NASA Technical Memorandum 112190

A practical guide to interpretation of large collections of incident narratives using the QUORUM method

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

March 1997

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

Table of Contents

A practical guide to interpretation of large collections of incident narratives using the QUORUM method

MICHAEL W. MCGREEVY

Ames Research Center

Summary

Analysis of incident reports plays an important role in aviation safety. Typically, a narrative description, written by a participant, is a central part of an incident report. Because there are so many reports, and the narratives contain so much detail, it can be difficult to efficiently and effectively recognize patterns among them. Recognizing and addressing recurring problems, however, is vital to continuing safety in commercial aviation operations.

A practical way to interpret large collections of incident narratives is to apply the QUORUM method of text analysis, modeling, and relevance ranking. In this paper, QUORUM text analysis and modeling are surveyed, and QUORUM relevance ranking is described in detail with many examples. The examples are based on several large collections of reports from the Aviation Safety Reporting System (ASRS) database, and a collection of news stories describing the disaster of TWA Flight 800, the Boeing 747 which exploded in mid-air and crashed near Long Island, New York, on July 17, 1996. Reader familiarity with this disaster should make the relevance-ranking examples more understandable. The ASRS examples illustrate the practical application of QUORUM relevance ranking.

Introduction

Problematic incidents in commercial aviation operations are more numerous than accidents, so analysis of incidents can provide a broader view of potentially unsafe situations. Unfortunately, the large numbers of incident reports and the many details they contain can overwhelm analysts. This is especially true because the number of available incident reports is steadily increasing. As a result, critically important patterns of incidents can be overlooked, or not recognized in a timely manner.

To help incident analysts, a new automated method has been developed for analyzing, modeling, and relevance-ranking incident narratives. This method has been applied to hundreds of reports from the Aviation Safety Reporting System (ASRS) database. It could also be applied to reports from incident databases being developed by commercial carriers, and other aviation organizations.

The method is called QUORUM, and it was developed at NASA Ames Research Center. QUORUM consists of a collection of software that analyses, models, and relevance-ranks text documents. This paper surveys QUORUM analysis and modeling methods, which are described in detail elsewhere (McGreevy, 1996; McGreevy, 1995). The method of relevance ranking is described here in detail, using numerous examples. Relevance ranking appears to be the most practical way to bring the benefits of QUORUM analysis and modeling to the operational community.

Interpreting Incident Narratives

When a safety-related incident occurs in day-to-day commercial aviation operations, it usually involves several people who are well-positioned to observe the incident and the related circumstances. These participants are typically members of flight or ground crews, air traffic controllers, or other professionals. Sometimes an incident is of such concern that one or more of the participants file formal incident reports. A key part of such a report is the narrative, in which the participants describe the episode in their own words.

The Aviation Safety Reporting System (ASRS) database contains tens of thousands of incident narratives, and many other organizations maintain, or are developing, similar databases. While the information in aviation safety reporting systems has been useful for identifying problems, the narratives themselves have not been fully utilized. This is true despite the fact that the narratives are considered to be the most useful part of the data. According to the late Bill

Reynard, former director of the ASRS (Reynard, undated), "[T]he real power of ASRS lies in the report narratives. Here pilots, controllers, and others tell us about aviation safety incidents and situations in detail."

The main reason why narratives are not fully utilized is that there are many thousands of them, each describing a particular situation, each with a wealth of detail that is contained in seemingly unstructured form. Even the most dedicated and knowledgeable analyst will have difficulty deriving an objective and comprehensive model of the incidents when faced with hundreds of reports. How much more difficult this becomes when a similar number of reports must be analyzed daily, which is the case at the ASRS.

Fortunately, there is a way to deal with the complexity inherent in a large collection of narratives describing particular incidents. Herbert Simon (1969), one of the seminal thinkers in computer science and complexity theory, asserts that "reality" can be adequately modeled by eliminating almost all of the detail, while retaining only the truly essential: "[F]or a tolerable description of reality, only a tiny fraction of all possible interactions needs to be taken into account." The basis of Simon's hypothesis is the redundancy of interactions in complex systems and situations. In his view, many associations are weak and can be ignored, while only a few associations are strong enough to demand consideration. The challenge is to be able to identify the essential interactions from among the blizzard of particular details.

In response to the demand for manageable representations of complex, real-world activities, operationally-oriented researchers have become increasingly interested in "situated" models. A situated model is one in which the significant elements of situational context, the things and events of real-world operations, are explicitly represented. This contrasts sharply with generic mental models applied uniformly to any setting or situation. In situated models, problems occurring in commercial aviation operations would be represented in their full situational contexts. Elements of these contexts would include, for example: systems and automation, crew factors, contingencies of air traffic, mechanical difficulties, safety and security, economic pressures, and the other daily, practical concerns of the operational community.

According to Nardi (1992), an advocate of situated modeling in operational contexts, "Taking context seriously means finding oneself in the thick of the complexities of particular situations at particular times with particular individuals." Since narrative descriptions of incidents contain a wealth of particular details about problems in day-to-day operations, they provide the kind of data that is required for development of situated models. Such comprehensive models have the potential to aid recognition of patterns among operational problems, and support development of well-integrated solutions.

The QUORUM Method

QUORUM is a method of text analysis, modeling, and relevance-ranking. It analyzes narratives to produce situated models. It also applies these models to show the analyst which narratives, and parts of narratives, describe recurring patterns. QUORUM, which stands for QUantitative, Objective, Representative, Unambiguous Modeler, was developed by the Aviation Operations Branch of the Flight Management and Human Factors Division at NASA Ames Research Center (McGreevy, 1996; McGreevy, 1995).

QUORUM has been applied to numerous collections of incident reports from the ASRS database, including: 300 mode-related reports, 101 altitude deviation reports, 200 ATC-related reports, 325 "crew pressure" reports, 185 training reports, and 313 automation error reports. In addition to ASRS reports, QUORUM has been applied to various other kinds of text, including: technical papers, news stories, political speeches, literary works, monthly reports, software specifications, and interviews.

There are four steps in the interpretation of large collections of incident narratives using the QUORUM method.

1) The first step is the selection of reports from the database. QUORUM does not currently influence this step, but it has the potential to do so, as will be shown later.

2) The second step is narrative analysis, the breaking down of narratives into their component parts. Central to this process is the QUORUM metric, a proximity-weighted measure of co-occurrence between words. Using the metric, QUORUM measures the patterns of words in a text to obtain a structural model. It first identifies prominent words in the text. It then measures the proximity between those words and any words in their contexts. Words which are frequently found close together are considered to have a greater degree of association than those found together infrequently or farther apart.

3) The third step is situational modeling. QUORUM models represent the prominent elements of situations, and their prominent interactions. The basic form of a QUORUM model is a list of the most prominently associated word pairs, each with a number representing their degree of association. This list form of the model can also be represented as a matrix of association weights or a network of weighted links. The list, matrix, and network represent the individual components of association explicitly, and these can be inspected and modified in detail. This explicitness is not available in models based on neural networks or hyperdimensional similarity metrics.

4) The fourth step is relevance ranking, the sorting of narratives or sentences according to their relevance to the interests of the analyst. This ranking shows the analyst which narratives, and parts of narratives, describe recurring patterns. In this step, the associated word pairs in QUORUM models are used as relevance criteria.

These four steps of narrative interpretation are described in the following sections.

Step 1: Report Selection

Report selection is the first step of incident report interpretation. The process of selecting reports from a database varies according to how each database is managed. Since the work described here utilized the ASRS database, that method of selection will be described. These methods, however, apply in general terms to the selection of reports from any incident database.

In order to obtain a collection of incident reports for analysis, analysts typically provide selection criteria to the ASRS, often in the form of topics or key words of interest. Selection is required because the ASRS database contains tens of thousands of diverse reports, and analyses must address a much smaller number of reports in order to be focused and manageable. Selection criteria are usually interpreted by an ASRS database specialist who extracts the requested reports. This can be a very effective process because ASRS specialists are knowledgeable about the nature of the database and the concerns of requesters.

In practice, selection is based on the contents of ASRS-provided data fields associated with each narrative, the contents of the narrative itself, or both. For example, if desired, any appearance of a particular word in a narrative can trigger selection. Alternatively, if two or more key words appear in a narrative, that report might be considered appropriate for selection. Similarly, if two or more key words appear in a sentence within a narrative, that report could be selected. Other such selection criteria are also possible, as long as they are supported by the database software.

Report selection from the ASRS database can be an informal, verbal, and iterative process, or it can be based on a single, formal, written specification. Since selection criteria are typically interpreted by ASRS experts, precise details of the selection criteria are not always documented by those requesting the reports, as long as they get a collection which meets their needs. This, however, can make it difficult later if it is necessary to refine the search. In order to take full advantage of the QUORUM method, the report requester should take responsibility for knowing the exact nature of the selection criteria used to select the reports. In fact, the QUORUM method makes precise selection easier, as will be shown later.

Step 2: Narrative Analysis

Once the incident reports are collected, they are ready for analysis. In practice, some minor adjustments to the text are needed to aid computer-based analysis. These include, for example, converting any abbreviations containing punctuation, such as changing F/O (first officer) to FO. Both forms are found in ASRS narratives. All ASRS abbreviations used in this paper are expanded in the glossary.

The QUORUM method of narrative analysis is based on the supposition that the structure of narratives describing incidents reflects the structure of the incidents themselves. So, by measuring the text, the incidents are measured. This assumption is formally stated as a working hypothesis. The general form of the hypothesis is: The structure of a text reflects the structure of the domain described in the text, as indicated by the concerns of the author(s). The text in this case is one or more ASRS incident report narratives. The domain in this case is problematic episodes in commercial aviation operations. The authors of the narratives include airline pilots, air traffic controllers, and others. Their concerns include the details of specific incidents, the situations in which the incidents occurred, aviation safety in general, and personal responsibility in particular.

QUORUM narrative analysis consists of taking a large number of simple measurements of the text. Two kinds of measurements are taken. The first is a measurement of word frequency. The second is a measurement of contextual relatedness between words.

In the first set of measurements, instances of each distinct word are counted to determine how many times the word is found in the whole collection of narratives. This count is called the frequency of occurrence. For a collection of hundreds of ASRS narratives, this results in a list of thousands of words, each one with its frequency of occurrence. It might be, for example, that the word MODE occurs 368 times in 300 automation-oriented reports, while the word AILERON occurs 8 times. Words which appear more often in the collection of narratives are interpreted as representing important things, concepts, actions, attributes, or other aspects of the situations described. The counts of so-called "stop words" such as THE, AND, and TO are typically excluded from the list.

The second set of simple measurements indicates the degree to which pairs of words occur in the same context. For every occurrence of a word A, the proximity of a word B is added to the total proximity between A and B if word B is no farther than a certain distance (typically one average sentence length) from word A. This is called a proximity-weighted co-occurrence metric, and its magnitude is represented by a relational metric value (RMV). A typical analysis might involve computing this context metric for 125,000 word pairs. For example, in one collection of automation-oriented ASRS reports, the word DISCONNECTED is often found in the context of the word AUTOPLT (i.e., autopilot), so this word pair has a high relational metric value of 659. The word FO (i.e., first officer) is also found in the context of AUTOPLT, but to a lesser degree, as shown by the lower RMV of 248. The precise derivation and meaning of the RMV values are described in McGreevy (1996) and McGreevy (1995).

What is important here is that, in addition to a measure of the prominence of words, the method of analysis includes a quantitative measurement of the degree to which pairs of words occur in the same context in the narrative text. The magnitude of this relation, the RMV, is larger for word pairs which are closely associated in the narratives, and smaller for those which are less closely associated. The measurement is interpreted as being descriptive of the degree to which two concerns (represented by a pair of associated words) occur in the same situational context. So, for example, the close proximity of the words AUTOPLT and DISCONNECTED is found by measurement of the text, but it is interpreted as a measure of the close proximity of the system known as the "autopilot" and the action "disconnected" in the situation described by the text.

Step 3: Situational Modeling

The purpose of modeling a collection of incident narratives is to provide an accurate, explicit, and simplified representation of the incidents and situations described in the narratives. Interpreting the model can aid in understanding recurring patterns among the incidents themselves.

In general, QUORUM produces a sparse model of the prominent associations in a body of text. These prominent associations are interpreted as being indicative of the prominent concerns of the authors. When applied to incident narratives, the model is interpreted as a model of the incidents themselves. The model represents the aspects of the incidents which concerned the incident reporters.

The QUORUM situational model can take a variety of forms. Its most basic form is that of a table containing three columns: 1) prominent words from the text, 2) words which are often found in close proximity to the prominent words, and 3) the relational metric value (RMV), which indicates the magnitude of the relation between the words in columns 1 and 2. Here is a small example:

PANEL	CTL	589
MODE	ALT	504
ILS	RWY	480
AUTOPLT	ALT	478
MODE	CTL	472
AUTOPLT	DISCONNECTED	448
MODE	SELECTED	409
FMS	DSCNT	404

Larger RMVs indicate a greater degree of situational association. Each row in the table represents a proximity-weighted co-occurrence relation between the two words. Useful models typically contain hundreds of relations, but this is only a tiny fraction of the total number of possible inter-word relations in the analyzed texts.

The table of relations can also be represented as a matrix. In this form, it can be subjected to dimensional analysis, such as singular value decomposition (e.g., Deerwester, Dumais, Furnas, Landauer, and Harshman, 1990), or network

reduction, such as Pathfinder analysis (Schvaneveldt, Durso, and Dearholt, 1989). See McGreevy (1995) for a discussion of Pathfinder analysis applied to text analysis data.

The table of relations can also be represented as a network, providing a diagrammatic model of the situations. Each word in the table is represented as a node in the network, and the relation between each word pair is shown as a link. The network model can be displayed as an aid to analysis. An analyst familiar with the method can identify prominent sections of the model which represent prominent aspects of the situations in the collection of incident narratives. To make interpretation easier, each of the links in the network model (where each link represents a relationship between word pairs) can be illustrated with sentences taken directly from the original reports, as was done in McGreevy (1996), Appendix 2.

In any of its forms, the QUORUM situational model is very abstract, however, and its meaning can be difficult to appreciate. Even when illustrated with sentences from the narratives, the model can be difficult to interpret. Further, for detailed models, the list of relations is very long, the matrix is unwieldy, and the network is too complex to be neatly drawn.

One solution is to organize the detailed word-oriented model as an object-oriented model (McGreevy 1995, McGreevy 1996). This increases the clarity of the model by grouping related information in an intuitive, situation-based structure. To create the object-oriented model, however, the analyst must perform a semantic interpretation which requires significant knowledge of aviation operations and object-oriented analysis. Even then, the model is highly abstract.

Step 4: Relevance Ranking

While explicit models are useful for some researchers, the operationally-oriented analyst might find it cumbersome to use the models themselves. Relevance ranking, a step beyond modeling, enables QUORUM models to more effectively benefit the operational community. Using QUORUM models to relevance-rank text composed by the incident reporters themselves allows analysts to focus on descriptions of concrete incidents rather than diagrams of abstract models.

Relevance ranking is a process of sorting a list of items so that those likely to be of greater relevance to one's concerns and interests appear closer to the top of the list. Relevance ranking can help the analyst to efficiently read and interpret large collections of narratives. For example, in order to find episodes of greatest interest, it is useful to relevance-rank the narratives from a collection of narratives. Alternatively, to find complete thoughts of greatest interest, it is useful to relevance-rank all of the sentences from a collection of narratives. Sentences can also be ranked within a single narrative, as a way to summarize each narrative by presenting its most relevant sentences. Relevance ranking is further explained, and illustrated with examples, in the following sections.

Relevance Criteria

Relevance criteria determine what is considered to be of greater interest. Analysts usually start with an approximate notion of their relevance criteria, that is, of what constitutes relevance to their concerns. QUORUM's explicit model of relevance can help analysts develop very explicitly defined criteria, and allows them to readily refine those criteria.

In the QUORUM method of relevance ranking, sets of proximity-weighted co-occurrence relations are used as relevance criteria for ranking text items (e.g., narratives, paragraphs, sentences). Any set of QUORUM relations can be used as relevance criteria, but the criteria are typically a model or sub-model of a collection of text. Derivation of QUORUM models is explained in detail in McGreevy (1996) and McGreevy (1995).

Relevance criteria can be selected and fine-tuned to achieve various kinds of relevance ranking. These include: ranking by typicality, ranking by topical focus, ranking by multiple sets of criteria, ranking by externally derived criteria, ranking by example, and ranking by "outsider" criteria.

For any ranked text item, QUORUM can show the analyst the components of relevance and their relative contributions. That way, the analyst can decide whether the item is appropriately ranked, and, if necessary, can modify the relevance criteria accordingly.

