Analysing Air Incident Reports: Workshop Challenge

Derek Bridge[1], Paulo Gomes[2], and Nuno Seco[2]

[1] University College Cork, Cork, Ireland
d.bridge@cs.ucc.ie
[2] Universidade de Coimbra, Coimbra, Portugal
{pgomes|nseco}@dei.uc.pt

1 Introduction

To promote discussion at the Fourth Workshop on Textual Case-Based Reasoning (TCBR), we set a *workshop challenge*. We encouraged all potential workshop participants, individually or in their research groups, to submit a short paper that addressed this challenge. The purpose of this paper is to explain the challenge.

We challenged participants to do the following: analyse the corpus of Air Investigation Reports available from the Transportation Safety Board of Canada [3]; imagine that they were to use the corpus to build a TCBR system to support human investigators; and then to write either:

– a short paper describing some of the problems the corpus poses for the development of a TCBR system;
– a longer paper describing the problems and proposing ways of overcoming some of these problems; or
– a longer paper reporting actual experience in addressing the problems.

In Section 2, we will give some background to incident analysis; in Section 3, we will describe the air incident reports in the corpus; in Section 4 we will highlight some of the relevant themes to be found in past TCBR research.

2 Incident Analysis

An *accident* is an event that unintentionally causes damage or injury. An *incident* is an event that unintentionally causes damage or injury or has the potential to do so. Hence, by these definitions, the set of incidents includes the set of accidents, but it includes many other events too. For reasons of good practice or legislative compliance, many organisations use computer-based *incident reporting and analysis systems*, with the goal of detecting and remedying problems before they result in accidents.

The following (from [1]) summarises the arguments in favour of incident reporting and analysis systems:

[3] `http:www.tsb.gc.ca/en/reports/` and `http://www.tsb.gc.ca/fr/reports/`

– *Incident reports help to find out why accidents* don't *happen.* They identify barriers that prevented an incident being an accident. Similar barriers might then be legislated for, strengthened, or applied more widely.
– *The higher frequency of incidents permits quantitative analysis.* Incidents are more frequent than accidents. Contributing factors can be analysed statistically.
– *They provide a reminder of hazards.* These systems make it more likely that recurrent problems will be detected and remedied.
– *Feedback keeps staff 'in the loop'.* These systems encourage staff participation in safety improvement.
– *Data (and lessons) can be shared.* Common problems, solutions and good practice can be disseminated within and between organisations.
– *Incident reporting schemes are cheaper than the costs of an accident.* The costs of running such systems, if successful, should be compared with the benefits in accidents prevented.
– *May be required to do it.* As mentioned above, such systems may be required by regulatory bodies.

A computer-based incident reporting and analysis system might support the staff in an organisation in both the *acquisition* and *analysis* of reports. The tasks that the system might support include:

– authoring a new investigation report;
– proposing safety actions in response to a new incident; and
– discovering recurring unsolved problems.

But there is often a trade-off between ease of acquisition and ease of analysis. For example, free-text reports may be easy to author but they are likely to be incomplete, inconsistent in their level of detail, and hard to process automatically. Structured documents (with prescribed fields and field values) or semi-structured documents (with prescribed fields or headers, some of which might be filled by free-text) place more strictures on human authors but are more amenable to analysis.

In some cases, incidents are labelled with terms from taxonomies, coding schemes or other controlled vocabularies. Johnson [1] gives an example of the Eindhoven classification scheme for reporting incidents in the Accident and Emergency Departments of UK hospitals: an analyst assigns the label O-EX to an incident, for example, if the incident is the result of an organisational factor (rather than, e.g., a technical or human behaviour factor) that is external to the reporting organisation.

Taxonomies, coding schemes and other controlled vocabularies can suffer from the following problems [1]:

– *How do we avoid confirmation bias?* The scheme may encode current preconceptions and be unamenable to alternative interpretations of events.
– *How to identify an appropriate level of abstraction?* If the level of abstraction is too high, analysis may be vapid and recommendations may be vague; if the level of abstraction is too low, authoring effort increases and the level of reporting consistency between experts decreases.

