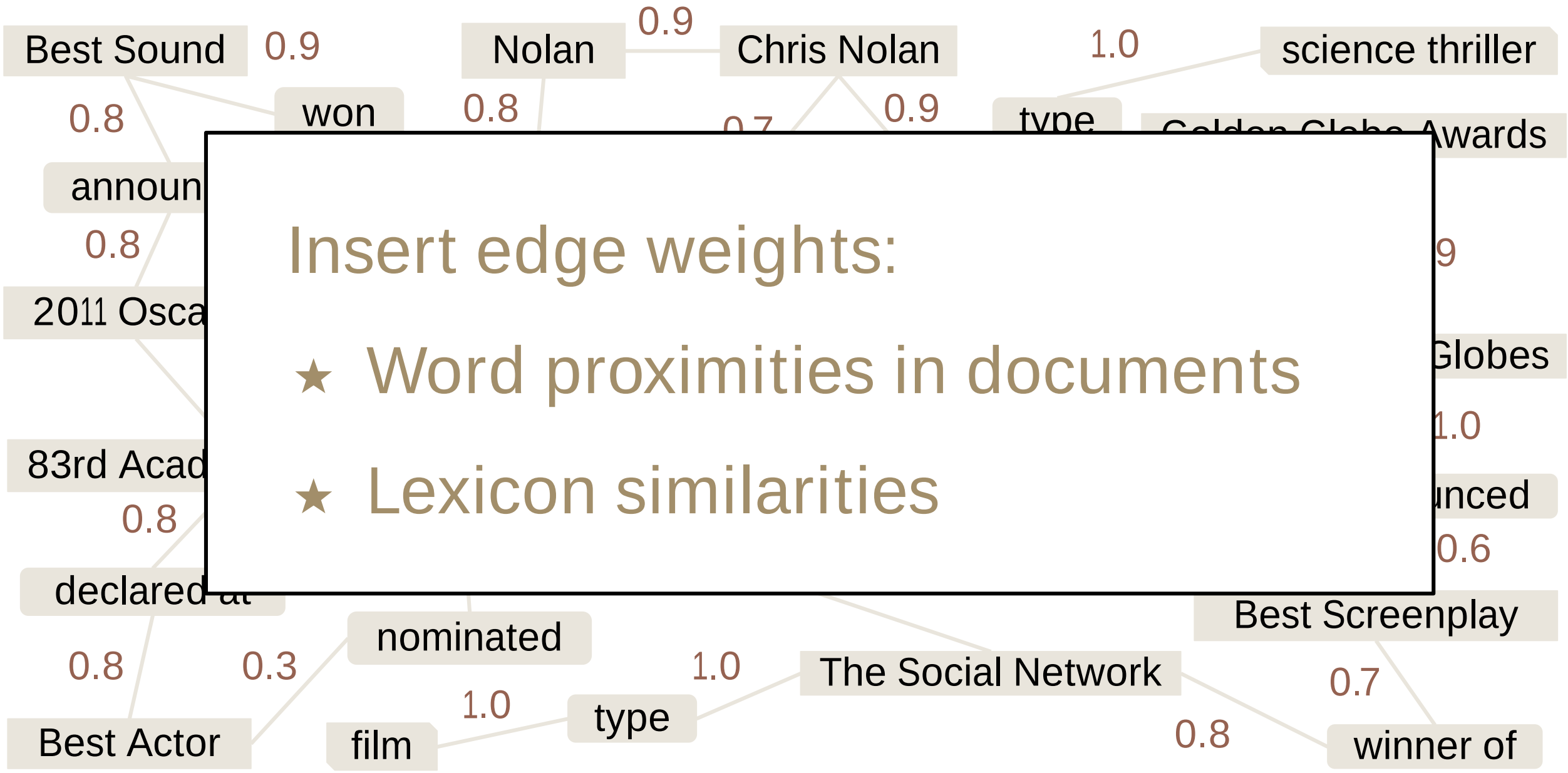
Question: Which Nolan films won an Oscar but missed a Golden Globe?