Calculating Relevance Ranking Value (RRV)

The relevance ranking value (RRV) is the number by which text items (e.g., narratives, paragraphs, sentences) are ranked. The following equations are used to calculate the relevance ranking value for each text item:

$$RRV(t) = A * \sum_{r=0}^{N-1} RCV(r,t) \qquad (1)$$

$$RCV(r,t) = \frac{RMV(r,c) * RMV(r,t)}{B} \qquad (2)$$

where:

RRV(t): relevance ranking value of text item t

t: index of the text item to be ranked

N: total number of criterion relations (i.e., relevance criteria)

RCV(r,t): relevance component value of relation r in text item t

r: index of criterion relation, R(r), whose form is:

R(r): [PT, TIC]

PT: probe term, one of the most prominent words in the text

TIC: term-in-context, a word that is prominent in the context of the probe term

RMV(r,c): relational metric value whose magnitude indicates the degree of proximity-weighted co-occurrence between the two words in relation r, as measured in text collection c

c: index of the collection of text from which the criterion relations are derived

RMV(r,t): relational metric value whose magnitude indicates the degree of proximity-weighted co-occurrence between the two words in relation r, as measured in text item t

T(t): number of tokens in text item t; used to measure the length of the item; For narratives, T(t) is the number of words. For sentences, T(t) is the number of words and stand-alone punctuation marks(but could just as well be the number of words).

A: For sentences, A = 1. For narratives, A = 2000/T(t). For narratives, the parameter T(t) is applied here so that RCVs are all integers. The factor 2000 ensures that RRVs are all integers. The constant 2000 is used because it is larger than the number of words found in any narrative processed so far.

B: For sentences, B = T(t), the number of tokens in the sentence. For narratives, B = 1.

For details of deriving relational metric values (RMVs), whose magnitudes indicate the degree of proximity-weighted co-occurrence between words in a text, see McGreevy (1996) and McGreevy (1995).

Example of Relevance Ranking

Relevance ranking is illustrated here using a collection of 102 news stories. The stories were sampled between November 5 and December 16, 1996 using the San Jose Mercury News service, NewsHound. To be selected for analysis, each story had to contain at least one instance of the acronym "TWA." Most but not all of these stories are on

topics closely related to the explosion of TWA Flight 800 near Long Island, New York, on July 17, 1996. This collection is used because the story of Flight 800 is likely to be familiar to the reader, and this familiarity will make the examples more understandable. In particular, the reader is likely to have a sense of the relative prominence of various topics associated with the disaster. If a collection of ASRS incident reports had been used here, the reader would be unfamiliar with the events described and would find it more difficult to interpret the relevance criteria and judge the results of relevance ranking. Once the method is presented, subsequent examples illustrate the operational benefit of the QUORUM method by applying it to ASRS incident reports.

Relevance criteria— Relevance criteria are the relations which determine how text items will be ranked. The relations can come from a variety of sources. In this example they are derived from the 102 news stories, most of which are about Flight 800. Shown below is a sample of the 280 relations that are most prominent in the collection of news stories. (The line containing "..." indicates that some of the relations are not shown.) The 280 relations constitute a QUORUM model of the news stories. For details on deriving QUORUM models from text, see McGreevy (1996) and McGreevy (1995).

In the table below, the number in the third column is the relational metric value (RMV) of each relation. Its magnitude indicates the degree of association between the two words, PT and TIC, in column 1 and 2, as measured in the collection of news stories. For example, the contextual association between "Flight" and "800" is very prominent in the text, as indicated by the RMV of 1725, while the contextual association between "safety" and "board," with its RMV of 158, is only somewhat prominent, and is the least prominent of the relations in the model. (Notes: If both words in a relation are probe terms, the most frequently occurring word in the word pair is shown in the probe term column. Words like "Board" and "board" are treated as distinct words because capitalizations found in the text are retained in this analysis.)

probe term (PT)	term in context (TIC)	RMV(r,c)
Flight	800	1725
TWA	Flight	1486
TWA	800	1461
fuel	tank	1115
New	York	990
fuel	center	894
United	States	865
fuel	tanks	849
bomb	missile	752
Long	Island	720
tank	center	702
Aviation	Federal	693
air	traffic	684
Federal	Administration	668
Aviation	Administration	662
Airport	International	656
July	17	640
TWA	crash	604
National	Transportation	602
Safety	Transportation	600
year	last	594
National	Safety	589
airport	security	584
Safety	Board	580
people	killed	555
TWA	explosion	554
people	230	541
Transportation	Board	532
National	Board	522
Airport	Kennedy	518
bomb	mechanical	455
...		
tanks	cooler	159
safety	flight	159
people	died	159
last	flight	159
Kennedy	minutes	159
FBI	Kallstrom	159

Airport	takeoff	159
Air	British	159
safety	board	158

Ranked text items— When used as relevance criteria, the relations of any QUORUM model can be used to relevance-rank any collection of text items, such as narratives, paragraphs, or sentences. As an example, the 280 relations of the preceding model are used here to relevance-rank all of the sentences contained in the 102 news stories. Because the model represents the most prominent relations in the whole collection, it ranks sentences according to typicality. That is, sentences with the highest relevance ranking values (RRV) are most representative of the entire collection of text. As such, they contain the main themes in the collection of stories.

Shown below are the 10 most typical sentences from the 102 news stories, according to QUORUM relevance ranking. Review of these sentences suggests that these sentences are, in fact, typical of the sentences contained in the whole collection of stories. The first sentence ("Mysterious explosion on TWA Flight 800 to Paris kills 230.") does seem to be representative of the entire collection in that it contains the main points of the news stories. It could serve as a headline for the whole collection.

RRV	line#	index	sentence
11967	_1763_	1996Dec032_18	Mysterious explosion on TWA Flight 800 to Paris kills 230 .
10281	_3110_	1996Nov054_18	And now TWA Flight 800 is the latest unsolved crash .
9995	_2973_	1996Dec103_6	Security concerns have been heightened since the July explosion of TWA Flight 800 , killing 230 people off New York's Long Island .
9197	_3487_	1996Dec142_1	Static electricity latest focus of TWA Flight 800 probe .
8767	_2750_	1996Nov222_8	Attention was riveted on airport security after TWA Flight 800 blew up in July and killed 230 people .
8283	_3618_	1996Dec144_2	Families of victims of last summer's crash of TWA Flight 800 said Saturday they will press TWA for immediate compensation .
8130	_3596_	1996Dec112_2	The National Transportation Safety Board's ongoing investigation into the explosion of TWA Flight 800 has landed at Honeywell .
8101	_3585_	1996Nov057_24	That night , TWA Flight 800 crashed off the coast of Long Island , killing all 230 people on board .
7664	_3435_	1996Nov243_1	NTSB Has Yet To Interview Ground Crews From TWA Flight 800 .
7445	_2614_	1996Nov272_3	The information it has been able to develop so far about ValuJet Flight 592 and TWA Flight 800 is staggering .

(If any sentence appeared more than once among the news stories, some of which appeared on the news wires more than once, only one representative is shown. Column 2, line number, counts lines within the collection. The index, column 3, is of the form YYYYMMMDDN_L, where YYYY is year (e.g., 1996), MMM is month (e.g., Dec), DD is day (e.g., 03), N is story number that day, L is line number within the story. Punctuation is spaced for processing, and stand-alone punctuation marks such as commas and periods count here as tokens in evaluating T(t), although one could just as well count only words.)

QUORUM relevance ranking is based on all of the relations in the list of relevance criteria. Collectively, the relations detect all sorts of proximities among words, such as pairs of words which are often found in the same general vicinity, several words appearing in loose clusters, and tightly coupled groups of words which are always found right next to each other in the same order. An example of the latter is the word group, "TWA Flight 800." QUORUM analysis recognizes this cluster of words as important in the collection. That recognition is reflected in the prominence of the relations [Flight,800], [TWA, Flight], and [TWA, 800] in the model (shown earlier) of the 102 news stories. In fact, these three relations are the most prominent ones in the model.

It is important to appreciate the fact that QUORUM relevance ranking uses all of the pairwise relations in the model as relevance criteria when ranking text items. The presence or absence of one particular group of words does not, by itself,

determine relevance. Here, for example, are the six most relevant sentences that do not contain "TWA Flight 800." Despite having no explicit mention of "TWA Flight 800," these sentences are still recognized by QUORUM as being highly relevant to the main themes of the story about Flight 800.

RRV	line#	index	sentence
4999	_3621_	1996Dec144_5	On Friday the National Transportation Safety Board said a buildup of volatile vapors in the 747's partially full center fuel tank could have triggered the explosion .
4870	_3617_	1996Dec144_1	TWA 800 families want compensation .
4837	_3460_	1996Dec1310_3	The National Transportation Safety Board cautioned that no conclusions have been reached in the July 17 midair explosion that killed 230 people on their way to Paris from New York .
4835	_3360_	1996Dec124_16	The FBI and the National Transportation Safety Board are still investigating three theories : a missile , a bomb and mechanical failure .
4391	_3209_	1996Nov055_32	Meanwhile , National Transportation Safety Board officials are pursuing and testing their own theories : that a defective fuel pump , fuel probe or other source of a spark ignited the center fuel tank and destroyed the plane .
4267	_3241_	1996Dec137_13	All 230 people aboard were killed .

Further, the fact that a sentence does contain a prominent word group such as "TWA Flight 800" is not sufficient to make that sentence highly relevant. To illustrate this, the two least relevant sentences that contain "TWA Flight 800" are shown below. The magnitudes of the relevance ranking values (RRV) indicate that these sentences have some relevance, but that they are not as relevant as those above.

RRV	line#	index	sentence
2451	_3325_	1996Nov201_2	When Pierre Salinger charged that TWA Flight 800 was brought down by " friendly fire , " he bolstered it with a claim that an Air France jet had to swerve wildly to avoid a missile that same night .
1991	_2368_	1996Dec153_8	And as much as you may be put off by the author's heavy-handed plotting , you are elated to be overcoming the frustrations recently visited by the mystery of TWA Flight 800 , and to be able to track the elusive cause of an air disaster .

Recall that for a news story to be included in the collection, it was only necessary that the word "TWA" appear somewhere in the story. The most prominent topic among the 102 news stories is the disaster of Flight 800, but some of the stories focus on aviation safety, the airline business, or other airline issues. Some stories make only a passing reference to TWA. Thus, not all of the stories in the collection are about TWA Flight 800 itself. Accordingly, the sentences contained in the 102 news stories vary greatly in relevance.

To further illustrate that QUORUM properly ranked the sentences on their relevance to the main themes of the collection, ten sentences, each containing "TWA," are shown below. The relevance ranking values associated with these sentences span a wide range. Sentences toward the top of the list have higher relevance ranking values (RRV), while those toward the bottom of the list have lower relevance ranking values. The sentences toward the top of the list are more relevant to the disaster of Flight 800, which is the most prominent theme in the collection. Sentences toward the bottom of the list are less relevant to the disaster of Flight 800. This further demonstrates that QUORUM does indeed rank the sentences according to relevance.

RRV	line#	index	sentence
6273	_3024_	1996Dec057_9	President Clinton launched the Aviation Safety and Security Commission last summer after the unexplained explosion of TWA Flight 800 off New York's Long Island coast .

5426	_2457_	1996Nov215_8	InVision saw its stock rise sharply after TWA Flight 800 plunged into the Atlantic Ocean off Long Island , NY , on July 17 .
4573	_3352_	1996Dec124_8	East Hampton is about 30 miles east of Center Moriches , the point of land closest to where TWA Flight 800 went down July 17 .
3499	_3346_	1996Dec124_2	A Saudi Arabian Airlines pilot flying in the area where TWA Flight 800 exploded reported seeing " a green flare " in the sky early Thursday that authorities could not immediately identify .
2716	_3079_	1996Dec043_3	James Kallstrom , who is leading the criminal probe into the explosion of TWA Flight 800 , said terrorism has come a long way since the 1970s , when bombs often were directed at real estate , " bricks and mortar . "
1624	_173_	1996Dec091_12	In 1960 , 134 people were killed when a United Air Lines DC-8 and a TWA Super Constellation collided over New York City .
776	_3588_	1996Nov057_27	Crawford said , though , most customers don't blame TWA for the crash .
348	_2909_	1996Nov217_5	The pilot of the small plane , former TWA flight engineer Neal Reinwald , was giving a lesson to an Illinois woman at the time of the crash , the St Louis Post-Dispatch reported for Thursday editions .
151	_1543_	1996Dec082_31	It is because of the trend toward regional domination that TWA , United and other airlines are calling on the US government to take a close look at the long-term impact of an American-BA alliance .
0	_1260_	1996Nov211_27	Douglas' largest sales this year have been a 15-plane sale of new MD-80s to TWA and a five-plane sale of MD-11 freighters to Lufthansa German Airlines .

Finally, shown below are some of the sentences that QUORUM ranked as having the least relevance. Of all the sentences in the collection of news stories, the first sentence, "The committee... ," has the lowest non-zero relevance ranking value. Its RRV is 4. That minimal value is due to the relation between the words "security" and "airports." (The components of relevance of particular sentences are discussed in the next two sections.) The sentence, "Let's hope... ," is also related to airport security, but it is verbose, contains little useful information, and is barely relevant to Flight 800. QUORUM, using the 280 most prominent relations in the news stories, finds no relevance in this sentence. The remaining sentences are clearly irrelevant. The last two sentences were contained in news stories consisting of several short summaries on various topics. Since one of the summaries mentioned TWA, the whole story was included in the collection. QUORUM correctly recognizes that these sentences are irrelevant to the main themes of the 102 news stories.

RRV	line#	index	sentence
4	_2494_	1996Dec123_6	The committee suggested greater use of high-tech equipment , bomb-sniffing dogs and trained security managers to detect explosive devices and materials among cargo , mail , baggage , carry-ons and travelers at US airports .
0	_2652_	1996Nov272_41	Let's hope the electronic sniffer can tell the differences among Faberge's Babe perfume , the fragrance of wet Converse shoes , the subtle aroma of plastic explosives and the personal redolence that's an inevitable consequence of hastening to the airport two hours early , schlepping the bags past curb-side porters , running the gantlet of metal detectors and getting bumped from the last flight to your home town .
0	_1074_	1996Nov263_21	Oakland Airport , where the number of passengers has jumped 148 percent since 1988 , is planning to build a new

```
                                multilevel parking garage and expand roadways into the
                                airport to cope with the growing demand .

0    _3249_    1996Dec012_8     The twister struck as a dangerous weather front moved
                                eastward Saturday across the South , spawning several other
                                tornadoes that caused major property damage .

0    _1660_    1996Dec166_50    Many of us didn't really understand the new law that
                                dismantled the 6-decade-old welfare system that has long
                                guaranteed a federal safety net to needy people .
```

The examples shown above suggest that QUORUM appropriately ranks text items according to relevance. This ranking is based on the relevance criteria, typically a QUORUM model. In this example, the relevance criteria represent the main themes in the collection of news stories. In general, relevance ranking is most useful when the relevance criteria reflect the particular concerns and interests of the analyst using the ranking. Six examples of this are provided later in the section, "Options: Choosing how Text is Ranked." After that, refinement of relevance criteria is discussed in the section, "A Closer Look at QUORUM Relations." First, however, it is important to understand how the relevance ranking value is calculated for each text item.

Description of calculation of a relevance ranking value— Using equations 1 and 2, and the 280 relevance criteria, the relevance ranking value (RRV) can be found for any text item. As an example, the RRV is found here for the most typical sentence in the collection of news stories:

```
"Mysterious explosion on TWA Flight 800 to Paris kills 230 ."
```

QUORUM first determines that the sentence contains 10 of the 280 relevance criteria. Each criterion relation, R(r), includes its degree of association, RMV(r,c), as measured in collection c, the 102 news stories. See McGreevy (1996) and McGreevy (1995) for details of how to measure RMVs in collections of text.

R(r)	PT	TIC	RMV(r,c)
R(0)	Flight	800	1725
R(1)	TWA	Flight	1486
R(2)	TWA	800	1461
R(3)	TWA	explosion	554
R(5)	Flight	explosion	408
R(4)	800	explosion	415
R(7)	800	230	249
R(6)	TWA	230	274
R(8)	Flight	230	237
R(9)	explosion	230	205

QUORUM then measures the degree of association, RMV(r,t), of each relation R(r) in text item t, in this case the sentence, "Mysterious explosion... ." These values are shown in the first column of the table below. See McGreevy (1996) for details of how to measure RMVs in a single sentence.

RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
20	R(0)	Flight	800	1725
20	R(1)	TWA	Flight	1486
19	R(2)	TWA	800	1461
19	R(3)	TWA	explosion	554
18	R(5)	Flight	explosion	408
17	R(4)	800	explosion	415
17	R(7)	800	230	249
15	R(6)	TWA	230	274
16	R(8)	Flight	230	237
13	R(9)	explosion	230	205

After counting the number of tokens in the sentence (in this case, T(t)=11), QUORUM finds the relevance component value (RCV) for each of the relations. This value is the product of RMV(r,c) and RMV(r,t) for each relation, divided by T(t). So for example,

$$RCV(0,t) = RMV(0,t) * RMV(0,c) / T(t) = 1725 * 20 / 11 = 3136$$

The rest of the relevance component values are computed in a similar manner, and all are shown in the first column of the table below.