– *What to do when incidents change?* Technical, managerial, operational and legal change may render parts of the scheme obsolete and may demand the extension of the scheme to new circumstances.
– *How do we decide what is relevant?* The salience of certain factors may only be appreciated at some point after reporting of an incident, e.g. in the aftermath of some other accident. Then, incidents may need to be recoded; the scheme itself may even need to be extended.
– *How do we keep track of complex causal relationships?* These schemes typically detach factors from their context.

In the next section, we will describe the corpus of incident reports that we are using for our workshop challenge.

3 Air Investigation Reports

Our workshop challenge uses the corpus of Air Investigation Reports available from the Transportation Safety Board of Canada. The Transportation Safety Board of Canada (TSB) is the agency, independent of government and the transportation industry, that investigates incidents in Canadian marine, pipeline, rail and air transportation, with the goal of enhancing transportation safety. When an incident occurs, those involved are required to report it to the TSB. The TSB determines whether an investigation is warranted. When this is completed, an investigation report is written, agreed and published. In the case of air transportation incidents, these are the reports that comprise the corpus for our workshop challenge. The TSB periodically also publishes quantitative and qualitative analyses of incidents and recommendations.

Investigation reports are grouped on the TSB's web site by transportation mode (air, marine, pipeline, rail) and by year. An excerpt from one such Air Investigation Report (Report Number A05O0147) is given in Fig. 1. This is a very small excerpt: the full report is over 2,600 words long.

It is easy to imagine a rudimentary search facility for these investigation reports: either one might use a search engine that carries out full-text indexing and evaluates Boolean queries against the index, or one might use a document retrieval system, e.g. one that represents documents and queries in vector-space. In fact, the TSB web site already uses Microsoft Index Server to offer a simple index-based search facility.

To support a user in the tasks we identified in Section 2 (i.e. authoring new reports, proposing safety actions, and discovering recurring problems) requires more than these simple search engines currently offer. But to offer more than this, i.e. to offer the deeper level of automated analysis that is required, is extremely challenging.

The following are some of the characteristics of these investigation reports that make automated support more challenging:

– *Section headers are not prescribed.* Most reports have the same section headings as the one depicted in Fig. 1, i.e. Summary, Other Factual Information, Analysis, Findings as to Causes and Contributing Factors, Findings as to Risk, Other Findings, Safety Actions Taken, and Safety Concern. But many have only a subset of these headings, e.g. Safety Concern is often missing; sometimes Safety Actions Taken is missing; sometimes one or more of the sections of findings is absent or there is just a single Findings section. Some have appendices containing maps or glossaries.

Other reports use completely different section headings. For example, there seems to be a distinct subset of the reports, including, e.g., report number A05C0187, in which the reports have a Synopsis and then sections entitled Factual Information, Analysis, Conclusions (which may contain subsections for Findings as to Causes and Contributing Factors, Findings as to Risk, and Other Findings) and Safety Action. Furthermore, the reports that fall into this subset, unlike most other reports, are not stored in a single HTML file. Rather, each section is stored in a separate HTML file.

In many of the reports, long sections are split into subsections that deal with different aspects of the incident (e.g. weather, personnel, etc.), and their headings are incident-specific and cannot be expected to recur through the corpus.

Finally, Document A05H0002 is also unlike the others. It is designated as an investigation *update*. Accordingly, it has completely different section headings.

– *Section headers are only indicative of content.* Authors may find it hard to decide what content to assign to the different sections of the report: different authors may disagree; they may not even be self-consistent. It is not unusual for the content of earlier descriptive sections to anticipate the content of the later more evaluative sections; it is not unusual for the later more evaluative sections to repeat information given in the descriptive sections (e.g. to explain causal relationships).
– *High linguistic complexity.* The reports are long compared to the texts used in many previous TCBR systems; vocabulary is uncontrolled; and grammatical constructions can be complex (e.g. negation, causal and conditional relationships). Domain-specific technical terms, names and abbreviations are common. The text is 'well-considered' (it has clearly been well-reviewed), but there is still the potential for errors in typing, spelling, and grammar.
– *Multimedia content.* Maps are common. Occasionally, there are digital photos or diagrams, e.g. of aircraft components. In all cases, the text will refer to these images, often highlighting particular parts of the image.
– *Multilingual content.* Canada is bilingual. Accordingly, each report is availabel in both French and English. The translation from one to the other appears to be of high quality. The implication is that a computer system for Canadian air incident reporting would have to support both languages.
– *Inter-linked documents.* The corpus is not simply a *set* of documents. First, some reports explicitly reference earlier reports. We can see an example of this in Fig. 1, which twice cross-references report number A04W0114. Second, reports are arranged chronologically by report date. (Their text may contain the times and dates of events too.) This sets up temporal relation-

ships between reports and the events they describe. This may be especially relevant to consideration of, for example, the possible obsolescence of reports due to, e.g., changes in aircraft manufacture, in airport infrastructure, or in policies and legal requirements. Indeed, some of this obsolescence is a desirable feature: it shows that safety recommendations made in earlier reports have been heeded and therefore causal factors of later incidents have changed.

