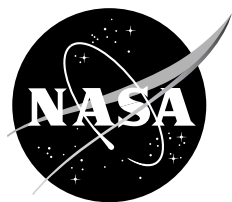


NASA/TM-2005-213469



Using Perilog to Explore “Decision Making at NASA”

*Michael W. McGreevy
Ames Research Center
Moffett Field, California*

October 2005

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the Lead Center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

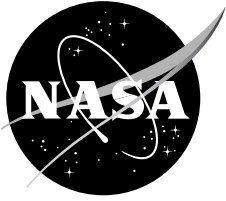
- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at <http://www.sti.nasa.gov>
- E-mail your question via the Internet to help@sti.nasa.gov
- Fax your question to the NASA Access Help Desk at (301) 621-0134
- Telephone the NASA Access Help Desk at (301) 621-0390
- Write to:
NASA Access Help Desk
NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076-1320

NASA/TM-2005-213469



Using Perilog to Explore “Decision Making at NASA”

*Michael W. McGreevy
Ames Research Center
Moffett Field, California*

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

October 2005

Available from:

NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
(703) 487-4650

Table of Contents

SUMMARY	1
INTRODUCTION.....	1
APPLYING PERILOG	1
FROM FLIGHT TO A PIVOTAL DECISION.....	2
VOCABULARY REVIEW	2
KEYWORD-IN-CONTEXT SEARCH	3
PHRASE REVIEW.....	4
PHRASE SEARCH	4
SUMMARY OF THIS EXAMPLE	5
FROM FOAM TO IN-FLIGHT ANOMALY 1983-2003	5
VOCABULARY REVIEW	5
PHRASE GENERATION.....	5
PHRASE SEARCH	5
ABOUT PHRASE MATCHING	6
SEARCH BY EXAMPLE (1 OF 2)	6
SEARCH BY EXAMPLE (2 OF 2)	7
SUMMARY OF THIS EXAMPLE	7
CONDITIONS AND CONTINGENCIES: IF P THEN Q	7
EXAMPLE OF IF P THEN Q.....	8
ANALYSIS OF THE ASSERTION	9
INTERPRETATION OF THE LOGIC OF THE ASSERTION.....	11
SUMMARY REGARDING "IF-THEN" ANALYSIS.....	12
CONCLUSION	12
REFERENCES.....	12
APPENDIX 1. PROCESSING THE CAIB TEXT FOR INPUT TO PERILOG	14
OBTAINING THE PDF FILE AND CONVERTING IT TO TEXT.....	14
EDITING PROBLEMATIC OUTPUT.....	14
CHARACTER PROCESSING	16
EDITING TEXT TO CREATE UNITS OF TEXT BASED ON PARAGRAPHS.....	16
LABELING UNITS OF TEXT	17
SELECTING A STOPLIST	18
APPENDIX 2. EXAMPLES OF PASSAGES ABOUT BELIEF.	20
APPENDIX 3. EXCERPTS CONTAINING THE WORD "IF", SORTED ON SOURCE PARAGRAPHS.....	21
APPENDIX 4. EXCERPTS CONTAINING THE WORD "IF", SORTED ON EXCERPTS.....	22

List of Figures

FIGURE 1. USING PERILOG TOOLS IN CONCERT, SHOWING TYPICAL SEQUENCES OF TOOL USE...	2
FIGURE 2. ANALYSIS OF THE ASSERTION: "IF IT [I.E., FOAM] HIT THE RCC, [THEN] WE DON'T SEE ANY [SAFETY-OF-FLIGHT] ISSUE"	10

Using Perilog to Explore “Decision Making at NASA”

MICHAEL W. MCGREEVY

NASA Ames Research Center

Summary

Perilog, a context intensive text mining system, is used as a discovery tool to explore topics and concerns in "Decision Making at NASA," chapter 6 of the Columbia Accident Investigation Board (CAIB) Report Volume 1. Two examples illustrate how Perilog can be used to discover highly significant safety-related information in the text without prior knowledge of the contents of the document. A third example illustrates how "if-then" statements found by Perilog can be used in logical analysis of decision making. In addition, in order to serve as a guide for future work, the technical details of preparing a PDF document for input to Perilog are included in an appendix.

Introduction

On February 1, 2003, Space Shuttle Columbia on STS-107 disintegrated during re-entry, killing its crew of seven astronauts, grounding the Space Shuttle, calling into question NASA's decision making and organizational culture, and testing the commitment of the United States to human space flight.

The Columbia Accident Investigation Board released its final report on the causes of the accident on August 26, 2003, citing both physical and organizational causes. According to the report, decision making during the mission was a central factor among the causes of the accident. The CAIB concluded that "Management decisions made during Columbia's final flight reflect missed opportunities, blocked or ineffective communications channels, flawed analysis, and ineffective leadership." (CAIB 2003: 170).

Upon reading the CAIB report, it becomes clear that it is a very rich repository of highly detailed information of fundamental importance to safety at NASA. In order to maximize the benefit of this and similar documents, it can be helpful to apply text mining methods to them, so that essential details and patterns among them are readily accessible. For example, NASA's Perilog text mining methods and software have been shown to be very helpful in exploring large collections of aviation incident reports and finding incident reports relevant to accident reports (e.g., McGreevy 2001; McGreevy and Statler 1998). As

an illustration of Perilog's utility as a discovery tool and an aid to analysis of decision making, it has been applied here to a chapter of the CAIB report.

The CAIB report is available as electronic text in Portable Document Format (PDF). Given the limitations of converting complex PDF files to text, and the desire to reliably label paragraphs with page numbers, the text preparation is somewhat demanding. Consequently, a single chapter of the CAIB report, "Decision Making at NASA" (chapter 6), was selected to serve as a test case. The details of the text preparation are described in appendix 1.

After an introduction to Perilog, the emphasis of this paper is on using Perilog to explore "Decision Making at NASA." The first two examples illustrate how Perilog can be used to find highly significant safety-related information in the text, even without prior knowledge of the contents of the document. The third example shows how searching for the word "if" in the context of words importantly related to it, and investigating statements of the form "if p then q," can be of benefit to a logical analysis of decision making.

Applying Perilog

Perilog is a collection of patented text mining methods (McGreevy 2004a-d, McGreevy 2001) that have been created and implemented in software by the author. The methods are based on context intensive analysis, modeling, and relevance ranking of text to exploit latent linguistic structure in seemingly unstructured text (McGreevy 2005). The methods include keyword-in-context search, flexible phrase search, phrase generation, phrase discovery, and search by example. They are supplemented by phrase review and vocabulary review.

A collection of these text mining tools has been integrated by the author into a user-friendly Perilog software package that runs on Unix and has a browser-based graphical user interface for accessibility from any computer platform. This package has been in operational use at the NASA/FAA Aviation Safety Reporting System and at a major airline since early 2002. A second major airline began to use Perilog in late 2004. A Perilog server has been used via network by a dozen researchers at NASA Ames Research Center since early 2003. Perilog

methods and software have also been distributed under government agreements and licenses to private companies, universities, and federal laboratories since mid-2001.

Perilog has been extensively applied to aviation safety incident narratives (McGreevy 2001, McGreevy 1997, McGreevy 1996) and to accident reports (McGreevy & Statler 1998). Other applications of Perilog have involved text as diverse as industrial safety incident reports, financial disclosures, political speeches, and Space Shuttle maintenance reports.

Application of Perilog to chapter 6 of the CAIB report is a departure from more typical applications to large collections of incident narratives in plain text format. The CAIB report is topically and graphically complex, with the main text augmented by many illustrations, captions, sidebars, email excerpts, viewgraphs, and the like. Consequently, converting it from PDF to coherent plain text for input to Perilog is not straightforward. Further, it is necessary to make decisions about what constitutes a unit of text within a large document and how best to label these units, and then to implement these decisions. In addition, preparation of any text for input to Perilog requires careful consideration of so-called "stopwords."

Given this complexity, a single chapter of the report, "Decision Making at NASA" (chapter 6), was selected and processed in order to develop and document a method of preparing the text for input to Perilog. That method is documented in appendix 1. Not only does the method apply to the CAIB report, it also applies to many other PDF documents, especially large and complex ones.

Perilog's processing of text is extensively documented in McGreevy (2001), so that is not discussed here. Further, Perilog's graphical user interface is thoroughly documented in McGreevy (2003), and that document is provided to all users of the Perilog system, so those details are not addressed here, either. Instead, the emphasis of this paper is on illustrating the use of Perilog for knowledge discovery.

The first two examples illustrate the synergy of the Perilog tools and their use in discovering information about the contents of a database, in this case, one containing the paragraphs of a chapter of the CAIB report. Figure 1 suggests some typical sequences of Perilog tool use, several of which are illustrated in these examples.

From flight to a pivotal decision

This example shows how Perilog can be used to quickly obtain the answers to these questions:

- What topics are prominent in the document?

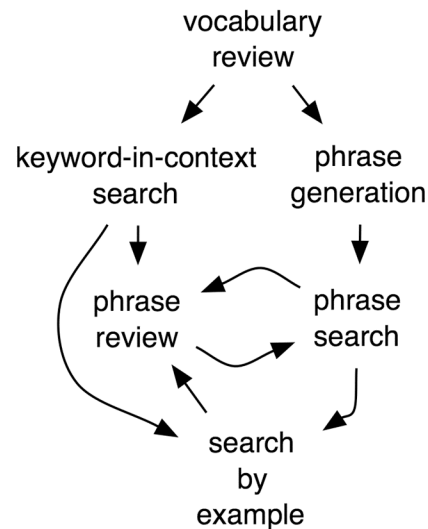


Figure 1. Using Perilog tools in concert, showing typical sequences of tool use, several of which are illustrated in the text.

- What are the most relevant sections of the document regarding a particular topic?
- What concerns are associated with this topic?
- What are the most relevant sections of the document regarding particular concerns?

Vocabulary review

When searching a database of text, it is essential to be familiar with its vocabulary. The Perilog vocabulary review tool provides that familiarity. It can be used to list the words that are most commonly found in the database, or words containing particular patterns of letters. The output of vocabulary review does not include so-called stopwords (such as "the", "and", "an").

Shown below are the most prominent 26 words (in one case, a word fragment) in the CAIB chapter on "Decision Making at NASA," and their frequencies of occurrence.

313	FLIGHT	142	MANAGEMENT
310	FOAM	128	MANAGERS
291	DEBRIS	128	ORBITER
264	MISSION	116	BIPOD
248	NOT	116	LAUNCH
239	NASA	115	NO
217	'S	108	STRIKE
212	TEAM	107	COLUMBIA
192	SHUTTLE	107	LOSS
191	SPACE	104	ANALYSIS
174	DAMAGE	102	SAFETY
165	PROGRAM	101	EXTERNAL
143	WOULD	101	TANK

This list of vocabulary words provides a sense of the main topics in the chapter and their relative prominence. Note that "flight" is the most prominent word, followed by

"foam", "debris", and "mission". The list also includes words like "not", "would", and "no" since they are often found in narratives involving problematic situations. The listed item "'S" (apostrophe S) is a word fragment that typically indicates possession. (Note that Perilog currently processes all text as upper case.)