RCV(r,t)	RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
3136	20	R(0)	Flight	800	1725
2701	20	R(1)	TWA	Flight	1486
2523	19	R(2)	TWA	800	1461
956	19	R(3)	TWA	explosion	554
667	18	R(5)	Flight	explosion	408
641	17	R(4)	800	explosion	415
384	17	R(7)	800	230	249
373	15	R(6)	TWA	230	274
344	16	R(8)	Flight	230	237
242	13	R(9)	explosion	230	205

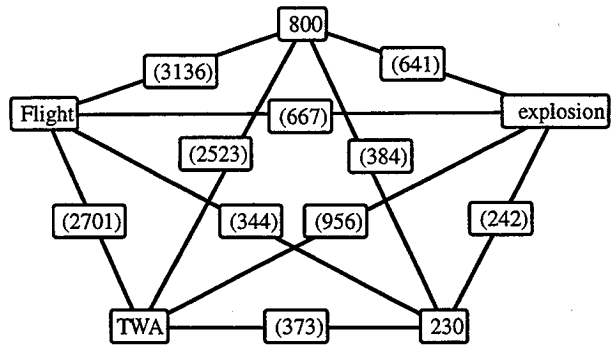
RRV = 11967

The relevance ranking value (RRV) for the sentence is then found by taking the sum of the values in the first column. This results in an RRV value of 11967 for the sentence, "Mysterious explosion... ."

By calculating the RRV for each text item, and sorting the items on the RRV, the text items are ranked according to their relevance to the relevance criteria. Since the example sentence had the highest RRV (11967), it is considered to be the text item that is most relevant to the relevance criteria.

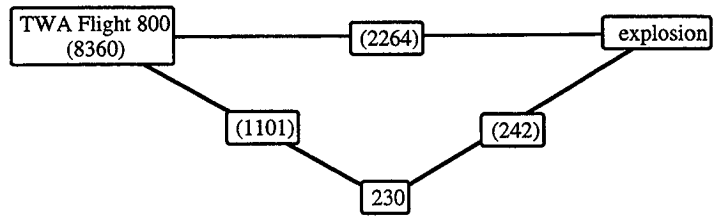
Components of relevance— The table developed in the previous section contains the "components of relevance" of the sentence, "Mysterious explosion... ." The components of relevance indicate the exact nature of the measured relevance. For that reason, they are shown in a number of examples throughout the rest of the paper.

For a more intuitive view of the components of relevance, a network can be drawn to represent them. The network below represents the components of relevance developed in the previous section. The values shown on the links are the RCV(r,t) values from column 1 of the table, associating each pair of nodes, PT and TIC.



Shown below is a simplified network representation, created by treating "TWA Flight 800" as a single unit and combining link weights. The link weights among "TWA," "Flight," and "800" are summed and shown within the "TWA Flight 800" node. Link weights involving the three words are summed, so that, for example, the link weight of the relation [TWA Flight 800, explosion] is 956 plus 667 plus 641, which is equal to 2264. This diagram clearly illustrates the relevance components of the most typical sentence among the 102 news stories,

"Mysterious explosion on TWA Flight 800 to Paris kills 230 ."



Here is another example of a text item, its components of relevance, and its relevance ranking value (RRV). The sentence is:

"Attention was riveted on airport security after TWA Flight 800 blew up in July and killed 230 people ."

This sentence has 19 tokens, counting the period, so T(t) equals 19.

Shown below are the components of relevance of this sentence, based on the 280 relevance criteria, and equations 1 and 2. The relevance ranking value (RRV) of the sentence is the sum of the values of column 1.

RCV(r,t)	RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
1815	20	R(0)	Flight	800	1725
1564	20	R(1)	TWA	Flight	1486
1461	19	R(2)	TWA	800	1461
614	20	R(3)	airport	security	584
569	20	R(4)	people	230	541
555	19	R(5)	people	killed	555
283	20	R(6)	killed	230	269
227	18	R(7)	July	230	240
214	17	R(8)	people	July	240
202	15	R(9)	TWA	July	257
193	17	R(10)	800	July	216
183	14	R(11)	800	230	249
173	12	R(12)	TWA	230	274
169	16	R(13)	Flight	July	201
162	13	R(14)	Flight	230	237
141	12	R(15)	people	Flight	224
138	13	R(16)	people	800	203
104	11	R(17)	TWA	people	181

RRV= 8767					

Relevance Density

Strictly speaking, when the lengths of text items are taken into consideration, as in the preceding sections, text items are ranked on "relevance density." The default used in QUORUM relevance ranking is to rank on relevance density. As a result, more concise text items are considered more relevant.

The most relevant sentence, based on typicality and relevance density, was found in the preceding section to be:

"Mysterious explosion on TWA Flight 800 to Paris kills 230 ."

This sentence was a headline contained in an item on December 3 describing candidates for the top stories of 1996. Since this is a headline, it is both concise and representative of the main points of the collection of stories about Flight 800.

Some analyses, however, can benefit from measuring relevance without consideration of the length of the text item. In this case, longer, possibly more detailed items are considered more relevant.

The most relevant sentence, based on typicality but without consideration of the number of tokens in the text item, is the lead sentence from a story on December 13.

"The National Transportation Safety Board on Friday issued several urgent recommendations to the Federal Aviation Administration to protect fuel tanks from heat sources that could touch off the kind of explosion that occurred with TWA Flight 800 ."

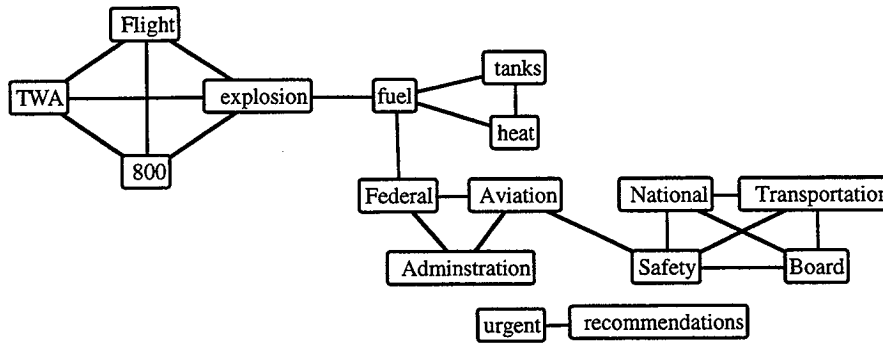
Since this is the lead sentence in a story, it is not as constrained as a headline with respect to length, but it contains the main points of the particular story. Thus, this sentence is longer than the headline "Mysterious explosion..." and has more details.

Here are the components of relevance of the longer sentence.

RCV(r,t)	RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
907	20	R(0)	Flight	800	1725
782	20	R(1)	TWA	Flight	1486
730	19	R(2)	TWA	800	1461
446	20	R(3)	fuel	tanks	849
364	20	R(4)	Aviation	Federal	693
348	20	R(5)	Aviation	Administration	662
334	19	R(6)	Federal	Administration	668
316	20	R(7)	National	Transportation	602
315	20	R(8)	Safety	Transportation	600
305	20	R(9)	Safety	Board	580
294	19	R(10)	National	Safety	589
266	19	R(11)	Transportation	Board	532
247	18	R(12)	National	Board	522
247	17	R(13)	TWA	explosion	554
171	16	R(14)	Flight	explosion	408
163	15	R(15)	800	explosion	415
132	20	R(16)	recommendations	urgent	252
98	19	R(17)	tanks	heat	197
90	18	R(18)	fuel	heat	190
78	9	R(19)	fuel	explosion	333
72	16	R(20)	fuel	Federal	171
49	10	R(22)	Aviation	Safety	187

RRV=6754					

Here is a network diagram of the components of relevance shown in the table above. Note how QUORUM automatically detects recurring clusters of words.



To find RRV' for a sentence, the relevance without consideration of the length of the text item, RRV can be multiplied by T(t), the number of tokens in the text item. In this case:

```
RRV' = RRV * T(t) = 6754 * 38 = 256652
```

Unless specifically noted, relevance ranking in this paper is based on relevance density.

Options: Choosing how Text is Ranked

Relevance ranking of text using QUORUM is a flexible process. It is designed to be adaptable to a wide variety of particular interests. The greatest flexibility is in the selection and use of relevance criteria. Relevance criteria characterize the interests and concerns of the analyst, and determine how the text is ranked. In the sections which follow, these relevance ranking options are illustrated:

1. Ranking by typicality
2. Ranking by topical focus
3. Ranking by multiple sets of criteria
4. Ranking by externally derived criteria
5. Ranking by example
6. Ranking by "outsider" criteria

The other significant flexibility in relevance ranking is the selection of what text items to rank. In the sections which follow, these examples are illustrated:

- ranking sentences within a collection of narratives
- ranking narratives within a collection of narratives
- ranking sentences within each narrative

Option 1: Ranking by Typicality

An analyst investigating a thematically related collection of text items might want the most typical ones listed first because they would be highly representative of the whole collection. Knowing which text items are most typical of a collection can greatly increase the efficiency of the analyst in interpreting the collection. The QUORUM model of a whole collection represents the relevance criteria which determine "typicality." That is, if a text item is relevant to the prominent concerns expressed in the whole collection, it is said to be typical of the collection.

Flight 800 example— In an earlier section of this paper, "Example of Relevance Ranking," an example was used to demonstrate the process of relevance ranking, and to show how relevance ranking values are computed. In that example, the relevance criteria were 280 relations representing the whole collection of 102 news stories about Flight 800. These relations were used to rank all of the sentences contained in the collection. Thus, that example illustrates ranking of sentences by typicality.

ASRS example— The rest of this section is an example of ranking sentences and narratives from ASRS reports according to typicality. First, a model of the whole collection is obtained for use as relevance criteria. Next, the sentences are ranked according to the relevance criteria. Finally, the narratives are ranked according to the relevance criteria. (All ASRS abbreviations are expanded in the glossary.)

Shown below are some of the 300 relations of a QUORUM model representing a collection of 313 automation-error incidents from the ASRS database. (Lines containing "..." indicate that some of the relations are not shown.) In this example, these relations are used as the relevance criteria.

This model includes many relations which might be called "domain generic" in that they are exceedingly common in commercial aviation situations. For example, the relation [FT, ALT] indicates that the most closely associated words in this collection of narratives are FT (i.e., feet) and ALT (i.e., altitude). While generic, this relation does indicate a pervasive concern in these narratives with specific altitude in feet. Scattered among these relations are some which are more specific to automation, such as the relation [AUTOPLT, DISCONNECTED]. That is, there is a prominent association between "autopilot" and "disconnected" in the analyzed reports. Relations in which one of the words is NOT or BUT are often associated with problematic situations.

probe term (PT)	term in context (TIC)	RMV(r,c)
FT	ALT	2462
ACFT	NOT	1534
FT	10000	1488
ACFT	FT	1131
NOT	BUT	1003
FT	MSL	963
FT	DSCNT	870

...		
LNDG	GEAR	518
FT	DEP	515
FT	1000	514
ACFT	FO	507
ALT	MODE	504
CAPT	FLYING	498
...		
FT	ATC	456
AUTOPLT	DISCONNECTED	448
ALT	CLB	444
...		
ACFT	FUEL	249
FT	SELECTED	248
ACFT	WITHOUT	248
ACFT	TKOF	248
NOT	2	247
FT	2	247
DSCNT	FO	247
NAV	VERT	246

When used as relevance criteria, the 300 relations sampled in the above table can be used to rank text items from the 313 reports on typicality. Here are the five most typical sentences. That is, these sentences are most representative of the concerns expressed in the whole collection.

- APCH THEN CLRED US FROM 11000 FT MSL TO 10000 FT MSL . (rpt# 295961)
- AT ABOUT 30000 FT THE CABIN ALT REACHED 10000 FT WITH A WARNING LIGHT . (rpt# 260523)
- ACFT DID NOT CAPTURE ALT . (rpt# 314310)
- ALT BUST - ASSIGNED 9000 FT DSNDED THROUGH 9000 FT TO ABOUT 8600 FT . (rpt# 312900)
- JUST PRIOR TO TOUCHDOWN WIND GUST AND/OR THERMAL ACTIVITY CAUSED ACFT TO CLB FROM 10 FT RADIO ALT TO 30 FT RADIO ALT . (rpt# 274159)

Shown in the table below are the report numbers (in the last column) of the top five narratives, ranked on typicality (column one, the relevance ranking value, RRV). RRV' is the relevance ranking value without consideration of the number of tokens in the text item. T(t) is the number of tokens in text item t (i.e., each narrative). For narratives, T(t) is the number of words.

RRV	RRV'	T(t)	rpt#
14938322	231544	31	312900
9424981	499524	106	264689
8896336	422576	95	309840
8848010	822865	186	162356
8793094	233017	53	315410

Here are the three most typical and concise narratives in the collection.

narrative from ASRS report number 312900:
ALT BUST -- ASSIGNED 9000 FT DSNDED THROUGH 9000 FT TO ABOUT 8600 FT. FO FLYING AND LOOKING FOR ATC CALLED TFC, HAND FLYING. FLT MGMNT COMPUTER DID NOT ALERT BUSTING ALT.

narrative from ASRS report number 264689:
WHILE FLYING FLT FROM MSP TO SAN WE WERE GETTING NUMEROUS ALT CHANGES AND TA'S. THE CAPT WAS FLYING. WE WERE CLRED TO 12000 FT. I SAW HE HAD SELECTED 11900 FT BUT WAS REACHING UP TO CORRECT IT. I PROCEEDED TO GET THE NEW ATIS. THE CAPT SET 12000 FT IN THE ALT WINDOW, BUT ON THE A-320, SETTING A NEW ALT WHEN WITHIN 300 FT OF THE OLD ALT PUTS YOU IN A VERT SPD MODE AND YOU WILL MISS YOUR ALT. WE CAUGHT IT AND CORRECTED AT 11600 FT -- 400 FT BELOW ASSIGNED. THE CTLR DID NOT INDICATE A CONFLICT OR ANY CONCERN.

narrative from ASRS report number 309840:
DURING CLBOUT FROM MDT, WE WERE GIVEN 8000 FT ALT ASSIGNMENT. NEARING 5000 FT DEP TOLD US TO MAINTAIN 5000 FT FOR A SINGLE ENG LIGHT ACFT AT ABOUT 6000 FT VFR. WE LEVELED AT 5000 FT AND THEN I RPTED TFC. DEP THEN TOLD US TO CLB AND MAINTAIN 8000 FT. AS WE CLBED WE HAD

```
A TA THEN AN RA FROM TCASII. THE OTHER ACFT WAS IN SIGHT THE ENTIRE TIME AND PASSED VERT
AND HORIZ AS STATED ABOVE. WE IGNORED THE RA SINCE THE ACFT WAS IN SIGHT AND PASSED
BEHIND US.
```

Shown below are the components of relevance of narrative 312900, based on the 300 relevance criteria and equations 1 and 2. Narrative 312900 is a good representative of the whole collection largely because of the prominent "domain generic" relations, such as [FT, ALT], and the fact that the narrative is very concise. The next section, "Ranking by Topical Focus," shows how to rank text items according to a more specific, selected set of criteria.

RCV(r,t)	RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
73860	30	R(0)	FT	ALT	2462
40068	84	R(1)	FT	9000	477
14792	43	R(2)	FT	DSNDED	344
12888	36	R(3)	FT	ASSIGNED	358
10944	24	R(4)	FT	ATC	456
10440	18	R(5)	FT	TFC	580
10192	26	R(6)	FO	FLYING	392
10140	39	R(7)	FT	FO	260
8904	21	R(8)	FT	CALLED	424
8680	14	R(9)	NOT	FLT	620
7230	15	R(10)	NOT	ALT	482
5712	16	R(11)	ALT	ASSIGNED	357
4708	11	R(12)	FLT	ALT	428
4410	10	R(13)	FT	FLT	441
2700	9	R(14)	NOT	ATC	300
2440	8	R(15)	FLT	FO	305
1896	3	R(16)	FT	NOT	632
1540	4	R(17)	NOT	FO	385

```
RRV'=231544
RRV =231544 * 2000/31 = 14938322
```

In the case of narratives, the sum of the relevance component values (RCVs) is RRV', the relevance ranking value without consideration of length of the narrative. Recall that RRV is the relevance ranking value based on relevance density. For this narrative, RRV is equal to RRV' divided by 31, the number of words in the narrative, times 2000, a numerical factor used for narratives, as explained in the section, "Calculating relevance ranking value (RRV)."

Option 2: Ranking by Topical Focus

QUORUM relevance criteria can be focused on a particular topic by retaining topical relations and deleting the others. Topically focused relevance criteria are used to rank text items according to their relevance to particular topics.