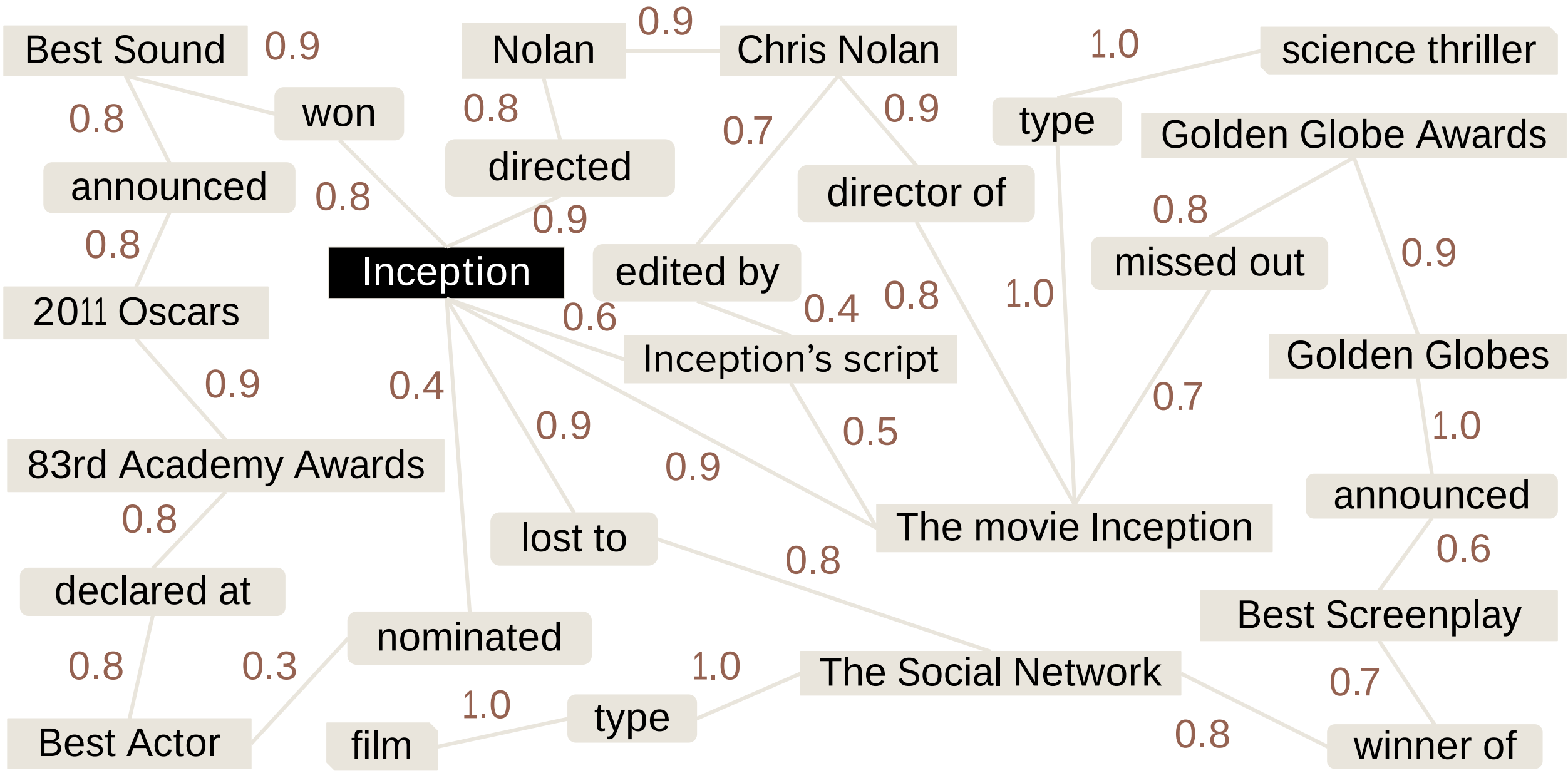
Question: Which Nolan films won an Oscar but missed a Golden Globe?



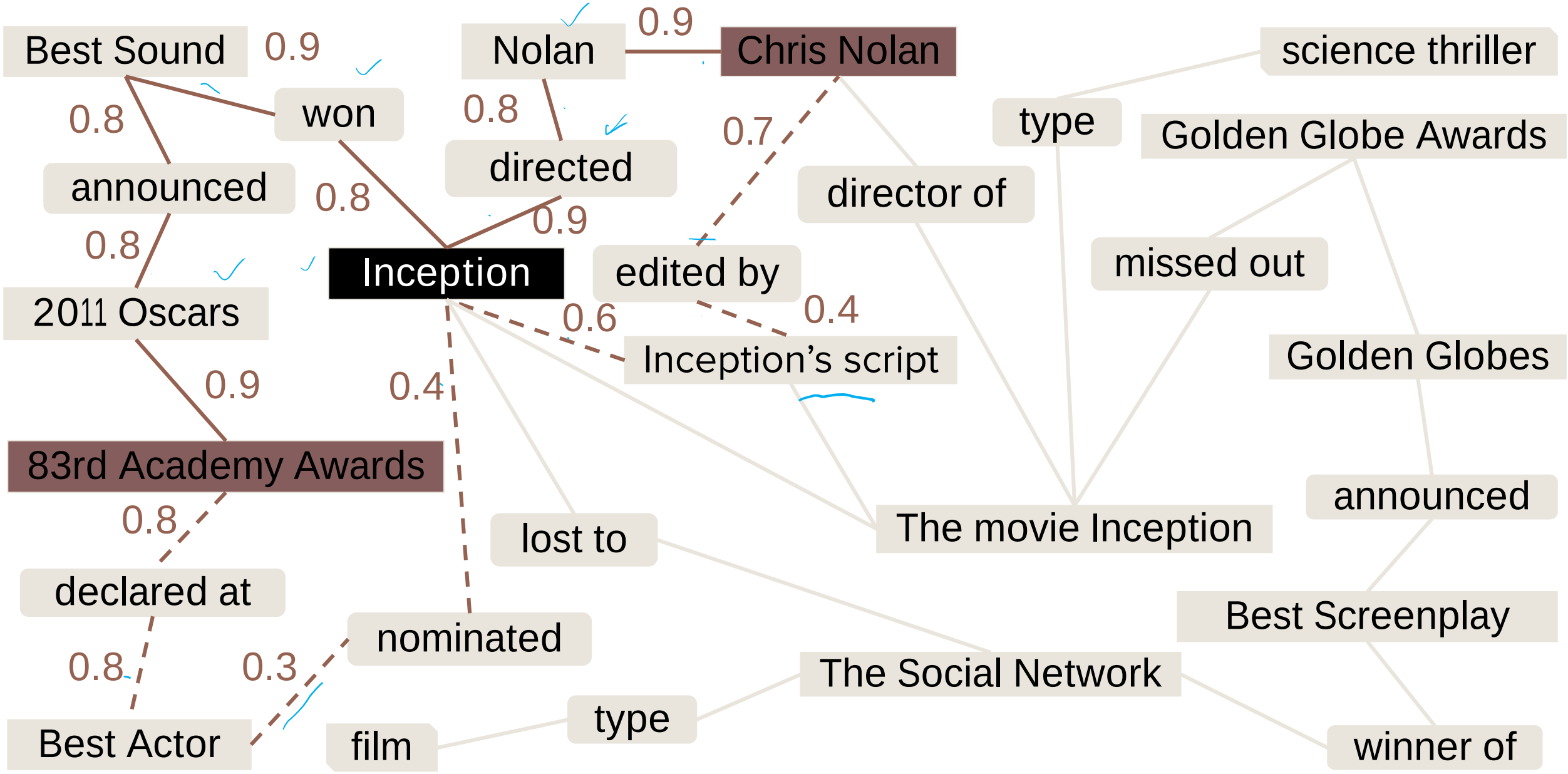
Question: Which Nolan films won an Oscar but missed a Golden Globe?



