Discourse Complements Lexical Semantics for Non-factoid Answer Reranking

Peter Jansen[1], Mihai Surdeanu[1], and
Peter Clark[2]

[1]University of Arizona
[2]Allen Institute for Artificial Intelligence

THE UNIVERSITY OF ARIZONA

NLP

A12 ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Discourse Complements Lexical Semantics for Non-factoid Answer Reranking

Searching for "Answer Structures" in Text

Peter Jansen[1], Mihai Surdeanu[1], and
Peter Clark[2]

[1]University of Arizona
[2]Allen Institute for Artificial Intelligence

THE UNIVERSITY OF ARIZONA

NLP

AI2 ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

Motivation

- Do people write answers differently than other text?

- Over and above lexical matches, what kind of "deep structure" do good answers generally contain?

- Can we detect "answer structures" or explanatory text that is suggestive of high-quality answers in a corpus?

Example

Q: How does myelination affect action potentials?

A: The speed of conduction increases with the diameter of the axon and, in many vertebrate axons, with myelination. ...

Elaboration

YAHOO! ANSWERS

Answers Home

All Categories

Arts & Humanities

Beauty & Style

Business & Finance

Cars & Transportation

Computers & Internet

Consumer Electronics

Dining Out

Education & Reference

Entertainment & Music

Environment

Family & Relationships

Food & Drink

Games & Recreation

Health

Home & Garden

Local Businesses

News & Events

Pets

Politics & Government

How can you make your straight hair wavy without using a curler? ☆

Best Answer `Asker's Choice`

melissa p answered 7 years ago

Salt water is your best "Natural look" to achieve this use a spray bottle to mix about 8oz h2o to 1tbs kosher salt spray in hair all over as even as possible when damp them you have to work..... take hair in fairly large but not to large of section twists with fingers and blow dry each twist for best results remember the smaller the section the tighter the wave.Good luck

Other Answers (17)   Rated Highest ⌄

shoujomaniac101 answered 7 years ago

Back in the 1940s, they would make 'pin-curls': women would wet down their hair and twist sections into curls, then pin down. The tighter the twist, the tighter the curl. They'd do this before they went to bed and in the morning would have waves or curls! You can do the same thing today, except instead of just wetting and twisting your hair, use a setting lotion to help maintain the wave throughout the next day.

Mychelle K answered 7 years ago

While your hair is damp, braid it then put it in a bun. It's better if you sleep on it. Then let it loose in the morning when it's dry, and just tease it a little with your fingers. don't brush it! Put a little hair spray to set.

Don't do it too often though, it's not too healthy.

gabriellepeek answered 7 years ago

You can have beautiful mermaid waves by towel-drying your hair and putting in mousse from the roots to the ends and braiding it. make sure to separate the individual braids the way you want the waves to dry(meaning, don't use just two braids with a center part, as when it dries, you will have a big straight section in the middle of the back of your head) i would make two braids on each front side, 2 on each back side and 2 for the back. make them as tight as you can for best results and braid all the way down to the ends. secure with braiding hair ties(tinier and rubber or plastic). let it air dry. don't sleep on it or blow dry it or you will get fuzzies.

Cheryl~* answered 7 years ago

I think the best way to get best-looking wavy hair is.... crossing your hair and tie it at the end. to get fine curl, cross a little amount of hair each time; for bigger wave, just make it 4-5 bunch. Spray some moose or stylling gel on it. Left for a few hours or over night for the best result(at least 6-9 hours depends on individuals' hair structure).

Related

Straight to

How to ma
without usi

How can I

How can y
without usi

How do I g
curler?

Today o

World Cup

Yahoo! Answers Community Question Answering Corpus

10,000 "How" QA pairs (5k train, 2.5k dev/test)
Minimum 4 answers per question (average 9)

instructions for making proteins and RNAs and by replicating with each cell division, DNA ensures faithful inheritance of genetic information from generation to generation.

Genomics: Large-Scale Analysis of DNA Sequences

The entire "library" of genetic instructions that an organism inherits is called its **genome**. A typical human cell has two similar sets of chromosomes, and each set has DNA totaling about 3 billion nucleotide pairs. If the one-letter abbreviations for the nucleotides of one strand were written in letters the size of those you are now reading, the genetic text would fill about 600 books the size of this one. Within this genomic library of nucleotide sequences are genes for about 75,000 kinds of proteins and an as yet unknown number of RNA molecules that do not code for proteins.

Since the early 1990s, the pace at which we can sequence genomes has accelerated at an almost unbelievable rate, enabled by a revolution in technology. The development of new methods and DNA-sequencing machines, such as those shown in **Figure 1.12**, have led the charge. The entire sequence of nucleotides in the human genome is now known, along with the genome sequences of many other organisms, including bacteria, archaea, fungi, plants, and other animals.

The sequencing of the human genome was heralded as a scientific and technological achievement comparable to landing the *Apollo* astronauts on the moon in 1969. But it was only the beginning of an even bigger research endeavor, an effort to learn how the activities of the myriad proteins encoded by the DNA are coordinated in cells and whole organisms. To make sense of the deluge of data from genomesequencing projects and the growing catalog of known protein functions, scientists are applying a systems approach at the cellular and molecular levels. Rather than investigating a single gene at a time, these researchers have shifted to studying whole sets of genes of a species as well as comparing genomes between species—an approach called **genomics**.
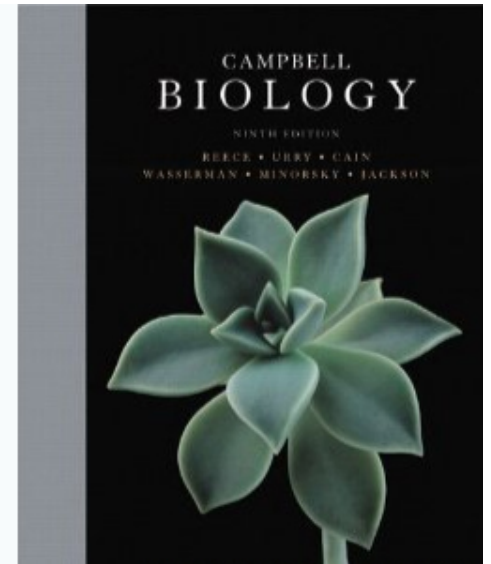
Three important research developments have made the genomic approach possible. One is "high-throughput" technology, tools that can analyze biological materials very rapidly and produce enormous amounts of data. The automatic DNA-sequencing machines that made the sequencing of the human genome possible are examples of high-throughput devices (see **Figure 1.12**). The second major development is **bioinformatics**, the use of computational tools to store, organize, and analyze the huge volume of data that result from high-throughput methods. The third key development is the formation of interdisciplinary research teams—melting pots of diverse specialists that may include computer scientists, mathematicians, engineers, chemists, physicists, and, of course, biologists from a variety of fields.

