

Radosław Klimek*, Paweł Skrzyński*, Michał Turek*

Weryfikacja procesów biznesowych metodą tablic semantycznych

1. Wprowadzenie

Notacja BPMN (*Business Process Modeling Notation*) zyskuje obecnie popularność i jest już szeroko wykorzystywana do opisu procesów biznesowych dla przedsiębiorstw. Można ją także traktować – co jednak nie jest do końca poprawne – jako dodatkową warstwę dostarczającą mechanizmów wizualizacji dla języków wykonywalnych takich jak BPEL (*Business Process Execution Language*). Zaletą BPMN i BPEL jest zgodność z architekturą SOA umożliwiającą łatwą integrację różnego rodzaju działalności biznesowej. Pomimo wzrostu znaczenia BPMN nie ma obecnie żadnego formalnego mechanizmu, który umożliwiłby weryfikację procesów biznesowych zapisanych w BPMN. Możliwość walidacji procesu daje symulacja, która jednak wymaga translacji do języka wykonywalnego takiego jak BPEL i wykonania procesu na silniku procesowym. Problem stanowi także to, iż nie każdy proces zapisany w BPMN może być odwzorowany do BPEL. Kolejnym problemem jest fakt, iż translacja taka rzadko jest możliwa do przeprowadzenia w pełni automatycznie (różne narzędzia CASE narzucają tutaj różne ograniczenia), a ingerencja specjalistów wiąże się z dodatkowymi kosztami.

Żeby wypełnić tę lukę, zostało zaproponowane podejście dedukcyjne oparte na metodzie tablic semantycznych umożliwiające automatyczną weryfikację procesów biznesowych opisanych w BPMN. Ponieważ specyfikacje dostarczane przez OMG mają na celu maksymalizację zwrotu kapitału (ROI) poprzez adaptację MDA (*Model Driven Architecture*) w procesie wytwarzania i integracji oprogramowania wprowadzenie dodatkowego ognia w łańcuch wytwarzania oprogramowania umożliwiającego walidację prac na etapie modelowania wydaje się szczególnie celowe.

Artkuł dotyczy zagadnień formalnej weryfikacji modeli procesów biznesowych z wykorzystaniem wnioskowania dedukcyjnego. Do specyfikacji systemu i żądanych własności

* Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie

została wykorzystana logika temporalna jako najlepiej opisująca kwestie żywotności i bezpieczeństwa systemów [6]. Została zastosowana metoda wnioskowania metodą tablic semantycznych [2], która może przebiegać automatycznie i jest ciekawą alternatywą dla podejścia tradycyjnego (np. rezolucja), umożliwiając m.in. względnie łatwe wskazanie błędów w specyfikacji logicznej systemu. Zaproponowane także zostały metody pozyskiwania formuł logiki temporalnej bezpośrednio z modeli zapisanych w notacji BPMN, a zazwyczaj właśnie uzyskiwanie takiej logicznej specyfikacji systemu stanowi wąskie gardło podejścia dedukcyjnego.

Zastosowanie przedstawionego podejścia formalnego ma następujące zalety:

- redukcja kosztów wytwarzania oprogramowania,
- automatyzacja procesu wnioskowania,
- dobra skalowalność metody,
- ekstrakcja formuł logiki temporalnej bezpośrednio z modelu biznesowego,
- możliwość wykrywania anomalii w modelu biznesowym przedsiębiorstwa (sprzeczności w poszczególnych gałęziach drzewa wnioskowania),
- możliwość sprawdzenia osiągalności wszystkich aktywności w modelu.

