

Przydatność języka Gello i standardu RIM do obiektowego modelowania systemów medycznych

P. KOSIUCZENKO

e-mail: piotr.kosiuczenko@wat.edu.pl

Instytut Systemów Informatycznych
Wydział Cybernetyki WAT
ul. Kaliskiego 2, 00-908 Warszawa

Współczesne systemy medyczne charakteryzują się dużą złożonością i heterogenicznością. Wymaga to ich ujednoczenia, także za pomocą standardów. Health Level 7 (HL7) jest szeroko zaakceptowanym standardem pozwalającym na wymianę danych medycznych pomiędzy różnymi instytucjami i systemami informatycznymi. Stowarzyszony z nim standard Reference Information Model (RIM) dotyczy składni i semantyki wiadomości przesyłanych pomiędzy różnymi systemami medycznymi. Tak jak diagramy klas języka UML, modele oparte o RIM wymagają uszczegółowienia. W tym celu powstał klon języka OCL zwany Gello.

W tej pracy scharakteryzowane zostały wyżej wspomniane standardy oraz powiązania pomiędzy nimi. Zbadana została także ich przydatność do modelowania i projektowania systemów medycznych.

Słowa kluczowe: HL7, RIM, Gello, UML, medycyna, ścieżki kliniczne

1. Wprowadzenie

System opieki zdrowotnej, podobnie jak wiele innych systemów istotnych dla społeczeństwa, jest w swojej naturze bardzo złożony i wymaga właściwego wsparcia ze strony informatyki. Procesy diagnozowania i leczenia wiążą się z wytwarzaniem i przesyłaniem różnorodnych dokumentów. Mogą one być bardzo skomplikowane ze względu na różnorodność terminologii medycznej oraz posiadać złożoną strukturę. W związku z tym powinny one mieć zestandaryzowaną formę. Potrzebne są tu standardy dotyczące struktury, znakowania i semantyki wymienianych dokumentów w celu zapewnienia interoperacyjności systemów informatycznych. Heterogeniczne systemy używane dziś w medycynie często nie są w stanie tego zapewnić. Potrzebny jest też standard wymiany dokumentów.

Najbardziej rozpowszechnionym standardem w tej dziedzinie jest Health Level 7 (HL7) [2] i stowarzyszony z nim standard Reference Information Model (RIM) [3] (zob. także [10]). HL7 jest standardem ANSI dotyczącym wymiany danych pomiędzy systemami informatycznymi opieki zdrowotnej. Definiuje on gramatykę i słownictwo dokumentów, tak aby umożliwić ich wymianę pomiędzy heterogenicznymi systemami opieki zdrowotnej. Standard RIM jest obiektowo zorientowany. Definiuje on składnię i semantykę przesyłanych wiadomości. RIM jest powiązany

z językiem UML [6] będącym standardem do modelowania systemów obiektowo zorientowanych.

Język UML jest zdefiniowany na meta poziomie poprzez tak zwany meta-model będący pewnym szczególnym diagramem klas. Diagramy klas, jako takie, nie posiadają wystarczającej ekspresywności i muszą być uszczegółowione poprzez dodatkowe wymagania. Do tego celu używany jest język Object Constraint Language (OCL) [5]. Jego odpowiednikiem w przypadku HL7 i RIM jest Gello [9].

W tym artykule omawiana jest przydatność standardu RIM i języka Gello do obiektowego modelowania repozytoriów medycznych. RIM i Gello porównane są z językami UML i OCL.

2. Charakterystyka HL7 i RIM

Health Level 7 (HL7, [2]) jest standardem służącym do wymiany danych medycznych i opisu tzw. workflows związanych z diagnozowaniem i leczeniem. Ma on zapewnić interoperacyjność pomiędzy różnymi systemami informatycznymi używanymi w opiece medycznej. Liczba 7 odnosi się do siódmej, najwyższej warstwy w ISO/OSI protokole przesyłu danych. Zwana jest ona „warstwą aplikacji”. Najniższa warstwa zwana jest warstwą fizyczną i dotyczy wymiany danych na poziomie maszyn i bitów. Najnowsza wersja standardu ma numer 3 (oznaczana jest tu jako

HL7 V3). W przeciwieństwie do wersji numer 2, jest ona oparta na formalnej, obiektowo zorientowanej metodologii. HL7 V3 jest też znacznie bardziej dopracowanym dokumentem. Warto tu zaznaczyć, że istnieje polska wersja interfejsu do wymiany danych pomiędzy szpitalami wyspecyfikowana zgodnie ze standardem HL7 (zob. [4]).