Keyword-in-context search

Using the most frequently occurring word, "flight", as a query to Perilog's keyword-in-context search, 156 relevant paragraphs are found and presented in order of their relevance to "flight" in context, that is, "flight" in the context of words importantly related to it. Here is the most relevant paragraph (having capitalization and extra spacing introduced by Perilog) with instances of the word "flight" highlighted:

THE THIRD BIPOD FOAM LOSS OCCURRED ON JUNE 25 , 1992 , DURING THE LAUNCH OF COLUMBIA ON STS-50 , WHEN AN APPROXIMATELY 26 - BY 10-INCH PIECE SEPARATED FROM THE LEFT BIPOD RAMP AREA . POST - MISSION INSPECTION REVEALED A 9-INCH BY 4.5-INCH BY 0.5-INCH DIVOT IN THE TILE , THE LARGEST AREA OF TILE DAMAGE IN SHUTTLE HISTORY . THE EXTERNAL TANK PROJECT AT MARSHALL SPACE **FLIGHT** CENTER AND THE INTEGRATION OFFICE AT JOHNSON SPACE CENTER CITED SEPARATE IN - **FLIGHT** ANOMALIES . THE INTEGRATION OFFICE CLOSED OUT ITS IN - **FLIGHT** ANOMALY TWO DAYS BEFORE THE NEXT **FLIGHT** , STS-46 , BY DEEMING DAMAGE TO THE THERMAL PROTECTION SYSTEM AN " ACCEPTED **FLIGHT** RISK . " 12 IN INTEGRATION HAZARD REPORT 37 , THE INTEGRATION OFFICE NOTED THAT THE IMPACT DAMAGE WAS SHALLOW , THE TILE LOSS WAS NOT A RESULT OF EXCESSIVE AERODYNAMIC LOADS , AND THE EXTERNAL TANK THERMAL PROTECTION SYSTEM FAILURE WAS THE RESULT OF " INADEQUATE VENTING . " 13 THE EXTERNAL TANK PROJECT CLOSED OUT ITS IN - **FLIGHT** ANOMALY WITH THE RATIONALE THAT FOAM LOSS DURING ASCENT WAS " NOT CONSIDERED A **FLIGHT** OR SAFETY ISSUE . " 14 NOTE THE DIFFERENCE IN HOW EACH PROGRAM ADDRESSED THE FOAM - SHEDDING PROBLEM : WHILE THE INTEGRATION OFFICE DEEMED IT AN " ACCEPTED RISK , " THE EXTERNAL TANK PROJECT CONSIDERED IT " NOT A SAFETY - OF - **FLIGHT** ISSUE . " HAZARD REPORT 37 WOULD FIGURE IN THE STS-113 **FLIGHT** READINESS REVIEW , WHERE THE CRUCIAL DECISION WAS MADE TO CONTINUE FLYING WITH THE FOAM - LOSS PROBLEM . THIS INCONSISTENCY WOULD REAPPEAR 10 YEARS LATER , AFTER BIPOD FOAM - SHEDDING DURING STS-112 . (CAIB 2003: 124)

The above paragraph, which is highly relevant to "flight", repeatedly refers to foam loss and foam shedding in the context of flight: as separate In-Flight Anomalies cited by the External Tank Project at Marshall Space Flight Center and by the Integration Office at JSC; as an anomaly that was closed prior to the next flight (STS-46); as an accepted flight risk; as a closed out in-flight anomaly that

was "not considered a flight or safety issue" and "not a safety-of-flight issue;" and as the subject of a report that "would figure in the STS-113 Flight Readiness Review."

This "flight" paragraph immediately suggests that "foam" is a significant concern. The prominence of "foam" in the vocabulary list shown earlier strongly supports this suggestion. By using "foam" as a new query to Perilog's keyword-in-context search, the 121 relevant paragraphs are easily found and sorted on relevance. For example, here is the most relevant paragraph (grouped with its immediately preceding section header, as explained in appendix 1), with instances of the word "foam" highlighted:

BIPOD RAMP **FOAM** LOSS EVENTS - CHUNKS OF **FOAM** FROM THE EXTERNAL TANK ' S FORWARD BIPOD ATTACHMENT , WHICH CONNECTS THE ORBITER TO THE EXTERNAL TANK , ARE SOME OF THE LARGEST PIECES OF DEBRIS THAT HAVE STRUCK THE ORBITER . TO PLACE THE **FOAM** LOSS FROM STS-107 IN A BROADER CONTEXT , THE BOARD EXAMINED EVERY KNOWN INSTANCE OF **FOAM** - SHEDDING FROM THIS AREA . **FOAM** LOSS FROM THE LEFT BIPOD RAMP (CALLED THE - Y RAMP IN NASA PARLANCE) HAS BEEN CONFIRMED BY IMAGERY ON 7 OF THE 113 MISSIONS FLOWN . HOWEVER , ONLY ON 72 OF THESE MISSIONS WAS AVAILABLE IMAGERY OF SUFFICIENT QUALITY TO DETERMINE LEFT BIPOD RAMP **FOAM** LOSS . THEREFORE , **FOAM** LOSS FROM THE LEFT BIPOD AREA OCCURRED ON APPROXIMATELY 10 PERCENT OF FLIGHTS (SEVEN EVENTS OUT OF 72 IMAGED FLIGHTS) . ON THE 66 FLIGHTS THAT IMAGERY WAS AVAILABLE FOR THE RIGHT BIPOD AREA , **FOAM** LOSS WAS NEVER OBSERVED . NASA COULD NOT EXPLAIN WHY ONLY THE LEFT BIPOD EXPERIENCED **FOAM** LOSS . (SEE FIGURE 6 . 1-1) (CAIB 2003:123)

The above paragraph is interpreted by Perilog as being highly relevant because the word "foam" is strongly contextually associated with words like "loss", "bipod", "ramp", "left", "shedding", "area", "external", "events", "tank", "STS-107", "chunks", "observed", etc. These contextual associations are situationally important in the CAIB chapter on "Decision Making at NASA," and are automatically made part of the query network used by Perilog when the user searches on the word "foam". That is why this search method is called keyword-in-context search. It does not merely find the keyword in the text, but finds the keyword in the contexts of situationally important words. Further, the user can enter multiple keywords, if desired.

Perilog's keyword-in-context provides the option to highlight just the query terms in relevant text, as above, or to highlight the query terms and their contextually associated terms. Here is the same paragraph as above, but this time with "foam" and its contextually associated terms highlighted.

BIPOD RAMP FOAM LOSS EVENTS - CHUNKS OF FOAM FROM THE **EXTERNAL TANK 'S FORWARD BIPOD** ATTACHMENT , WHICH CONNECTS THE ORBITER TO THE EXTERNAL TANK , ARE SOME OF THE LARGEST PIECES OF DEBRIS THAT HAVE STRUCK THE ORBITER . TO **PLACE** THE **FOAM LOSS** FROM **STS-107** IN A **BROADER CONTEXT** , THE **BOARD EXAMINED EVERY KNOWN INSTANCE OF FOAM - SHEDDING** FROM THIS **AREA** . **FOAM LOSS** FROM THE **LEFT BIPOD RAMP** (CALLED THE - Y **RAMP** IN NASA PARLANCE) HAS BEEN CONFIRMED BY IMAGERY ON 7 OF THE 113 MISSIONS FLOWN . HOWEVER , ONLY ON 72 OF THESE MISSIONS WAS AVAILABLE IMAGERY OF SUFFICIENT QUALITY TO **DETERMINE LEFT BIPOD RAMP FOAM LOSS** . **THEREFORE , FOAM LOSS** FROM THE **LEFT BIPOD AREA** OCCURRED ON APPROXIMATELY 10 PERCENT OF FLIGHTS (SEVEN EVENTS OUT OF 72 IMAGED FLIGHTS) . ON THE 66 FLIGHTS THAT IMAGERY WAS AVAILABLE FOR THE RIGHT **BIPOD AREA** , **FOAM LOSS** WAS **NEVER OBSERVED** . **NASA COULD NOT EXPLAIN WHY** ONLY THE **LEFT BIPOD EXPERIENCED FOAM LOSS** . (SEE FIGURE 6.1-1 .) (CAIB 2003:123)

The automatic highlighting shown above makes it very easy for the reader to see quickly what concerns are relevant to "foam".

Phrase review

Another way to identify concepts importantly related to "foam" is to extract and list phrases from relevant text using Perilog's phrase review tool. Here are phrases that occur at least twice in the preceding "foam" paragraph.

- 7 FOAM LOSS
- 4 LEFT BIPOD
- 3 BIPOD RAMP
- 2 BIPOD AREA
- 2 BIPOD RAMP FOAM LOSS
- 2 EXTERNAL TANK
- 2 LEFT BIPOD RAMP

Shown below are the phrases that occur at least 5 times in the 25 paragraphs that are most relevant to "foam". The 25 paragraphs are the most relevant among the 121 paragraphs found earlier by a keyword-in-context search on the word "foam".

- 50 FOAM LOSS
- 35 BIPOD RAMP
- 30 EXTERNAL TANK
- 15 BIPOD FOAM
- 15 BIPOD RAMP FOAM
- 15 FLIGHT ANOMALY
- 15 LEFT BIPOD
- 12 LEFT BIPOD RAMP
- 10 EXTERNAL TANK PROJECT
- 10 FLIGHT READINESS REVIEW
- 10 FOAM - SHEDDING
- 10 SHUTTLE PROGRAM
- 9 BIPOD RAMP FOAM LOSS
- 7 BIPOD FOAM LOSS
- 7 FLIGHT ISSUE

- 7 FOAM STRIKE
- 7 SAFETY - OF - FLIGHT
- 7 SAFETY OF FLIGHT
- 7 THERMAL PROTECTION SYSTEM
- 6 LEFT BIPOD RAMP FOAM
- 6 PROGRAM REQUIREMENTS CONTROL BOARD
- 6 RIGHT BIPOD
- 5 ACCEPTED RISK
- 5 BIPOD LOSS
- 5 INTEGRATION OFFICE
- 5 KNOWN BIPOD
- 5 STS-113 FLIGHT READINESS REVIEW

Phrase search

The above list of phrases, extracted from text relevant to "foam", suggests query phrases that can be used to find text relevant to particular "foam"-related concerns. Many of the phrases involve the physical circumstances of the accident, such as "foam loss" and "bipod ramp". Others represent broad safety concerns, such as "safety of flight" and "accepted risk". Two other phrases stand out due to their length, specificity, and references to the human and organizational circumstances of the accident. These are: "program requirements control board" and "STS-113 flight readiness review".

When a Perilog phrase search is performed using the latter two phrases as a query, 35 relevant paragraphs are found. Shown below is the most relevant paragraph (including its immediately preceding section title, as explained in appendix 1):

STS-113 FLIGHT READINESS REVIEW : A PIVOTAL DECISION - BECAUSE THE BIPOD RAMP SHEDDING ON STS-112 WAS SIGNIFICANT , BOTH IN SIZE AND IN THE DAMAGE IT CAUSED , AND BECAUSE IT OCCURRED ONLY TWO FLIGHTS BEFORE STS-107 , THE BOARD INVESTIGATED NASA 'S RATIONALE TO CONTINUE FLYING . THIS DECISION MADE BY THE **PROGRAM REQUIREMENTS CONTROL BOARD** AT THE **STS-113 FLIGHT READINESS REVIEW** IS AMONG THOSE MOST DIRECTLY LINKED TO THE STS-107 ACCIDENT . HAD THE FOAM LOSS DURING STS-112 BEEN CLASSIFIED AS A MORE SERIOUS THREAT , MANAGERS MIGHT HAVE RESPONDED DIFFERENTLY WHEN THEY HEARD ABOUT THE FOAM STRIKE ON STS-107 . ALTERNATELY , IN THE FACE OF THE INCREASED RISK , STS-107 MIGHT NOT HAVE FLOWN AT ALL . HOWEVER , AT **STS-113 'S FLIGHT READINESS REVIEW** , MANAGERS FORMALLY ACCEPTED A FLIGHT RATIONALE THAT STATED IT WAS SAFE TO FLY WITH FOAM LOSSES . THIS DECISION ENABLED , AND PERHAPS EVEN ENCOURAGED , MISSION MANAGEMENT TEAM MEMBERS TO USE SIMILAR REASONING WHEN EVALUATING WHETHER THE FOAM STRIKE ON STS-107 POSED A SAFETY - OF - FLIGHT ISSUE . (CAIB 2003:125)

The retrieved text is highly relevant to concerns about "foam" in the context of "program requirements control board" and "STS-113 flight readiness review". It describes "a pivotal decision," despite foam losses, to

continue flying. As stated in the above CAIB paragraph, "This decision made by the Program Requirements Control Board at the STS-113 Flight Readiness Review is among those most directly linked to the STS-107 accident." (CAIB 2003:125).

(Note that Perilog's flexible phrase search recognizes the match between the phrase "STS-113's flight readiness review" with the possessive "STS-113's" and the query phrase "STS-113 flight readiness review" without the possessive. While this might seem unremarkable, many state-of-the-art phrase search tools cannot recognize near-matches.)

Summary of this example

Using Perilog, it was just a few steps from first identifying "flight" and "foam" as significant words to finding an important passage about decision making. The example does not suggest, however, that this sequence of searches is the only or the best way to explore the text. One might well have started with a keyword-in-context search on "decision", a key word from the chapter title. The important point here is that significant insights can be gained by using Perilog text mining tools in concert, following one of the suggested sequences illustrated in figure 1.

In this example, the vocabulary review tool identified the two most prominent words as "flight" and "foam". Keyword-in-context searches found paragraphs that are highly relevant to "flight" and "foam" in context. Phrase review found a collection of phrases that are contextually and situationally associated with "foam". Finally, the selected phrases "program requirements control board" and "STS-113 flight readiness review" were used as queries in Perilog's flexible phrase search to find text relevant to these "foam"-related concerns.