W procesie adaptacji MDA (*Model Driven Architecture*) duży nacisk powinien zostać położony na możliwość automatycznych przejść pomiędzy kolejnymi krokami. Jak już wspomniano wcześniej, istnieje możliwość opisanie w BPMN procesu w sposób, który później uniemożliwi jego automatyczne wykonanie na silnikach procesowych poprzez BPEL. Rozwiązaniem jest systematyczne zastosowanie wzorców procesowych [1] (stanowiących pewną analogię do wzorców projektowych używanych podczas tworzenia oprogramowania zgodnie z paradygmatem obiektowym). W związku z tym, iż teoretycznie każdy proces biznesowy może zostać opisany jako złożenie poszczególnych wzorców procesowych, artykuł dostarcza opisu automatycznej transformacji (wybranych) wzorców procesowych do formuł logiki temporalnej co umożliwi potem formalną weryfikację metodą dedukcyjną. W przyszłości zostaną stworzone narzędzia umożliwiające zbudowanie formuł logicznych tworzących specyfikację dla dowolnego procesu biznesowego i dzięki czemu proces taki może być formalnie weryfikowany. Artykuł przedstawia przykład procesu biznesowego, dla którego zostaje wygenerowana formuła logiczna, która następnie jest przetwarzana przy pomocy metody tablic semantycznych.

2. Notacja BPMN

BPMN jest standardową notacją graficzną przeznaczoną do opisu procesów biznesowych dostarczoną przez *Business Process Management Initiative* (BPMI), która została zaakceptowana przez OMG. Zasadniczym celem powstania BPMN było dostarczenie notacji, która by była łatwa i zrozumiała dla analityków biznesowych, którzy znają i tworzą procesy

biznesowe oraz specjalistów technicznych, którzy odpowiadają za ich implementację i automatyzację, a także kierowników i menedżerów, którzy tymi procesami zarządzają i je monitorują.

Kolejnym celem było zapewnienie możliwości wizualizacji językom opartym na XML, które umożliwiają wykonanie procesów – takich jak BPEL4WS (*Business Process Execution Language for Web Services*). Zanim podjęto próbę standaryzacji notacji dla takich potrzeb zauważono, że analitycy biznesowi mają tendencję do posługiwania się różnego rodzaju grafami przepływu. Z drugiej strony notacja musiała brać pod uwagę rosnącą popularność języków opartych na usługach sieciowych i XML, które umożliwiały automatyczne wykonanie procesów. Języki te były zgodne z paradygmatem strukturalnym natomiast BPMN musiał mieć równocześnie korzenie w diagramach grafowych co było trudne do pogodzenia. Języki takie jak BPEL są zoptymalizowane pod kątem interoperacyjności systemów zarządzania procesami (BPM, *Business Process Management systems*) co czyni je trudnymi do zrozumienia przez człowieka niebędącego specjalistą technicznym. W związku z tym pojawił się rozdźwięk pomiędzy tym, co było używane przez analityków biznesowych do opisu procesów (nieformalne notacje grafowe) oraz językami używanymi przez techników do implementacji systemów. Notacja BPMN stanowi nie tylko próbę zasypania tej luki, ale uwzględnia też ludzki poziom interoperabilności, który nie był uwzględniany przez języki typu BPEL, nastawione na interoperabilność systemów informatycznych.

W procesie standaryzacji notacji graficznej należy mieć na uwadze złożoność interakcji zachodzących zarówno wewnątrz przedsiębiorstwa, jak i pomiędzy przedsiębiorstwami. Przedsiębiorstwa powinny móc się porozumieć nawzajem, podobnie uczestnicy procesów zachodzących pomiędzy nimi również powinni mieć taką możliwość, aby szybko dostosowywać się do warunków (B2B). Aby to osiągnąć, BPMN bazuje na grafowych diagramach przepływu, dzięki czemu jest bardzo czytelny, a równocześnie dostarcza mechanizmów odwzorowania do postaci wykonywalnej. BPMN odziedziczył mechanizmy B2B takie jak procesy prywatne i publiczne i równocześnie zaadaptował zaawansowane konstrukcje takie jak wyjątki i transakcje.