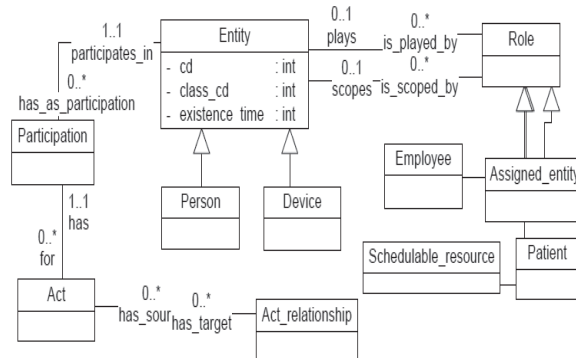
HL7 V3 jest oparty na tzw. Reference Information Model (RIM). RIM jest obiektowo zorientowanym meta-modelem pozwalającym na specyfikację danych (np. danych laboratoryjnych, danych klinicznych, wyników badań klinicznych itp.), a szczególnie wiadomości, zarówno pod względem ich gramatyki, jak i semantyki, w celu zapewnienia interoperacyjności systemów medycznych (zob. np. [1]).

RIM służy też do specyfikacji takich elementów systemów medycznych, jak procesy i workflows. Kręgosłupem RIM są klasy Act, Role i Entity. Są one połączone ze sobą za pomocą asocjacji będących jednocześnie klasami (ang. *association class*), takimi jak: Act_relationship, Participation i Relationship_link (zobacz np. rys. 1).

Klasa Act odgrywa fundamentalną rolę w RIM. Jej elementy nazywane są tu działaniami (ang. *acts*). Działanie jest zamierzonym postępowaniem w dziedzinie opieki zdrowotnej. Działania mogą być powiązane ze sobą za pomocą asocjacji z klasy Act_relationship. Każde działanie występuje w pewnym modusie, który opisuje jego stan od zdefiniowanego, poprzez planowany do zakończonego. Każde działanie ma tylko jeden modus w aktualnym momencie swego cyklu życia.

Klasa Entity odpowiada jednostkom, takim jak ludzie, miejsca i rzeczy. Jej elementy, zwane tu jednostkami, są powiązane poprzez role z innymi jednostkami, a poprzez tzw. partycypacje, tj. elementy klasy Participation, biorą udział w działaniach, tj. elementach klasy Act.

Klasa Role odpowiada rolom, które mogą być odgrywane przez jednostki. Tak jak w języku UML, role są jakby kapeluszkami lub funkcjami, które mogą być „zakładane” lub przyjmowane przez jednostki.



Rys. 1. Podstawowe klasy RIM za [3]

Participation (uczestnictwo) również jest jedną z podstawowych klas. Określa ona, w jaki sposób obiekt z klasy Entity, odgrywający pewną specyficzną rolę, funkcjonuje w zakresie działania. Uczestnictwo jest ograniczone do zakresu działania, w przeciwieństwie do roli określającej kompetencje jednostki niezależnie od jakiegokolwiek działania. Podmioty mogą uczestniczyć w działaniu na wiele sposobów. Na przykład osoby w roli Certified_practitioner mogą uczestniczyć w działaniu będącym elementem klasy Patient_encounter jako lekarze.

Podobnie jak UML, RIM definiuje pojęcie aktora. Aktorzy mogą uczestniczyć w działaniach na różne sposoby. Na przykład dentysta i technik dentystyczny są podmiotami w zabiegu dentystycznym. Są oni mniej lub bardziej bezpośrednio zaangażowani w zabieg. Pacjent może być sam aktorem wykonującym samodzielnie zabieg (np. dezynfekcja i bandażowanie rany, czy auto-iniekcja insuliny)

RIM różni się od UML przede wszystkim tym, że z jednej strony nie posiada tak rozbudowanego meta-modelu i zbioru diagramów, zaś z drugiej strony posiada liczny zbiór dedykowanych klas, które są przeznaczone do modelowania różnych artefaktów medycznych. Zamiast ograniczeń zapisanych w języku naturalnym lub OCL [5], używane są tu tak zwane diagramy blokowe (ang. *Block Diagrams*). Zdaniem autorów standardu RIM (zob. [7]) są one prostsze w użyciu niż OCL, czy komentarze w języku naturalnym umieszczone w prostokątach połączonych z elementami modelu, jak ma to miejsce w przypadku UML-a.