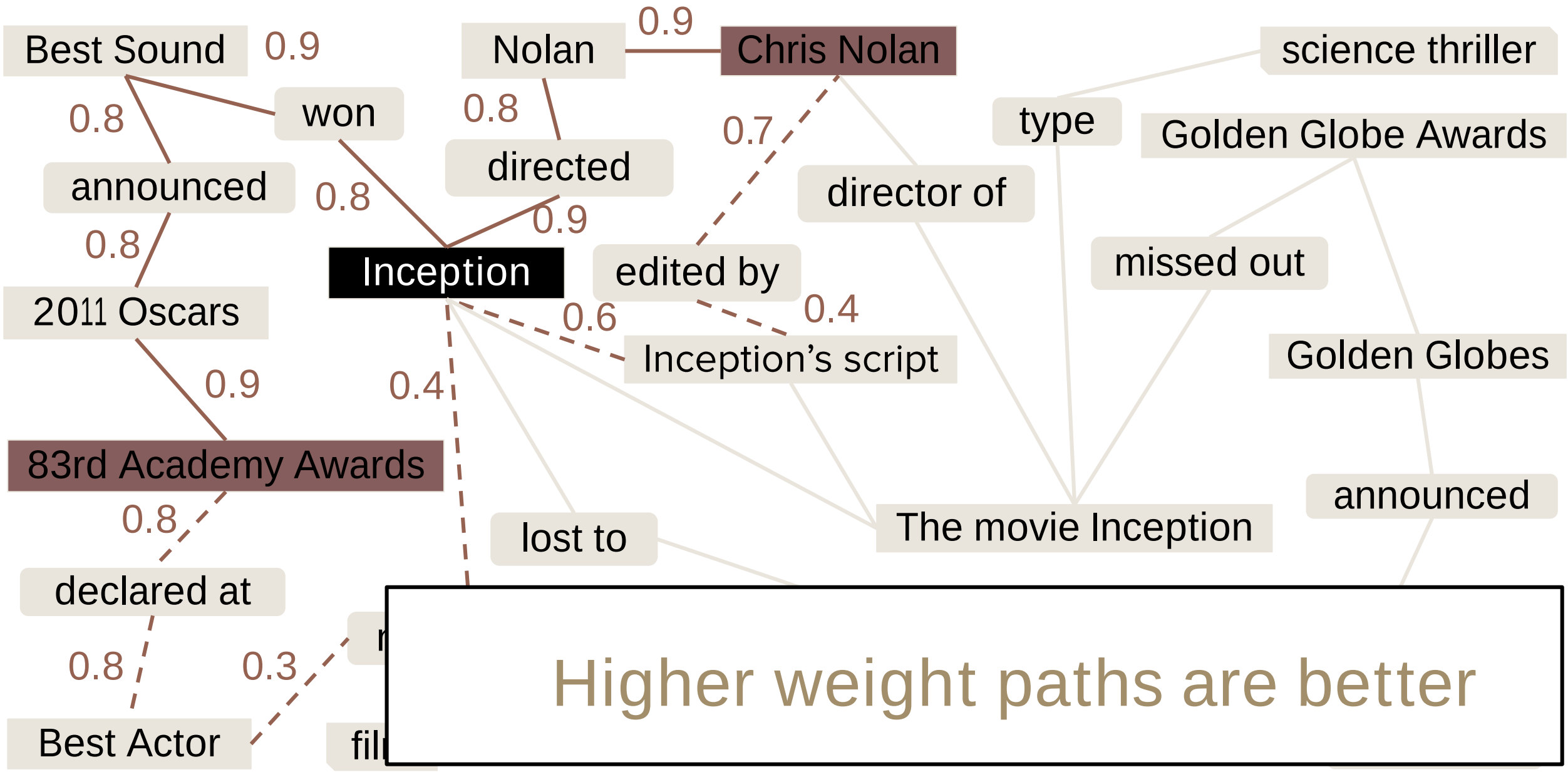
Question: Which Nolan films won an Oscar but missed a Golden Globe?



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: $G = (\overbrace{N, E}, \overbrace{M_N^T, M_E^T}, \overbrace{M_N^W, M_E^W})$

Nodes and edges

Mapping nodes and edges to types

Mapping nodes and edges to weights

- Cornerstone groups: Set of sets $C = \{C_1, C_2, \ldots, C_n\}$
- At least one instance from each group is necessary in optimal subgraph

Formalizing intuitions: What we have

1. Answers on paths connecting cornerstones
2. Cost of path = Sum of edge costs (1 - edge weights)
3. Higher weight paths have lower costs
4. Shorter paths have lower costs

