

業務ログから得られた業務フローの標準化レベルと 業務フローの正規化

Standardization level of business process flow based on the business log and
normalization as removing noise of business process

笹川 文義† 川村 旭‡ 金井 剛†
Fumiyoshi Sasagawa Akira Kawamura Tsuyoshi Kanai

1. 背景と論文の概要

企業の各種の業務への IT システムの導入の加速度的な拡大と、大企業における内部統制の確立の必要性から IT システムのログが企業の IT システムに蓄積されている。基幹系システムだけでなく生産系システムにおいても製品管理の観点から IT システムの導入とログの管理が進展している。このようにして IT システムに蓄積されたログに残された証跡に基づいて、実際に行なわれた業務を有る程度推測することができる。すなわちログに残された業務（イベント）の名称（イベント名）とそれが発生した時間を示すタイムスタンプ及び業務と別の業務の関連を媒介するキー（ID）によって一連の業務の流れを推測できる。こうした推測によって得られた個々の業務の流れをまとめ上げることによって有向グラフが得られる。この有効グラフを業務フローと呼ぶ[1]。

IT システムに蓄積された証跡から業務フローを推測して出力・分析を実施するアプリケーション、サービスは大学での研究ツール (Prom) [2]としてまた企業では顧客の業務分析のツール (IDS Scheer 社の PPM) [3]またはサービスとして (Fujitsu Interstage Analytics) [4]としていくつか開発されている。これらのツールの主な目的の一つは業務の可視化である。これに基づいて業務フローから業務の問題点を洗い出し業務の改善の基礎データを得ることが重要であると考えられている。一般的には業務の標準化が

遅れていればそれだけ業務の改善が必要であると考えられている。すなわち、業務フローの標準化の程度を計測すれば良いことになる。そこで本論文では事例に基づいて業務フローから標準化のレベルを求める方法を報告する。さらに実際の業務データを扱って業務フローを推測する場合の問題点の一つは実際の業務の流れや IT システムログの内容について十分な検討無しに生成した業務フロー図には、実際の業務自体の複雑さに由来しない見かけ上の複雑さが紛れ込む場合があることである。見かけ上の複雑さが発生する原因は種々のものが考えられるが、ここでは主に統計的（確率論的）な手法で検出可能な見かけ上の複雑さの除去方法について提案する。

2. 業務フローを推定する方法の概要

企業の業務フローを推定する方法は、担当者へのインタビューに基づく方法、システムのマニュアルや仕様書等から業務フローを推定する方法が知られているが、ここで説明するのは企業のデータベースに蓄積されている業務ログに基づいた方法である。ここでは実際のデータではなく説明用に作成した単純化したデータを使用して説明する（表1参照）。

表 1 企業の事業本部のイベントデータテーブル（一部）の例

処理 ID	テーブル名	日時	見積り金額	注文金額	円換算額	契約金額
0001	購入依頼	2010/5/7 13:24:42	---	---	---	---
0001	注文	2010/5/8 17:12:11	---	28615	28615	---
0001	見積り回答	2010/5/10 8:27:54	28615	---	28615	28615
0001	入荷	2010/5/11 20:22:17	---	---	---	---
0001	検収	2010/5/14 17:37:25	28615	28615	28615	28615
0002	購入依頼	2010/7/11 21:01:32	---	---	---	---
0002	注文	2010/7/14 14:40:30	---	39576	39576	---
0002	見積り回答	2010/7/16 14:21:24	39576	---	39576	39576
0002	入荷	2010/7/17 7:14:58	---	---	---	---
0002	検収	2010/7/18 13:21:29	39576	39576	39576	39576

表1は更新ログファイルのデータ構造例を示している。表1の更新ログファイルにはタイムスタンプを示す日時の項目、テーブル名を示す項目、実行者のIDを示す項目、処

† 株式会社富士通研究所, FUJITSU LABORATORIES LTD.

‡ 富士通株式会社, FUJITSU LTD.

理IDを示す項目、その他のフィールド、見積り金額、注文金額、円換算合価、契約金額を示す項目が設けられている。もちろん、実際の更新ログファイルは表1のように7項目で終わることは稀である。ここでテーブル名をイベントの名称と解釈するとテーブル名の項目と日時の項目の対応関係から日時項目をイベントの発生時刻と解釈できる。イベントとイベントを結びつける証跡としては処理IDの項目を選択すればイベントを日時順に並べることでアクティビティの列を得ることが出来る(表2を参照)。このアクティビティの列をフローインスタンスと呼ぶ。イベントの発生日時が異なるが、イベントの発生順が同じフローインスタンスの集まりをフロー種別と呼び、イベントの順序付けられた矢印で遷移を表す(表2のフロー種別を参照)。

表2のシーケンスはアクティビティを発生日時順に並べた

表2 イベントデータテーブルから得られたアクティビティのシーケンスの例

フロー種別ID	フロー種別	件数	平均実行時間(h)	最大時間(h)	最小時間(h)
1	購入依頼→注文→見積り回答→入荷→検収	100	188.1	255.7	101.4
2	購入依頼→入荷→検収	50	102.7	155.7	36
3	購入依頼→入荷(2)→検収	40	151.9	214.1	63.2

