

Eine Untersuchung der Potentiale automatisierter Abstraktionsansätze für Geschäftsprozessmodelle im Hinblick auf die induktive Entwicklung von Referenzprozessmodellen

Jana-Rebecca Rehse, Peter Fettke, und Peter Loos

Institut für Wirtschaftsinformatik (IWi)
im Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI),
Universitätscampus, Geb. D3.2, 66123 Saarbrücken, Germany
{Jana-Rebecca.Rehse, Peter.Fettke, Peter.Loos}@iwi.dfki.de

Abstract. In der Literatur wurde bereits eine Reihe von Ansätzen zur Abstraktion für Geschäftsprozessmodelle identifiziert und darüber hinaus verschiedene allgemeine Anwendungsszenarien vorgestellt. Die induktive Entwicklung von Referenzmodellen stellt eine weitere mögliche Anwendung bereits bestehender Abstraktionsansätze dar. Dieser Beitrag zielt darauf ab, Abstraktion in Geschäftsprozessmodellen im Allgemeinen sowie konkret einige dieser Ansätze bezüglich ihrer Eignung zur Entwicklung von induktiven Referenzprozessmodellen zu bewerten. Dazu wird ein generischer Algorithmus entwickelt, der aus einer Menge von Prozessmodellen ein Referenzprozessmodell gewinnt, indem die Modelle gemäß der Vorgehensweise des verwendeten Ansatzes abstrahiert und die abstrahierten Modelle anschließend zusammengefügt werden. Dieses Vorgehen wird genutzt, um acht verschiedene Abstraktionsansätze hinsichtlich ihrer Potentiale im Rahmen der induktiven Entwicklung von Referenzmodellen zu evaluieren, indem sie auf eine Menge von Modellen angewendet werden. Die resultierenden Referenzmodelle werden bewertet und verglichen, um einen Überblick über die verschiedenen Ansätze und ihre Eignung zu erhalten.

Keywords: induktive Entwicklung von Referenzprozessmodellen, Abstraktion in Geschäftsprozessmodellen

1 Einführung

Abstraktion als Idee der Vereinfachung oder Verallgemeinerung von Sachverhalten existiert im wissenschaftlichen Diskurs nicht nur als erkenntnistheoretisches Konzept, sondern auch als spezifische Anwendung in einzelnen Disziplinen. Eine mögliche Interpretation der Abstraktion in Geschäftsprozessmodellen ist, relevante Eigenschaften eines Modelles zu bewahren, während unwesentliche Details ausgeblendet werden können. Diese Interpretation der Abstraktion soll im Folgenden anstelle des allgemeinen Abstraktionsbegriffs verwendet werden. Da bereits vorhandene Arbeiten aufge-

griffen und der Fokus auf die Automatisierung gelegt wird, wird eine mathematisch-formale Definition des Begriffs zugrunde gelegt (vgl. [16]).

Das Konzept ist bereits von verschiedenen Autoren untersucht worden. Neben einer Reihe konkreter Ansätze wie diese Art der Abstraktion zu realisieren ist [4], [10-11], [13] gibt es verschiedene konzeptuelle Arbeiten zum Thema [14], [10] sowie eine Veröffentlichung, die nicht nur ein entsprechendes Rahmenwerk ausführlich definiert, sondern auch Anwendungsszenarien und ausgewählte Ansätze vorstellt [16]. In [14] werden außerdem eine Reihe praktischer Anwendungsfälle für Abstraktionsansätze identifiziert. Diese Publikationen stellen entweder die spezifische Vorgehensweise ihres Ansatzes vor oder ausgewählte, bestehende Ansätze allgemein gegenüber. Die induktive Entwicklung von Referenzprozessmodellen ist dagegen bisher noch nicht als mögliches Anwendungsgebiet von Abstraktionsansätzen identifiziert worden. Da ein induktives Vorgehen, bei dem ausgehend von einer Reihe unterschiedlicher Prozessmodelle ein Referenzprozessmodell entwickelt wird, gezielt Gemeinsamkeiten der Modelle betont und von unternehmensspezifischen Besonderheiten absieht, bieten sich zur Realisierung dieser Modifikationen und zur Generalisierung eines Modells Abstraktionsansätze an, die gegebenenfalls auch weitgehend automatisiert werden können. Die Potentiale bestehender Ansätze bezüglich dieser Anwendung sollen hier betrachtet werden. Dabei ist das primäre Ziel des Beitrags nicht, einen weiteren Abstraktionsansatz zu definieren, sondern detailliert zu evaluieren, wie die Vielzahl an bestehenden Ansätzen allgemein für die induktive Entwicklung von Referenzprozessmodellen verwendet werden kann und was die Vor- und Nachteile der einzelnen Ansätze sind. Ausgehend von dieser Evaluation können dann allgemeine Anforderungen an einen Abstraktionsansatz, der für die induktive Entwicklung von Referenzprozessmodellen verwendet werden soll, formuliert und konkrete Handlungsempfehlungen abgeleitet werden.

Dazu werden zunächst in Abschnitt 2 einige Grundlagen der Abstraktion in Geschäftsprozessmodellen aufgegriffen und kurz erläutert. Abschnitt 3 geht auf die Evaluation der einzelnen Ansätze ein, indem die Auswahl- und Bewertungskriterien erklärt, ein generischer Algorithmus zur Anwendung der Ansätze definiert und die verwendeten Beispielmodelle illustriert werden. In Abschnitt 4 wird die konkrete Anwendung am Beispiel von drei exemplarischen Abstraktionsansätzen vorgestellt, deren Resultate in Abschnitt 5 verglichen und diskutiert werden. Abschnitt 6 beschließt den Beitrag mit einer Schlussfolgerung und der Betrachtung weiterführender Fragen.