Z czasem zauważono jednak, że niezależnie od tego czy diagramy blokowe są naprawdę prostsze w użyciu, czy też nie, nie wystarczają one do specyfikacji wielu aspektów. Ogólnie rzecz biorąc tego rodzaju diagramy nie mają

odpowiedniej ekspresywności i wiele istotnych ograniczeń nie może być wyrażonych za ich pomocą. Odpowiedzią na te problemy było powstanie języka Gello [9].

3. Gello versus OCL

Wiele ograniczeń dotyczących obiektów i powiązań między nimi nie może być wyrażonych za pomocą samych diagramów klas. Z drugiej strony diagramy klas nie nadają się do specyfikacji metod. Nie są one w stanie wyrazić warunków, które muszą być spełnione podczas funkcjonowania systemu. Jest to możliwe dopiero za pomocą tzw. kontraktów. Potrzebne są formuły wyrażające warunki, kiedy metoda może być wykonana i jaki jest jej rezultat oraz warunki, które muszą być zawsze spełnione.

Są trzy rodzaje formuł tworzących kontrakt. Warunki wstępne (*pre-conditions* w terminologii OCL) pozwalają opisać kiedy metoda może być wywołana, a w szczególności nakładają one ograniczenia na aktualne parametry metody. Warunki końcowe (*post-conditions* w terminologii OCL) pozwalają opisać stan w jakim się ma znaleźć system po wykonaniu metody oraz relację pomiędzy nim a stanem poprzedzającym wykonanie metody. Niezmienniki (*invariants* w terminologii OCL) pozwalają na wyrażenie warunków spójności, które muszą być spełnione podczas działania systemu. OCL jest czystym językiem zapytań, tj. sprawdzenie warunku z założenia nie zmienia stanu obiektów. OCL jest językiem o silnej typizacji wyrażań, tj. każde wyrażenie tego języka musi mieć swój typ.

Gello jest wersją języka OCL skonstruowaną specjalnie dla standardu RIM [3]. W 2005 roku został on zaakceptowany jako standard ANSI pod nazwą HL7 Gello. W zasadzie jest on podzbiorem OCL-a. Tak jak OCL jest to czysty język zapytań, który pozwala na kontraktową specyfikację systemów. Typy danych są zdefiniowane podobnie jak w OCL, jednak w odróżnieniu od OCL-a nie ma tu automatycznego spłaszczenia kolekcji składających się z innych kolekcji. Kolekcja kolekcji może być spłaszczona za pomocą specjalnego operatora *flatten*.

4. CDA

Clinical Document Architecture (CDA) [2] jest standardem służącym do znakowania dokumentów klinicznych, specyfikacji ich struktury i semantyki. Jest to część standardu HL7 V3. Jego celem jest ujednoczenie dokumentów klinicznych umożliwiające ich wymianę.

Dokument kliniczny specyfikuje obserwacje i badania pacjenta, udzielane usługi, trwałość dokumentów (tj. czas, przez który dokument musi być zachowany w niezmienionej formie), organizację odpowiedzialną za dokument, dane dotyczące wiarygodności (ang. *authentication*) i jego kontekst. Standard ten jest oparty na języku XML. Zawartość i semantyka dokumentów są zdefiniowane za pomocą standardu RIM. Dane są zapożyczone ze standardu HL7 V3. Dokumenty CDA mogą być przesyłane w wiadomościach zgodnych ze standardem HL7 V3.

5. Ewaluacja RIM i Gello

HL7 i w szczególności RIM ma swoich krytyków, którzy podważają celowość rozwijania tego standardu, jak na przykład B. Smith, W. Ceusters w pracy [8]. Mankamenty tego standardu wspomniani autorzy przedstawiają następująco:

- Interoperacyjność – pomimo że standard RIM został przyjęty już w 1996 roku, po ponad dziesięciu latach starań i inwestowania w RIM interoperacyjność nie została tak naprawdę osiągnięta.
- Przydatności w nowych dziedzinach zastosowań – RIM opiera się na zestawie „normatywnych” klas (takich jak wyżej wspomniane *Act*, *Role* i *Participation*), z których każda posiada liczny zbiór odpowiednio zinterpretowanych atrybutów. Gdy RIM jest stosowany do nowej dziedziny (jak np. farmaceutyka), należy wybrać tylko pewne z nich. Nie jest jednak jasne, czy wybór atrybutów we wszystkich przypadkach wystarcza.
- Zakres – z klas podstawowych tylko klasa *Entity* może być użyta do modelowania tych rzeczy, które nie są działaniami. Jednak nie wiadomo, jak na tej podstawie można w przejrzysty sposób użyć RIM do modelowania np. aspektów związanych z procesem choroby, czy interakcji pomiędzy lekami, gdyż te ostatnie nie są ani jednostkami, ani działaniami.
- Dokumentacja – dokumentacja oparta na RIM jest niejasna i słabo zintegrowana z innymi częściami składowymi HL7 V3, dla których RIM jest podstawą. Ma ona wiele wewnętrznych niespójności, na przykład w przypadku użycia takich pojęć, jak działanie, instancja klasy *Act*, obiekt klasy *Act*. Jest to szczególnie niepożądane

w przypadku standardów dotyczących wiadomości.