Flight 800 example— In the Flight 800 news stories, it is possible to focus on topics such as the FBI and its chief spokesman, James Kallstrom. Here are some of the 241 relations of a QUORUM model of that topic. (The line containing "..." indicates that some of the relations are not shown.) Note that the most prominent words in the context of "FBI" are "crash" and "Kallstrom."

probe term (PT)	term in context (TIC)	RMV(r,c)
FBI	crash	205
FBI	Kallstrom	159
FBI	James	146
FBI	TWA	135
FBI	investigating	126
Kallstrom	criminal	115
FBI	Assistant	100
FBI	missile	98
FBI	interviews	97
FBI	Director	95
FBI	NY	93
FBI	director	92
FBI	criminal	89
FBI	board	83
FBI	assistant	77

FBI	spokesman	73
FBI	theories	71
Kallstrom	evidence	70
FBI	800	66
Kallstrom	crash	66
...		
FBI	conducted	18
FBI	cost	18
FBI	incriminating	18
FBI	investigated	18
FBI	law	18
FBI	law-enforcement	18
FBI	nonsense	18
FBI	officer	18
FBI	planning	18
FBI	record	18

Shown below are the five sentences that are most focused on the topic of FBI + Kallstrom, based on relevance density. That is, these are the most relevant and concise sentences on the topic of FBI + Kallstrom among the 102 news stories about Flight 800.

- FBI considers sabotage in TWA crash .
- FBI considers saboteur theories on TWA crash .
- Early in the investigation , FBI Assistant Director James Kallstrom was asked if a meteor could have downed Flight 800 .
- James Kallstrom , the FBI assistant director who is leading the criminal investigation of the crash , said only that the bureau is pursuing every scenario .
- FBI Assistant Director James Kallstrom , who heads the agency's criminal probe of the disaster , said he remains confident there will be an answer .

Here are the five sentences that are most focused on the topic of FBI + Kallstrom, based on relevance without consideration of sentence length. That is, these are the most relevant to the topic of FBI + Kallstrom, but not the most concise sentences among the 102 news stories about Flight 800.

- The FBI is still investigating whether a bomb or missile downed TWA flight 800 even though transportation officials lean toward mechanical failure as the cause of the crash , a FBI spokesman said Saturday .
- James Kallstrom the FBI assistant director who is leading the criminal investigation of the crash , said only that the bureau is pursuing every scenario .
- James Kallstrom , who is heading the FBI criminal probe into the crash , said Saturday he agrees with the NTSB recommendations , but is critical of those who are " speculating publicly on what caused this horrific tragedy . "
- The remarks by FBI Assistant Director James Kallstrom reflect the growing belief among investigators that a mechanical malfunction caused the center fuel tank to explode July 17 before the jet smashed into the Atlantic Ocean off Long Island shortly after takeoff from John F Kennedy International Airport .
- FBI Assistant Director James Kallstrom , who heads the agency's criminal probe of the disaster , said he remains confident there will be an answer .

The table below contains the relevance components of the sentence, "The FBI is still investigating... ." The top five words found in the context of "FBI" in the 102 news stories are "crash," "Kallstrom," "James," "TWA," and "investigating." The prominence of three of these relations in this sentence (i.e., [FBI, crash], [FBI, investigating], and [FBI, TWA]), as well as others, make it highly relevant to the topical focus on FBI + Kallstrom.

RCV(r,t)	RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
105	18	R(0)	FBI	crash	205
64	18	R(1)	FBI	investigating	126
50	13	R(2)	FBI	TWA	135
41	20	R(3)	FBI	spokesman	73
36	13	R(4)	FBI	missile	98

24	15	R(6)	FBI	bomb	57
24	13	R(7)	FBI	800	66
22	13	R(8)	FBI	mechanical	60
19	13	R(9)	FBI	failure	52
16	15	R(10)	FBI	cause	39
13	13	R(11)	FBI	flight	35
13	13	R(12)	FBI	downed	36
10	17	R(13)	FBI	whether	21
7	13	R(14)	FBI	lean	19

ASRS example— The rest of this section is an example of finding, in a collection of ASRS reports, the sentences and narratives which are most relevant to a particular topic. For example, relevance criteria can be focused on automation concerns by selecting only automation-oriented relations. Shown below are some of the 256 relations in an automation-oriented model of 313 automation-error narratives from the ASRS database. (Lines containing "..." indicate that some of the relations are not shown.) Unlike the typicality model of this collection that was discussed in a preceding section, this model is focused on the topic of automation. When used to relevance-rank text items, the relations in this model serve as topically focused relevance criteria.

probe term (PT)	term in context (TIC)	RMV(r,c)
PANEL	CTL	589
MODE	ALT	504
ILS	RWY	480
AUTOPLT	ALT	478
MODE	CTL	472
AUTOPLT	DISCONNECTED	448
MODE	SELECTED	409
FMS	DSCNT	404
MODE	SPD	400
CAPTURE	ALT	399
FMC	NOT	393
...		
MODE	VERT	268
AUTOPLT	FLT	265
SYS	PWR	262
SYS	NOT	259
SYS	FLT	258
MODE	NAV	251
...		
AUTOPLT	CAPTURE	166
SELECTED	CAPT	165
SELECTED	LOC	163
AUTOPLT	TRIM	162
...		
FMC	ENTERED	91
DME	RWY	91
DATA	BEFORE	91
FMC	CTL	90
DISPLAY	ALT	90
LOC	NOT	89
FMS	WDB	89
DME	APCH	89

Here are the 5 most representative automation-oriented sentences in narratives of the 313 reports, based on the automation-focused relevance criteria.

- AT APPROX FL320 THE FO SELECTED THE VERT SPD MODE ON THE MODE CTL PANEL AND SELECTED A HIGHER SPD , THUS SLOWING THE CLB RATE . (rpt# 304278)
- I SELECTED FLT LEVEL CHANGE ON THE MODE CTL PANEL TO CONTINUE DSCNT . (rpt# 294000)
- UPON DISENGAGING THE AUTOPLT THE ALT SELECT INDICATOR AND EADU WOULD NOT INDICATE SELECTED ALT OR ANY MODE OF THE FLT DIRECTOR . (rpt# 270213)
- WHEN LOC WAS SELECTED ON THE MODE CTL PANEL THE ACFT BANKED L . (rpt# 186479)

- IN AUTOPLT MODE FLCH FLT LEVEL CHANGE, INSTEAD OF VNAV , FO DIALED IN 1900 FT ON ALT WINDOW IN MODE CTL PANEL . (rpt# 318230)

Shown in the table below are the report numbers of the top five narratives, ranked on relevance to the automation-oriented relevance criteria. The relevance ranking value (RRV) is shown in column one. RRV' is the relevance ranking value without consideration of the number of tokens in the text item. T(t) is the number of tokens in text item t (i.e., each narrative). In narratives, T(t) is the number of words.

RRV	RRV'	T(t)	rpt#
2371063	278600	235	317930
1814151	203185	224	294000
1771101	252382	285	259800
1745406	55853	64	91522
1614465	104133	129	314310

Here are the two most relevant narratives among the 313 narratives of the collection, based on the automation-focused relevance criteria.

narrative from ASRS report number 317930:
CAPT FLYING, AUTOPLT ON, AUTOTHROTTLES ON, DIGITAL FLT GUIDANCE SYS #1 USED FOR VERT AND LATERAL NAV, SPD 310 KTS AT FL240 PWR MGMNT SYS PROGRAMMED BUT NOT ENGAGED. ATC GAVE DSCNT CLRNC TO FL220. CAPT SELECTED 'PERF' ON THE DIGITAL FLT GUIDANCE SYS TO ENGAGE THE PWR MGMNT SYS. ON THE DSCNT PAGE OF THE PWR MGMNT SYS HE SELECTED VERT SPD. TO START 1000 FPM DSCNT THE PWR MGMNT SYS RECALCULATED THE OPTIMUM SPD TO BE 320 KTS. AS THE THROTTLES BEGAN TO ADVANCE HE TURNED THEM OFF TO PREVENT THE SPD INCREASE. HE THEN TRIED TO CHANGE THE SPD IN THE PWR MGMNT SYS TO 310 KTS, BUT IT WOULD NOT ACCEPT IT. DURING THIS TIME EITHER THE CAPT SELECTED OR AUTOPLT AUTOMATICALLY REVERTED TO IAS. I SAW ON THE FMA WE WERE IN IAS AND BECAUSE THE AUTOTHROTTLES WERE OFF BUT HAD BEEN ADVANCED WHEN PWR MGMNT SYS WAS SELECTED, WE WERE IN A CLB OF APPROX 1000 FPM AND AT APPROX 24500 FT. I CALLED OUT ALT AND CAPT INITIATED DSCNT. CONTRIBUTING FACTORS: NORMAL CRUISE SPD 320 KTS/.76 MACH CAPT USED 310 KTS. PWR MGMNT SYS SHOULD HAVE BEEN PROGRAMMED WITH DESIRED SPDS TO PREVENT USE OF OPTIMUM SPDS. PWR MGMNT SYS SHOULD HAVE BEEN SELECTED WHILE IN CRUISE OR AFTER DSCNT STARTED NOT IN THE MIDDLE OF A 'MODE' CHANGE. AUTOTHROTTLES TURNED OFF. ALT ALERTER ARMED FOR FL220.

narrative from ASRS report number 294000:
IN CRUISE AT FL350, ATC CLRED US TO CROSS JAXSN AT FL330. A FEW MINS LATER ATC RECLRED US TO CROSS 15 NM S OF JAXSN AT FL330. I WROTE DOWN THE CLRNC ALT AND DISTANCE. THEN SET 31000 FT IN THE MODE CTL PANEL (B757), BUILT THE FIX AND ENTERED FL330 IN THE FMS. AS WE APCHED THE ASSIGNED XING FIX I GLANCED AT THE 31000 FT I HAD ENTERED IN THE MODE CTL PANEL AND DETERMINED WE WOULD HAVE TO INCREASE OUR RATE OF DSCNT TO COMFORTABLY MAKE FL310. THE FMS SHOWED US LOW ON THE DSCNT PROFILE BUT I IGNORED IT AS THEY CAN OCCASIONALLY BE OFF. THE ACFT TRIED TO LEVEL AT FL330 IN VNAV BUT I WAS CONVINCED OUR ASSIGNED ALT WAS FL310. I SELECTED FLT LEVEL CHANGE ON THE MODE CTL PANEL TO CONTINUE DSCNT. AT APPROX FL320 ZDV TOLD US TO CLB TO FL330 AND TURN L APPROX 30 DEGS L OF COURSE. WE COMPLIED. THE ONLY REASON I CAN THINK OF FOR HAVING SELECTED DIFFERENT ALT FOR THE FMS AND MODE CTL PANEL IS A NOTE COMMONLY PLACED ON OUR ATL-DCA FLT PLANS ADVISING US ZDC ROUTING REQUIRES XING JAXSN AT OR BELOW FL310. WE ALSO WERE ON THE FINAL DAY OF A 4 DAY TRIP HAVING BEGUN THE DAY WITH AN XA30 WAKE UP.

There are 49 non-zero components of relevance in narrative 317930. The first 14 are shown below. What makes this narrative relevant are the prominent automation-oriented relations. For example, the prominent relations [SYS, PWR] and [SYS, MGMNT] are largely due to the many references to PWR MGMNT SYS in the narrative.

RCV(r,t)	RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
48208	184	R(0)	SYS	PWR	262
42336	189	R(1)	SYS	MGMNT	224
11610	45	R(2)	SYS	FLT	258
11472	48	R(3)	SELECTED	DSCNT	239
10591	89	R(4)	SYS	SPD	119
10320	86	R(5)	SYS	DSCNT	120
10285	85	R(6)	SYS	SELECTED	121
10106	31	R(7)	AUTOPLT	CAPT	326
8806	34	R(8)	SYS	NOT	259

8008	28	R(9)	SELECTED	NOT	286
6825	65	R(10)	SELECTED	PWR	105
6552	13	R(11)	MODE	ALT	504
5738	19	R(12)	SELECTED	SPD	302
5610	34	R(13)	SELECTED	CAPT	165

Option 3: Ranking by Multiple Sets of Criteria

It is possible to rank text items according to the intersection of multiple sets of relevance criteria. The first step is to separately rank the text items according to each of the sets of criteria. Then, the relevance ranking values (RRV) of each item are combined by multiplying them together.

Flight 800 example— In this example, the first set of criteria contains 241 relations on the topic of FBI + Kallstrom. These were described in the preceding section, "Ranking by topical focus." As a reminder, here are the top six relations:

probe term (PT)	term in context (TIC)	RMV(r,c)
FBI	crash	205
FBI	Kallstrom	159
FBI	James	146
FBI	TWA	135
FBI	investigating	126
Kallstrom	criminal	115

The second set of criteria contains 300 relations on the topic of the National Transportation Safety Board (NTSB). Here are the top 20 relations in this set of criteria:

probe term (PT)	term in context (TIC)	RMV(r,c)
National	Transportation	602
Safety	Transportation	600
National	Safety	589
Board	Safety	580
Board	Transportation	532
Board	National	522
NTSB	FAA	275
NTSB	safety	224
Safety	Flight	200
NTSB	recommendations	197
NTSB	investigators	196
Safety	Aviation	187
NTSB	tanks	181
NTSB	crash	142
NTSB	fuel	142
Safety	Foundation	140
Transportation	Department	131
NTSB	made	127
NTSB	agency	125
NTSB	investigation	124

All of the sentences in the 102 news stories about Flight 800 were ranked separately on the two sets of criteria. Then, the two relevance ranking values (RRV) of each sentence were combined by multiplying them together. Sentences were ranked according to the magnitude of the product. These sentences are the five which are most relevant to both topics, FBI + Kallstrom and National Transportation Safety Board (NTSB).

- The FBI and the National Transportation Safety Board are still investigating three theories : a missile , a bomb and mechanical failure .
- As a result , the possibility that a missile struck Flight 800 remains one of three theories being investigated by the FBI and the National Transportation Safety Board .
- James Kallstrom , who is heading the FBI criminal probe into the crash , said Saturday he agrees with the NTSB recommendations , but is critical of those who are " speculat publicly on what caused this horrific tragedy . "

- The transcripts of the FBI interviews were turned over to the NTSB last week .
- NTSB investigators were invited to conduct dual interviews with the FBI at the time , but the chose not to participate , " according to one criminal investigator who spoke on the condition of anonymity .

ASRS example— Text items from the ASRS database can also be ranked on multiple sets of relevance criteria. The relevance criteria used in this example are 936 automation relations and 982 training relations derived from 185 training-oriented narratives from the ASRS database.

Here are the 20 most prominent automation relations among the 936 relations used as the first set of relevance criteria:

probe term (PT)	term in context (TIC)	RMV(r,c)
AUTOPLT	ACFT	360
AUTOPLT	FT	345
AUTOPLT	ALT	312
AUTOPLT	MODE	307
MODE	ALT	288
AUTOPLT	NOT	276
SYS	ACFT	276
GLASS	COCKPIT	256
COMPUTER	FLT	230
AUTOPLT	CAPT	204
AUTOPLT	ENGAGED	201
MODE	NOT	190
AUTOPLT	FO	185
FMC	NOT	182
AUTOPLT	DISCONNECTED	177
COMPUTER	ACFT	170
MODE	HDG	170
FMC	PAGE	162
COMPUTER	MGMNT	158
FMC	DISPLAY	157

Here are the 20 most prominent training relations among the 982 relations used as the second set of relevance criteria:

probe term (PT)	term in context (TIC)	RMV(r,c)
SIMULATOR	TRAINING	340
TRAINING	ACFT	291
TRAINING	FO	266
TRAINING	NOT	247
TRAINED	NOT	188
TRAINING	RECEIVED	185
EXPERIENCE	ACFT	180
TRAINING	PLT	174
TYPE	RATING	172
TRAINING	APCH	160
LINE	TRAINING	157
EXPERIENCE	TRAINING	155
TRAINING	CAPT	151
RECURRENT	TRAINING	148
TRAINING	COMPANY	148
TRAINING	FLT	146
EXPERIENCE	FO	145
QUALIFIED	ACFT	126
TRAINING	RPTR	120
SIMULATOR	NOT	118

These automation and training relations were used as relevance criteria to separately rank all of the sentences in the narratives of 185 training-oriented ASRS reports. The two relevance ranking values for each sentence were then multiplied together to produce a combined rank. Shown below are the five most relevant sentences, based on the combined ranking. As expected, each sentence involves some connection between automation and training.

