Predicting Prerequisite Structure in Wikipedia

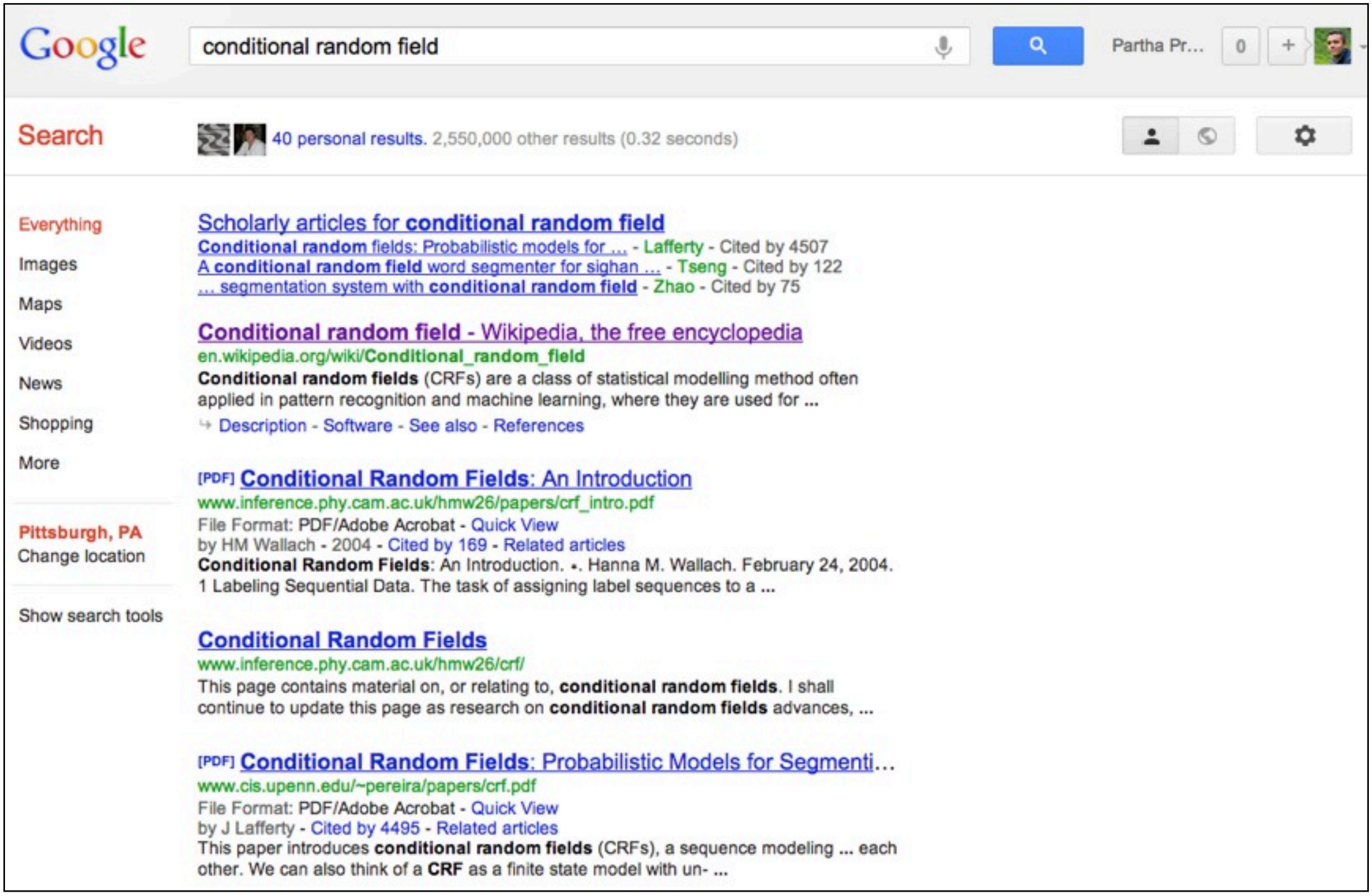
Partha Pratim Talukdar
(CMU)

William W. Cohen
(CMU)

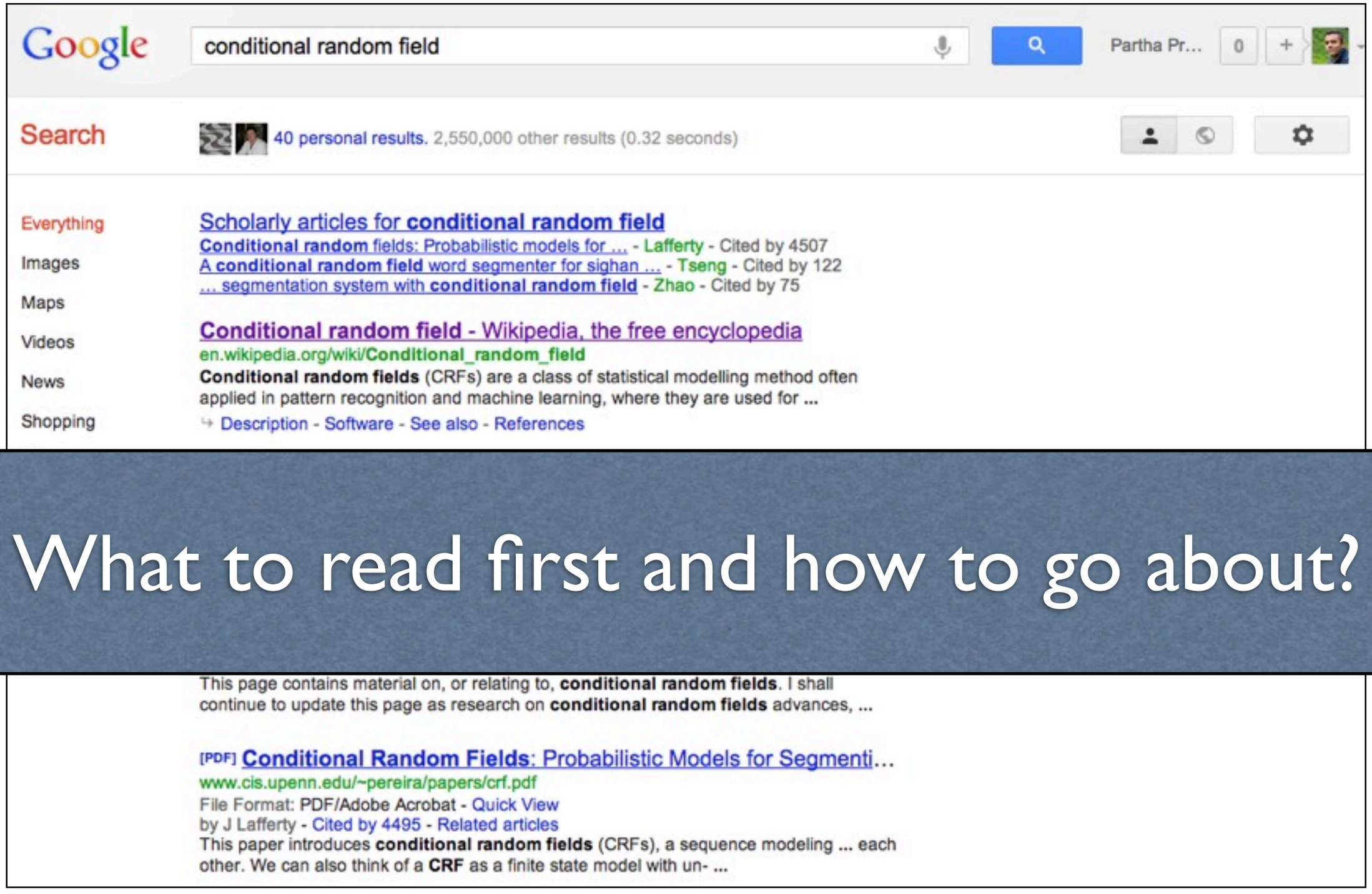
7th Workshop on Innovative Use of NLP for Building Educational Applications (BEA7)
NAACL-HLT 2012, June 7, 2012

Need to Comprehend: *Conditional Random Field*

Need to Comprehend: *Conditional Random Field*



Need to Comprehend: *Conditional Random Field*



What to read first and how to go about?

Conditional random field

From Wikipedia, the free encyclopedia

This article **provides insufficient context for those unfamiliar with the subject**. Please help improve the article with a good introductory style. *(June 2011)*

Conditional random fields (CRFs) are a class of statistical modelling method often applied in pattern recognition and machine learning, where they are used for structured prediction. Whereas an ordinary classifier predicts a label for a single sample without regard to "neighboring" samples, a CRF can take context into account; e.g., the linear chain CRF popular in natural language processing predicts sequences of labels for sequences of input samples.

CRFs are a type of discriminative undirected probabilistic graphical model. It is used to encode known relationships between observations and construct consistent interpretations. It is often used for labeling or parsing of sequential data, such as natural language text or biological sequences[1] and in computer vision.[2] Specifically, CRFs find applications in shallow parsing,[3] named entity recognition[4] and gene finding, among other tasks, being an alternative to the related hidden Markov models. In computer vision, CRFs are often used for object recognition and image segmentation.

Contents [hide]

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

Print/export

Languages
Deutsch
Español
日本語
中文

Article Talk Read

Conditional random field

From Wikipedia, the free encyclopedia

This article **provides insufficien**
subject. Please help improve the
2011)

Conditional random fields (CRFs) are a class of statistical modelling method often applied in pattern recognition and machine learning, where they are used for structured prediction. Whereas an ordinary classifier predicts a label for a single sample without regard to "neighboring" samples, a CRF can take context into account; e.g., the linear chain CRF popular in natural language processing predicts sequences of labels for sequences of input samples.

CRFs are a type of discriminative undirected probabilistic graphical model. It is used to encode known relationships between observations and construct consistent interpretations. It is often used for labeling or parsing of sequential data, such as natural language text or biological sequences[1] and in computer vision.[2] Specifically, CRFs find applications in shallow parsing,[3] named entity recognition[4] and gene finding, among other tasks, being an alternative to the related hidden Markov models. In computer vision, CRFs are often used for object recognition and image segmentation.

Contents [hide]

Visit the main page

Log in / create account

WIKIPEDIA
The Free Encyclopedia

Article Talk Read Edit View history Search

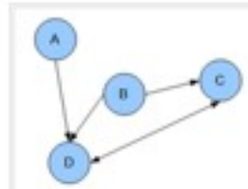
Graphical model

From Wikipedia, the free encyclopedia

A **graphical model** is a probabilistic model for which a graph denotes the conditional independence structure between random variables. They are commonly used in probability theory, statistics—particularly Bayesian statistics—and machine learning.

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Toolbox
Print/export
Languages
Español
한국어
Русский
中文

Contents [hide]



An example of a graphical model. Each arrow indicates a dependency. In this example: D depends on A, D depends on B, D depends on C, C depends on B, and C depends on D.

Types of graphical models [edit]

Generally, probabilistic graphical models use a graph-based representation as the foundation for encoding a complete distribution over a multi-dimensional space and a graph that is a compact or factorized representation of a set of independences that hold in the specific distribution. Two branches of graphical representations of distributions are commonly used, namely, Bayesian networks and Markov networks. Both families encompass the properties of factorization and independences, but they differ in the set of independences they can encode and the factorization of

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Toolbox
Print/export
Languages
Deutsch
Español
日本語
中文

Conditional random fields (CRFs) are a class of statistical modelling method often applied in pattern recognition and machine learning, where they are used for structured prediction. Whereas an ordinary classifier predicts a label for a single sample without regard to "neighboring" samples, a CRF can take context into account; e.g., the linear chain CRF popular in natural language processing predicts sequences of labels for sequences of input samples.

CRFs are a type of discriminative undirected probabilistic graphical model. It is used to encode known relationships between observations and construct consistent interpretations. It is often used for labeling or parsing of sequential data, such as natural language text or biological sequences[1] and in computer vision.[2] Specifically, CRFs find applications in shallow parsing,[3] named entity recognition[4] and gene finding, among other tasks, being an alternative to the related hidden Markov models. In computer vision, CRFs are often used for object recognition and image segmentation.

Contents [hide]

Random variable

From Wikipedia, the free encyclopedia

In probability and statistics, a **random variable** or **stochastic variable** is a variable whose value is subject to variations due to chance (i.e. randomness, in a mathematical sense). As opposed to other mathematical variables, a random variable conceptually does not have a single, fixed value (even if unknown); rather, it can take on a set of possible different values, each with an associated probability. The interpretation of a random variable depends on the interpretation of probability:

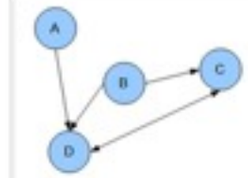
- The objectivist viewpoint: as the outcome of an experiment or event where randomness is involved (e.g. the result of rolling a die, which is a number between 1 and 6, all with equal probability; or the sum of the results of rolling two dice, which is a number between 2 and 12, with some numbers more likely than others).
- The subjectivist viewpoint: the formal encoding of one's beliefs about the various potential values of a quantity that is not known with certainty (e.g. a particular person's belief about the net worth of someone like Bill Gates after Internet research on the subject, which might have possible values ranging between about $50 billion and $100 billion, with values near the center more likely).

Graphical model

From Wikipedia, the free encyclopedia

A graphical model is a probabilistic model for which a graph denotes the conditional independence structure between random variables. They are commonly used in probability theory, statistics—particularly Bayesian statistics—and machine learning.

Contents [hide]



An example of a graphical model. Each arrow indicates a dependency. In this example: D depends on A, D depends on B, D depends on C, C depends on B, and C depends on D.

Types of graphical models [edit]

Generally, probabilistic graphical models use a graph-based representation as the foundation for encoding a complete distribution over a multi-dimensional space and a graph that is a compact or factorized representation of a set of independences that hold in the specific distribution. Two branches of graphical representations of distributions are commonly used, namely, Bayesian networks and Markov networks. Both families encompass the properties of factorization and independences, but they differ in the set of independences they can encode and the factorization of

Log in / create account

Article Talk Read Edit View history Search

Random variable

From Wikipedia, the free encyclopedia

In probability and statistics, a **random variable** or **stochastic variable** is a variable whose value is subject to variations due to chance (i.e. randomness, in a mathematical sense). As opposed to other mathematical variables, a random variable conceptually does not have a single, fixed value (even if unknown); rather, it can take on a set of possible different values, each with an associated probability. The interpretation of a random variable depends on the interpretation of probability:

- The objectivist viewpoint: as the outcome of an experiment or event where randomness is involved (e.g. the result of rolling a die, which is a number between 1 and 6, all with equal probability; or the sum of the results of rolling two dice, which is a number between 2 and 12, with some numbers more likely than others).
- The subjectivist viewpoint: the formal encoding of one's beliefs about the various potential values of a quantity that is not known with certainty (e.g. a particular person's belief about the net worth of someone like Bill Gates after Internet research on the subject, which might have possible values ranging between about $50 billion and $100 billion, with values near the center more likely).

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Log in / create account

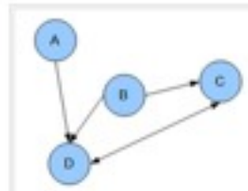
Article Talk Read Edit View history Search

Graphical model

From Wikipedia, the free encyclopedia

A graphical model is a probabilistic model for which a graph denotes the conditional independence structure between random variables. They are commonly used in probability theory, statistics—particularly Bayesian statistics—and machine learning.