Theme: Feedback Mechanisms Regulate Biological Systems

Just as a coordinated control of traffic flow is necessary for a city to function smoothly, regulation of biological processes is crucial to the operation of living systems. Consider your muscles, for instance. When your muscle cells require more energy during exercise, they increase their consumption of the sugar molecules that serve as fuel. In contrast, when you rest, a different set of chemical reactions converts surplus sugar to storage molecules.

Like most of the cell's chemical processes, those that either decompose or store sugar are accelerated, or catalyzed, by proteins called enzymes. Each type of enzyme catalyzes a specific chemical reaction. In many cases, these reactions are linked into chemical pathways, each step with its own enzyme. How does the cell coordinate its various chemical pathways? In our example of sugar management, how does the cell match fuel supply to demand, regulating its opposing pathways of sugar consumption and storage? The key is the ability of many biological processes to self-regulate by a mechanism called feedback.

In feedback regulation, the output, or product, of a process regulates that very process. The most common form of regulation in living systems is **negative feedback**, in which accumulation of an end product of a process slows that process. For example, the cell's breakdown of sugar generates chemical energy in the form of a substance called ATP. When a cell makes more ATP than it can use, the excess ATP "feeds back" and inhibits an enzyme near the beginning of the pathway (**Figure 1.13a**).

protists. Although protists were once ...
species. And recent evidence shows t...
fungi than they are to each other. Th...

▼ Figure 1.15 The three domai...

(a) Domain Bacteria

Bacteria are the most diverse and wi...
now classified into multiple kingdoms...
in this photo is a bacterial cell.

(c) Domain Eukarya

▲ **Kingdom Plantae** consists of
terrestrial multicellular

All these RNAs are specified by genes, and the process of their transcription is also referred to as gene expression By carrying the instructions for making proteins and RNAs and by replicating with each cell division, DNA ensures faithful inheritance of genetic information from generation to generation.

Genomics: Large-Scale Analysis of DNA Sequences

The entire "library" of genetic instructions that an organism inherits is called its **genome**. A typical human cell has two similar sets of chromosomes, and each set has DNA totaling about 3 billion nucleotide pairs. If the one-letter abbreviations for the nucleotides of one strand were written in letters the size of those you are now reading, the genetic text would fill about 600 books the size of this one. Within this genomic library of nucleotide sequences are genes for about 75,000 kinds of proteins and an as yet unknown number of RNA molecules that do not code for proteins.

Since the early 1990s, the pace at which we can sequence genomes has accelerated at an almost unbelievable rate, enabled by a revolution in technology. The development of new methods and DNA-sequencing machines, such as those shown in **Figure 1.12**, have led the charge. The entire sequence of nucleotides in the human genome is now known, along with the genome sequences of many other organisms, including bacteria, archaea, fungi, plants, and other animals.

The sequencing of the human genome was heralded as a scientific and technological achievement comparable to landing the *Apollo* astronauts on the moon in 1969. But it was only the beginning of an even bigger research endeavor, an effort to learn how the activities of the myriad proteins encoded by the DNA are coordinated in cells and whole organisms. To make sense of the deluge of data from genome-sequencing projects and the growing catalog of known protein functions, scientists are applying a systems approach at the cellular and molecular levels. Rather than investigating a single gene at a time, these researchers have shifted to studying whole sets of genes of a species as well as comparing genomes between species—an approach called **genomics**.
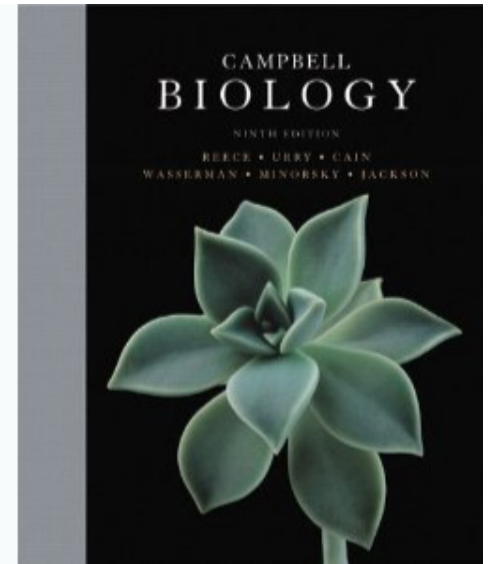
Three important research developments have made the genomic approach possible. One is "high-throughput" technology, tools that can analyze biological materials very rapidly and produce enormous amounts of data. The automatic DNA-sequencing machines that made the sequencing of the human genome possible are examples of high-throughput devices (see **Figure 1.12**). The second major development is **bioinformatics**, the use of computational tools to store, organize, and analyze the huge volume of data that result from high-throughput methods. The third key development is the formation of interdisciplinary research teams—melting pots of diverse specialists that may include computer scientists, mathematicians, engineers, chemists, physicists, and, of course, biologists from a variety of fields.

Theme: Feedback Mechanisms Regulate Biological Systems

Just as a coordinated control of traffic flow is necessary for a city to function smoothly, regulation of biological processes is crucial to the operation of living systems. Consider your muscles, for instance. When your muscle cells require more energy during exercise, they increase their consumption of the sugar molecules that serve as fuel. In contrast, when you rest, a different set of chemical reactions converts surplus sugar to storage molecules.

Like most of the cell's chemical processes, those that either decompose or store sugar are accelerated, or catalyzed, by proteins called enzymes. Each type of enzyme catalyzes a specific chemical reaction. In many cases, these reactions are linked into chemical pathways, each step with its own enzyme. How does the cell coordinate its various chemical pathways? In our example of sugar management, how does the cell match fuel supply to demand, regulating its opposing pathways of sugar consumption and storage? The key is the ability of many biological processes to self-regulate by a mechanism called feedback.

