

顧客要件に合致するパッケージシステムの ノウハウ事例抽出支援システムの開発

Development of Design Know-how Cases Retrieval Support System on a Package System which corresponds to Customer Requirements

中村覚¹ 稗方和夫¹ 岡田伊策^{1,2} 齋藤稔² 笈田佳彰² 増田幸司³ 阿部雅昭³ 堀口隆三³
Satoru Nakamura¹, Kazuo Hiekata¹, Isaac Okada^{1,2}, Minoru Saito², Yoshiaki Oida², Koji Masuda³,
Masaaki Abe³, and Ryuzo Horiguchi³

¹ 東京大学

¹The University of Tokyo

² 富士通株式会社

² FUJITSU LIMITED.

³ 株式会社富士通システムズ・ウエスト

³Fujitsu Systems West Limited

Abstract:

パッケージによるシステム導入はコストや保守性の観点から、実績ノウハウの再利用が重要である。パッケージ導入ベンダーは、顧客の要件に合致するように既存ノウハウを詳細かつ最大限に再利用し、システム設計・構築することが求められる。しかしノウハウ事例の探索・選別プロセスに要する時間および探索結果の網羅性は、探索者の能力・暗黙知に依存する。

本研究ではパッケージによるシステム導入時の設計におけるコスト削減や構築時間の短縮を目的とし、実績ノウハウ情報を構造化し、要件に合致するノウハウ事例の抽出を支援するシステムを開発した。そのシステムをERPパッケージに適用することで、有用性の検証を行った。

1. 緒言

1.1. パッケージシステム

パッケージシステム・ソフトウェア（以下、パッケージ）は、業種・規模を問わず多数の企業・組織で共通する業務（財務会計や給与計算など）や、業種内（製造業、卸売業など）で共通するビジネスモデルについて、多数の企業で汎用的に利用できるソフトウェアである。本研究では特に、一般化された業務フロー、処理、標準機能を持ち、主にそれらの機能の組み合わせとパラメータ設定を行うことで、業務を構築するようなパッケージを対象とする。

したがってパッケージ導入ベンダー企業は、RFP等の顧客要件から抽出した業務要件について、パッケージが提供する標準機能群を選択・組み合わせることでシステム設計を行う。ここで各用語の定義を含めて、「業務」「業務機能」「標準機能」の関係を図1に示す。例えば「債務支払/消込」という「業務」

は、「現金支払」や「会計伝票相殺」という「業務機能」によって構築され、各業務機能はパッケージが提供する「仕入先明細」や「支払消込」等の「標準機能」群の組み合わせによって構成される。

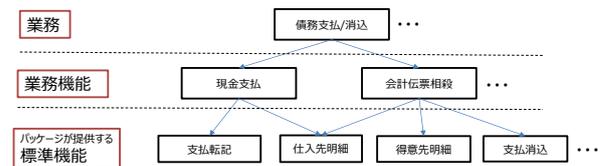


図1: 業務、業務機能、標準機能の関係例

1.2. 設計ノウハウ再利用の必要性和課題

このシステム設計において、パッケージが提供する膨大な標準機能と、それらの組み合わせによって構築される業務機能および業務内容を網羅的に理解しておくことが必要である。したがって、システム設計に携わる者はマニュアル等の文書を学習してパッケージに関する知識を補強する。しかし、文書から膨大な情報をすべて知識として蓄えることは実質

困難であり、システム設計の実経験を通じて獲得していく他、過去のシステム構築（以下、過去プロジェクト）で作成された文書に記述された設計履歴等の既存ノウハウを探索・参照して設計を行う。特に一般的なパッケージ導入において作成される業務機能の一覧表と業務フロー図は、システム設計におけるノウハウを豊富に含む文書である。業務機能の一覧表は過去プロジェクトにおいて構築した業務内容や業務機能と、それらの機能の構築に必要な標準機能群を一覧形式にした文書である。また業務フローはそれらの関係性をダイアグラムで表現した文書である。しかしこれらの既存ノウハウは個々の文書に分散して記述されており、また探索者が文書内容を都度確認し、取捨選択して必要な情報を探索する必要がある。先に述べたようにパッケージが提供する標準機能、およびその組み合わせによって構築される業務機能や業務内容は多岐に渡るため、人手による探索プロセスには多大な労力を要し、また必要な情報の抜け漏れが生じるリスクが存在する。

1.3.研究目的

本研究では 1.2 で述べた課題の解決を目的とし、機能要件に対応する業務や機能、およびそれらの構築に必要な標準機能の組み合わせ等、システム設計に必要な情報を自動的に提示するシステムを開発する。具体的には過去プロジェクトで作成された各種文書に記述された設計に関する既存ノウハウを機械可読な形で記述し、与えられた機能要件に合致する標準機能を機械的に抽出する手法を提案する。またある ERP パッケージに本手法を適用し、提案手法の有用性の評価を行う。

