

**KWESTIONARIUSZ-RECENZJA ROZPRAWY DOKTORSKIEJ
DLA RADY WYDZIAŁU MATEMATYKI, INFORMATYKI I MECHANIKI
UNIwersytetu Warszawskiego**

Tytuł rozprawy: Implementacja systemu wnioskującego za pomocą DNA

Autor rozprawy: mgr Łukasz Rogowski

- 1. Jakie zagadnienie naukowe jest rozpatrywane w pracy (teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?**

Recenzowana praca dotyczy tzw. obliczeń DNA (ang. DNA computing) będącą nową dziedziną wiedzy, leżącą na pograniczu informatyki i biologii molekularnej. Celem tych badań jest opracowanie nowego paradygmatu przetwarzania informacji na poziomie molekularnym wykorzystując zdobyte intensywnie obecnie rozwijanej inżynierii genetycznej. Końcowym produktem tych badań ma być nowa generacja komputerów biomolekularnych, które mogą w przyszłości stanowić alternatywę dla tradycyjnych systemów komputerowych elektronicznych. Komputery takie mają pracować jako maszyny przepływowe, wykorzystując w swym działaniu reakcje biochemiczne. W tym przypadku mają być wykorzystane cząsteczki kwasów nukleinowych DNA i RNA, które w przyrodzie w organizmach żywych są podstawą pamięci genetycznej oraz najważniejszych procesów w biologii: transkrypcji i translacji, dzięki którym istnieje życie.

Wykorzystanie tych kwasów nukleinowych dla potrzeb techniki wnosi nowe spojrzenie i perspektywy na możliwości przetwarzania informacji. Rozwiązania takie mogą mieć istotne zalety, gdyż oferują one dużą masowość i równoległość obliczeń, które trudne są do osiągnięcia w tradycyjnych komputerach. Jako biokomputer DNA rozumiany jest tutaj odpowiednio dobrany zbiór molekuł i enzymów reagujących w określonych warunkach laboratoryjnych. Wykorzystuje się w tym celu cząsteczki DNA jednoniciowe lub dwuniciowe tak zwane oligonukleotydy oraz odpowiednie enzymy restrykcyjne i enzym ligazy. Poszczególne nukleotydy adenina, tymina, guanina i cytozyna oznaczane są odpowiednio literami A, T, G, C. Tworzą one alfabet czteroliterowy, z którego tworzone są łańcuchy DNA.

Doktorant w swojej rozprawie podchodzi do zagadnienia wnioskowania z punktu widzenia logiki. Dany jest zbiór zdań logicznych **A** problemem jest czy dane zdanie logiczne **p** można otrzymać ze zdań **A** za pomocą ciągu wnioskowań

korzystających z reguły odrywania (modus ponens). Następnie schemat ten jest odzwierciedlany za pomocą łańcuchów DNA i znanych operacji na nich.

Dzięki masowej równoległości przeprowadzanych operacji problemy o dużej złożoności przestają być niemożliwe do rozwiązania w skończonym czasie. Ważną zaletą tego podejścia jest również możliwość implementacji w nanotechnologii. Dlatego tematyka poruszana w tej pracy jest jak najbardziej aktualna, bardzo nowatorska, szeroka i mało jeszcze poznana.

W recenzowanej rozprawie nie sformułowano wyraźnie tezy pracy. Jakkolwiek określono główny jej cel, który brzmi: „przedstawienie nowej koncepcji systemu wnioskującego, która w przeciwieństwie do istniejących uwzględnia implementację większości praw klasycznego rachunku zdań i jego tautologii”. Tak więc cel ten został dostatecznie jasno sformułowany przez autora.

Rozprawa ma charakter teoretyczny. Autor wychodzi z klasycznego rachunku zdań logiki matematycznej i metodami obliczeń DNA próbuje stworzyć system do programowania w tej logice. Natomiast należy zauważyć, że w podejściu tym została całkowicie zaniedbana weryfikacja eksperymentalna w laboratorium inżynierii genetycznej. A przecież o przyszłym sukcesie tej dziedziny wiedzy będą decydować walory praktyczne, gdyż przyszłe komputery biomolekularne muszą rzeczywiście pracować i przewyższać ich odpowiedniki elektroniczne. Tłumaczenie autora we Wstępie, że „implementacja laboratoryjna omawianej koncepcji opiera się w całości na doświadczeniach przeprowadzonych już wcześniej i zaprezentowanych przez innych” badaczy nie jest przekonująca. Jak zaznacza autor w Podsumowaniu jego praca zdecydowanie rozszerza możliwości poprzednich systemów, wobec tego wiele koncepcji wprowadzonych przez niego (a stanowiących główny przyczynek tej rozprawy) nie zostało jeszcze zweryfikowanych eksperymentalnie, więc powstaje pytanie czy potwierdzą się one praktycznie, co zdaniem recenzenta jest bardzo istotne.