Jednym z głównych wymagań BPMN była prostota tworzenia modelu biznesowego i równoczesne umożliwienie opis złożoności, która występuje w rzeczywistości przedsiębiorstwa. Żeby spełnić te sprzeczne wymagania, zdecydowano się pogrupować graficzne aspekty notacji w odpowiednie kategorie. W rezultacie dostarczono tylko kilka kategorii notacyjnych, żeby osoba czytająca diagram mogła szybko rozpoznać odpowiednie elementy i go zrozumieć. W ramach tych kategorii można dodawać różne rozszerzenia i wariacje w celu sprostania bardziej złożonym wymaganiom bez zasadniczej zmiany wyglądu diagramu. Cztery podstawowe elementy to:

- Obiekty sterujące przepływem: zdarzenia, aktywności, bramy.
- Obiekty łączące: przepływy sekwencyjny, wiadomość, asocjacja.
- Elementy grupujące.
- Artefakty: dane, adnotacje, grupy.

Wygląd tych elementów jest dosyć intuicyjny i w dużej mierze oparty na stosowanych w UML diagramach aktywności. Głównym problemem staje teraz się opracowanie sposobu automatycznej budowy formuły logicznej na potrzeby weryfikacji takiego procesu.

3. Wzorce procesowe

Riehle i Zullighoven w swojej pracy [8] określają wzorzec jako: „abstrakcyjna forma, której konkretne wystąpienia powtarzają się w określonym kontekście”. Wzorce procesowe można traktować jak przykłady, które ukazują sposób łączenia czynności w celu rozwiązania konkretnego problemu. Analogicznie do podejścia znanego z podejścia obiektowego do tworzenia oprogramowania (Gamma i inni [7] jako pierwsi skatalogowali 23 wzorce występujące w systemach obiektowych), wzorce procesowe zostały zidentyfikowane i opisane w artykule zespołu Van der Alsta [1]. Wzorce te zostały pogrupowane w sześć kategorii:

- podstawowe wzorce sterowania przepływem,
- zaawansowane rozgałęzienia,
- wzorce strukturalne,
- wiele instancji,
- wzorce stanowe,
- wzorce anulowania.

Niniejszy artykuł ogranicza rozważania do podstawowych wzorców sterowania przepływem natomiast dalsze badania prowadzone są nad budową transformacji dla pozostałych wzorców. Do podstawowych wzorców procesowych zaliczamy zatem:

- Sekwencja: aktywność w procesie pracy jest wykonywana po zakończeniu aktywności, która ją poprzedza.
- Rozwidlenie: punkt w którym pojedynczy wątek wykonania się rozwidla na kilka wątków wykonywanych równocześnie lub w jakimś porządku.
- Synchronizacja: punkt, którym wiele równoległych wątków łączy się w jeden z założeniem, że każda z nadchodzących gałęzi wykona się dokładnie raz.
- Wybór wyłączny: punkt, w którym jedna lub wiele gałęzi zostaje wybrana do dalszego wykonania.
- Proste złączenie: punkt, w którym kilka alternatywnych gałęzi się łączy z założeniem, że żadna z równoległych gałęzi nie jest wykonywana równolegle.

Powyższa lista będzie podstawą dalszych rozważań na temat procesów budowy formuł logicznych dla modeli przepływu sterowania w różnych konfiguracjach.

4. Weryfikacja procesów biznesowych

Procesy biznesowe wymagają weryfikacji podobnie jak wszystkie inne procesy i systemy informatyczne. Formalna weryfikacja pełni szczególną rolę w dowodzeniu poprawno-