In the next section, we briefly discuss relevant Case-Based Reasoning systems for handling incident reports.

4 Case-Based Reasoning for Incident Reporting and Analysis

Case-Based Reasoning has long been applied to 'lessons learned' systems and therefore it seems that it may be well-suited to incident reporting and analysis. Indeed, as we will now describe, Textual Case-Based Reasoning and Conversational Case-Based Reasoning have both been applied in this domain already.

Johnson used the US Navy's NaCoDAE system to build *Conversational CBR* systems based on a subset of US Aviation Safety Reporting System reports [1–3]. These particular reports are much more structured than those from the Canadian TSB. Each report may contain structured fields with values that categorise, for example, the type, location, consequences and resolutions of anomalies detected. In one system, these fields were used to encode the questions associated with the cases. In a second system, additional questions were associated with cases by manually classifying them using the Eindhoven classification scheme mentioned in Section 2.

Before reviewing work in which *Textual CBR* has been applied in this domain, it is worth contrasting textual CBR with IR, information retrieval (i.e. document retrieval) because the distinction between the two recurs throughout a lot of the past TCBR research.

The task of an IR system is to retrieve all documents from some collection that are relevant to a user's need (hence, precision and recall are primary evaluation measures). IR systems are typically domain-independent [4]: any knowledge they use (e.g. term weights) tends to be obtained in a general way from the document collection, occasionally supplemented by other generic resources such as thesauruses.

By contrast, CBR systems are task- and domain-dependent [4]. The cases (textual or otherwise) will represent experiences related to the task and domain, and they will be a major part of the knowledge that is used to draw inferences [5] in carrying out the reasoning within the task (be it classification, diagnosis, interpretation, or whatever) [6]. The CBR system will integrate other task- and domain-specific knowledge to improve proficiency at the task [4]. In the case of textual CBR systems, this might include task- and domain-specific keyword dictionaries, phrase dictionaries, thesauruses, glossaries, etc. [7] as well as similarity measures and adaptation knowledge. The system should be evaluated by its proficiency at the task: for many, if not most, tasks, precision and especially recall may not be relevant measures. In summary, relative to IR systems, a textual CBR system will tend to “eschew flexibility and generality for precision and utility for a given group of users” [8].

From this standpoint, textual CBR systems seem much better suited to incident reporting and analysis than do more generic IR or text-mining systems. Interestingly, over the years, we can chart an ever greater incorporation of task- and domain-specific knowledge into textual CBR systems for incident reporting and analysis.

Researchers at University College Dublin (UCD) have developed a number of textual CBR systems using reports from the Air Investigation Unit of Ireland. These reports contain eighteen structured fields and a free-text description. In their early work, the UCD researchers compare retrieval based solely on a similarity measure defined over twelve of the fields, with retrieval using cosine similarity on a vector representation of keywords taken from the free-text descriptions, and with the combination of the two [9]. In a small-scale experiment, the system using the structured fields performed best; the retrieval over the vector representation performed poorly and impaired the performance of the combined system. However, a similar piece of work carried out independently (not in UCD) in the case of a Medical Event Reporting System, where the free-text was more detailed, found that retrieval over the vector representation outperformed retrieval using structured fields, and the combination performed best [10]. These conflicting results perhaps emphasise the value of taking a domain-specific approach.