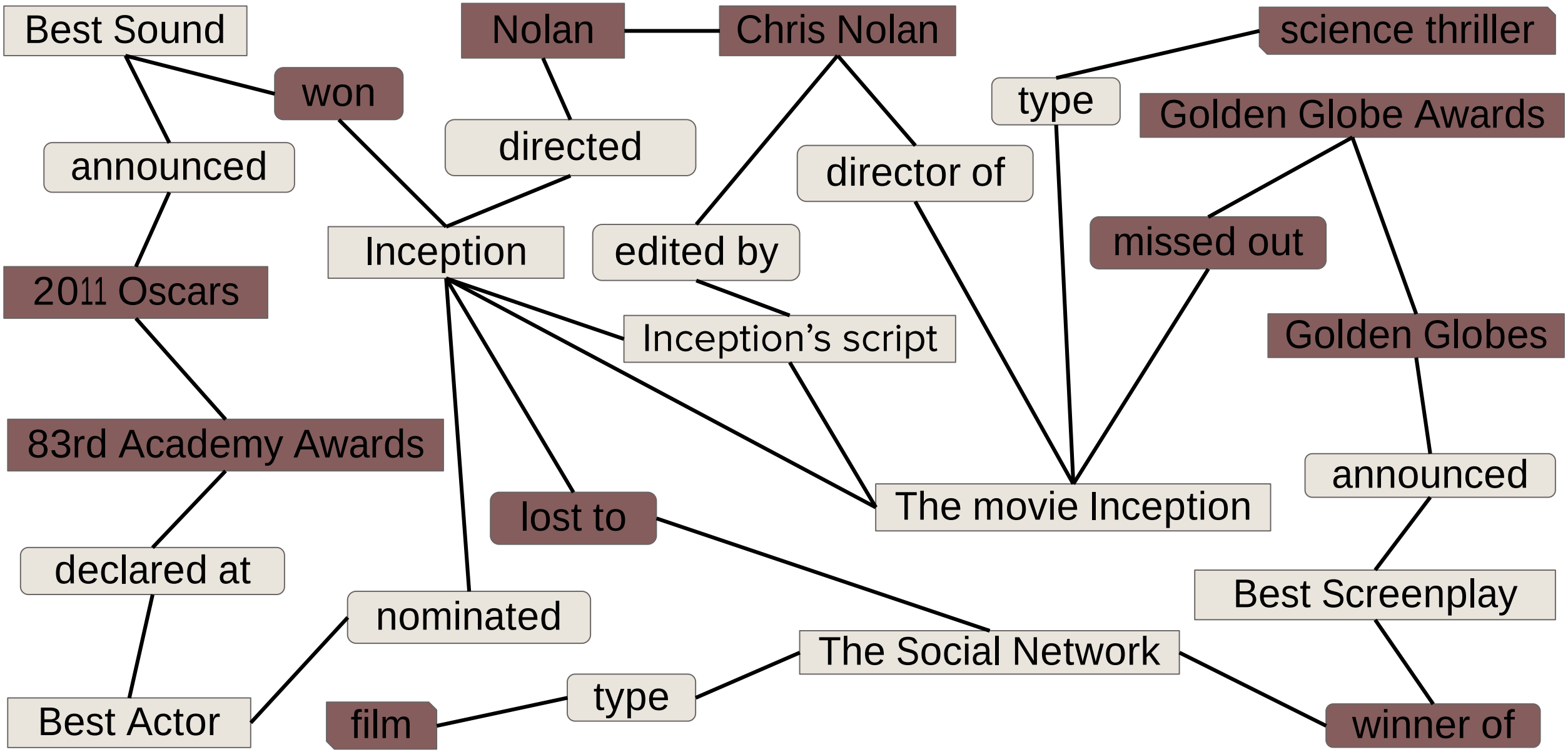
Steiner Trees

- Given:
  - Undirected and weighted graph G
  - Subset of vertices = Terminals T_1, T_2, ..., T_n
- Find: Minimum weight tree containing all terminals
- Cornerstones = Terminals

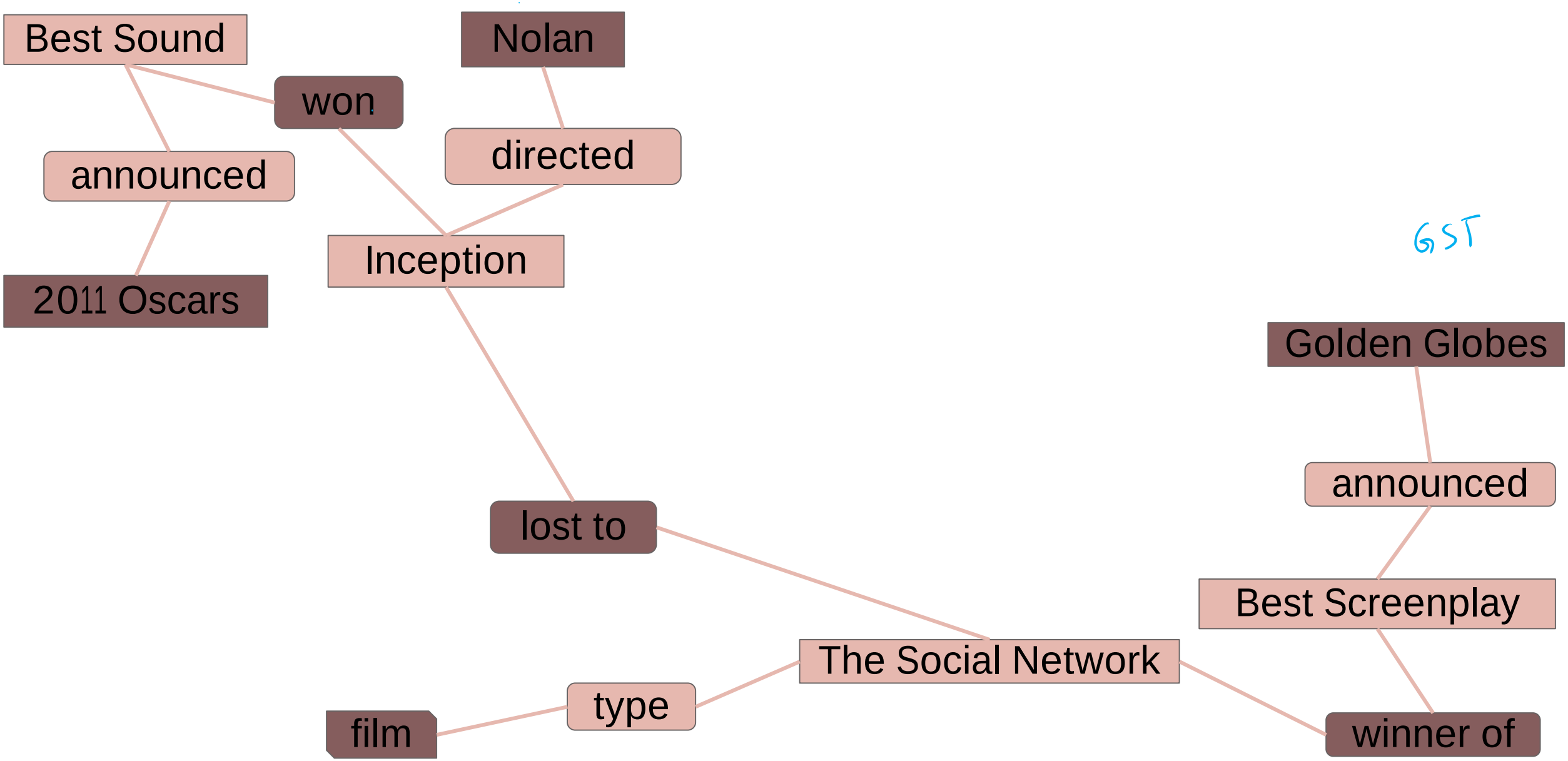
Group Steiner Trees

- Terminals occur as groups C_i (cornerstones)
- Group Steiner Tree contains at least one terminal from each group
- Compute Group Steiner Trees on Quasi-KG
- We use method from Ding et al. [2007]
- Dynamic programming exponential in #terminals but $O(n \log n)$ in graph size