もので、件数はシーケンスのインスタンスの件数を示している。平均実行時間は対応するシーケンスの個々のインスタンスの始まりから終わりまでかかった時間の全インスタンスでの平均値を示している。同様にして最大時間と最小時間も対応するシーケンスの個々のインスタンスの始まりから終わりまでかかった時間の全インスタンスでの最大値、最小値である。このようにして得られたアクティビティのシーケンスからフロークラスの図を描くことができる(表2を参照)。図1で括弧の数字は遷移の順序を示している。またフローインスタンス数とは表2の件数を意味しシーケンスに含まれるフローインスタンスの数である。これらのフロークラスを重ね合わせると業務フローが求められる(図1を参照)。

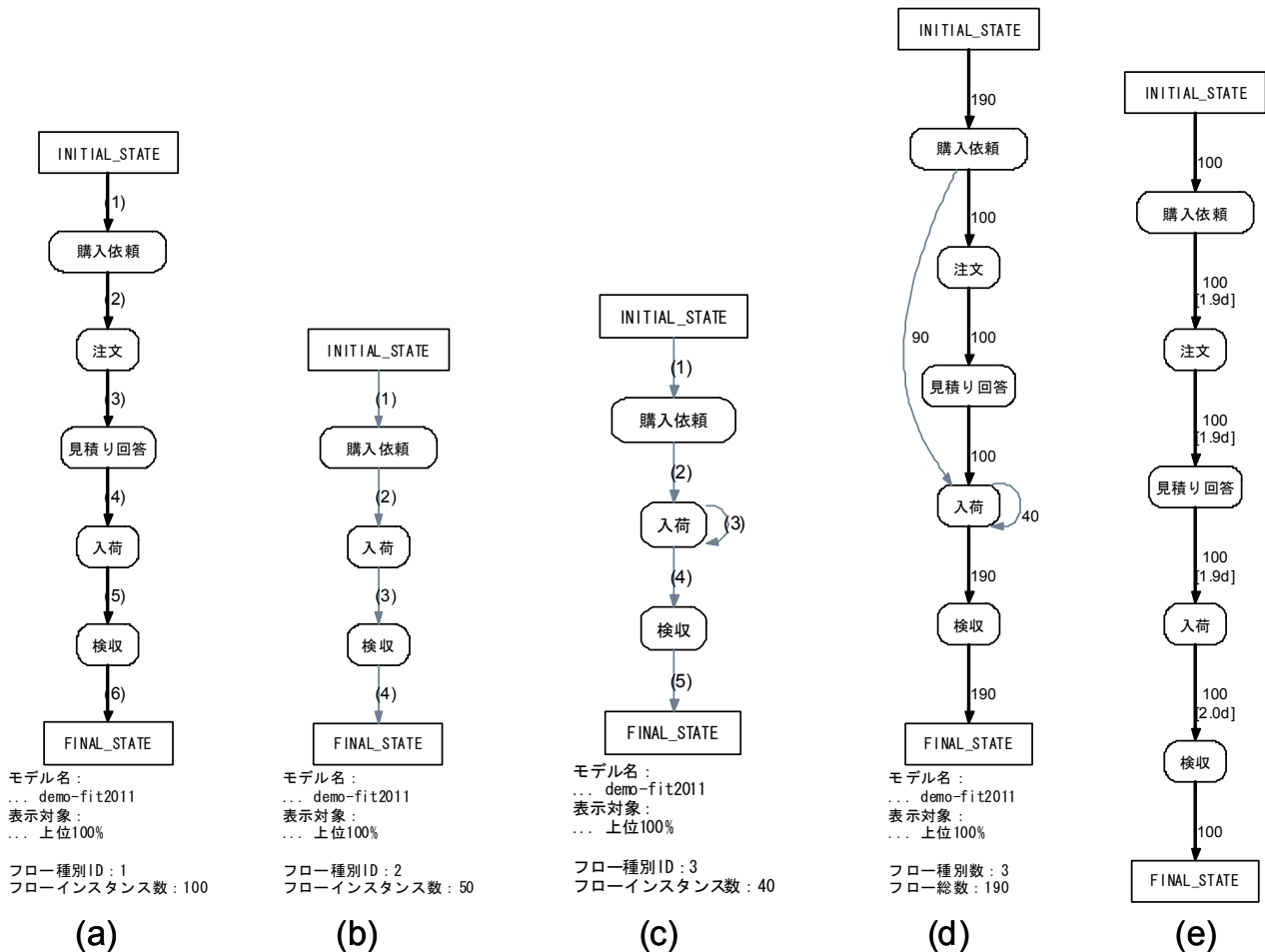


図1 (a), (b), (c)のフロー図は表2のフロー種別を示し、(d)はこれらを重ね合わせた業務フローを示す。(e)のフロー図は標準化された業務フローの説明のためのフロー図である。

3. 業務フローの標準化の指標 (KPI)