Contents [hide]
1 Types of graphical models
1.1 Bayesian network
1.2 Markov random field
1.3 Other types
2 Applications
3 See also
4 Notes
5 References and further reading
5.1 Books and book chapters
5.2 Journal articles
5.3 Other



An example of a graphical model. Each arrow indicates a dependency. In this example: D depends on A, D depends on B, D depends on C, C depends on B, and C depends on D.

Types of graphical models [edit]

Generally, probabilistic graphical models use a graph-based representation as the foundation for encoding a complete distribution over a multi-dimensional space and a graph that is a compact or factorized representation of a set of independences that hold in the specific distribution. Two branches of graphical representations of distributions are commonly used, namely, Bayesian networks and Markov networks. Both families encompass the properties of factorization and independences, but they differ in the sets of independences they can encode and the factorization of

Current events
Random article
Donate to Wikipedia
Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Toolbox
Print/export
Languages
Deutsch
Español
日本語
中文

Conditional random fields (CRFs) are a class of statistical modelling method often applied in pattern recognition and machine learning, where they are used for structured prediction. Whereas an ordinary classifier predicts a label for a single sample without regard to "neighboring" samples, a CRF can take context into account; e.g., the linear chain CRF popular in natural language processing predicts sequences of labels for sequences of input samples.

CRFs are a type of discriminative undirected probabilistic graphical model. It is used to encode known relationships between observations and construct consistent interpretations. It is often used for labeling or parsing of sequential data, such as natural language text or biological sequences[1] and in computer vision.[2] Specifically, CRFs find applications in shallow parsing,[3] named entity recognition[4] and gene finding, among other tasks, being an alternative t
image segm

1 Descripti
1.1 Inf
1.2 Pa
1.3 Ex
1.4 Hig
2 Software
3 See also
4 Referenc

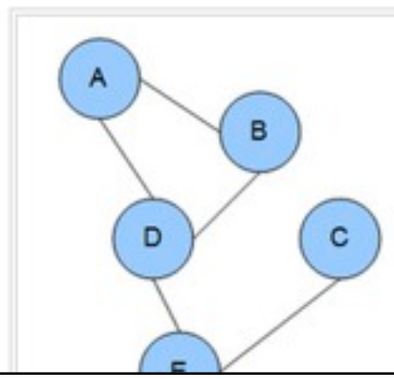
Log in / create account

Article Talk Read Edit View history Search

Markov random field

From Wikipedia, the free encyclopedia
(Redirected from Markov network)

In the domain of physics and probability, a **Markov random field** (often abbreviated as MRF), **Markov network** or **undirected graphical model** is a set of random variables having a Markov property described by an undirected graph. A Markov random field is similar to a Bayesian network in its representation of dependencies; the differences being that Bayesian networks are directed and acyclic, whereas Markov networks are undirected and may be cyclic. Thus, a Markov network can represent certain dependencies that a Bayesian network cannot (such as cyclic dependencies); on the other hand, it can't represent certain dependencies that a Bayesian network can (such as induced dependencies).



Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Interaction
Help
About Wikipedia
Community portal
Recent changes

Log in / create account

Article Talk Read Edit View history Search

Random variable

From Wikipedia, the free encyclopedia

In probability and statistics, a **random variable** or **stochastic variable** is a variable whose value is subject to variations due to chance (i.e. randomness, in a mathematical sense). As opposed to other mathematical variables, a random variable conceptually does not have a single, fixed value (even if unknown); rather, it can take on a set of possible different values, each with an associated probability. The interpretation of a random variable depends on the interpretation of probability:

- The objectivist viewpoint: as the outcome of an experiment or event where randomness is involved (e.g. the result of rolling a die, which is a number between 1 and 6, all with equal probability; or the sum of the results of rolling two dice, which is a number between 2 and 12, with some numbers more likely than others).
- The subjectivist viewpoint: the formal encoding of one's beliefs about the various potential values of a quantity that is not known with certainty (e.g. a particular person's belief about the net worth of someone like Bill Gates after Internet research on the subject, which might have possible values ranging between about $50 billion and $100 billion, with values near the center more likely).

Log in / create account

Article Talk Read Edit View history Search

Graphical model

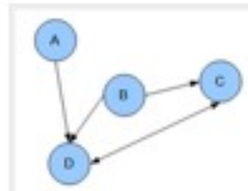
From Wikipedia, the free encyclopedia

A graphical model is a probabilistic model for which a graph denotes the conditional independence structure between random variables. They are commonly used in probability theory, statistics—particularly Bayesian statistics—and machine learning.

Contents [hide]

1 Types of graphical models
- 1.1 Bayesian network
- 1.2 Markov random field
- 1.3 Other types

2 Applications
3 See also
4 Notes
5 References and further reading
- 5.1 Books and book chapters
- 5.2 Journal articles
- 5.3 Other



An example of a graphical model. Each arrow indicates a dependency. In this example: D depends on A, D depends on B, D depends on C, C depends on B, and C depends on D.

Types of graphical models [edit]

Generally, probabilistic graphical models use a graph-based representation as the foundation for encoding a complete distribution over a multi-dimensional space and a graph that is a compact or factorized representation of a set of independences that hold in the specific distribution. Two branches of graphical representations of distributions are commonly used, namely, Bayesian networks and Markov networks. Both families encompass the properties of factorization and independences, but they differ in the sets of independences they can encode and the factorization of

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

Print/export

Languages
Deutsch
Español
日本語
中文

Conditional random fields (CRFs) are a class of statistical modelling method often applied in pattern recognition and machine learning, where they are used for structured prediction. Whereas an ordinary classifier predicts a label for a single sample without regard to "neighboring" samples, a CRF can take context into account; e.g., the linear chain CRF popular in natural language processing predicts sequences of labels for sequences of input samples.

CRFs are a type of discriminative undirected probabilistic graphical model. It is used to encode known relationships between observations and construct consistent interpretations. It is often used for labeling or parsing of sequential data, such as natural language text or biological sequences[1] and in computer vision.[2] Specifically, CRFs find applications in shallow parsing,[3] named entity recognition[4] and gene finding, among other tasks, being an alternative t
image segm

1 Descripti
1.1 Inf
1.2 Pa
1.3 Ex
1.4 Hig
2 Software
3 See also
4 Referenc

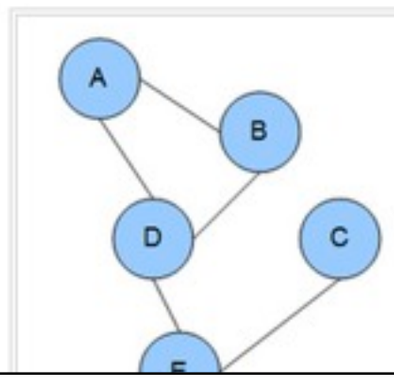
Log in / create account

Article Talk Read Edit View history Search

Markov random field

From Wikipedia, the free encyclopedia
(Redirected from Markov network)

In the domain of physics and probability, a **Markov random field** (often abbreviated as MRF), **Markov network** or **undirected graphical model** is a set of random variables having a Markov property described by an undirected graph. A Markov random field is similar to a Bayesian network in its representation of dependencies; the differences being that Bayesian networks are directed and acyclic, whereas Markov networks are undirected and may be cyclic. Thus, a Markov network can represent certain dependencies that a Bayesian network cannot (such as cyclic dependencies); on the other hand, it can't represent certain dependencies that a Bayesian network can (such as induced dependencies).



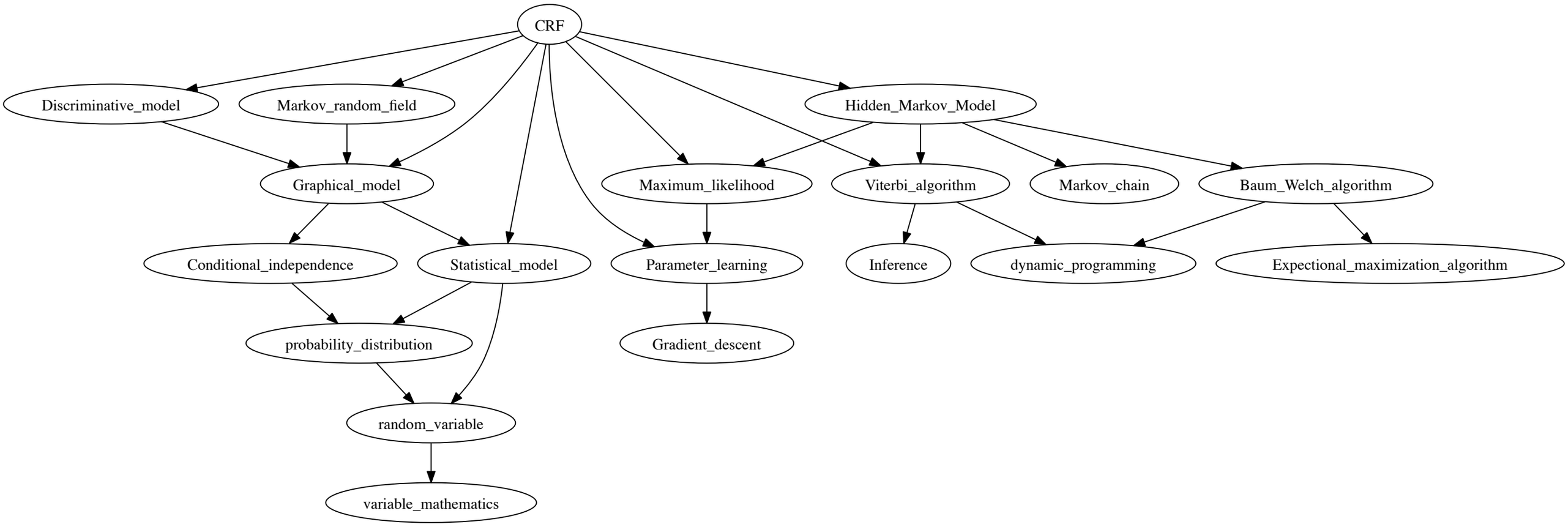
Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes

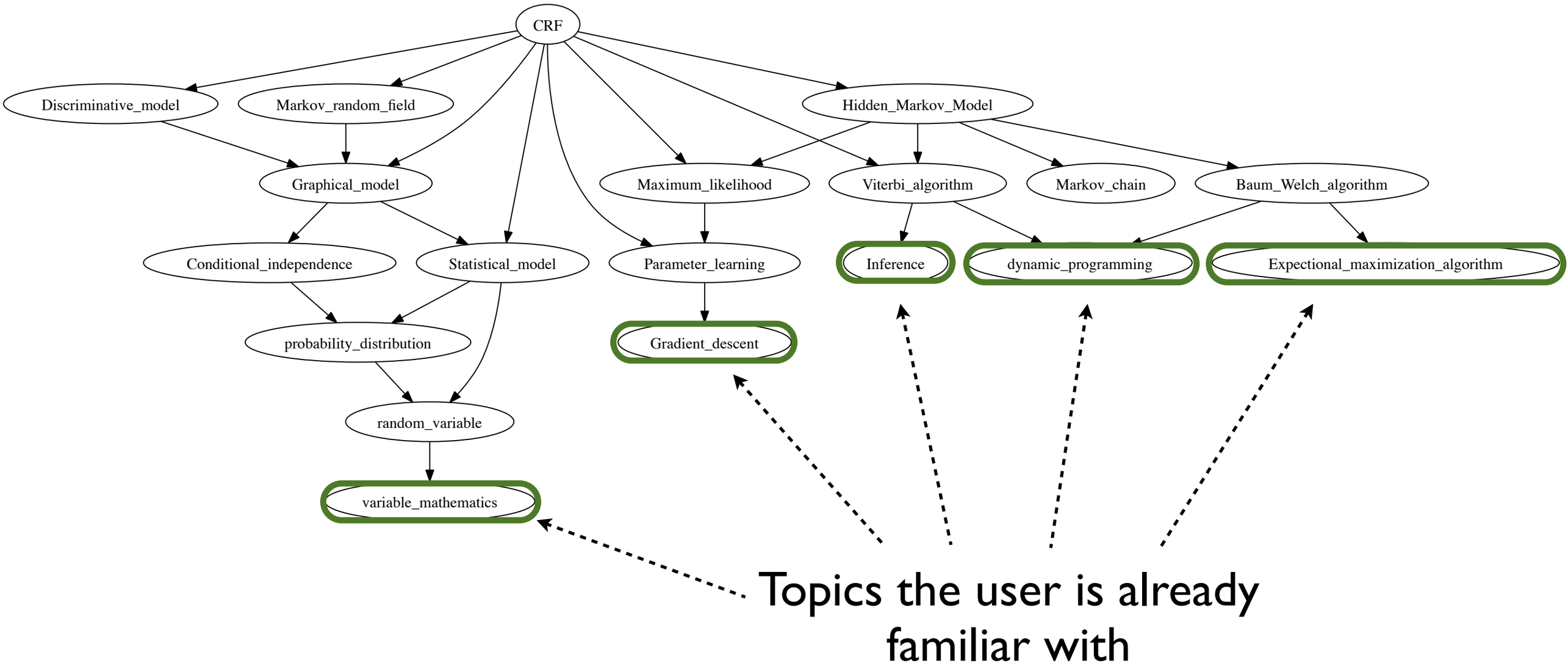
Attempting to comprehend a topic top-down can easily become difficult to manage: need better tools!

