

Ansätze zum Ähnlichkeitsabgleich von deklarativen Geschäftsprozessmodellen

Michaela Baumann¹, Michael Heinrich Baumann^{2,3}, Lars Ackermann¹, Stefan Schönig¹, Stefan Jablonski¹

Abstract: Prozesse in Organisationen lassen sich grundsätzlich in zwei Klassen unterteilen: Strikte Routineprozesse und flexible Prozesse mit einem sich dynamisch entwickelnden, variablen Kontrollfluss. Aufgrund der unterschiedlichen Natur strikter und flexibler Prozesse entstanden auch zwei verschiedene Klassen von Prozessmodellierungssprachen: imperative und deklarative. Während imperative Modelle die exakte, schrittweise Reihenfolge der Aufgaben in einem Prozess beschreiben, formulieren deklarative Modelle lediglich Regeln, welchen die Prozessausführung folgen muss. Aufgrund immer größer werdender Prozessmodellsammlungen ist es notwendig, unabhängig vom Modellierungsansatz ähnliche Modelle zu erkennen und zusammenzuführen, was mit dem Begriff Ähnlichkeitsabgleich bezeichnet wird. Für imperative Prozessmodelle existiert bereits eine Vielzahl von Methoden zur Lösung dieses Problems. Da es für deklarative Prozessmodelle jedoch keine allgemein gültigen Lösungen gibt, werden in der vorliegenden Arbeit grundlegende Möglichkeiten des Ähnlichkeitsabgleichs diskutiert.

Keywords: Geschäftsprozessmanagement, Prozessmodellähnlichkeit, Deklarative Prozessmodelle

1 Einleitung

Das Geschäftsprozessmanagement (GPM) stellt eine etablierte Methode zur Modellierung, Ausführung und Analyse von Prozessen dar [Du13]. Prozesse in Organisationen lassen sich grundsätzlich in zwei Klassen unterteilen [Ja94]: strikte Routineprozesse mit eindeutig vorgeschriebenem und vorab spezifizierbarem Kontrollfluss und flexible, entscheidungsintensive Prozesse mit einem sich dynamisch entwickelnden und von Fall zu Fall unterschiedlichen Kontrollfluss. Folgerichtig entstanden auch zwei Paradigmen von Prozessmodellierungssprachen: Imperative und deklarative. Erstere fokussieren das exakte Beschreiben der Reihenfolge der Arbeitsschritte in einem Prozess. Deklarative Modelle hingegen formulieren lediglich Regeln für die Prozessausführung, wobei grundsätzlich jede Ausführung, die keine Regel verletzt, valide ist. Derartige Regeln können die Reihenfolge der Prozessschritte, aber auch die Abhängigkeiten derselben von Daten, Personen und Werkzeugen repräsentieren. Während sich strikte Prozesse häufig besser mittels imperativer Modellierungssprachen (z.B. BPMN) repräsentieren lassen, eignen sich für flexible Prozesse meist eher deklarative Notationen (z.B. CMMN) [Pi12].

¹ Universität Bayreuth, Institut für Informatik, Universitätsstraße 30, D-95447 Bayreuth, {michaela.baumann, lars.ackermann, stefan.schoenig, stefan.jablonski}@uni-bayreuth.de

² Universität Bayreuth, Institut für Mathematik, Universitätsstraße 30, D-95447 Bayreuth, michael.baumann@uni-bayreuth.de

³ Die Arbeit von Michael Heinrich Baumann wird durch ein Begabtenstipendium der „Hanns-Seidel-Stiftung e.V. (HSS)“ aus Mitteln des „Bundesministeriums für Bildung und Forschung (BMBF)“ gefördert.

Aufgrund der wegen Mehrfachmodellierungen immer größer werdenden Sammlungen von Prozessmodellen wird es zunehmend schwieriger, diese zu verwalten. Mehrfachmodellierung kann dabei unter anderem durch verschiedene Versionen desselben Prozessmodells im Rahmen des Business Process Reengineering [TG12], zielgruppenspezifische Modellanpassungen im Zuge der Internationalisierung [Di09] oder Unternehmensfusion verursacht werden. Zur Verwaltung großer Prozesssammlungen ist es notwendig, ähnliche Modelle zu erkennen und zusammenzuführen, was mit dem Begriff Ähnlichkeitsabgleich bezeichnet wird [Di11]. Ein solcher Abgleich dient daneben auch der Wiederverwendbarkeit von Teilprozessen sowie dem Vergleich mit Referenzprozessen (Konformitätsprüfung) und Vorgaben (Complianceprüfung) [BL12]. Die größte Herausforderung für einen solchen Abgleich ist die Heterogenität in Konzeption (Granularität der Aktivitäten) und Realisation (Repräsentation des Verhaltens, Formulierung der Beschriftungen und verwendete Terminologie) der abzugleichenden Modelle [WDM10].