Those already familiar with the findings of the CAIB report will not be surprised that "foam" is highly associated with "foam loss", "bipod ramp", "external tank", "safety of flight", "thermal protection system", "program requirements control board", "STS-113 flight readiness review", etc. Further, careful readers will already appreciate the important connection between the Program Requirements Control Board and the STS-113 Flight Readiness Review. What Perilog offers even the most careful reader, however, is a tool kit for achieving this level of insight even when documents are unfamiliar, very long, and very detailed. This example illustrates how quickly and easily it can be done.

From foam to in-flight anomaly 1983-2003

Another sequence of Perilog tool use is illustrated in this second example: vocabulary review, phrase generation, phrase search, and search by example.

Vocabulary review

As observed earlier, a review of the vocabulary of the CAIB chapter on "Decision Making at NASA" shows the word "foam" to be particularly prominent. (See table on page 2.)

Phrase generation

To gain a better understanding of a document, it can be helpful to find specific concepts involving a particularly prominent word such as "foam". Perilog's phrase generation tool provides this assistance. Shown below are the top 20 generated phrases containing the word "foam".

102	FOAM STRIKE
102	FOAM LOSS
102	BIPOD FOAM
70	FOAM SHEDDING
67	RAMP FOAM
64	BIPOD RAMP FOAM
57	BIPOD FOAM LOSS
43	STS-112 FOAM
32	FOAM DEBRIS
30	FOAM LOSSES
30	ET FOAM
29	RAMP FOAM LOSS
29	FOAM DEBRIS STRIKE
22	FOAM STRIKES
21	FOAM DEBRIS STRIKES
20	STS-112 FOAM LOSS
15	MISSING FOAM
14	FOAM OCCURRED
14	FOAM EVENT
14	BIPOD FOAM SHEDDING

Generated phrases are not extracted from the text, but are inferred from a network model of the text. As a consequence, the output of phrase generation is a list of phrases that are likely to occur in the text, with the more likely phrases appearing toward the top of the list. The numerical values are a rough estimate of frequency within the entire database of paragraphs. In addition, the phrase list is compressed, so that, for example, "foam shedding" represents both "foam shedding" and "foam-shedding".

These phrases provide significant insights into the nature of "foam"-related concerns in the document. They can be used directly to discover, for example, that "foam strike", "foam loss", and "bipod foam" are among the greatest "foam"-related concerns. The other phrases provide additional specificity. Thus, phrase generation converts a simple input word to a list of specific concepts involving the word.

Phrase search

As a next step, any number of these phrases can be used as a query to find paragraphs containing them. To do so, the phrases can be copied, with or without the numerical values, from the output of phrase generation to the input of phrase search.

When the first three phrases, "foam strike", "foam loss", and "bipod foam" are used as a query to phrase search, 84 relevant paragraphs are found. The most relevant of these paragraphs is the following:

COLUMBIA : CAUGHT IN THE MIDDLE - STS-112 FLEW IN OCTOBER 2002 . AT 33 SECONDS INTO THE FLIGHT , A PIECE OF THE **BIPOD FOAM** FROM THE EXTERNAL TANK STRUCK ONE OF THE SOLID ROCKET BOOSTERS . AS DESCRIBED IN SECTION 6.1 , THE STS-112 **FOAM STRIKE** WAS DISCUSSED AT THE PROGRAM REQUIREMENTS CONTROL BOARD FOLLOWING THE FLIGHT . ALTHOUGH THE INITIAL RECOMMENDATION WAS TO TREAT THE **FOAM LOSS** AS AN IN - FLIGHT ANOMALY , THE SHUTTLE PROGRAM INSTEAD ASSIGNED IT AS AN ACTION , WITH A DUE DATE AFTER THE NEXT LAUNCH . (THIS WAS THE FIRST INSTANCE OF **BIPOD FOAM LOSS** THAT WAS NOT DESIGNATED AN IN - FLIGHT ANOMALY .) THE ACTION WAS NOTED AT THE STS-113 FLIGHT READINESS REVIEW . THOSE FLIGHT READINESS REVIEW CHARTS (SEE SECTION 6.1) PROVIDED A FLAWED FLIGHT RATIONALE BY CONCLUDING THAT THE **FOAM LOSS** WAS " NOT A SAFETY - OF - FLIGHT " ISSUE . (CAIB 2003:137)

When the top 20 phrases returned by phrase generation are used as a query to phrase search, along with their numerical weights, 136 paragraphs are found to be relevant. The following text, which is from a table (CAIB 2003: 128, figure 6.1-7), is the most relevant. It summarizes bipod foam loss incidents dating from 1983 to 2003.

MISSION DATE COMMENTS STS-1 APRIL 12 , 1981 LOTS OF DEBRIS DAMAGE . 300 TILES REPLACED . STS-7 JUNE 18 , 1983 FIRST KNOWN LEFT **BIPOD RAMP FOAM SHEDDING EVENT** . STS-27R DECEMBER 2 , 1988 DEBRIS KNOCKS OFF TILE ; STRUCTURAL DAMAGE AND NEAR BURN THROUGH RESULTS . STS-32R JANUARY 9 , 1990 SECOND KNOWN LEFT **BIPOD RAMP FOAM EVENT** . STS-35 DECEMBER 2 , 1990 FIRST TIME NASA CALLS **FOAM DEBRIS** " SAFETY OF FLIGHT ISSUE , " AND " RE - USE OR TURNAROUND ISSUE . " STS-42 JANUARY 22 , 1992 FIRST MISSION AFTER WHICH THE NEXT MISSION (STS-45) LAUNCHED WITHOUT DEBRIS IN - FLIGHT ANOMALY CLOSURE / RESOLUTION . STS-45 MARCH 24 , 1992 DAMAGE TO WING RCC PANEL 10-RIGHT . UNEXPLAINED ANOMALY , " MOST LIKELY ORBITAL DEBRIS . " STS-50 JUNE 25 , 1992 THIRD KNOWN **BIPOD RAMP FOAM EVENT** . HAZARD REPORT 37 : AN " ACCEPTED RISK . " STS-52 OCTOBER 22 , 1992 UNDETECTED **BIPOD RAMP FOAM LOSS** (FOURTH BIPOD EVENT) . STS-56 APRIL 8 , 1993 ACREAGE TILE DAMAGE (LARGE AREA) . CALLED " WITHIN EXPERIENCE BASE " AND CONSIDERED " IN FAMILY . " STS-62 OCTOBER 4 , 1994 UNDETECTED **BIPOD RAMP FOAM LOSS** (FIFTH BIPOD EVENT) . STS-87 NOVEMBER 19 , 1997 DAMAGE TO ORBITER THERMAL PROTECTION SYSTEM SPURS NASA TO BEGIN 9 FLIGHT TESTS TO RESOLVE **FOAM - SHEDDING** . FOAM FIX INEFFECTIVE . IN - FLIGHT ANOMALY

EVENTUALLY CLOSED AFTER STS-101 AS " ACCEPTED RISK . " STS-112 OCTOBER 7 , 2002 SIXTH KNOWN LEFT **BIPOD RAMP FOAM LOSS** . FIRST TIME MAJOR DEBRIS EVENT NOT ASSIGNED AN IN - FLIGHT ANOMALY . EXTERNAL TANK PROJECT WAS ASSIGNED AN ACTION . NOT CLOSED OUT UNTIL AFTER STS-113 AND STS-107 . STS-107 JANUARY 16 , 2003 COLUMBIA LAUNCH . SEVENTH KNOWN LEFT **BIPOD RAMP FOAM LOSS EVENT** . (CAIB 2003:128)

(See appendix 1 regarding table formats and units of text. While all of the text from the table is treated as a single paragraph in this exercise, it could have been treated as multiple paragraphs when setting up the database.)

About phrase matching

In the above CAIB paragraph, phrases that match the query phrases are highlighted. For example, the highlighted phrase "foam debris" directly matches the query phrase "foam debris", and the highlighted phrase "foam-shedding" (with a dash) flexibly matches the query phrase "foam shedding" (without a dash). Sometimes two or more short matches reveal longer phrases, as when phrases in a paragraph match query phrases by overlapping one or more of the query words. For example, this long, highlighted phrase from the paragraph:

bipod ramp foam event

matches the following query phrases (spaced for word-by-word alignment with the long phrase above):

bipod foam
ramp foam
bipod ramp foam
foam event

With longer phrases, matching can be even more complex. For example, this long, highlighted phrase from the paragraph:

bipod ramp foam shedding event

flexibly matches the following query phrases (spaced for word-by-word alignment with the long phrase above):

bipod foam
foam shedding
ramp foam
bipod ramp foam
foam event
bipod foam shedding

These results illustrate the power and benefit of flexible matching in Perilog phrase search.

Search by example (1 of 2)

Any text, such as a CAIB report paragraph containing "foam" related text, can be used as a query to Perilog's search by example. One option is to base the search on all

of the phrase structure in the query text (not just the highlighted phrases). For example, the paragraph found earlier, from a table which summarizes bipod foam losses from 1983 to 2003, can be used as a query. That paragraph, shown in its entirety on page 6, begins:

MISSION DATE COMMENTS STS-1 APRIL 12 , 1981 LOTS OF DEBRIS DAMAGE . 300 TILES REPLACED . STS-7 JUNE 18 , 1983 FIRST KNOWN LEFT **BIPOD RAMP FOAM SHEDDING EVENT** . (CAIB 2003:128)

A search by example using the entire paragraph as a query finds 199 relevant paragraphs. The most relevant paragraph is, of course, the one used as a query. The second most relevant paragraph is shown below. It describes the first bipod foam loss incident in 1983 and shows that it was considered to be an in-flight anomaly at the time. This paragraph is most like the query paragraph with respect to the matching phrases it contains. The phrases highlighted by Perilog in this paragraph are also found in some form in the query paragraph.

THE **FIRST KNOWN BIPOD RAMP FOAM LOSS** OCCURRED DURING STS-7 , CHALLENGER 'S SECOND MISSION (SEE FIGURE 6.1-2) . IMAGES TAKEN AFTER **EXTERNAL TANK** SEPARATION REVEALED THAT A 19 - BY 12-INCH PIECE OF THE **LEFT BIPOD RAMP** WAS MISSING , AND THAT THE **EXTERNAL TANK** HAD SOME 25 SHALLOW DIVOTS IN THE FOAM JUST FORWARD OF THE BIPOD STRUTS AND ANOTHER 40 DIVOTS IN THE FOAM COVERING THE LOWER **EXTERNAL TANK** . AFTER THE MISSION WAS COMPLETED , THE PROGRAM REQUIREMENTS CONTROL BOARD CITED THE **FOAM LOSS** AS AN IN - **FLIGHT ANOMALY** . CITING AN EVENT AS AN IN - **FLIGHT ANOMALY** MEANS THAT BEFORE THE NEXT LAUNCH , A SPECIFIC NASA ORGANIZATION MUST RESOLVE THE PROBLEM OR PROVE THAT IT DOES NOT THREATEN THE SAFETY OF THE VEHICLE OR CREW . (CAIB 2003:123)

This paragraph not only contains "foam" phrases from the full text of the query paragraph, including "first known bipod ramp foam loss", "left bipod ramp", and "foam loss", it also contains the matching phrases "external tank" and "flight anomaly", which are related to the topic of "foam" but do not contain the word "foam".

The longer phrase "in-flight anomaly" is only matched as "flight anomaly" in this example because the word "in" was included among the stopwords. Phrases whose first or last word is a stopword are properly accounted for in the Perilog methods, as documented in the patents, but they remain to be fully addressed by the current implementation in software.

Search by example (2 of 2)

The above CAIB paragraph, which describes the first bipod foam loss incident, can itself be used as a query to search by example. The following paragraph is found to

be the most relevant (after the original query paragraph). It is about the most recent foam loss incident, on STS-107, and indicates that the occurrence was considered to be an in-flight anomaly.