ści systemów i istnieją tutaj zasadniczo dwa podejścia [5]. Pierwsze jest oparte na eksploracji przestrzeni stanów. Podejście to ma dziś ugruntowaną pozycję i można stwierdzić, że wyprzedziła ona inne strategie weryfikacji co warte jest podkreślenia tym bardziej, że jeszcze niedawno sama metoda była niedoceniana. Olbrzymi postęp wynika m.in. ze znalezienia szeregu efektywnych algorytmów rozwiązujących problem narastającej w trakcie weryfikacji przestrzeni stanów [4]. Obecnie dostępne są już liczne pakiety oprogramowania umożliwiające tzw. weryfikację modelową. Na odmiennym biegunie znajdują się metody dedukcyjne, które zakładają weryfikację w oparciu o formuły przyjętego systemu logicznego. Wnioskowanie logiczne pozwala w zasadzie jako jedyne weryfikować nieskończone obliczenia. Jest to ponadto podejście naturalne dla człowieka. Podejście to ma jednak pewne ograniczenia, o czym będzie mowa w dalszej części pracy. Ważnym i uznanym systemem wnioskowania jest z kolei logika temporalna [6], która może być wykorzystana zarówno w weryfikacji modelowej, jak i podejściu dedukcyjnym. Logika ta pozwala bardzo wygodnie specyfikować wszystkie najważniejsze własności systemów informatycznych należących zarówno do klasy bezpieczeństwa, jak i żywotności. W dalszych rozważaniach ograniczymy się jednak do odmiany liniowej logiki temporalnej w tzw. wersji najmniejszej logiki temporalnej (por. [3]) będącej rozszerzeniem klasycznego rachunku zdań o aksjomat $\Box(P \Rightarrow Q) \Rightarrow (\Box P \Rightarrow \Box Q)$ oraz regułę wnioskowania $\vdash \neg P \Rightarrow \vdash \neg \Box P$.

Wnioskowanie dedukcyjne zakłada stworzenie modelu logicznego weryfikowanego systemu i na tej podstawie dowodzone są specyfikowane własności. Dużą trudnością jest tutaj stworzenie modelu logicznego systemu, który w przypadku zwykłego systemu informatycznego obejmować może bardzo dużą liczbę formuł specyfikowanych niejednokrotnie „ręcznie” (co do których i tak nie ma pewności, czy nie został popełniony błąd wpływający na proces dowodzenia własności). Drugą ważną kwestią jest dobranie właściwego systemu wnioskowania. Systemy klasyczne bazujące na rezolucji zakładają m.in. tworzenie wielu formuł rozrastających się z każdym krokiem dowodzenia, przy czym część z tak utrzymywanych kierunków wnioskowania może w dalszej fazie okazać się niepotrzebna.

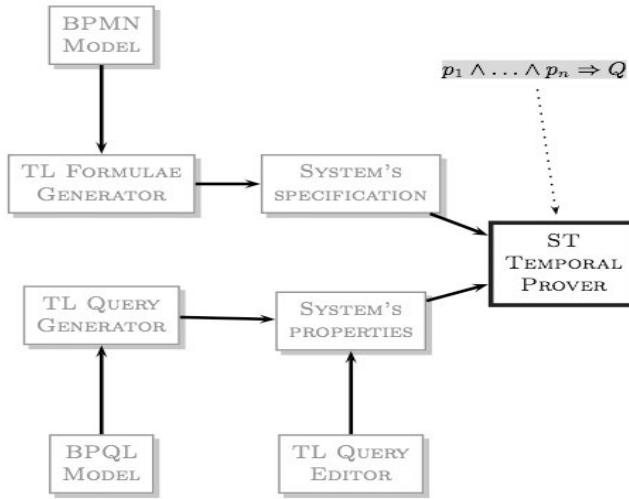
Proponowane w pracy podejście jest odmienne. Przede wszystkim odnosi się do dziedziny procesów biznesowych modelowanych w języku BPMN. Wydaje się, że jest to język szczególnie dogodny do ekstrakcji formuł logicznych wyrażonych w logice temporalnej. Wynika to z faktu, że odpowiednie modele opisują przepływy sterowania, jakie mają miejsce w modelu biznesowym, a same przepływy danych nie są zawarte w takim opisie. Dla modeli biznesowych bardzo silnym i powszechnie przyjętym standardem są wspomniane wzorce projektowe. To właśnie na bazie tych wzorców projektowych może się dokonywać ekstrakcja formuł logicznych. Z każdym wzorcem związana jest formuła lub skończona liczba formuł dobrze opisująca własności wzorca. Kolejną ważną kwestią proponowanego rozwiązania jest dobór odpowiedniego systemu wnioskowania. Postanowiono odejść od systemów tradycyjnych na rzecz wnioskowania metodą tablic semantycznych. Metoda ta jest dobrze znana z logiki klasycznej i może być także z powodzeniem stosowana na gruncie logik modalnych [2, 9]. Generalnie metoda polega na umieszczeniu formuły logicznej w korzeniu drzewa wnioskowania. Formuła jest następnie dekomponowana na formuły