2. 関連研究

あるドメイン内の概念とそれら概念間の関係に関する知識を形式的に表現する方法としてオントロジーがある。古崎ら[1]は「法造」というオントロジー構築・利用環境を用いて、石油精製プラントに関するオントロジーを構築した。また構築したオントロジーを用いたアプリケーション開発も行っている。本研究ではパッケージの標準機能に関する情報を RDF によって機械可読な形で記述し、システム設計に必要な知識や各種文書等の既存ノウハウをメタデータとして構造化し、情報の関連付けを行った。

また那須川[2]はコールセンターにおける顧客とのやり取りの記録をテキストマイニングし、抽出した知識に基づくデータ分析を行った。彼は構文解析を用いて、自由記述形式の文章から得られる主語や述語、目的語といった構造を抽出し、文書間の類似度や頻出する内容の抽出を行っている。本研究でも

構文解析を利用し、文書に記述された文から述語や目的語を抽出することで、機能に関する情報の構造化を行った。

3. 提案手法

本研究ではパッケージが提供する標準機能に関する情報を構造化し、入力された機能要件に合致する機能群を機械的に抽出することを目的とする。これを実現するため、本章では「標準機能情報の構造化」「標準機能と業務の関連付け」「機能要件に合致する標準機能の抽出」の順に説明する。

3.1.標準機能情報の構造化

パッケージが提供する各標準機能の意味を計算機が理解可能とするため、機能に関する情報を構造化する。パッケージが提供する標準機能には、名称とその機能内容に関する説明文が付与されている場合を想定する。例えば名称が「仕入先マスタ登録」である標準機能には「仕入先マスタを登録する」という説明文が付与されている。このように機能の説明は「○○（対象）に対して××（処理）を行う」という形で記述される。この説明文を構文解析し、各標準機能を「対象」と「処理」に分割して記述する。なお、構文解析器には「Cabocha」を利用した。

本研究では RDF を用いて構造化した標準機能の情報を記述する。各標準機能にそれを一意に定める URI を付与し、抽出した「対象」および「処理」に関する情報をそのメタデータとして付与する。具体的には「対象」を表すメタデータフィールド（`ex:object`）と「処理」を表すフィールド（`ex:process`）を定義し、抽出した値をそのバリューとして付与する。またそれらの階層関係も定義した。例えば「仕入先マスタ」はクラス「マスタ」のインスタンスであり、クラス「マスタ」はクラス「対象」の一部である、という記述を行う。さらに RDF スキーマの一つである `rdfs:range` を用いて、「対象」クラスと「処理」クラスの間を定義する。例えば「処理」クラスの下位概念である「マスタ管理」の目的語は、対象クラス「マスタ」のインスタンスである、という関係を定義した。この例を図 2 に示す。

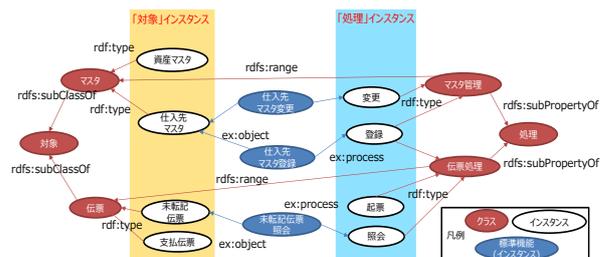


図 2: 標準機能に関する情報の構造化

3.2.標準機能と業務の関連付け

次に標準機能と業務の関連付けを行う。1章で述べたように、業務は業務機能の組み合わせで実現され、各業務機能は複数の標準機能の組み合わせによって構築される。これらの情報は過去プロジェクトで作成される業務機能の一覧表や業務フロー図等の文書に記述されている。したがってこれらの文書内に記述された標準機能や業務機能、および業務に関する情報を抽出し、3.1章で定義した標準機能のメタデータとして記述する。さらにそれぞれ抽出した情報の階層関係も合わせて記述する。たとえば標準機能「仕入先マスタ登録」は業務機能「仕入先マスタ」を構成する一要素である。よって標準機能の「機能」フィールド (ex:function) のバリューとして、クラス「業務機能」のインスタンスである「仕入先マスタ」を記述する。同様に業務機能と業務の関係についても、業務に関するメタデータフィールド (ex:operation) を用いて関連づけた。