THE SEVENTH AND FINAL **KNOWN BIPOD RAMP FOAM LOSS OCCURRED** ON JANUARY 16 , 2003 , DURING THE LAUNCH OF COLUMBIA ON STS-107 . AFTER THE COLUMBIA **BIPOD LOSS** , THE **PROGRAM REQUIREMENTS CONTROL BOARD** DEEMED THE **FOAM LOSS** AN IN - **FLIGHT ANOMALY** TO BE DEALT WITH BY THE **EXTERNAL TANK** PROJECT . (CAIB 2003:126)

This and the previously retrieved paragraphs show that despite the characterization of bipod ramp foam loss as an in-flight anomaly as early as STS-7, launched June 18, 1983, it was again characterized as an in-flight anomaly on the ill-fated flight of Columbia on STS-107 nearly 20 years later. During the intervening years, despite numerous similar incidents of foam loss, NASA did not "resolve the problem or prove that it does not threaten the safety of the vehicle or crew." (CAIB 2003: 123)

Summary of this example

This significant insight was quickly achieved by using Perilog tools in concert. The word "foam" was found by vocabulary review to be prominent, and phrase generation produced a list of prominent phrases containing the word "foam". These phrases were input to phrase search to find paragraphs in which the "foam" phrases are prominent. The most relevant of these paragraphs, which summarizes bipod foam loss incidents, was then used in its entirety as input to search by example in order to find paragraphs that not only contain the previously identified "foam" phrases but are also relevant to associated concerns such as "flight anomaly" and "external tank". The most relevant of the retrieved paragraphs describes bipod ramp foam loss on STS-7 in 1983 as an in-flight anomaly. This paragraph was itself used as another input to search by example to find a paragraph showing that bipod ramp foam loss was found to be an in-flight anomaly on STS-107 in 2003. Thus, in a series of simple steps, Perilog found significant safety-critical concerns about "foam" within the context of "Decision Making at NASA."

Conditions and contingencies: *if p then q*

While the two previous examples demonstrate how Perilog tools can be used in concert for knowledge discovery, this final example demonstrates Perilog's utility in a different way. One difference is that the previous examples use Perilog to find text passages that are highly relevant to various kinds of things present in the world, such as "foam", "bipod ramp foam shedding event", and "Program Requirements Control Board". In contrast, this final example uses Perilog to find instances of a particular kind of statement about the world,

conditional statements, and shows how one such statement might be evaluated. Conditional statements, which are typically found in the form "if p then q," are an essential part of the structure of decision making.

Further, while the first two examples concentrate on the Perilog tools and the text that they retrieve, this final example concentrates on a detailed analysis of a single example of a conditional statement found by Perilog. The purpose of this example is to show the potential value of finding and analyzing conditional statements. Rather than merely continuing to show the already demonstrated power of Perilog to retrieve highly relevant text, this example shows how that power can support analysis of decision making.

A Perilog keyword-in-context search on the word "if" in the CAIB report chapter "Decision Making at NASA" is useful for finding important conditional statements, and the conditions and contingencies that they contain, which are factors in decision making. These factors can be used in modeling the network of logical inferences that led to particular decisions. This example shows a small but representative part of such a network.

Performing the keyword-in-context search using the word "if" as a query produces paragraphs ranked on relevance to "if" in the context of words like "ask", "asked", "asking", "launch", "determine", "could", "see", "damage", "RCC", and "managers". For example, here are the two most relevant paragraphs, with occurrences of "if" highlighted.

MCCORMACK : " RIGHT , IT COULD POTENTIALLY HIT THE RCC AND WE DON'T INDICATE ANY OTHER POSSIBLE COATING DAMAGE OR SOMETHING , WE DON'T SEE ANY ISSUE **IF** IT HIT THE RCC . ALTHOUGH WE COULD HAVE SOME SIGNIFICANT TILE DAMAGE **IF** WE DON'T SEE A SAFETY - OF - FLIGHT ISSUE . " (CAIB 2003:161)

ACCORDING TO BOEING ANALYSTS WHO WERE MEMBERS OF THE DEBRIS ASSESSMENT TEAM , SCHOMBURG CALLED TO ASK ABOUT THEIR RATIONALE FOR PURSUING IMAGERY . THE BOEING ANALYSTS TOLD HIM THAT SOMETHING THE SIZE OF A LARGE COOLER HAD HIT THE ORBITER AT 500 MILES PER HOUR . PRESSED FOR ADDITIONAL REASONS AND NOT FULLY UNDERSTANDING WHY THEIR ORIGINAL JUSTIFICATION WAS INSUFFICIENT , THE ANALYSTS SAID THAT AT LEAST THEY WOULD KNOW WHAT HAPPENED **IF** SOMETHING WERE TO GO TERRIBLY WRONG . THE BOEING ANALYSTS NEXT ASKED WHY THEY WERE WORKING SO HARD ANALYZING POTENTIAL DAMAGE AREAS **IF** SHUTTLE PROGRAM MANAGEMENT BELIEVED THE DAMAGE WAS MINOR AND THAT NO SAFETY - OF - FLIGHT ISSUES EXISTED . SCHOMBURG REPLIED THAT THE ANALYSTS WERE NEW AND WOULD LEARN FROM THIS EXERCISE . (CAIB 2003:160)

Text relevant to the word "if" in context, such as these two CAIB report paragraphs, can be useful, as shown below, for: 1) identifying conditions and contingencies that are explicitly at issue in a document; 2) comparing the explicitly given contingencies with implicit contingencies; 3) building decision networks; and 4) evaluating the validity of conditional statements relating conditions and contingencies.

Example of if p then q

As an example, the passage from the first paragraph, "we don't see any issue if it hit the RCC", will be analyzed here in some detail. For this exercise, detailed analysis will not include the entire passage, but only the part that is underlined here: "IT COULD POTENTIALLY HIT THE RCC AND WE DON'T INDICATE ANY OTHER POSSIBLE COATING DAMAGE OR SOMETHING , WE DON'T SEE ANY ISSUE IF IT HIT THE RCC . ALTHOUGH WE COULD HAVE SOME SIGNIFICANT TILE DAMAGE IF WE DON'T SEE A SAFETY - OF - FLIGHT ISSUE . " It could be argued that the meaning of the whole passage is, "If it hit the RCC and we don't indicate any other possible coating damage or something, then we don't see any safety-of-flight issue, but we might still have significant tile damage." The analysis of the selected passage does not, however, depend upon the actual or possible lack of indications of damage because indications are causally posterior to the state of the RCC, and it is the question of whether the RCC was actually damaged that is central to the analysis, as discussed below.

Before getting into the example, some explanation is needed. In the paragraph from which the analyzed passage is taken, "McCormack" is the manager of the Mission Evaluation Room (MER), "a support function of the Shuttle Program Office that supplies engineering expertise for missions in progress" (CAIB 2003: 141). The quote is from his briefing on STS-107 flight day nine (Friday, January 24, 2003) to the Mission Management Team. "RCC" is reinforced carbon-carbon, which protects the leading edge of the Orbiter's wing from thermal damage during re-entry. The "it" in "if it hit the RCC," refers to foam insulation that was shed from the External Tank during launch.

In addition, a few definitions are needed. Contingencies are possible situations or events that are dependent upon potentially variable conditions, which are themselves situations or events. A conditional statement of the form "if p then q," "if p, q," or "q if p" asserts that contingency q is situationally related to, and contingent upon, condition p. Thus, if condition p is true or exists or occurs, then one can justifiably infer that contingency q is also true or exists or occurs, as long as the conditional statement "if p then q," "if p, q," or "q if p," a conditional relation between p and q, is a valid relation.

Given these definitions, the example passage is a conditional statement that relates the contingency q ("we don't see any issue") and the condition p ("it hit the RCC"), in the form of a logical relation "q if p." The passage can also be rewritten in the form "if p then q" as follows: "if it hit the RCC [then] we don't see any issue." In either form, the meaning is the same.

The passage is an assertion that a consequence of foam hitting the RCC would be that the MER wouldn't see a safety-of-flight issue. Is this a valid assertion?

As is now known, the condition was true: foam had indeed hit the RCC. The contingency was also true: the MER, as managed by McCormack, did in fact conclude that there was no safety-of-flight issue. The problem is that the contingency is not logically implied by the condition, given the facts of the situation, as will be shown in detail below. The assertion is therefore invalid.

The invalidity is not due to the fact that the MER's conclusion was not a *direct* consequence of the fact that foam hit the RCC. The assertion was offered as a summary of a more detailed chain of reasoning. It reports an input and an output of the MER's deliberations. That is, given the condition of a foam strike on the RCC as input to the MER, an output of the MER was a conclusion that it was not seen as a safety-of-flight issue. Such a summary is perfectly acceptable, even desirable. The invalidity is that, given the facts of the situation, there is no logical chain of reasoning that leads from the input to the output. This is not news, as the CAIB has essentially come to that conclusion. The purpose of addressing the invalidity here is to explicitly show how a plausible fact-based chain of reasoning can be subverted by introduction of belief-based reasoning.

Analysis of the assertion

As shown in figure 2 on page 10 and the corresponding discussion below, directly relating the condition "it [i.e., foam] hit the RCC" and the contingency "we don't see any [safety-of-flight] issue" obscures and bypasses intervening situational possibilities, vitally important possible states of the world, which are fundamentally important contingencies of the condition "it hit the RCC." These obscured situational possibilities should have been the conditions upon which safety-of-flight judgments were based.

Specifically, there are several situational possibilities that are immediately and directly contingent upon the condition that foam hit the RCC, including: 1) the foam damaged the RCC (and this fact is known); 2) the foam did not damage the RCC (and this fact is known); and 3) it is uncertain whether the foam damaged the RCC.

These possibilities can be shown as mutually exclusive conditional relations (r_i) between conditions (p_i) and contingencies (q_i):

- r1: if p_1 then q_1 : If foam hit the RCC, then foam damaged the RCC (and this fact is known).
- r2: if p_2 then q_2 : If foam hit the RCC, then foam did not damage the RCC (and this fact is known).
- r3: if p_3 then q_3 : If foam hit the RCC, then it is uncertain whether foam damaged the RCC.

Note that $p_1=p_2=p_3$. This single condition has three contingencies: q_1 , q_2 , and q_3 . Thus, an alternative form of the above relations would be: if p_1 then q_1 , q_2 , or q_3 , where "or" is an exclusive "or."

Recall that the broader passage appeared to contain a condition that is the logical conjunction of "it [foam] hit the RCC" and "we don't indicate any other possible coating damage or something." Thus, according to that conjunction, there are three possible situations:

- 1) foam damaged the RCC (and this fact is known) and we don't indicate any other possible coating damage or something, or
- 2) foam did not damage the RCC (and this fact is known) and we don't indicate any other possible coating damage or something, or
- 3) it is uncertain whether foam damaged the RCC and we don't indicate any other possible coating damage or something.

That is to say, the conjunction of "we don't indicate any other possible coating damage or something" with "it [foam] hit the RCC" does not change the situation with respect to the actual condition of the RCC. At the risk of explaining the obvious, this is because an indication is causally posterior to that which is indicated. For example, an indication of a sensor measuring heat is caused by the heat; the indication does not cause the heat. Thus, it is the condition of the RCC that causally determines the safety-of-flight, as shown below, not the indications.

The contingencies q_1 , q_2 , and q_3 of the preceding relations r_1 , r_2 , and r_3 are themselves situational conditions upon which safety-of-flight judgments are contingent. For example, contingency q_1 ("foam damaged the RCC") in relation r_1 is contingent upon p_1 ("foam hit the RCC"), whereas the condition "foam damaged the RCC" has the contingency "it is a safety-of-flight issue." Accordingly, in the following relations (r_4 , r_5 , r_6), contingencies q_1 , q_2 , and q_3 above are renumbered as