3.1 業務フローの理想と現実

通常、企業等の業務は定められた手順に従って進行すると考えられている。このため企業のITシステムに蓄積される業務の証跡に基づいて推測された業務フローも少数の例外を除いて予め定められた手順に従うために単調な、直線状の業務フローになるものと思われるかもしれない(図1の(e)参照)。しかしながら実際の業務フローはこのように単純ではない。業務フローが単純でない原因は次のような理由によると推測できる。企業内での業務フロー変更の原因としては次のようなものが考えられる。基幹系では法令の変更、社外の取引先との契約変更、会社の事業方針の変更などが原因として考えられる。製造系では新製品開発によるラインの変更や製品の機能向上にともなうラインの追加が考えられる。

部品及び原料の供給の社外事情による企業単独では制御できない変化によるライン変更も原因として考えられる。これらの変化が年月とともに積み重なることによって、個々

の業務のフローインスタンスの重ねあわせとして表現される業務フローは複雑さを増加させていくことになる。企業内でのコンプライアンスが如何に徹底されようとも、これらの原因によって業務フローは当初想定した業務フローから逸脱し変化し始めると推測される。時間がたてばたつほど当初想定した業務の流れとは驚くほど異なる業務フローが出現して担当者を驚かすことがあるかもしれない(図2を参照)。図2は基幹系のシステムのログの実際のデータに基づいて作成したフロー図である。

さらに、業務フローが自動的に進行するのではなく、人間の判断によって進行する場合を考えてみる。この場合は担当者が同じだとしても前述したような各種の外部の環境変化を考慮して独自の方針及び判断で業務を遂行するために業務フローも変化しこれにより業務フローが複雑化すると推定できる。さらに普通は担当が替わるのは珍しいことではないので、これにより状況判断の結論が担当者に大きく依存する場合は業務フローも大きく変化することが推測される。

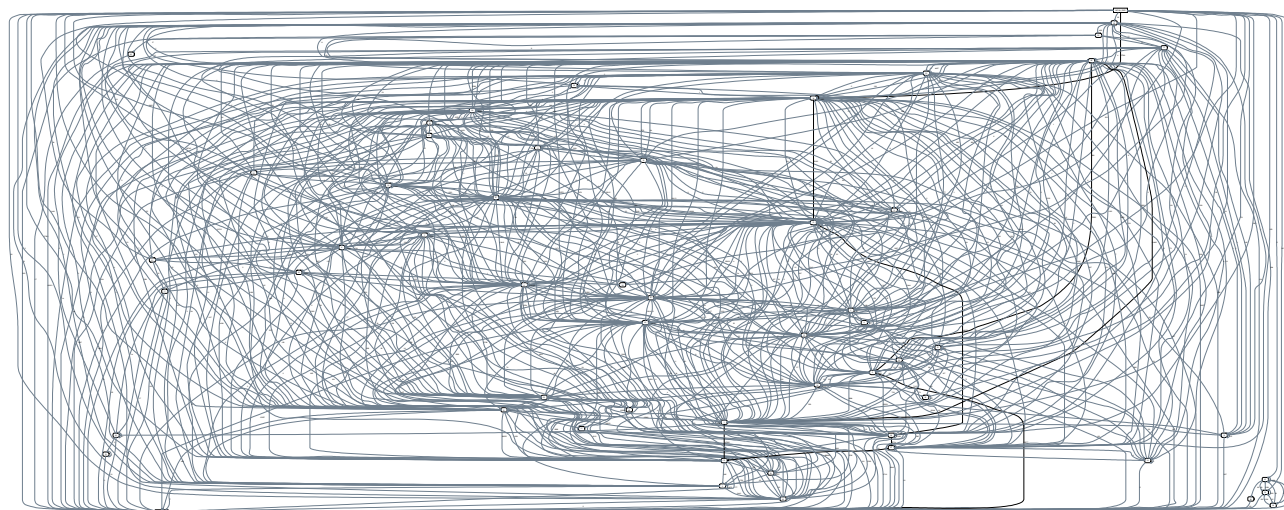


図2 実際の基幹系の業務フローの1例。イベント数が多く、さらに件数が1件程度のフロー種別が多く業務フローを複雑にしている。

3.2 業務フローの見かけ上の複雑さ

業務フローが複雑になる原因としては前述したように企業活動の本質とも言える内的な原因によるものがある。それ以外に、業務フローを複雑にしている別の原因としてITシステムに蓄積されたデータから業務フローを推定する際に業務と対応しないデータを抽出して、誤って業務上のイベントと見做したことによって余分な複雑さが現れる場合がある。さらに幾つかの業務が実際は並列して処理されているのだが、データを抽出する際に並列処理を見分ける有

3.3 業務フローの標準化レベル

業務フローが標準化しているとは、対象としている業務のフローインスタンスの大部分が同一のフロー種別に属していることである。このような同一のフロー種別を標準フローと呼ぶことにしよう。標準フローは一般には複数あっても構わない。現実の業務には例外が生じるのが一般的である。実際の業務フローは全てのフローインスタンスが少数の同一のフロー種別に属するという場合から全てのフ

効な手段がないために業務フローが見かけ上複雑になってしまうといった場合が起こりえる。実際の業務フローを推測する上でこのような複雑さの発生を抑えることが実際上は重要になってくる。これについては論文の後半である程度詳しく論じることにして、ここで仮にそのようなデータ処理上の問題は生じなかったとして業務の標準化の指標について考察することにする。