2 Grundlagen der Abstraktion in Geschäftsprozessmodellen

Im Folgenden werden die in [16] eingeführten und formal definierten Begrifflichkeiten für die Beschreibung von Abstraktionsansätzen verwendet. Diese sollen hier noch einmal kurz erklärt und in Abbildung 1 illustriert werden. In den betrachteten Arbeiten werden Modelle auf zwei Arten abstrahiert: Entweder wird der Abdeckungsgrad der Modelle reduziert, indem man Objekte entfernt (*Eliminierung*) oder die Granularität wird erhöht, indem man Objekte zusammenfasst (*Aggregation*), sodass das abstrahierte Modell weniger Informationen zur Verfügung stellt als das Ausgangsmodell.

Abstraktionsobjekte sind die Teile eines Modells, die relevante Informationen beschreiben und daher für die Abstraktion in Frage kommen, hier also Knoten und Kanten. *Abstraktionskriterien* dienen dazu, Abstraktionsobjekte zu beschreiben, indem sie damit assoziierte Informationen bereitstellen. Die dazugehörigen *Signifikanzfunktionen* treffen basierend auf den Kriterien die Entscheidung, ob Objekte abstrahiert werden sollen.

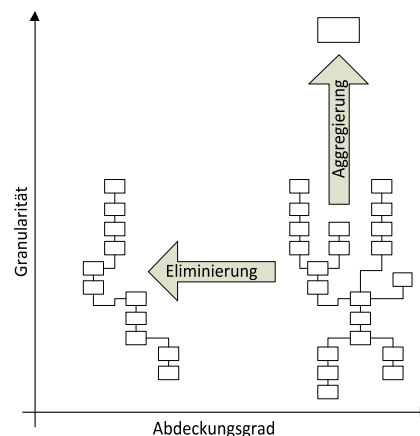


Abb. 1. Illustration von Aggregation und Eliminierung (vgl. [14])

3 Konzeption der Evaluation

3.1 Einordnung der ausgewählten Abstraktionsansätze

Die ausgewählten Abstraktionsansätze werden miteinander verglichen, um die Eignung für die hier vorgestellte Anwendung einzuschätzen, eine detaillierte Analyse durchzuführen und die Automatisierung zu ermöglichen. Diese Evaluation erfolgt anhand des im Folgenden erklärten Kriterienkatalogs, mit dem die Ansätze eingeordnet und bewertet wurden. Diese Kriterien entstammen zum einen dem verwendeten Rahmenwerk zur Beschreibung der Abstraktion [16], sodass sie einen Vergleich der Vorgehensweisen ermöglichen, zum anderen den identifizierten Anforderungen hinsichtlich der induktiven Entwicklung von Referenzprozessmodellen, sodass die tatsächlichen Potentiale der Ansätze bewertet werden können.

1. Prinzip: Welches der möglichen Prinzipien wird verwendet? Dabei wurden bei der Betrachtung der Ansätze das syntaktische Prinzip, das allein anhand syntaktischer Kriterien Abstraktionsentscheidungen trifft, das strukturelle Prinzip, das Informationen über die Prozessstruktur beachtet, und das semantische Prinzip, das versucht, Abstraktionsentscheidungen anhand der Semantik des Modells oder der Modellierungssprache zu treffen, identifiziert.
2. Nutzerorientierung: Wie gut lässt sich die Abstraktion auf die unterschiedlichen Bedürfnisse einzelner Nutzer anwenden?

3. Übergeordnetes Ziel: Wie motivieren die Autoren die Abstraktion und wofür soll sie verwendet werden?
4. Abstraktionsobjekte: Welche Objekte des Modells sollen abstrahiert werden?
5. Abstraktionskriterien: Welche Eigenschaften des Objekts betrachtet man?
6. Signifikanzfunktionen: Wie wird entschieden, ob Objekte abstrahiert werden?
7. Abstraktionsoperationen: Wird Aggregation oder Eliminierung verwendet?
8. Implementierung: Ist eine Implementierung im Forschungsprototyp vorhanden?
9. Anpassung: Wie gut lässt sich der Ansatz auf verschiedene Modellierungssprachen anwenden?
10. Benötigte Informationen: Welche Informationen werden zusätzlich zum Prozessmodell verwendet, um den Ansatz anzuwenden?
11. Mögliche Ergänzungen: Wie kann der vorgestellte Ansatz sinnvoll ergänzt werden?

Neben diesen allgemeinen Kriterien wurden einige Maßstäbe ergänzt, die konkret die Eignung zur induktiven Entwicklung von Referenzprozessmodellen bewerten sollen.

12. Vernachlässigung von Besonderheiten: Wie gut können unternehmensspezifische Besonderheiten ausgeblendet werden?
13. Betonung gemeinsamer Strukturen: Wie gut werden Gemeinsamkeiten der einzelnen Modelle identifiziert und ins Referenzmodell übernommen?
14. Sprachliche Aspekte: Wird auf die Gestaltung sprachlicher Konstrukte (bspw. Knotenlabels) eingegangen?

3.2 Generischer Algorithmus zur Evaluation

Aus einer Menge von Prozessmodellen soll heuristisch ein Referenzmodell entwickelt werden, indem mithilfe eines beliebigen Abstraktionsansatzes jedes Modell einzeln abstrahiert und die Modelle anschließend zusammengefügt werden. Die jeweiligen Abstraktionsobjekte, Abstraktionskriterien und Abstraktionsoperationen müssen also bestimmt werden. Dazu wird ein generischer Algorithmus definiert, der, gegeben eine Menge von Prozessmodellen, mithilfe von Abstraktionsschritten ein Referenzmodell generiert. Ein Abstraktionsschritt besteht aus einem Prozessmodell, einer Menge von darin enthaltenen Abstraktionsobjekten und einer Abstraktionsoperation, die auf diesen Objekten durchzuführen ist. Der generische Algorithmus stellt den Abstraktionsschritt als Schnittstelle bereit, führt die entsprechenden Operationen aus und fügt die Modelle anschließend zusammen. Letzteres kann sowohl manuell als auch mithilfe eines geeigneten Algorithmus wie beispielsweise [8] geschehen.