Praca składa się z siedmiu rozdziałów, bibliografii oraz dodatku. Pierwsze cztery rozdziały mają charakter kompilacyjny. Ich celem jest ukazanie stanu wiedzy w omawianej dziedzinie. Największą wartość poznawczą mają rozdziały 4 i 5, gdzie autor przedstawia swoją koncepcję rozszerzonego systemu wnioskowania oraz analizę tego systemu. Z kolei w dodatku zamieszczono zestaw łańcuchów DNA lepkich końców o długości 4.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł (w tym literatury światowej i stanu wiedzy na opisywany temat) świadcząca o dostatecznej wiedzy autora oraz czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Zamieszczony w rozprawie wykaz publikacji obejmuje 40 pozycji przy czym w dwóch autorem jest doktorant a w trzech – współautorem. Nie jest to dużo, zważywszy że dziedzina ta rozwijana jest od ponad dwudziestu lat i w związku z tym pojawiło się już na ten temat bardzo wiele różnych publikacji. Wśród zamieszczonych w rozprawie pozycji gro z nich dotyczy podstawowych koncepcji DNA obliczeń, które pozwalają nie zaznajomionemu z tą dziedziną czytelnikowi zorientować się z aktualnym stanem wiedzy. Temu celowi służy głównie rozdział 2, gdzie zawarto wprowadzenie do tej tematyki. Opisano podstawowe operacje przeprowadzane na oligonukleotydach DNA oraz przedstawiono modele filtrowania, splatania a także modele oparte na samoskładaniu i na membranach. Z drugiej strony w pracy zamieszczono też kilka

pozycji związanych z wnioskowaniem molekularnym oraz z implementacją systemów logicznych za pomocą DNA.

We Wstępie rozprawy przedstawiono ogólne tło rozważanych zagadnień oparte na pracach Adelmiana, Ogihary, Shapiro, Benensona, Rozenberga, Unolda, Trocia, Rodriguez-Patona, Sainz de Muriety, Lee, Parka i innych. Z kolei rozdział 3 zawiera przegląd znanych implementacji systemów wnioskowania. Dużo miejsca poświęcono systemowi Unolda i Trocia. System ten był główną inspiracją do napisania tej pracy. Wiele uwagi zwrócono na system wnioskowania Shapiro, Kaplana, Rana. W dalszej kolejności opisano koncepcję Lee, Parka, po czym Sakakibary, Suyamy, a następnie Rodriguez-Patona a także Sainz de Muriety.

Nigdzie w rozprawie nie odniesiono się do prac prof. Jana Mulawki, który wraz ze swoim zespołem już dwadzieścia lat temu prowadził badania w dziedzinie wnioskowania molekularnego we współpracy z czołowymi eksperymentatorami z inżynierii genetycznej (jak prof. Piotr Węgleński z Wydziału Biologii Uniwersytetu Warszawskiego w Warszawie oraz prof. Andrzej Płucienniczak z Instytutu Biotechnologii i Antybiotyków w Warszawie) i miał wiele osiągnięć o zasięgu międzynarodowym.

W związku z powyższym przedstawioną analizę literaturową uznać należy za nie wyczerpującą i dlatego należy ją rozszerzyć przynajmniej o pozycje wskazane autorowi przez recenzenta.

3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Zdaniem recenzenta tego typu rozprawa powinna składać się z dwóch części - teoretycznej i eksperymentalnej. W części teoretycznej autor rozwiązał postawione zagadnienia a metody i założenia wydają się właściwe. Natomiast nie jest uzasadnione założenie autora, że można nie wykonywać sprawdzenia doświadczalnego w laboratorium inżynierii genetycznej a jedynie odwołać się do prac eksperymentalnych innych badaczy.

