CSTBank Annotation Guidelines

March 22, 2003

Contents

Chapter 1

Introduction

This document describes the annotation conventions used in the CSTBank corpus. Cross-document Structure Theory (CST) [Rad00] attempts to characterize the relationships that exist between pairs of sentences that come from one or more documents in a cluster of topically related articles. This document is intended to serve as a guide for annotators who will be marking up text for CST relationships, as well as for researchers using the CSTBank corpus in their work. The theory itself as well as its relevance to automatic summarization will be described in the next chapter.

The remainder of this document will be organized as follows. Section 2 will offer a formal description of each of the 18 CST relationships that have been established thus far. In Section 3, an example cluster of nine documents will be presented. The documents will be shown both in their raw formats as well as in fully annotated form. Next, the fourth section will present the XML scheme for CST judgement files. Lastly, Section 5 will give examples of summaries produced from three source documents from the Milan cluster.

1.0.1 Cross-document Structure Theory (CST)

Cross-document Structure Theory attempts to describe the relationships that exist between two or more sentences from different source documents that are related to the same topic, as well as those that come from a single source document. For example, consider a collection of news articles related to the same event or story, that have been collected over a period of a few days or weeks and are ordered with respect to their publication dates. We know that in their writing, journalists try to include sufficient background information in every article, such that even readers who have not kept up with a given story from its beginning may understand the article. Therefore, we would expect some of the information in the earlier articles to be repeated again in later documents. This information may appear as the identical sentence previously used or may be reworded or paraphrased. Such relationships may exist between a set of two or more sentences. Another example of a commonly encountered CST relationship is that of contradiction. If we collect a cluster of news stories from different news agencies and that are written by different authors, they frequently will not agree on all details of an event. They may contradict one another as to the particular time of an event, the exact location, the persons involved, and many other specific details.

Knowing the CST relationships that exist between sentences in a document cluster can be useful in a number of natural language processing tasks. One such task is that of multi-document summarization (MDS) [ZBGR02] . MDS is the process of automatically creating a summary from a set of topically-related documents. For instance, many online news delivery services, such as NewsInEssence (http://www.newsinessence.com) and NewsBlaster (http://www.cs.columbia.edu/nlp/newsblaster) create summaries of news stories, using multiple source articles related to the given topic. Such summaries have the advantage of providing different perspectives to readers, since source documents may be from different newspapers and written by different authors. Also, in the case where the source documents have publication dates that span the time period of a breaking news story, summaries can capture an entire story and it present it to the user, such that he or she does not have to read each document separately to learn about a current event or topic of interest.

However, creating a summary automatically from multiple source documents is quite challenging. Often times, the summaries may contain redundant messages, or even sentences that contradict one another. Knowing the CST relationships between the sentences could help researchers to build better summarization systems. For example, if it were known that a given pair of sentences were paraphrases of one another, such that they contained the same information, only one of them could be included in the summary. Such procedures might limit redundancy and generally improve the overall quality of automatically produced multi-document summaries.

Chapter 2

Definitions of CST Relationships

17 major CST relationships have been identified by [Rad00] and are presented in Figure 2.1. In addition, an eighteenth relationship has been recently added. The definitions and examples are based on those in [ZBGR02] . Note that some CST relationships have directionality. In the examples below, this means that they are expressed in terms of the relationship of the first sentence (S1) to the second sentence (S2). This will be explained further in the sections to follow.

ID	Relationship	Description	Span 1 (S1)	Span 2 (S2)
1	Identity	The same text appears in more than one location	Tony Blair was elected for (Repetition) a second term today.	Tony Blair was elected for a second term today.
2	Equivalence (Paraphrase)	Two text spans have the same information content	Derek Bell is experiencing a resurgence in his career.	Derek Bell is having a "comeback year."
3	Translation	Same information content in different languages	Shouts of "Viva la revolucion!" echoed through the night.	The rebels could be heard shouting, "Long live the revolution".
4	Subsumption	S1 contains all information in S2, plus additional information not in S2	With 3 wins this year, Green Bay has the best record in the NFL.	Green Bay has 3 wins this year.
5	Contradiction	Conflicting information	There were 122 people on the downed plane.	126 people were aboard the plane.
6	Historical Background	S1 gives historical context to information in S2	This was the fourth time a member of the Royal Family has gotten divorced.	The Duke of Windsor was divorced from the Duchess of Windsor yesterday.
7	Citation	S1 explicitly cites document S2	Prince Albert then went on to say, "I never gamble."	An earlier article quoted Prince Albert as saying "I never gamble."
8	Modality	S1 presents a qualified version of the information in S2, e.g., using "allegedly"	Sean "Puffy" Combs is reported to own several multimillion dollar estates.	Puffy owns four multimillion dollar homes in the New York area.
9	Attribution	S1 presents an attributed version of information in S2, e.g. using "According to CNN,"	According to a top Bush advisor, the President was alarmed at the news.	The President was alarmed to hear of his daughter's low grades.
10	Summary	S1 summarizes S2.	The Mets won the Title in seven games.	After a grueling first six games, the Mets came from behind tonight to take the Title.
11	Follow-up	S1 presents additional information which has happened since S2	102 casualties have been reported in the earthquake region.	So far, no casualties from the quake have been confirmed.
12	Indirect speech	S1 indirectly quotes something which was directly quoted in S2	Mr. Cuban then gave the crowd his personal guarantee of free Chalupas.	"I'll personally guarantee free Chalupas," Mr. Cuban announced to the crowd.
13	Elaboration / Refinement	S1 elaborates or provides details of some information given more generally in S2	50% of students are under 25; 20% are between 26 and 30; the rest are over 30.	Most students at the University are under 30.
14	Fulfillment	S1 asserts the occurrence of an event predicted in S2	After traveling to Austria Thursday, Mr. Green returned home to New York.	Mr. Green will go to Austria Thursday.
15	Description in S2	S1 describes an entity mentioned and father of two, has declined	Greenfield, a retired general yesterday. to comment.	Mr. Greenfield appeared in court
16	Reader Profile	S1 and S2 provide similar information written for a different audience.	The Durian, a fruit used in Asian cuisine, has a strong smell.	The dish is usually made with Durian.
17	Change of perspective	The same entity presents a differing opinion or presents a fact in a different light.	Giuliani criticized the Officer's Union as "too demanding" in contract talks.	Giuliani praised the Officer's Union, which provides legal aid and advice to members.
18	Overlap (partial equivalence)	S1 provides facts X and Y while S2 provides facts X and Z; X, Y, and Z should all be non-trivial.	A small plane crashed into the *25th floor of a skyscraper in downtown* Milan today.	A small plane crashed into the tallest building in downtown Milan Thursday evening, *causing smoke to pour out of the top floors of the skyscraper.*

Figure 2.1: CST relationships and examples

Chapter 3

Fully annotated examples

The following nine source documents are from the ”Milan cluster” of CSTBank. This cluster, which we have used as our training and development cluster for designing the CSTBank annotation scheme and guidelines, is about the crash of a small plane into a skyscraper in Milan, Italy that occurred on April 18, 2002 and the events surrounding it. They were collected live from the Web and were published by five different news sources. Each document is presented with its source and time of publication (the local time at the location of publication) so that the changes in facts over time and across information sources can be considered.