In feedback regulation, the output, or product, of a process regulates that very process. The most common form of regulation in living systems is **negative feedback**, in which accumulation of an end product of a process slows that process. For example, the cell's breakdown of sugar generates chemical energy in the form of a substance called ATP. When a cell makes more ATP than it can use, the excess ATP "feeds back" and inhibits an enzyme near the beginning of the pathway (**Figure 1.13a**).

CAMPBELL
BIOLOGY
NINTH EDITION
REECE · URRY · CAIN
WASSERMAN · MINORSKY · JACKSON

protists. Although protists were once
species. And recent evidence shows th
fungi than they are to each other. Th

▼ Figure 1.15 The three domai

(a) Domain Bacteria

Bacteria are the most diverse and wi
now classified into multiple kingdoms
in this photo is a bacterial cell.

(c) Domain Eukarya

▲ **Kingdom Plantae** consists of
terrestrial multicellular

...instructions for making proteins and RNAs and by replicating with each cell division, DNA ensures faithful inheritance of genetic information from generation to generation.

Genomics: Large-Scale Analysis of DNA Sequences

The entire "library" of genetic instructions that an organism inherits is called its genome. A typical human cell has two similar sets of chromosomes, and each set has DNA totaling about 3 billion nucleotide pairs. If the one-letter abbreviations for the nucleotides of one strand were written in letters the size of those you are now reading, the genetic text would fill about 600 books the size of this one. Within this genomic library of nucleotide sequences are genes for about 75,000 kinds of proteins and an as yet unknown number of RNA molecules that do not code for proteins.

Since the early 1990s, the pace at which we can sequence genomes has accelerated at an almost unbelievable rate, enabled by a revolution in technology. The development of new methods and DNA-sequencing machines, such as those shown in Figure 1.12, have led the charge. The entire sequence of nucleotides in the human genome is now known, along with the genome sequences of many other organisms, including bacteria, archaea, fungi, plants, and other animals.

The sequencing of the human genome was heralded as a scientific and technological achievement comparable to landing the Apollo astronauts on the moon in 1969. But it was only the beginning of an even bigger research endeavor, an effort to learn how the activities of the myriad proteins encoded by the DNA are coordinated in cells and whole organisms. To make sense of the deluge of data from genomesequencing projects and the growing catalog of known protein functions, scientists are applying a systems approach at the cellular and molecular levels. Rather than investigating a single gene at a time, these researchers have shifted to studying whole sets of genes of a species as well as comparing genomes between species—an approach called genomics.

Three important research developments have made the genomic approach possible. One is "high-throughput" technology, tools that can analyze biological materials very rapidly and produce enormous amounts of data. The automatic DNA-sequencing machines that made the sequencing of the human genome possible are examples of high-throughput devices (see Figure 1.12). The second major development is bioinformatics, the use of computational tools to store, organize, and analyze the huge volume of data that result from high-throughput methods. The third key development is the formation of interdisciplinary research teams—melting pots of diverse specialists that may include computer scientists, mathematicians, engineers, chemists, physicists, and, of course, biologists from a variety of fields.
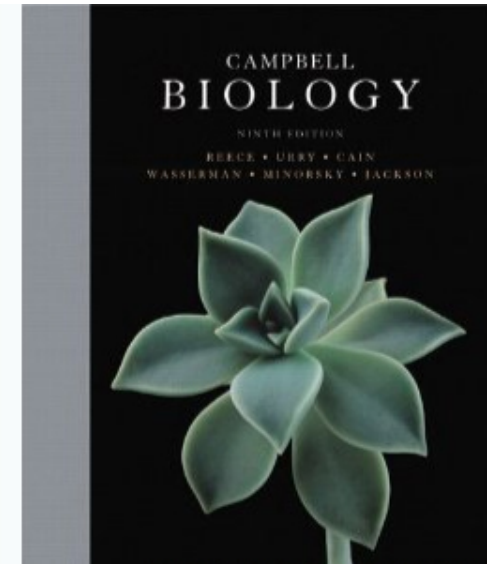
Theme: Feedback Mechanisms Regulate Biological Systems

Just as a coordinated control of traffic flow is necessary for a city to function smoothly, regulation of biological processes is crucial to the operation of living systems. Consider your muscles, for instance. When your muscle cells require more energy during exercise, they increase their consumption of the sugar molecules that serve as fuel. In contrast, when you rest, a different set of chemical reactions converts surplus sugar to storage molecules.

Like most of the cell's chemical processes, those that either decompose or store sugar are accelerated, or catalyzed, by proteins called enzymes. Each type of enzyme catalyzes a specific chemical reaction. In many cases, these reactions are linked into chemical pathways, each step with its own enzyme. How does the cell coordinate its various chemical pathways? In our example of sugar management, how does the cell match fuel supply to demand, regulating its opposing pathways of sugar consumption and storage? The key is the ability of many biological processes to self-regulate by a mechanism called feedback.

CAMPBELL
BIOLOGY
NINTH EDITION
REECE · URRY · CAIN
WASSERMAN · MINORSKY · JACKSON

protists. Although protists were once...
species. And recent evidence shows th...
fungi than they are to each other. Th...

▼ Figure 1.15 The three domai...

(a) Domain Bacteria

Bacteria are the most diverse and wi...
now classified into multiple kingdoms...
in this photo is a bacterial cell.

(c) Domain Eukarya

Campbell's Biology Corpus

185 "How" and 193 "Why" QA pairs crafted by domain expert
Gold answers located in textbook

Reranking Model Architecture

Question

Answer Collection
(Campbell's Biology, or
Yahoo! Answers DB)

Candidate Retrieval

A1

Answer Reranking

A2   +++
      ++
      +

Ranked List of
Answer Candidates

Models

Discourse Markers
(shallow)

Discourse Parser
(deep)