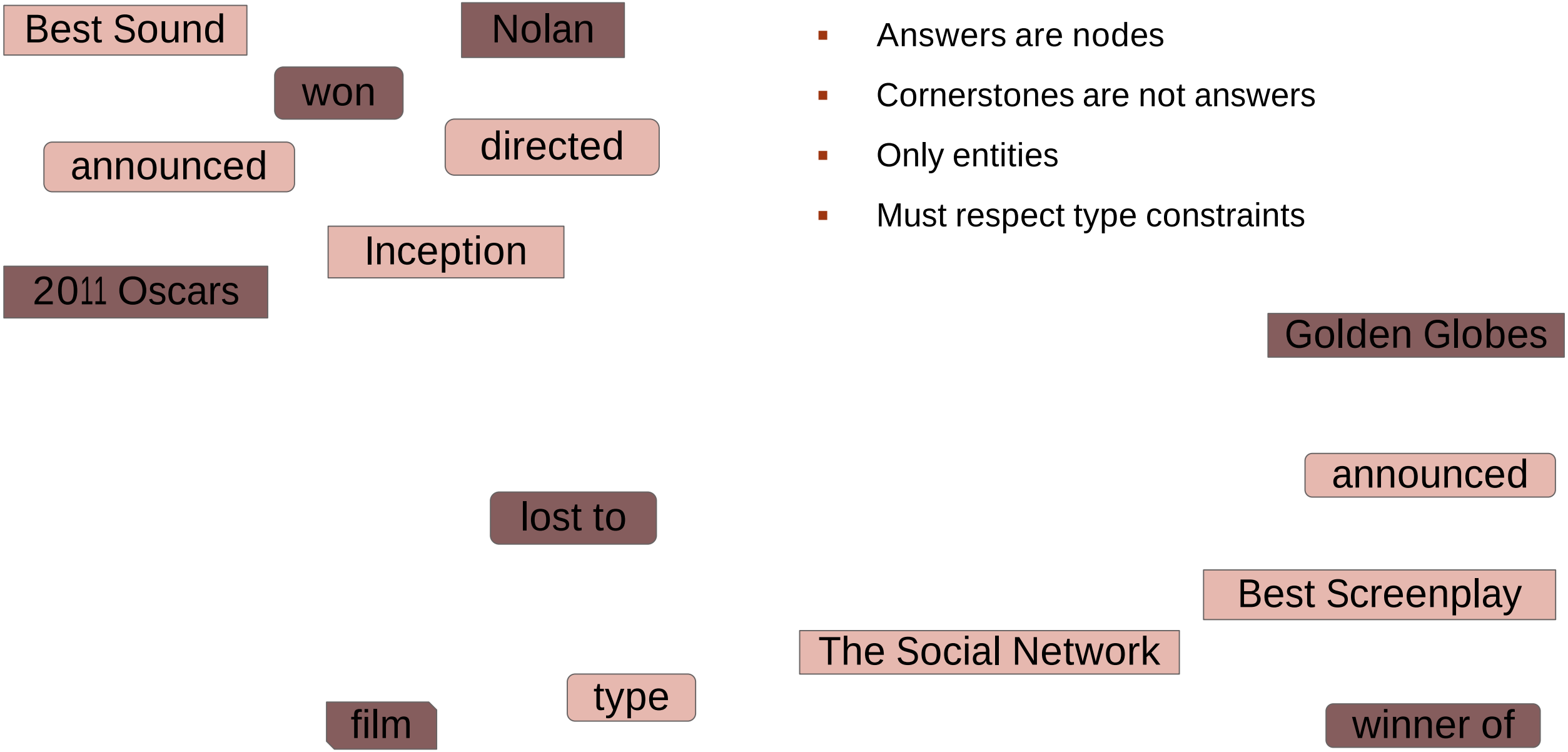
Question: Which Nolan films won an Oscar but missed a Golden Globe?



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

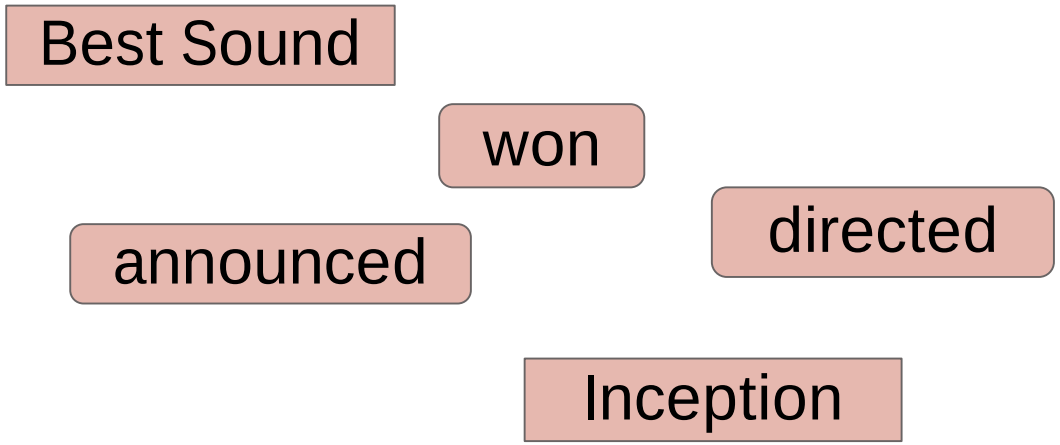


Question: Which Nolan films won an Oscar but missed a Golden Globe?