Das Ziel ist, mithilfe von Abstraktion gezielt unternehmensspezifische Besonderheiten auszublenden, Gemeinsamkeiten der Modelle zu erkennen und die Granularität der Modelle zu vereinheitlichen. Dies ist in Abbildung 2 zu sehen: Während sich Modell A und B nicht sofort zusammenfügen lassen, kann man mittels Abstraktion die Funktionen „Auftrag fachlich prüfen“ und „Auftrag kaufmännisch prüfen“ zu der höher granularen Funktion „Auftrag prüfen“ aggregieren. Da die beiden Modelle danach gleich sind, ergibt sich unmittelbar ein Referenzmodell (die Zuordnung der einzelnen Knoten ist farblich gekennzeichnet).

Der Abstraktionsgrad der Modelle kann vereinheitlicht werden, indem man das Abstraktionskriterium entsprechend wählt. So kann bei gleicher Granularität die gleiche Anzahl Abstraktionsschritte durchgeführt werden. Allerdings ist der Abstraktionsgrad eines Modelles automatisiert nur schwer zu erkennen, daher muss dies im Zweifelsfall manuell festgelegt werden.

Programm: Generischer Algorithmus zur Generierung von Referenzmodellen mittels Abstraktion

```

Menge von Prozessmodellen  $p_1, \dots, p_n$ 
begin
  for all  $p_i$  in  $p_1, \dots, p_n$ 
    while ( $p_i$  contains insignificant objects)
       $S = \text{find insignificant objects}$  in  $p_i$ ;
       $op = \text{find operation}$  in  $p_i$ ;
      Apply  $op$  onto  $S$  in  $p_i$ ;
    merge  $p_1, \dots, p_n$  into  $p$ ;
  return  $p$ ;
end

```

Die unterstrichenen Operationen müssen durch eine Schnittstelle Abstraktionsalgorithmus definiert werden.

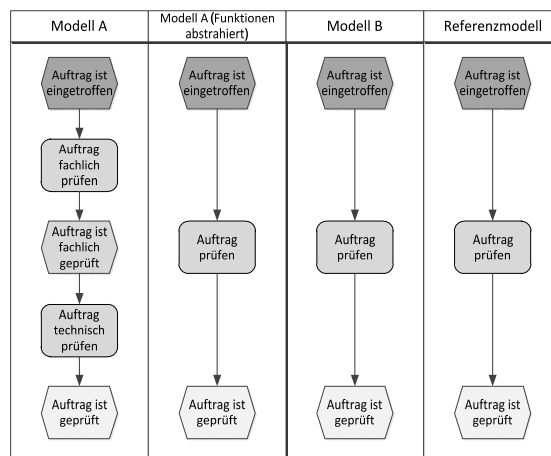


Abb. 2. Beispiel zur Erläuterung des generischen Algorithmus

3.3 Auswahl der betrachteten Algorithmen

In der Literatur existiert eine Reihe von Veröffentlichungen, die sich mit Abstraktion in Geschäftsprozessmodellen beschäftigt, wobei Abstraktion auch implizit zur Generierung von Prozesssichten [2], [7] oder zur Prozessvereinfachung [11] verwendet wird. Eine solche Auswahl befindet sich bereits in [16], allerdings wurde hier nur ein

Teil der dort betrachteten Ansätze übernommen und um andere, dort nicht erwähnte Ansätze ergänzt. Schlussendlich wurden insgesamt acht Ansätze nach folgenden Kriterien ausgewählt:

- **Automatisierung:** Die gezielte Automatisierung der Abstraktion und damit der induktiven Entwicklung der Referenzmodelle steht hier im Mittelpunkt. Die Abstraktion ist daher möglichst mathematisch-formal definiert worden und es wurden solche Ansätze ausgewählt, die dieser Definition entsprechen und die Automatisierung ermöglichen.
- **Anwendung auf Geschäftsprozessmodelle bzw. ereignisgesteuerte Prozessketten [5]:** Da die verschiedenen Ansätze später anhand von als EPK modellierten Prozessmodellen aus [1] evaluiert werden sollen, wurde die Auswahl so eingeschränkt, dass nur solche Ansätze betrachtet wurden, die auf einfachen, allgemeinen Modellierungssprachen oder EPKs definiert oder die leicht auf solche anzupassen sind.
- **Möglichst allgemeine Anwendung:** Um die Vergleichbarkeit der Algorithmen zu gewährleisten, sollten sie möglichst allgemein motiviert und somit auch im hier gewählten Kontext anwendbar sein. Daher wurden solche Ansätze, die ein sehr spezifisches Ziel (bspw. Abstraktion von Prozesssichten für externe Partner, vgl. [3]) verfolgen, hier nicht betrachtet.
- **Definition von Abstraktionsobjekten, -kriterien und -operationen:** Damit sie überhaupt sinnvoll angewendet werden können, sollten die gewählten Ansätze dem hier verwendeten Rahmenwerk der Abstraktion in Geschäftsprozessmodellen entsprechenden und dessen formale Anforderungen möglichst gut erfüllen.
- **Gleichmäßige Verteilung der Prinzipien:** Syntaktische, strukturelle und semantisch eingeordnete Ansätze sollten gleichermaßen betrachtet werden.

3.4 Vorgehensweise

Das verwendete Beispiel besteht aus den Modellen drei verschiedener Rechnungsprüfungsprozesse, die in [1] dokumentiert sind und hier als Modell A, B und C bezeichnet werden (siehe Abbildung 3). Die Zuordnung der einzelnen Prozessbestandteile wird durch die Identität der Knotenlabels beschrieben, das heißt, dass Knoten mit gleicher Beschriftung auch als identisch betrachtet werden. In Abbildung 3 sind exemplarisch die Knoten, die in allen Modellen enthalten sind, farblich hervorgehoben. Hier wurde die Zuordnung aufgrund der Knotenlabels gewählt, für die vorgestellte Anwendung kann allerdings auch eine beliebige andere Zuordnung der Modellteile gewählt werden. Aufgrund der gewählten Zuordnung weisen die Modelle denselben Abstraktionsgrad auf. Sie werden im Folgenden mithilfe eines bzw. mehrerer Schritte abstrahiert und anschließend möglichst passend zusammengefügt.