Comprehension Plan for Conditional Random Field (CRF)

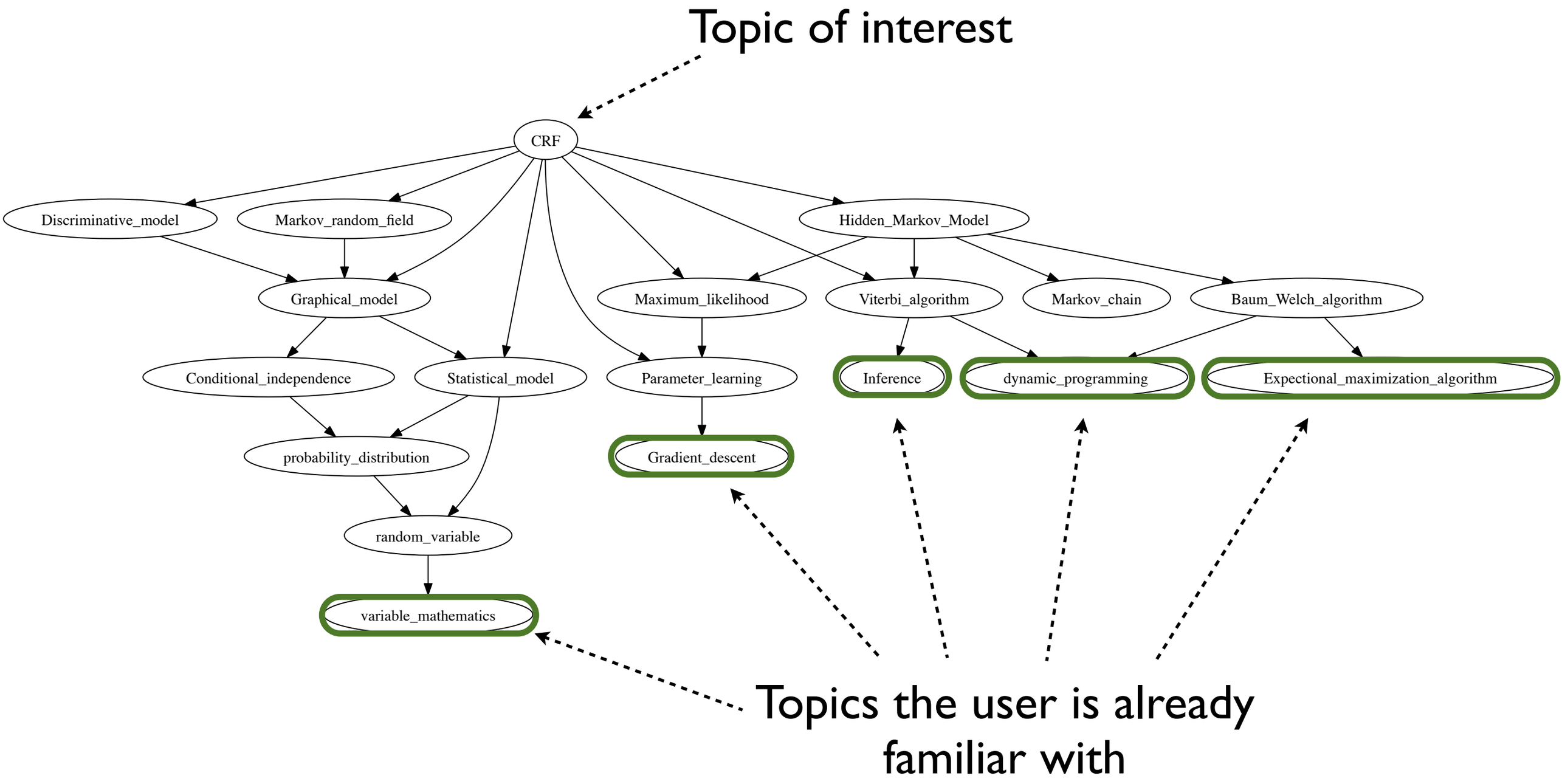
Comprehension Plan for Conditional Random Field (CRF)



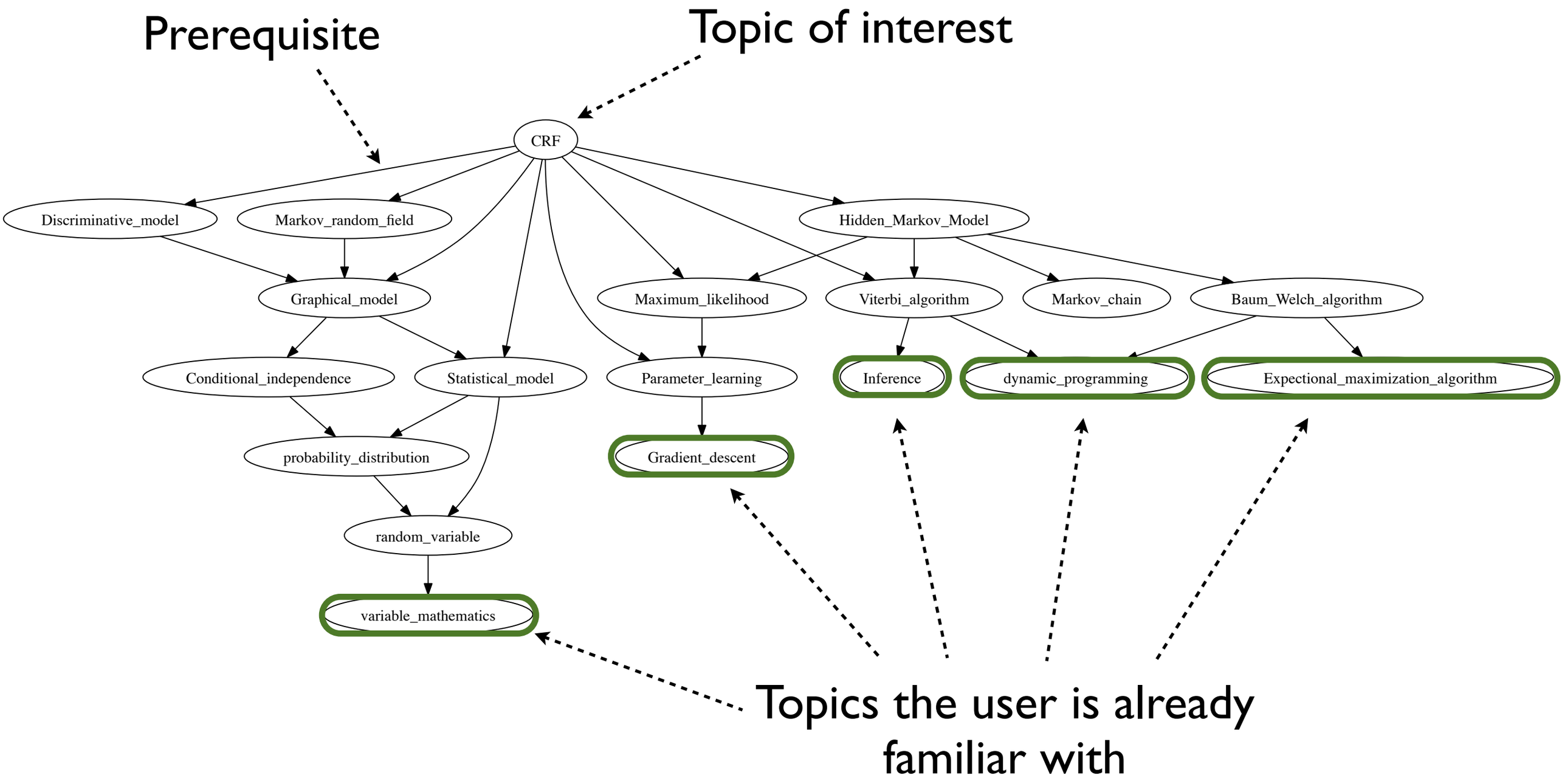
Comprehension Plan for Conditional Random Field (CRF)



Comprehension Plan for Conditional Random Field (CRF)



Comprehension Plan for Conditional Random Field (CRF)



Ingredients for a Comprehension Plan

Ingredients for a Comprehension Plan

- Estimation of user's current knowledge

Ingredients for a Comprehension Plan

- Estimation of user's current knowledge
- Determining prerequisite structure among topics

Ingredients for a Comprehension Plan

- Estimation of user's current knowledge
- Determining prerequisite structure among topics
- Plan over prerequisite structure to connect user's current knowledge with target topic

Ingredients for a Comprehension Plan

- Estimation of user's current knowledge

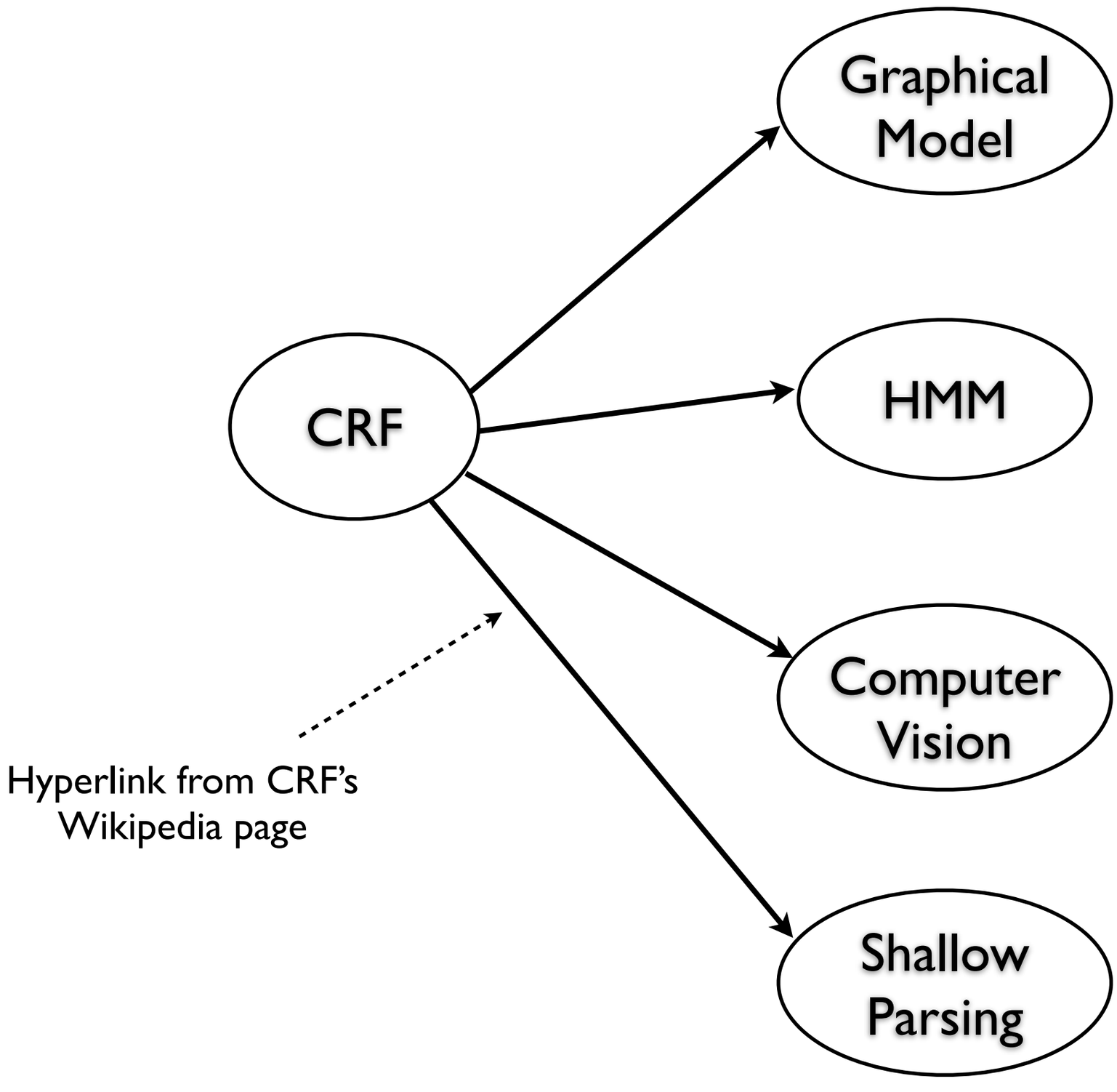
This Talk

Determining prerequisite structure among topics

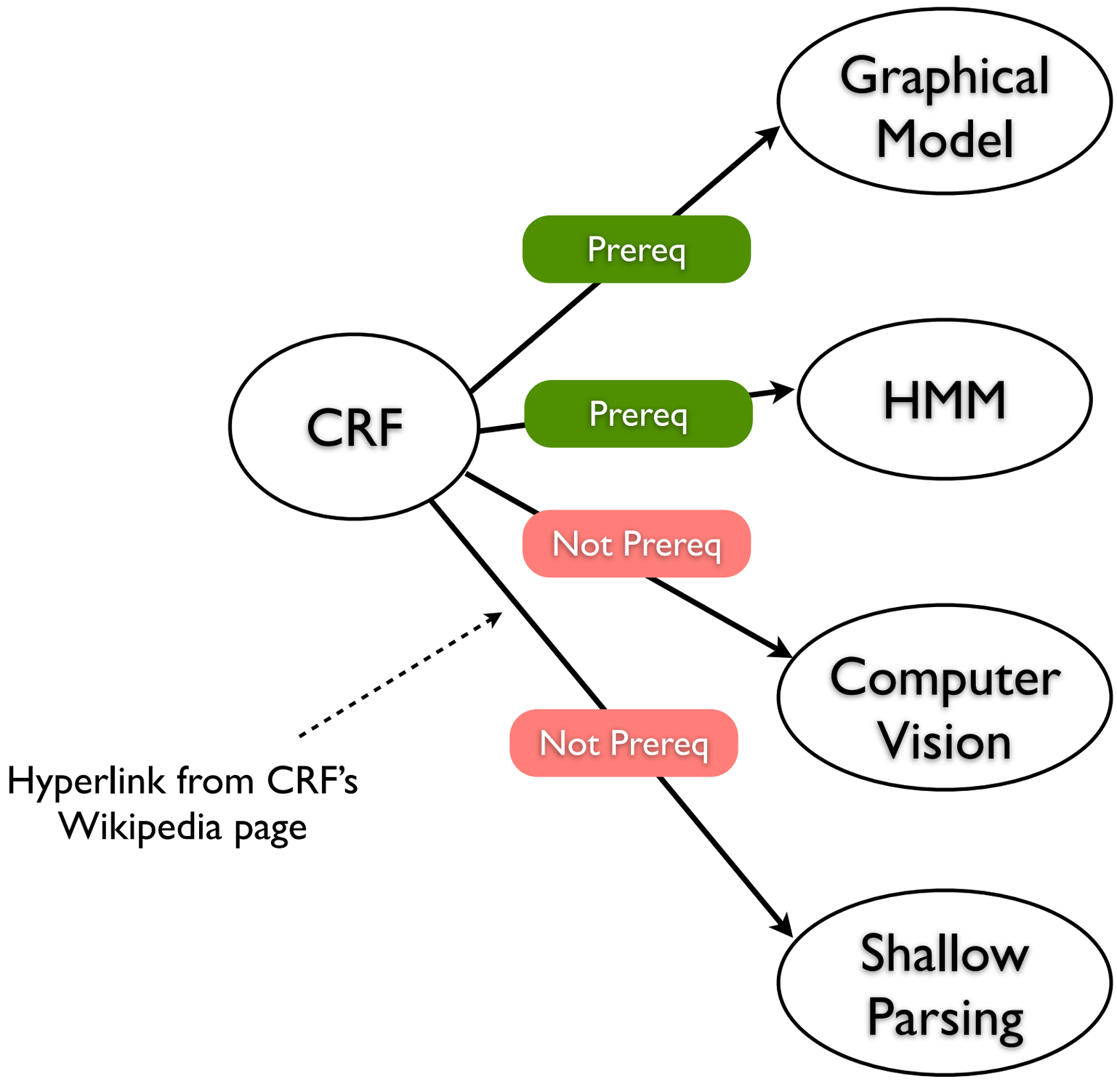
- Plan over prerequisite structure to connect user's current knowledge with target topic