Since the Milan cluster contains articles collected live from the Web, it should be noted that some remnants of the original HTML captured from the news agencies' sites can be seen in the following documents. For example, in the first and second sentences of Document 1, one can observe some source information or advertisement (CNN.com and CNNenEspanol.com, respectively). Technically, CST relationships can occur between such elements. However, we have not shown such examples in this document, since they are trivial for our purposes.

```
Document 1
Source: CNN
Publication: April 18, 2002 at 12:22 p.m.

[1] CNN.com - Plane hits skyscraper in Milan - April 18, 2002

[2] CNNenEspanol.com A small plane has hit a skyscraper in
central Milan, setting the top floors of the 30-story building
on fire, an Italian journalist told CNN.

[3] The crash by the Piper tourist plane into the 26th floor
occurred at 5:50 p.m. (1450 GMT) on Thursday, said journalist
Desideria Cavina.

[4] The building houses government offices and is next to the
city's central train station.

[5] Several storeys of the building were engulfed in fire, she said.

[6] Italian TV says the crash put a hole in the 25th floor of the
Pirelli building, and that smoke is pouring from the opening.

[7] Police and ambulances are at the scene.

[8] Many people were on the streets as they left work for the evening
at the time of the crash.

[9] Police were trying to keep people away, and many ambulances were
on the scene.

[10] There is no word yet on casualties.

[11]  U.N. envoy horror at Jenin camp U.S. bombing kills Canadians
Chinese missiles concern U.S. 2002 Cable News Network LP, LLLP.

[12]  An AOL Time Warner Company.

[13] All Rights Reserved. under which this service is provided to you.
```

```
Document 2
Source: CNN
Publication: April 18, 2002 at 12:32 p.m.

[1] CNN.com - Plane hits skyscraper in Milan - April 18, 2002

[2] CNNenEspanol.com A small plane has hit a skyscraper in central
Milan, setting the top floors of the 30-story building on fire, an
Italian journalist told CNN.

[3] The crash by the Piper tourist plane into the 26th floor occurred
at 5:50 p.m. (1450 GMT) on Thursday, said journalist Desideria Cavina.

[4] The building houses government offices and is next to the city's
central train station.

[5] Several storeys of the building were engulfed in fire, she
said.

[6] Italian TV showed a hole in the side of the Pirelli building with
smoke pouring from the opening.

[7] RAI state TV reported that the plane had apparently radioed an SOS
because of engine trouble.

[8] Earlier though, in Rome, the senate's president, Marcello Pera, said
it "very probably" appeared to be a terrorist attack.

[9] Police and ambulances are at the scene.

[10] Many people were on the streets as they left work for the evening at
the time of the crash.

[11] Police were trying to keep people away, and many ambulances were on
the scene.

[12] There is no word yet on casualties.

[13] TV pictures from the scene evoked horrific memories of the September
11 attacks on the World Trade Center in New York and the collapse of the
building's twin towers.

[14] "I heard a strange bang so I went to the window and outside I saw the
windows of the Pirelli building blown out and then I saw smoke coming from
them," said Gianluca Liberto, an engineer who was working in the area told
Reuters.

[15] The building is known as the Pirelli skyscraper but the Italian tyre
and cable company does not operate out of the building.

[16] It is one of the symbols of Italy's financial capital and is one of
the world's tallest concrete buildings, designed between 1955 and
1960.
```

```
Document 2, con't

[17] U.N. envoy horror at Jenin camp U.S. bombing kills Canadians Chinese
missiles concern U.S.   2002 Cable News Network LP, LLLP.

[18] An AOL Time Warner Company.

[19] All Rights Reserved. under which this service is provided to
you.
```

```
Document 16
Source: ABC
Publication: April 18, 2002 at 1:00 p.m.

[1] ABCNEWS.com : Plane Slams Into Milan Skyscraper

[2] The Pirelli Building in Milan, Italy, was hit by a small
plane.

[3] ABCNEWS.com  A small plane crashed into a skyscraper in
downtown Milan today, setting several floors of the 30-story building on
fire.

[4] The plane crashed into the 25th floor of the Pirelli building in
downtown Milan.

[5] The weather was clear at the time of the crash.

[6] Smoke poured from the opening as police and ambulances rushed to the
area.

[7] The president of the Italian Senate, Marcello Pera, told Italian
television it "very probably" appeared to be a terrorist attack but soon
afterwards his spokesman said it was probably an accident.

[8] A transport official told Reuters the plane had reported problems with
its undercarriage and was circling the city ahead of trying to land at a
local airport.

[9] The Pirelli building houses the administrative offices of the local
Lombardy region and sits next to the city's central train station.

[10] It is constructed of concrete and glass.

[11] The crash happened just before rush hour, as office workers were
closing their day.

[12] ABCNEWS' Pierre Thomas contributed to this report.

[13] Copyright   2002 ABCNEWS Internet Ventures.
```

```
Document 19
Source: MSNBC
Publication: April 18, 2002 at 1:00 p.m.

[1] Small plane hits Milan building

[2] Smoke rises from the Milan skyscraper.

[3] A small airplane crashed into a government building in heart of Milan,
setting the top floors on fire, Italian police reported.

[4] There were no immediate reports on casualties as rescue workers
attempted to clear the area in the city s financial district.

[5] Few details of the crash were available, but news reports about it
immediately set off fears that it might be a terrorist act akin to the
Sept. 11 attacks in the United States.

[6] Those fears sent U.S. stocks tumbling to session lows in late morning
trading.

[7] WITNESSES REPORTED hearing a loud explosion from the 30-story office
building, which houses the administrative offices of the local Lombardy
region and sits next to the city s central train station.

[8] Italian state television said the crash put a hole in the 25th floor
of the Pirelli building.

[9] News reports said smoke poured from the opening.

[10] Police and ambulances rushed to the building in downtown
Milan.

[11] No further details were immediately available.

[12] Check back for additional details on this breaking story.

[13] The Associated Press and Reuters contributed to this report.
```

```
Document 21
Source: USA Today
Publication: April 18, 2002 at 1:03 p.m.

[1] USATODAY.com - Small plane slams into Milan skyscraper

[2] Smoke billows from the Pirelli building.

[3] MILAN, Italy (AP)   A small plane with only the pilot
on board on Thursday crashed into a landmark skyscraper
in downtown Milan.

[4] Smoke poured out of the 30-story building and there was no
immediate word of any victims.

[5] The building houses the regional government offices,
authorities said.

[6] Police officer Celerissimo De Simone said the pilot of the
Piper aircraft sent out a distress call at 5:50 p.m. just before
the crash near Milan's main train station.

[7] RAI state TV reported that the pilot said the SOS was because
of engine trouble.

[8] The weather was clear at the time.

[9] Police cordoned off the area as people gawked at the skyscraper.

[10] At least one ambulance was dispatched.

[11] Initially, Marcello Pera, president of Italy's senate, said
it "very probably" appeared to be a terrorist attack.

[12] Later, his spokesman said Pera had spoken with the Interior
Minister and the crash didn't appear to be any kind of an attack.

[13] The crash put a hole in the 25th floor of the Pirelli building,
and smoke was seen pouring from the opening.

[14] Police and ambulances rushed to the building.

[15] The Pirelli skyscraper, located near the central train station,
is Italy's first skyscraper and one of the world's tallest concrete
buildings.

[16] It was built in 1958 and designed by architects Gio Ponti and
Pier Luigi Nervi.
```

```
Document 21, con't

[17] The building is one of the main symbols of Milan, along with the
city's cathedral.

[18] "The president just moments ago was informed about the incident
in Milan.

[19] I have no additional information at all for you at this time.

[20] This is a breaking story and we don't have anything else
beyond that."

[21] It was the second time since the Sept. 11 terror attacks on
New York and Washington that a plane has struck a high-rise building.

[22] On Jan. 5, a 15-year-old boy crashed a stolen plane into a
building in Tampa, Fla.

[23] He was the only casualty.

[24] Copyright 2002 The Associated Press.

[25] All rights reserved.

[26] This material may not be published, broadcast,rewritten or
redistributed.

[27] Gannett Co. Inc.
```

```
Document 23
Source: ABC
Publication: April 18, 2002 at 1:42 p.m.

[1] ABCNEWS.com : Plane Slams Into Milan Skyscraper

[2] The Pirelli Building in Milan, Italy, was hit by
a small plane.

[3] A small plane crashed into a skyscraper in downtown
Milan today, setting several floors of the 32-story
building on fire.

[4] At least two people are dead, including the pilot,
Italy's ANSA wire service said.

[5] The plane crashed into the 25th floor of the Pirelli
