

## RECENZJA

rozprawy doktorskiej mgra Łukasza Rogowskiego  
„Implementacja systemu wnioskującego za pomocą DNA”  
Promotor: prof. dr hab. inż. Tadeusz Krasieński

### 1. Tematyka rozprawy

Obliczenia DNA (ang. DNA computing) stanowią dzisiaj prężnie rozwijającą się dziedzinę leżącą na pograniczu informatyki, matematyki, biologii i chemii. Odkąd prawie 30 lat temu przedstawiono teoretyczne możliwości obliczeń z wykorzystaniem łańcuchów molekularnych, dziedzina ta rozwinęła się w kilka równoległych wątków (jak komputery DNA, molekularne systemy wnioskowania, programowanie logiczne z DNA, biosensory i inne) i wchodzi dzisiaj w skład tzw. nanonauki, obejmującej również komputery kwantowe. To właśnie w obliczeniach biomolekularnych i komputerach kwantowych upatruje się dzisiaj główny sposób na przełamanie ograniczeń stawianych przez szeroko rozumiane technologie krzemowe. Przełomowy dla obliczeń na molekułach okazał się rok 1994 i słynny dzisiaj eksperyment Adlemana, chociaż należy stwierdzić, że raczej mało kto już dzisiaj upatruje zastosowanie biokomputerów w znajdowaniu rozwiązań dla trudnych obliczeniowo problemów. To raczej molekularny poziom tych obliczeń (DNA, RNA, białka) stanowi magnes przyciągający naukowców. Autonomiczne, molekularne, reagujące na obecność określonych ciągów molekularnych i działające bezpośrednio w organizmie bioautomaty czy też bioczujniki zdają się być główną siłą napędową tej dziedziny.

Recenzowana rozprawa podejmuje właśnie tematykę konstrukcji molekularnego systemu wnioskującego, który w zamierzeniu ma istotnie rozszerzyć zbiór tautologii realizowanych przez dostępne obecnie tego typu systemy. Tak postawiony cel rozprawy wpisuje się bardzo dobrze w aktualny nurt badań w dziedzinie biomolekularnych obliczeń, co więcej, zwiększając też wkład polskiej nauki w rozwój tej nowoczesnej dziedziny.

W świetle powyższych uwag można stwierdzić, że **problematyka rozprawy jest ważna i w pełni uzasadniona.**

## 2. Kompozycja, zawartość i redakcja rozprawy

Recenzowana rozprawa składa się z pięciu rozdziałów poprzedzonych Wstępem i zakończonych Podsumowaniem, Dodatku, Bibliografii i obejmuje 96 numerowanych stron. Nie zawiera, a szkoda, spisu rysunków, wykazu skrótów i oznaczeń i/lub indeksu ważniejszych pojęć,

W krótkim Wstępie Doktorant zarysowuje dosłownie problematykę klasycznego rachunku zdań, formułuje nieformalnie cel pracy, by zaraz po tym przejść do tematyki tzw. obliczeń DNA, czy może formułując to precyzyjniej obliczeń realizowanych za pomocą ciągów DNA. Obliczenia DNA są wskazane jako alternatywny do obecnych systemów komputerowych model obliczeń, który podobnie jak obliczenia na procesorach graficznych czy komputerach kwantowych umożliwia przekroczenie bariery stawianej przez prawo Moore'a. Doktorant krótko opisuje historię obliczeń DNA, zaczynając od klasycznego już dzisiaj eksperymentu Adlemana z 1994 roku (choć może zasadniej byłoby wspomnieć na początku o pracy Heada z 1987 roku – cytowanej zresztą w dysertacji na innym miejscu), następnie wymieniając komputer Shapiry oraz kilka innych znanych rozwiązań DNA komputerów, wszystkie zresztą oparte na enzymach restrykcyjnych. Jako alternatywna gałąź obliczeń na cząsteczkach DNA wskazane są systemy logiczne, w tym ważne dla dysertacji prace współautorstwa Recenzenta i zespołu Shapiry. Po tym passusie następuje druga próba zdefiniowania celu rozprawy, po której Doktorant wspomina bardzo interesującą i stosunkowo świeżą próbę modelowania wnioskowania za pomocą obliczeń DNA niewykorzystujących enzymy restrykcyjne (czego Doktorant zresztą nie zaznacza). Wstęp kończy wymienienie autorskich publikacji, schemat dysertacji i podziękowania.