Prerequisites $\subseteq$ Hyperlinks

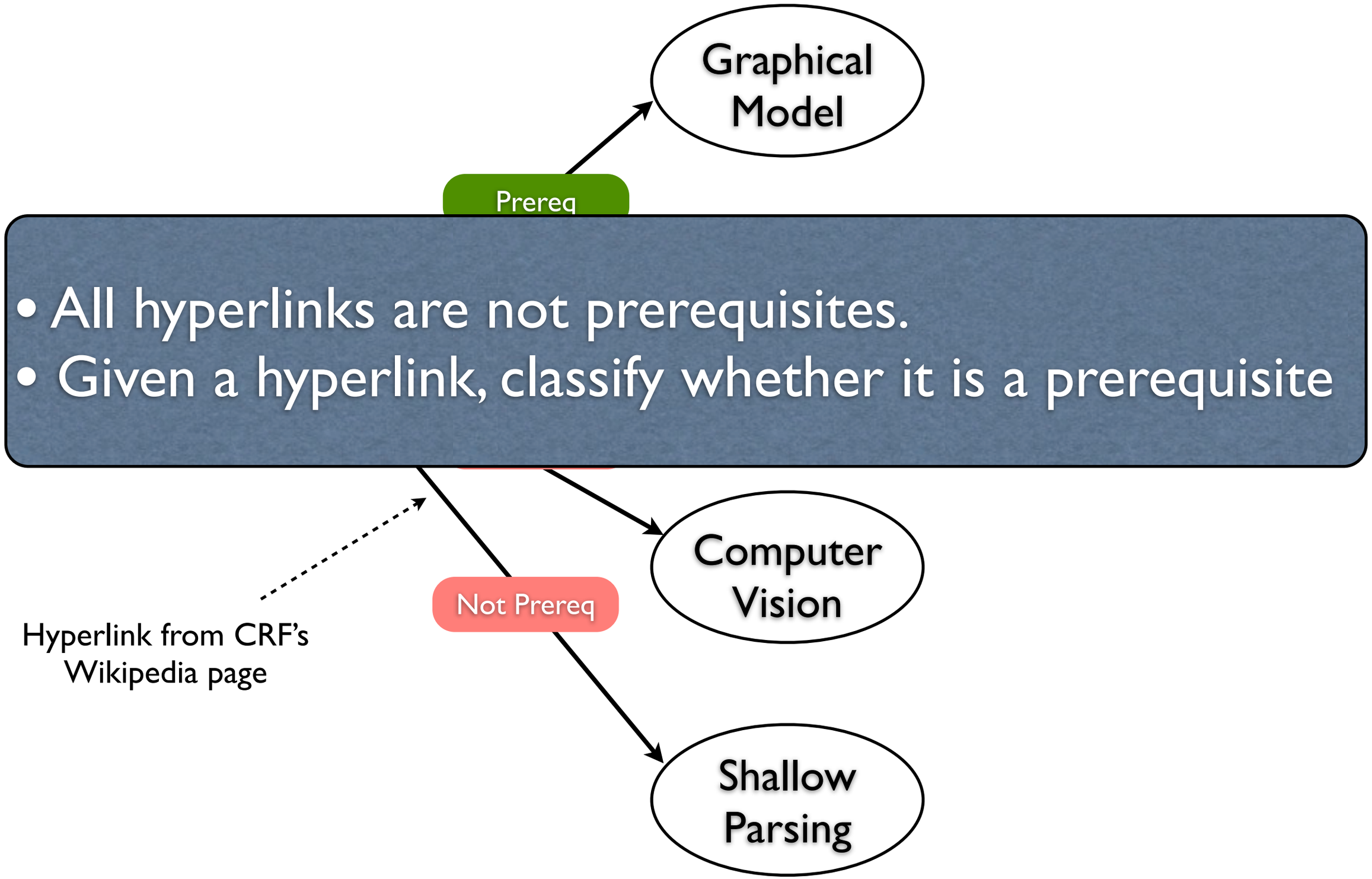
Prerequisites $\subseteq$ Hyperlinks



Prerequisites $\subseteq$ Hyperlinks



Prerequisites $\subseteq$ Hyperlinks



Outline

- Motivation
- Prerequisite Classification
  - Data Preprocessing
  - Classifier
  - Prerequisite Judgements using MTurk
  - Features
- Experiments
- Conclusion

Data Preprocessing: Target Concepts

Data Preprocessing: Target Concepts

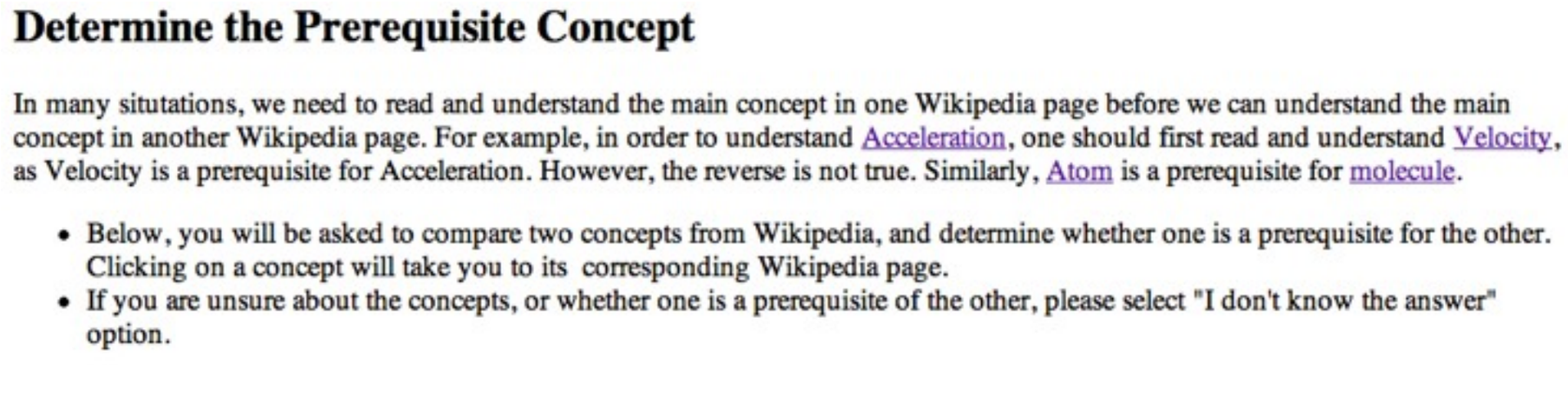
Target Concept	#Nodes	#Edges	#Edits
Global Warming	19,170	501,608	1,490,967
Meiosis	19,811	444,100	880,684
Newton's Laws of Motion	15,714	436,035	795,988
Parallel Postulate	14,966	363,462	858,785
Public-key cryptography	16,695	371,104	1,003,181

Table 1: Target concepts used in the experiments.

Sample up to 400 edges from each subgraph

Gold Prerequisite Judgements using Mechanical Turk

Gold Prerequisite Judgements using Mechanical Turk

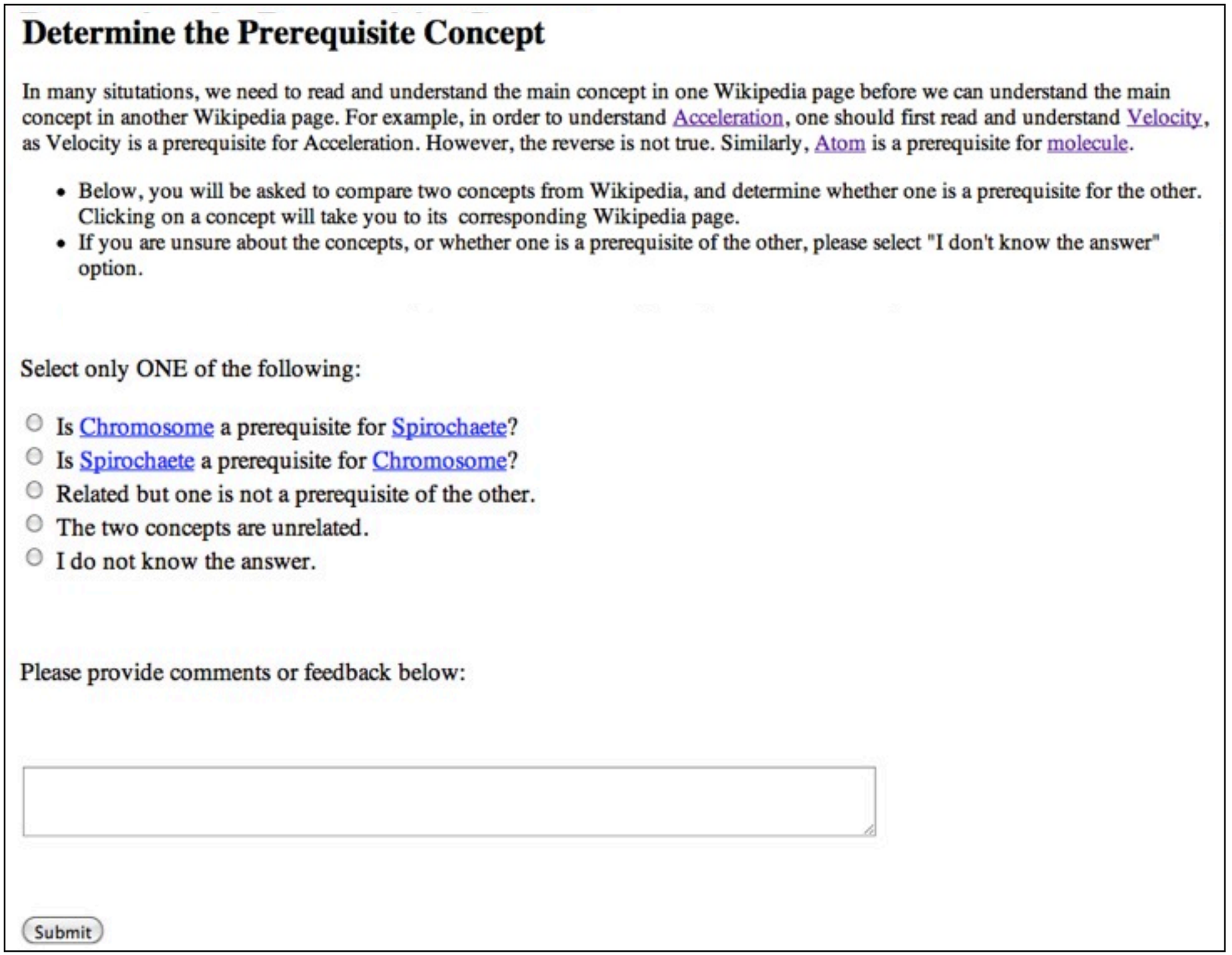


Determine the Prerequisite Concept

In many situtations, we need to read and understand the main concept in one Wikipedia page before we can understand the main concept in another Wikipedia page. For example, in order to understand Acceleration, one should first read and understand Velocity, as Velocity is a prerequisite for Acceleration. However, the reverse is not true. Similarly, Atom is a prerequisite for molecule.

- Below, you will be asked to compare two concepts from Wikipedia, and determine whether one is a prerequisite for the other. Clicking on a concept will take you to its corresponding Wikipedia page.
- If you are unsure about the concepts, or whether one is a prerequisite of the other, please select "I don't know the answer" option.

Gold Prerequisite Judgements using Mechanical Turk



Determine the Prerequisite Concept

In many situtations, we need to read and understand the main concept in one Wikipedia page before we can understand the main concept in another Wikipedia page. For example, in order to understand Acceleration, one should first read and understand Velocity, as Velocity is a prerequisite for Acceleration. However, the reverse is not true. Similarly, Atom is a prerequisite for molecule.

- Below, you will be asked to compare two concepts from Wikipedia, and determine whether one is a prerequisite for the other. Clicking on a concept will take you to its corresponding Wikipedia page.
- If you are unsure about the concepts, or whether one is a prerequisite of the other, please select "I don't know the answer" option.

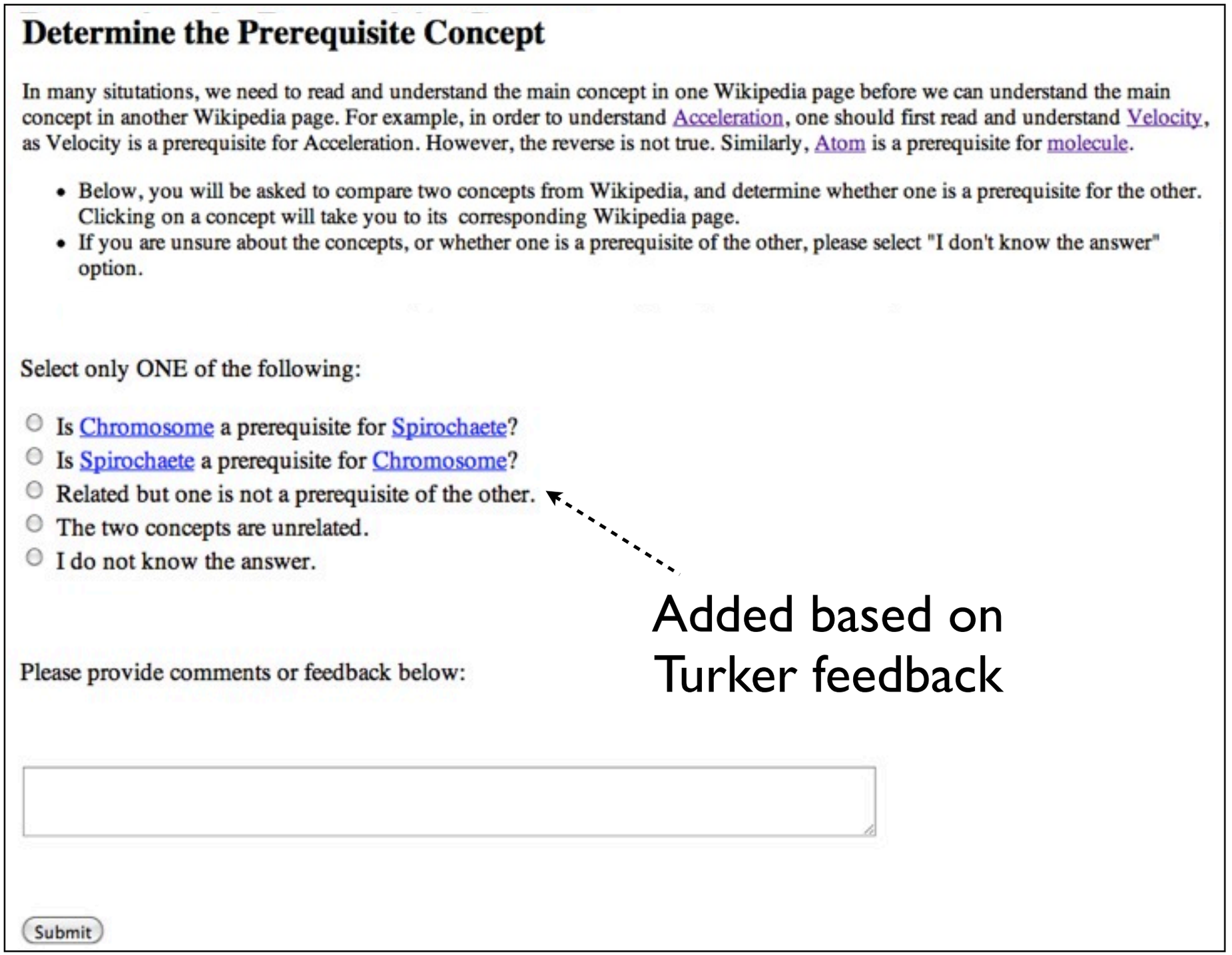
Select only ONE of the following:

- ○ Is Chromosome a prerequisite for Spirochaete?
- ○ Is Spirochaete a prerequisite for Chromosome?
- ○ Related but one is not a prerequisite of the other.
- ○ The two concepts are unrelated.
- ○ I do not know the answer.