Lexical
Semantics

Marcu's (1997)
Discourse Markers

RST TreeBank

Gigaword
Bio Wikipedia
Bio Textbook
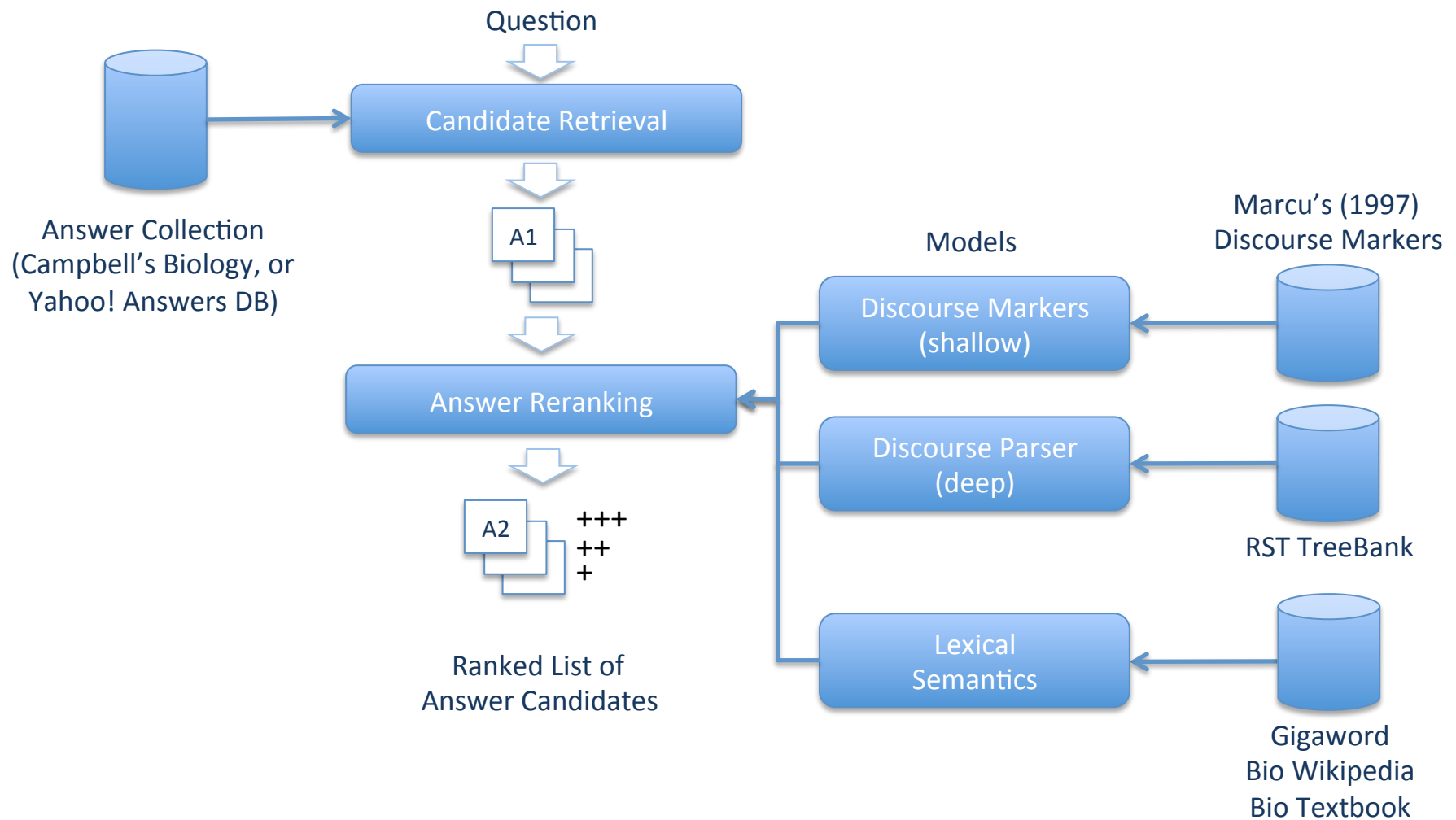
Reranking Model Architecture
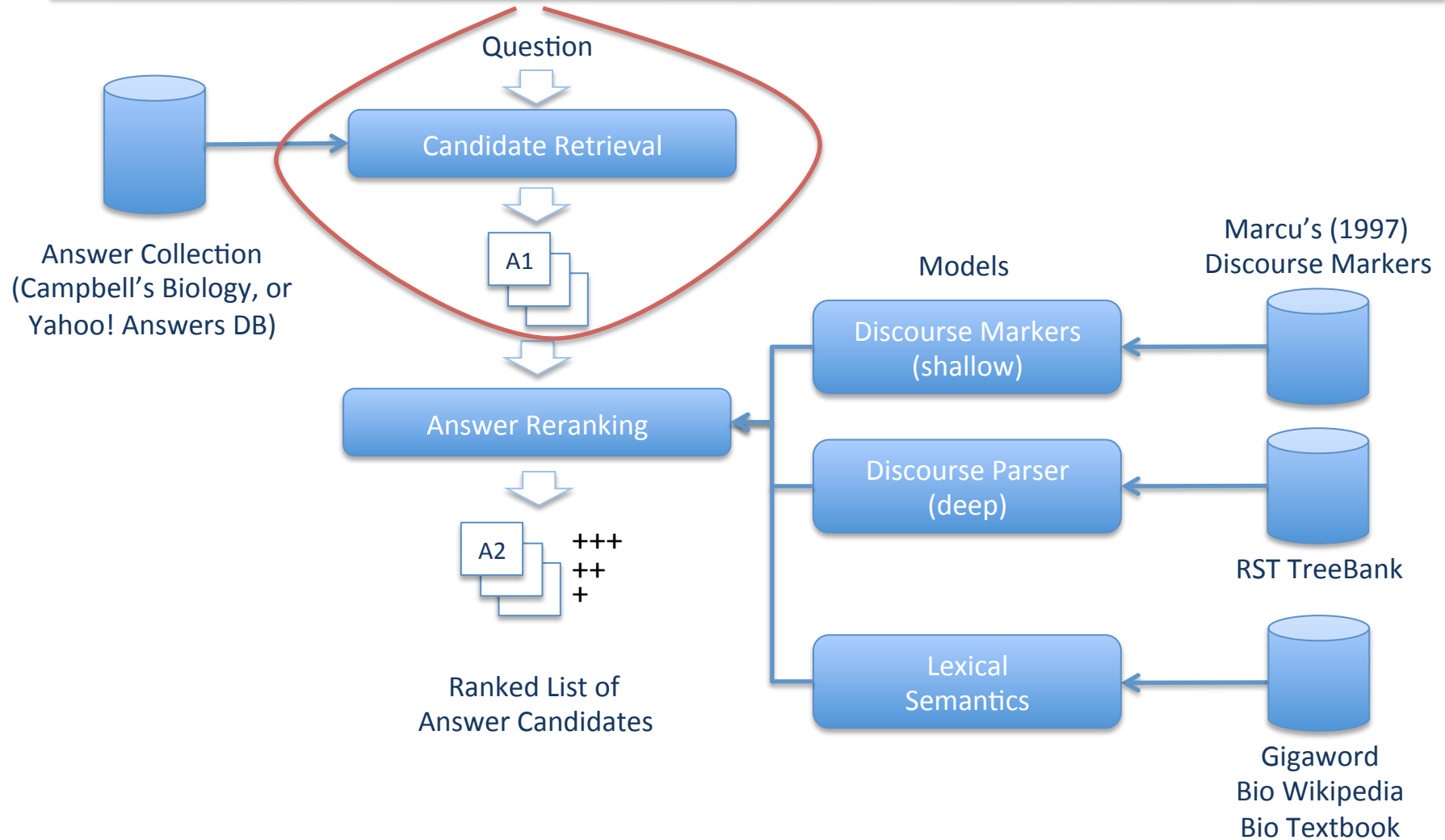
Reranking Model Architecture