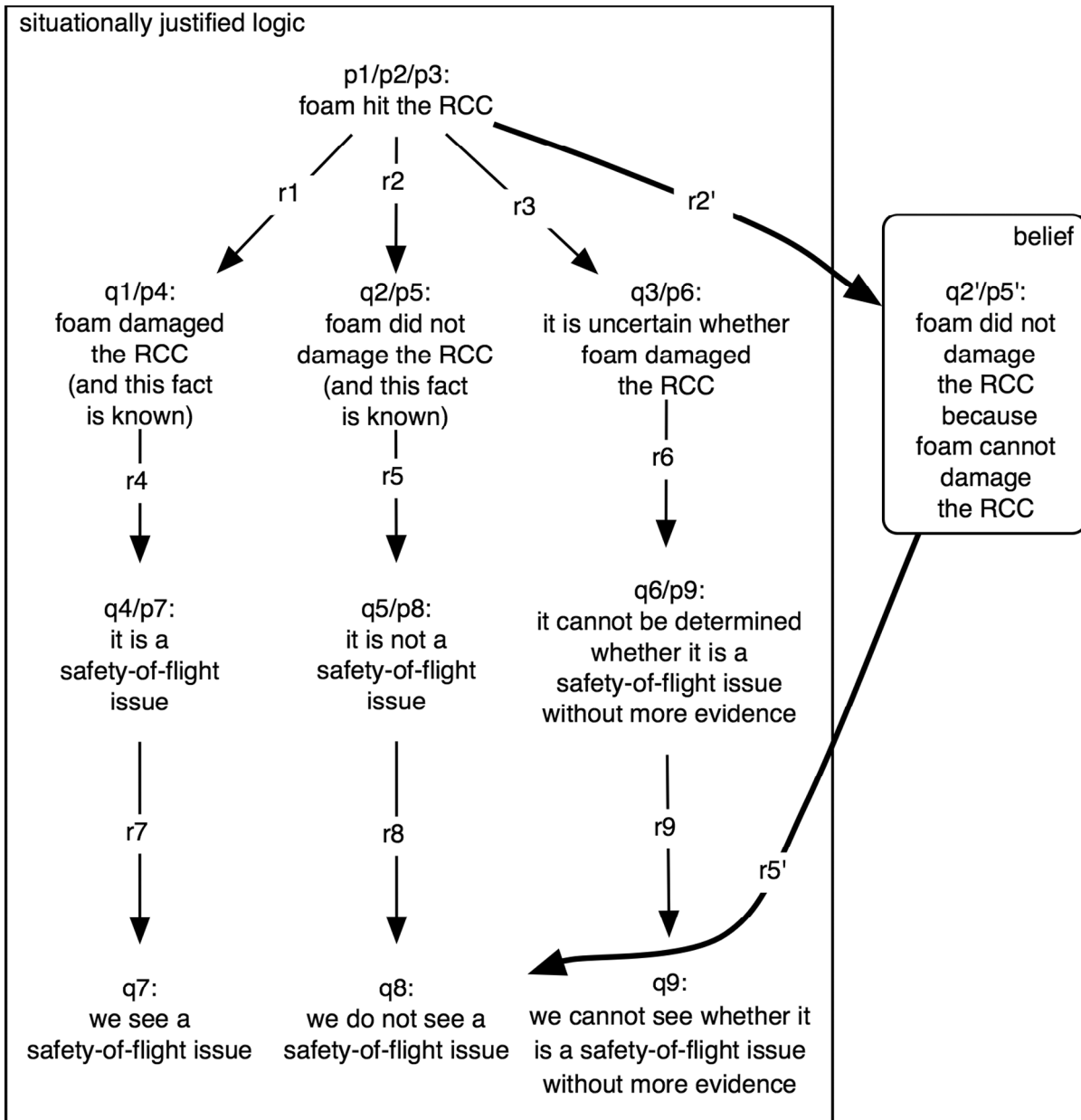


Figure 2. Analysis of the assertion: "If it [i.e., foam] hit the RCC, [then] we don't see any [safety-of-flight] issue" and its implicit conditions and contingencies, showing how belief can short-circuit logic. The conditions and contingencies all correspond to actual or potential states of the world. Each link r_i is a logical relation between a condition p_i and a contingency q_i ($i=1..9$). Links $r_{2'}$ and $r_{5'}$ are belief relations. Contingency $q_{2'}$ is a belief related to condition $p_1=p_2=p_3$. Condition $p_{5'}$ is a belief related to contingency q_8 .

conditions p4, p5, and p6 respectively, each with its own contingency (q4, q5, q6):

- r4: if p4 then q4: If foam damaged the RCC (and this fact is known), then it is a safety-of-flight issue.
- r5: if p5 then q5: If foam did not damage the RCC (and this fact is known), then it is not a safety-of-flight issue.
- r6: if p6 then q6: If it is uncertain whether foam damaged the RCC, then it cannot be determined whether it is a safety-of-flight issue without more evidence.

There must also be relations which relate states of the world regarding safety to states of the world regarding perceptions of safety. For example, in an ideal world:

- r7: if p7 then q7: If it is a safety-of-flight issue, then we see a safety-of-flight issue.
- r8: if p8 then q8: If it is not a safety-of-flight issue, then we do not see a safety-of-flight issue.
- r9: if p9 then q9: If it cannot be determined whether it is a safety-of-flight issue without more evidence, then we cannot see whether it is a safety-of-flight issue without more evidence.

Note that the contingencies q4, q5 and q6 in relations r4, r5, and r6 became conditions p7, p8 and p9 respectively in relations r7, r8, and r9.

The nine relations r1-r9 are diagrammed in figure 2 on page 10.

Interpretation of the logic of the assertion

In order to be justified in asserting that "we do not see a safety-of-flight issue" based on the network of logical relations r1 through r9, the conditions p2, p5, and p8 would have to be true. That is, given that foam actually hit the RCC (p2), it would also have to be true that foam really did not damage the RCC (q2). Further, if foam really did not damage the RCC (p5), then by definition (r5) it is not a safety-of-flight issue (q5), and if it is not a safety-of-flight issue (p8), then we do not see a safety-of-flight issue (q8).

Unfortunately, it was not known, at the time the assertion was made, whether foam damaged the RCC.

In fact, conditions p3, p6, and p9 were true. That is, foam actually hit the RCC (p3), but it was uncertain whether foam damaged the RCC (q3/p6), so it could not be determined whether it was a safety-of-flight issue without more evidence (q6/p9). One can only logically conclude, according to this set of situationally justified logical relations, that "we cannot see whether it is a safety-of-flight issue without more evidence" (q9). There is no

logical path from condition p6, not knowing whether the RCC was damaged, to q8 ("we do not see a safety-of-flight issue").

How is it possible to come to the conclusion that "we do not see a safety-of-flight issue" without knowing whether foam damaged the RCC? According to the CAIB's conclusions (CAIB 2003:168),

"Program managers ... gradually become inured to external tank foam losses and on a fundamental level did not believe foam striking the vehicle posed a critical threat to the orbiter. In particular, Shuttle managers exhibited a belief that RCC panels are impervious to foam impacts."

(See appendix 2 for this and other excerpts relevant to belief.)

Thus, given the condition "foam hit the RCC," the contingency q2 "foam did not damage the RCC" was *believed* to be true, even in the absence of sufficient evidence about the current state of the world, because it was *believed*, based on prior experience, that foam *cannot* damage the RCC. Thus the belief is: "foam did not damage the RCC because foam cannot damage the RCC." This belief seems to have resulted from an unjustified faith in induction such that a small set of repeated experiences was interpreted as a universal law.

The pre-existing belief could have been reinforced by conjunction with the condition, if it were the case, that "we don't indicate any other possible coating damage or something," but the belief was not based on these indications. Further, not all available indications were examined, and the remaining subset of available indications was insufficient to judge the state of the RCC. Thus, in this case, a lack of indication of damage would not imply an actual lack of damage, although it might seem reassuringly supportive of the belief. An unambiguous positive indication of damage, however, such as a clear photograph of the leading edge of Columbia's left wing, would have strongly challenged the belief.

The assertion that "if foam hit the RCC, then we don't see a safety-of-flight issue" can be expanded to make explicit the underlying belief, and the belief relations, as shown below. The belief relations are designated r2' and r5'.

- r2': if p2 then q2': If foam hit the RCC, then foam did not damage the RCC because foam cannot damage the RCC.
- r5': if p5' then q5: If foam did not damage the RCC because foam cannot damage the RCC, then we do not see a safety-of-flight issue.

(Note that the contingency q_2' of relation r_2' becomes the condition p_5' of relation r_5').

The belief relations r_2' and r_5' are not situationally justified logical assertions but belief-based assertions. These relations and the intervening belief node (q_2'/p_5') cause a short circuit in the structure of situationally justified logical assertions, as illustrated in figure 2, by illogically linking the condition "if foam hit the RCC" and the contingency "we don't see a safety-of-flight issue." The perception of safety is achieved through belief rather than by determining the situational state of the world regarding damage to the RCC and judging accordingly.

Finally, it is important to understand that this example serves only as a simple illustration. It is not, nor is it intended to be, a definitive analysis. The actual network of logical and belief relations operative during STS-107 was vastly more subtle and complex, even for the subset of concerns addressed here.

Summary regarding "if-then" analysis

This example illustrates an approach to constructing and evaluating a decision making network based on the facts of the accident and the conditional statements found in the text of the accident report. A Perilog keyword-in-context search on the query word "if" was conducted to find and relevance-rank the passages that contain conditional statements, and the top-ranking one was used as the basis of the example analysis. That analysis shows how a belief can bypass essential conditions and contingencies. A more detailed analysis would include more of the available conditional statements. Accordingly, some of the most relevant "if" clauses (excerpts containing "if" and a condition) found by Perilog are listed in appendices 3 and 4. The corresponding contingencies are to be found in context in the full text of the report. In addition, belief-related passages found by Perilog are listed in appendix 2. These elements can serve as input for further elaboration of the structure of decision making at NASA during the final flight of Columbia.

Conclusion

Accident reports provide a valuable safety resource when they are read and understood, but they have significant additional value when they can be analyzed and explored using text mining operations. Perilog text mining tools can be very usefully applied for this purpose, as has been illustrated by several examples of its application to chapter 6, "Decision Making at NASA," of the final report of the Columbia Accident Investigation Board (CAIB). Perilog makes it easy for users to find highly significant safety-related topics and concerns, even if a user has no initial knowledge of those topics and concerns. That is, Perilog search capabilities do not require a user to know in advance what they are looking

for, because Perilog is highly effective as a knowledge discovery tool.

In addition, as suggested by the third example, Perilog can be used to find significant narrative evidence, in the form of conditional statements, for analysis of decision making. These statements, typically of the form "if p then q," are assertions that condition p implies contingency q. While the judgments of experts and expert groups such as the Mission Evaluation Room (MER) have the effect of asserting relations between conditions and contingencies that serve as the basis of decision making, the example analysis shows how belief relations can bypass logical relations. To make better decisions, the logical structure of decision making must be examined, and belief relations must be identified and eliminated at the very time when decisions are being made. Given such a requirement, the logic of mission-critical decision making should be rigorously diagrammed and analyzed in real time. Thus, when forming or taking the advice of expert individuals or groups in these situations, verbal assertions relating conditions to contingencies should only be formally accepted after thorough, documented analysis by an unbiased team of decision making experts. Such a team would most likely be managed by the independent NASA Engineering Safety Center (NESC).

In support of this and related uses, Perilog can contribute to mission success by providing unique and valuable capabilities for text mining of mission documentation. In fact, at the request of the Systems Engineering Office of the NESC, a copy of the Perilog text mining software was transferred to NESC for application to their databases containing unstructured text. (Perilog was delivered under a Federal Government Transfer Letter, Reference No. SUA2-000904, on August 2, 2004.) The NESC databases include Problem Reporting and Corrective Action (PRACA) records and mission non-conformance records for the Space Station and Space Shuttle, and will contain similar records from missions conducted under NASA's new Vision for Space Exploration (NASA 2004). Perilog is particularly appropriate for application to these databases because its unique methods and software exploit latent linguistic structure in seemingly unstructured text, and work together synergistically, enabling highly effective search and retrieval of information contained in narrative data.

References

- CAIB (Columbia Accident Investigation Board Report), Volume 1. Washington, D.C.: 2003.
<http://www.caib.us/news/report/volume1/default.html>
- McGreevy, M. W.: Reporter concerns in 300 mode-related incident reports from NASA's Aviation Safety Reporting System. NASA TM-110413, 1996.

McGreevy, M. W.: A practical guide to interpretation of large collections of incident narratives using the QUORUM method. NASA TM-112190, 1997.

McGreevy, M. W.: Searching the ASRS database using QUORUM keyword search, phrase search, phrase generation, and phrase discovery. NASA TM-2001-210913, 2001.

McGreevy, M. W. (2003). User Guide to QUORUM Perilog's Graphical User Interface. Unpublished software documentation.

McGreevy, M. W. (2004a). System, method and apparatus for generating phrases from a database. U. S. Patent number 6,697,793.

McGreevy, M. W. (2004b). System, method and apparatus for discovering phrases in a database. U. S. Patent number 6,721,728.

McGreevy, M. W. (2004c). System, method and apparatus for conducting a phrase search. U. S. Patent number 6,741,981.

McGreevy, M. W. (2004d). System, method and apparatus for conducting a keyterm search. U. S. Patent number 6,823,333.

McGreevy, M. W. (2005). Approaching experiential discourse iconicity from the field. *Text*, 25, 1, pp. 67-105.

McGreevy, M. W.; and Statler, I. C.: Rating the relevance of QUORUM-selected ASRS incident narratives to a 'Controlled Flight Into Terrain' accident. NASA TM-1998-208749, 1998.

NASA (2004). The Vision for Space Exploration. P-2004-01-334-HQ.

Appendix 1. Processing the CAIB text for input to Perilog.