Um zu evaluieren, inwieweit sich die ausgewählten Ansätze zur Entwicklung eines Referenzprozessmodells eignen, werden im folgenden Abschnitt drei Abstraktionsansätze auf ein Beispiel angewendet. Es wurden ein syntaktischer, ein semantischer und ein „hybrider“ Ansatz, dessen Prinzip durch die Kriterien bestimmt werden kann, gewählt.

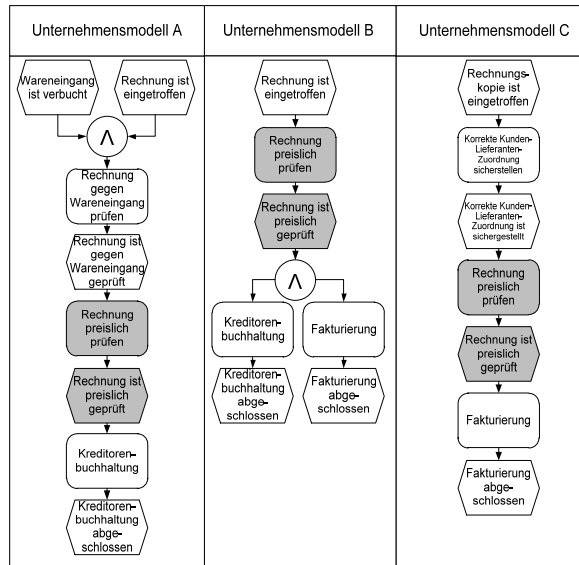


Abb. 3. Beispiele für individuelle Unternehmensmodelle aus [1]

4 Evaluierung der Abstraktionsansätze bezüglich induktiver Referenzprozessmodellentwicklung

4.1 Polyvyanyy et al.: Reducing the Complexity of Large EPCs [11]

Die Operationen dieses Ansatzes sind bereits auf EPKs definiert und können entsprechend der Modelltopologie angewendet werden. Prinzipiell sind bei unterschiedlicher Reihenfolge der verschiedenen Operationen unterschiedliche Modelle zu erwarten. Als Abstraktionskriterium dient hier die relative Häufigkeit der Knoten innerhalb der drei Modelle. In jedem Modell werden die Knoten mit der niedrigsten relativen Häufigkeit identifiziert und die Modellteile, in denen sie sich befinden, gemäß den vorgegebenen Regeln abstrahiert, wobei die Ereignisse entsprechend an die neugebildeten Funktionen angepasst werden müssen.

Diese Knoten in Modell A sind „Wareneingang ist verbucht“ und „Rechnung gegen Wareneingang prüfen“. Ersterer bleibt bestehen, da es keine entsprechende Abstraktionsoperation in diesem Ansatz gibt. Letzterer wird innerhalb der Sequenz mit „Rechnung preislich prüfen“ in die allgemeinere Funktion „Rechnung prüfen“ aggregiert. In Modell B tritt zwar jeder Knoten in mindestens zwei Modellen auf, der Operatorblock aus „Kreditorenbuchhaltung“ und „Fakturierung“ ist dabei am seltensten. Er wird in die Funktion „Rechnung verbuchen“ aggregiert. Modell C enthält den Knoten „Korrekte Kunden-Lieferanten-Zuordnung sicherstellen“, der ebenfalls mit „Rechnung preislich prüfen“ zu „Rechnung prüfen“ aggregiert wird.

Hier ergibt sich direkt eines der Probleme des generischen Algorithmus. Zwar lässt sich durch die elementar gehaltenen Abstraktionsschritte der Abstraktionsgrad des

resultierenden Modells gut kontrollieren. Ist aber in einem Modell nur ein Teil eines größeren Prozesses enthalten, können diese weder abstrahiert noch miteinander identifiziert werden. Hier wurde stattdessen eine linguistische Abstraktion durchgeführt, die Knoten also nach der syntaktischen Abstraktion gemäß dem generischen Prozess umbenannt. Dieser Schritt verfälscht die Abstraktion und ist somit kritisch zu betrachten, allerdings ist dies die einzige Möglichkeit, die abstrahierten Modelle überhaupt zusammenzufügen. Diese Schwierigkeiten ergeben sich auch aus dem hier gewählten Abstraktionskriterium und aus der geringen Größe der Modelle. Das vorgeschlagene Kriterium der Transitionswahrscheinlichkeiten konnte hier nicht angewendet werden, sodass weitere Kriterien überprüft werden sollten. Auf die Benennung der neuen Knoten wird hier ebenfalls nicht eingegangen, sodass die Vorgänge nur schwer automatisiert werden können.

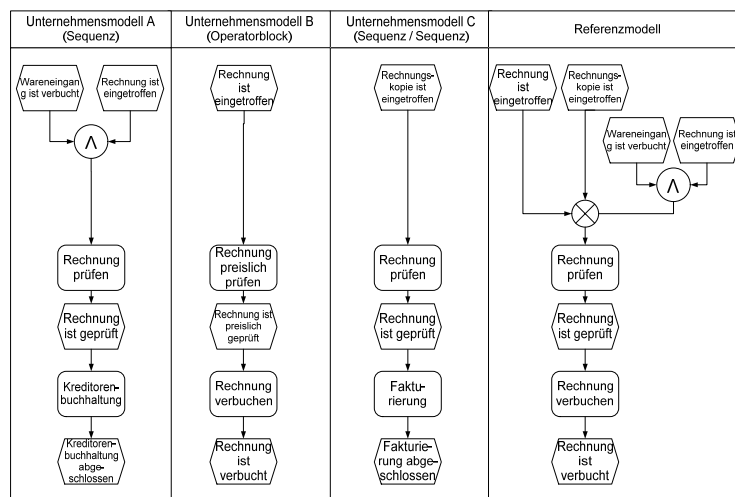


Abb. 4. Illustration des Ansatzes aus [11]

4.2 Smirnov et al.: Meronymy-Based Aggregation of Activities in Business Process Models [13]

Die Anwendung dieses Ansatzes benötigt zunächst die Einordnung der einzelnen Funktionen der Modelle in einem kontrollierten Vokabular, das bestimmt, aus welchen Funktionen sich eine einzelne Funktion zusammensetzt. Dies ist hier durch einen Funktionsbaum (in Abbildung 5) dargestellt.