ローインスタンスが独立したフロー種別となる場合の中間にある。例外の少ない場合に偏れば業務フローが大部分のフローインスタンスが同一のフロー種別に属する、比較的目視で特徴を把握しやすい業務フローになる。そうではなく例外の多い側に偏れば、例外が多く目視で特徴を把握するのが難しい業務フローになる。例外のない場合に接近す

るほど標準化レベルが高く、逆に例外だけの場合に接近するほど標準化レベルが低くなることになる。

このように標準化レベルを標準フローへのフローインスタンスの集中の度合い、例外の多寡を基準にして計測できる。すなわち、フロー種別に属するフローインスタンスの件数をグラフとしてプロットした時に、フローインスタンスの件数が比較的少数のフロー種別に集中している場合には標準化レベルが高く、標準フローの定義が意味を持っている。このようにして得られて標準フローが元来の業務の手引書などの記述と一致していれば業務の標準化が成功していると考えられる。また非常に分類が難しい例として少数のフロー種別に多数のフローインスタンスが集中しているが、これら標準フローに属さない例外的なフロー種別に属するフローインスタンスの総和が標準フローに属するフローインスタンスの総和と同程度の場合には標準化に関して判断するのが難しくなる。このようなグレーゾーンについては標準化の程度について数値化しておく標準化の程度について議論しやすくなる。

標準化の数値化は次のようにして行なう。フロー種別に含まれるフローインスタンスの件数の順にフロー種別をソートして並べたとする。フローインスタンスの件数を上位のフロー種別から積み上げて加算していき全フローインスタンスに対するパーセンテージ、頻度を計算しておく。この頻度がX%のときに対応するフロー種別の数を勘定して全フロー種別に対する比率を計算する。仮に対応するこの比率をY%とする。標準化の程度はフロー種別(一般には複数)の出現頻度X%が大きい場合に対応するフロー種別の比率Y%が小さければ小さいほど標準化レベルが高いことになる。この比率をフロー種別比率と呼ぶ。

フロー種別をフローインスタンスの大きい順番に並べるとする。最もフローインスタンスが多い最初のフロー種別から着目したフロー種別までの累積したフローインスタンスの総和を計算してパレート図とする。フロー種別の総数もフローインスタンスの総数も事例毎に違うために全体の総数に対するパーセンテージで規格化しておくと比較する時に便利である。