Example

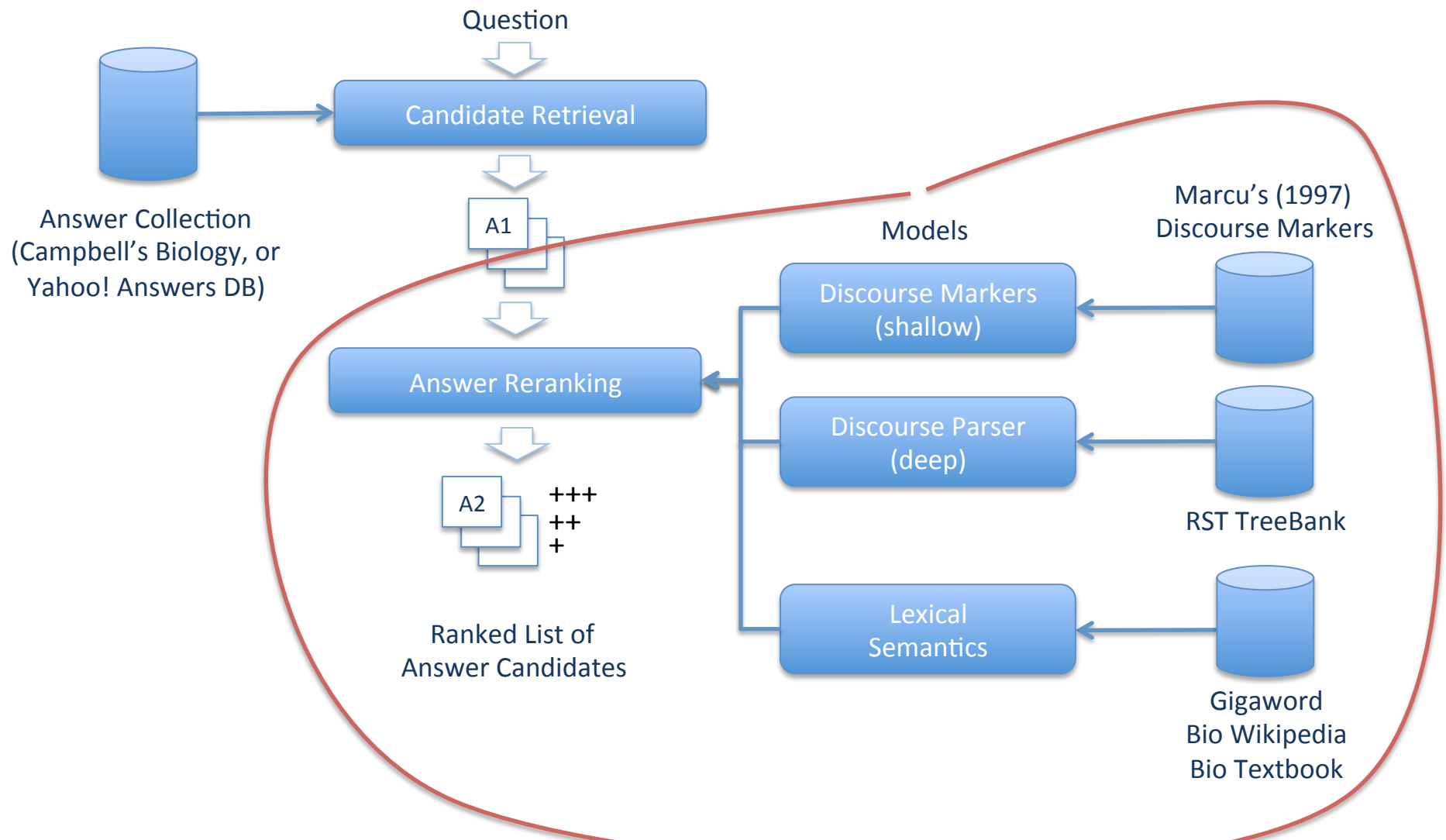Q: How is CO_2 first incorporated into the Calvin Cycle?

The Calvin cycle uses the chemical energy of ATP and NADPH to reduce CO_2 to sugar.

After IR

Example

Q: How is CO2 first incorporated into the Calvin Cycle?

❌ The Calvin cycle uses the chemical energy of ATP and NADPH to reduce CO2 to sugar.

After IR

✓ …The cycle begins by incorporating CO2 from the air into organic molecules already present in the chloroplast. This initial incorporation of carbon…

After Reranking

Representing Discourse: Shallow and Deep Features

Shallow representation:

- **Discourse Markers:** "a word or phrase whose function is to organize discourse into segments"
- Examples: and, in, that, for, if, as, not, by, but, because, …

Deep representation:

- **Discourse Parsing:** "a description of how two pieces of text connect"
- Examples from Rhetorical Structure Theory (RST): Elaboration, Attribution, Contrast, Background, Evaluation, …

Discourse Marker Features

Q: How do I compare two strings in Java?

Markers: and, in, that, for, if, as, not, by, but, …

A: There are several ways in Java. The easiest is by using the equals() method of String. The…

Discourse Marker Features

Q: How do I compare two strings in Java?

QSEG

Markers: and, in, that, for, if, as, not, by, but, …

A: There are several ways in Java. The easiest is by using the equals() method of String. The…

Discourse Marker Features

Q: How do I compare two strings in Java?