- Trudność nauczenia się – HL7 V3 jest trudny do opanowania dla szerokiego grona użytkowników ze względu na raczej amatorską dokumentację, dużą liczbę przypadków specjalnych, częste zmiany i uzupełnienia oraz trudności w definicji wiadomości.
- Marketing – wiele spośród marketingowych haseł propagujących HL7 V3 jako „standard danych dla biomedycznej informatyki” nie jest uzasadnione, biorąc pod uwagę, że wciąż istnieje wiele nierozwiązanych problemów technicznych.

W pracy [7] przedstawiona jest odpowiedź głównych projektantów RIM na krytykę tego standardu. Twierdzą oni, że krytyka standardu jest niekonstruktywna i przypominają główne założenia RIM, nie wdając się jednak w szczegółową dyskusję.

Trzeba tu zaznaczyć, że powyższa krytyka w punktach dotyczących interoperacyjności i marketingu nie jest czymś wyjątkowym w przypadku HL7 i RIM. Występuje ona także w przypadku wielu innych standardów. Pozostałe punkty wydają się jednak w dużej mierze trafione.

Zaletą RIM jest dopasowanie do potrzeb modelowania i projektowania systemów medycznych, w tym bardzo bogaty zbiór odpowiednich klas. Warto tu dodać, że jeśli chodzi o nowe dziedziny zastosowań, to zasadniczo można tu użyć standardowych metod rozszerzania i dopasowywania meta-modelu tak, jak są one używane w przypadku języka UML. Dokładniej możliwa jest definicja odpowiednich profili tego języka. Każdy model UML-a jest instancją (ang. *instantiation*) meta-modelu UML-a i ma się do niego tak, jak obiekt do klasy. Istnieje tu możliwość użycia stereotypów nadających pewne szczególne znaczenie konkretnym klasom (np. do oznaczenia choroby) oraz innym elementom modelu. Istnieje też możliwość wprowadzania nowych meta-klas i atrybutów, które mogą modelować konkretne klasy, odpowiadające pewnym szczególnym zastosowaniom oraz powiązania między nimi. Ta ostatnia opcja nazywana jest rozszerzeniem wagi ciężkiej, gdyż modyfikuje oryginalny meta-model UML-a.

Jednakże autorzy standardu RIM nie przewidują tego rodzaju rozszerzeń i wymagają, aby klasy wprowadzane przy specyfikacji rozszerzały klasy występujące w modelu RIM (zob. [7], rozdział drugi), co znacznie ogranicza możliwość adaptacji tego standardu. Autorzy nie

wspominają o stereotypizacji klas i ogólnie o profilach. Nie jest to też zgodne z ideą UML-a, że każdy model jest instancją meta-modelu. Zamiast tego proponowany jest mechanizm rzutowania polegający na pominięciu niepotrzebnych atrybutów – raczej na zasadzie ich nieużywania, jeśli nie są potrzebne, niż dosłownego ich usuwania z modelu – jak również poprzez dodatkowe warunki uszczegóławiające dany model. Takie podejście ma zapewnić interoperacyjność, ogranicza ono jednak znacznie możliwość adaptacji modelu RIM, a przez to także jego zastosowania.

Wspomniane diagramy blokowe mają z jednej strony umożliwić bardziej szczegółową specyfikację systemów, niż za pomocą samych diagramów klas, z drugiej strony mają być bardziej czytelne niż formuły OCL-a. Jak wspomniano w poprzednim rozdziale, diagramy blokowe w żadnym razie nie mogą zastąpić OCL-a, gdyż nie mają wystarczającej mocy opisowej. Ponadto osobie mającej pewną wprawę w czytaniu diagramów UML-a zdecydowanie łatwiej jest czytać diagramy klas z dodatkowymi ograniczeniami zapisanymi np. w języku OCL niż diagramy blokowe.