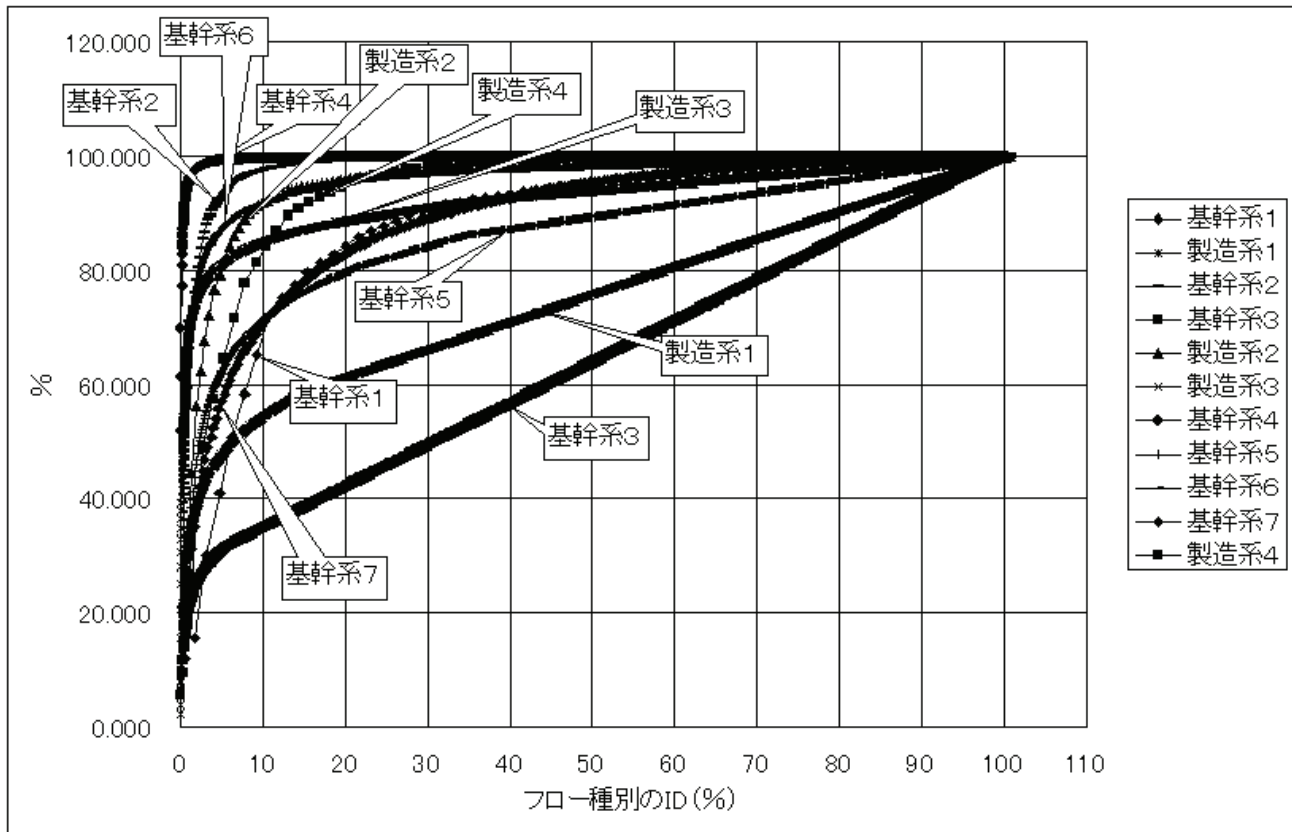


図 3 種々の実際の事例でのフロー種別の件数の累積比率

実際の業務フローの複雑さと標準化レベルが対応するように次のように標準化レベルを定義する。

定義1. (1) 頻度 90%のフロー種別比率 < 10%の時、標準化レベルは5である。(2) 10% ≤ 頻度 90%のフロー種別比率 < 20% かつ 頻度 80%のフロー種別比率 < 10%の場合、標準化レベルは4である。(3) 20% ≤ 頻度 90%のフロー種別比率 < 40% かつ 頻度 80%のフロー種別比率 < 10% または

10% ≤ 頻度 90%のフロー種別比率 < 20% かつ 10% ≤ 頻度 80%のフロー種別比率 < 20% の場合、標準化レベルは3である。(4) 20% ≤ 頻度 90%のフロー種別比率 < 40% かつ 10% ≤ 頻度 80%のフロー種別比率 < 40% かつ 頻度 70%のフロー種別比率 < 10% の場合、標準化レベルは4である。(5) 20% ≤ 頻度 90%のフロー種別比率 かつ 10% ≤ 頻度 80%のフロー種別比率 < 40% かつ 10% ≤ 頻度

70%のフロー種別比率 < 20% かつ 頻度 60%のフロー種別比率 < 10% の場合標準化レベルは1である。

(6) (1)から(5)以外は標準化レベルは0である。経験上の事例に合わせているために見かけ上定義が複雑になっている。標準化レベル5は標準化レベルが「たいへん非常に高い」、以下同様に「非常に高い」、「高い」、「普通」、「低い」、「非常に低い」と表現できる。標準化レベルを定義1のように定義すると図3から表3のような結果になる。表3から基幹系、製造系といった相違は標準化レベルの相違の原因ではないことが判る。今後、標準化レベルを決めているより深い要因の分析が必要である。

表3 定義1による標準化レベル

データ名	標準化レベル
基幹系 1	1
基幹系 2	5
基幹系 3	0
基幹系 4	5
基幹系 5	1
基幹系 6	5
基幹系 7	2
製造系 1	0
製造系 2	5
製造系 3	3
製造系 4	4

4. 業務フローに混入するノイズとその除去

4.1 業務フローにノイズが混入する理由

イベント数が多くともイベント間の遷移の種類が少なれば業務フローは単純になる。

遷移の頻度の多寡は業務フローの複雑さには関係しないが、遷移の種類が多いと業務フローは複雑になる。業務フローが複雑になる場合には主に二つの原因があると考えられる。一つは第3.1節で述べたように分析しようとしている業務の本来の複雑さを反映している場合である。この場合の複雑さはノイズではなく、業務の複雑さを反映している。他の一つは本来の業務とは関係の無い遷移が何らかの理由で業務フローに生じたために複雑になった場合である。後者の複雑化の原因としては本来の業務とは無関係なデータがイベントとして業務フローに紛れ込んでしまった場合が考えられる。例えば、業務で共通に利用するデータベースのデータがイベントとして紛れ込んでしまう場合には、業務とは無関係な遷移が生じてしまう。このような業務とは無関係なイベントを非イベントと呼ぶことにする。このように各業務で共通に利用されるデータから生じた非イベントには他のイベントからの遷移、さらには他のイベントから非イベントへの遷移が発生しやすい。このようなノイズを除去する方法を以下で提案する。以下では提案された方法を検証のために用意した単純化したデータに適用した場合に説明する。

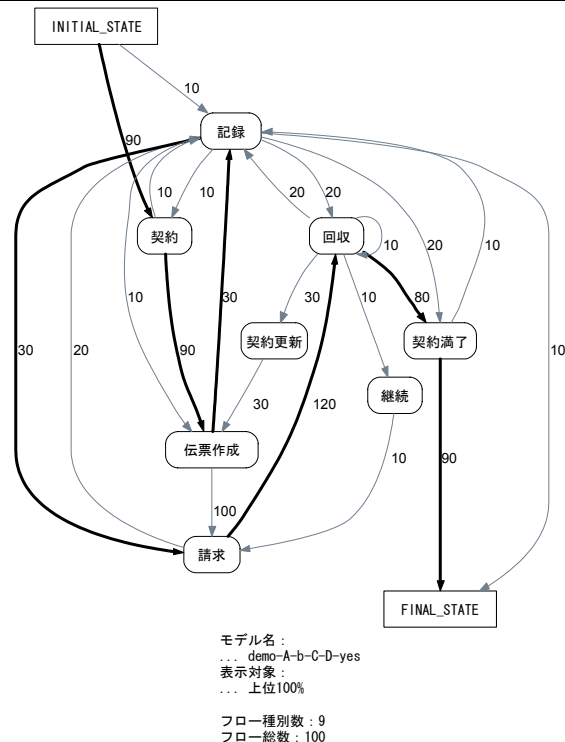


図4 非イベントのある業務フローの一例。この場合は「記録」が非イベントである。

4.2 非イベントの存在の判定法

イベント間の遷移確率は遷移元のイベント(発イベント)を決めたときの遷移先のイベント(着イベント)が発生する条件付確率である。もし着目したイベントが非イベントならば、その発生が発イベントに依存するとは考え難い。すなわち、発イベントが何であって、非イベント発生