prostsze – aż do pozbycia się wszystkich spójników logicznych formuły początkowej. Dekomponowanie odbywa się poprzez zastosowanie dobrze zdefiniowanych operacji podziału na formuły prostsze. Należy przy tym pamiętać, że w korzeniu drzewa wnioskowania jest umieszczana negacja formuły początkowej. Jeżeli teraz w wyniku procesu dekomponowania, który kończy się po uzyskaniu zdań atomowych w liściach drzewa wnioskowania, otrzymamy w każdej gałęzi sprzeczności, to poszczególne gałęzie (a następnie całe drzewo) uznajemy za domknięte i tym samym formuła początkowa jest spełniona dla każdego swojego wartościowania. Należy zwrócić uwagę na zalety takiego postępowania. W praktyce formuła umieszczona w korzeniu drzewa może być co prawda bardzo długa i rozbudowana, jako np. zawierająca w istocie opis szeregu własności systemu (gdzie poszczególne własności są łączone spójnikiem koniunkcji), jednakże w każdym kroku postępowania formuła staje się prostsza. Ponadto sytuacja nieuzyskania sprzeczności w jakiegokolwiek gałęzi drzewa jest równoznaczna ze wskazaniem wartościowania, które nie spełnia formuły zadanej.

Na rysunku 1 został przedstawiony ogólny schemat proponowanego systemu wnioskowania w środowisku procesów biznesowych i wnioskowania opartego na metodzie tablic semantycznych. System składa się zasadniczo z dwóch części. Pierwsza z nich odpowiada za dostarczenie formuł logiki temporalnej uzyskanych z modeli procesów biznesowych zapisanych w BPMN. Ekstrakcja formuł następuje jednak zawsze z wzorców procesów biznesowych, gdyż przyjmuje się, że cały model jest skomponowany z takich wzorców. Zalety stosowania wzorców są oczywiste i zostały przedstawione w literaturze dziedzinowej. Wszystkie uzyskane w ten sposób formuły logiki temporalnej, przechowywane w osobnym module systemu, składają się na specyfikację samego systemu i będą tym samym tworzyć zasadniczą część formuły poddawanej procesowi wnioskowania, którą można wyrazić logicznie jako koniunkcję własności $p_1 \dots p_n$. Druga część systemu odpowiada za przygotowanie badanych własności modelu. Formuły takie mogą być w najprostszym przypadku wprowadzane poprzez zwykły edytor tekstowy jako badane i żądane własności, ale też mogą być generowane z języka zapytań procesów biznesowych BPQL. Ten drugi sposób pozyskiwania formuł nie będzie w tej pracy dalej rozważany i stanowić będzie przedmiot osobnych badań. Uzyskana formuła Q stanowiąca pożądaną i badaną własność modelu tworzy wraz z formułami pozostałymi – składającymi się na specyfikację modelu i wygenerowanymi z modelu BPMN, następującą formułę wynikową:

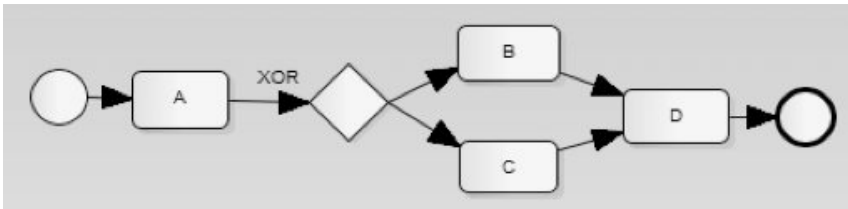
$$p_1 \wedge \dots \wedge p_n \Rightarrow Q \quad (1)$$

Formuła ta stanowi wejście do proponowanego modułu automatycznego wnioskowania metodą tablic semantycznych i po jej zanegowaniu jest umieszczana w korzeniu drzewa wnioskowania. Proces wnioskowania jest zakończony, gdy albo drzewo uznajemy za zamknięte po znalezieniu sprzeczności we wszystkich gałęziach, albo po stwierdzeniu w jakiegokolwiek gałęzi braku takiej sprzeczności po uprzednim uzyskaniu w całym procesie dekompozycji zdań atomowych.