building in downtown Milan.

[6] The weather was clear at the time of the crash.

[7] Italian TV reported the bottom floors of the building
had been evacuated, and that many people were injured.

[8] The Associated Press reported only the pilot was on board.

[9] The plane was destined for Italy's capital Rome, but
there were conflicting reports as to whether it had come
from Locarno, Switzerland or Sofia, Bulgaria.

[10] Smoke poured from the opening as police and ambulances
rushed to the area.

[11] The fire devastated two floors of the building, but is
now under control, reported ABCNEWS' Rome bureau.

[12] There were initial fears that the crash could have been
a terrorist attack, but local officials now say it was probably
an accident.

[13] A transport official told Reuters the plane had reported
problems with its undercarriage and was circling the city ahead
of trying to land at a local airport.

[14] Police officer Celerissimo De Simone told The AP the pilot
of the Piper aircraft had sent out a distress call at 5:50 p.m.
just before crashing.

[15] Italian TV said the pilot SOS involved a technical problem.
The crash happened during rush hour, as office workers were ending
their day.

[16] A small tourist plane crashed into the tallest building in Milan.
```

```
Document 23, con't

[17] The Pirelli building houses the administrative offices of the local
Lombardy region and sits next to the city's central train station.

[18] It is one of the world's tallest concrete buildings, and one of the
 main symbols of Milan.

[19] At his lunchtime press briefing today, White House press
secretary Ari Fleischer told reporters: "The president just
moments ago was informed about the incident in Milan.

[20] I have no additional information at all for you at this time.

[21] This is a breaking story and we don't have anything else beyond
 that; the president has been informed."

[22] "I think you can presume that we will be; if we are not already;
 in touch with Italian authorities and will ascertain precisely what
the facts are," he said.

[23] The world has been on edge about airplane crashes into buildings
since the Sept. 11 terror attacks.

[24] On Jan. 5, a 15-year-old boy crashed a stolen plane into a
building in Tampa, Fla.

[25] He was the only casualty.

[26] April 18; The 32-story Pirelli Building, struck today by a
small airplane, is Milan's tallest building, and is one of the
world's highest concrete skyscrapers.

[27] Built in the 1950s, the office building stands 32 stories
tall, or about 417 feet, according to the Web site,
SkyScraperPage.com.

[28] By comparison, New York's Empire State Building is 102
stories, or 1,455 feet.

[29] Centrally located near the city's train station, the
Pirelli Building was designed by the architect Gio Ponti,
and one of Italy's most important 20th-century structural engineers,
Pier Luigi Nervi (1891-1979),known for his use of reinforced concrete.

[30]  Copyright  2002 ABCNEWS Internet Ventures.
```

```
Document 25
Source: Fox News
Publication: April 18, 2002 at 1:44 p.m.

[1] FOXNews.com

[2] MILAN, Italy   A small plane crashed into the tallest
building in downtown Milan Thursday evening,
causing smoke to pour out of the top floors of the skyscraper.

[3] Three people were reported to have been killed, and rescue
workers said dozens of people had been taken to a nearby hospital.

[4] Officials said the crash appeared to have been an accident
and that only the pilot was aboard.

[5] The weather was clear at the time of the crash.

[6] Police officer Celerissimo De Simone said the pilot of the
Piper aircraft had sent out a distress call at 5:50 p.m. just
before the crash near Milan's main train station.

[7] RAI state TV reported that the pilot said the SOS was
because of engine trouble.

[8] The plane had taken off from Locarno, Switzerland, and
was heading to Milan's Linate airport, De Simone said.

[9] Earlier, in Rome, the senate's president, Marcello Pera,
said it "very probably" appeared to be a terrorist attack.

[10] But Pera's spokesman later said he had spoken with the
Interior Minister and the crash didn't appear to be any kind
of an attack.

[11] There were holes on both sides of the slim skyscraper.

[12] "It sounded like a bomb.

[13] The pavement shook like an earthquake," said a woman
 identifying herself only as Lucia.

[14] "It was shocking," said Luccheta Antonio, 52, a barber
 down the block.

[15] "The windows shook and the mirrors."

[16] "It was a violent explosion," said Stefano Bottazzi, 35,
who works in a skyscraper 500 yards from building.

[17] "The clock fell to the floor."
```

```
Document 25, con't

[18] On the streets, rescue workers in orange uniforms helped
the injured including a man with a bloody
shirt holding his head.

[19] Ambulances streamed into the area and pedestrians
peered upward.

[20] The crash put a hole in the 25th floor of the Pirelli building.

[21] Police cordoned off the area as people gawked
at the skyscraper.

[22] At 30 stories high, the Pirelli skyscraper, located near
the central train station, is Italy's first skyscraper and one
of the world's tallest concrete buildings.

[23] It was built in 1958 and designed by architects Gio Ponti
and Pier Luigi Nervi.

[24] The building is one of the main symbols of Milan, along
with the city's cathedral.

[25] White House press secretary Ari Fleischer said National
Security Adviser Condoleezza Rice and White House chief of
staff Andrew Card had informed Bush of the crash.

[26] "The president just moments ago was informed about the
incident in Milan," he said at his afternoon press briefing.

[27] "I have no additional information at all for you at this time.

[28] This is a breaking story and we don't have anything else
beyond that the president has been informed.

[29] "I think you can presume that we will be if we are not
already in touch with Italian authorities and will ascertain
precisely what the facts are," Fleischer said.

[30] In Washington, the FBI was trying to learn more about
the incident.

[31] One official said bureau personnel were assisting their
Italian counterparts in the investigation.
```

```
Document 25, con't

[32] It was the second time since the Sept. 11 terror attacks
on New York and Washington that a plane has struck a high-rise
building.

[33] On Jan. 5, a 15-year-old boy crashed a stolen plane into
a building in Tampa, Fla.

[34] He was the only casualty.

[35] The Associated Press contributed to this report.

[36] Privacy Statement.

[37] For FoxNews.com comments write to foxnewsonline
foxnews.com comments foxnews.com http://www.foxsports.com/
Associated Press.

[38] All rights reserved.

[39] This material may not be published, broadcast,
rewritten, or redistributed.

[40] Fox News Network, LLC 2002.

[41] All rights reserved.

[42] All market data delayed 20 minutes.

[43] Outside the U.S.

[44] U.S.

[45] Regional Maps / Forecasts   U.S.
```

```
Document 36
Source: ABC
Publication: April 18, 2002 at unknown time

[1] ABCNEWS.com : Plane Slams Into Milan Skyscraper

[2] The Pirelli Building in Milan, Italy, was hit by
a small plane.

[3] (ABCNEWS.com) A small plane crashed into the 25th
floor of a skyscraper in downtown Milan today.

[4] At least three people, including the pilot, were
dead, Italy's ANSA wire service said.

[5] Dozens of people in the Pirelli building were injured
after several floors of the 32-story building caught
fire, local reports said.

[6] At least 15 were taken to the hospital.

[7] Only the pilot was on board the plane, reported The
Associated Press.

[8] The weather was clear at the time of the crash.

[9] The crash happened during rush hour, as office workers
were ending their day, so few people were in the building
when it happened.

[10] The building has since been cleared, and the fire is
now under control, reported ABCNEWS' Rome bureau.

[11] It damaged the 23rd through the 29th floors, which
were being restored and weren't occupied at the time,
local reports said.

[12] A man who was on the 21st floor at the time the
plane hit said the building was evacuated in an orderly
fashion.

[13] The Pirelli building houses the administrative
offices of the local Lombardy region and sits next to
the city's central train station.

[14] A small tourist plane crashed into the tallest
building in Milan.

[15] The plane was destined for Italy's capital Rome,
and departed from Lucerne airport in Switzerland,
sources from the National Air Safety board told ANSA.

[16] It was the second time since the Sept. 11 terror
attacks that a plane has struck a high-rise building,
and the crash raised fears of another attack.
```

```
Document 36, con't

[17] But local officials now say it was probably
an accident.

[18] A transport official told Reuters the plane
had reported problems with its undercarriage and
was circling the city ahead of trying to land at
a local airport.

[19] Police officer Celerissimo De Simone told
The AP the pilot of the aircraft had sent out a
distress call at 5:54 p.m. just before crashing.

[20] Italian TV said the pilot's SOS involved a
technical problem.

[21] Police say the plane was an Air Commando;
a small plane similar to a Piper.

[22] Initial reports described the plane as a Piper,