条件付確率は発イベントに依存しなくて、着イベントの発生確率に等しいかそれに近い値になるはずである。この発生確率は着イベントの頻度から計算できる。逆に、発生確率が発イベントに依存するならば、そのイベントは業務に関連したイベントと解釈してよさそうである。着目したイベントの発生確率は着イベントの発生頻度から推定できる。

表4 図4の業務フローのイベント間遷移頻度表。

発	着	INITIAL STATE	契約満了	回収	継続	契約更新	請求	記録	伝票作成	契約	FINAL STATE
INITIAL STATE	—	—	0	0	0	0	0	10	0	90	0
契約満了	—	—	0	0	0	0	0	10	0	0	90
回収	—	—	80	10	10	30	0	20	0	0	0
継続	—	—	0	0	0	0	10	0	0	0	0
契約更新	—	—	0	0	0	0	0	0	30	0	0
請求	—	—	0	120	0	0	0	20	0	0	0
記録	—	—	20	20	0	0	30	0	10	10	10
伝票作成	—	—	0	0	0	0	100	30	0	0	0
契約	—	—	0	0	0	0	0	10	90	0	0
FINAL STATE	—	—	—	—	—	—	—	—	—	—	—

表5 表4の頻度表でイベント毎の計と総計から計算したイベントの発生確率の推定値 (発生確率=イベント毎の頻度の計/総計)

契約満了	回収	継続	契約更新	請求	記録	伝票作成	契約	FINAL STATE
0.12	0.17	0.01	0.03	0.16	0.12	0.15	0.12	0.12

表6 表4と表5から計算したイベント発生条件付確率P(着|発)の表

発	着	INITIAL STATE	契約満了	回収	継続	契約更新	請求	記録	伝票作成	契約	FINAL STATE
INITIAL STATE	—	—	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.90	0.00
契約満了	—	—	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.90
回収	—	—	0.80	0.07	1.00	1.00	0.00	0.20	0.00	0.00	0.00
継続	—	—	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00
契約更新	—	—	0.00	0.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00
請求	—	—	0.00	0.80	0.00	0.00	0.00	0.20	0.00	0.00	0.00
記録	—	—	0.20	0.13	0.00	0.00	0.21	0.00	0.08	0.10	0.10
伝票作成	—	—	0.00	0.00	0.00	0.00	0.71	0.30	0.00	0.00	0.00
契約	—	—	0.00	0.00	0.00	0.00	0.00	0.10	0.69	0.00	0.00
FINAL STATE	—	—	—	—	—	—	—	—	—	—	—

表7 表5と表6とから計算した非イベントである場合の確率からの外れの程度の指標 $|1 - P(\text{着}|\text{発})/P(\text{着})|$ 。

発	着	INITIAL STATE	契約満了	回収	継続	契約更新	請求	記録	伝票作成	契約	FINAL STATE
INITIAL STATE	—	—	1.00	1.00	1.00	1.00	1.00	0.14	1.00	6.74	1.00
契約満了	—	—	1.00	1.00	1.00	1.00	1.00	0.14	1.00	1.00	6.74
回収	—	—	5.88	0.62	85.00	27.67	1.00	0.72	1.00	1.00	1.00
継続	—	—	1.00	1.00	1.00	1.00	0.56	1.00	1.00	1.00	1.00
契約更新	—	—	1.00	1.00	1.00	1.00	1.00	1.00	0.53	1.00	1.00
請求	—	—	1.00	3.59	1.00	1.00	1.00	0.72	1.00	1.00	1.00
記録	—	—	0.72	0.24	1.00	1.00	0.32	1.00	0.49	0.14	0.14
伝票作成	—	—	1.00	1.00	1.00	1.00	3.39	1.58	1.00	1.00	1.00
契約	—	—	1.00	1.00	1.00	1.00	1.00	0.14	3.58	1.00	1.00
計			13.60	10.44	93.00	35.67	10.27	6.44	10.60	13.88	13.88

図4の業務フローを例として非イベントの判定方法を具体的に説明する。図4の業務フローで遷移を示す矢印が伴っている数字は遷移の頻度を示している。これらの数字を使ってイベント間の遷移頻度表(表4を参照)が得られる。着イベントの発生頻度T(着)と発生頻度の総計(Total)がこの表から得られる。これらの計算から各イベントの発生確率が $P(\text{着})=T(\text{着})/\text{Total}$ から求められる。表4から求めた各イベントの発生確率を表5にまとめている。例えば、「契約満了」の発生確率は $P(\text{契約満了})=100/860$ となる。さらに発イベントが指定されたときの着イベントの発生確率、条件付確率 $P(\text{着}|\text{発})$ はイベントの発生頻度T(着)と表4の遷移頻度F(着|発)から $P(\text{着}|\text{発})=F(\text{着}|\text{発})/T(\text{着})$ と計算できる。例えば、イベント「回収」から「契約満了」への遷移では、 $F(\text{契約満了}|\text{回収})=80$ 、 $T(\text{契約満了})=100$ となる。よって $P(\text{契約満了}$