In terms of the characterisation given in [11], UCD’s later work [12, 13] falls at the *weakly-textual* end of the spectrum of approaches to textual CBR. Textual CBR is considered to be strongly-textual when the text is the focus of the reasoning and requires sophisticated treatment; it is considered weakly-textual when the text offers only limited reasoning support and does not require sophisticated processing [11]. By this characterisation, the Canadian TSB corpus would seem to require strongly-textual CBR. But the successful use of the structured fields reported in [9] suggests that weakly-textual CBR may suffice for the Irish corpus. UCD’s later systems made further efforts to combine similarity measures defined on the structured fields with simple vector space approaches for the free-text fields [12, 13]. Most notably, in small-scale experiments Wilson et al. obtain some success with an approach that uses the structured fields but also uses separate vectors for each section of the free-text (in their case, Witness Account, Determined Cause, and Safety Recommendation) [13].

Interestingly, the most recent Workshop on Textual Case-Based Reasoning (2006) contains several papers related to the field of incident reporting but they use somewhat more sophisticated text processing than the papers mentioned so far. In the case of anomaly reporting for the European Space Agency’s Operations Centre, for example, Wiratunga et al. describe their CAM prototype that, given a report, aims to retrieve similar reports both for detecting recurring problems and for solving new problems [14]. The CAM system uses thirteen structured fields as features and extracts further features from four sections of text. Considerable effort goes into selecting the textual features based on part-of-speech tagging, stop word removal, stemming, frequency pruning, word co-occurrence profiling and the use of feature extraction rules.

Another example from the same workshop is the paper by Bentebibel and Despres, describing the SAARA system [15]. In SAARA, cases describe road accidents and the aim is to assist users in constructing road safety proposals. The cases, although textual, appear to be highly structured, and this allows for a focus in the research on compositional and hierarchical case adaptation.

5 Conclusion

In this paper, we have presented our workshop challenge. We have given some background information about incident reporting and analysis; we have described the challenge corpus, highlighting some of the problems that it poses for the Textual CBR community; and we have reviewed research in which CBR, and especially textual CBR, has been applied to similar tasks and domains.

We hope that by encouraging multiple research groups to work on common problems, such as those posed by the corpus that we selected for our workshop challenge, we will see greater synergy between different strands of future research.

References

1. Johnson, C.: Using case-based reasoning to support the indexing and retrieval of incident reports. In Cottam, M., et al., eds.: Foresight and Precaution: Proceedings of European Safety and Reliability Conference ESREL 2000, Rotterdam, The Netherlands, Balkema (2000) 1387–1394
2. Johnson, C.: A brief overview of novel computational techniques for incident reporting. In: Procs. of AAAI Workshop on Intelligent Lessons Learnt Systems, Naval Research Laboratory; Navy Centre for Applied Research in Artificial Intelligence (2000) 20–25
3. Johnson, C.: Software support for incident reporting systems in safety-critical applications. In Koornneed, F., van der Meulen, M., eds.: Computer Safety, Reliability and Security: Procs. of 19th International Conference SAFECOMP 2000. LNCS 1943, Springer-Verlag (2000) 96–106
4. Lenz, M., Hübner, A., Kunze, M.: Textual CBR. In Lenz, M., et al., eds.: Case-Based Reasoning Technology: From Foundations to Applications. LNAI 1400. Springer-Verlag (1998) 115–137
5. Ashley, K.D., Brüninghaus, S.: Developing mapping and evaluation techniques for textual case-based reasoning. In Lenz, M., Ashley, K., eds.: Procs. of the First Workshop on Textual Case-Based Reasoning, AAAI Press (1998) 20–23
6. Weber, R.O., Ashley, K.D., Brüninghaus, S.: Textual case-based reasoning. The Knowledge Engineering Review **20** (2006) 255–260
7. Lenz, M.: Knowledge sources for textual CBR applications. In Lenz, M., Ashley, K., eds.: Procs. of the First Workshop on Textual Case-Based Reasoning, AAAI Press (1998) 24–29
8. Burke, R.D.: Surveying the opportunities for textual CBR: A position paper. In Lenz, M., Ashley, K., eds.: Procs. of the First Workshop on Textual Case-Based Reasoning, AAAI Press (1998) 13–19

9. Cassidy, D., Carthy, J., Drummond, A., Dunnion, J., Sheppard, J.: The use of data mining in the design and implementation of an incident report retrieval system. In Jones, M.H., et al., eds.: Procs. of the Systems and Information Engineering Design Symposium. (2003) 13–18
10. Tsatsoulis, C., Amthauer, H.A.: Finding clusters of similar events within clinical incident reports: a novel methodology combining case based reasoning and information retrieval. Quality and Safety in Health Care **12** (2003) 24–32
11. Wilson, D.C., Bradshaw, S.: CBR textuality. Expert Update **3** (2000) 28–37
12. Carthy, J., Wilson, D.C., Wang, R., Dunnion, J., Drummond, A.: Using T-Ret system to improve incident report retrieval. In Gelbukh, A., ed.: Procs. of Computational Linguistics and Intelligent Text Processing 2004. LNCS 2945, Springer-Verlag (2004) 468–471
13. Wilson, D.C., Carthy, J., Abbey, K., Sheppard, J., Wang, R., Dunnion, J., Drummond, A.: Textual CBR for incident report retrieval. In Kumar, V., et al., eds.: Procs. of Computational Science and its Applications 2003. LNCS 2667, Springer-Verlag (2003) 358–367
14. Wiratunga, N., Massie, S., Craw, S., Donati, A., Vicari, E.: Case based reasoning for anomaly report processing. In Wiratunga, N., Lamontagne, L., eds.: Procs. of the Second Workshop on Textual Case-Based Reasoning (TCBR — Reasoning with Text). (2006) 44–54 (Workshop Programme at the 8th European Workshop on Case-Based Reasoning, ed. Mirjam Minor).
15. Bentebibel, R., Despres, S.: Using compositional and hierarchical adaptation in the SAARA system. In Wiratunga, N., Lamontagne, L., eds.: Procs. of the Second Workshop on Textual Case-Based Reasoning (TCBR — Reasoning with Text). (2006) 98–108 (Workshop Programme at the 8th European Workshop on Case-Based Reasoning, ed. Mirjam Minor).

Summary

. . . the aircraft cartwheeled on the lake, . . . then came to rest on the bottom of the lake, . . . [the pilot] drowned.

Other Factual Information

. . . winds were estimated to be from the southwest . . . the most common direction for the wind is from the north, . . . it could not be determined whether the pilot had [satisfied Canadian Aviation Recency Requirements] . . . there was nothing to indicate that the pilot's performance was degraded by physiological factors . . . He did not sustain any immobilizing or incapacitating injuries during the crash . . . By design, when the door is locked from the inside, it cannot be opened from the outside . . . a previous TSB investigation and report (A04W0114) . . . suggested that [Transport Canada] consider additional methods to facilitate rapid emergency exits from seaplanes in the event that the cabin becomes submerged.

Analysis

. . . In this occurrence, although this pilot's flying activity exceeded the minimum requirements. . . , it is unlikely that critical flight skills and procedures were practised to ensure proficiency. The current recency requirements in Canada allow pilots to go for extended periods without any retraining in critical flight skills, presenting the risk that pilots will not be prepared to deal with unusual or critical flight situations when they arise. . .

If the aircraft doors are locked as required during flight, the design of the door lock mechanism does not permit access to the cabin from the outside by using the exterior door handles. . .

Findings as to Causes and Contributing Factors

1. For undetermined reasons, the aircraft cartwheeled after contacting the water and came to rest in an inverted position.
2. The pilot was unable to exit the aircraft and he drowned.

Findings as to Risk

1. The pilot had not flown a training flight with an instructor for more than four years. This likely resulted in a degradation of his skills and decision-making processes.
2. The current recency requirements in Canada allow pilots to go for extended periods without retraining on critical flight skills, presenting a risk that pilots will be ill-prepared to deal with unusual or critical flight situations when they arise.
3. The design of the door lock mechanism prevents opening of the doors from the outside when locked from the inside. This same design is currently being used in all of Cessna's new production single-engine aircraft.
4. The exterior door handles are not easily discernable when the handles are closed and the visibility is poor.
5. . . .

Other Finding

It could not be determined whether the pilot had complied with the recency requirements . . . of the Canadian Aviation Regulations.

Safety Action Taken

Transport Canada has undertaken a Risk Assessment. . . to identify the extant risks related to egress from submerged seaplanes and to identify the most effective means of mitigating those risks.

Safety Concern

The following safety concern is similar to the one published in report A04W0114, referenced previously.

Based on historical data, . . . the TSB is concerned that seaplane occupants may not be adequately prepared to escape the aircraft after it becomes submerged. . .

Fig. 1. An excerpt from Air Investigation Report Number A05O0147