W swojej koncepcji rozszerzonego molekularnego systemu wnioskującego autor wykorzystał prace Unolda i Trocia a także Kaplana, Rana, Shapiro. Podejście to opiera się na sprawdzonym już modelu splatania, które polega na naprzemiennym działaniu enzymu restrykcyjnego oraz ligazy. W tym przypadku użyty został enzym BseXI, który rozcina dwuniciowe łańcuchy DNA pozostawiając zawsze lepkie końce o długości 4.

Działanie proponowanego rozwiązania opiera się na koncepcji tak zwanego biomolekularnego systemu autonomicznego, a więc takiego którego działanie nie wymaga udziału laboranta w trakcie obliczeń. W rozwiązaniu takim reakcje zachodzą sekwencyjnie, samoczynnie podobnie jak obliczane są instrukcje w czasie wykonywania programu komputerowego. Całość reakcji sprowadza się tutaj do trzech kroków: 1) przygotowanie molekuł wejściowych, 2) właściwa reakcja biochemiczna, 3) odczytanie wyniku. Tak więc działanie takiego systemu jest właściwie pewnego rodzaju programowaniem ale wykonywanym w oparciu o reakcje biochemiczne a nie słowa binarne.

Do kodowania informacji w proponowanym podejściu wykorzystuje się dwuniciowe cząsteczki DNA. Pojedyncze literały reprezentowane są za pomocą unikalnych sekwencji o długości czterech nukleotydów, wyznaczanych w taki sposób, że literały reprezentujące zmienną atomową i jej negację są do siebie

komplementarne. Z kolei fakty z jednym literałem wykorzystują molekuly zawierające w sobie sekwencję 3'-CGTCG-5' rozpoznawalną przez enzym restrykcyjny BseXI. Natomiast fakty o dwóch literałach wykorzystują molekuly zawierające w sobie sekwencję 3'-GTCGCGAC-5' rozdzielającą łańcuchy reprezentujące literały. W procesie wnioskowania wykorzystywana jest jeszcze dwuniciowa molekula terminalna o charakterystycznej strukturze. Do zadania zapytania wykorzystuje się specjalną molekulę dwuniciową, która zawiera unikalną czteroelementową sekwencję komplementarną z samym sobą.

Odpowiedź systemu rozróżnia się po długości powstałych cząsteczek za pomocą elektroforezy w żelu. Interesująca jest reakcja wykrycia sprzeczności. W tym przypadku powstaje molekula nie zawierająca lepkich końców oraz unikalnej długości 104 nukleotydów.

W dalszej części autor demonstruje różne aplikacje systemu w odniesieniu do klasycznego rachunku zdań. Pokazuje implementacje wielu praw rachunku zdań. Różne operacje wykonywane podczas działania systemu są ilustrowane licznymi przykładami reakcji na cząsteczkach DNA. Przedstawione reakcje przebiegają w sposób właściwy.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Oryginalność rozprawy polega głównie na rozszerzeniu molekularnych systemów wnioskujących wymienionych w poprzednim punkcie tej recenzji, dzięki czemu możliwa stała się implementacja większości praw klasycznego rachunku zdań i jego tautologii.

Do najważniejszych osiągnięć rozprawy, stanowiących samodzielny i oryginalny dorobek autora, należy zaliczyć opracowanie rozwiązania odznaczającego się następującymi zaletami:

- aksjomaty początkowe, wyciągane wnioski czy zapytania do systemu mogą być w formie koniunkcji czy alternatyw,
- system umożliwia wprowadzić pojęcie negacji wraz z możliwością wyciągania wniosków również na podstawie literałów o wartości fałszu jak i wniosków o wartości fałszu,
- w trakcie wnioskowania korzysta się z praw klasycznego rachunku zdań, które są naturalną konsekwencją wprowadzenia spójnika negacji, między innymi regułę kontrapozycji czy prawa de Morgana,
- system sygnalizuje o ewentualnym pojawieniu się sprzeczności niezależnie czy wynikła ona z błędnie dobranych aksjomatów początkowych czy powstała w trakcie reakcji, w wyniku kolejnych wnioskowań.