$|\text{回収})=80/100=0.8$ となる。表4と表5から計算したイベント発生条件付確率 $P(\text{着}|\text{発})$ の結果を表6にまとめている。次は非イベントの候補を見つけるには、非イベントは業務に関連がないために他のイベントとは独立に発生すると思われる。すなわち非イベント発生条件付確率は非イベントの発生確率に等しい。

$$P(\text{非イベント}|\text{他のイベント})=P(\text{非イベント})$$

次式で定義される着イベントが非イベントである指標を導入する。この指標は着イベントが非イベントの場合に理想的な場合は0となると期待される。

$$|1 - P(\text{着}|\text{発})/P(\text{着})|$$

図4の業務フローについて上記の指標は表5と表6から計算できて表7のように求められる。上記の指標で着イベントを固定して可能な全ての発イベントについて和を取る。この図4の業務フローについて、この指標の総和は表7の「計」欄に表示されている。この値が最も小さい着イベントが非イベントの第一候補となる。図4の業務フローの例では表7からイベント「記録」が非イベントの候補となる。これが実際に非イベントであるかどうかを確認する一つの方法は元の業務フローで非イベント候補になったイベントを仮に削除して、得られた業務フローが比較的単純な業務フローである場合には削除したイベントを非イベントと考えて問題ないであろう。実際、図4の業務フローでイベント「記録」を削除すると対応する業務フローとして図5を得る。図5の業務フロー図は図4の業務フローと比較して単純であることからイベント「記録」を非イベントと設定して問題ないことが判る。

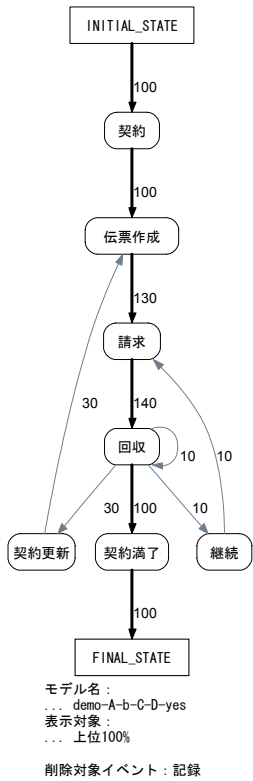


図5 非イベントを削除したフロー図

4.3 並列処理の検出法

図6のように業務フローに並列で処理されるイベントがある場合にも、業務フローは見かけ上複雑になってしまう。なぜなら並列に処理される部分フローが同じ処理IDを持つ場合には、これらの部分フローを並列であると判定するのは一般には難しい。すなわち並列に処理されている部分フローが同じ処理IDを持つために、部分フローのイベント間に本来は無いはずの遷移が発生し、このために部分フローの融合が起こってしまい並列処理を見分けるのが困難になる。部分フローの融合、すなわちイベント間の見かけ上の遷移が生じるために業務フローが複雑になる。見かけ上の遷移が発生するために業務フローが複雑になる点は前

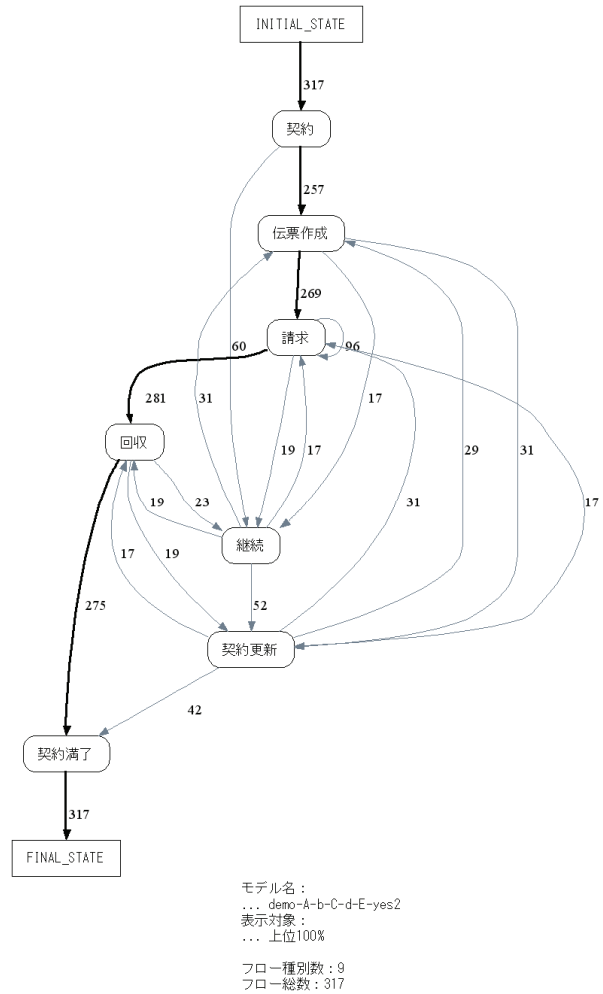


図6 並列処理が混入した業務フローの例

節で議論した非イベントの存在によって業務フローが複雑になる場合と同じである。

このため前節の議論を並列処理の検知に適用して見かけ上の遷移を推定すれば、並列処理の候補を検知することが出来る。実際の例で見かけ上の遷移を特定して並列処理の候補を検知できるかどうかを試してみる。図6の業務フローを例として取り上げる。図6からイベント間の遷移頻度表を得ることができる。