Please provide comments or feedback below:

Submit

Gold Prerequisite Judgements using Mechanical Turk



Determine the Prerequisite Concept

In many situtations, we need to read and understand the main concept in one Wikipedia page before we can understand the main concept in another Wikipedia page. For example, in order to understand Acceleration, one should first read and understand Velocity, as Velocity is a prerequisite for Acceleration. However, the reverse is not true. Similarly, Atom is a prerequisite for molecule.

- Below, you will be asked to compare two concepts from Wikipedia, and determine whether one is a prerequisite for the other. Clicking on a concept will take you to its corresponding Wikipedia page.
- If you are unsure about the concepts, or whether one is a prerequisite of the other, please select "I don't know the answer" option.

Select only ONE of the following:

- Is Chromosome a prerequisite for Spirochaete?
- Is Spirochaete a prerequisite for Chromosome?
- Related but one is not a prerequisite of the other.
- The two concepts are unrelated.
- I do not know the answer.

Added based on Turker feedback

Please provide comments or feedback below:

Submit

Agreement among Turkers

Agreement among Turkers

Domain	Time (s) / Evaluation	Worker / HIT	# HITs	κ
Meiosis	38	3	400	0.50
Public-key Cryp.	26	3	200	0.63
Parallel Postulate	41	3	200	0.55
Newton's Laws	20	5	400	0.47
Global Warming	14	5	400	0.56
Average	27.8	-	-	0.54

Table 2: Statistics about the Gold-standard data prepared using Amazon Mechanical Turk. Also shown are the averaged κ statistics-based inter-annotator agreement in each domain. The last row corresponds to the κ value averaged across all five domains.

Agreement among Turkers

Domain	Time (s) / Evaluation	Worker / HIT	# HITs	κ
Meiosis	38	3	400	0.50
Public-key Cryp.	26	3	200	0.63
Parallel Postulate	41	3	200	0.55
Newton's Laws	20	5	400	0.47
Global Warming	14	5	400	0.56
Average	27.8	-	-	0.54

Table 2: Statistics about the Gold-standard data prepared using Amazon Mechanical Turk. Also shown are the averaged κ statistics-based inter
each domain. The last row cor
averaged across all five domains.

Total cost: $278, completed in a week.

Classifier: Maximum Entropy

$$p(y|x) = \frac{\exp(\mathbf{w} \cdot \phi(x, y))}{\sum_{y' \in Y} \exp(\mathbf{w} \cdot \phi(x, y'))}, \; y \in Y = \{-1, +1\}$$

Classifier: Maximum Entropy

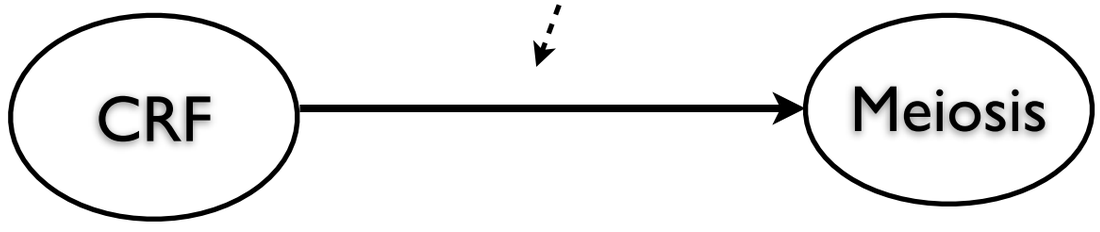
$$p(y|x) = \frac{\exp(\mathbf{w} \cdot \phi(x, y))}{\sum_{y' \in Y} \exp(\mathbf{w} \cdot \phi(x, y'))}, \; y \in Y = \{-1, +1\}$$

Prerequisite

Classifier: Maximum Entropy

$$p(y|x) = \frac{\exp(\mathbf{w} \cdot \phi(x, y))}{\sum_{y' \in Y} \exp(\mathbf{w} \cdot \phi(x, y'))}, y \in Y = \{-1, +1\}$$

Prerequisite

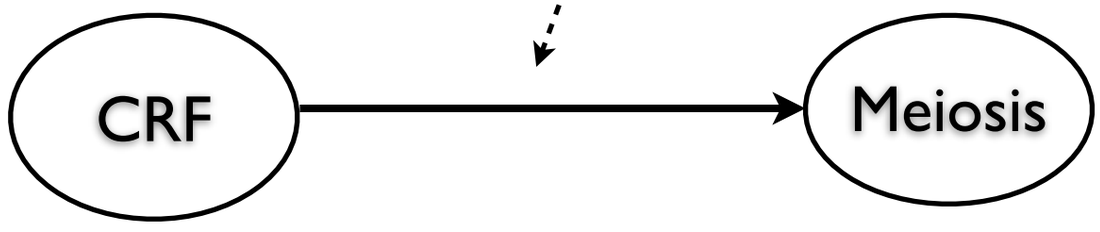


Classifier: Maximum Entropy

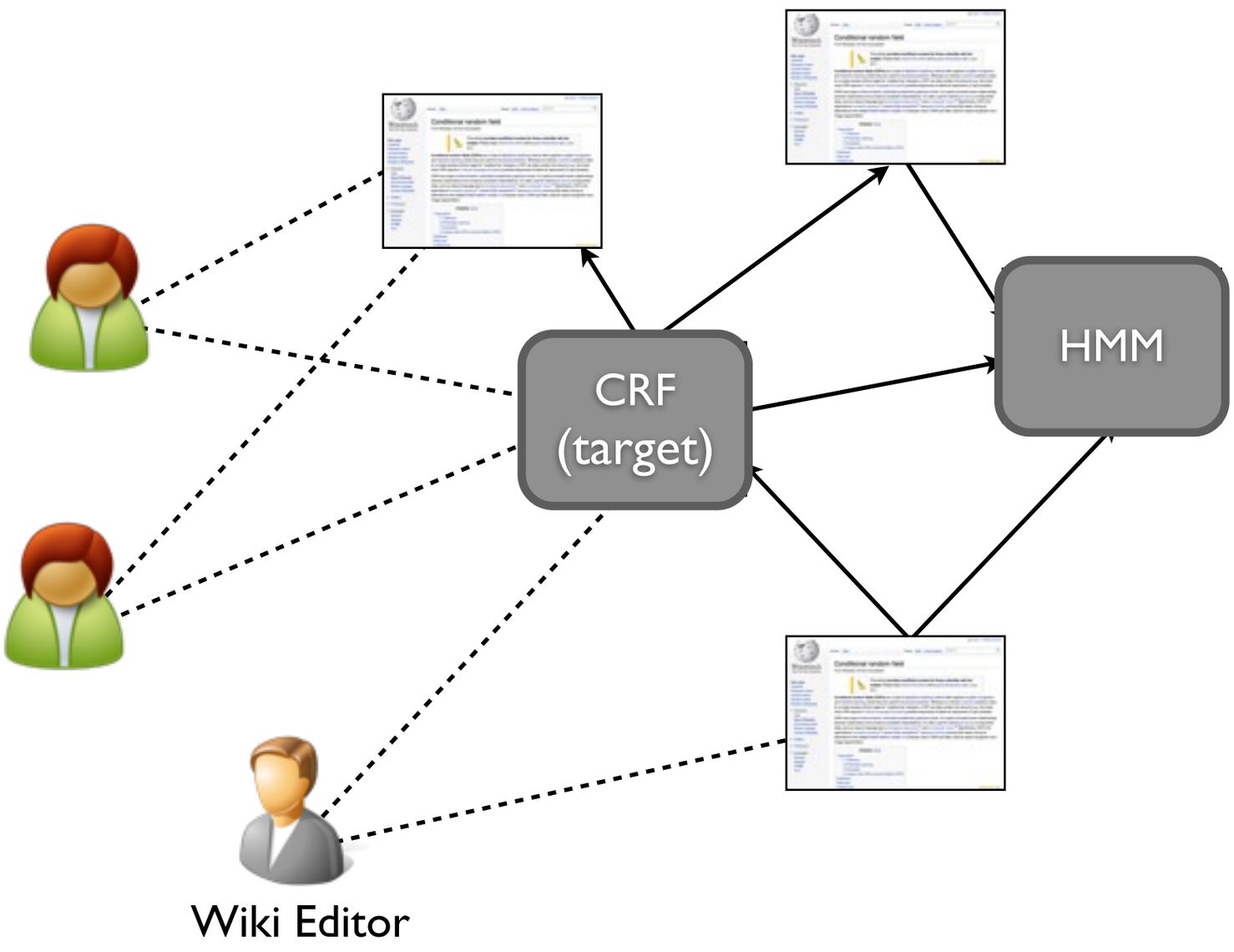
Feature Function

$$p(y|x) = \frac{\exp(\mathbf{w} \cdot \phi(x, y))}{\sum_{y' \in Y} \exp(\mathbf{w} \cdot \phi(x, y'))}, y \in Y = \{-1, +1\}$$

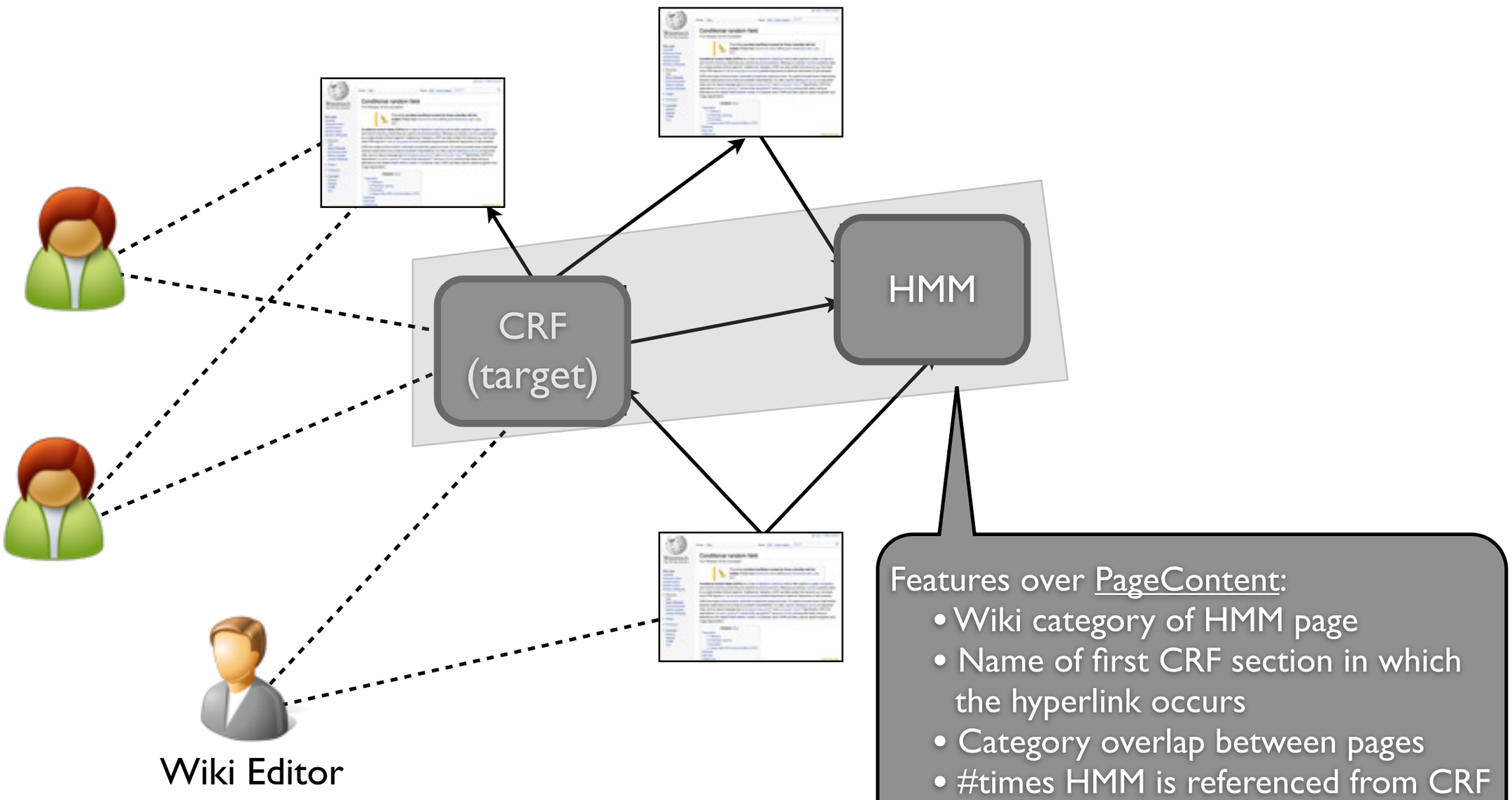
Prerequisite



Features

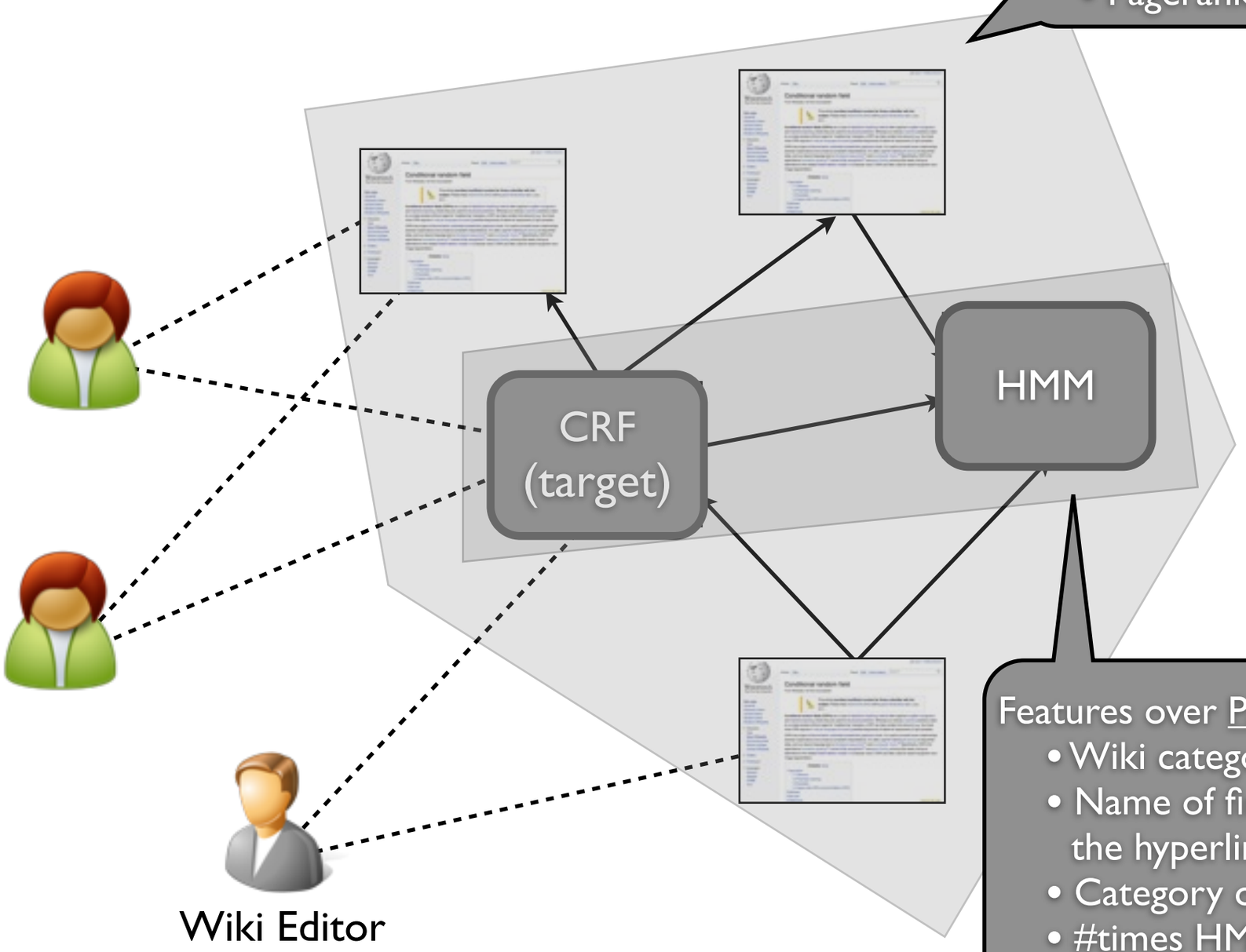


Features


CRF
(target)
HMM
Wiki Editor
Features over PageContent:
• Wiki category of HMM page
• Name of first CRF section in which the hyperlink occurs
• Category overlap between pages
• #times HMM is referenced from CRF
...