Auf dieser Grundlage kann nun mithilfe der im Artikel beschriebenen Metrik berechnet werden, welche Menge von Funktionen in eine übergeordnete Funktion aggregiert werden sollen. Dafür berechnet man, wie gut eine Menge von Knoten einen übergeordneten Knoten repräsentiert, wobei die Knotenmenge mit der besten Bewertung in ihren gemeinsamen Vorgängerknoten aggregiert wird. In Modell C ist das der Wurzelknoten, der einen Abdeckungsgrad von 0,6 aufweist. Damit ist der Abstraktionsgrad dieses Modells und damit der des Referenzmodelles sofort sehr hoch, die

anderen Modelle können nur noch daran angepasst werden. Dies geschieht sowohl in Modell A als auch in Modell B durch jeweils zwei Abstraktionsschritte, die die Bewertung 1 bzw. 0,6 aufweisen. Im Modell A werden zunächst „Rechnung gegen Wareneingang prüfen“ und „Rechnung preislich prüfen“ zu „Rechnung prüfen“, in Modell B „Kreditorenbuchhaltung“ und „Fakturierung“ zu „Rechnung verbuchen“ aggregiert. Danach werden in beiden Modellen die verbliebenen Funktionen zum Wurzelknoten „Rechnungsprüfung“ aggregiert. Damit bestehen alle Modelle nur noch aus dieser sehr generischen Funktion und den entsprechenden Startereignissen.

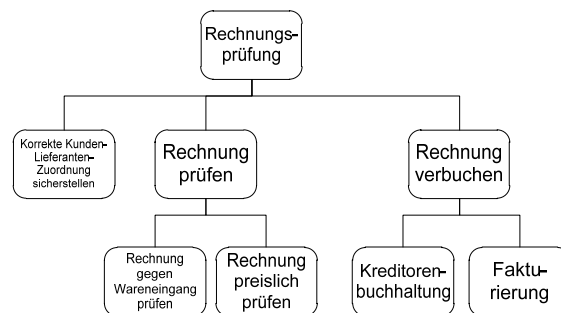


Abb. 5. Funktionsbaum der Modelle A, B und C

Aufgrund des klar definierten Abstraktionskriteriums ist dieser Ansatz am besten zu automatisieren. Allerdings ist dadurch auch der Abstraktionsgrad der Modelle nur schwer zu kontrollieren. Man könnte hier eventuell das Signifikanzkriterium anpassen, sodass nicht immer die Operation mit der besten höchsten Bewertung durchgeführt werden muss. Gleichzeitig ist hier das Ziel, semantisch sinnvoll zu abstrahieren. Dadurch, dass das kontrollierte Vokabular nur auf Funktionen definiert ist, bleiben Ereignisse entweder so bestehen oder werden gemäß der Funktion ausgetauscht, sodass sie an Relevanz verlieren. Außerdem stellt das geforderte kontrollierte Vokabular einen erheblichen Mehraufwand dar, der vor allem bei großen Modellmengen nicht realisierbar ist. Liegt es vor, ist die Abstraktion einfach zu berechnen. Auch die Bezeichnung des neu aggregierten Prozesselements ist durch das kontrollierte Vokabular gegeben.

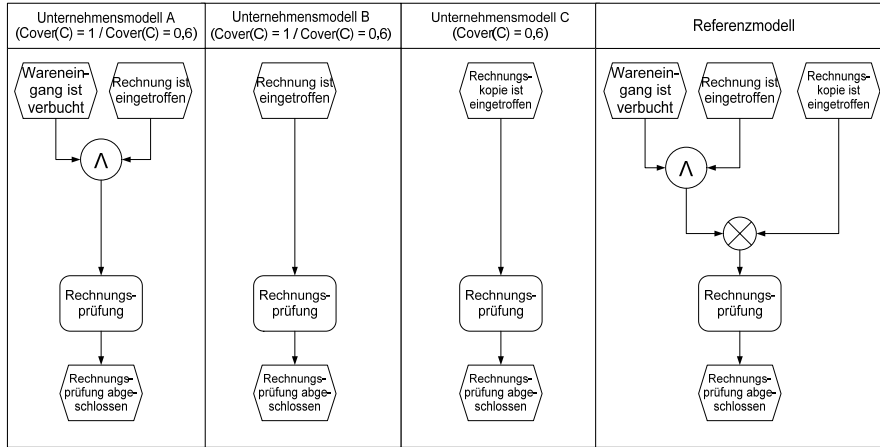


Abb. 6. Illustration des Ansatzes aus [13]

4.3 Günther und van der Aalst: Fuzzy Mining - Adaptive Process Simplification Based on Multi-perspective Metrics [4]

Die hier vorausgesetzte relative Häufigkeit von Knoten und Kanten wird als der Anteil der Modelle definiert, in dem der Knoten bzw. die Kante zwischen zwei Knoten auftritt. Weiterhin sind ein Signifikanz- und ein Korrelationsniveau zu definieren. Um die Vorgehensweise hier anschaulich zu gestalten, wurden beide Zahlen auf 50% gesetzt, sodass alle Objekte, die nur in einem Modell auftreten, abstrahiert werden sollen. Da hier auch das Korrelationsniveau bei 50% liegt, ist keins der Objekte stark mit einem anderen korreliert, alle Objekte werden also aus dem Modell eliminiert. Dies sind drei Knoten im Modell A („Wareneingang ist verbucht“, „Rechnung gegen Wareneingang prüfen“ sowie „Rechnung ist gegen Wareneingang geprüft“). Da alle Knoten aus Modell B in mindestens zwei Modellen auftreten, wird keiner von ihnen abstrahiert. Aus Modell C dagegen werden zwei Knoten („Korrekte Kunden-Lieferanten-Zuordnung sicherstellen“ und „Korrekte Kunden-Lieferanten-Zuordnung ist sichergestellt“) eliminiert.