表5と同様にしてイベントの発生確率の推定値を計算する。得られたイベントの発生確率の推定値を計算する。前と同様に着イベントの条件付き発生確

表 8 見かけ上の遷移である場合の確率からの外れの程度の指標 $|1 - P(\text{着}|\text{発})/P(\text{着})|$

発	着	INITIAL STATE	伝票作成	契約	回収	請求	契約更新	継続	契約満了	FINAL STATE
INITIAL STATE	—	1.00	6.05	1.00	1.00	1.00	1.00	1.00	1.00	1.00
伝票作成	—	1.00	1.00	1.00	2.53	3.89	1.68	1.00	1.00	1.00
契約	—	4.72	1.00	1.00	1.00	1.00	8.47	1.00	1.00	1.00
回収	—	1.00	1.00	1.00	1.00	2.00	2.63	5.12	1.00	1.00
請求	—	1.00	1.00	5.25	0.26	1.68	2.00	1.00	1.00	1.00
契約更新	—	0.35	1.00	0.62	0.59	1.00	1.00	0.07	1.00	1.00
継続	—	0.31	1.00	0.58	0.78	7.21	1.00	1.00	1.00	1.00
契約満了	—	1.00	1.00	1.00	1.00	1.00	1.00	1.00	6.05	1.00

率 $P(\text{着}|\text{発})$ を計算する。表 8 の評価指標が 1 より小さい遷移を、互いに並列処理の関係にあるイベント間の見かけ上の遷移として判定すると、

表 8 からイベント「契約更新」から「伝票作成」、「契約更新」から「回収」、「契約更新」から「請求」、「継続」から「伝票作成」、「継続」から「回収」、「継続」から「請求」といった遷移が見かけ上の遷移と判定される。ただし自分自身への遷移は、並列処理ではありえないので、対象外としている。得られた見かけ上の遷移の候補の発イベントと着イベントからイベント「継続」と「契約更新」は他のイベントの作る部分フローと並列に処理されている可能性がある。そこで業務フローでこれらのイベントだけを抜き出して部分フローを作成してみる。イベント「継続」、「契約更新」を仮に削除してイベント「伝票作成」、「契約」、「回収」、「請求」、「契約満了」で構成されている業務フローを作成すると図 7 を得る。得られた業務フローは元の業務フロー（図 6 参照）と比較して複雑でない。削除したイベント「継続」と「契約更新」は図 7 の業務フローとは並列に処理されていると推定される。

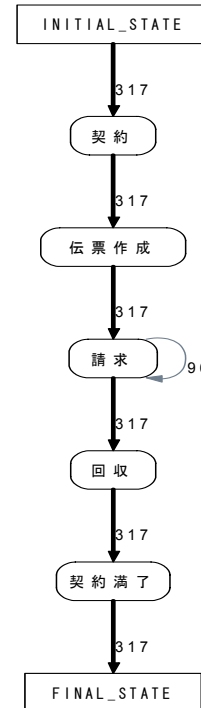


図 7 図 6 の業務フローから並列処理と推定される部分フローを除いたフロー

<http://www.fujitsu.com/global/services/software/interstage/solutions/bpm/apd.html>

- [5] W.M.P. van der Aalst, A.J.M.M. Weijters, and L. Maruster, "Workflow Mining: Discovering Process Models from Event Logs", IEEE Transactions on Knowledge and Data Engineering, Volume 16, Issue 9 (2004).

5. まとめと今後

この論文では企業のデータベースに残されている業務の証跡から業務フローを推測する方法を紹介した。得られた業務フローがどの程度標準化されているかを計測するための尺度として標準化レベルを提案し、実際の業務で使用されているシステムのログに適用した。また複雑な業務フローを簡素化する一つの方法として、非イベント検知及び並列処理検知といった方法を適用して業務フローを簡素化する方法を提案した。

参考文献

- [1] 矢野 啓介, 野村 佳秀, 金井 剛, "業務プロセス自動可視化のための実践的アプローチ", FIT2011
- [2] B.F. van Dongen, A.K.A. de Medeiros, H.M.W. Verbeek, A.J.M.M. Weijters, W.M.P. van der Aalst, "The ProM Framework: A New Era in Process Mining Tool Support", Applications and Theory of Petri Nets 2005, Lecture Notes in Computer Science (2005).
- [3] IDS Scheer PPM, http://www.ids-scheer.com/jp/index_jp.html
- [4] Business Process Management Automated Process Discovery: Fujitsu Global,