Features

Features over Hyperlink Graph:
- Random Walk with Restart (RWR) score of a topic
- Pagerank score of source
- Pagerank score of target



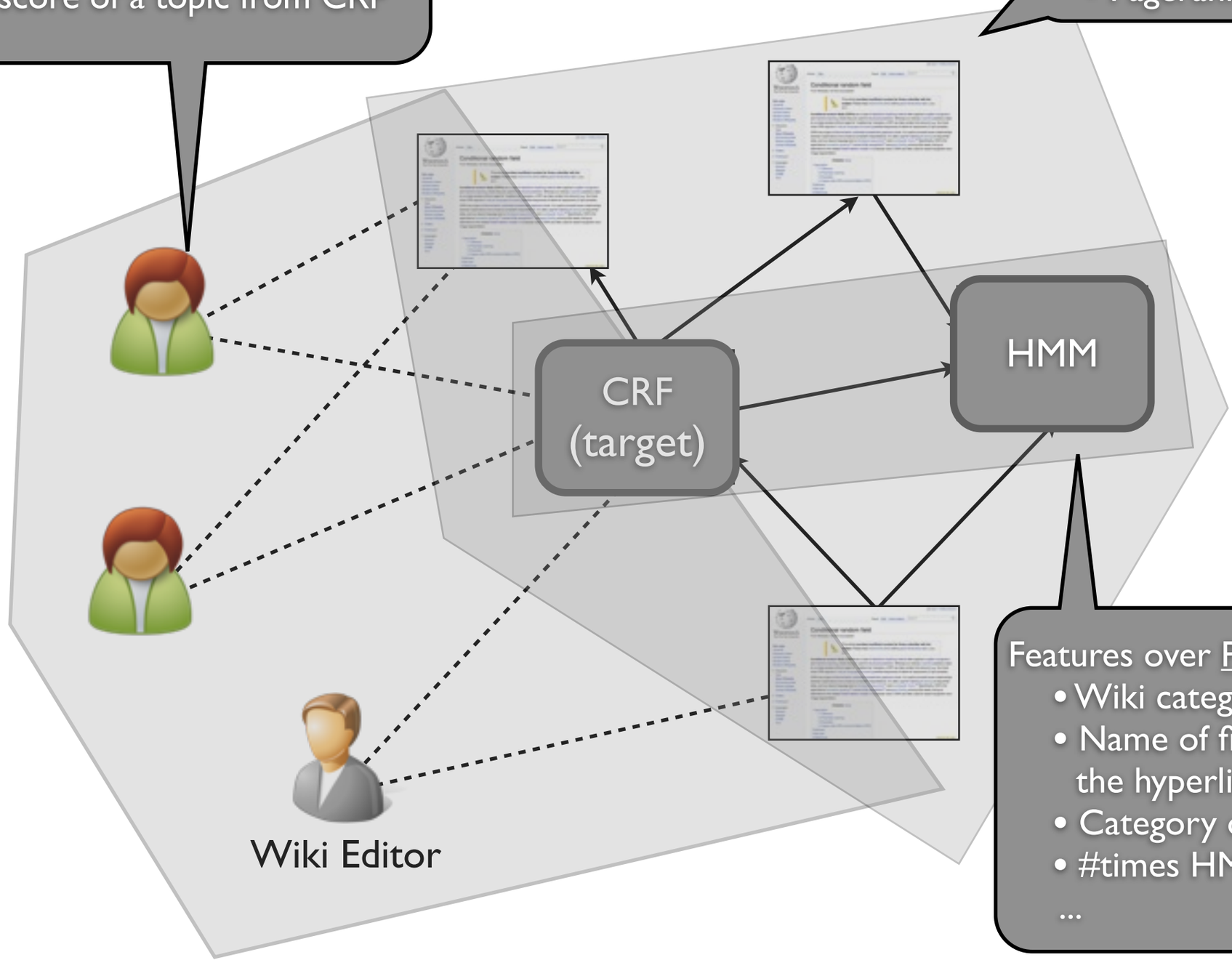
Features over PageContent:
- Wiki category of HMM page
- Name of first CRF section in which the hyperlink occurs
- Category overlap between pages
- #times HMM is referenced from CRF

...

Features

Feature over WikiEdits Graph:
- RWR score of a topic from CRF

Features over Hyperlink Graph:
- Random Walk with Restart (RWR) score of a topic
- Pagerank score of source
- Pagerank score of target



Features over PageContent:
- Wiki category of HMM page
- Name of first CRF section in which the hyperlink occurs
- Category overlap between pages
- #times HMM is referenced from CRF

...

Experiments

Experiments

- Can we train a classifier for prerequisite classification?

Experiments

- Can we train a classifier for prerequisite classification?
- How effective is out-of-domain training?

Experiments

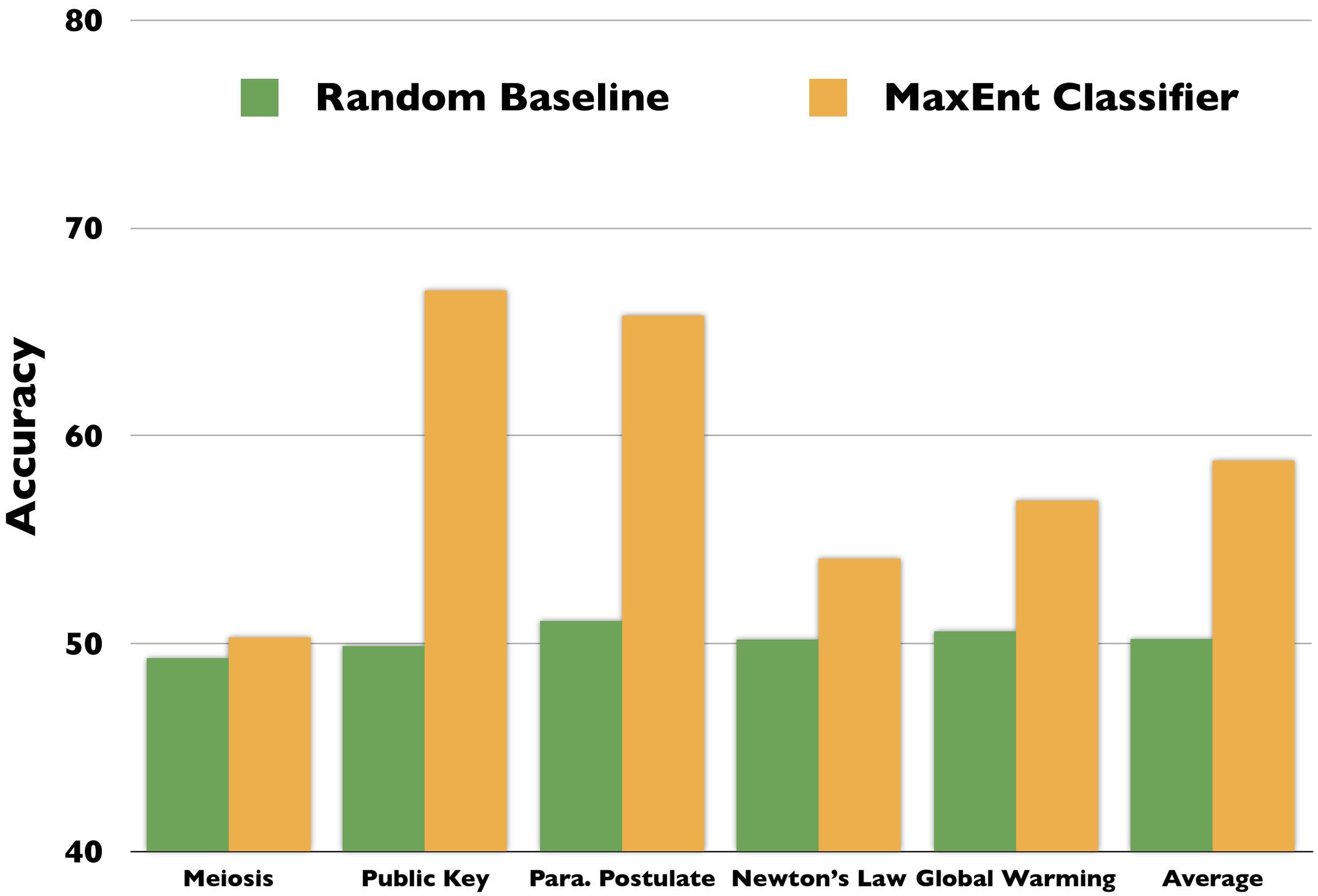
- Can we train a classifier for prerequisite classification?
- How effective is out-of-domain training?
- What are the effects of different features?

Experiments

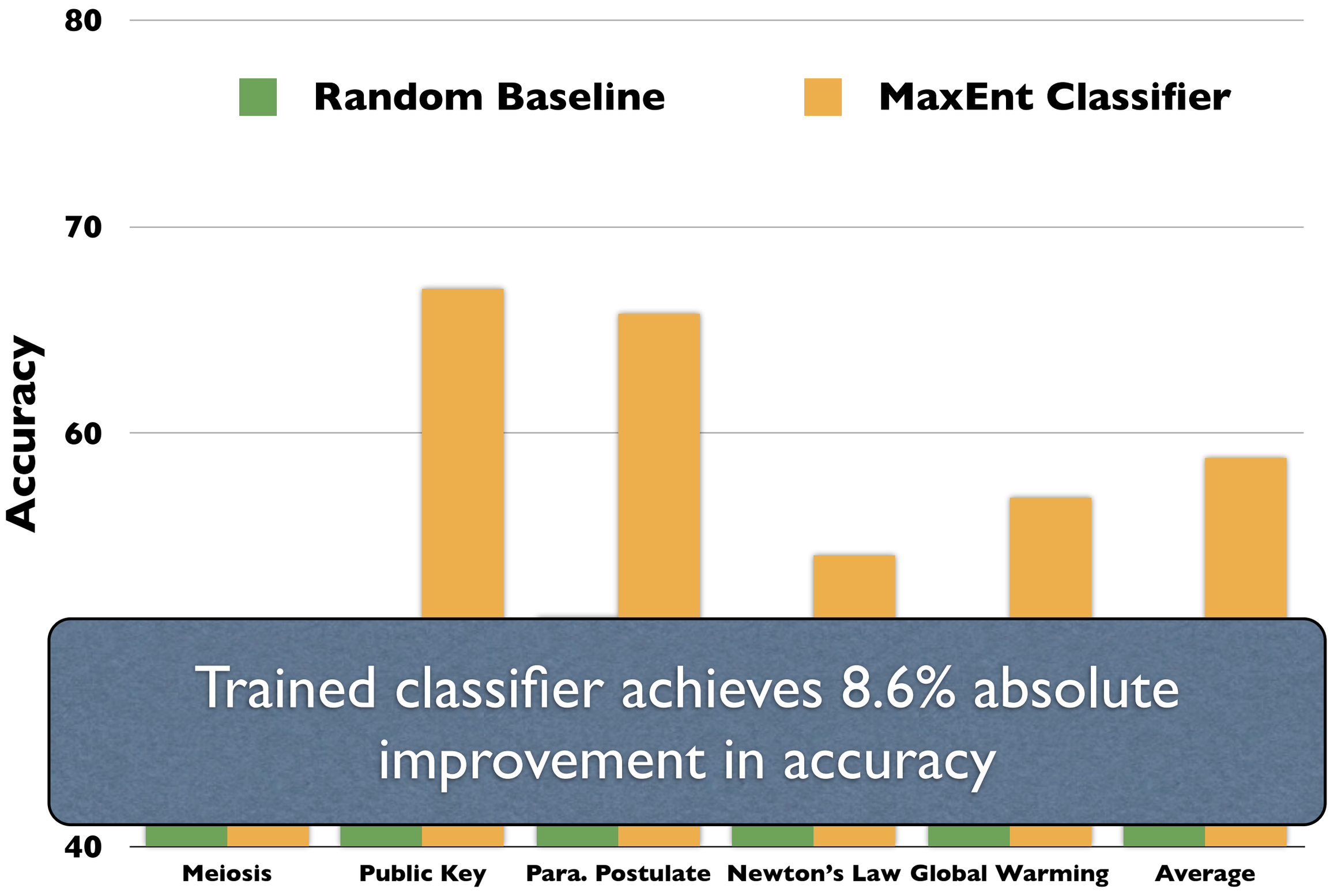
- Can we train a classifier for prerequisite classification?
- How effective is out-of-domain training?
- What are the effects of different features?