QSEG

Markers: and, in, that, for, if, as, not, by, but, …

Text not in question

Marker

Matches question

A: There are several ways in Java. The easiest is by using the equals() method of String. The…

OTHER by QSEG

Discourse Marker Features

Q: How do I compare two strings in Java?

QSEG

Markers: and, in, that, for, if, as, not, by, but, …

A: There are several ways in Java. The easiest is by using the equals() method of String. The…

OTHER in QSEG   OTHER by QSEG

OTHER in QSEG

QSEG by QSEG

Discourse Parser Features: Rhetorical Structure Theory

- Rhetorical Structure Theory (RST)

Elaboration

There are several ways in Java.

Elaboration

The easiest is by using the equals() method of String.

The == operator compares Strings for reference equality.
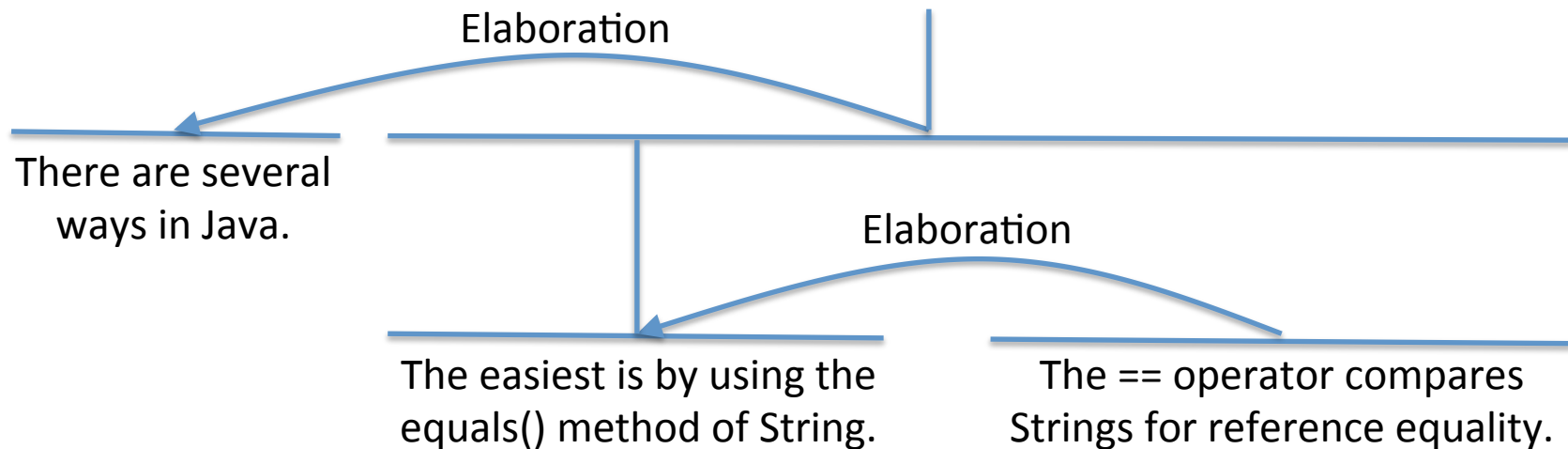
- **Discourse Relations:** Elaboration, Contrast, Comparison, Background, Cause, Attribution, Evaluation, etc…
- We use Feng & Hirst's (2012) discourse parser
  - Soon to release our own

Discourse Parser Features: Example

Discourse Parser Features: Example

(Campbell's Biology)
Q: How does myelination affect action potentials?

❌

A1: 1. How do action potentials and graded potentials differ?

Answer Collection

Elaboration

✔ A20: ... The speed of conduction increases with the diameter of the axon and in many vertebrate axons, with myelination. ...

OTHER   **QSEG**

Lexical Semantics

- Recursive Neural Network Language Models (Mikolov et al., 2013)

"photosynthesis"

"plant"

"dog"
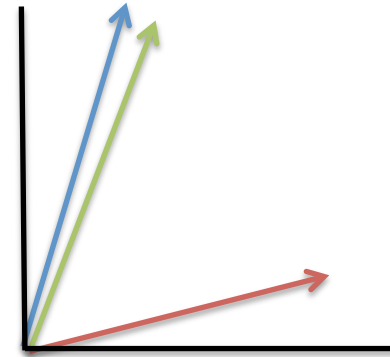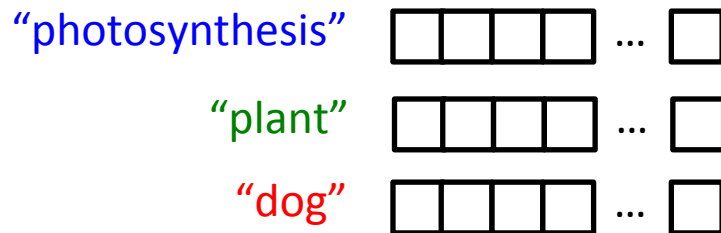
- Lexical semantics have been shown to help "bridge the lexical chasm" and greatly increase QA performance (Yih et al., 2013)

- **Control condition:** If our discourse features are truly detecting deep answer structures, we would expect the performance to stack with lexical matching models.