Im Gegensatz zu imperativen Modellen existieren für deklarative Prozessmodelle bislang jedoch kaum Lösungen für den Ähnlichkeitsabgleich, bzw. die vorhandenen können nur sehr fallspezifisch angewendet werden. Ausgehend von einer Einführung zu deklarativen Prozessmodellen stellt die vorliegende Arbeit einen Überblick über mögliche Verfahren des Ähnlichkeitsabgleichs für deklarative Prozessmodelle dar.

2 Deklarative Prozessmodelle

Fahland et al. [Fa09] sehen den wesentlichen Unterschied zwischen imperativen und deklarativen Prozessmodellen darin, wie ermittelt werden kann, ob ein gegebenes Verhalten konform bezüglich des Prozessmodells ist oder nicht. Im Falle der imperativen Modelle muss ein Verhalten mittels eines Pfades durch das Modell rekonstruiert werden können. Im Gegensatz dazu wird bei deklarativen Prozessmodellen geprüft, dass zu keinem Zeitpunkt eine der enthaltenen Regeln verletzt worden ist. Ein deklaratives Prozessmodell ist immer gegeben durch eine Menge an Aktivitäten und eine Menge an Regeln, die die Ausführung der Aktivitäten bedingen. Diese Regeln können beispielsweise in linearer temporaler Logik oder Prädikatenlogik formuliert sein [Fa09]. Wir verwenden im Folgenden keine festgelegte deklarative Modellierungssprache, um die Idee unseres Ansatzes so einfach wie möglich darzustellen. Ein Modell, bestehend aus den Aktivitäten A , B und C , könnte beispielsweise folgende Regeln umfassen: „ A muss mind. einmal und kann max. zweimal ausgeführt werden“, „ B muss genau einmal ausgeführt werden“, „ C darf max. einmal ausgeführt werden“, „ B kann nur dann ausgeführt werden, wenn A bereits ausgeführt wurde“, „Wenn C ausgeführt wird, muss auch B später noch ausgeführt werden“

3 Stand der Wissenschaft

Für imperative Prozessmodelle werden in der Literatur zahlreiche Möglichkeiten für einen Modellabgleich vorgeschlagen. Zusammenfassungen und Auswertungen verschiedener Ansätze finden sich beispielsweise in [Di11]. Die Methoden reichen dabei von strukturellen, graphbasierten Ansätzen über labelbasierte Ansätze bis hin zu verhaltensbasierten

Ansätzen [BBJ15, Di09, LR15]. Informationen, die verwendete Dokumente, zugewiesene Agenten und Rollen oder zu benutzende Werkzeuge betreffen, können ebenfalls für einen Ähnlichkeitsabgleich herangezogen werden [Ba14, WDM10].

Alle Perspektiven, die weder Struktur noch Verhalten betreffen, z. B. Label, können für deklarative wie für imperative Modelle gleich, oder zumindest ähnlich, behandelt werden. Folglich sind im Weiteren vor allem Verhalten und Struktur von Interesse. Für die Ähnlichkeitsbestimmung zweier deklarativer Prozessmodelle können zwei Ansätze der Literatur entnommen werden. Zum einen ist es in manchen Fällen möglich, deklarative Prozessmodelle in imperative zu überführen [PDCM14], für die dann die oben erläuterten Methoden angewendet werden können. Wie in [GH01] gezeigt, ist es auch möglich, bestimmte deklarative Prozessmodelle auf endliche Zustandsautomaten abzubilden und dann einen Editierabstand (Graph-Edit-Distance) zu berechnen [Wo06]. Da in einem Zustandsautomat jedoch Zustände, und nicht Aktivitäten, die Knoten darstellen, ist eine Abwandlung der Abbildung von Aktivitäten auf eine Abbildung von Zuständen erforderlich. Auch ist zu beachten, dass deklarative Modelle, die Variablen benutzen, nicht zwingend in einen endlichen Automaten überführt werden können [ZSJ14]. Es müssen also Ansätze auch für solche Prozessmodelle gefunden werden.