- WITH REGARD TO TRAINING RECEIVED ON THIS ACFT , THE RPTR STATED THAT THE SIMULATOR DID NOT HAVE THIS CHARACTERISTIC OR PROB SO WAS NOT TRAINED IN THE EXACT AUTOPLT DESIGN THAT IS IN THE ACFT . (rpt# 284362)
- MORE EMPHASIS ON HAND FLYING RATHER THAN AUTOMATED SYSTEMS DURING SIMULATOR TRAINING WOULD HELP THE LINE CREWS . (rpt# 66636)
- OUR ACR TRAINING INDICATES PF FMC DISPLAY UNIT BE ON PROGRESS PAGE AND PNF'S FMC DISPLAY UNIT BE ON LEGS PAGE , BECAUSE THE PROGRESS PAGE 1 HAS A TOP OF DSCNT ADVISOR DISPLAY . (rpt# 272507)
- WAS NOT AWARE OF MISTAKE IN WAYPOINT INSERTION IN FMC DUE TO LACK OF EXPERIENCE IN ACFT . (rpt# 71850)
- THIS PROB WAS DISCOVERED WHILE TRAINING IN AN ACR AIRLINES FLT SIMULATOR USING AN MD88 FMS . (rpt# 294429)

Combining separate rankings produces a logical intersection of multiple topics, in this case, automation AND training. If all of the relations had been combined as a single set of criteria, the result would have been a logical union, that is, automation OR training. Combining separate rankings ensures that text items meeting both relevance criteria are ranked highest.

Although the text items ranked in this example are sentences, narratives can also be ranked according to their relevance to multiple sets of criteria.

Option 4: Ranking by Externally Derived Criteria

As mentioned earlier, any set of relations can be applied to any set of text items. The model need not represent the text items being ranked. This could be useful for a variety of applications, including finding text in a collection B that is similar to that in a collection A.

Flight 800 example— It is possible, for example, to use the 280-relation QUORUM model of the 102 news stories on Flight 800 as relevance criteria for ranking the sentences in a collection of ASRS reports. The model of the news stories was described in the section, "Example of relevance ranking calculation." As a reminder, here are the top 10 relations:

probe term (PT)	term in context (TIC)	RMV(r,c)
Flight	800	1725
TWA	Flight	1486
TWA	800	1461
fuel	tank	1115
New	York	990
fuel	center	894
United	States	865
fuel	tanks	849
bomb	missile	752
Long	Island	720

This exercise is only meaningful if there is come overlap of content, so the criteria were applied to 325 reports which include incidents involving fuel. Since ASRS narratives are abbreviated and capitalized, the relevance criteria were also abbreviated and capitalized.

Here are the 5 sentences that are most relevant to the QUORUM model of the Flight 800 news stories.

- WHILE THE REFUELERS WERE TRANSFERRING FUEL OUT OF THE CTR AUX TANK , THEY ACCIDENTALLY ALSO REMOVED FUEL FROM THE #1 TANK . (rpt# 242855)
- FUEL ON FINAL 1350 LBS L TANK , 950 LBS R TANK , 7000 LBS CTR TANK . (rpt# 301328)
- IN CRUISE FL350 , CAPT NOTICED FUEL IN MAIN TANKS DECREASED TO 1400 / 1100 LBS AND AFTER CHKING FUEL PANEL , NOTICED CTR TANK PUMP SWITCHES IN MID POS INSTEAD OF ON , ACFT CONFIGN WITH AUX TANKS AND 3 POS CTR FUEL TANK SWITCHES WITH UPPER POS PLACARDED DEACTIVATED . (rpt# 301328)
- I MISREAD THE MEL AND DISPATCHED THE FLT WHICH REQUIRED USE OF THE FUEL IN THE CTR TANK . (rpt# 288905)
- THE MEL STATED THAT FUEL IN CTR AND AUX TANKS CONSIDERED UNUSABLE . (rpt# 288905)

This table contains the components of relevance of the first sentence, "While the refuelers... ."

RCV(r,t)	RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
2133	44	R(0)	FUEL	TANK	1115
1011	26	R(1)	FUEL	CTR	894
733	24	R(2)	TANK	CTR	702

Shown below are the fuel-related relations from the 280-relation model of the Flight 800 news stories:

probe term (PT)	term in context (TIC)	RMV(r,c)
FUEL	TANK	1115
FUEL	CTR	894
FUEL	TANKS	849
TANK	CTR	702
FUEL	EXPLOSION	333
EXPLOSION	TANK	314
FUEL	AIR	252
FUEL	PUMP	234
FAA	TANKS	223
FUEL	ANY	215
FUEL	VAPORS	211
TANKS	HEAT	197
FUEL	FAA	193
AIR	TANKS	191
FUEL	HEAT	190
TANKS	REQUIRE	186
NTSB	TANKS	181
FUEL	EXPLOSIVE	175
TANKS	UNDERGROUND	174
FUEL	FEDERAL	171
TANKS	PREVENT	170
FUEL	IGNITED	168
FUEL	STATIC	167
TANKS	COOLER	159

Even this small collection of relations, gleaned from a non-technical source, could be useful for retrieving and relevance-ranking ASRS or other incident reports. If a more technical and comprehensive model of Flight 800 concerns were applied, it would be possible to retrieve and rank incident reports which are even more relevant to that disaster.

This example suggests the potential benefit of using the relations of a QUORUM model of one collection as relevance criteria in another collection.

ASRS example— Narratives from the ASRS database can also be ranked on externally derived criteria. The relevance criteria used in this example are the 256 relations of an automation-oriented model of 313 automation-error narratives from the ASRS database. This model was described in an earlier ASRS example in the section "Ranking by Topical Focus." As a reminder, here are the top ten relations:

probe term (PT)	term in context (TIC)	RMV(r,c)
PANEL	CTL	589
MODE	ALT	504
ILS	RWY	480
AUTOPLT	ALT	478
MODE	CTL	472
AUTOPLT	DISCONNECTED	448
MODE	SELECTED	409
FMS	DSCNT	404
MODE	SPD	400
CAPTURE	ALT	399

The collection of 300 narratives to be ranked are those analyzed and modeled in McGreevy (1996). Each of the narratives contain the word "mode."

The mode collection was successfully ranked on the topic of cockpit automation, even though the relevance criteria came from a different collection of narratives. The three most relevant narratives, based on relevance density, are shown below. This supports the notion that QUORUM relevance ranking criteria are reusable. (Note: The criterion collection and ranked collection overlapped by 18 reports. Report 218897, containing the third narrative shown below, was one of the eighteen. Only 4 others are among the top 50 most relevant reports, so overlap had little effect.)

narrative from ASRS report number 204756:
AUTOPLT ON IN 'PERF' MODE, CRUISE CONDITIONS. ACFT STARTED A SLIGHT DSCNT TO ABOUT 300 FT BELOW ASSIGNED ALT, WHEREUPON CAPT SELECTED 'VERT SPD' MODE AND A 500 FPM CLB. BUT ACFT STARTED TO CLB AT 2000 FPM AND WENT RIGHT THROUGH SELECTED ALT OF FL350 TO ABOUT 450 FT HIGH, WHEREUPON CAPT DISCONNECTED AUTOPLT AND RETURNED TO FL350. NO CONFLICT. I'M STILL NOT SURE IF THIS WAS DUE TO MOUNTAIN WAVE ACTIVITY OR AUTOPLT MALFUNCTION OR BOTH. CAPT ASSUMED MOUNTAIN WAVE AND INSTRUCTED ME TO RPT IT TO CTR. THIS PARTICULAR AUTOPLT, WHEN USED IN THE 'PERF CRZ' MODE (WHICH IS SOP) CONSISTENTLY DEVIATES FROM SELECTED ALT BY + OR - 100 TO 200 FT. THIS MAKES IT AT TIMES DIFFICULT TO DETERMINE IF AUTOPLT IS FUNCTIONING 'NORMALLY' OR MALFUNCTIONING UNTIL IT IS TOO LATE. STILL, IF WE HAD BEEN MORE AGGRESSIVE IN DISCONNECTING AUTOPLT SOONER AND FLYING PROPER ALT, WE MIGHT HAVE DIMINISHED THE ALT EXCURSION.

narrative from ASRS report number 252165
WE WERE GIVEN A DSCNT FROM 310000 FT BY CTR. THE ACFT WAS ON AUTOPLT WITH LNAV AND VNAV ENGAGED, USING THE FMC AND AREA NAV. I WAS THE PF. THE FO SET MODE CTL PANEL ALT TO 28000 FT WITH THE AUTOPLT ENGAGED. THE FMC DID NOT ACCEPT 28000 FT INTO THE PROGRAM AND IT TOOK 3 ENTRIES TO ACTIVATE IT. AFTER ENTERING THE ALT IN THE FMC, I LOOKED AT THE INSTS AND THOUGHT THAT THE ALTIMETER HAD FAILED BECAUSE IT WAS SHOWING A CLB THROUGH 31600 FT. SHORTLY THEREAFTER, CTR CALLED FOR OUR ALT AS I WAS TAKING THE ACFT OFF AUTOPLT AND CORRECTING THE CLB. WE WERE CLOSE TO 31900 FT BEFORE WE COULD LEVEL AND START DOWN MANUALLY. THE FO WAS INVOLVED IN PAPERWORK AND WAS CAUGHT BY SURPRISE ALSO. IT APPEARED THAT WHEN THE FMC WOULD NOT ACCEPT 28000 FT THAT THE VNAV LOST THE ALT INPUT. THIS PROBABLY CAUSED THE AUTOPLT TO TRIP FROM COMMAND TO CTL WHEEL STEERING PITCH, WHICH WAS THE INDICATION WHEN I TOOK OVER MANUALLY. WHY THE AUTOPLT WENT INTO A CLB WHEN TRIPPED TO CTL WHEEL STEERING PITCH IS A MYSTERY. NEITHER THE FO NOR MYSELF HAD FELT THE ACFT GO INTO A CLB. OUR NORMAL ALT WARNING DID NOT GIVE ANY SIGNAL IN THIS CASE BECAUSE THE MODE CTL PANEL HAD BEEN SET TO 28000 AND WE HAD ENTERED A CLB OUT OF 31000 INSTEAD OF A DSCNT. THE AUTOPLT DID NOT GIVE AN AURAL WARNING BECAUSE IT DID NOT TRIP OFF COMPLETELY, BUT ONLY SWITCHED TO CTL WHEEL STEERING IN PITCH MODE. ON AUTOFLT ACFT, ANY PROB WITH PROGRAMMING THE FMC CAN DISTRACT THE PLTS' ATTN FROM THE FLT INSTS. USUALLY, NORMAL INST SCAN OR ANY ONE OF THE WARNING DEVICES WOULD HAVE BROUGHT MY ATTN TO THE ERROR IN THE FLT CTL BEFORE ALT COULD CHANGE BY 600 FT. SUPPLEMENTAL INFO FROM ACN 252364: THE CAPT (THE PF) HAD THE AUTOPLT ENGAGED, IN THE 'CTL WHEEL STEERING' MODE. WE RECEIVED AND ACKNOWLEDGED A DSCNT CLRNC TO FL280. WE WERE ALSO ASKED TO KEEP OUR SPD UP. THE CAPT SELECTED A HIGHER SPD IN THE AUTOPLT MODE CTL PANEL, THEN PROCEEDED TO LEAN DOWN OVER THE COMPUTER TO SET IN THE LOWER ALT. MEANWHILE, WITH THE FASTER SPD DIALED IN, THE AUTOTHROTTLES ADVANCED, WHICH MUST HAVE PITCHED THE NOSE OF THE AIRPLANE UP AND CAUSED IT TO CLB. THE 'CTL WHEEL STEERING' MODE OF THE AUTOPLT ONLY HOLDS WHATEVER FLT ATTITUDE THE ACFT IS PRESENTLY HOLDING. NEVER USE 'CTL WHEEL STEERING' MODE OF THE AUTOPLT UNDER NORMAL LINE OPS.

narrative from ASRS report number 218897
AT ATC REQUEST, DOING MACH .82 OR BETTER DSCNT FOR SPACING INTO JFK. FMC PROGRAMMED FOR .82 DSCNT, BE LEVEL 10 NM W STW FL230, THEN CROSS LINDY FL190 AT 250 KTS. FULL VNAV DSCNT. ACFT MADE 10 NM W STW AT FL230, BUT WENT INTO ALT HOLD AND SPD MODE. INSERTED FL230 INTO FMC (CRUISE PAGE) AND 300 KTS DSCNT SPD AS ACFT WAS AT ABOUT 325 KTS. THE FMC MACH/AIRSPD CHANGEOVER DID NOT OCCUR. WHEN FL230 INSERTED INTO CRUISE PAGE AND VNAV SELECTED, ACFT STARTED TO CLB. AT FL233, I DISCONNECTED AUTOPLT AND EASED NOSE BACK DOWN FOR DSCNT TO FL230. WE WERE ABOVE FL233 ABOUT 15 SECONDS, REACHING ABOUT 23450 FT. ATC DID NOT QUESTION OR COMMENT ON THE ALTDEV. ONCE BACK AT FL230, THE COMPUTER WAS CHKED THAT ALL ENTRIES WERE CORRECT. NO CORRECTIONS WERE NEEDED. AUTOPLT, LNAV, AND VNAV RE-ENGAGED. LINDY SPD/ALT XING MADE WITH NO MANUAL INTERVENTION.

Shown below are the many automation-oriented components of relevance of narrative 204756.

RCV(r,t)	RMV(r,t)	R(r)	PT	TIC	RMV(r,c)
18164	38	R(0)	AUTOPLT	ALT	478
17108	52	R(1)	SELECTED	ALT	329

14616	29	R(2)	MODE	ALT	504
12679	31	R(3)	MODE	SELECTED	409
12062	37	R(4)	AUTOPLT	CAPT	326
8064	18	R(5)	AUTOPLT	DISCONNECTED	448
7200	18	R(6)	MODE	SPD	400
5994	27	R(7)	AUTOPLT	MODE	222
5134	17	R(8)	SELECTED	SPD	302
4930	34	R(9)	SELECTED	CLB	145
4896	18	R(10)	AUTOPLT	NOT	272
4824	18	R(11)	MODE	VERT	268
4455	27	R(12)	SELECTED	CAPT	165
2628	18	R(13)	DISCONNECTED	CAPT	146
2628	18	R(14)	MODE	DSCNT	146
2622	23	R(15)	MODE	CLB	114
2151	9	R(16)	SELECTED	DSCNT	239
2057	17	R(17)	SELECTED	BUT	121
2052	18	R(18)	MODE	CAPT	114
2044	14	R(19)	AUTOPLT	SELECTED	146
1962	18	R(20)	SELECTED	VERT	109
1808	8	R(21)	AUTOPLT	DSCNT	226
1008	9	R(22)	DISCONNECTED	ALT	112

Ranking Sentences within each Narrative

In addition to relevance ranking all of the sentences in a collection of narratives, or all of the narratives in a collection of narratives, it can also be useful to rank the sentences within each narrative. This idea is introduced here because it will simplify the illustrations in the section, "Ranking by Example," which follows this one.

By displaying all sentences in a narrative in the order they appear, but with their relevance ranking values in the left column, it is possible to quickly see the most relevant sentences in the full context of the rest of the narrative. Two examples of this are shown in this section.

The relevance criteria used in these examples represent the topic of automation. They are the 256 relations of an automation-oriented model of 313 automation-error narratives from the ASRS database. This model was described in an earlier ASRS example in the section "Ranking by Topical Focus." As a reminder, here are the top ten relations:

probe term (PT)	term in context (TIC)	RMV(r,c)
PANEL	CTL	589
MODE	ALT	504
ILS	RWY	480
AUTOPLT	ALT	478
MODE	CTL	472
AUTOPLT	DISCONNECTED	448
MODE	SELECTED	409
FMS	DSCNT	404
MODE	SPD	400
CAPTURE	ALT	399

Shown below are the sentences of the narrative of ASRS report number 264689. They are shown in the order they appear in the narrative. Only two of the sentences are relevant to the automation concerns in the model. The relevance ranking value (RRV) of each sentence appears in the left column. The components of relevance are shown below each relevant sentence.

RRV	sentence
0	WHILE FLYING FLT FROM MSP TO SAN WE WERE GETTING NUMEROUS ALT CHANGES AND TA'S .
0	THE CAPT WAS FLYING .
0	WE WERE CLRED TO 12000 FT .

121 I SAW HE HAD SELECTED 11900 FT BUT WAS REACHING UP TO CORRECT IT .

RCV(r,t)	RMV(r,t)	PT	TIC	RMV(r,c)
121	15	SELECTED	BUT	121

0 I PROCEEDED TO GET THE NEW ATIS .

712 THE CAPT SET 12000 FT IN THE ALT WINDOW , BUT ON THE A320 , SETTING A NEW ALT WHEN WITHIN 300 FT OF THE OLD ALT PUTS YOU IN A VERT SPD MODE AND YOU WILL MISS YOUR ALT .

RCV(r,t)	RMV(r,t)	PT	TIC	RMV(r,c)
319	26	MODE	ALT	504
165	17	MODE	SPD	400
124	32	SETTING	ALT	159
104	16	MODE	VERT	268

0 WE CAUGHT IT AND CORRECTED AT 11600 FT - 400 FT BELOW ASSIGNED .

0 THE CTLR DID NOT INDICATE A CONFLICT OR ANY CONCERN .

The most relevant sentences represent a "topical summary by selection" of the whole narrative. That is, the narrative is summarized with respect to the topic. In the examples shown here, the relevance criteria focus on the topic of automation, so the most relevant sentences are those which focus on the topic of automation. If the relevance criteria were based only on whatever topics happened to be contained in each narrative, then the most relevant sentences would be an "abstract by selection."