Ein Vorteil dieser Vorgehensweise ist, dass der Anwender durch die Definition der Niveauewerte entscheiden kann, ob aggregiert oder eliminiert werden soll, während die Operation bei allen anderen Ansätzen vorgegeben war. Bei der hier gewählten relativen Häufigkeit kommt in einem Schritt allerdings nur eine Operation zur Geltung, da eine Kante nicht häufiger sein kann als ihre zugehörigen Knoten. Da die Abstraktionskriterien frei wählbar sind, kann dies aber anders gestaltet werden. Tatsächlich ist dieser Ansatz von allen vorgestellten am flexibelsten: Es werden zwar Kriterien definiert, die konkrete Auswahl bleibt aber dem Nutzer überlassen.

5 Überblick und Diskussion der Ergebnisse

Zum abschließenden Vergleich der Abstraktionsansätze werden die Ergebnisse der Evaluation in Tabelle 1 präsentiert und die drei resultierenden Referenzprozessmodelle in Abbildung 8 dargestellt. Es fällt auf, dass sie sich nicht nur im Abstraktionsgrad, sondern auch in der Modellstruktur erheblich unterscheiden. Dies liegt zum einen an den verschiedenen angelegten Abstraktionskriterien, zum anderen an den unterschiedlichen angewendeten Abstraktionsoperationen. Ausschließlich der Ansatz aus [4] erlaubt die Eliminierung von Startereignissen, in den anderen Ansätzen bleiben diese erhalten und müssen kombiniert werden. Dies steht nicht immer im Einklang mit dem Abstraktionsgrad des restlichen Modells, was besonders beim Modell aus [13] zu sehen ist. Da dort das Abstraktionskriterium ausschließlich Funktionen betrachtet, müssen die Ereignisse manuell angepasst werden. Die strenge Definition von Kriterium und Signifikanzfunktion bestimmt hier den Abstraktionsgrad, der daher nur schwer kontrolliert werden kann. Dieser ist in den beiden anderen Modellen niedriger gewählt worden, könnte aber weiter erhöht werden.

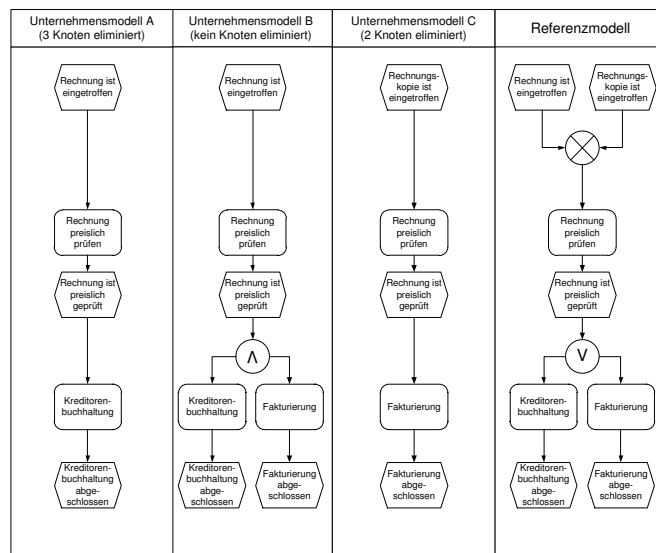


Abb. 7. Illustration des Ansatzes aus [4]

Dagegen decken nur die Modelle, die mithilfe der Konzepte aus [11] und [13] entstehen, den ganzen Prozess ab, was in der angewendeten Operation begründet ist. Hier wurde ausschließlich die Aggregation verwendet, sodass nicht die Granularität sondern eben der Grad der Abdeckung modifiziert wird. Betrachtet man die oben definierten Kriterien zur Bewertung der Abstraktionsansätze, so stellt man fest, dass die Identifizierung gemeinsamer Strukturen bei Polyvyanyy et. al. [11] vor allem manuell durch die Benennung der aggregierten Funktionen und die Identifizierung von Teilprozessen geschieht. Ansonsten werden die Knoten nur durch ihre Labels miteinander

identifiziert, hier könnte eine vorherige manuelle Zuordnung Abhilfe schaffen. Dies würde vielleicht auch die Benennung der Funktionen vereinfachen, da auf sprachliche Aspekte ansonsten nicht eingegangen wird. Dagegen können Unternehmensspezifika der einzelnen Modelle hier gezielt abstrahiert werden, da Kriterium und Operation relativ frei gewählt werden können.