Aby osiągnąć wymienione korzyści, doktorant wprowadził kilka innowacji związanych z budową cząsteczek DNA używanych podczas procesu wnioskowania. Wymienić tu należy następujące propozycje:

- zastosowanie cząsteczek symetrycznych, co pozwoliło na wykonanie różnych przekształceń tej samej molekuly,
- fragment rozpoznawalny przez enzym restrykcyjny umieszczony został w dwóch końcach tej samej cząsteczki, dzięki temu cięcie w obrębie takiej molekuly może pojawić z obu jej stron,

-- cząsteczki mogą posiadać lepkie końce dobrane w taki sposób, że są komplementarne z samymi sobą i w wyniku połączenia dwóch takich samych molekuł tworzą cząsteczkę bez lepkich końców, która oznacza konkretną odpowiedź na zadane pytanie.

Poza tym w rozprawie przedstawiono dwa interesujące zastosowania opracowanego systemu wnioskowania. Są to częściowe rozwiązanie problemu SAT oraz obliczanie wartości funkcji logicznych.

Zdaniem recenzenta pozycja rozprawy w stosunku do stanu wiedzy jest znacząca. Wyniki tych prac zostały opublikowane w pięciu artykułach przy czym dwa w czasopismach o zasięgu międzynarodowym. Niepokój budzi jednak brak jakiegokolwiek weryfikacji doświadczalnej, co znacznie obniża walory tej pracy. Recenzent ma jednak nadzieję, że eksperymenty potwierdzą przyjęte przez doktoranta założenia.

Reasumując tą część recenzji należy stwierdzić, że przedstawione w pracy wyniki stanowią istotne rozszerzenie stanu wiedzy w dziedzinie wnioskowania molekularnego w oparciu o cząsteczki DNA.

5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia wyników (zwięzłość, jasność, poprawność redakcyjna)?

Recenzowana rozprawa liczy 96 stron i ogólnie zredagowana jest w sposób jasny, zwięzły i poprawny językowo. Treść rozprawy jest ściśle powiązana z przedmiotem badań. Przyjęty układ treści oraz jej podział pomiędzy rozdziały wydaje się właściwy. Uzasadnienia prowadzące do wyprowadzenia rezultatów różnych koncepcji są przekonujące. Autor potrafi krótko i syntetycznie omówić złożone zagadnienie.

Podczas czytania pracy nie natrafiono na większe uchybienia redakcyjne. Znalaziono kilka błędów literowych, które zaznaczono w tekście rozprawy.

Rozdział drugi poświęcony jest wprowadzeniu do obliczeń DNA. Ma on charakter dydaktyczny, gdyż wyjaśnia trudne operacje laboratoryjnego działania na DNA. Dzięki temu niewątpliwą zaletą pracy jest jej łatwe zrozumienie przez czytelnika. Wynika to także stąd, że omawiane operacje wnioskowania są szeroko ilustrowane reakcjami na przykładowych molekułach DNA. Szkoda, że omawiane reakcje nie zostały potwierdzone za pomocą eksperymentów w laboratorium inżynierii genetycznej.

Pewnym mankamentem redakcji pracy jest brak przejrzystego sformułowania wniosków analizy teoretycznej dla różnych operacji wnioskowania w postaci wydzielonych, ponumerowanych części tekstu.

6. Jakie są słabe strony rozprawy oraz jej główne wady?

Główną słabością rozprawy jest ograniczenie jej tylko do części teoretycznej, rezygnując z jakichkolwiek eksperymentów laboratoryjnych. Wzmiankowano już o tym wielokrotnie w tej recenzji przy różnych okazjach. Praca przecież jest z pogranicza logiki oraz genetyki molekularnej, gdzie kluczową rolę odgrywają doświadczenia wykonywane właśnie w laboratorium inżynierii genetycznej. Skoro w tytule są słowa implementacja systemu wnioskującego za pomocą DNA, to należało pokazać, że taki system rzeczywiście działa! Poza tym doktorant wspomina w Podsumowaniu „ze względu na techniczne ograniczenia narzędzi biologicznych,

system ten wciąż nie odwzorowuje pełnej zgodności z logiką matematyczną", więc powstaje pytanie, jak się on w rzeczywistości zachowa, implementując go w typowym laboratorium. Tak więc usprawiedliwienie się autora w Podsumowaniu o nie przeprowadzeniu weryfikacji eksperymentalnej nie jest zasadne.