but did not note the specific model.

[23] According to ABCNEWS aviation expert John Nance,
Piper planes have no history of mechanical troubles
or other problems that would lead a pilot to lose control.

[24] "Those aircraft have direct mechanical controls,"
 says Nance.

[25] "If you have a control problem, it would be the
loss of an engine, and those are either controllable
or not."

[26] U.S. officials told ABCNEWS they had no indication
the crash was a terrorist attack.

[27] Neither American nor Italian forces were on
heightened alert, they said.

[28] The FBI has offered help in the crash investigation,
but it remains to be seen if Italy will accept.

[29] At his lunchtime press briefing today, White House
press secretary Ari Fleischer told reporters:
"The president just moments ago was informed about the
incident in Milan.

[30] I have no additional information at all for you
at this time.
```

```
Document 36, con't

[31] This is a breaking story and we don't have anything
 else beyond that; the president has been informed."

[32] "I think you can presume that we will be; if we
are not already; in touch with Italian authorities and
will ascertain precisely what the facts are," he said.

[33] Italian Prime Minister Silvio Berlusconi was on his
way back to Italy from Bulgaria, and a crisis
center is being set up.

[34] Airspace at the local Linate airport was reportedly
cleared, but Milan's Malpensa international airport,
further away from the city center, was not affected.

[35] On Jan. 5, a 15-year-old boy crashed a stolen plane
into a building in Tampa, Fla.

[36] He was the only casualty.

[37] ABCNEWS' Rome bureau and John McWethy at the Pentagon
contributed to this report.

[38] April 18; The Pirelli building, struck
today by a small airplane, is Milan's tallest building,
and is one of the world's highest concrete skyscrapers.

[39] Built in the late 1950s, the office building stands
32 stories tall, or about 417 feet, according to the Web
site, SkyScraperPage.com.

[40] Other sources say the building is 30 stories tall.

[41] By comparison, Paris' Eiffel Tower is about twice
as tall, and New York's Empire State Building
is 102 stories, or 1,455 feet tall.

[42] Nevertheless, with its location near the city's central
train station and its height relative to the buildings around
it, the Pirelli building is considered a landmark and symbol of
the commercial city of Milan.

[43] The building was designed by the architect Gio Ponti,
along with one of Italy's most important 20th-century structural
engineers, Pier Luigi Nervi (1891-1979), known for his use of
reinforced concrete.

[44] Pirelli, the tire and cable company, does not operate
out of the Pirelli building.

[45] However, the office building houses the government
administration of the Lombardy region of Italy.

[46] Copyright   2002 ABCNEWS Internet Ventures.
```

```
Document 43
Source: Fox News
Publication: April 18, 2002 at 3:36 p.m.

[1] FOXNews.com

[2] Thursday: Smoke billows from the Pirelli building
after a small plane crashed into it.

[3] MILAN, Italy   A small plane smashed into the
tallest building in Milan Thursday evening, causing
smoke to pour out of the heavily damaged top floors
of the skyscraper.

[4] Three people were reported killed and dozens of
others were injured.

[5] There were initial fears that the midtown crash
was a terrorist attack   it was the second time since
Sept. 11 that a plane has struck a high-rise building
but officials later said that it appeared to have been
a tragic accident.

[6] "It sounded like a bomb.

[7] The pavement shook like an earthquake," said a woman
identifying herself only as Lucia.

[8] Some eyewitnesses reported that the plane was on
fire before crashing into the 30-story Pirelli
building, Milan Police Office Celerissimo De Simone said.

[9] The Rockwell Commander plane, en route from
Switzerland on a 20-minute flight to Milan, punched
a hole in the 25th floor of the building, sparking a
smoky fire that was quickly put out.

[10] Rescuers helped bloodied men in business suits
evacuate the building.

[11] The weather was clear at the time of the crash,
which occurred near the end of the work day and
left gaping holes on both sides of the slim skyscraper.

[12] A large section of an entire floor lost its walls,
and smoke and liquid poured from the gash in one side
of the building.

[13] "The initial information that the Interior Ministry
has leads us to lean toward an accident,"
Interior Minister Claudio Scajola said.

[14] The pilot, identified as 75-year-old Luigi Sasulo
of Pregassona, Switzerland, had sent out a distress
call at 5:54 p.m. just before the crash near Milan's main
train station, said De Simone.
```

```
Document 43, con't

[15] RAI state TV reported that the pilot said he was
experiencing engine trouble.

[16] "We believe it isn't a terrorist attack," said police Sgt.

[17] Vincenzo Curto, who was reached at the Carabinieri
headquarters in Milan.

[18] "The pilot might have taken ill or it was an engine
problem."

[19] In Rome, a spokesman for the senate president, Marcello
Pera, said the interior minister had informed him that the crash
didn't appear to be a terror attack.

[20] Earlier, Pera had said it "very probably" was an attack.

[21] The plane had taken off from Locarno, Switzerland, 50 miles
northwest of Milan, and was heading to Milan's Linate airport.

[22] Patrick Herr, spokesman for the Swiss air traffic control
office SKYGUIDE, said the plane left Locarno at 5.15 p.m.

[23] A woman who worked on the eighth floor, well below the
crash, said she saw 10 people injured and bleeding.

[24] News reports said at least 30 were taken to the hospital.

[25] One Milan hospital, Fatebene Fratelli, said it had
received 20 injured, including a woman with burns.

[26] "It was shocking," said Luccheta Antonio, 52, a barber
down the block.

[27] "The windows shook, and the mirrors."

[28] "It was a violent explosion," said Stefano Bottazzi,
35, who works in a skyscraper 500 yards from building.

[29] "The clock fell to the floor."

[30] On the streets, rescue workers in orange uniforms
helped the injured.

[31] Ambulances streamed into the area and pedestrians
peered upward.

[32] As ambulance crews worked, a man with his shirt
splattered with blood and his hand covering a gash
on his head was rushed from the scene.
```

```
Document 43, con't

[33] Police cordoned off the area as passersby gawked
at the skyscraper.

[34] At 30 stories high, the Pirelli structure, located
near the central train station, is Italy's first skyscraper
and one of the world's tallest concrete buildings.

[35] It was built in 1958 and designed by architects Gio
Ponti and Pier Luigi Nervi.

[36] The building is one of the main symbols of Milan,
along with the city's cathedral.

[37] The skyscraper, built of concrete and glass with
a diamond-shaped floor plan, has inspired design around
the world including the MetLife building in New York.

[38] National Security Adviser Condoleezza Rice and
White House chief of staff Andrew Card broke the news
of the crash to President Bush, White House press
secretary Ari Fleischer said.

[39] "I think you can presume that we will be
if we are not already in touch with Italian authorities
and will ascertain precisely what the facts are," he said.

[40] FBI personnel were assisting their Italian counterparts
in the investigation, an FBI official said.

[41] U.S. authorities had no intelligence suggesting any
kind of terrorist attack was imminent in Milan, a U.S.
official said.

[42] On March 27, the State Department issued a warning
for American citizens traveling in four Italian cities,
including Milan, during Easter.

[43] The warning said the possible threat was based on
information about "extremist groups."

[44] It was the second time since the Sept. 11 terror
attacks that a plane has struck a high-rise building.

[45] On Jan. 5, a 15-year-old boy flying alone crashed
a stolen plane into a building in Tampa, Fla.

[46] The boy, Charles Bishop, left behind a suicide
note saying that Al Qaeda terrorists had tried to recruit
him, but police said there was no truth to the claim.

[47] Relatives of the boy, who was the only fatality, have
filed a lawsuit claiming the acne drug Accutane was behind
his suicide.
```

```
Document 43, con't

[48] U.S. officials have called a mosque and cultural center
in Milan "the main Al Qaeda station house in Europe."

[49] Since Sept. 11, several individuals have been arrested
in Milan as part of a crackdown on suspected Islamic militants.

[50] Italian authorities uncovered an alleged Al Qaeda plot
in January, 2001, to attack the U.S.

[51] Embassy in Rome and are investigating whether a second
plot was in the works earlier this year.

[52] The Associated Press contributed to this report.

[53] Privacy Statement.

[54] For FoxNews.com comments write to foxnewsonline
foxnews.com comments foxnews.com
http://www.foxsports.com/  Associated Press.

[55] All rights reserved.

[56] This material may not be published, broadcast,
rewritten, or redistributed.

[57] Fox News Network, LLC 2002.

[58] All rights reserved.

[59] All market data delayed 20 minutes.

[60] Crash Brought Memories of Sept. 11   Outside the U.S.?

[61] U.S.

[62] Regional Maps / Forecasts   U.S.
```