Reranking Model Architecture
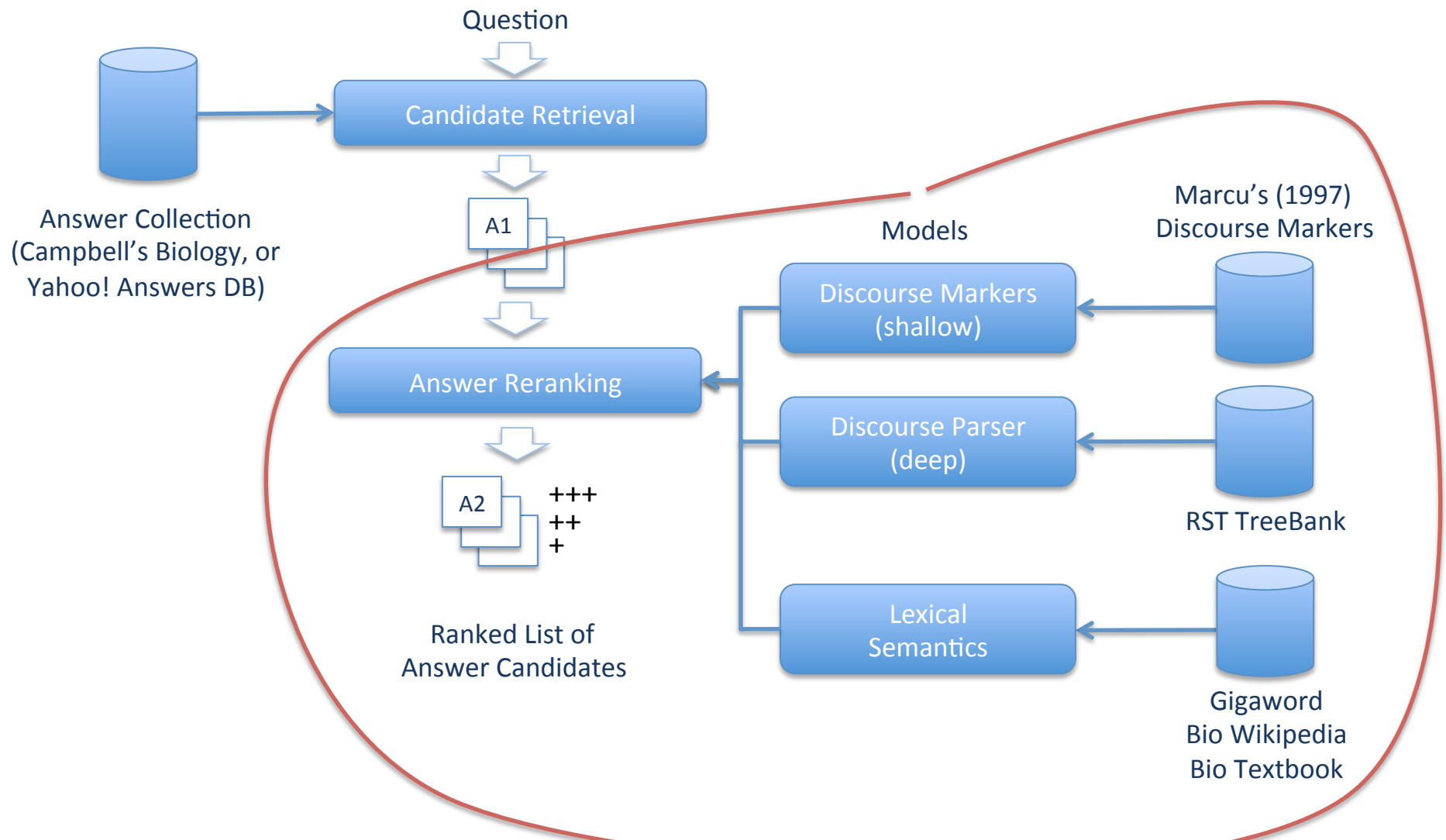
Performance

Discourse Model Performance

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1	Precison@1 w/LS features
Random Baseline	14.3%	14.3%
IR Baseline	19.6%	27.6%
IR + Shallow	24.1% (+23%)	29.3% (+10%)
IR + Deep	24.3% (+24%)	28.7% (+8%)
IR + Shallow + Deep	**24.8% (+27%)**	30.5% (+15%)

In all cases, $p < 0.05$ over baseline

Discourse Model Performance

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1	Precison@1 w/LS features
Random Baseline	14.3%	14.3%
IR Baseline	19.6%	27.6%
IR + Shallow	24.1% (+23%)	29.3% (+10%)
IR + Deep	24.3% (+24%)	28.7% (+8%)
IR + Shallow + Deep	24.8% (+27%)	30.5% (+15%)

In all cases, $p < 0.05$ over baseline

Discourse Model Performance

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1	Precison@1 w/LS features
Random Baseline	14.3%	14.3%
IR Baseline	19.6%	27.6%
IR + Shallow	24.1% (+23%)	29.3% (+10%)
IR + Deep	24.3% (+24%)	28.7% (+8%)
IR + Shallow + Deep	**24.8% (+27%)**	**30.5% (+15%)**

In all cases, p<0.05 over baseline

Discourse Model Performance

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1	Precison@1 w/LS features
Random Baseline	14.3%	14.3%
IR Baseline	19.6%	27.6%
IR + Shallow	24.1% (+23%)	29.3% (+10%)
IR + Deep	24.3% (+24%)	28.7% (+8%)
IR + Shallow + Deep	24.8% (+27%)	30.5% (+15%)

In all cases, $p<0.05$ over baseline

Discourse Model Performance

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1	Precison@1 w/LS features
Random Baseline	14.3%	14.3%
IR Baseline	19.6%	27.6%
IR + Shallow	24.1% (+23%)	29.3% (+10%)
IR + Deep	24.3% (+24%)	28.7% (+8%)
IR + Shallow + Deep	**24.8% (+27%)**	30.5% (+15%)

In all cases, $p < 0.05$ over baseline

Discourse Model Performance

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1	Precison@1 w/LS features
Random Baseline	14.3%	14.3%
IR Baseline	19.6%	27.6%
IR + Shallow	24.1% (+23%)	29.3% (+10%)
IR + Deep	24.3% (+24%)	28.7% (+8%)
IR + Shallow + Deep	**24.8% (+27%)**	**30.5% (+15%)**

+ LEXICAL SEMANTICS

+7.0%
GigaWord

In all cases, p<0.05 over baseline

Discourse Model Performance

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1 (+relative %)	Precison@1 w/LS features
Random Baseline	14.3%	14.3%
IR Baseline	19.6%	27.6%
IR + Shallow	24.1% (+23%)	29.3% (+10%)
IR + Deep	24.3% (+24%)	28.7% (+8%)
IR + Shallow + Deep	**24.8% (+27%)**	**30.5% (+15%)**

+5.2% (Abs)

+2.9% (Abs)

cases, p<0.05 over baseline

Discourse Model Performance

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1 (+relative %)	Precison@1 w/LS features
Random Baseline	14.3%	14.3%
IR Baseline	19.6%	27.6%
IR + Shallow	24.1% (+23%)	29.3% (+10%)
IR + Deep	24.3% (+24%)	28.7% (+8%)
IR + Shallow + Deep	**24.8% (+27%)**	**30.5% (+15%)**

+5.2% (Abs)

+2.9% (Abs)

cases, p<0.05 over baseline

Discourse features are approximately **60%** orthogonal to Lexical Semantics features

Discourse Model Performance

- Campbell's Biology "How" (185 Questions)