Shown below are the sentences of the narrative of ASRS report number 156875. They are shown in the order they appear in the narrative. Only two of the sentences are relevant to the automation concerns in the model. The relevance ranking value (RRV) of each sentence appears in the left column. The components of relevance are shown below each relevant sentence.

RRV	sentence
0	I WAS HAND FLYING A WDB STRETCH OUT OF LGA .
0	WE WERE BEING GIVEN NUMEROUS VECTORS AND STEP CLBING RESTRICTIONS BY NY DEP CTL .
0	WE WERE LEVEL AT 15000 FT .
0	WERE GIVEN A HDG CHANGE AND INSTRUCTED TO CLB TO 17000 FT .
2457	CAPT SET WRONG ALT IN MODE CTL PANEL AND WE EXCEEDED THE 17000 FT RESTRICTION .

RCV(r,t)	RMV(r,t)	PT	TIC	RMV(r,c)
625	17	PANEL	CTL	589
504	16	MODE	ALT	504
501	17	MODE	CTL	472
369	16	MODE	PANEL	369
177	14	PANEL	ALT	203
96	14	MODE	SET	110
93	12	PANEL	SET	124
92	13	MODE	CAPT	114

0 DEP CALLED US AS WE CLBED THROUGH 17500 FT AND TOLD US TO DSND BACK TO 17000 FT WHICH WE DID .

0 SHORTLY AFTER THAT WE WERE SWITCHED TO NY CTR WITHOUT ANY FURTHER COMMENT .

758 THE PROCS WE USE WHEN FLYING THE ADVANCED COCKPIT AIRPLANES PUTS SO MUCH ATTN ON THE MODE CTL PANEL AND THE FD THAT WE GET A LITTLE LAX IN KEEPING THE RAW DATA IN OUR SCAN .

RCV(r,t)	RMV(r,t)	PT	TIC	RMV(r,c)
270	17	PANEL	CTL	589
216	17	MODE	CTL	472
159	16	MODE	PANEL	369
113	17	DATA	RAW	246

```
0    HAD I BEEN FLYING BY THE ALTIMETER INSTEAD OF THE FD THIS WOULD NOT HAVE
     HAPPENED .
```

A minor problem with non-standard usage appears in this example. Note that the last two sentences each contain FD (i.e., flight director), which refers to automation. Among the 313 reports, there is only one other occurrence of FD. All other references to the flight director are spelled out as FLT DIRECTOR. While the relation [DIRECTOR, FLT] is, in fact, included among the relevance criteria, the non-standard usage, FD, is unrecognized. If the sentence, "Had I been flying... ," contained FLT DIRECTOR instead of FD, the relevance ranking value (RRV) of this sentence would have been 382 * 17 / 17 = 382, instead of zero. Similarly, the second to last sentence would also have a higher relevance ranking value. Related concerns about non-standard usage are discussed in more detail in the next section. In general, QUORUM focuses on common usage and ignores words used only a few times in a collection.

Option 5: Ranking by Example

If the analyst has in hand one or more interesting narratives, it is possible to use relevance ranking to find other, similar narratives. As a first step, a QUORUM model is derived from the example narratives. This model is then used to rank a collection of reports. The reports with the highest ranking will be most similar to the examples. This is known as ranking by example. It allows analysts to find more reports like the ones of interest.

When ranking by example, the larger the number of examples, the easier it is for QUORUM to concentrate on the commonalities and ignore the differences. If there are only a few example narratives, it is important that non-standard vocabulary in the examples is changed to standard usage. For example, VERT SPD is far more commonly used than VS, so any occurrences of VS in the examples should be changed to VERT SPD. Similarly, the commonly used DSCNT replaces DSNT in the example narratives, and DEV replaces DEVIATION. If there are two widely used forms (usually with one being more common, however), both forms can be included. For example, MODE CTL PANEL is more common than MCP, but both are widely used. To deal with this, any instances of either form in the example narratives are replaced with "MODE CTL PANEL (MCP)."

If the collection of example narratives is small, the relevance criteria derived from them should be edited. This is done to ensure that only the relations of interest to the analyst are used in the ranking. For instance, the fact that an example incident occurred near LAX (Los Angeles) might not be of interest, depending on the goals of the analyst. If LAX is not of interest, all relations pertaining to LAX should be deleted from the relevance criteria.

The following is an illustration of ranking by example. In a project for the ASRS, a set of 313 automation-error narratives from the ASRS database were ranked by example. The examples were selected from among the 313 narratives to be ranked, but that is not necessary for the method to work. The set of examples consisted of two ASRS reports numbered 139884 and 163566. These reports contain 375 and 472 words, respectively. The analysts said that these reports were representative of human-automation incidents.

A QUORUM model was derived from these two narratives. Since there were only two examples, non-standard usages were changed to standard ones, and the relevance criteria were edited to delete those involving specific geographic locations, specific altitudes, and units of measure. Selections and deletions are discussed in more detail in the section, "Selecting Relations for Use as Relevance Criteria."

Here are the top 14 of 134 relations derived from the two narratives.

probe term (PT)	term in context (TIC)	RMV(r,c)
MODE	SPD	355
MODE	SELECTED	217
SPD	SELECTED	173
ALT	RESTRICTIONS	135
ALT	DSCNT	134
SPD	VERT	133
SPD	DSCNT	131
FMC	PROGRAMMED	123
DSCNT	FMC	122
MODE	FMC	121
MODE	RESELECTED	118
ALT	SELECTOR	112

MODE	DSCNT	111
DSCNT	PROFILE	109

To indicate the contents of the example narratives (139884 and 163566), here are the three most relevant sentences from each of them, in order of relevance. In fact, these are also the most relevant sentences in the whole collection of 313 narratives. (The method of ranking the relevance of sentences within each narrative was presented in the preceding section, "Ranking Sentences within each Narrative.") These sentences suggest that use of cockpit automation to control altitude changes, especially descents, is a prominent concern in the example narratives.

RRV	line#	index	sentence
2163	_6_	139884	I RESELECTED THE VERT SPD MODE .
2146	_29_	163566	THEN ATC REQUESTED AN EXPEDITED DSCNT THROUGH FL200 AND I SELECTED SPD DSCNT MODE ON THE DSCNT PAGE OF THE FMC AND A SPD OF 250 KTS , WHICH NO LONGER AFFORDS ALT PROTECTION FOR RESTRICTIONS ON THE PROFILE DSCNT .
2139	_5_	139884	THE NEXT TIME I LOOKED UP THE MODE CTL PANEL (MCP) WAS OPERATING WITH THE SPD MODE SELECTED , WHICH CONFUSED ME BECAUSE I HAD NOT SELECTED THAT MODE .
1314	_33_	163566	I FAILED TO REALIZE THAT THE ALT RESTRICTIONS ARE NOT IN EFFECT DURING A SPD MODE DSCNT .
1313	_42_	163566	AFTER A QUICK DISCUSSION WITH THE CAPT WE REALIZED THAT IN MY ABSENCE HE HAD SELECTED THE SPD MODE INSTEAD OF THE PATH MODE ON THE FMC .
1243	_10_	139884	WHEN THE CAPT SELECTED SPD HE HAD ALSO SET 10000 FT IN THE MODE CTL PANEL (MCP) , NOT UNDERSTANDING THAT THE FMC WOULD NOT CAPTURE AT 14000 FT .

Shown below are the report numbers (in the last column) of the 15 narratives that are most relevant to the concerns in reports 139884 and 163566. Notice that the two example reports show up at the top of the list because they are most relevant to the relevance criteria. This is exactly what one expects, given that the relevance criteria were derived from these reports.

RRV	RRV'	T(t)	rpt#
1178037	220882	375	139884
1143254	269808	472	163566
931496	72191	155	218897
572595	67280	235	317930
568156	72724	256	304278
443754	23519	106	264689
401973	45021	224	294000
395125	55120	279	302317
392157	14902	76	306764
383367	44279	231	317620
368716	22123	120	303544
295727	51013	345	261452
283609	17442	123	297905
272412	17162	126	318230
269145	35258	262	218329

Shown below are the most relevant sentences from each of the top 10 reports (apart from the two example reports). Recall that the whole narratives are ranked, not just these sentences. The topics contained in these sentences suggest that the narratives from which these sentences came are indeed relevant to the topics in the example reports (139884 and 163566). These sentences suggest that use of cockpit automation to control altitude changes, especially descents, is a prominent concern in these narratives, just as in the example narratives. Thus, these reports are indeed similar to the examples.

RRV	line#	index	sentence
761	_5_	218897	INSERTED FL230 INTO FMC (CRUISE PAGE) AND 300 KTS DSCNT SPD AS ACFT WAS AT ABOUT 325 KTS .
904	_18_	317930	ON THE DSCNT PAGE OF THE PWR MGMNT SYS HE SELECTED VERT SPD .

2590	_34_	304278	AT APPROX FL320 THE FO SELECTED THE VERT SPD MODE ON THE MODE CTL PANEL AND SELECTED A HIGHER SPD , THUS SLOWING THE CLB RATE .
504	_50_	264689	THE CAPT SET 12000 FT IN THE ALT WINDOW , BUT ON THE A320 , SETTING A NEW ALT WHEN WITHIN 300 FT OF THE OLD ALT PUTS YOU IN A VERT SPD MODE AND YOU WILL MISS YOUR ALT .
884	_60_	294000	I SELECTED FLT LEVEL CHANGE ON THE MODE CTL PANEL TO CONTINUE DSCNT .
1096	_75_	302317	I DID NOT SEE THE FO CHANGE THE MODE CTL PANEL OR INITIATE THE VERT SPD CLB .
245	_85_	306764	THE FMC WAS PROGRAMMED FOR THE XING RESTR BUT WOULD NOT ACCEPT IT .
1170	_95_	317620	THE FO INITIATED THE DSCNT BY SELECTING A VERT SPD IN PROFILE MODE .
664	_115_	303544	IF NOT , MANUAL ACFT CTL , OR USE MODE CTL PANEL TO ACHIEVE DESIRED RESULTS .
734	_140_	261452	ONE ALLOWS YOU TO CONTINUE TO ROTATE THE VERT SPD WHEEL IN THE ALT CAPTURE MODE .

Option 6: Ranking by "Outsider" Criteria

In a sense, analysts are "outsiders" while incident reporters are "insiders." To select and rank reports based on "outsider" criteria, it is necessary to map these criteria to the language of the "insiders."

In their own words— To understand the concerns of incident reporters, it is important to take special note of the fact that reporters describe incidents in their own words. These words do not necessarily translate directly into the concerns of incident analysts. The people who write and submit commercial aviation incident reports to the ASRS include cockpit crews, air traffic controllers, cabin crews, and ground crews. These reporters share a common vocabulary, the jargon of day-to-day commercial aviation operations. Even within this commonality, however, different groups of reporters tend to use somewhat different vocabularies because their roles, particular equipment, and immediate environments differ. The words and concepts found in narratives written by pilots, for example, tend to differ from those found in narratives written by controllers.

The people who seek to understand commercial aviation incident reports include airline managers, union representatives, federal regulators, human factors researchers, and others. Analysts in each of these groups have their own sets of concerns and their own professional vocabularies. The words and concepts used by these analysts are often different from those used by the incident reporters themselves. Human factors researchers, for example, might be concerned with "decision making," "crew pressures," or "mode confusion" but these concepts, and other such theory-oriented ideas, are not explicitly described in incident narratives.

Another issue in vocabulary development is created by the way text is entered into the database. The ASRS, for example, capitalizes all words in narratives, and abbreviates many of them. Typical abbreviations include ACFT for aircraft, TFC for traffic, CAPT for captain (but occasionally for capture), FO or F/O for first officer, AUTOPLT for autopilot, and FMC for flight management computer. It is usually the responsibility of the ASRS database expert who retrieves reports to translate an analyst's queries into the vocabulary used in the database.

To effectively use QUORUM models for relevance ranking, it is necessary to develop an understanding of the lexicon, the specialized vocabulary, of the domain and the database being analyzed. One must also understand the specialized vocabularies and concepts of those who analyze narratives. This understanding can be captured in QUORUM models by collecting narrative-based relations that represent analysts' concerns. As illustrated in the following example, QUORUM provides a mechanism for mapping between the words and concepts of the analysts and those of the incident reporters.

ASRS example— In one recent project, an analyst was interested in "crew pressure." A search was conducted for ASRS incident reports containing such words as PRESSURE, TIME, and SCHEDULE. The search returned 325 reports. Analysis of the reports revealed, however, that most of these reports involved references to mechanical pressure rather than "crew pressure." For example, among the top 300 relations in the 325 narratives there were 11 relations involving PRESSURE, shown in the table below. These relations indicate that in the 325 reports, the most prominent words in the

context of PRESSURE are OIL, ENG (i.e., engine), LOW, LIGHT, TIME, #2, ACFT, FUEL, SCHEDULE, CABIN, and NOT.

relation N	probe term	term in context	RMV
21	PRESSURE	OIL	881
37	ENG	PRESSURE	698
57	PRESSURE	LOW	605
85	PRESSURE	LIGHT	526
116	TIME	PRESSURE	469
124	PRESSURE	#2	453
164	ACFT	PRESSURE	400
184	FUEL	PRESSURE	377
224	PRESSURE	SCHEDULE	343
271	PRESSURE	CABIN	311
283	NOT	PRESSURE	306

These relations were used to relevance-rank the sentences contained in the narratives of the 325 ASRS reports. Shown below are the five most relevant sentences. The number following each sentence is the ASRS report number. Clearly, these sentences have little to do with "crew pressure."

- AT LL10Z , THE #2 ENG GEAR BOX OIL PRESSURE FLUCTUATED AND ENG LOW PRESSURE OIL LIGHT ILLUMINATED . (rpt# 211276)
- SHORTLY THEREAFTER THE #2 CSD OIL PRESSURE LOW LIGHT AND #2 HYD PRESSURE LOW LIGHT ILLUMINATED . (rpt# 186702)
- CLBING THROUGH FL180 #2 ENG LOW OIL PRESSURE ANNUNCIATOR ILLUMINATED . (rpt# 248466)
- #2 OIL PRESSURE GAUGE WAS INDICATING 0 OIL QUANTITY . (rpt# 248466)
- ENG OIL PRESSURE FINALLY DROPPED BELOW NORMAL . (rpt# 266668)

To focus on "crew pressure" reports, the complete QUORUM model of the 325 reports, consisting of thousands of relations, was edited to retain only prominent relations likely to involve "crew pressure." Accordingly, relations such as [PRESSURE,OIL], [ENG,PRESSURE], and [PRESSURE,LOW] were deleted, while relations such as [TIME,PRESSURE], [PRESSURE,SCHEDULE], and [PRESSURE,FELT] were retained. A total of 300 "crew pressure" relations are contained in this focused model. Here is a sample of the relations in the "crew pressure" model:

probe term (PT)	term in context (TIC)	RMV(r,c)
PRESSURE	TIME	406
PRESSURE	SCHEDULE	366
LATE	FLT	158
PRESSURE	FELT	151
...		
ERROR	FUEL	93
LATE	DEP	91
DELAYED	FLT	90
LATE	TIME	89
PRESSURE	UNDER	89
...		
COMPANY	INVESTIGATION	36
COMPANY	MORALE	36

The 300 "crew pressure" relations were used to relevance-rank the sentences contained in the narratives of the collection of 325 reports. Shown below are the five most typical sentences, which prominently contain many of the "crew pressure" relations. The number following each sentence is the ASRS report number. These sentences are indeed focused on "crew pressure." In the project, the "crew pressure" relations were also used to relevance rank the narratives.

- CAPT FELT SCHEDULE PRESSURE AND FELT RUSHED DURING SHORT TAXI. (rpt# 108496)
- HE FELT THAT TIME WAS A FACTOR IN A SCHEDULE PRESSURE SIT. (rpt# 242855)
- FLC THEN DEPARTS LATE AND HAS SCHEDULE PRESSURE TO MAKE UP THE TIME. (rpt# 308450)

- FIRST, WE WERE RUNNING BEHIND SCHEDULE AND WERE THEREFORE EXPERIENCING TIME PRESSURE. (rpt# 110644)
- WE WERE UNDER A PRESSURE OF SHORT TIME TO MAKE SCHEDULE (APPROX 18 MINS). (rpt# 85293)

Once the focused set of "crew pressure" relations is available, it can be applied to any collection of ASRS reports in order to rank the text on this particular collection of "crew pressure" concerns. Further, if the database were structured to support the QUORUM method, it would be possible to retrieve reports based on this or any other focused model. In addition, it is possible to further fine-tune the focused set of relations by adding or deleting relations. After repeated analysis of "crew pressure" reports, one or more standardized models of "crew pressure" could be developed. Such models could provide standardized retrieval and ranking criteria for use by others.