4 Ansätze zum Ähnlichkeitsabgleich

In deklarativen Modellen wird der Kontrollfluss implizit über Regeln eingeschränkt. Letztere bilden demnach in deklarativen Modellen die Grundlage für einen Ähnlichkeitsabgleich für die Verhaltensperspektive. Die Ähnlichkeitsmaße der anderen Perspektiven sollen analog zu denen imperativer Prozessmodelle übernommen werden. Wie für imperative Prozessmodelle wird auch für deklarative Prozessmodelle eine Abbildung M , die zwei Modelle G_1 und G_2 aufeinander abbildet, festgelegt. Für die im Folgenden aufgeführten Ideen zur Ähnlichkeitsberechnung wird zunächst eine 1:1-Abbildung angenommen. Das Finden der besten Abbildung M^* erfolgt dann über einen Optimierungsansatz und wird hier nicht weiter behandelt. Die Ansätze müssen dahingehend bewertet werden, als dass sie Ähnlichkeitswerte liefern, die mit dem „Common Sense“ übereinstimmen, aber auch in der Praxis berechenbar sind.

4.1 Ähnlichkeitsabgleich via Ausführungspfade und Konformitätsprüfung

Die erste vorgestellte Möglichkeit geht ähnlich vor, wie es bei Process-Mining-Verfahren üblich ist, um die Fitness von Logdaten bezüglich der erzeugten Modelle zu bestimmen [RvdA06]. Es werden simulierte, im besten Fall alle möglichen, Ausführungspfade für ein Modell verwendet und dann mit dem zweiten Modell abgeglichen, d. h. auf Ausführbarkeit im zweiten Modell getestet. Da in deklarativen Prozessmodellen jedoch prinzipiell auch unendlich lange Ausführungspfade möglich sind, kann über Simulation nur eine endliche Menge an möglichen, endlichen Ausführungspfaden erzeugt werden. Dieses Thema der Automated Sequence Generation findet sich beispielsweise in [Ha12]. Die Re-

präsentativität des Ergebnisses der Ähnlichkeitsprüfung hängt von der Menge an simulierten Pfaden ab. Im Gegenzug muss auch verglichen werden, ob Ausführungspfade, die in einem Modell nicht erlaubt sind, im anderen auch nicht erlaubt sind. Insgesamt ergäben sich also vier Gruppen von Pfaden (erlaubt-erlaubt, erlaubt-nicht erlaubt, nicht erlaubt-erlaubt, nicht erlaubt-nicht erlaubt). Als Ähnlichkeitsmaß bietet sich der Quotient aus gleichartigen und allen Pfaden an. Sind alle diese Zahlen endlich, kann das Maß einfach berechnet werden, sind die Anzahlen jedoch unendlich, existiert lediglich ein Grenzwert. Es kann dabei vorkommen, dass dieser Grenzwert immer aus $\{0,1\}$ ist.