3.3.機能要件に対する標準機能群の抽出

機能要件に対して、それに合致する標準機能群を抽出する流れを図3に示す。ここでは「小切手を印刷する」という機能要件を例として説明を行う。

まず機能要件を「対象」と「処理」に分割する (図中①)。機能要件を形態素解析し、各形態素を「対象」か「処理」を判別する。この例では対象が「小切手」、処理が「印刷」として構造化される。

そして構造化した機能要件に合致する標準機能を抽出する (図中①、以下「処理①」とする)。RDF クエリ言語である SPARQL を用いて、「対象フィールド (ex:object)」に「小切手」、「処理フィールド (ex:process)」に「印刷」という文字列を含む標準機能

能を抽出する。その結果、標準機能「小切手再印刷」等が抽出される。

次に抽出した標準機能が含まれる業務機能を抽出し、その業務機能に関連する標準機能群の抽出を行う (図中②、以下「処理②」とする)。先に抽出した「小切手再印刷」は業務機能「小切手管理」がメタデータとして付与されているため、この関係を持つ他の標準機能を抽出する。この結果、「小切手無効化」や「小切手番号」等の標準機能が抽出される。

さらに抽出した業務機能が含まれる業務を抽出し、その業務に含まれる他の業務機能群や、その業務機能を構築するための標準機能群を抽出する (図中③、以下「処理③」とする)。業務機能「小切手管理」には、業務「債務管理」の一つであるというメタデータが付与されている。そのため、業務「債務管理」に含まれる業務機能、その業務機能に含まれる標準機能を抽出する。その結果、業務機能「債務伝票計上」を構成する標準機能「未転記伝票変更」や、業務機能「仕入先マスタ」を構成する標準機能「仕入先マスタ登録」等が抽出される。

最後に関連する標準機能群を抽出する (図中④、以下「処理④」とする)。例えば標準機能「未転記伝票変更」は「対象」が「未転記伝票」であり、「処理」が「変更」として構造化される。そして「伝票」に関する「処理」には「登録」や「照会」等が存在するため、対象が「未転記伝票」であり、処理として「登録」や「照会」を行う標準機能群を抽出する。その結果、標準機能「未転記伝票照会」等が抽出される。

ここまでの処理で抽出された全ての標準機能群を、入力された機能要件に対する業務機能の一覧として出力する。

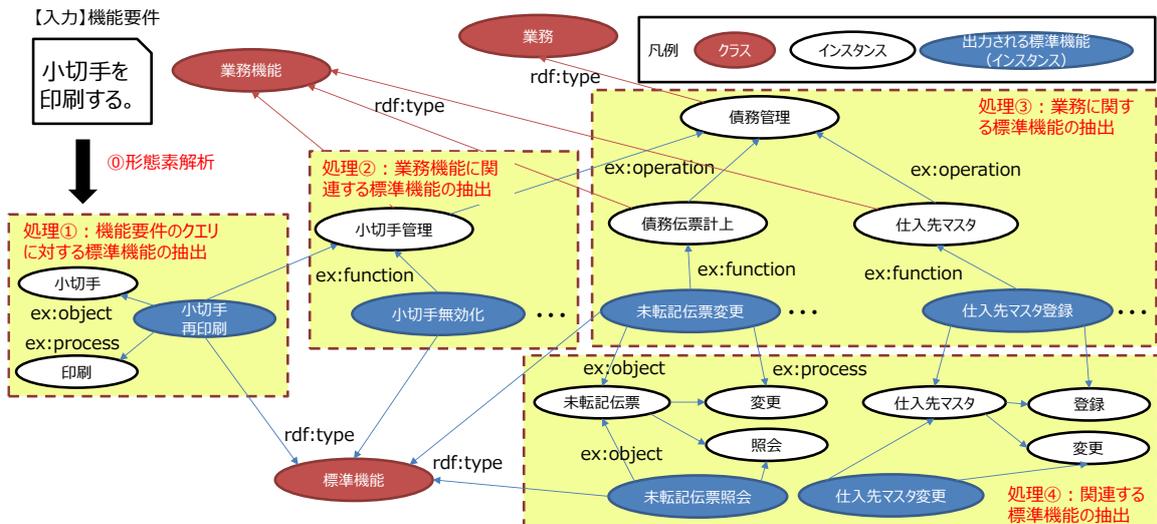


図3: 機能要件に対する標準機能群の抽出の流れ

4. ケーススタディ

4.1. ケーススタディ概要

ある ERP パッケージを対象とし、提案手法に基づいて開発したシステム上でケーススタディを行った。提案手法の抽出精度を計測するために、「小切手を印刷する」という機能要件をサンプルとして用いた。この機能要件に対する業務フロー図を図 4 に示す。小切手の印刷や決済等の小切手管理に関する機能に加え、債務伝票や支払処理等の機能群が必要となる。