Rys. 1. Schemat systemu weryfikacji procesów biznesowych metodą tablic semantycznych

Jak już wspomniano, ekstrakcja formuł z modeli BPMN następuje przy założeniu, że model taki jest zbudowany od początku do końca z wzorców projektowych opracowanych dla procesów biznesowych. Dla każdego z tych wzorców zostaje wygenerowany zestaw formuł logiki temporalnej specyfikujący własności danego wzorca projektowego. Przykładowo na rysunku 2 zostały pokazane dwa proste wzorce projektowe wraz z bramką xor wybierającą przepływ sterowania.



Rys. 2. Wzorec „Proste złączenie” i „Wybór wyłączny”

Dla takiej sytuacji uzyskamy następujące formuły:

$$(true \vee true) \Rightarrow \diamond A \quad (2)$$

$$A \Rightarrow \wedge(\diamond B \wedge \sim \diamond C) \vee (\sim \diamond B \wedge \diamond C) \quad (3)$$

$$\square \sim (B \wedge C) \quad (4)$$

$$(B \vee C) \Rightarrow \diamond D \quad (5)$$

Po dodaniu formuły własnościowej

$$A \Rightarrow \Diamond \tag{6}$$

otrzymamy formułę:

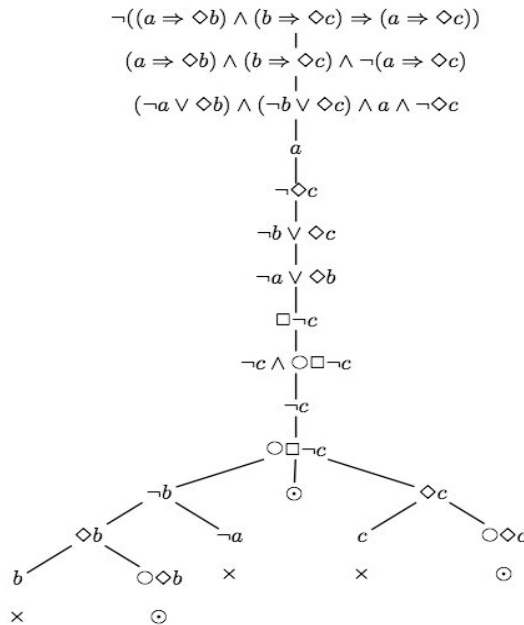
$$((\text{true} \vee \sim \text{true}) \Rightarrow \Diamond A) \wedge (A \Rightarrow (\Diamond B \wedge \sim \Diamond C) \vee (\sim \Diamond B \wedge \Diamond C)) \wedge (\Box \sim (B \wedge C)) \wedge ((B \vee C) \Rightarrow \Diamond D) \Rightarrow (A \Rightarrow \Diamond D) \tag{7}$$

która może już stanowić wejście dla modułu wnioskowania metodą tablic semantycznych. Należy pamiętać, że formuła taka, aby mogła być umieszczona w korzeniu drzewa wnioskowania, musi zostać jeszcze cała zanegowana.