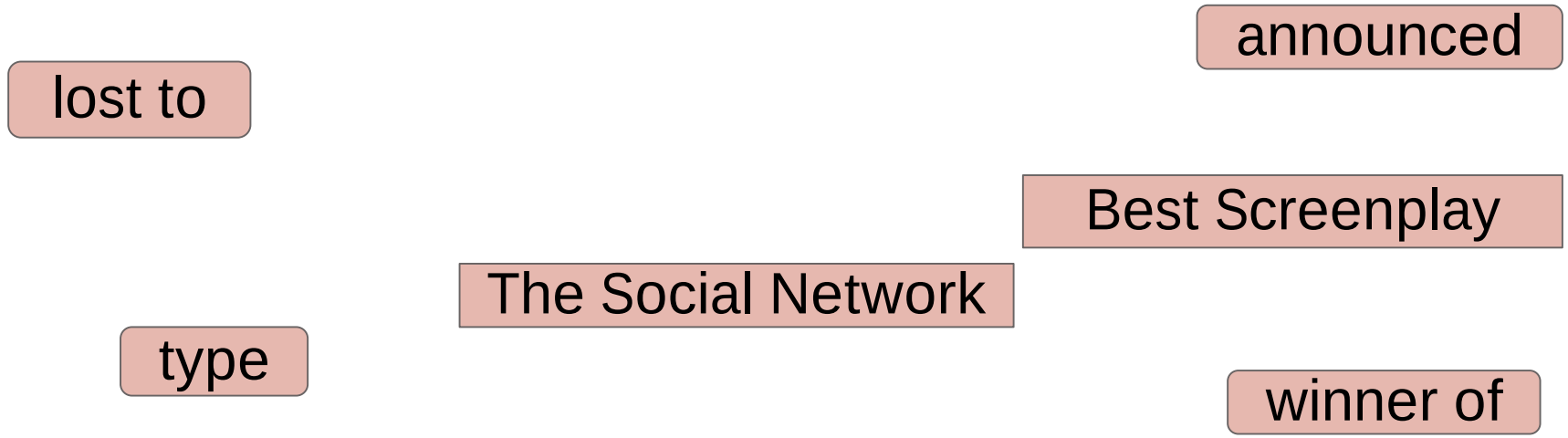


- Answers are nodes
- Cornerstones are not answers
- Only entities
- Must respect type constraints

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

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

Inception

Best Screenplay

The Social Network

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 multiple criteria
- Best answer chosen

Inception

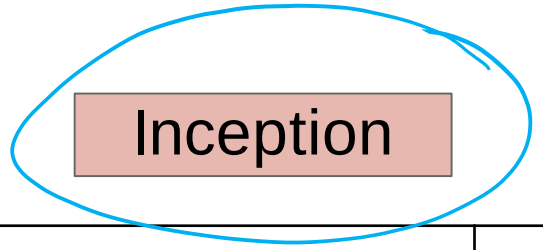
Sum of node weights	9.75
Number of GSTs	5
Sum of tree costs	6.78
Distance to cornerstones	19

The Social Network

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

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

- Answers ranked by multiple criteria
- Best answer chosen

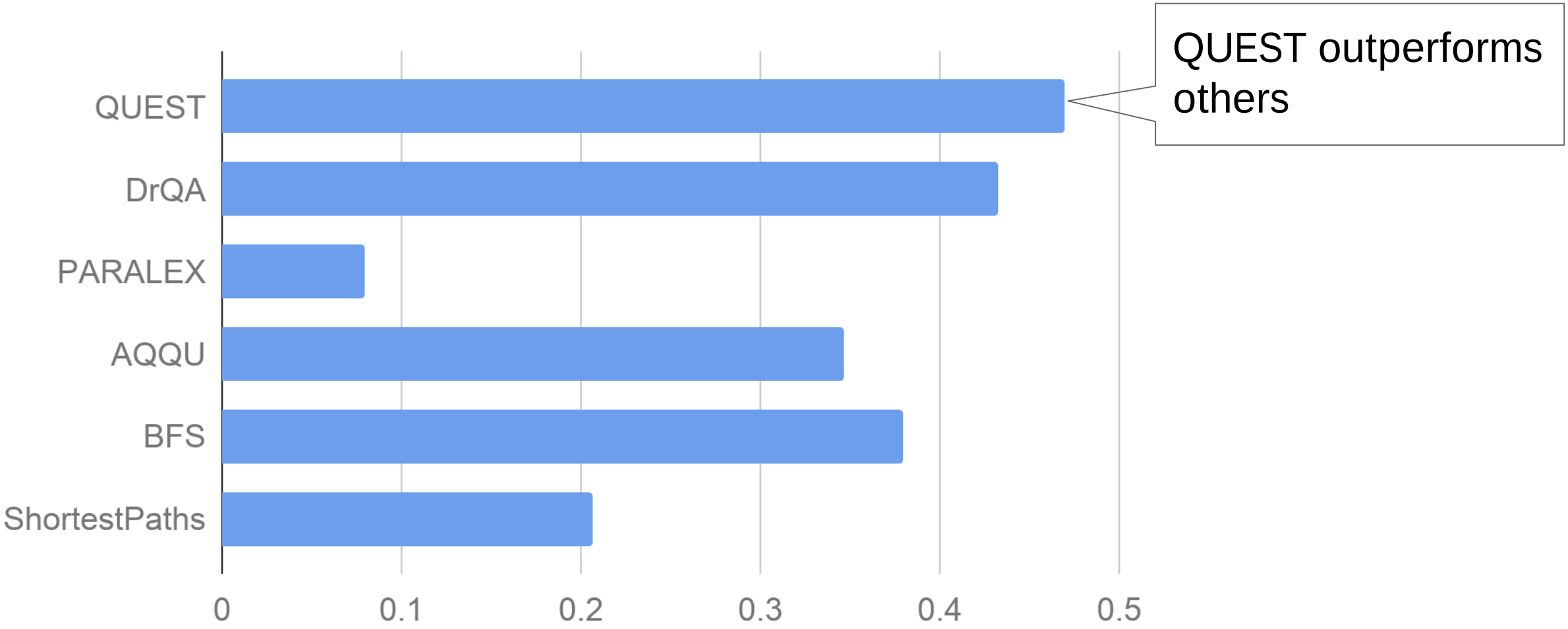


Sum of node weights	9.75
Number of GSTs	5
Sum of tree costs	6.78
Distance to cornerstones	19