Features	Precision@1 (+relative %)	Precison@1 w/LS features
IR Baseline	24.1%	25.6%
IR + Shallow	29.9% (+24%)	30.1% (+19%)
IR + Deep	28.9% (+20%)	28.5% (+12%)
IR + Shallow + Deep	**30.4% (+27%)**	**30.7% (+21%)**

+ LEXICAL SEMANTICS

+1.5% BioWiki

Discourse Model Performance

- Campbell's Biology "How" (185 Questions)

Features	Precision@1 (+relative %)	Precison@1 w/LS features
IR Baseline	24.1%	25.6%
IR + Shallow	29.9% (+24%)	30.1% (+19%)
IR + Deep	28.9% (+20%)	28.5% (+12%)
IR + Shallow + Deep	**30.4% (+27%)**	**30.7% (+21%)**

+6.3% (Abs)

+5.1% (Abs)

Discourse features are approximately **80%** orthogonal to Lexical Semantics features

Discourse Model Performance

- Campbell's Biology "Why" (193 Questions)

Features	Precision@1 (+relative %)	Precison@1 w/LS features
IR Baseline	28.6%	31.7%
IR + Shallow	38.0% (+33%)	38.6% (+22%)
IR + Deep	38.6% (+35%)	**39.5% (+24%)**
IR + Shallow + Deep	**39.4% (+38%)**	**39.3% (+24%)**

+10.8% (Abs)

+7.6% (Abs)

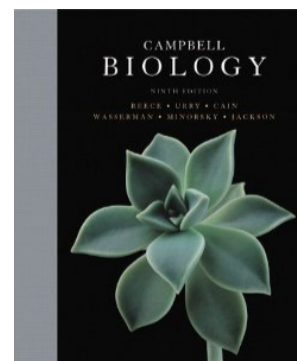Discourse features are approximately **70%** orthogonal to Lexical Semantics features

Analysis 1:
Domain Transfer

- How general are discourse features?

Train on Open Domain
CQA Reranking Task

Test on Biology Domain
QA Task

	Bio How (In-Domain)	Transfer (YA->Bio)
IR + Shallow + Deep	30.4% (+26%)	
IR + LS + Shallow + Deep	30.7% (+21%)	

Precision @1

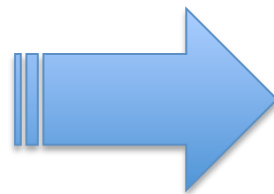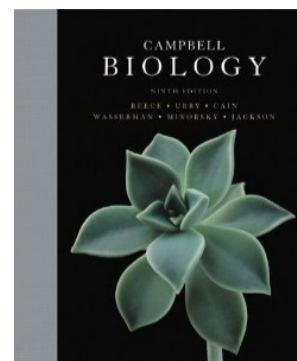# Analysis 1:
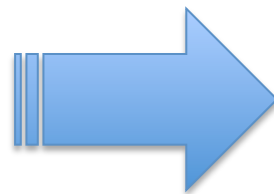# Domain Transfer

- How general are discourse features?

**Train on Open Domain
CQA Reranking Task**

**Test on Biology Domain
QA Task**

	Bio How (In-Domain)	Transfer (YA->Bio)
IR + Shallow + Deep	30.4% (+26%)	30.1% (+25%)
IR + LS + Shallow + Deep	30.7% (+21%)	29.5% (+17%)

Precision @1

Transfer Performance is nearly identical to in-domain training!

Analysis 2: Integrating Discourse and Lexical Semantics

- Can we create discourse features that include lexical semantics?

Q: How do dogs learn to bury bones?

QSEG

Markers: and, in, that, for, if, as, not, by, but, …

A: Puppies are natural diggers, and don't have to be taught by their parents. This is due to …

OTHER and OTHER

Analysis 2: Integrating Discourse and Lexical Semantics

- Can we create discourse features that include lexical semantics?

Q: How do dogs learn to bury bones?

QSEG

puppy

pet

chew

Markers: and, in, that, for, if, as, not, by, but, …

A: Puppies are natural diggers, and don't have to be taught by their parents. This is due to …

OTHER and OTHER

A: Puppies are natural diggers, and don't have to be taught by their parents. This is due to …

QSEG and OTHER

Analysis 2: Integrating Discourse and Lexical Semantics

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1
Random Baseline	14.3%
IR Baseline	19.6%
IR + LS Baseline	26.6%
IR + LS + Shallow + Deep	30.5% (+15%)
IR + LS + Shallow + Deep + Shallow_LS + Deep_LS	**32.9% (+24%)**

In all cases, $p<0.05$ over IR + LS Baseline

Analysis 2: Integrating Discourse and Lexical Semantics

- Yahoo! Answers Corpus CQA (10k Questions)

Features	Precision@1
Random Baseline	14.3%
IR Baseline	19.6%
IR + LS Baseline	26.6%
IR + LS + Shallow + Deep	30.5% (+15%)
IR + LS + Shallow + Deep + Shallow_LS + Deep_LS	**32.9% (+24%)**

In___ __es, p<0.05 over IR + LS Baseline

1 in 3 questions answered correctly

Summary

1. First to systematically explore structured *cross-sentence* features driven by both shallow and deep discourse representations for QA

2. Good performance improvements (up to +24% over a strong baseline) across thousands of questions in *two domains* and *two question types*

3. Discourse features are complementary to lexical semantic similarity, and attending to "answer structures"

4. Excellent domain transfer, suggesting our discourse features have general utility to non-factoid QA

Thank you!

- Source code (will be) available at:
  http://nlp.sista.arizona.edu/releases/acl2014

- Discourse parser available soon

Analysis:
Intra vs. Inter-sentence Features

- How important are cross-sentence features?

	Biology (How)	Biology (Why)	Yahoo! Answers
Within-sentence	+0.8%	+8.4%	+13.1%
Full model	+21.0%*	+23.9%*	+14.8%

Model: IR + LS + Shallow + Deep   Relative increases over baseline

Analysis:
Intra vs. Inter-sentence Features

- How important are cross-sentence features?

	Biology (How)	Biology (Why)	Yahoo! Answers
Within-sentence	+0.8%	+8.4%	+13.1%
Full model	+21.0%*	+23.9%*	+14.8%

Model: IR + LS + Shallow + Deep   Relative increases over baseline

+20%   +16%   +2%

39%   1 sentence
57%   1-2 sentences