4.2 Ähnlichkeitsabgleich via Verhaltensmuster und Wahrheitstabellen

Die zweite Möglichkeit ist eine Approximation der Methoden aus 4.1. Es werden zunächst für die abgebildeten Aktivitäten vorher festgelegte Verhaltensmuster bestimmt, die eine bestimmte Anzahl an Aktivitäten umfassen. Bei bis zu 3-stelligen Mustern für die Aktivitäten A und B ergeben sich beispielsweise folgende Muster, wobei nur positive Zusammenhänge berücksichtigt werden: $-, A, B, AA, AB, \dots, BBB$. Es wird dann mit Hilfe der Regelmengen beider Prozessmodelle jedem Muster ein Wahrheitswert zugewiesen, d. h., es wird geprüft, ob das Muster im jeweiligen Modell so ausführbar ist. Ein einzelnes A würde in diesem Fall bedeuten, dass A bzw. $M(A)$ irgendwann einmal ausgeführt wird und davor und danach weder A noch B (bzw. $M(A)$, $M(B)$) ausgeführt werden. Es wird hier zunächst nicht unterschieden, ob die Ausführung direkt oder irgendwann nacheinander erfolgen muss. Sind die Wahrheitswerte für die Muster in beiden Modellen bestimmt, werden sie abgeglichen: Ist ein Muster in beiden Modellen erlaubt oder in beiden Modellen nicht erlaubt, wird ihm der Wert 1, ist ein Muster in einem Modell erlaubt, im anderen nicht, der Wert 0 zugewiesen. Diese Werte werden addiert und durch die Anzahl der Muster geteilt, was einen Mittelwert im Intervall $[0,1]$ liefert. Erlauben beide Modelle die gleichen Muster, ist der Wert 1. Je weniger Muster übereinstimmen, desto mehr tendiert der Mittelwert zu 0. Solch ein Mittelwert wird für je zwei abgebildete Aktivitäten berechnet. Für eine Abbildung mit drei abgebildeten Aktivitäten werden also drei Mittelwerte gebildet, die noch einmal gemittelt werden. Die wichtigste Frage, die sich für diese Möglichkeit stellt, ist die, welche Verhaltensmuster für den Abgleich verwendet werden sollen, um ein möglichst objektives und verallgemeinerbares Ergebnis zu erhalten. Ein Vorteil dieser Methode ist, dass Transitivität in den Regelmengen berücksichtigt wird.

4.3 Ähnlichkeitsprüfung via Prädikatenlogik und Regelähnlichkeit

Eine weitere Möglichkeit, die sich im Zusammenhang mit einem Ähnlichkeitsabgleich von zwei Prozessmodellen ergibt ist die, die Regeln der Prozessmodelle auf Ebene der zugrunde liegenden Logik direkt miteinander zu vergleichen. Es stellt sich die Frage, ob aus der Logik bekannte Erfüllbarkeits- bzw. Entscheidbarkeitskriterien für eine Ähnlichkeitsbestimmung benutzt werden können, wie zum Beispiel in [Bi92] oder [SS94] für statistische Lernverfahren verwendet. Nach einer ersten Einschätzung ergibt sich für die Regelmengen deklarativer Prozessmodelle, zumindest dann, wenn sie Aussagen über komplexere

Zusammenhänge erlauben, dass sie unentscheidbar werden. Der Vergleich kann (gegeben entscheidbare Regelmengen) so erfolgen: Kann eine Regel aus G_1 aus den Regeln von G_2 abgeleitet werden, wird sie mit 1 bewertet, sonst mit 0, und andersherum. Dann werden diese Werte aufsummiert und durch die Anzahl der Regeln geteilt. Ein Problem stellt die Anzahl der Regeln dar, da Regelmengen, die zu einem gleichen Verhalten führen, nicht identisch, also vor allem nicht gleich mächtig, sein müssen. Das heißt, sowohl Dividend als auch Divisor in diesem Ähnlichkeitsmaß sind nicht eindeutig bestimmt.

5 Schluss und Ausblick

Die vorliegende Arbeit zeigt einige Ansätze und Schwierigkeiten beim Ähnlichkeitsabgleich deklarativer Prozessmodelle auf. Da die bisher in der Wissenschaft verfügbaren Techniken nicht für alle flexiblen Prozesse verallgemeinerbar sind, werden neue Ansätze vorgestellt, die jedoch noch weiter untersucht werden müssen. Nach dem derzeitigen Stand der Forschung scheint die Methode *Verhaltensmustervergleich* am praktikabelsten und vielversprechendsten, sowohl von der Konzeption als auch der technischen Umsetzbarkeit. Jedoch sind zunächst alle Ansätze auszuarbeiten und zu evaluieren. Aufgrund mangelnder Abgleichsansätze für deklarative Prozessmodelle, muss für einen Referenzwert zunächst auf das Urteil von Modellierungsexperten für eine Evaluation zurückgegriffen werden. Die vorgestellten Ansätze sollen dann im Weiteren mit Ähnlichkeitswerten anderer Perspektiven, beispielsweise den Aktivitätenbeschreibungen [LR15] oder den Agenten und Ressourcen [Ba15], kombiniert werden.