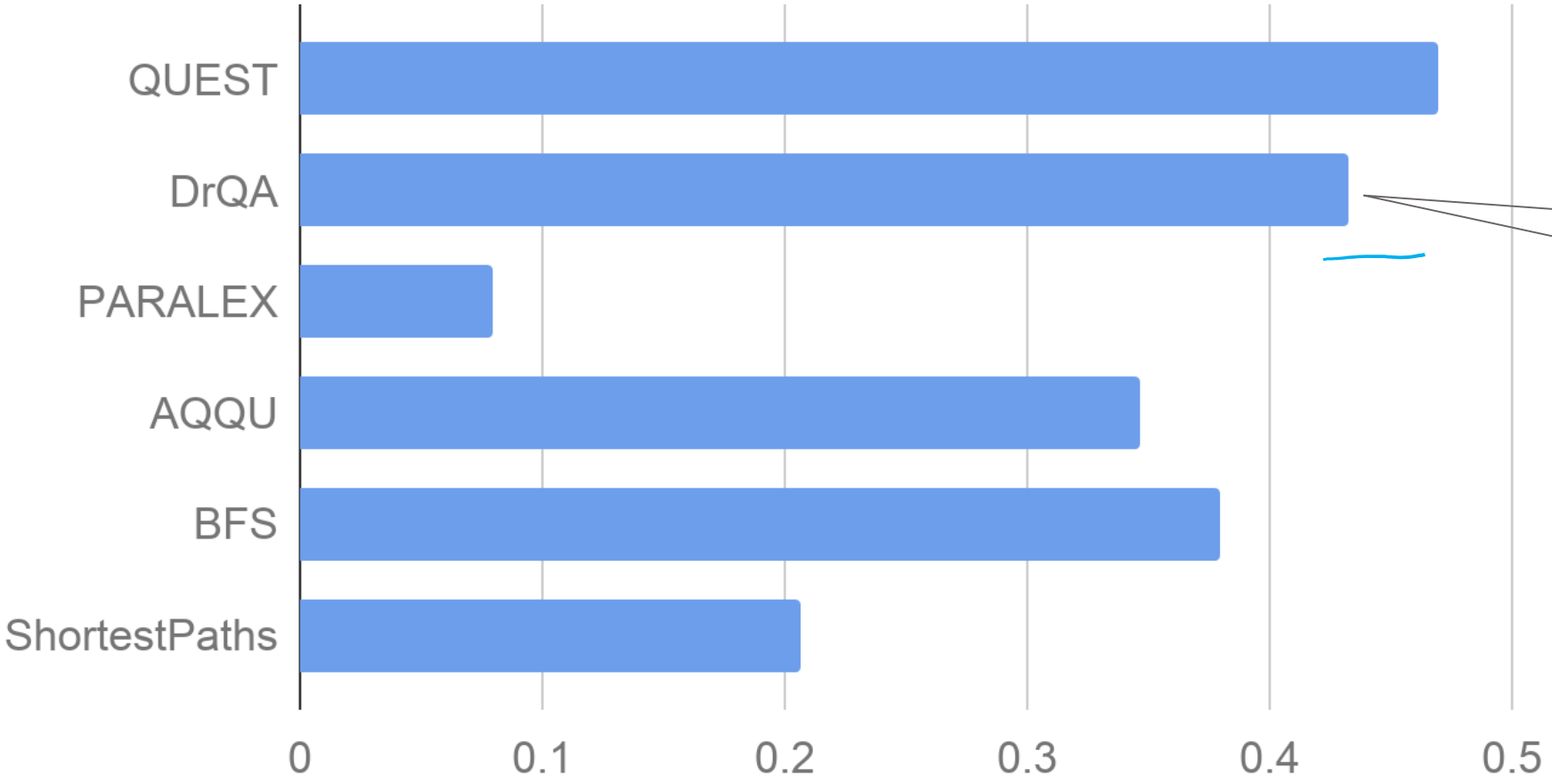
Experimental Results: Setup

- Benchmark: 300 complex questions
- Metric: Hit scores (correct answer in top-5)
- Baselines: QA algorithms and graph methods