Innym mankamentem pracy jest rozważenie zbyt skromnej bibliografii (tylko 40 pozycji). Brakuje na przykład wielu artykułów napisanych przez recenzenta i jego zespół. Jak zaznaczono w tej recenzji wcześniej, zajmował się on zagadnieniami wnioskowania metodami obliczeń DNA już dwadzieścia lat temu, przez co publikacje te są mniej znane dla szerszego kręgu badaczy.

W swojej rozprawie doktorant pokazał podstawowe operacje (reakcje biochemiczne) działania różnych elementów systemu, natomiast brak jest rozważenia wnioskowania w przypadku większej bazy wiedzy.

Chociaż w rozprawie w rozdziale 5.1 rozważa się niektóre ograniczenia systemu, to jednak można by było dogłębniej rozpatrzeć właściwości proponowanego systemu z punktu widzenia ograniczeń laboratorium DNA. Brakuje też porównania z innymi podobnymi rozwiązaniami.

Niektóre pojęcia, wzory i rysunki w rozprawie powtarzają się (na przykład na stronach 33 i 71, rysunki Rys. 91f i Rys.93d etc.).

Nie wprowadzono numeracji wzorów. Podobnie nie oznaczono numerami dwóch rozdziałów Wstępu i Podsumowania.

Rozdział 1 Wiadomości wstępne z logiki matematycznej nic istotnego do przedmiotu rozprawy nie wnosi, dlatego zdaniem recenzenta powinien on być usunięty. W ewentualnych odwołaniach wykorzystywać pozycję literaturową A. Grzegorzcyka.

W świetle powyższych uwag krytycznych zdaniem recenzenta powinien być wprowadzony nowy rozdział na przykład pod tytułem Sprawdzenie eksperymentalne. Z kolei rozdział 3. Znane implementacje systemów wnioskowania powinien być znacznie rozwinięty.

7. Jaka jest przydatność rozprawy dla nauk technicznych?

Przedstawiony przez doktoranta system wnioskujący uwzględnia implementację większości praw klasycznego rachunku zdań i jego tautologii. Stanowi on istotny przyczynek w dziedzinie programowania w logice w oparciu o molekuly DNA. Praca rozszerza więc ogólną wiedzę na temat możliwości wykorzystania zdobyczy inżynierii genetycznej dla potrzeb obliczeń DNA. Stwarza to nadzieję, że może powstać w przyszłości nowa generacja komputerów biomolekularnych, przepływowych, wykorzystujących reakcje chemiczne na kwasach nukleinowych o konkurencyjnych właściwościach w stosunku do klasycznych komputerów elektronicznych, których możliwości dalszego rozwoju są ograniczone.

Należy zaznaczyć, że komputery biomolekularne (szczególnie autonomiczne, a do takich zalicza się system proponowany przez doktoranta) mają przed sobą dużą potencjalną przyszłość ze względu na ich masowość i równoległość obliczeń. Cząsteczki DNA mogą być wykorzystane do rozwiązywania problemów algorytmicznych, matematycznych czy logicznych; a same obliczenia przeprowadzane są w odpowiednich warunkach laboratoryjnych za pomocą operacji biochemicznych na nich. Tematyka ta jest więc bardzo interesująca, szeroka oraz otwarta, a zarazem mało jeszcze poznana i wymaga wobec tego współpracy badaczy z różnych kierunków naukowych a zwłaszcza informatyki, logiki, biologii,

biochemii, elektroniki, fizyki. Dlatego wszelkie badania w dziedzinie DNA obliczeń mogą mieć dużą przydatność dla nauk technicznych i powinny być dalej kontynuowane.

8. Do której z następujących kategorii recenzent zalicza rozprawę:

- a) Nie spełniającej wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy,
- b) Wymagająca wprowadzenia poprawek i ponownego recenzowania,
- c) Spełniająca wymagania,
- d) Spełniająca wymagania z wyraźnym nadmiarem,
- e) Wybitnie dobra, zasługująca na wyróżnienie.

Biorąc pod uwagę przytoczone w powyższej recenzji argumenty, zaliczam pracę do kategorii b), a więc wymagającej wprowadzenia poprawek i ponownego recenzowania.