Wspomniane powyżej mankamenty nie wiążą się bezpośrednio z zastosowaniami w medycynie. Główny problem ze standardem RIM polega na tym, że pod względem metodologicznym jest on wciąż mało dopracowany. Wiąże się to z faktem, iż powstawał w dużej mierze równoległe z językiem UML i podczas jego tworzenia niektóre istotne ustalenia i rozwiązania metodologiczne dotyczące tego ostatniego języka nie były znane.

Stowarzyszony z RIM język Gello jest klonem języka OCL, niewiele odbiegającym od swego pierwowzoru i jako taki ma mocne podstawy metodologiczne. Jeśli chodzi o stosowalność HL7, a w szczególności RIM, to kluczową rolę odgrywają tu dwa fakty. Po pierwsze powszechność użycia – jeśli standard jest szeroko przyjęty, to jego użyciu trudno się oprzeć, nawet jeśli ma poważne mankamenty natury metodologicznej. Po drugie istotne jest ulepszanie standardu. Wielu wspomnianym problemom można byłoby zaradzić poprzez przyjęcie odpowiednich zmian w standardzie.

6. Konkluzja

W niniejszej pracy omówiony został standard HL7, blisko związany z nim standard RIM, standard CDA oraz język Gello. Zaletą HL7 i RIM są bliski związek z zastosowaniami medycznymi i doświadczenie w ich użyciu.

Za przyjęciem HL7 i RIM przemawia ich popularność. Przeciw ich przyjęciu przemawiają niedoskonałości metodologiczne oraz nie-precyzyjność specyfikacji. Język Gello jest klonem języka OCL i jako taki ma dobre podstawy metodologiczne oraz teoretyczne, nie wprowadza on jednak żadnych istotnych zmian, czy adaptacji w celu specyfikacji systemów medycznych.

Mimo technicznych niedoskonałości RIM, trudno byłoby zalecać przyjęcie nowego standardu, gdyż o użyteczności standardu decydują nie tylko jego strona metodologiczna, narzędzia wspierające i formalne, ale przede wszystkim powszechność jego użycia. W przypadku kiedy nie zbierze się konsorcjum o odpowiednim potencjale finansowo-politycznym, mogące przeforsować szerokie użycie nowego standardu, jego wprowadzanie jest raczej skazane na niepowodzenie.

Inaczej ma się sprawa z Gello i OCL. Jak wspomnieliśmy, Gello jest zaledwie klonem OCL-a i właściwie stanowi jego fragment. OCL jest szeroko znany i bardzo dobrze dopracowany od strony metodologicznej. Z drugiej strony istnieją liczne narzędzia UML-a ten język wspierające, w tym szeroko stosowane, jak na przykład Rational Architect i Together. Dlatego też język OCL może być rozważony jako alternatywa dla Gello.

7. Bibliografia

- [1] B. Blobel, G. Stassinopoulos, P. Pharow, „Model-Based Design and Implementation of Secure, Interoperable EHR Systems”, *Proc. Of AMIA Annual Symposium*, 96-100, 2003.
- [2] HL7 Board, HL7, <http://www.hl7.org/>.
- [3] HL7 Board, „HL7 Reference Information Model”, V 01-12, 1.2.2002, http://www.miforum.net/distillate/rim/rim0112_body.htm.
- [4] CompuGroup, „Projekt HL7PL”, Polska Wersja Standardu HL7: http://www.uhc.com.pl/index.php?option=com_content&task=view&id=23&Itemid=75.
- [5] OMG, „Object Constraint Language”, Version 2.0, formal/06-05-01, OMG, 2006.
- [6] OMG: „Unified Modeling Language Specification”, Superstructure, Version 2.2, formal/2009-02-02, 2009.
- [7] G. Schadow, C. Mead, D. Walker, „The HL7 Reference Information Model Under Scrutiny”, *Studies in health technology and informatics*, Vol. 124, 2006, str.151-156.
- [8] B. Smith, W. Ceusters, „HL7 RIM: An Incoherent Standard”, *Studies in Health Technology and Informatics*, vol. 124, 133-138, 2006.
- [9] M. Sordo, O. Ogunyemi, A. Boxwala, R. Greene, S. Tu, „GELLO: An Object-Oriented Query and Expression Language for Clinical Decision Support”, *Technical Report*, DSG-TR-2003-02, Harvard Medical School, 2003.
- [10] Y. Sumita, M. Takata, K. Ishitsuka, Y. Tominaga, K. Ohe, „Building a reference functional model for EHR systems”, *International Journal of Medical Informatics*, vol. 76(9), 688-700, 2007.