3.0.2 CST relationships judgements

Table 3.0.2 shows some of the CST relationship judgements for the source documents in the example cluster. Note that CST relationships are not mutually exclusive - there may be more than one relation for any given pair of sentences. Also note that not all pairs of sentences will be related. CST relationships will be referred to by their associated numbers given previously in the description section.

3.0.3 Discussion of Examples

In the above example annotation, two judges, working independently, found 15 pairs of sentences in three source documents that had CST relationships. Let us discuss the reasons each of these relations were assigned.

1. [2,2] [16,3]: Paraphrase (2) and Attribution (9)

(S1) CNNenEspanol.com A small plane has hit a skyscraper in central Milan, setting the top floors of the 30-story building on fire, an Italian journalist told CNN.

Source doc	Source sent num	Target doc	Target sent num	CST relation(s)
2	2	16	3	2,9
2	2	19	3	2
2	3	16	4	4,5,9
2	3	19	8	5
2	4	16	9	2
2	6	19	9	4
2	8	16	7	2
2	10	16	11	2
2	16	16	10	6
16	3	2	5	4
16	6	19	9	4
16	7	2	8	4,9
19	3	16	3	2,9
19	4	2	12	4
19	11	2	12	4

Table 3.1: CST Judgement Example

(S2) ABCNEWS.com A small plane crashed into a skyscraper in downtown Milan today, setting several floors of the 30-story building on fire.