Experimental Results: Performance

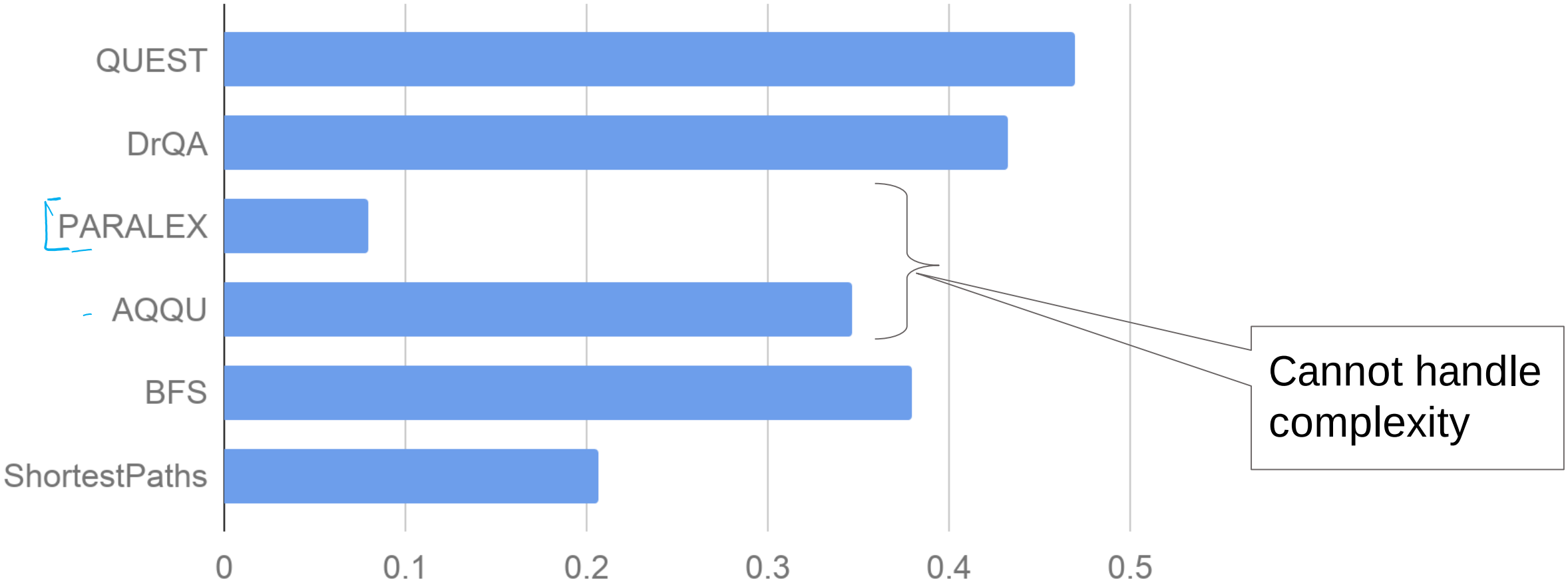


Experimental Results: Performance

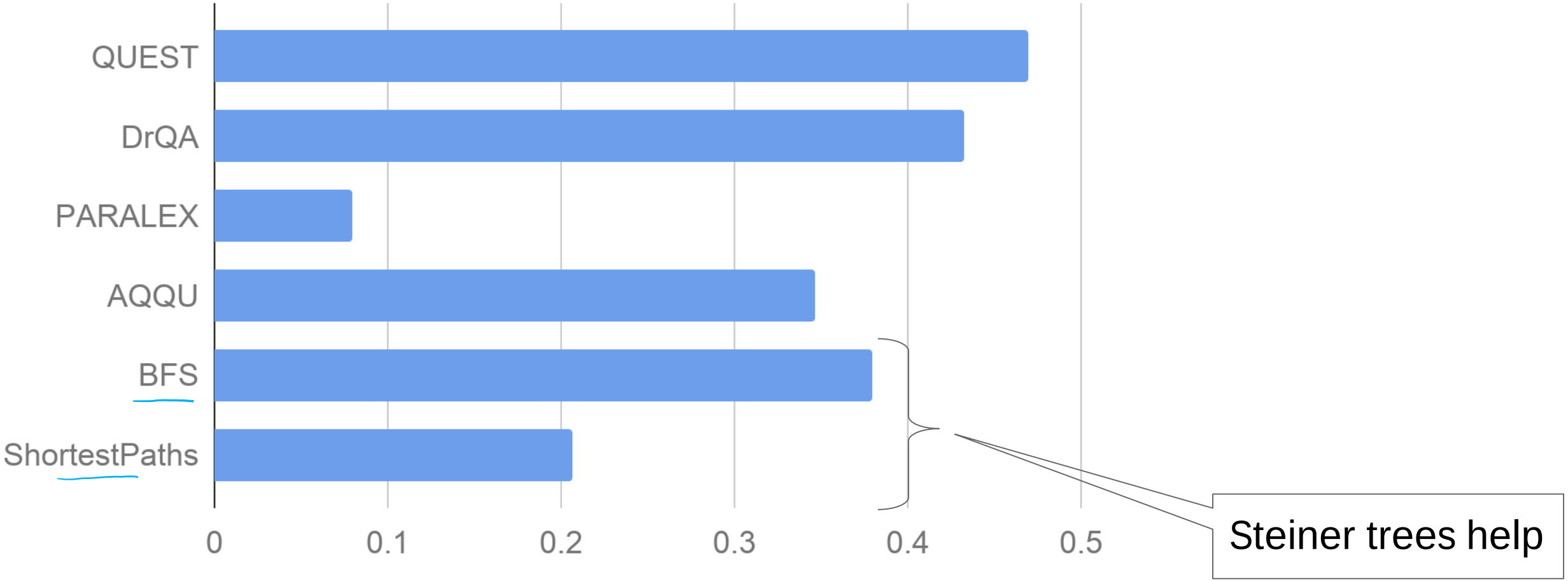


Cannot join multi-doc evidence

Experimental Results: Performance



Experimental Results: Performance

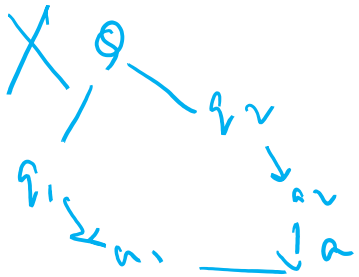


Experimental Results: Anecdotes

- 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

- Answers complex questions directly over text (+KG)
- On-the-fly joining of multi-document evidence
- Robust to high degrees of noise
- Robust to ungrammatical constructs
- No need for question decomposition



Challenges in QA

1. Diversity in question formulation 
2. Complexity in information needs 

Conclusions

- State-of-the-art techniques for QA
- Open KGs are noisy but useful
- Paraphrase-driven learning is vital
- Utility of templates goes beyond curated KGs → text + open KGs
- Graph-based methods are powerful for complex questions
- Group Steiner trees on noisy quasi-KGs enables answering complex questions

Find out more at ga.mpi-inf.mpg.de