This appendix describes how a PDF version of chapter 6 of the CAIB report was pre-processed to convert it into proper format for input to Perilog. It can serve as a guide for pre-processing of other PDF files, such as the rest of the CAIB report, for input to Perilog.

Obtaining the PDF file and converting it to text

Chapter 6 of the CAIB report was downloaded from:

<http://www.caib.us/news/report/pdf/vol1/chapters/chapter6.pdf>

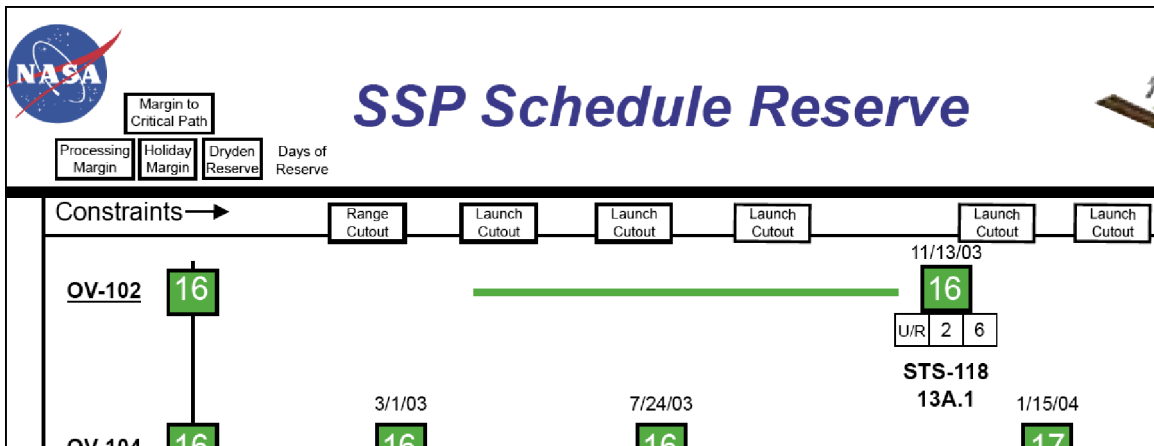
Using Adobe Reader 6, this file was converted from PDF to text by using the "Save as text..." command in the "File" menu. The text file uses the octal 015 character, carriage return, to indicate an end of line, while Unix systems such as Mac OS X use the octal 012 character, newline. Consequently, the next step was to convert carriage returns in the text to newlines.

Editing problematic output

It was discovered that the text output by Adobe Reader from the PDF file contained some sequences of words that were concatenated without spaces, including some very long strings. These problems were found to involve text extracted from presentation slides ("vugraphs"). The longest such string is the following, which contains 142 continuous alphabetic characters and no spaces or line breaks:

```
ofReserveProcessingMarginHolidayMarginDrydenReserveRangeCutoutLaunchCutoutLaunchCutout
LaunchCutoutLaunchCutoutLaunchCutoutConstraintsCritical
```

The source of this string is the top of the vugraph illustrated in on page 137 of the CAIB report, in chapter 6, figure 6.2-5, as shown in the following excerpt:



Given this and other problematic text output from Adobe Reader, the output was hand edited. Most of the editing involved deleting all text extracted from vugraphs. In addition, the page headers and footers, which were mixed in with the text, were removed. Thus, instances of the header "COLUMBIA ACCIDENT INVESTIGATION BOARD" and footer "<page number> Report Volume I August 2003" were deleted. Further, fragments of email messages were found to be repeated, so they were removed. Extracted emails that were out of proper sequence with the text were restored to their proper positions. For example, the following three passages appeared without the referenced emails, so the emails were repositioned to restore the proper sequence.

The following reply from Campbell to Daugherty was sent at 4:49 p.m.

On the next day, Tuesday, Daugherty sent the following to Campbell.

Campbell's reply:

Captions and sidebars were placed by Adobe Reader into the main text in approximately the position where they occurred in the original document, and those positions were accepted as placing the information into its appropriate context, but if they appeared in the middle of a paragraph, the integrity of the paragraph was restored by moving the inserted text out of the paragraph. In some cases, extracted text from a figure caption was placed in the middle of a word in the main text. These were repaired during editing.

Several other edits were also made. The strings "[continued on next page] and [continued from previous page]" were removed. Lines ending in "-" were inspected for word breaks, and broken words were rejoined. As indicated in the Errata for Volume 1 (appendix D.b), "Trish Petite" was corrected to "Trish Petete" on page 156. An instance of "Ron Dittermore" was corrected to "Ron Ditemore."

In the process of saving PDF files as text, Adobe Reader's mapping of tables to linearized text gives rows priority, so that entire rows of text are mapped, one after the other. This is compatible with tables organized like CAIB report figure 6.1-7 in which each row contains a unit of information, a record consisting of several fields, because the mapping to linearized text maintains the contextual adjacency of the row information. In this case, the extracted text starts as shown below, with the first three field types, "MISSION," "DATE," and "COMMENTS," followed by three-line records consisting of a mission identifier, a date, and comments:

```
MISSION
DATE
COMMENTS
STS-1
April 12, 1981
Lots of debris damage. 300 tiles replaced.
STS-7
June 18, 1983
First known left bipod ramp foam shedding event.
STS-27R
December 2, 1988
Debris knocks off tile; structural damage and near burn through results.
STS-32R
...
```

Adobe Reader's row-priority order is not as compatible with tables like CAIB report figure 6.1-1 in which each column contains a unit of information, a record consisting of several fields. In this case, the extracted text starts as shown below, with the first field type, "Flight," followed by all of the flight identifiers from the table, then the second field type, "ET #," followed by all of the external tank numbers from the table, etc.:

```
Flight
STS-7
STS-32R
STS-50
STS-52
STS-62
STS-112
STS-107
ET #
06
25
45
55
62
115
93
ET Type
SWT
LWT
...
```

Adobe Reader's linearization of PDF tables having a column-oriented format causes the resulting text to be contextually grouped by category, as here when flight identifiers appear together, then external tank numbers appear together, etc., rather than maintaining the contextual association of information for each flight.

In summary, of the text that appears in the PDF version of the CAIB report, chapter 6, "Decision Making at NASA," text from vignettes is excluded due to its poor formatting when converted by Adobe Reader, and page headers, page footers, and continuation notes are also excluded, while the main text, sidebars, emails, figure captions, tables (regardless of orientation), and end notes are all retained.

Character processing

Earlier, carriage return characters were mapped to newline characters for compatibility with Unix. At this point, additional character processing was performed. Octal characters with codes less than 040 (with the exception of 012 and 015) or greater than 176 were mapped to spaces (octal code 040) prior to Perilog processing because some characters in these ranges have been found to have undesirable consequences in the Unix processing environment.

Characters with octal codes outside the acceptable range were found in chapter 6 with the frequencies and encoded characters shown in the following table. The standard used for character encoding in the PDF file was not available, so the encoded characters were determined by inspection of the text and its corresponding numerical codes.

<u>octal code</u>	<u>frequency</u>	<u>description and mapping</u>
216	6	the letter "e" with <i>accent grave</i> (better mapped to unaccented "e")
245	358	both a true curled apostrophe and a bullet
247	1	a stray beta character in an email message header
311	43	ellipsis (...)
320	65	a long dash
322	243	starting curled double quotes
323	261	ending curled double quotes
324	24	starting straight single quote
325	46	ending straight single quote

In addition, all letters were mapped to upper case because the current implementation of the Perilog software works best that way. This is a legacy of the software's original application to the ASRS database in which all letters are capitalized.

Editing text to create units of text based on paragraphs

The text resulting from the previous processes can be processed as a single document by Perilog, but for present purposes the approach taken was to treat individual paragraphs as units of text so that they could be retrieved and relevance-ranked individually. (There are other ways to break up long documents into units of text, such as treating sentences or passages as sub-documents.) The text was edited to ensure that paragraphs could be easily recognized by the software as individual units of text. To mark paragraphs, the text was edited to ensure that a single empty line followed each paragraph. In addition, the text (especially email) was edited to remove any blank lines that would mark an undesired paragraph break.

While ordinary paragraphs were treated as units of text, various other text configurations were also treated as units of text.

- Quoted sub-paragraphs that illustrate points made in preceding text were combined with the text that introduces them. For example, the numbered subparagraphs shown below were combined with the paragraph that precedes them to form a single unit of text.

design requirements in the Shuttle's "Flight and Ground System Specification-Book 1, Requirements," precluded foam-shedding by the External Tank. Specifically:

3.2.1.2.14 Debris Prevention: *The Space Shuttle System, including the ground systems, shall be designed to preclude the shedding of ice and/or other debris from the Shuttle elements during prelaunch and flight operations that would jeopardize the flight crew, vehicle, mission success, or would adversely impact turnaround operations.¹*

3.2.1.1.17 External Tank Debris Limits: *No debris shall emanate from the critical zone of the External Tank on the launch pad or during ascent except for such material which may result from normal thermal protection system recession due to ascent heating.²*

- Each section header was included with the first paragraph of its section to form a single unit of text. Since headers become in-line text, a dash (-) was added after each header in order to distinguish it. For example (from page 145):

MISSED OPPORTUNITY 1 - On Sunday, Rodney Rocha e-mailed a Johnson Space Center Engineering Directorate manager to ask if a Mission Action Request was in progress for Columbia's crew to visually inspect the left wing for damage. Rocha never received an answer.

- In transcripts, a unit of text was defined as continuous speech by a single speaker. Further, if a section header preceded a speaker's quote, the header and quote were combined as a single unit of text, with a dash added after the header. For example (from page 147):

Transcript Excerpts from the January 21, Mission Management Team Meeting - Ham: " Alright, I know you guys are looking at the debris."

- Email messages were treated as units of text. Text introducing an email message, such as the following from page 150:

An e-mail that Lieutenant Colonel Timothy Lee sent to Don McCormack the following day shows that the Defense Department had begun to implement Austin's request.

was combined with the message it introduces to create a single unit of text.

- Figure captions were treated as units of text.
- Sidebars were handled on a case-by-case basis. Some paragraph boundaries were ignored, as when the whole sidebar on page 124 was treated as a single unit of text. In contrast, each definition in the sidebar on page 122 was treated as a separate unit of text. Paragraphs in the long sidebar "The Crater Model" were retained, and each was treated as a unit of text.
- Each end note was treated as a unit of text, except those consisting of "Ibid.," which were grouped with the full reference.

Labeling units of text

Once the editing processes of the preceding section were complete, the paragraphs and paragraph-like units of text were labeled. Each one was automatically preceded by a header line, which is required by Perilog. For example,

```
NONWORD v01_ch06_pg000_item0002
This chapter connects Chapter 5's analysis of NASA's
broader policy environment to a focused scrutiny of Space
Shuttle Program decisions that led to the STS-107 accident.
Section 6.1 illustrates how foam debris losses that violated
design requirements came to be defined by NASA management...
```

A Perilog header line starts with the string "NONWORD" and after a space or tab has an identifier that is unique for each unit of text. The identifier can be as simple as a number, or can be more complex and informative like the one shown above. Identifiers may include numbers, letters, and underscores.

The identifier above has four parts, separated by an underscore: volume number, chapter number, page number, and item number. "v01" refers to volume 1 of the CAIB report. "ch06" refers to chapter 6, "Decision Making at NASA." "pg000" refers to the page number, where 000 is a placeholder for the actual page number. "item0002" refers to the second unit of text, where the item number starts with 0001 and ends with the total number of units.

The identifier used here is very helpful when relevant text is later retrieved by Perilog. When the relevant text is shown with its identifier, or a list of relevant identifiers is provided, the identifiers themselves indicate the location of the text item within the CAIB document. While the demonstration database contains only chapter 6 of volume 1, when numerous volumes and chapters are included, these identifiers would be even more helpful in locating the original text of each item.

Since there is no way to automatically and reliably assign page numbers to the headers, each item of text was located in the source document and each header was edited to provide the correct page number. For example, in the identifier v01_ch06_pg000_item00020 the placeholder 000 was changed to the actual page number, 121, to produce the identifier v01_ch06_pg121_item0002.

The resulting file containing all of chapter 6, with each paragraph and paragraph-like unit of text having a header line, was now in the proper format for input to Perilog. This format and the subsequent processing steps are described in the documentation provided with the Perilog software distributions and are beyond the scope of this appendix.