CST type: 2, 9
Directionality: -, >

The two sentences contain the same information, with only slight differences in their wordings. In addition, S1, which is the second sentence from document 2, attributes its information to a source, specifically "an Italian journalist." S2, the third sentence from document 16, does not contain any such attribution, but rather simply states the facts.

2. [2,2] [19,3]: Paraphrase (2)

(S1) CNNenEspanol.com A small plane has hit a skyscraper in central Milan, setting the top floors of the 30-story building on fire, an Italian journalist told CNN.

(S2) A small airplane crashed into a government building in heart of Milan, setting the top floors on fire, Italian police reported.

CST type: 2
Directionality: -

The two sentences contain the same information, that a small plane crashed into a building in Milan and set it on fire. Also, since both sentences contain an attribution, we can label the relation as paraphrase alone.

3. [2,3] [16,4]: Subsumption (4), Contradiction (5) and Attribution (9)

(S1) The crash by the Piper tourist plane into the 26th floor occurred at 5:50 p.m. (1450 GMT) on Thursday, said journalist Desideria Cavina.

(S2) The plane crashed into the 25th floor of the Pirelli building in downtown Milan.

CST type: 4, 5, 9
Directionality: >, -, >

S1 contains more information than S2 since it indicates the time of the crash as well as the kind of plane involved in the incident. Therefore, there is subsumption. The sentences contradict one another as to the floor of the building that the plane crashed into. Finally, S1 attributes its information to a journalist, Desideria Cavina, while S2 contains no such attribution.

4. [2,3] [19,8]: Contradiction (5)

(S1) The crash by the Piper tourist plane into the 26th floor occurred at 5:50 p.m. (1450 GMT) on Thursday, said journalist Desideria Cavina.

(S2) Italian state television said the crash put a hole in the 25th floor of the Pirelli building.

CST type: 5
Directionality: -

The sentences contradict one another as to the floor of the building that was hit by the plane. Since they both contain attributions, the only CST relation assigned is that of contradiction.

5. [2,4] [16,9]: Paraphrase (2)

(S1) The building houses government offices and is next to the city's central train station.

(S2) The Pirelli building houses the administrative offices of the local Lombardy region and sits next to the city's central train station.

CST type: 2
Directionality: -

Both sentences contain information about the Pirelli building, the offices it houses and its location, so the CST relationships is paraphrase.

6. [2,6] [19,9]: Subsumption (4)

(S1) Italian TV showed a hole in the side of the Pirelli building with smoke pouring from the opening.

(S2) News reports said smoke poured from the opening.

CST type: 4
Directionality: >

Both sentences contain an attribution. However, S1 contains more information than S2, in particular that the hole is in the side of the Pirelli building. Therefore, the appropriate relationship is subsumption.

7. [2,8] [16,7]: Paraphrase (2)

(S1) Earlier though, in Rome, the senate's president, Marcello Pera, said it "very probably" appeared to be a terrorist attack.

(S2) The president of the Italian Senate, Marcello Pera, told Italian television it "very probably" appeared to be a terrorist attack but soon afterwards his spokesman said it was probably an accident.

CST type: 2
Directionality: -

The information contained in S1 and S2 is approximately the same. However, S2 contains more information than S1, since it indicates the believes about the situation at a later time. In addition, S2 attributes its information to Italian television. (Note that the relations of subsumption and attribution are indicated in example 12 rather than here. This is because these relations have directionality and must go from [16,7] to [2,8] rather than the other way around.)

8. [2,10] [16,11]: Paraphrase (2)

(S1) Many people were on the streets as they left work for the evening at the time of the crash.

(S2) The crash happened just before rush hour, as office workers were closing their day.

CST type: 2
Directionality: -

The sentences contain the same information, only they are worded differently.

9. [2,16] [16,10]: Historical Background (6)

(S1) It is one of the symbols of Italy's financial capital and is one of the world's tallest concrete buildings, designed between 1955 and 1960.

(S2) It is constructed of concrete and glass.

CST type: 6
Directionality: >

Although both sentences describe the Pirelli building, S1 provides historical background information as to when the building was designed.

10. [16,3] [2,5]: Subsumption (4)

(S1) ABCNEWS.com A small plane crashed into a skyscraper in downtown Milan today, setting several floors of the 30-story building on fire.

(S2) Several storeys of the building were engulfed in fire, she said.

CST type: 4
Directionality: >

Although both sentences indicate that the building was on fire, S1 provides additional information as to the reason it was on fire (a plane crashed into it) as well as where the accident took place.

11. [16,6] [19,9]: Subsumption (4)

(S1) Smoke poured from the opening as police and ambulances rushed to the area.

(S2) News reports said smoke poured from the opening.

CST type: 4
Directionality: >

Both sentences indicate that smoke was coming out of the building. However, S1 also indicates that emergency personel rushed to the scene.

12. [16,7] [2,8]: Paraphrase (2) and Attribution (9)

(S1) The president of the Italian Senate, Marcello Pera, told Italian television it ”very probably” appeared to be a terrorist attack but soon afterwards his spokesman said it was probably an accident.

(S2) Earlier though, in Rome, the senate’s president, Marcello Pera, said it ”very probably” appeared to be a terrorist attack.

CST type: 2, 9
Directionality: -,>

See example 7.

13. [19,3] [16,3]: Paraphrase (2) and Attribution (9)

(S1) A small airplane crashed into a government building in heart of Milan, setting the top floors on fire, Italian police reported.

(S2) ABCNEWS.com A small plane crashed into a skyscraper in downtown Milan today, setting several floors of the 30-story building on fire.

CST type: 2, 9
Directionality: -,>

The sentences contain the same information, although S1 contains an attribution as well.

14. [19,4] [2,12]: Subsumption (4)

(S1) There were no immediate reports on casualties as rescue workers attempted to clear the area in the city s financial district.

(S2) There is no word yet on casualties.

CST type: 4
Directionality: >

Both S1 and S2 state that there is no word yet on casualties, but S1 contains additional information about the attempts of the rescue workers to clear the scene, as well as the location of the incident.

15. [19,11] [2,12]: Subsumption (4)

(S1) No further details were immediately available.

(S2) There is no word yet on casualties.

CST type: 4
Directionality: >

S1 suggests that no details were available (including details about casualties). Therefore, the relationship is that of subsumption, since S1 contains more information than S2.

3.0.4 Tricky CST Judgements

Historical Background (6) versus Description (15)

The crucial difference between the Historical Background relation and the Description relation, is that the former holds when one sentence provides *historical* information about an entity, situation or event described in the second sentence. To contrast, Description holds when one sentence describes an *entity* that is mentioned in the other sentence. The information contained in the descriptive sentence is that of a current nature.

Examples:

[36-2]: The Pirelli Building in Milan, Italy, was hit by a small plane.
[21-16]: It was built in 1958 and designed by architects Gio Ponti and Pier Luigi Nervi.