Literaturverzeichnis

- [Ba14] Baumann, M. H. et al.: Towards Multi-perspective Process Model Similarity Matching. In: Enterprise and Organizational Modeling and Simulation, Jgg. 191 in LNBP, S. 21–37. Springer Berlin Heidelberg, 2014.
- [Ba15] Baumann, M. et al.: Resource-Aware Process Model Similarity Matching. In (Toumani, Farouk; et al., Hrsg.): Service-Oriented Computing - ICSOC 2014 Workshops, Jgg. 8954 in Lecture Notes in Computer Science, S. 96–107. Springer International Publishing, 2015.
- [BBJ15] Baumann, M.; Baumann, M. H.; Jablonski, S.: On Behavioral Process Model Similarity Matching: A Centroid-based Approach. In: ICCGI 2015, S. 125–131. IARIA, 2015.
- [Bi92] Bisson, G.: Learning in FOL with a similarity measure. In: Proc. of the Nat. Conf. on Artificial Intelligence. S. 82–82, 1992.
- [BL12] Becker, M.; Laue, R.: A comparative survey of business process similarity measures. Computers in Industry, 63(2):148–67, 2012.
- [Di09] Dijkman, R. et al.: Aligning Business Process Models. In: Int. Enterprise Distributed Object Computing Conf., S. 45–53. IEEE, 2009.
- [Di11] Dijkman, R. et al.: Similarity of business process models: Metrics and evaluation. Information Systems, 36(2):498 – 516, 2011.

- [Du13] Duman, M. et al.: *Fundamentals of Business Process Management*. Springer, 2013.
- [Fa09] Fahland, D. et al.: Declarative versus Imperative Process Modeling Languages: The Issue of Understandability. In: *Enterprise, Business-Process and Information Systems Modeling*, Jgg. 29 in LNBIP, S. 353–366. Springer Berlin Heidelberg, 2009.
- [GH01] Giannakopoulou, D.; Havelund, K.: Automata-based verification of temporal properties on running programs. In: *16th Annual Int. Conf. on Automated Software Engineering (ASE)*. S. 412–416, 2001.
- [Ha12] Hallé, S. et al.: A Logical Approach to Data-Aware Automated Sequence Generation. In: *Transactions on Computational Science XV*. Springer Berlin Heidelberg, S. 192–216, 2012.
- [Ja94] Jablonski, S.: MOBILE: A modular workflow model and architecture. In: *Proc. of Int. Working Conf. on Dynamic Modelling and Information Systems*. Citeseer, 1994.
- [LR15] La Rosa, M. et al.: Detecting approximate clones in business process model repositories. *Information Systems*, 49:102 – 125, 2015.
- [PDCM14] Prescher, J.; Di Ciccio, C.; Mendling, J.: From Declarative Processes to Imperative Models. In: *SIMPDA*. S. 162–173, 2014.
- [Pi12] Pichler, P. et. al.: Imperative versus Declarative Process Modeling Languages: An Empirical Investigation. In: *Business Process Management Workshops*. Springer Berlin Heidelberg, S. 383–394, 2012.
- [RvdA06] Rozinat, A.; van der Aalst, W. M. P.: Conformance Testing: Measuring the Fit and Appropriateness of Event Logs and Process Models. In: *Business Process Management Workshops*. LNBIP, Springer Berlin Heidelberg, S. 163–176, 2006.
- [SS94] Sebag, M.; Schoenauer, M.: A rule-based similarity measure. In: *Topics in Case-Based Reasoning: First European Workshop, EWCBR-93*. Springer Berlin Heidelberg, S. 119–131, 1994.
- [TG12] Tka, M.; Ghannouchi, S. A.: Comparison of Business Process Models as Part of BPR Projects. *Procedia Technology*, 5:427 – 436, 2012.
- [WDM10] Weidlich, M.; Dijkman, R.; Mendling, J.: The ICoP Framework: Identification of Correspondences between Process Models. In: *Advanced Information Systems Engineering*, Jgg. 6051 in LNCS, S. 483–498. Springer Berlin Heidelberg, 2010.
- [Wo06] Wombacher, A.: Evaluation of Technical Measures for Workflow Similarity Based on a Pilot Study. In: *On the Move to Meaningful Internet Systems (OTM)*, Jgg. 4275 in LNCS, S. 255–272. Springer Berlin Heidelberg, 2006.
- [ZSJ14] Zeising, M.; Schönig, S.; Jablonski, S.: Towards a common platform for the support of routine and agile business processes. In: *Int. Conf. on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*. S. 94–103, 2014.