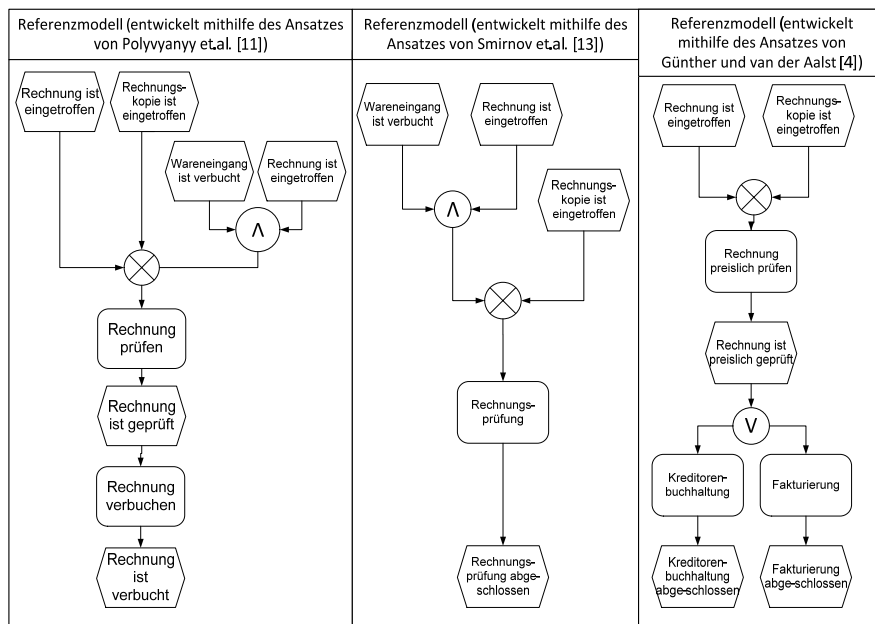


Abb. 8. Vergleich der verschiedenen Referenzmodelle

Bei Smirnov et. al. [13] dagegen ist die Identifikation von Gemeinsamkeiten durch das kontrollierte Vokabular (im Original *ontology*, hier realisiert als Funktionsbaum), die alle in den Modellen auftretenden Funktionen identifiziert und miteinander in Relation setzt, möglich. Diese ermöglicht auch den Umgang mit den entsprechenden Sprachkonstrukten. Insgesamt ist die Existenz eines solchen kontrollierten Vokabulars natürlich nicht immer zu garantieren und nimmt einen Teil der Arbeit auch schon vorweg.

Günther und van der Aalst beschreiben einen sehr allgemein gehaltenen Ansatz, wo viel über die Wahl der Abstraktionskriterien konfiguriert werden kann. Allerdings beziehen auch sie nicht mit ein, wie man mit der Verwendung natürlicher Sprache, wie etwa in Knotenlabels, umgeht. Durch die freie Wahl der Abstraktionskriterien können aber unternehmensspezifische Besonderheiten gezielt ausgeblendet werden. Die Identifizierung der Besonderheiten, die wie bei den anderen Ansätzen auch, insbesondere beim Zusammenfügen der Modelle wichtig ist, erfolgt auch hier nur über gleiche Knotenlabels.

Tabelle 1. Vergleich von acht Abstraktionsansätzen

Kriterien	Ansätze							
	Polyvyanyy et al. [11]	Sadiq / Orlowska [12]	Polyvyanyy et al. [9]	Liu / Shen [7]	Smirnov et al. [15]	Smirnov et al. [13]	Bobrik et al. [2]	Günther / v. d. Aalst [4]
Prinzip	Syntaktisch	Syntaktisch	Strukturell	Strukturell	Semantisch	Semantisch	Semantisch	„Hybrid“
Nutzero.	o	-	+	-	+	o	o	+
Übergeordnetes Ziel	Komplexitätsreduzierung	Aufdeckung struktureller Konflikte	Verwaltung verschiedener Modelle	Generierung von Prozesssichten	Verwaltung verschiedener Modelle	-	Generierung von Prozesssichten	Verdichtung von Modellen
Abstraktionsobjekte	Prozessfragmente	Prozessfragmente	SESE-Regionen	Knotenmengen	Knotencluster	Knotenmengen	Prozessfragmente	Knoten & Kanten
Abstraktionskriterien	Transitions-wahrsch.	Strukturelle Korrektheit	-	-	Ähnlichkeit des Clusters	Grad der Abdeckung	-	Signifikanz & Korrelation
Signifikanzfunktion	Schwellenwerte	Existenzfunktion	-	-	Schwellenwerte	Höchste Abdeckung	-	Schwellenwerte
Abstr.-Oper.	Aggr. & Elim.	Elim.	Aggr.	Aggr.	Aggr.	Aggr.	Aggr. & Elim.	Aggr. & Elim.
Anpassung	o	+	+	o	-	+	+	-
Benötigte Informationen	Erweiterte EPKs	Korrekte Strukturen	Prozessstrukturbaum	Umsetzung der Regeln	Zusätzliche Attribute	kontrolliert. Vokabular	-	Relative Häufigkeiten
Mögliche Ergänzungen	Semantische Kriterien	Kontrolle des Abstr.-grades	Abstraktionskriterien	Abstraktionskriterien	-	Konkrete Anwendung	Abstraktionskriterien	-
Besond.	+	o	o	+	+	o	+	+
Gemein.	-	-	o	-	o	+	o	-
Sprache	-	-	-	-	o	o	o	-

6 Schlussfolgerung

Insgesamt hat jeder der acht hier evaluierten Ansätze ein bezüglich des betrachteten Prozesses sinnvolles Referenzprozessmodell ergeben, sodass sie alle weiter verfolgt werden könnten. Allerdings wiesen einige Ansätze Schwächen wie mangelnde Flexibilität oder die nicht spezifizierte Benennung neu aggregierter Knoten auf. Gerade

letztere ist notwendige Voraussetzung um Abstraktion zu automatisieren, und kann Grundlage für weitere Betrachtungen sein. Weiterhin wäre eine Studie zur konkreten Anwendung der Algorithmen auf weit verbreitete Modellierungssprachen ebenso interessant wie hilfreich, um die praktische Umsetzbarkeit der Ansätze zu beurteilen.

Durch die im Rahmen dieser Evaluation gewonnenen Erkenntnisse können folgende allgemeine Anforderungen an Abstraktionsansätze, die zur Entwicklung von Referenzprozessmodellen verwendet werden sollen, formuliert werden:

- Ausnutzung der verfügbaren Daten: Häufig sind in Prozessmodellen zusätzliche Informationen gespeichert, beispielsweise in Form von Input- und Outputinformationen, die im Zuge der Abstraktion soweit wie möglich ausgenutzt werden sollten.
- Einbeziehung natürlicher Sprache: Zur Entwicklung von Referenzprozessmodellen ist es unabdingbar, natürliche Sprache zu betrachten.
- Unabhängigkeit vom Zusammenfügen: Hier wird das Referenzmodell im Zuge des Zusammenfügens, beispielsweise durch linguistische Abstraktion, verfälscht. Daher sollte ein Abstraktionsansatz möglichst unabhängig vom Zusammenfügen sein.
- Hierarchisierung: Insbesondere bei der Aggregation besteht ein Zusammenhang der Abstraktion mit hierarchisierten Prozessmodellen. Diese Beziehungen zwischen Knoten sollte ein solcher Ansatz ausnutzen.