Selecting a stoplist

Most of the Perilog processes depend upon recognition of a class of words that are called "stopwords." These are typically rather generic words, including articles such as "a" and "the", nouns such as "something", pronouns such as "anybody", verbs such as "are", adverbs such as "generally", and other words that are generic from the point of view of the topics of interest. For this reason, stopwords are sometimes called noncontent words. Clearly, words like "foam" and "damage" are content words rather than stopwords, but there is no automatic and definitive test that can ensure that categorizing a particular word as a stopword is appropriate for a given text.

Stoplists must be selected carefully, and tuned for compatibility with the full range of topics of interest within the text being processed. A stoplist for "general text" (Frakes & Baeza-Yates 1992: 114-5) was initially applied to the CAIB text. When reviewing extracted phrases, some of them seemed odd. To diagnose the problem, phrase search was used to find instances of the odd phrases. For example, the extracted phrases "mission evaluation" and "intercenter photo" seemed truncated. Reviewing the results of the phrase search showed that "mission evaluation" always occurs in the phrase "mission evaluation room", and "intercenter photo" always appears in "intercenter photo working group".

Reviewing the initially applied stoplist provided the explanation. The words "room", "working", and "group" were among the words in the stoplist, so those words were restricted within phrases. After customizing the stoplist to remove words such as "room", "working", "group", and others that seemed potentially problematic, much better results were obtained. This illustrates the importance of matching a stoplist to the full range of topics of interest within a text. Shown below are the words contained in the stoplist that was ultimately applied to the CAIB text.

A	E	HOWEVER	OTHERS
ALREADY	EACH	I	OUR
ALSO	EITHER	IN	P
ALTHOUGH	EVEN	IS	Q
AN	EVENLY	IT	QUITE
AND	EVER	ITS	R
ANOTHER	EVERY	ITSELF	RATHER
ANY	EVERYBODY	J	REALLY
ANYBODY	EVERYONE	JUST	S
ANYONE	EVERYTHING	K	SHALL
ANYTHING	EVERYWHERE	L	SHE
ANYWHERE	F	LARGELY	SINCE
ARE	FOR	LET	SO
AS	FROM	LETS	SOME
AT	FURTHERED	LIKELY	SOMEBODY
B	FURTHERING	M	SOMEONE
BE	FURTHERS	ME	SOMETHING
BEEN	G	MOSTLY	SOMEWHERE
BEINGS	GENERALLY	MR	STILL
BOTH	H	MRS	SUCH
BY	HAD	MUCH	T
C	HAS	MY	THAN
D	HAVE	MYSELF	THAT
DID	HAVING	N	THE
DO	HE	O	THEIR
DOES	HER	OF	THEM
DONE	HERSELF	ON	THESE
DOWNED	HIM	ONLY	THEY
DOWNING	HIMSELF	OR	THIS
DOWNNS	HIS	OTHER	THOSE

THOUGH	WAS	WHILE	YOUR
TO	WE	WHOSE	YOURS
TOO	WELL	WILL	Z
U	WELLS	WITH	
US	WENT	X	
V	WERE	Y	
VERY	WHETHER	YET	
W	WHICH	YOU	

Perilog methods, as documented in the patents (McGreevy 2004a-d), also include provision for numerous special purpose stoplists based on the primary application-specific stoplist, such as the one shown above. These stoplists provide fine-tuning for processing special phrases such as "in-flight anomaly" or "take off", which begin or end with words that in most other contexts are still to be treated as stopwords. The current software does not fully implement these methods. As a consequence, Perilog software is now delivered with several stoplists, including a safety-related stoplist that does not include prepositions such as "in" and "off", and a nonsafety-related stoplist that includes prepositions.

Appendix 2. Examples of passages about beliefs.

Highlighting was added (manually) for readability. The order of the passages shown here is determined by the relevance of the source paragraphs to the words "belief", "beliefs", "believe", and "believed" in context. The passages are from CAIB Report Volume 1, chapter 6, "Decision Making at NASA." Source page numbers in CAIB Report Volume 1 are shown in parentheses.

ENGINEERS WHO ATTENDED THIS BRIEFING INDICATED A BELIEF THAT MANAGEMENT FOCUSED ON THE ANSWER - THAT ANALYSIS PROVED THERE WAS NO SAFETY - OF - FLIGHT ISSUE - RATHER THAN CONCERNS ABOUT THE LARGE UNCERTAINTIES THAT MAY HAVE UNDERMINED THE ANALYSIS THAT PROVIDED THAT ANSWER . (160)

AFTER REVIEWING AVAILABLE FILM , INTERCENTER PHOTO WORKING GROUP ENGINEERS BELIEVED THE ORBITER MAY HAVE BEEN DAMAGED BY THE STRIKE IT WAS AT THIS POINT , BEFORE ANY ANALYSIS HAD STARTED , THAT SHUTTLE PROGRAM MANAGERS OFFICIALLY SHARED THEIR BELIEF THAT THE STRIKE POSED NO SAFETY ISSUES , AND THAT THERE WAS NO NEED FOR A REVIEW TO BE CONDUCTED OVER THE WEEKEND . (142)

PROGRAM MANAGERS , FROM RON DITTEMORE TO INDIVIDUAL MISSION MANAGEMENT TEAM MEMBERS , HAD , OVER THE COURSE OF THE SPACE SHUTTLE PROGRAM , GRADUALLY BECOME INURED TO EXTERNAL TANK FOAM LOSSES AND ON A FUNDAMENTAL LEVEL DID NOT BELIEVE FOAM STRIKING THE VEHICLE POSED A CRITICAL THREAT TO THE ORBITER . IN PARTICULAR , SHUTTLE MANAGERS EXHIBITED A BELIEF THAT RCC PANELS ARE IMPERVIOUS TO FOAM IMPACTS . (168)

DEBRIS ASSESSMENT TEAM MEMBERS BELIEVED THAT IMAGING OF POTENTIALLY DAMAGED AREAS WAS NECESSARY EVEN AFTER THE JANUARY 24 , MISSION MANAGEMENT TEAM MEETING , WHERE THEY HAD REPORTED THEIR RESULTS . (168)

SECTION 6.3 NOTES THE DECISIONS MADE DURING STS-107 IN RESPONSE TO THE BIPOD FOAM STRIKE , AND REVEALS HOW ENGINEERS ' CONCERNS ABOUT RISK AND SAFETY WERE COMPETING WITH - AND WERE DEFEATED BY - MANAGEMENT 'S BELIEF THAT FOAM COULD NOT HURT THE ORBITER , AS WELL AS THE NEED TO KEEP ON SCHEDULE . (121)

CALVIN SCHOMBURG STATED A BELIEF THAT IF THERE WAS SEVERE DAMAGE TO THE TILES , " NOTHING COULD BE DONE . " (160)

AT THIS POINT , TILE EXPERT CALVIN SCHOMBURG STATES HIS BELIEF THAT NO SAFETY - OF - FLIGHT ISSUE EXISTS . (162)

THE BOEING ANALYSTS NEXT ASKED WHY THEY WERE WORKING SO HARD ANALYZING POTENTIAL DAMAGE AREAS IF SHUTTLE PROGRAM MANAGEMENT BELIEVED THE DAMAGE WAS MINOR

AND THAT NO SAFETY - OF - FLIGHT ISSUES EXISTED . (160)

SCHOMBURG , THOUGH AWARE OF THE DEBRIS ASSESSMENT TEAM 'S REQUEST FOR IMAGING , TOLD SHACK AND PETETE THAT HE BELIEVED ON - ORBIT IMAGING OF POTENTIALLY DAMAGED AREAS WAS NOT NECESSARY . (156)

THE DEBRIS ASSESSMENT TEAM FOCUSED ON ANALYZING THE IMPACT AT LOCATIONS OTHER THAN THE RCC LEADING EDGE . THIS MAY HAVE BEEN DUE , AT LEAST IN PART , TO THE TRANSPORT ANALYSIS PRESENTATION AND THE LONG - STANDING BELIEF THAT FOAM WAS NOT A THREAT TO RCC PANELS . (145)

A TILE EXPERT TOLD MANAGERS DURING FREQUENT CONSULTATIONS THAT STRIKE DAMAGE WAS ONLY A MAINTENANCE - LEVEL CONCERN AND THAT ON - ORBIT IMAGING OF POTENTIAL WING DAMAGE WAS NOT NECESSARY . MISSION MANAGEMENT WELCOMED THIS OPINION AND SOUGHT NO OTHERS . THIS CONSTANT REINFORCEMENT OF MANAGERS ' PRE - EXISTING BELIEFS ADDED ANOTHER BLOCK TO THE WALL BETWEEN DECISION MAKERS AND CONCERNED ENGINEERS . (169)

SHUTTLE PROGRAM MANAGERS ENTERED THE MISSION WITH THE BELIEF , RECENTLY REINFORCED BY THE STS-113 FLIGHT READINESS REVIEW , THAT A FOAM STRIKE IS NOT A SAFETY - OF - FLIGHT ISSUE . (171)

CONTE ASKED ROCHA IF HE WANTED HER TO PURSUE SUCH A REQUEST THROUGH MISSIONS OPERATIONS DIRECTORATE CHANNELS . ROCHA SAID NO , BECAUSE HE BELIEVED PROGRAM MANAGERS WOULD STILL HAVE TO SUPPORT SUCH A REQUEST . SINCE THEY HAD ALREADY DECIDED THAT IMAGING OF POTENTIALLY DAMAGED AREAS WAS NOT NECESSARY , ROCHA THOUGHT IT UNLIKELY THAT THE DEBRIS ASSESSMENT TEAM COULD CONVINCE THEM OTHERWISE WITHOUT DEFINITIVE DATA . (158)

NO TESTS WERE PERFORMED WITH LARGER DEBRIS OBJECTS BECAUSE IT WAS NOT BELIEVED SUCH DEBRIS COULD EVER IMPACT THE ORBITER . THIS RESULTED IN A VERY LIMITED SET OF CONDITIONS UNDER WHICH CRATER 'S RESULTS WERE EMPIRICALLY VALIDATED . (144)

AS IT BECAME CLEAR DURING THE MISSION THAT MANAGERS WERE NOT AS CONCERNED AS OTHERS ABOUT THE DANGER OF THE FOAM STRIKE , THE ABILITY OF ENGINEERS TO CHALLENGE THOSE BELIEFS GREATLY DIMINISHED . (169)

Appendix 3. Excerpts containing the word "if", sorted on source paragraphs.

The excerpts are from CAIB Report Volume 1, chapter 6, "Decision Making at NASA." These excerpts are from the 40 paragraphs that are most relevant to the word "if". The order of the excerpts shown here is determined by the relevance of the 40 paragraphs to the word "if" in context, which is determined by a Perilog keyword-in-context search on the word "if". Source page numbers in CAIB Report Volume 1 are shown in parentheses. Also see appendix 4, in which the excerpts are sorted differently.