Some analysts might be interested in any reference, no matter how rare, to certain "hot-button" words appearing in incident reports. Words like FIRED and SUSPENDED, for example, appear among the "crew pressure" reports, but FIRED occurs only twice and SUSPENDED occurs only once. As a result, these words do not appear among the top 300 "crew pressure" relations. Even so, such words indicate significant crew pressure: in the context of FIRED, the word PLT (i.e., pilot) is the most closely related word, and UNION is the second most related word. To look for patterns among reports containing such rare terms, the analyst should select reports from the database which contain even one of their hot-button words (e.g., FIRED, SUSPENDED, TERMINATION, ILLEGAL, UNSAFE, MORALE). The QUORUM method can then be used to see if there is a pattern among the reports.

A Closer Look at QUORUM Relations

The QUORUM method of text analysis, modeling, and relevance ranking is based on proximity-weighted co-occurrence relations between words in the text. Derivation of these relations, and their use in modeling, is described in detail elsewhere (McGreevy, 1996; McGreevy, 1995). For relevance ranking, it is necessary to select and delete QUORUM relations in order to develop and refine sets of relevance criteria. It is important to understand the basis of these selections and deletions.

Once the method is more mature, it is likely that some analysts will be able to utilize standard sets of criteria for such topics as training, automation, crew pressure, and the like. Other analysts, however, will want to be able to precisely shape and refine the relevance criteria they use for relevance ranking. Accordingly, it is necessary to understand the kinds of relations encountered and how to use them.

In the section immediately following, these kinds of QUORUM relations are surveyed:

- Relations involving only rarely occurring words
- Not-too-distant relations
- Relations with "stop words"
- Reciprocal or reflexive relations
- Relations with pronouns
- Relations with units of measure
- Domain-generic relations
- Situation-generic relations
- Location-specific relations
- Infrastructure relations
- Object relations
- Off-topic relations

Some QUORUM relations are closely associated with prominent word groups in the text. Because of this, word group analysis can help the analyst to interpret some of the prominent relations. This is briefly discussed in the section, "Word Groups," following the discussion of the various kinds of relations.

Finally, QUORUM relations can be used to improve the selection of text for analysis. This is discussed in the section, "Relevance Criteria versus Selection Criteria."

Selecting Relations for Use as Relevance Criteria

In order to use the QUORUM method effectively, one must understand how to appropriately select and delete relations, the contextual associations of word pairs. Relevance ranking is particularly sensitive to proper selection and deletion of relations. To make sense of this task, it can be useful to categorize relations.

Here is a list of some of the important kinds of relations, with a brief discussion of the nature and uses of each type, which ones should be selected or deleted, and the reasons for doing so.

Relations involving only rarely occurring words— There are more relations in this category than in any other. In order to find the essence of a text, and to drastically reduce the potential complexity of the model, the most important step is to eliminate relations involving only rarely occurring words. This is done by limiting relations to those involving at least one of the most frequently occurring words in the text. This does not preclude relations in which one of the words occurs infrequently. Relations involving only infrequently occurring words provide many interesting details, but according to the Simon approach to complexity management, these details are not essential for a "tolerable description of reality."

Not-too-distant relations— The words in the neighborhood of an instance of a probe term are considered to be terms-in-context. Words farther from that instance have a weaker claim on that designation until, at some point, it becomes meaningless. After conducting sensitivity analyses, it appears that a distance of one average sentence length is an appropriate, though somewhat arbitrary, cut-off point (McGreevy, 1995). So, for example, if the average sentence length is 20 words, words beyond that distance from the probe terms are considered to be too distant to be in the same context. This cut-off is used in order to achieve computing efficiencies. It appears likely that having no cut-off at all would yield comparable —if more precise and costly— results.

Relations with "stop words"— So-called "stop words" are those which do not refer to things, concepts, actions, attributes, or attribute values. Words such as "the," "that," "and," and the like are stop words. QUORUM models which include stop words have the potential to be valuable for grammatical analysis or subtle domain analysis. For example, words often found in the context of the word "the" typically represent things of importance. Words in the context of prepositions can yield information about spatial relations in a domain. For most domain models, however, there is greater interest in first discovering the prominent things, concepts, actions, attributes, and attribute values, and how they relate to one another. For this reason, relations involving stop words are usually deleted from QUORUM models.

Reciprocal or reflexive relations— In the current QUORUM method, the relation between a word X and a word Y is the same as that between Y and X. For that reason, if X and Y are both probe terms, the reciprocal relations [X,Y] and [Y,X] will both be found. If so, only one of them is retained, as the relational metric values will be equal. The most frequently occurring word gets the privilege of appearing first, so that if word X is more common than word Y, the form [X,Y] is retained and [Y,X] is deleted. (See McGreevy (1995) for a discussion of using asymmetric and symmetric relations.) It is also common for two instances of the same word X to be found in close proximity, resulting in the reflexive relation [X,X]. The magnitude of this relation for various words might be of interest in some analyses, but this relation is deleted from current QUORUM models.

Relations with pronouns— Relations with pronouns have great potential to aid in domain analysis. QUORUM models which include pronouns can be very useful. For example, the differences between the kinds of things associated with "I" and those associated with "we" suggest the limits of teamwork involving the use of automation and other cognitive activities (McGreevy, 1996). Unfortunately, the frequent references to "I" and "we" in many incident narratives causes these words to become the hub around which all other words revolve. To avoid having most relations in a model involve pronouns, relations with pronouns are often, but not always, deleted from QUORUM models.

Relations with units of measure— In ASRS incident reports, the word "FT" (i.e., feet) is dominant. This is because specific altitude is a significant factor in the context of an incident, and it often plays a central role in the incident itself. This very prominence, however, causes FT to be related to many, many other words in the domain. While this relatedness is meaningful, these relations can crowd out more specific relations. After an initial QUORUM model is made, and any pervasive concern with certain units of measure is noted, it can be useful to delete relations with units of measure from subsequent models. In fact, relations with units of measure are but one example of "domain-generic relations," the next category of relation.

Domain-generic relations— Domain-generic relations are those which are common to the domain, but shed little light on the specifics of the incidents. This is a gray area. An initial model of aviation incident narratives should include such relations as [FT, ALT], which indicates that the word "altitude" is often found in the context of the word "feet." Another example, [ALT, ACFT], indicates a contextual association of "altitude" and "aircraft." Such relations are probably useful in projects such as domain identification (determining the domain of the text) and domain mapping/modeling (modeling the overall structure of the domain). Domain-generic relations are not particularly useful, however, to operational analysts who are interested in the structure of specific, problematic situations. They are likely to take it for granted that "feet" and "altitude" or "altitude" and aircraft" are closely associated. Similarly, generic relations involving "said," "say," or "says" in news stories are of limited value if one is interested in the structure of the domain described in the stories, as opposed to the domain of news gathering. For these reasons, it can be useful to delete prominent domain-generic relations.

Situation-generic relations— Situation-generic relations are those which are common to particular situations, but shed little light on the specifics of these situations. This is another gray area. For example, it can be useful, initially, to find that there are many relations involving "approach" in a collection of aviation incidents, as this could indicate that the approach phase is prominently represented among the reports. Once this fact is established, however, relations such as [APCH,RWY], indicating that "approach" and "runway" are closely associated, are obvious and could be deleted, depending on the goals of the project. Such relations are a subset of domain-generic relations.

Location-specific relations— Location-specific relations are those which involve specific airports, runways, airway intersections, altitudes, or other spatial locations. Some examples are: [RWY, 4L], indicating the prominence of "runway 4L," [10000, FT], indicating the prominence of "10000 feet," and [LAX, APCH], indicating the prominence of such phrases as "LAX (Los Angeles) approach" or "on approach to LAX". Location-specific relations might be of significant interest to some analysts because they are useful for finding patterns among incidents involving certain airports, runways, intersections, altitudes, and the like. Other analysts are more interested in certain kinds of problems wherever they occur. For the latter group, location-specific relations should be deleted.

Infrastructure relations— ASRS narratives sometimes contain text that has been added by an ASRS database specialist, such as the phrase:

```
"CALLBACK CONVERSATION WITH RPTR REVEALED THE FOLLOWING INFO."
```

If this and similar text is not deleted in pre-processing, relations such as [CALLBACK, CONVERSATION] can be prominent in the model. This problem is worse in analyses of news stories, which are cluttered with widely varying information in addition to the news stories themselves, including bylines, disclaimers, promotion of Internet sites of the news organizations, notes to editors from the wire service, and the like. It is useful to delete as much of this clutter as possible, and to delete any relations which slip through, such as the relations [Press, Writer], [World, Wide], and [World, Web] in the collection of Flight 800 stories.

Object relations— After all of the kinds of relations above have been resolved, what remains are intra-object relations and inter-object relations. Object relations, regardless of their specific type, are the most useful relations for situation and domain modeling (McGreevy, 1996; McGreevy, 1995), and relevance ranking. Being explicit about the kind of object relation is unnecessary for most analyses, but at the very least, an intuitive feeling for what constitutes a useful relation is important.

Intra-object relations are relations within one object. Here are the main kinds of intra-object relations that can exist between two words:

word X	word Y
object A	object A
object A	action of object A
object A	attribute/part of object A
object A	attribute value of object A
action of object A	action of object A
action of object A	attribute/part of object A
action of object A	attribute value of object A
attribute/part of object A	attribute/part of object A
attribute/part of object A	attribute value of object A

```
attribute value of object A   attribute value of object A
```

Examples of intra-object relations include: [Flight, 800], [TWA, Flight], [TWA, 800], [New, York], [Long, Island], [Aviation, Federal], [Aviation, Administration], [Federal, Administration], [tank, center].

Inter-object relations are relations between two different objects, or some aspects of two different objects. Here are the main kinds of inter-object relations that can exist between two words:

word X	word Y
object A	object B
object A	action of object B
object A	attribute/part of object B
object A	attribute value of object B
action of object A	object B
action of object A	action of object B
action of object A	attribute/part of object B
action of object A	attribute value of object B
attribute/part of object A	object B
attribute/part of object A	action of object B
attribute/part of object A	attribute/part of object B
attribute/part of object A	attribute value of object B
attribute value of object A	object B
attribute value of object A	action of object B
attribute value of object A	attribute/part of object B
attribute value of object A	attribute value of object B

Examples of inter-object relations include: [fuel, tank], [bomb, missile], [fuel, center], [plane, bomb].

Multi-word entities, such as "New York," "TWA Flight 800," "National Transportation Safety Board," "mode control panel," "level off," and "altitude window," require special consideration when categorizing object relations. Experience has shown that explicitly linking all such word groups degrades the performance of QUORUM, especially in relevance ranking. It is better to treat each word in a multi-word entity as separate element. For example, rather than link "TWA," "Flight," and "800" as a single word, "TWA_Flight_800," one can suppose that "TWA" is an object, "Flight" a part of the object "TWA," and "800" is an attribute value of "Flight." Alternatively, one can suppose that the words "TWA," "Flight," and "800" are all fragments of the composite object "TWA Flight 800," but without treating "TWA_Flight_800" as a single word. Similarly, a relation like [New, York] can be considered to be an intra-object relation of the type [object A, object A], that is, a relation between (a fragment of) object A and (another fragment of) object A.

Off-topic relations— Once the relations that are judged to be extraneous are deleted, and the object relations are collected, what remains is a well-scrubbed model of the collection of text. In the case of the Flight 800 stories, this can still include references to other airlines, other aviation concerns, and other disasters. An analyst might consider these to be of interest, or might consider them to be off the topic. Any relations which are not of interest can be deleted in order to focus the model on a particular topic.

It might by useful for some purposes to refine the model of the Flight 800 collection, for example, by removing such relations as [Delta, Continental], referring to the possible merger of these airlines, [crash, ValuJet], referring to the disaster in the Florida Everglades, and [Airlines, Delta], referring to incidents involving Delta Airlines. This would more tightly focus the model on Flight 800 itself, rather than including related airline issues.

In addition, it is possible to retain only such relations as those containing "FBI" or "Kallstrom" for a detailed look at one aspect of the Flight 800 story, as shown in the section, "Ranking by Topical Focus." Even among these relations, those such as [FBI, siege], [FBI, militants], [FBI, Ruby], and [FBI, Ridge] should be deleted if one wishes to maintain a tight focus on Flight 800.

Word Groups

Word group analysis can help in the interpretation of QUORUM relations. It can also show how particular systems, subsystems, or other details are named in the analyzed narratives. In addition, subsequent retrievals of reports can be based on word groups of particular interest.

Obviously, some of the contextual association of words, as analyzed by the QUORUM method, is due to word groups. For example, in narratives describing incidents involving the autopilot, the relation between MODE and PANEL is partly due to the frequency of the word group "MODE CTL PANEL" (i.e., mode control panel). Analysis of word groups provides insight into how prominently related words are actually grouped in the text. Analysis of word groups also helps in the development of multi-word vocabularies based on prominent relations. Here are some examples of word groups occurring among 313 ASRS reports describing automation errors in glass cockpits. The first column contains the frequency of occurrence. Lines containing "..." indicate that some of the word groups are not shown.

The first list has word groups containing MODE.

freq	word group
12	MODE CTL PANEL
8	SPD MODE
4	ALT CAPTURE MODE
4	VERT NAV MODE
3	ARC MODE
3	FLT MODE ANNUNCIATOR
3	MANUAL MODE
3	MAP MODE
3	MODE C
3	NAV MODE
3	PROFILE MODE
3	VERT SPD MODE
3	VNAV MODE
2	ACR Y'S MODE C
...	

This list has word groups containing PANEL:

freq	word group
12	MODE CTL PANEL
8	OVERHEAD PANEL
3	AUDIO PANEL
3	DIGITAL FLT GUIDANCE PANEL
2	AFDS PANEL
2	CENTRAL WARNING PANEL
2	CIRCUIT BREAKER PANEL
2	ELECTRICS PANEL
2	FLT CTL PANEL
2	FMS AND MODE CTL PANEL
2	FT IN THE MODE CTL PANEL
2	INST PANEL
2	LATERAL PANEL
2	LNDG GEAR INDICATING PANELS
...	

Finding word groups derived from prominent relations, such as those in the two tables above, can assist in focusing on particular modes, panels, or other specific elements of systems and automation. By understanding how incident reporters name things, it is possible to retrieve and analyze reports more effectively.

The low frequencies of the word groups in the two preceding tables indicate a lack of focus on any one mode or panel in the analyzed collection of reports. In retrieving a subsequent collection of reports from the database, however, an analyst could retrieve a large number of reports containing one or more of these word groups. The QUORUM method can then be applied to look for patterns among those incidents.

Relevance Criteria versus Selection Criteria

It is important not to confuse report selection criteria with relevance criteria. Selection criteria determine which reports are to be gathered from the database into the collection to be analyzed. Relevance criteria determine the order of presentation of text items from the collection. It is often useful to have broad selection criteria, but more focused relevance criteria. For example, one might gather several years of reports containing the word "mode" and then rank them according to particular automation concerns.

If the analyst already has a QUORUM model in hand, that model can be used to select relevant reports from the database. Even in current databases, the most closely related word pairs in the model can be used as selection criteria to obtain more reports. For example, if AUTOPLT and DISCONNECTED are found in the same sentence or within, say, 20 words of each other in a narrative, then the report containing that narrative could be selected. Reports that were selected by the largest number of prominent word pairs would be the most relevant ones.

If a database retrieval system were designed to fully utilize QUORUM models, then the relevance criteria could be used directly as selection criteria. That is, the relations of any QUORUM model, taken together, could be used to query the database. For this to work, a QUORUM model of the narrative of each report would be pre-computed, perhaps as part of the process of entering the report into the database. This would produce a table for each narrative containing a QUORUM model of that narrative. The selection and retrieval process would then consist of finding those reports that have prominent relations that are also prominent among the relevance criteria of the query. Candidate reports would be relevance ranked and the most relevant ones would be retrieved.

Thus, by building QUORUM models of each narrative into a database, other QUORUM models can be used as relevance criteria to query the database directly. This would allow analysts to select the reports that most closely match their interests and concerns.

Discussion

Take-Home Messages for the Operationally-Oriented Reader

Using the QUORUM method involves four steps:

1) Report selection,
2) Narrative analysis,
3) Situational modeling, and
4) Relevance ranking.

Relevance ranking using QUORUM is a way to quickly find patterns among large numbers of incident reports without having to interpret complex and abstract models. Instead, the analyst can stay focused on text from the narratives themselves.

The QUORUM method can help operational analysts to quickly locate relevant narratives and sentences from large collections of incident reports. In particular, QUORUM can find narratives and sentences that

- are typical of those in the current collection,
- involve a topic of interest,
- involve several topics of interest,
- are typical of those in some other collection,
- are similar to example incidents of particular interest, or
- are relevant to specialized interests defined by the analyst.