Erkenntnistheoretisch ist eine induktive Herangehensweise kritisch zu betrachten. Wenn ausgehend von einzelnen Sachverhalten auf allgemeine Gesetzmäßigkeiten geschlossen wird, besteht das Risiko eines Fehlschlusses. Allerdings stellt sich die Frage, inwieweit dies bei der Modellanalyse zu einem Problem wird. Im Kontext der induktiven Referenzmodellierung werden die (automatisiert) gewonnen Modelle nicht vorbehaltlos übernommen und verwendet, sondern kritisch betrachtet und überarbeitet, um sinnvolle Referenzmodelle sicherzustellen.

Abschließend lässt sich die anfangs gestellte Frage zur Eignung von Abstraktionsansätzen für die induktive Entwicklung von Referenzmodellen grundsätzlich positiv beantworten, wobei keiner der Ansätze alle notwendigen Kriterien erfüllt. An dieser Stelle könnte die Entwicklung eines neuen Abstraktionsansatzes speziell zum Zwecke der induktiven Entwicklung von Referenzmodellen gemäß den oben identifizierten Kriterien eine Lösung sein.

Danksagung. Die hier dargestellten Ergebnisse stammen unter anderem aus dem Projekt „Konzeptionelle, methodische und technische Grundlagen zur induktiven Erstellung von Referenzmodellen (Reference Model Mining)“, gefördert durch die DFG (GZ LO 752/5-1).

Literatur

1. Becker, J., Delfmann, P., Knackstedt, R., Kuropka, D.: Konfigurative Referenzmodellierung. In: Becker, J., Knackstedt, R. (eds.): Wissensmanagement mit Referenzmodellen. Konzepte für die Anwendungssystem- und Organisationsgestaltung. Springer, Berlin et al. (2002)
2. Bobrik, R., Reichert, M., Bauer, T.: View-Based Process Visualization. Alonso, G., Dadam, P., Rosemann, M. (eds.): BPM 2007. LNCS, Vol. 4714, pp. 88-95. Springer, Berlin Heidelberg (2007)
3. Chiu, D.K.W., Cheung, S.C., Till, S., Karlapalem, K., Li, Q., Kafeza, E.: Workflow view driven crossorganizational interoperability in a web service environment. *Information Technology and Management* 5 (3-4), 221-250 (2004)
4. Günther, C.W., van der Aalst, W.M.P.: Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. In: Alonso, G., Dadam, P., Rosemann, M. (Hg.): BPM 2007. LNCS, Vol. 4714, pp. 328-343. Springer, Berlin, Heidelberg (2007)
5. Keller, G., Nüttgens, M., Scheer, A.: Semantische Prozessmodellierung auf der Grundlage "Ereignisgesteuerter Prozessketten (EPK)". Technical Report Heft 89, Veröffentlichungen des Instituts für Wirtschaftsinformatik Universität des Saarlandes (1992)
6. Li, C., Reichert, M., Wombacher, A.: Discovering Reference Models by Mining Process Variants Using a Heuristic Approach. In: BPM'09. LNCS, Vol. 5701, pp. 344-362. Springer (2009)
7. Liu, D., Shen, M.: Workflow modeling for virtual processes: an order-preserving process-view approach. *Information Systems* 28 (6), 505-532 (2003)
8. La Rosa, M., Dumas, M., Uba, R., Dijkman, R.: Merging Business Process Models. In: Meersman, R., Dillon, T.S., Herrero, P. (eds.): Proceedings of the International Conference on Cooperative Information Systems. LNCS, Vol. 6426, pp. 96-113. Springer (2010)
9. Polyvyanyy, A., Smirnov, S., Weske, M.: On Application of Structural Decomposition for Process Model Abstraction. In: Proceedings of the 2nd International Conference on Business Process and Services Computing, pp. 110-122. Leipzig (2009)
10. Polyvyanyy, A., Smirnov, S., Weske, M.: Process Model Abstraction: A Slider Approach. In: EDOC 2008, pp. 325-331 (2008)
11. Polyvyanyy, A., Smirnov, S., Weske, M.: Reducing Complexity of Large EPCs. In: EPK'08 GI-Workshop, Saarbrücken. LNI, Vol. P-141, pp. 195-208. GI, Bonn (2008)
12. Sadiq, W., Orłowska, M.E.: Analyzing Process Models Using Graph Reduction Techniques. *Information Systems* 25 (2), 117-134 (2000)
13. Smirnov, S., Dijkman, R., Mendling, J., Weske, M.: Meronymy-Based Aggregation of Activities in Business Process Models. In: Parsons, J. et al. (Hg.): ER 2010. LNCS, Vol. 6412, pp. 1-14. Springer, Berlin Heidelberg (2010)
14. Smirnov, S., Reijers, H., Nugteren, T., Weske, M.: Business Process Model Abstraction: Theory and Practice. Technical Report 35, Hasso-Plattner-Institut, <http://bpt.hpi.uni-potsdam.de/pub/Public/SergeySmirnov/abstractionUseCases.pdf> (2010)
15. Smirnov, S., Reijers, H., Weske, M.: A Semantic Approach for Business Process Model Abstraction. In: Mouratidis, H., Rolland, C. (eds.): CAiSE 2011. LNCS, Vol. 6741, pp. 497-511, Springer, Berlin (2011)
16. Smirnov, S., Reijers, H., Weske, M., Nugteren, T.: Business process model abstraction: a definition, catalog, and survey. *Distributed and Parallel Databases* 30 (1), 63-99 (2012)