

情報処理学会ソフトウェア工学研究会要求工学ワークショップ
平成28年5月19日(木)～21日(土) 於 高崎

要求記述支援のための 事例分析

茨城工業高等専門学校
滝沢 陽三

背景・経緯

- 当初の研究目的・内容
(1990年代後半～2000年代前半)
 - 要求者自身による要求仕様化の支援手法開発
 - 要求者の意図を開発者に伝えることが主目的
 - 自然言語で仕様化することを想定
 - 自然言語は事前学習が必要ない記法
 - 限定的な自然言語処理技術による記述の洗練
 - 記述文章の単文化処理に基づく洗練が中心
 - 構文解析処理と単語辞書検索による情報導出
 - 用語出現に基づくドメイン固有の記述の追加支援
 - 支援システムの構成と運用
 - ドメイン（要求分野）ごとに定義された単語辞書構築
 - 開発者による支援と要求者の学習を組み入れ

背景・経緯

- その後の経緯
(2010年代前半)
 - 要求工学の観点でまとめ直し
 - 要求仕様化ではなく（顧客自身による）要求定義
 - 事例に基づく文法定義による記述の形式化
 - 単文化処理の改良
 - 形態素解析システム等による事例解析
 - 品詞情報を中心とした既存記述の分析

前回までの課題

- 事例記述の収集・分析
 - 「不適切な記述→失敗した開発」以外
 - 不適切な記述→成功した開発
 - 適切な記述→失敗した開発
 - 当面は「適切な記述」と判断された文章を中心に事例解析
- 「単文化」の有用性検討
 - 単文化することの意義
 - 元は「制限言語」による要求の仕様化を想定
 - 顧客による要求記述を想定し、別の観点での記述洗練を想定

用語の確認（定義）

- 要求者≡顧客
 - 利害関係者（ステークホルダ）の一部
- 要求者は提案依頼書（RFP）までは支援なしに作成するものとする。
 - 必要な項目等の枠組みは存在
- 要求記述≡機能要求仕様書
 - RFPは機能要求の提案依頼

研究目的の再定義

- 顧客自身が、提案依頼書に相当する記述から、機能要求を自然言語で定義する、手法および支援システムの開発
- 支援システムの対象
 - RFPの洗練・詳細化、機能要求の導出
- 開発者のドメイン経験に依存していた、開発者による支援の部分をもシステム化
 - ドメイン内部の知識・表現体系が必要
(これまでは単文化と記述追加支援を想定)

事例分析

- 記述文書中の品詞分布の傾向調査
 - 元は単文化のための文法定義改良に必要なものとして実施
 - 今回は単文化以外の記述洗練の手がかりを得るための分析として実施
- MeCabによるSubtypeまでの品詞情報を用いて記述内の割合を調査
 - 概ね「一割以上=多い」とみなす。
- 「要求」記述にこだわらず、様々な分野の記述を解析
 - システム仕様書、提案依頼書（RFP）、携帯アプリレビュー、文学作品 - 8記述、100語～6000語／記述

事例分析：分析結果・考察

- 形式的な記述（仕様書等）ほど「名詞＋する／できる」「名詞＋を行う」を使用。
 - 体言止めを「名詞（＋する）」等と解釈するならば、その傾向は更に強い。
- 日本語文化固有の意味限定用法？
 - 裏付を調査中。
 - 英語では「名詞は全て動詞となり得る」。
 - 要求記述における多様な「モード」の存在の可能性
 - 「ドメイン」だけでなく「モード」の観点でも記述チェックすることで、あいまいさ解消（やノイズ除去）の仕組みを構築可能？

用語の確認（定義）

- モード
 - フランス語のmode、英語のmood
 - 本来の意味は「流行」に近い。
 - 言語学においては（今回の場合は）記述者の意図を表現するための記述方法を指す。
 - 客観的な意味内容は変わらないが、記述者の意図によって追加される動詞・述語表現。
- 【例】客観的な意味内容「書く」
 - モードによる追加：「書け」「書きます」「書こう」
- ここでは、要求定義活動、または、ドメインごとに存在する表現体系と定義

事例分析：修正支援例

- 「トークが過去ログから表示される現象がまだ直ってない」
 - 修正／改善／回復
 - 以前の状態を考慮して選択
- 「2つ以上のボタンを持つマウスと、キーボードを有すること」
 - 含有／保有／携行／有（する）／保持／維持／残存／携帯／所持／所蔵／所有
 - より限定された意味による他、主語との共起率を考慮して選択（変更不要の判定を含む）
- 支援ツール実装の可能性
 - 選択候補は「類語辞典」情報から生成可能？
 - 用語の統一の観点も含む？（「先生／教員」等）

事例分析：分析の発展

- 品詞割合の平均情報量の比較
 - 品詞（Subtype）39種類を対象
 - 結果：形式的な記述ほど平均情報量が低い。
 - 出現した品詞の種類の数に関わらず
- 考察
 - 形式化に伴い、一般自立動詞からサ変接続名詞への使用の偏りが原因と解釈すれば、先の分析の解釈と矛盾しない。
 - 単語数が少ない記述の性質を知る手段として有効である可能性あり

用語の確認

- 平均情報量（エントロピー）
 - ある事象系の情報の無秩序さを示す尺度
- 記述変化に伴う平均情報量の変化の例
 - 上がる（最大エントロピーに向かう）場合
 - より強い暗号化が行われている。
 - 下がる（冗長度が低下する）場合
 - より強い圧縮が行われている。
- 特定の単語の出現頻度が高ければ平均情報量は下がり、品詞割合の平均情報量も下がる。

事例分析：分析の発展

- 形式的な記述の「改善」前後の品詞分布や平均情報量の変化の調査
 - RFPサンプル2記述で比較
 - 結果（改善前→改善後）
 - いずれも名詞約3%減・動詞1～3%増
 - 平均情報量は増加
 - 考察
 - 改善⇨形式化と解釈すると先の分析・解釈と矛盾
 - 「非形式的な記述の追加」に伴う変化と解釈？

事例分析：（余談）

- 要求記述に関する事例の収集は困難
 - 内容の機密性よりもむしろ「要求が明確になっていなかった」（⇨記述自体が存在しない）こと自体の機密性が大きい可能性
 - 既存の汎用ソフトウェア／システムパッケージの導入のみが「要求」として想定される（結果、導入のみで稼働できない）ケースも。
 - 要求記述支援以前に、「要求を明確にする」必要性を高めるための支援の仕組みが必要？
 - 「要求が明確になっていなかった場合の具体的な失敗／成功例」の収集・分析の必要性

今後の予定

- 更なる事例分析・考察
 - 別の観点からの分析と考察
 - 文型の傾向調査等
 - 要求分析・設計への発展の追跡
 - 分析ツールの整理・発展
- 手法の整理と支援ツールの開発
 - ドメイン情報に基づく記述の追加支援と、「モード」情報に基づく記述の洗練支援
 - 事例分析（ツール）に基づく上記情報のデータベース化

（事例分析（2））

- 文型の傾向調査
 - 単文、複文、重文の割合
 - 体言止めの箇条書きも文章として調査
 - 単文単位における格助詞数の割合
 - 単文ごとの複雑さの調査
- 今回の目的に沿った構文木生成ツールの開発