All evaluations in leave-one-target-out setting

Performance Comparison for Prerequisite Prediction

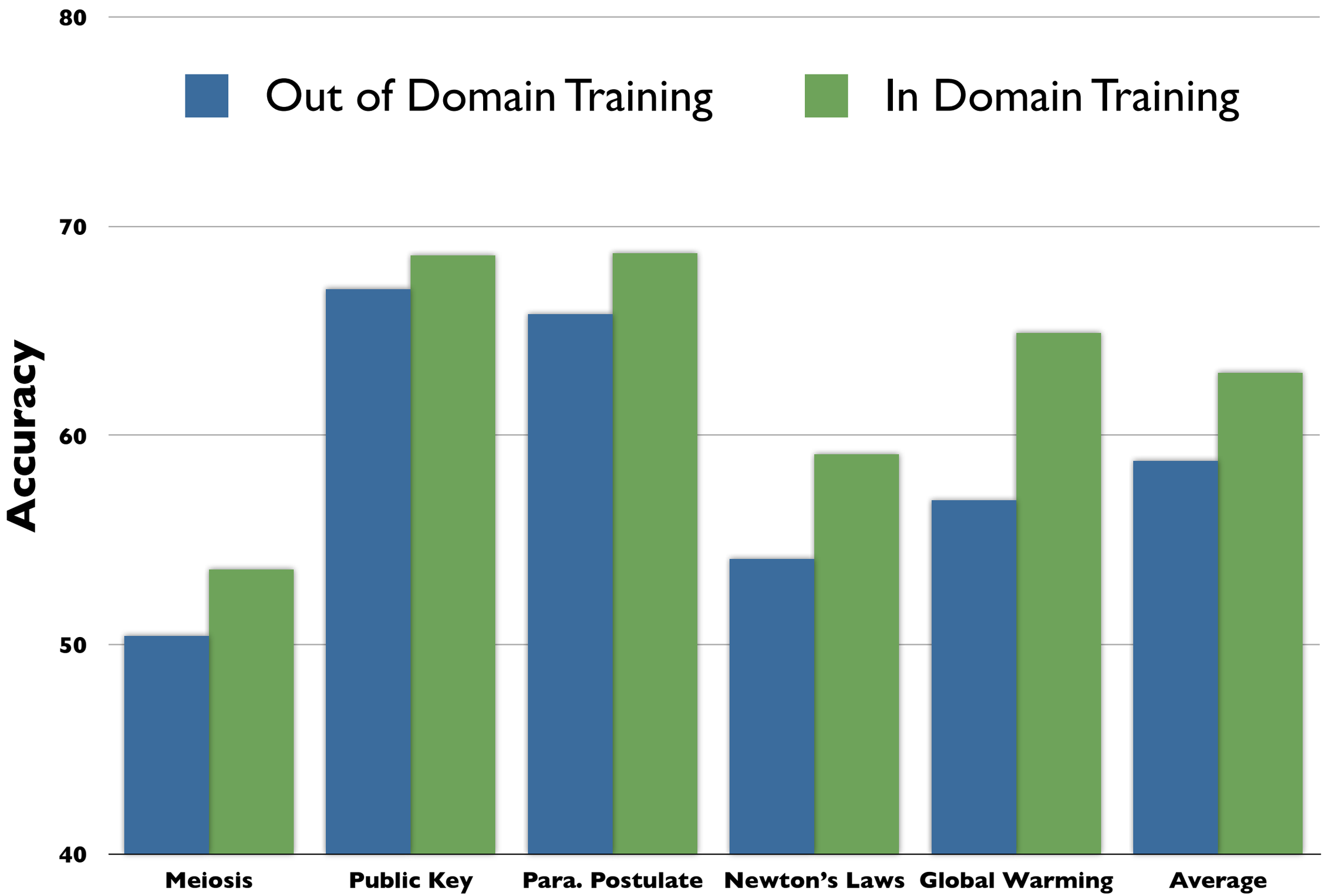


Performance Comparison for Prerequisite Prediction

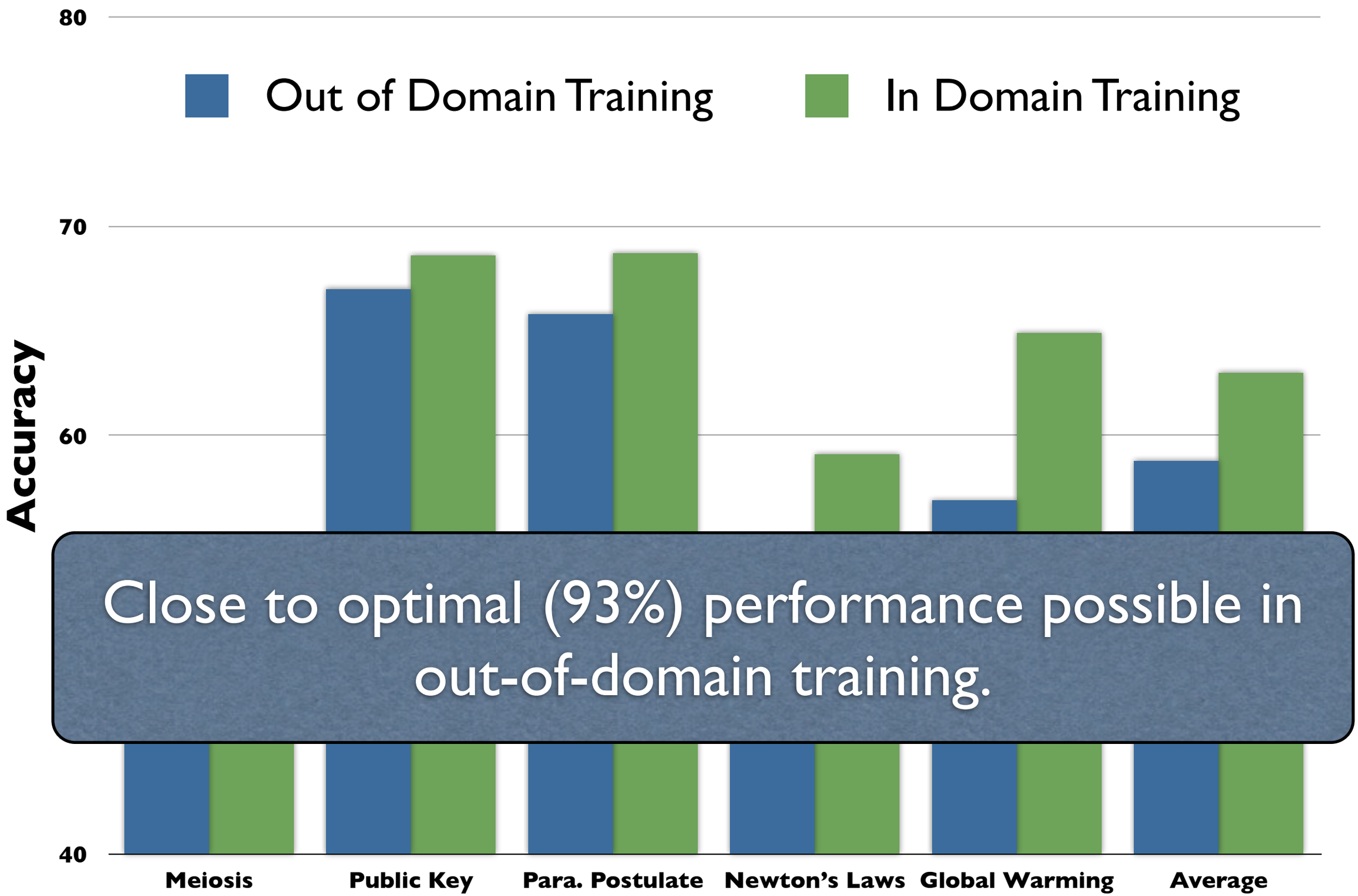


Trained classifier achieves 8.6% absolute improvement in accuracy

Effect of Out of Domain vs In Domain Training



Effect of Out of Domain vs In Domain Training

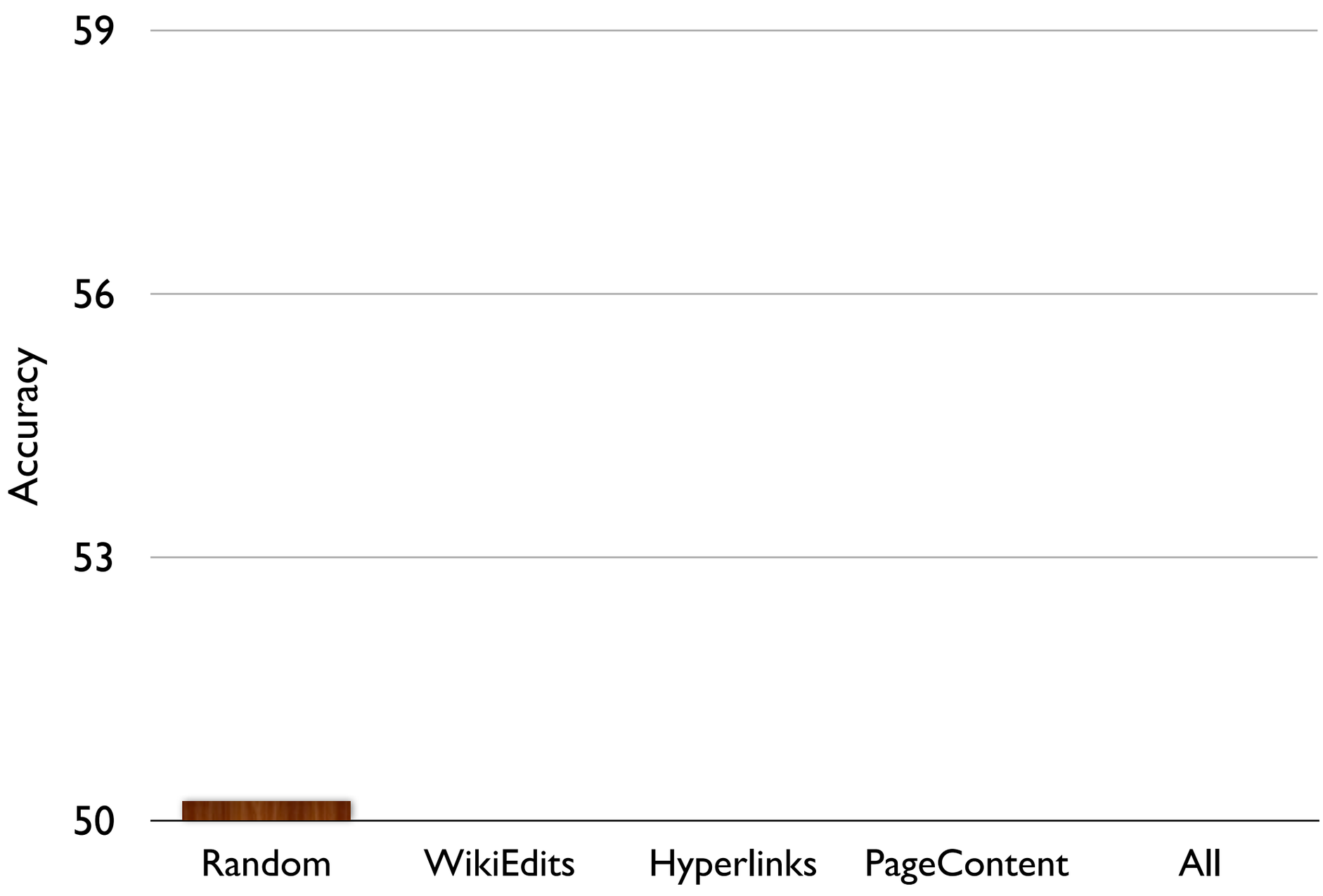


Close to optimal (93%) performance possible in out-of-domain training.