Other benefits of the method are that:

- QUORUM relevance criteria are explicit and can be refined and re-used; and
- QUORUM relevance criteria can be used to retrieve relevant reports from databases.

The QUORUM method is readily available for use by others. The descriptions of the method in this paper, and in McGreevy (1995) and McGreevy (1996), are sufficient to guide implementation by interested parties.

QUORUM software is designed for research use, not distribution. It is under constant development and is not currently available for use by others.

Related Research

The vast amount of text available in electronic form has contributed substantially to the "information glut." In response, researchers are generating their own overabundance of papers about ways to deal with it. Methods described in this literature are typically based on finding and exploiting patterns in collections of text. Variation among methods and factions is primarily due to varying allegiances to linguistics, quantitative analysis, representations of domain expertise, and the practical demands of the applications. Typical applications involve finding items of interest from large collections of text, having appropriate items routed to just the right people, and condensing the contents of many documents into summary form.

QUORUM methods and applications are related in a general way to research addressing the information overload. Related research includes work in: data mining (Fayyad, Piatetsky-Shapiro, and Smyth, 1996), search engines (Zorn, Emanoil, Marshall, and Panek, 1996), discourse analysis (Kitani, Eriguchi, and Hara, 1994), information extraction (Cowie and Lehnert, 1996), information filtering (Foltz and Dumais, 1992), and information retrieval (Salton, 1991). Cutting across these approaches are concerns about how to subdivide words and collections of words into useful pieces, how to categorize the pieces, how to detect and utilize various relations among the pieces, and how transform the many pieces into a smaller number of representative ones.

The QUORUM method has some specific similarities to work done by others. For example, Hawking and Thistlewaite (1996, 1995) are also developing a proximity metric to support relevance ranking. Chen, Hsu, Ortwig, Hoopes, and Nunamaker (1994) use a proximity metric to produce summary outlines of large bodies of text. Greffenstette (1993) and Jing and Croft (1994) use proximity metrics to extract clusters of related words from text to elicit word meanings and create thesauri. Kupiec, Pedersen, and Chen (1995) generate document extracts using various heuristics.

The work of Hawking and Thistlewaite (1996, 1995) on the PADRE system is apparently most similar to QUORUM relevance ranking, but the two methods were developed independently and differ substantially. QUORUM and PADRE are similar in that they both apply proximity metrics to determine the relevance of documents. Their definitions of proximity and relevance are very different, however.

- QUORUM measures the proximity-weighted co-occurrences of pairs of words, while PADRE measures the spans of text that contain clusters of any number of target words. Thus, QUORUM is based on binary relations and PADRE is based on multi-way ("N-ary") relations.
- QUORUM relations have a simple and clear definition, while PADRE spans and clusters have complex, non-intuitive, and somewhat arbitrary definitions.
- Each use of PADRE to rank documents requires specification of a small group of words that might be clustered in the text. In contrast, the many binary relations that constitute QUORUM relevance criteria automatically detect a wide variety of word clusters in the text.
- QUORUM relevance criteria represent a large number of the most prominent contextual relationships, ranging from the obvious to the subtle, and they may be systematically refined. PADRE criteria have much less resolution and potential for refinement.
- QUORUM relevance criteria consist of word pairs whose contributions to relevance are graded, while PADRE relevance criteria are based on the assumption that the greatest relevance is achieved when all of the target words are closest to each other.
- QUORUM relevance criteria are detailed models of the text from which they were derived, while PADRE relevance criteria are generated by "human free association."

The QUORUM proximity metric and QUORUM relevance ranking based on that metric seem to offer significant advantages over PADRE's metric and relevance ranking. QUORUM offers greater objectivity, efficiency, and versatility.

Finally, the application of QUORUM to narratives suggests some commonality with work in the field of qualitative narrative analysis ("narratology"). In contrast to the quantitative QUORUM method, however, the field of narratology is a more humanistic approach to the interpretation of stories contained in narratives (Berger, 1997; Riessmann, 1993). The goal of narratology is to understand the nature of stories in general. Narratology is concerned with the underlying symbolic structures, sources, motivations, and effects of stories, rather than their contents. While QUORUM objectively finds prominent commonalities among large numbers of incident narratives, narratology applies more specialized and

subjective analyses to one or a few stories. These few stories are often derived from interviews with individuals thought to be representative of some sociologically significant group, or they are selected from mass media, folk tales, or literature. Despite their differences, the qualitative and quantitative approaches to narrative interpretation have the potential to exchange conceptual and methodological insights. For example, narratology articulates the factors which influence the transformation of raw experience to written form, and its interpretation by the reader, while QUORUM provide a more objective measure of structure and emphasis.

Why QUORUM Works

QUORUM works because it is based on a model of presence. In that model, presence is viewed as a web of physical and logical adjacencies that are imposed by the structure of the working environment, engagement with that structure, and the demands of the domain in general and the situation in particular. Thus, presence is modeled as association by contiguity (i.e., metonymic relatedness) among the concerns of the person present (McGreevy, 1994; McGreevy, 1992). Since a domain can also be modeled as adjacencies among concerns, such a domain model is a model of presence in the domain (McGreevy, 1995).

This model of presence readily applies to the analysis of narratives. The fundamental reason for this applicability is that narratives are derived from presence. In constructing narratives, incident reporters represent their concerns about the working environment, their role in that environment, the demands of day-to-day commercial aviation operations in general, and the incidents in particular. Further, "the basic impulse of narrative prose is association by contiguity" (Jacobson, 1987, pg. 310), so the structure of narrative prose is similar to the structure of presence. Both narratives and presence can be modeled as association by contiguity, that is, logical and physical adjacencies, among the concerns of the person present in the situation.

QUORUM measures, models, and ranks narratives according to their characteristic associative structure. This structure is fundamentally based on the presence of the reporter in the problematic situation. That presence is transformed by the reporter, according to his or her concerns, into the linear array of words that constitutes the narrative. When incident reporters tell their stories in narrative form, importantly associated concerns tend to appear in closer proximity to one another than do less-importantly associated concerns. As a result, importantly associated domain words tend to be found in closer proximity to one another than do less-importantly associated domain words. Since there is a direct correspondence between the concerns of the reporters and the words they use to describe their concerns, the structure among reporter concerns can be measured by measuring the structure among the words used to describe those concerns.

By measuring and modeling the proximities among words in the domain vocabulary as they are distributed in narratives, QUORUM effectively measures and models the proximities among the concerns of the incident reporters. Further, since the structure among concerns is based on the presence of the reporters in particular situations, that structure can be interpreted as the structure of their presence in those situations. When large numbers of narratives are analyzed, the measured structure is based on many writers and many situations. This tends to make the commonalities among situations more prominent, while downplaying atypical details. The structure is sparsely modeled by including only the most prominent associations. This results in a tolerable model of reality because the many weak associations are safely ignored, and only the relatively few strong associations are retained. The model is "situated" because it includes the mutual contexts of all prominent elements of the situations.

QUORUM ranks text from incident narratives using relevance criteria that are based on QUORUM models. As a consequence, the highest ranking items represent the most prominent concerns of the reporters about the details of specific incidents, the situations in which the incidents occurred, aviation safety in general, and personal responsibility in particular.

Conclusion

The details of problematic incidents and situations are the raw data from which operational safety and efficiencies can be derived. In addition, the vitality and focus of aviation safety research depends upon up-to-date, detailed awareness of day-to-day operational problems. For these reasons, the interpretation of incident reports plays an important role in aviation safety. The QUORUM method of narrative analysis, modeling, and relevance-ranking has the potential to assist those who interpret large volumes of incident reports. In that way, QUORUM can contribute to improvements in aviation safety.

References

Berger, A. A.: Narratives in Popular Culture, Media, and Everyday Life. Sage, Thousand Oaks, Calif., 1997.

Chen, H.; Hsu, P.; Ortwig, R.; Hoopes, L.; and Nunamaker, J. F.: Automatic Concept Classification of Text from Electronic Meetings. Comm. ACM, vol. 37, no. 10, 1994, pp. 56-73.

Cowie, J. and Lehnert, W.: Information Extraction. Comm. ACM, vol. 39, no. 1, 1996, pp. 81-91.

Deerwester, S.; Dumais, S. T.; Furnas, G. W.; Landauer, T. K.; and Harshman, R.: Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science, vol. 41, no. 6, 1990, pp. 391-407.

Fayyad, U.; Piatetsky-Shapiro, G.; and Smyth, P: The KDD Process for Extracting Useful Knowledge from Volumes of Data. Comm. ACM, vol. 39, no. 11, 1996, pp. 27-34.

Foltz, P. W. and Dumais, S. T.: Personalized Information Delivery- An Analysis of Information Filtering Methods. Comm. ACM, vol. 35, no. 12, 1992, pp. 51-60

Greffenstette, G.: Automatic Thesaurus Generation from Raw Text using Knowledge-Poor Techniques. Technical Report MLTT-001, Rank Xerox Research Center, Grenoble, France, 1993.

Hawking, D. A. and Thistlewaite, P. B.: Proximity Operators - So Near And Yet So Far. To appear in D.K. Harman (Ed.) Proc. Fourth Text REtrieval Conf., Gaithersburg, Maryland, November 1995.

Hawking, D. A. and Thistlewaite, P. B.: Relevance Weighting Using Distance Between Term Occurrences. Technical Report TR-CS-96-08, Department of Computer Science, Australian National University, June 1996.

Jacobson, R.: Language in Literature. Belknap Press, Cambridge, Mass., 1987.

Jing, Y. and Croft, W. B.: An Association Thesaurus for Information Retrieval. Technical Report 94-17, University of Massachusetts, 1994.

Kitani, T.; Eriguchi, Y.; and Hara, M.: Pattern Matching and Discourse Processing in Information Extraction from Japanese Text. JAIR, vol. 2, 1994, pp. 89-100.

Kupiec, J.; Pedersen, J.; and Chen, F.: A Trainable Document Summarizer. In Proceedings of ACM SIGIR, 1995.

McGreevy, M. W.: The Presence of Field Geologists in Mars-Like Terrain. Presence, vol. 1, no. 4, 1992, pp. 375-403.

McGreevy, M. W.: An Ethnographic Object-Oriented Analysis of Explorer Presence in a Volcanic Terrain Environment. NASA TM-108823. Ames Research Center, Moffett Field, Calif., 1994.

McGreevy, M. W.: A Relational Metric, Its Application to Domain Analysis, and an Example Analysis and Model of a Remote Sensing Domain. NASA TM-110358. Ames Research Center, Moffett Field, Calif., 1995.

McGreevy, M. W.: Reporter Concerns in 300 Mode-Related Incident Reports from NASA's Aviation Safety Reporting System. NASA TM-110413. Ames Research Center, Moffett Field, Calif., 1996.

Nardi, B. A.: Studying Context: A Comparison of Activity Theory, Situated Action Models, and Distributed Cognition. In East-West Int'l. Conf. on Human-Computer Interaction, 1992, pp. 352-359.

Reynard, W.: Memorandum for Recipients of Aviation Safety Reporting System Data, Caveat Regarding Statistical Use of ASRS Information. Undated.

Riessman, C. K.: Narrative Analysis. Sage, Newbury Park, Calif., 1993.

Salton, G.: Developments in Automatic Text Retrieval, Science, vol. 253, 1991, pp. 974-980.

Schvaneveldt, R. W; Durso, F. T.; and Dearholt, D. W.: Network Structures in Proximity Data. In G. Bower (ed.), The Psychology of Learning and Motivation (pp. 249-284). Academic Press, New York, 1989.

Simon, H. A.: The Sciences of the Artificial. MIT Press, Cambridge, Mass., 1969.

Zorn, P.; Emanoil, M.; Marshall, L; and Panek, M.: Advanced Web Searching: Tricks of the Trade. ONLINE, vol. 20, no. 3, 1996, pp. 14-28.

Appendix. Glossary of ASRS abbreviations that appear in this paper.

A-320	Airbus 320
A320	Airbus 320
ACFT	aircraft
ACR	air carrier
AFDS	automatic flight director system
ALT	altitude
APCH	approach
APCHED	approached
APPROX	approximately
ARPT	airport
ATC	air traffic control
ATIS	automatic terminal information service
ATL	Atlanta
ATTN	attention
AUTOFLT	auto-flight
AUTOPLT	autopilot
AUX	auxiliary
B757	Boeing 757
CAPT	captain
CAPT'S	captain's
CHK	check
CHKING	checking
CLB	climb
CLBED	climbed
CLBING	climbing
CLBOUT	climb-out
CLRED	cleared
CLRNC	clearance
CRZ	cruise
CSD	constant speed drive (unit)
CTL	control
CTLR	controller
CTR	center
DCA	Washington National
DEGS	degrees
DEP	departure
DEV	deviate, deviation
DME	distance measuring equipment
DSCNT	descent
DSND	descend
DSNDED	descended
EADU	electronic attitude director unit
ENG	engine
ENRTE	enroute
FAF	final approach fix
FD	flight director
FL180	flight level 180 (18000 feet)
FL200	flight level 200 (20000 feet)
FL220	flight level 220 (22000 feet)
FL230	flight level 230 (23000 feet)
FL240	flight level 240 (24000 feet)
FL260	flight level 260 (26000 feet)
FL310	flight level 310 (31000 feet)
FL320	flight level 320 (32000 feet)
FL330	flight level 330 (33000 feet)
FL350	flight level 350 (35000 feet)
FLC	flight crew
FLCH	flight level change
FLT	flight
FMA	flight mode annunciator
FMC	flight management computer
FMS	flight management system
FO	first officer
FPM	feet per minute
FT	feet
HDG	heading
HORIZ	horizontal
HR	hour
HYD	hydraulic
IAS	indicated air speed
ILS	instrument landing system
INFO	information
INST	instrument
INTXN	intersection
JAXSN	name of an intersection
KIAS	knots indicated air speed
KTS	knots
L	left
LAV	lavatory
LAX	lax; Los Angeles
LBS	pounds
LGA	La Guardia
LL10Z	ASRS-encoded time of day
LNDG	landing
LOC	localizer
MCP	mode control panel
MD88	McDonnell-Douglas 88
MDT	an airport in Pennsylvania; medium transport
MEL	minimum equipment list
MGMNT	management
MI	mile
MINS	minutes, minimums
MSL	mean sea level
MSP	Minneapolis-St. Paul
NAV	navigation
NM	nautical mile
NY	New York
PERF	performance
PF	pilot flying
PLT	pilot
PNF'S	pilot-not-flying's
POS	position
PROB	probably
PROCS	procedures
PWR	power
QUANT	quantity
R	right
RA	resolution advisory
RECLRED	recleared
RESTR	restriction
RPT	report
RPTED	reported
RPTR	reporter
RTE	route
RWY	runway
S	south
SAN	San Diego
SIT	situation
SOMTO	name of an intersection
SOP	standard operating procedure
SPD	speed
SPDS	speeds
SYS	system
TA	traffic advisory
TA'S	traffic advisories
TCASII	traffic alert and collision avoidance system 2
TFC	traffic
TKOF	takeoff
VERT	vertical
VFR	visual flight rules
VNAV	vertical navigation
WDB	wide body (aircraft)
XA30	ASRS-encoded time of day
XCHK	cross check
XING	crossing
ZDC	an air route traffic control center
ZDV	an air route traffic control center

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1997	Technical Memorandum

4. TITLE AND SUBTITLE

A practical guide to interpretation of large collections of incident narratives using the QUORUM method

5. FUNDING NUMBERS

6. AUTHOR(S)

Michael W. McGreevy

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Ames Research Center
Moffett Field, CA 94035-1000

8. PERFORMING ORGANIZATION REPORT NUMBER

A-976313

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

NASA TM-112190

11. SUPPLEMENTARY NOTES

Point of Contact: Michael McGreevy, Ames Research Center, MS 262-A/4, Moffett Field, CA 94035-1000 (415) 604-5784

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified-Unlimited
Subject Category–03

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Analysis of incident reports plays an important role in aviation safety. Typically, a narrative description, written by a participant, is a central part of an incident report. Because there are so many reports, and the narratives contain so much detail, it can be difficult to efficiently and effectively recognize patterns among them. Recognizing and addressing recurring problems, however, is vital to continuing safety in commercial aviation operations.

A practical way to interpret large collections of incident narratives is to apply the QUORUM method of text analysis, modeling, and relevance ranking. In this paper, QUORUM text analysis and modeling are surveyed, and QUORUM relevance ranking is described in detail with many examples. The examples are based on several large collections of reports from the Aviation Safety Reporting System (ASRS) database, and a collection of news stories describing the disaster of TWA Flight 800, the Boeing 747 which exploded in mid-air and crashed near Long Island, New York, on July 17, 1996. Reader familiarity with this disaster should make the relevance-ranking examples more understandable. The ASRS examples illustrate the practical application of QUORUM relevance ranking.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Incident report analysis, Aviation safety, Relevance ranking, Knowledge acquisition from text, Virtual reality and presence	46
	16. PRICE CODE: A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified		