1D-4

ヒヤリハット事例収集データベースにおける特徴抽出の実験評価

二反田慎* 大塚雄市** 牧之内顕文*** 金子邦彦****

九州大学大学院システム生命科学府システム生命科学専攻* 工学府機械科学専攻**

久留米工業大学工学部情報ネットワーク工学科*** 九州大学大学院システム情報科学研究院知能システム学部門**

1．はじめに

ヒヤリハット事例とは，インシデント(軽微な不具合)の一種で，実際の損害には至らなかったが何らかの失敗が発生した事例である．インシデント報告[1]は安全管理において現場での実データを取得しうる有用なツールである．インシデント報告はその集計，分析結果が各種安全対策の基礎資料として活用される．医療安全管理において，報告システムは一般的に導入され，電子化[2]の動きも進められている．

インシデント報告は下記内容で構成される[3].

1) 発生状況（選択項目）:発生日時や発生場所および関係する医療従事者や被害者の情報.
2) 発生場面とその内容（選択項目）:インシデントの被害内容を分類する情報
3) 要因（選択項目）:インシデントの発生要因を表現する項目
4) 事実経過〔自由記述〕:選択項目で不十分な内容を記述するための，背景，発生経過，要因および事後処置を文章で記入する項目

これらの内容を管理者が読解して，対策のための分析を実践する．また，データマイニング等の活用により，発生背景の探索などに活用できる可能性が指摘されている[4].

インシデント報告はいわゆるコーパス（専門家によって作成されカテゴリ分類されたテキストデータ）とは異なる点を有する．それは，インシデントに関与した作業者が直接データを入力する点である．それにより，選択項目と事実経過文の不整合の問題が生じる．一例を挙げれば，要因に薬剤の間違いが選択されているにも関わらず，事実経過には患者の転倒インシデントの内容が書かれているという問題が実際に発生している．この不整合の問題は報告データの信頼性を損壊する重要な問題である．しかしながらインシデント報告は匿名報告のため，報告者に照会しデータを修正することは原理的に不可能である．

インシデントを分析して対策を講じるに際し，最も重視されるのは対策の着目点となる要因である．従って，要因とその事実経過に不整合が存在することは分析上の障害となる．そこで，著者らは過去に蓄積されたデータを活用して，新規に入力された報告の要因が事実経過に合致しているかを判定し，管理者が整合性を判断するための「要因分類確認支援システム」（2 章で論じる）を提唱している．ただし，その実装に当たっては，インシデント報告の特徴に応じた課題を実践的に抽出し，情報技術手法を適用するための指針を明確にする必要がある．

そこで本研究ではシステムの実現化に向け，事実経過中に存在する単語の出現傾向を調べることで，適切な要因が選択されているかを判定するための提案を行い，その実験と評価について述べる．

2．要因分類確認支援システム

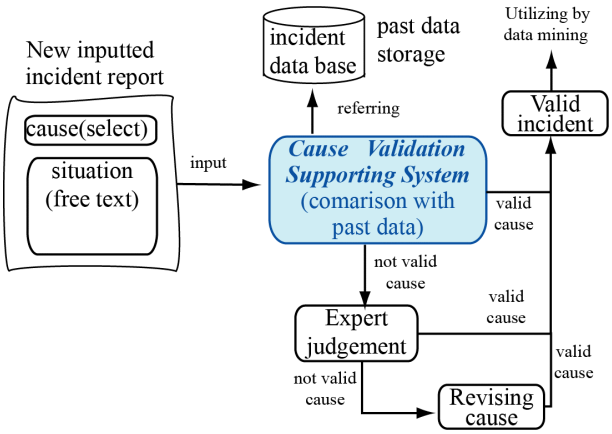

図 1 Cause validation supporting system

図 1に要因分類確認支援システムの構想を示す．過去蓄積されたデータを基に，新たに入力されたインシデント報告の要因と事実経過が整合しているかを判定するシステムを構築する．そして要因が整合しているか，不適合なら他の要因を推奨することを出力できるようにする．完全な信頼性を保障するのは困難であるため，管理者等の専門家に最終的に判断してもらい，妥当

Experiment and Evaluation of Feature Extraction for Hiyarihatto Report Collection Database
*Makoto Nitanda, Graduate School of Systems Life Science, Kyushu Univ.
**Yuichi Otsuka, Department of Mechanical Engineering Science Graduate School of Engieenring, Kyushu Univ.
***Akifumi Makinouchi, Kurume Institute. Tech.
****Kunihiko Kaneko, Department of Intelligent Systems,
Graduate School of Information Science and Electrical Engineering, Kyushu Univ.