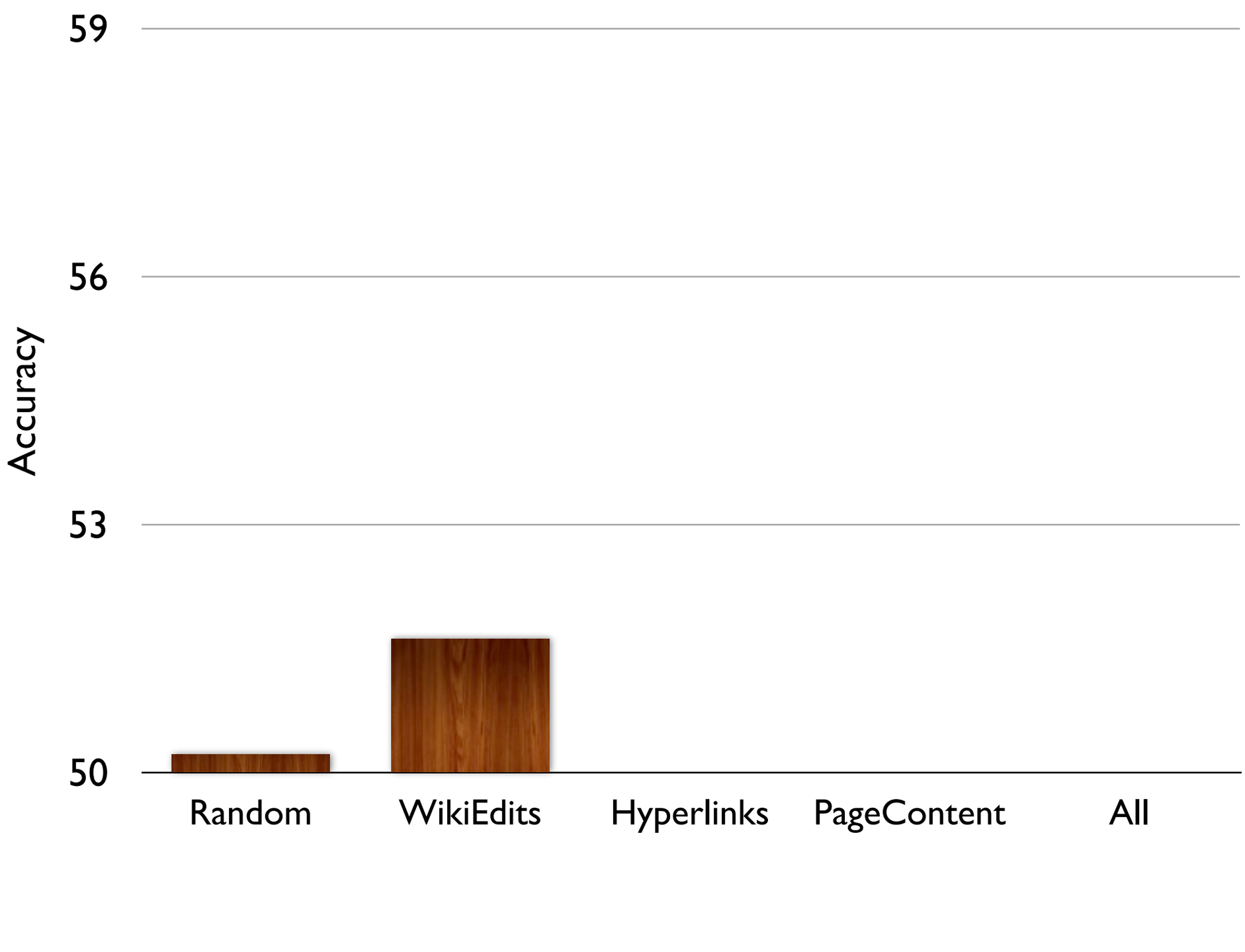
Effect of Different Features

Effect of Different Features

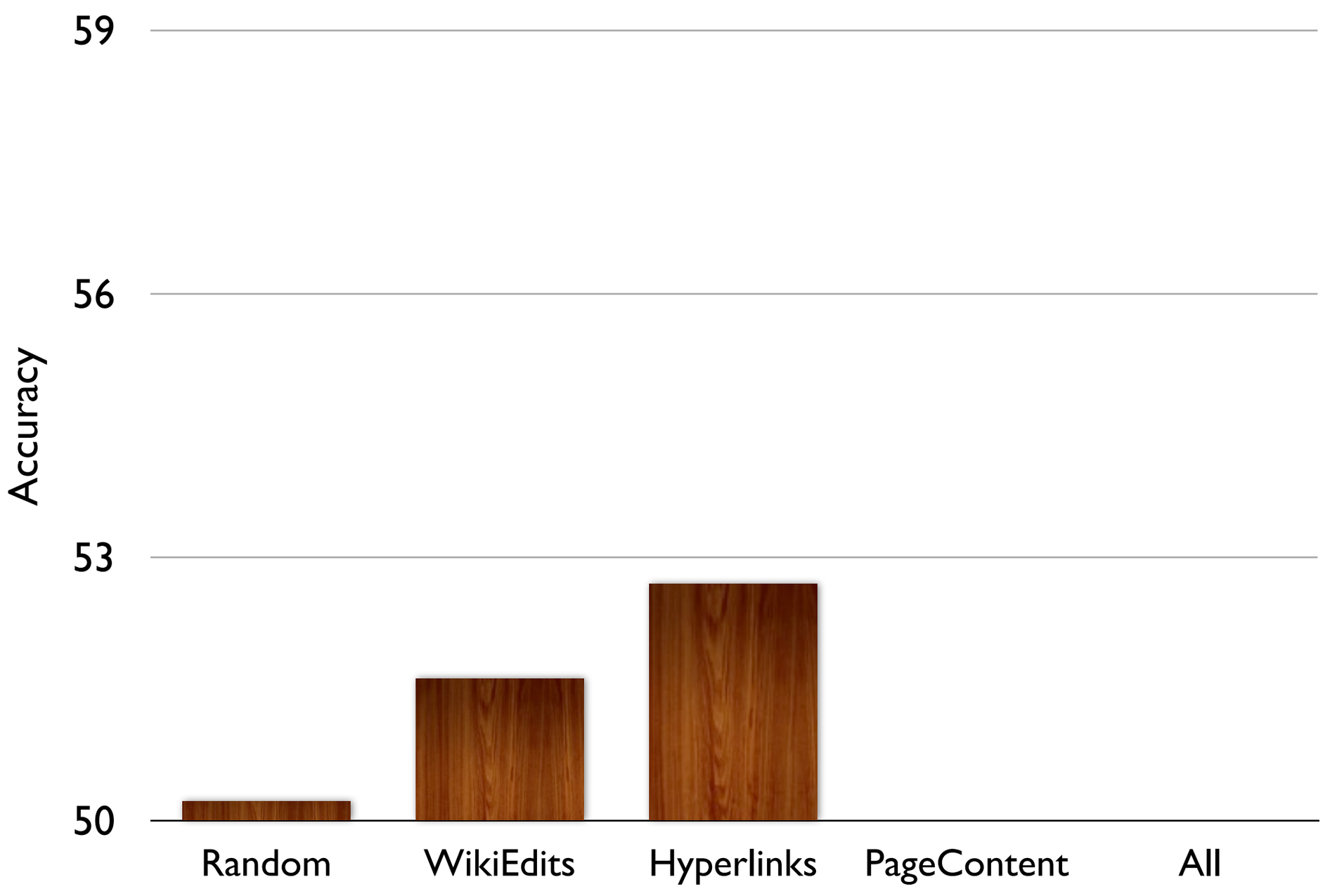
Effect of Different Features



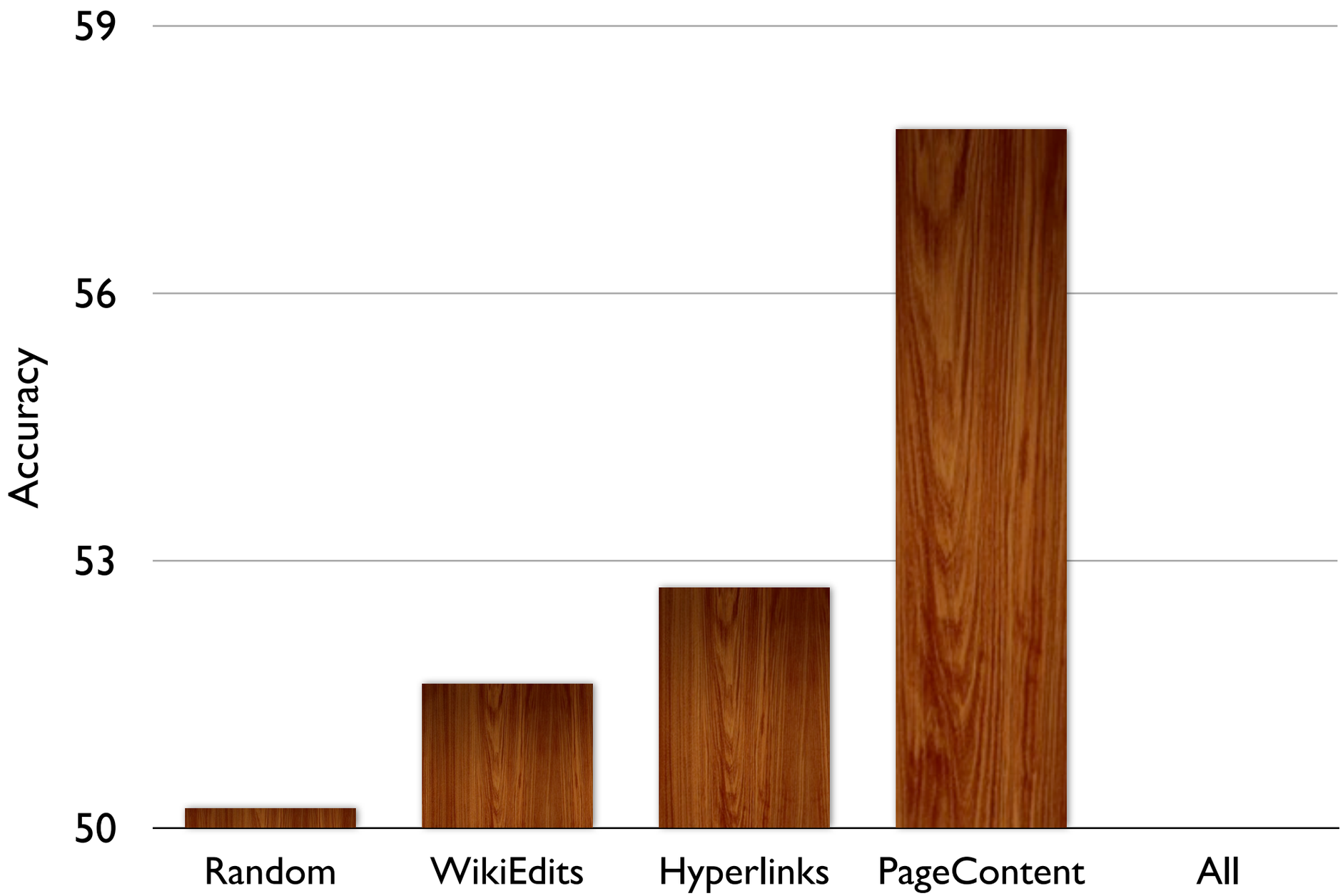
Effect of Different Features



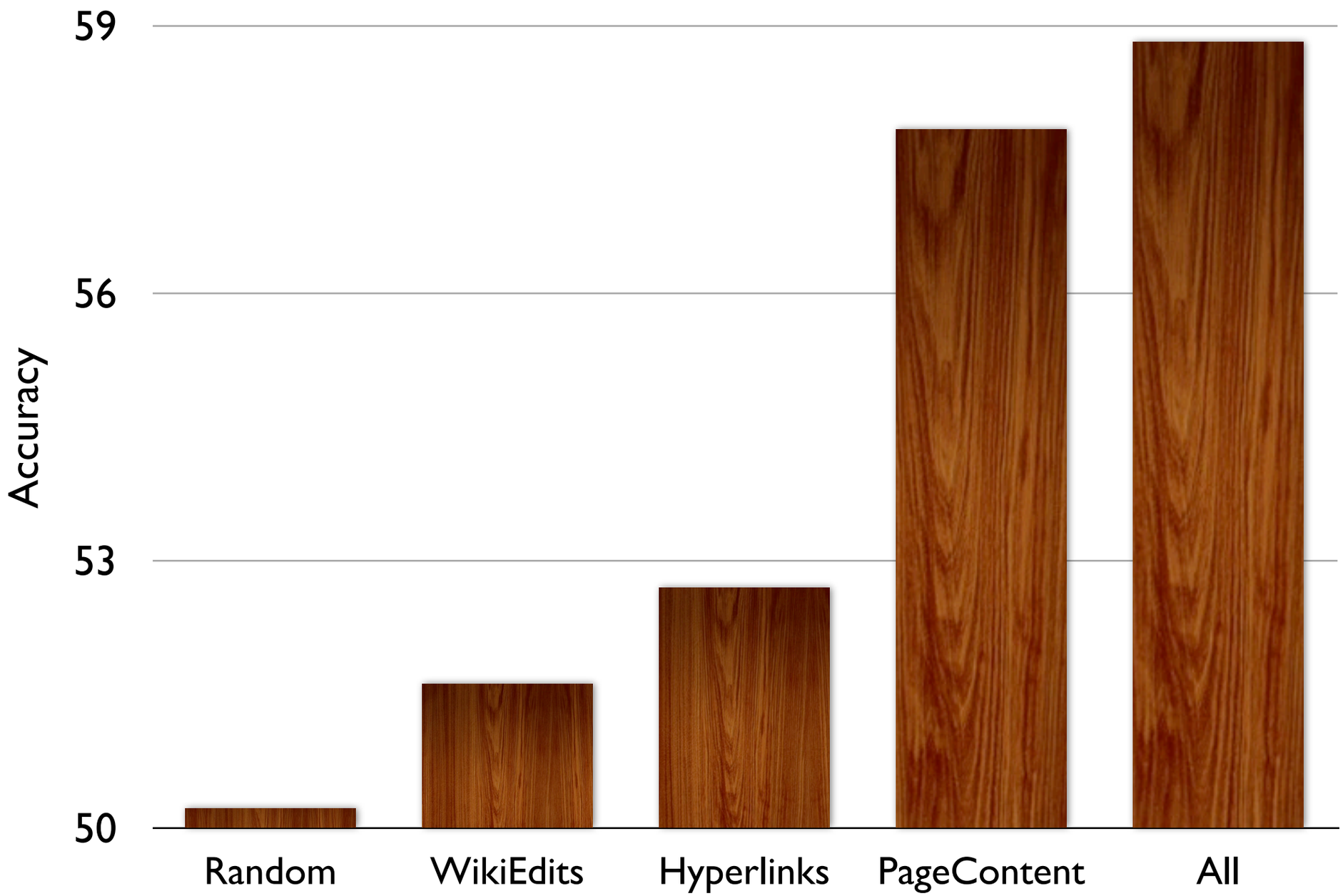
Effect of Different Features



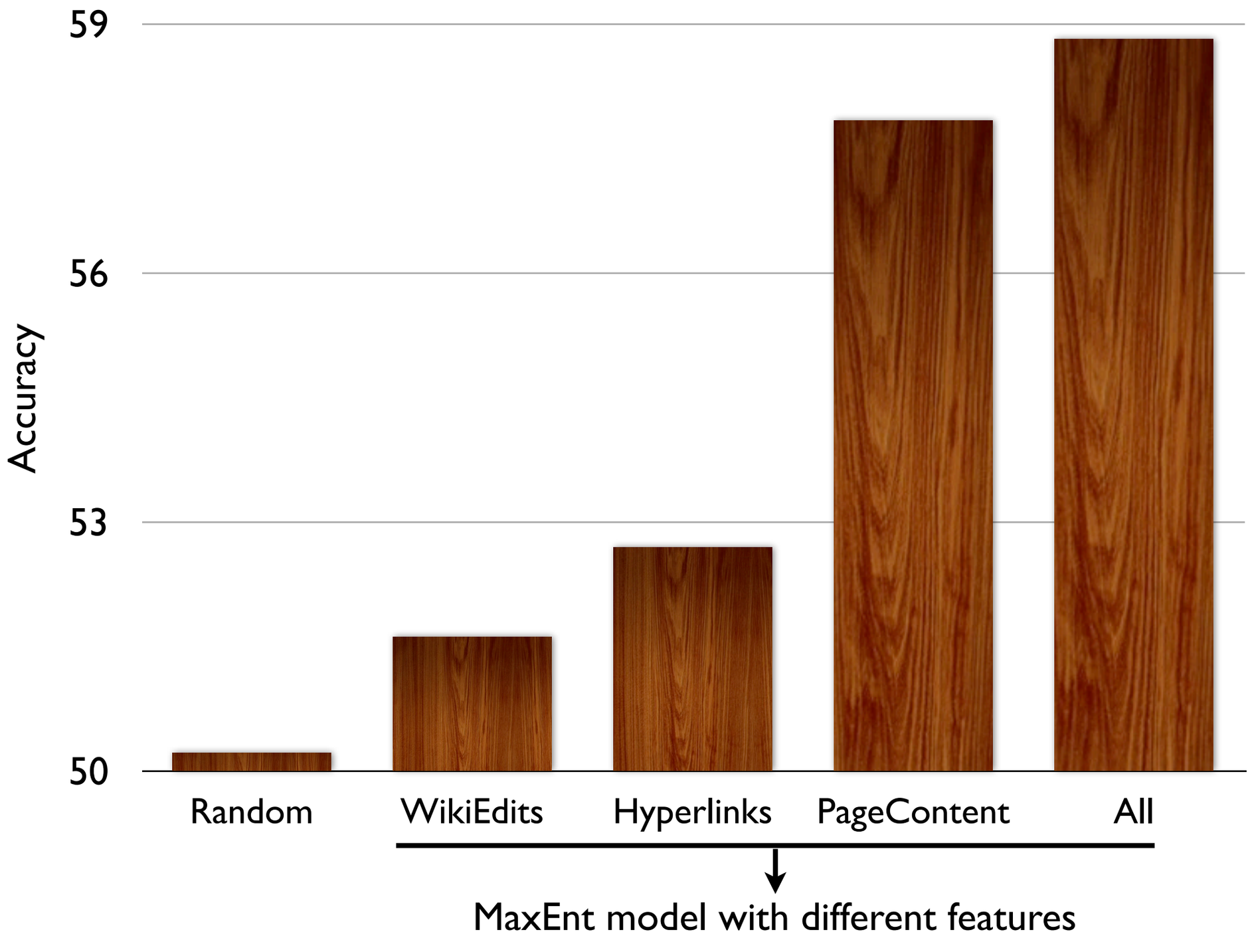
Effect of Different Features



Effect of Different Features



Effect of Different Features



Incorrect Classifications

Incorrect Classifications

- Mass is not a prerequisite for Physics
- Quantum mechanics is not a prerequisite for Bohr-Einstein debates
- Global warming is a prerequisite for Nitrous Oxide
- Carbon Dioxide is not a prerequisite for Carbon Sequestration

Examples of Discriminative Features

Examples of Discriminative Features

- Is Target title a substring of Source title?

Examples of Discriminative Features

- Is Target title a substring of Source title?
- Is there category overlap between source and target pages?

Examples of Discriminative Features

- Is Target title a substring of Source title?
- Is there category overlap between source and target pages?
- Identity of category overlap between pages.

Examples of Discriminative Features

- Is Target title a substring of Source title?
- Is there category overlap between source and target pages?
- Identity of category overlap between pages.
- Identity of source section from where target is linked.

Conclusion

Conclusion

- Novel task: prerequisite structure prediction
  - Demonstrated that relatively reliable features exist

Conclusion

- Novel task: prerequisite structure prediction
  - Demonstrated that relatively reliable features exist
- Future work
  - feature engineering
  - controlled data collection

Conclusion

- Novel task: prerequisite structure prediction
  - Demonstrated that relatively reliable features exist
- Future work
  - feature engineering
  - controlled data collection
- Ongoing work
  - full comprehension plan generation using predicted prerequisite structure

Thank You!