CST type: 6
Directionality: $<$

[36-5]: Dozens of people in the Pirelli building were injured after several floors of the 32-story building caught fire, local reports said.
[16-9]: The Pirelli building houses the administrative offices of the local Lombardy region and sits next to the city's central train station.

CST type: 15
Directionality: $<$

[16-9]: The Pirelli building houses the administrative offices of the local Lombardy region and sits next to the city's central train station.
[21-13]: The crash put a hole in the 25th floor of the Pirelli building, and smoke was seen pouring from the opening.

CST type: 15
Directionality: $>$

In the first pair of sentences, the second sentence, [21-16], provides historical information about an entity, the Pirelli Building. Specifically, the sentence states when the building was built and who designed it. These are historical facts about the building. To contrast, in the second pair of sentences, the second sentence, [16-9], provides a description of the entity, the Pirelli building, mentioned in the first sentence. The information given is descriptive rather than historical because it provides current information about this entity, specifically, what is housed in the building and its location. Likewise, in the third pair of sentences, sentence [16-9] gives current, descriptive information about the Pirelli building, which is the subject of the second sentence, [21-13].

Elaboration/Refinement (13) versus Contradiction (5) and Subsumption (4)

In sentence pairs that exhibit the Elaboration relationship, one sentence contains specific details not contained in the other sentence. However, they both convey the same main idea. In addition, it may be the case that the more detailed sentence was published after more details about a breaking news story were made available. Such sentence pairs should be marked as having an Elaboration relationship, rather than as Contradiction. In order to contradict one another, two sentences must have been published at the same time and contain conflicting information. In the Subsumption relationship, one sentence conveys only one idea, X, while the second sentence contains ideas X and Y. Therefore, we say that the second sentence subsumes the first.
Examples:

[36-20]: Italian TV said the pilot's SOS involved a technical problem.
[43-15]: RAI state TV reported that the pilot said he was experiencing engine trouble.

CST type: 13
Directionality: <

In the above sentence pair, the 'technical problem' mentioned in the first sentence becomes more specific, 'engine trouble' in the second sentence.

[36-4]: At least three people, including the pilot, were dead, Italy's ANSA wire service said.
[23-4]: At least two people are dead, including the pilot, Italy's ANSA wire service said.

CST type: 13
Directionality: >

Follow-up (11) versus Change of Perspective (17)

In sentence pairs that have a relationship of Follow-up, one should have been published later than the first and should contain details about a story that were not yet available when the first sentence was published. To contrast, in a Change of Perspective relationship, the facts that are believed to be true about a situation or breaking story should have changed. In this relationship, the first sentence should reflect what was believed at time one, while the second sentence should report what was believed to be true at the second point in time.

[36-28]: The FBI has offered help in the crash investigation, but it remains to be seen if Italy will accept.
[43-40]: FBI personnel were assisting their Italian counterparts in the investigation, an FBI official said.

CST type: 11
Directionality: <

The relationship between the above pair of sentences is clearly that of Follow-up. In the first sentence, it was not clear if the FBI would be helping Italy. However, by the time the second sentence was published, Italy had accepted the FBI's assistance.

[36-26]: U.S. officials told ABCNEWS they had no indication the crash was a terrorist attack.
[25-9]: Earlier, in Rome, the senate's president, Marcello Pera, said it "very probably" appeared to be a terrorist attack.

CST type: 17
Directionality: >

In the above sentence pair, the relationship is that of a Change of Perspective. In the second sentence, it was believed that the plane crash was a terrorist attack. However, in the first sentence, which was published after the second one, it is believed that the crash was not an attack.

3.0.5 Practice Sentences

(1)
[36-1]: ABCNEWS.com : Plane Slams Into Milan Skyscraper
[1-1]: CNN.com - Plane hits skyscraper in Milan - April 18, 2002

CST type:
Directionality:

(2)
[36-2]: The Pirelli Building in Milan, Italy, was hit by a small plane.
[25-20]: The crash put a hole in the 25th floor of the Pirelli building.

CST type:
Directionality:

(3)
[36-2]: The Pirelli Building in Milan, Italy, was hit by a small plane.
[25-23]: It was built in 1958 and designed by architects Gio Ponti and Pier Luigi Nervi.

CST type:
Directionality:

(4)
[36-2]: The Pirelli Building in Milan, Italy, was hit by a small plane.
[16-4]: The plane crashed into the 25th floor of the Pirelli building in downtown Milan.

CST type:
Directionality:

(5)
[23-2]: The Pirelli Building in Milan, Italy, was hit by a small plane.
[2-2]: CNNenEspanol.com A small plane has hit a skyscraper in central Milan, setting the top floors of the 30-story building on fire, an Italian journalist told CNN.

CST type:
Directionality:

(6)
[23-5]: The plane crashed into the 25th floor of the Pirelli building in downtown Milan.
[1-3]: The crash by the Piper tourist plane into the 26th floor occurred at 5:50 p.m. (1450 GMT) on Thursday, said journalist Desideria Cavina.

CST type:
Directionality:

(7)
[25-24]: The building is one of the main symbols of Milan, along with the city's cathedral.
[23-26]: The 32-story Pirelli Building, struck today by a small airplane, is Milan's tallest building, and is one of the world's highest concrete skyscrapers.

CST type:
Directionality:

(8)
[25-24]: The building is one of the main symbols of Milan, along with the city's cathedral.
[23-18]: It is one of the world's tallest concrete buildings, and one of the main symbols of Milan.

CST type:
Directionality:

(9)
[25-28]: This is a breaking story and we don't have anything else beyond that; the president has been informed.
[21-20]: This is a breaking story and we don't have anything else beyond that."

CST type:
Directionality:

(10)
[16-3]: A small plane crashed into a skyscraper in downtown Milan today, setting several floors of the 30-story building on fire.
[1-2]: CNNenEspanol.com A small plane has hit a skyscraper in central Milan, setting the top floors of the 30-story building on fire, an Italian journalist told CNN.

CST type:
Directionality:

(11)
[16-6]: Smoke poured from the opening as police and ambulances rushed to the area.
[19-9]: News reports said smoke poured from the opening.

CST type:
Directionality:

(12)
[23-13]: A transport official told Reuters the plane had reported problems with its undercarriage and was circling the city ahead of trying to land at a local airport.
[2-7]: RAI state TV reported that the plane had apparently radioed an SOS because of engine trouble.

CST type:
Directionality:

(13)
[25-4]: Officials said the crash appeared to have been an accident and that only the pilot was aboard.
[2-8]: Earlier though, in Rome, the senate's president, Marcello Pera, said it ”very probably” appeared to be a terrorist attack.

CST type:
Directionality:

(14)
[25-9]: Earlier, in Rome, the senate's president, Marcello Pera, said it ”very probably” appeared to be a terrorist attack.
[43-19]: In Rome, a spokesman for the senate president, Marcello Pera, said the interior minister had informed him that the crash didn't appear to be a terror attack.