IF IT HIT THE RCC ... IF WE DON'T SEE A SAFETY - OF - FLIGHT ISSUE (161)

IF SOMETHING WERE TO GO TERRIBLY WRONG ... IF SHUTTLE PROGRAM MANAGEMENT BELIEVED THE DAMAGE WAS MINOR AND THAT NO SAFETY - OF - FLIGHT ISSUES EXISTED (160)

IF HE HAD ADDITIONAL FOOTAGE OF EXTERNAL TANK SEPARATION ... IF HE HAD MORE EXTERNAL TANK SEPARATION FILM (148)

" WHAT - IF " LANDING SCENARIOS OF THE POTENTIAL OUTCOME IF THE MAIN LANDING GEAR DOOR SUSTAINED DAMAGED [sic] (169)

IF LAUNCH PROCESSING AND COUNTDOWN PROCEEDED SMOOTHLY ... IF MORE THOROUGH REPAIRS COULD BE DEVELOPED (173)

IF THERE WERE OPTIONS FOR THE SAFE RETURN OF THE STS-107 CREW ... IF THE WING STRUCTURE WAS PREDICTED TO FAIL ON LANDING (173)

IF FOAM SHEDDING COULD NOT BE PREVENTED ENTIRELY ... IF BURN - THROUGH OCCURS (129)

IF PROGRAM MANAGERS HAD UNDERSTOOD THE THREAT THAT THE BIPOD FOAM STRIKE POSED AND WERE ABLE TO UNEQUIVOCALLY DETERMINE BEFORE FLIGHT DAY SEVEN THAT THERE WAS POTENTIALLY CATASTROPHIC DAMAGE TO THE LEFT WING (174)

IF THERE IS A QUESTION ABOUT MAIN GEAR WELL BURN THRU ... IF YOU COME IN FAST AND AT SLIGHTLY LESS PITCH ATTITUDE (164)

IF SHACK 'S BOSS , JOHNSON SPACE CENTER ENGINEERING DIRECTOR FRANK BENZ , KNEW ABOUT THE REQUEST ... IF IT 'S NOT SAFE (157)

IF THE MISSION 'S ORBITER , ATLANTIS , SHOULD BE ROLLED FROM THE ORBITER PROCESSING FACILITY TO THE VEHICLE ASSEMBLY BUILDING , PER ITS PRE - LAUNCH SCHEDULE (139)

IF THE ORBITER LANDED AT KENNEDY ... IF BAD WEATHER AT KENNEDY FORCED THOSE TWO FLIGHTS TO LAND AT EDWARDS (136)

IF HE KNEW WHO WAS REQUESTING THE IMAGERY (153)

IF ANYBODY IS TALKING ABOUT EXTENSION DAYS OR GOING TO GO WITH THAT ... IF THEY GET ANY QUESTIONS AT THE PRESS CONFERENCES OR THAT SORT OF THING (161)

IF MONITORING WAS IMPROVED (127)

IF A MISSION ACTION REQUEST WAS IN PROGRESS FOR COLUMBIA'S CREW TO VISUALLY INSPECT THE LEFT WING FOR DAMAGE (145)

IF WE HAVE A GROUND OR SATELLITE ASSET THAT CAN TAKE A HIGH RESOLUTION PHOTO OF THE SHUTTLE WHILE ON - ORBIT - - TO SEE IF THERE IS ANY FOD DAMAGE ON THE WING (150)

IF ENGELAUF COULD HAVE THE FLIGHT DYNAMICS OFFICER AT JOHNSON SPACE CENTER MAKE AN OFFICIAL REQUEST TO THE CHEYENNE MOUNTAIN OPERATIONS CENTER (152)

IF AVAILABLE MONDAY MORNING ... IF ULF2 LAUNCH WERE 10 / 7 / 04 (135)

IF OTHER ASSUMPTIONS WERE USED (168)

IF HE WANTED HER TO PURSUE SUCH A REQUEST THROUGH MISSIONS OPERATIONS DIRECTORATE CHANNELS (158)

IF A DRYDEN LANDING WAS NOT REQUIRED (137)

IF THERE WAS SEVERE DAMAGE TO THE TILES (160)

IF THE ORBITER HAD BEEN DAMAGED (145)

IF THEY ALL AGREED WITH THE COMPLETED ANALYSES AND WITH THE CONCLUSION THAT NO SAFETY - OF - FLIGHT ISSUES EXISTED (163)

IF THE CRATER RESULTS WERE PROPERLY INTERPRETED (171)

IF IT WAS ACCEPTABLE (122)

IF ANY SR&QA PEOPLE WERE INVOLVED IN THE DECISION TO SAY THAT THE ASCENT DEBRIS HIT (LEFT WING) IS SAFE (166)

IF THE REQUEST WAS THEIRS (153)

IF A SHUTTLE FLIGHT HAD TO SLIP (134)

IF I SPENT MORE TIME THINKING ABOUT IT (155)

IF PROGRAM MANAGERS WERE ABLE TO UNEQUIVOCALLY DETERMINE BEFORE FLIGHT DAY SEVEN THAT THERE WAS POTENTIALLY CATASTROPHIC DAMAGE TO THE LEFT WING (174)

IF CIRCUMSTANCES WARRANTED A SLIP OF THAT DATE (133)

IF MANAGERS AND ENGINEERS WERE TO ARGUE THAT FOAM STRIKES ARE A SAFETY - OF - FLIGHT ISSUE (150)

IF THE SHUTTLE PROGRAM 'S RATIONALE TO FLY WITH FOAM LOSS WAS FOUND TO BE FLAWED (148)

IF THEY HAD ANY INTEREST / DESIRE IN REQUESTING RESOURCES OUTSIDE OF NASA TO VIEW THE ORBITER (158)

IF COLUMBIA HAD SUSTAINED DAMAGE (147)

IF THIS GOAL WAS NOT MET (131)

IF A REASONABLE LAUNCH SCHEDULE IS TO BE MAINTAINED (130)

IF HE HAD A " REQUIREMENT " FOR IMAGERY OF COLUMBIA 'S LEFT WING (153)

Appendix 4. Excerpts containing the word "if", sorted on excerpts.

The excerpts are from CAIB Report Volume 1, chapter 6, "Decision Making at NASA." These excerpts are from the 40 paragraphs that are most relevant to the word "if", which is determined by a Perilog keyword-in-context search on the word "if". The order of the excerpts shown here is determined by excerpt relevance to all of the concerns expressed in chapter 6, which is determined by using Perilog software tools to rank the excerpts according to their relevance to a model of the whole chapter. Source page numbers in CAIB Report Volume 1 are shown in parentheses. Also see appendix 3, in which the excerpts are sorted differently.

IF HE HAD ADDITIONAL FOOTAGE OF EXTERNAL TANK SEPARATION ... IF HE HAD MORE EXTERNAL TANK SEPARATION FILM (148)

IF THE SHUTTLE PROGRAM'S RATIONALE TO FLY WITH FOAM LOSS WAS FOUND TO BE FLAWED (148)

IF PROGRAM MANAGERS HAD UNDERSTOOD THE THREAT THAT THE BIPOD FOAM STRIKE POSED AND WERE ABLE TO UNEQUIVOCALLY DETERMINE BEFORE FLIGHT DAY SEVEN THAT THERE WAS POTENTIALLY CATASTROPHIC DAMAGE TO THE LEFT WING (174)

IF PROGRAM MANAGERS WERE ABLE TO UNEQUIVOCALLY DETERMINE BEFORE FLIGHT DAY SEVEN THAT THERE WAS POTENTIALLY CATASTROPHIC DAMAGE TO THE LEFT WING (174)

IF FOAM SHEDDING COULD NOT BE PREVENTED ENTIRELY ... IF BURN - THROUGH OCCURS (129)

IF MANAGERS AND ENGINEERS WERE TO ARGUE THAT FOAM STRIKES ARE A SAFETY - OF - FLIGHT ISSUE (150)

IF SOMETHING WERE TO GO TERRIBLY WRONG ... IF SHUTTLE PROGRAM MANAGEMENT BELIEVED THE DAMAGE WAS MINOR AND THAT NO SAFETY - OF - FLIGHT ISSUES EXISTED (160)

IF ENGELAUFGOULD HAVE THE FLIGHT DYNAMICS OFFICER AT JOHNSON SPACE CENTER MAKE AN OFFICIAL REQUEST TO THE CHEYENNE MOUNTAIN OPERATIONS CENTER (152)

IF SHACK 'S BOSS , JOHNSON SPACE CENTER ENGINEERING DIRECTOR FRANK BENZ , KNEW ABOUT THE REQUEST ... IF IT 'S NOT SAFE (157)

IF IT HIT THE RCC ... IF WE DON'T SEE A SAFETY - OF - FLIGHT ISSUE (161)

IF A SHUTTLE FLIGHT HAD TO SLIP (134)

IF A MISSION ACTION REQUEST WAS IN PROGRESS FOR COLUMBIA 'S CREW TO VISUALLY INSPECT THE LEFT WING FOR DAMAGE (145)

" WHAT - IF " LANDING SCENARIOS OF THE POTENTIAL OUTCOME IF THE MAIN LANDING GEAR DOOR SUSTAINED DAMAGED [sic] (169)

IF THEY ALL AGREED WITH THE COMPLETED ANALYSES AND WITH THE CONCLUSION THAT NO SAFETY - OF - FLIGHT ISSUES EXISTED (163)

IF THERE WAS SEVERE DAMAGE TO THE TILES (160)

IF HE HAD A " REQUIREMENT " FOR IMAGERY OF COLUMBIA'S LEFT WING (153)

IF HE WANTED HER TO PURSUE SUCH A REQUEST THROUGH MISSIONS OPERATIONS DIRECTORATE CHANNELS (158)

IF ANY SR&QA PEOPLE WERE INVOLVED IN THE DECISION TO SAY THAT THE ASCENT DEBRIS HIT (LEFT WING) IS SAFE (166)

IF THE MISSION 'S ORBITER , ATLANTIS , SHOULD BE ROLLED FROM THE ORBITER PROCESSING FACILITY TO THE VEHICLE ASSEMBLY BUILDING , PER ITS PRE - LAUNCH SCHEDULE (139)

IF I SPENT MORE TIME THINKING ABOUT IT (155)

IF LAUNCH PROCESSING AND COUNTDOWN PROCEEDED SMOOTHLY ... IF MORE THOROUGH REPAIRS COULD BE DEVELOPED (173)

IF COLUMBIA HAD SUSTAINED DAMAGE (147)

IF A DRYDEN LANDING WAS NOT REQUIRED (137)

IF A REASONABLE LAUNCH SCHEDULE IS TO BE MAINTAINED (130)

IF THERE IS A QUESTION ABOUT MAIN GEAR WELL BURN THRU ... IF YOU COME IN FAST AND AT SLIGHTLY LESS PITCH ATTITUDE (164)

IF THERE WERE OPTIONS FOR THE SAFE RETURN OF THE STS-107 CREW ... IF THE WING STRUCTURE WAS PREDICTED TO FAIL ON LANDING (173)

IF WE HAVE A GROUND OR SATELLITE ASSET THAT CAN TAKE A HIGH RESOLUTION PHOTO OF THE SHUTTLE WHILE ON - ORBIT - - TO SEE IF THERE IS ANY FOD DAMAGE ON THE WING (150)

IF THE CRATER RESULTS WERE PROPERLY INTERPRETED (171)

IF THE ORBITER LANDED AT KENNEDY ... IF BAD WEATHER AT KENNEDY FORCED THOSE TWO FLIGHTS TO LAND AT EDWARDS (136)

IF THE ORBITER HAD BEEN DAMAGED (145)

IF THEY HAD ANY INTEREST / DESIRE IN REQUESTING RESOURCES OUTSIDE OF NASA TO VIEW THE ORBITER (158)

IF HE KNEW WHO WAS REQUESTING THE IMAGERY (153)

IF THIS GOAL WAS NOT MET (131)

IF THE REQUEST WAS THEIRS (153)

IF AVAILABLE MONDAY MORNING ... IF ULF2 LAUNCH WERE 10 / 7 / 04 (135)

IF ANYBODY IS TALKING ABOUT EXTENSION DAYS OR GOING TO GO WITH THAT ... IF THEY GET ANY QUESTIONS AT THE PRESS CONFERENCES OR THAT SORT OF THING (161)

IF OTHER ASSUMPTIONS WERE USED (168)

IF CIRCUMSTANCES WARRANTED A SLIP OF THAT DATE (133)

IF IT WAS ACCEPTABLE (122)

IF MONITORING WAS IMPROVED (127)

REPORT DOCUMENTATION PAGE

*Form Approved
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 19-10-2005	2. REPORT TYPE Technical Memorandum	3. DATES COVERED (From - To)
--	---	-------------------------------------

4. TITLE AND SUBTITLE Using Perilog to Explore "Decision Making at NASA"	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S) Michael W. McGreevy	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER 21-078

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Ames Research Center Moffett Field, CA 94035-1000	8. PERFORMING ORGANIZATION REPORT NUMBER A-050002
---	---

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, D.C. 20546-0001	10. SPONSORING/MONITOR'S ACRONYM(S) NASA
	11. SPONSORING/MONITORING REPORT NUMBER NASA/TM—2005-213469

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified — Unlimited Distribution: Nonstandard
Subject Category: 16, 61, 81
Availability: NASA CASI (301) 621-0390

13. SUPPLEMENTARY NOTES
Point of Contact: Michael W. McGreevy, Ames Research Center, MS 262-4, Moffett Field, CA 94035-1000
(650) 604-5784

14. ABSTRACT
Perilog, a context intensive text mining system, is used as a discovery tool to explore topics and concerns in "Decision Making at NASA," chapter 6 of the Columbia Accident Investigation Board (CAIB) Report, Volume I. Two examples illustrate how Perilog can be used to discover highly significant safety-related information in the text without prior knowledge of the contents of the document. A third example illustrates how "if-then" statements found by Perilog can be used in logical analysis of decision making. In addition, in order to serve as a guide for future work, the technical details of preparing a PDF document for input to Perilog are included in an appendix.

15. SUBJECT TERMS
Information retrieval, Data mining, Flight safety, Columbia Accident Investigation Board (CAIB) Report, STS-107

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Michael W. McGreevy
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE (Include area code) (650) 604-5784
Unclassified	Unclassified	Unclassified	Unclassified	28	