Ilustrację samego procesu wnioskowania przeprowadzimy na prostszym przypadku dwóch złączeń (sekwencja trzech zadań: A, B, C). Sumaryczna formuła wyrażająca specyfikację oraz żadaną własność przyjmie postać:

$$(A \Rightarrow \Diamond B) \wedge (B \Rightarrow \Diamond C) \Rightarrow (A \Rightarrow \Diamond C) \tag{8}$$

Wynika ona z faktu, że po zadaniu A musi być wykonane zadanie B (pierwszy wzorzec „Proste złączenie”), natomiast po zadaniu B musi być wykonane zadanie C (drugi wzorzec „Proste złączenie”). Formuła własnościowa opisuje żądanie, aby po pierwszym z wymienionych zdań zostało kiedyś wykonane ostanie zadanie. Formuła po zanegowaniu zostaje umieszczona w drzewie wnioskowania, a sam proces takiego wnioskowania przebiega tak jak to pokazano na rysunku 3.



Rys. 3. Drzewo wnioskowania dla przykładowej formuły

Dekompozycja została przeprowadzona z uwzględnieniem wszystkich poziomów i rozgałęzień. Drzewo zostało domknięte – zawiera sprzeczności (x) we wszystkich swoich gałęziach. Należy przy tym zauważyć, że niektóre z gałęzi nie są dalej rozwijane, gdyż ich ekspansja nie prowadzi już do nowych zdań elementarnych.

5. Podsumowanie

Artykuł przedstawia innowacyjne podejście do formalnej weryfikacji procesów biznesowych za pomocą podejścia dedukcyjnego metodą tablic semantycznych i logiki temporalnej. W związku z tym, że każdy proces może być przedstawiony jako złożenie skończonej liczby wzorców procesowych, zaproponowano sposób translacji podstawowych wzorców projektowych do formuł logiki temporalnej, które są następnie przetwarzane za pomocą metody tablic semantycznych. Ponadto został omówiony sposób automatycznego generowania formuł logiki temporalnej składających się na specyfikację badanego modelu. W artykule został przedstawiony przykład obrazujący działanie metody. Dalsze badania obejmować będą zagadnienia związane z automatycznym generowaniem formuł logiki temporalnej dla pozostałych wzorców procesowych. Będzie także miało miejsce dokończenie implementacji oprogramowania wspomagającego modelowanie silnika automatycznego dowodzenia i przygotowanie wizualizatora formuł, wraz z konwerterem BPMN do formuł logiki temporalnej.

Na obecnym etapie można powiedzieć, że do ważnych wykazanych już zalet opracowanej metody należą:

- Możliwość automatycznej weryfikacji procesów biznesowych na etapie modelowania.
- Redukcja kosztów wytwarzania oprogramowania poprzez wcześniejszą detekcję błędów.
- Możliwość automatycznego generowania formuł logiki temporalnej bezpośrednio z modelu BPMN.
- Możliwość wskazania błędów i anomalii w analizowanym modelu biznesowym.

Literatura

- [1] van der Aalst W.M.P., ter Hofstede A.H.M., Kiepuszewski B., Barros A.P., *Workflow Patterns*. Distributed and Parallel Databases, 4(1), 2003, 5–51.
- [2] D’Agostino M., Gabbay D.M., Hahnle R., Posegga J. (eds), *Handbook of Tableau Methods*. Kluwer Academic Publishers, 1999.
- [3] van Benthem J., *Temporal Logic*. Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 4, Clarendon Press, 1993–95, 241–350.
- [4] Clarke E.M. Jr., Grumberg O., Peled D.A., *Model Checking*. MIT Press, 1999.
- [5] Clarke E.M., Wing J.M. et al., *Formal methods: State of the art and future directions*. ACM Computing Surveys, vol. 28, No. 4, 1996, 626–643.

- [6] Emerson E.A., *Temporal and Modal Logic*. Handbook of Theoretical Computer Science, vol. B: Formal Models and Semantics, Elsevier, MIT Press, 1990, 995–1072.
- [7] Gamma E., Helm R., Johnson R., Vlissides J., *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley, 1995.
- [8] Riehle D., Zullighoven H., *Understanding and Using Patterns in Software Development*. Theory and Practice of Object Systems, 1996, 2(1):3–13.
- [9] Wolper P., *The Tableau Method for Temporal Logic: An Overview*. Logique et Analyse, vol. 28, 1985, 119–136.