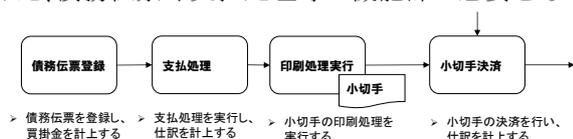


図 4: 「小切手印刷」に関する業務フロー図の例

またケーススタディを行うにあたり、対象とする ERP パッケージの導入・設計を行った過去プロジェクトで作成された文書を用いて、標準機能群の構造化を行う。ERP パッケージは「管理会計」や「販売管理」等の業務別にシステム設計が行われるが、本ケーススタディでは業務機能「小切手管理」を含む「財務会計」に関する文書を用い、計 560 個の標準機能に関する情報の構造化を行った。

さらに機能要件「小切手印刷」に必要な標準機能群の一覧を予め人手によって作成し、システムの出力結果における適合率、再現率を用いて評価した。

4.2. 結果と考察

システムの出力結果を表 1 に示す。この表は、3.3 章で説明した処理毎の抽出精度を示し、最終的な出力結果は適合率 43%、再現率 89%であった。この結果に対して、適合率、再現率の観点から考察を行う。

表 1: 抽出結果と精度

| | 正解数 | 検索結果数 | 検索結果に含まれる正解数 | 適合率 | 再現率 |
|--------|-----|-------|--------------|------|------|
| 処理①のみ | 18 | 1 | 1 | 1.00 | 0.06 |
| 処理①② | 18 | 5 | 5 | 1.00 | 0.28 |
| 処理①②③ | 18 | 37 | 16 | 0.43 | 0.89 |
| 処理①②③④ | 18 | 37 | 16 | 0.43 | 0.89 |

再現率に関しては、89%と高い値を示した。これは入力された顧客要件に対して、必要な標準機能群を網羅的に抽出できた点で評価できる。特に標準機能と業務の関連性に基づく「処理③」によって大きく再現率が向上した。これは業務機能の一覧表や業務フロー等の文書に記述された過去の設計ノウハウについて、各標準機能のメタデータとして関連付けた手法 (3.2 章参照) の有用性を示している。一方、再現率が 100%に達しない理由は、人手によって作成した正解データにおいて、対象とした過去プロジェクトでは使用されなかった 2 種類の標準機能 (小

切手の再連番を行うための機能等) が含まれていたためである。今後は標準機能に関する情報の構造化・関連付けにおいて、他の過去プロジェクトの文書を対象に含めることで、再現率の向上を目指す。

適合率に関しては、再現率に比べて低い値を示す結果となった。具体的には処理③によって、業務「債務管理」を構成する業務機能「借入金管理」に関する標準機能群が誤って抽出されたためである。しかし計算機による抽出結果を人手で取捨・選択することを前提としたならば、膨大な標準機能群から機能要件に合致するものを選択する作業に比べ、計算機が比較的網羅的に標準機能群を抽出し、それらから必要のない機能を削除する作業のほうが低コストと考えられるため、妥当な結果であると言える。一方、この適合率の向上、または人手による取捨選択時の労力低減を支援する方法として、標準機能の使用頻度などの利用履歴に基づく情報をメタデータとして付与し、利用者に提示すること等が考えられる。

5. 結言

本研究では、過去のパッケージ導入で作成された文書等から設計ノウハウを抽出し、それらをパッケージが提供する標準機能群に紐づけて構造化することにより、設計プロセスを支援するシステムを開発した。具体的には標準機能に関する情報を構造化し、それらの標準機能群によって構成される業務機能や業務内容を過去プロジェクトの文書から抽出・紐づけることで、機能要件に合致する標準機能群を機械的に抽出する手法を提案した。ある ERP パッケージを対象としたケーススタディでは、「小切手を印刷する」という機能要件に対して、適合率 43%、再現率 89%の精度の出力結果を得た。特に標準機能と、それらの組み合わせによって構築される業務内容を機械可読な形で記述することで、高い再現率を得ることができた。本研究では一つの機能要件に対する評価に留まっているため、今後は多様な機能要件に対する提案手法の有用性の評価を行う。さらに開発したシステムを用いて、実際のパッケージ導入プロセスを模擬した被験者実験を行い、設計プロセスにおける時間短縮や質の向上への影響について評価していく。

参考文献

- [1] 古崎晃司, et al. "オントロジー構築・利用環境「法造」の開発と利用.", 人工知能学会論文誌, Vol.7, No.4, pp.407-419, 2002.
- [2] 那須川哲哉. "コールセンターにおけるテキストマイニング (< 特集>「テキストマイニング」).", 人工知能学会誌, Vol. 16, No.2, pp.219-225, 2001.