な要因を最終的に決定する．これにより，報告データの信頼性向上および管理者の作業効率化に資することが期待できる．

３．要因合致判定に向けて

過去に蓄積されたデータの中から，ある要因 A が選択されているレポートに着目し，それらに記述されている事実経過の集合を収集したものを対象に，以下のような作業を行う(図２)．

1. キーワード抽出ソフト EKWords[5]を用いて単語の抽出を行う．但し，１文字の単語及びひらがなの単語は抽出されない為，原文から EKWords による結果を取り除き，残った単語を調査することで単語抽出結果に随時追加する．大文字と小文字，半角と全角の違いによって異なる単語として区別されてしまう為，抽出前に小文字は大文字に，半角は全角に変換しておく．扱う単語の種類は名詞のみに注目する事とし，その他の品詞は除外する．
2. 抽出した単語の中から同意語となる組み合わせを１つの代表語にまとめる．それ自身の単語以外に同意語が存在しない場合，その単語名をそのまま代表語とする．
3. 各代表語をカテゴリごとに分ける．カテゴリは一般的に事実経過文に記述される「時間」「場所」「人」「対象」「行動」「処置」の６つの項目を定義している．
4. カテゴリごとに存在する代表語を KJ 法によってサブカテゴリに分ける．サブカテゴリ名は代表語の集合を基に命名される．

このようにして得られた代表語とサブカテゴリを用いて，要因 A に属する事実経過にどのような傾向があるかを評価する．

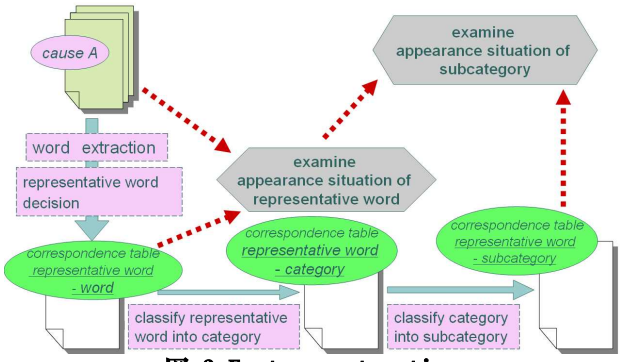

図 2 Feature extraction

４．過去事例による特徴抽出の実験・評価

計51項目ある要因のうち，ここでは「配薬時の薬剤確認不足・未実施」を対象にして実験を行った．報告件数は113件あり，それぞれに含まれる事実経過全てを対象にして単語抽出を行った結果，単語は全811種類存在し，単語を代表語に対応させた結果，全701種類となった．これらの代表語を6種のカテゴリに分け，そこからサブカテゴリに分割した結果，「時間帯」「日付」「曜日」「期間・頻度」「行動中」「行動前後」「室内設備」「院内施設」「患者部位」「院外施設」「医療従事者」「患者関係者」「患者意識」「患者病状」「患者バイタル」「薬剤名」「器具名」「用法名」「用量名」「患者行動」「配薬・与薬」「確認作業」「失敗検知」「確認用具」「処置名」「状態変化」「患者対応」「事後処置」の28種類のサブカテゴリが得られた．

各サブカテゴリに属する代表語の出現頻度に注目した結果，最も高い頻度から順に「薬剤名」「医療従事者」「容量名」「時間帯」「処置名」「配薬・予薬」「失敗検知」となった．但し「医療従事者」と「時間帯」はこの要因に特有ではなく，判定の際に考慮する必要性はない．それ以外の5つの項目に関しては要因に特有の為，要因合致の為に必須であると判断できる．一方，「用法名」や「確認作業」についても要因特有の代表語から形成されてはいるが，過去113件においては頻度が高くなく，要因合致判断の材料にどう含めるかが問題である．

この他に，サブカテゴリ・代表語の相関関係やストップワードについての検討も必要である．

５．まとめ

要因分類確認支援システム設計に向け，過去事例を対象に代表語とサブカテゴリの設定及びそれらの出現頻度について調べてきた．今後はその頻度を参考にして，要因合致のためにはどの代表語やサブカテゴリが必要かを決定し，新規の事実経過を対象にした，要因合致判定を支援するシステム設計を目指す．

参考文献

[1] Kohn, L. T., Corrigan, J. M., Donaldson, M. S., *TO ERR IS HUMAN: Building a Safer Health System*, National Academy Press, (2000)

[2] Otsuka, Y., Makinouchi,A., Noguchi, H., *Proc. of CJKMI 06*, p.133-135, (2006).

[3] Thomas, A.N., Pilkington, C.E., Greer, R., *Journal of Evaluation in Clinical Practice*, 9(1), 59-68, (2003).

[4] Bates, D.W., Evans, R.S., *et al.*, *Journal of the American Medical Informatics Association*, 10(2), 115-128, (2003).

[5] DJSOFT EKWords：http://www.djsoft.co.jp/products/ekwords.html