CST type:
Directionality:

(15)
[25-8]: The plane had taken off from Locarno, Switzerland, and was heading to Milan's Linate airport, De Simone said.
[43-21]: The plane had taken off from Locarno, Switzerland, 50 miles northwest of Milan, and was heading to Milan's Linate airport.

CST type:
Directionality:

Chapter 4

Examples in XML

Below is the XML version of the CST judgements for three of the Milan documents (the MI-3 set). Such XML files are called sentrel files.


```
<?xml version='1.0'?>
<!DOCTYPE TABLE SYSTEM "/clair/projects/cst/gulfair/judgement/sentrel.dtd" >
<TABLE>
<R SDID='2' SSENT='2' TDID='16' TSENT='3'>
<RELATION TYPE='2' JUDGE='J'/>
<RELATION TYPE='9' JUDGE='J'/>
<RELATION TYPE='2' JUDGE='Z'/>
<RELATION TYPE='9' JUDGE='Z'/>
</R>
<R SDID='2' SSENT='2' TDID='19' TSENT='3'>
<RELATION TYPE='2' JUDGE='J'/>
<RELATION TYPE='2' JUDGE='Z'/>
</R>
<R SDID='2' SSENT='3' TDID='16' TSENT='4'>
<RELATION TYPE='4' JUDGE='J'/>
<RELATION TYPE='5' JUDGE='J'/>
<RELATION TYPE='9' JUDGE='J'/>
<RELATION TYPE='5' JUDGE='Z'/>
</R>
<R SDID='2' SSENT='4' TDID='16' TSENT='9'>
<RELATION TYPE='4' JUDGE='J'/>
<RELATION TYPE='2' JUDGE='Z'/>
</R>
<R SDID='16' SSENT='3' TDID='2' TSENT='5'>
<RELATION TYPE='4' JUDGE='J'/>
<RELATION TYPE='4' JUDGE='Z'/>
</R>
<R SDID='2' SSENT='6' TDID='16' TSENT='4'>
<RELATION TYPE='4' JUDGE='J'/>
<RELATION TYPE='9' JUDGE='J'/>
</R>
```

```
Judgement file, con’t

<R SDID=’2’ SSENT=’6’ TDID=’19’ TSENT=’9’>
<RELATION TYPE=’4’ JUDGE=’J’/>
<RELATION TYPE=’4’ JUDGE=’Z’/>
</R>
<R SDID=’2’ SSENT=’8’ TDID=’16’ TSENT=’7’>
<RELATION TYPE=’4’ JUDGE=’J’/>
<RELATION TYPE=’2’ JUDGE=’Z’/>
</R>
<R SDID=’16’ SSENT=’7’ TDID=’2’ TSENT=’8’>
<RELATION TYPE=’4’ JUDGE=’J’/>
<RELATION TYPE=’4’ JUDGE=’Z’/>
</R>
<R SDID=’2’ SSENT=’10’ TDID=’16’ TSENT=’11’>
<RELATION TYPE=’2’ JUDGE=’J’/>
<RELATION TYPE=’2’ JUDGE=’Z’/>
</R>
<R SDID=’19’ SSENT=’4’ TDID=’2’ TSENT=’12’>
<RELATION TYPE=’4’ JUDGE=’J’/>
<RELATION TYPE=’4’ JUDGE=’Z’/>
</R>
<R SDID=’2’ SSENT=’16’ TDID=’16’ TSENT=’10’>
<RELATION TYPE=’6’ JUDGE=’J’/>
<RELATION TYPE=’15’ JUDGE=’J’/>
<RELATION TYPE=’6’ JUDGE=’Z’/>
<RELATION TYPE=’15’ JUDGE=’Z’/>
</R>
<R SDID=’16’ SSENT=’3’ TDID=’19’ TSENT=’3’>
<RELATION TYPE=’2’ JUDGE=’J’/>
<RELATION TYPE=’9’ JUDGE=’J’/>
<RELATION TYPE=’9’ JUDGE=’Z’/>
</R>
<R SDID=’2’ SSENT=’3’ TDID=’19’ TSENT=’8’>
<RELATION TYPE=’4’ JUDGE=’J’/>
<RELATION TYPE=’5’ JUDGE=’J’/>
<RELATION TYPE=’5’ JUDGE=’Z’/>
</R>
<R SDID=’19’ SSENT=’11’ TDID=’2’ TSENT=’12’>
<RELATION TYPE=’4’ JUDGE=’Z’/>
</R>
<R SDID=’16’ SSENT=’6’ TDID=’19’ TSENT=’9’>
<RELATION TYPE=’4’ JUDGE=’J’/>
</R>
</TABLE>
```

Chapter 5

Examples of summaries

Below are two examples of automatically produced summaries. Both were produced with MEAD using the cluster of documents that describe the Milan plane crash. The first one is a 10 percent summary and the second, a 20 percent summary. Both summaries are raw, and have not undergone any CST-enhancement procedure.

```
10% summary of MI-3 documents

[1]  The Pirelli Building in Milan, Italy, was hit by a small plane.
[2]  A small plane crashed into a skyscraper in downtown Milan today,
setting several floors of the 30-story building on fire.
[3]  The plane crashed into the 25th floor of the Pirelli building
in downtown Milan.
[4]  A small airplane crashed into a government building in heart of
Milan, setting the top floors on fire, Italian police reported.
[5]  CNN.com - Plane hits skyscraper in Milan - April 18, 2002
```

```
20% summary of MI-3 documents

[1]  The Pirelli Building in Milan, Italy, was hit by a small plane.
[2]  A small plane crashed into a skyscraper in downtown Milan today,
setting several floors of the 30-story building on fire.
[3]  The plane crashed into the 25th floor of the Pirelli building in
downtown Milan.
[4]  The Pirelli building houses the administrative offices of the local
Lombardy region and sits next to the city’s central train station.
[5]  A small airplane crashed into a government building in heart of Milan,
setting the top floors on fire, Italian police reported.
[6]  WITNESSES REPORTED hearing a loud explosion from the 30-story office
building, which houses the administrative offices of the local Lombardy
region and sits next to the city s central train station.
[7]  CNN.com - Plane hits skyscraper in Milan - April 18, 2002
[8]  The crash by the Piper tourist plane into the 26th floor occurred at
5:50 p.m. (1450 GMT) on Thursday, said journalist Desideria Cavina.
[9]  The building houses government offices and is next to the city’s
central train station.
```

Bibliography

[Rad00] Dragomir Radev. A common theory of information fusion from multiple text sources, step one: Cross-document structure. In *Proceedings, 1st ACL SIGDIAL Workshop on Discourse and Dialogue*, Hong Kong, October 2000.

[ZBGR02] Zhu Zhang, Sasha Blair-Goldensohn, and Dragomir Radev. Towards CST-enhanced summarization. AAAI 2002, August 2002.