W rozdziałach 1 i 2 Autor wprowadza podstawowe informacje na temat, odpowiednio wnioskowania formalnego i obliczeń DNA. W rozdziale 1 uwaga koncentruje się na klasycznym rachunku zdań, formalnych regułach wnioskowania i – podrzędnych w istocie do poprzednich sekcji – postaci normalnych logicznych zdań złożonych. Rozdział 2 podaje podstawowe informacje na temat struktury i operacji na ciągach DNA, w bardzo krótkiej sekcji dotyczącej modeli biokomputerów podany jest podział na systemy autonomiczne i nieautonomiczne, w kolejnej opisany jest skrótowo eksperyment Adlemana, w sekcji kolejnej wprowadzono klasyczny dla obliczeń DNA model splatania, wymieniając m.in. programowalny komputer/automat Shapiry i prezentując zasadę jego działania, rozszerzenia tego modelu (zespołów Unolda, Shapiry, Krasińskiego) oraz oparte na splataniu modele wnioskowania (zespołów Unolda i Shapiry). Rozdział kończy pojedynczy akapit ujęty w

oddzielną sekcję, a wspominający o pozostałych modelach bimolekularnych obliczeń, to jest bazujących na samoskładaniu i membranach. Sekcja ta niewątpliwie nie wyczerpuje tematu.

W rozdziale 3 Doktorant szerzej opisuje znane sobie modele wnioskowania na ciągach DNA. Rozdział zaczyna od prezentacji nie tyle samego modelu, co raczej zawartości publikacji z 2005 roku autorstwa Unolda i Trocia, która – jak Autor podkreśla – stanowiła jedną z głównych inspiracji rozprawy. W dwóch sekcjach omówione są: system wnioskowania, z godną podkreślenia własną listą krytycznych uwag odnośnie systemu, oraz model wartościowania funkcji logicznych. Również ta druga koncepcja doczekała się krytycznej analizy Autora. W kolejnym podrozdziale opisane są modele wnioskujące zespołu Shapiry z 2009 roku. W pierwszym możliwa jest identyfikacja przesłanek prowadzących do wniosku (implementacja wnioskowania w tył), w drugim zrealizowano wnioskowanie w przód. Obydwa modele zweryfikowano laboratoryjnie, obydwaj wykorzystują zjawisko fluorescencji do weryfikacji wyniku. Na uwagę zasługuje autorskie porównanie możliwości systemów wnioskujących opracowanych przez zespoły Unolda i Shapiry. Kolejny podrozdział prezentuje ostatnią już koncepcję, mającą wpływ na rozprawę. Została ona zawarta w publikacji In-Hee Lee i innych z 2003 roku. Interesującą jej własnością jest porzucenie restryktazy w modelu obliczeń na rzecz filtrowania oraz zrealizowanie dowodu nie wprost. Podobnie jak w poprzednich dwóch przypadkach i ta koncepcja doczekała się krytycznej analizy Doktoranta. Rozdział kończy omówienie pozostałych koncepcji, które sprowadza się do zacytowania dwóch podejść.

Kolejne dwa rozdziały, tj. 4 i 5, zawierają oryginalne wyniki badań Doktoranta i stanowią zasadniczą część dysertacji. W rozdziale 4 zaprezentowana jest nowa koncepcja systemu wnioskującego na ciągach DNA. Proponowany system indukcji oparty jest o splatanie, enzymy ligazy i restryktazy *BseXI*, ściśle określony molekularny zapis aksjomatów początkowych (literałów, klauzul), zawsze w koniunkcyjnej postaci normalnej, wreszcie molekułę terminalną i mechanizm wnioskowania *modus ponens* (regułę odrywania). Zapytania do systemu indukcji mogą dotyczyć klauzuli faktu zawartej w zbiorze aksjomatów początkowych lub możliwej do wywiedzenia, koniunkcji bądź alternatywy kilku faktów, zdania zapisanego w dysjunkcyjnej postaci normalnej, a nawet kilku zdań jednocześnie. Przyjęta reprezentacja molekularna pozwala również na zastosowanie spójnika negacji, a opisany w rozdziale system wnioskowania zdolny jest do wykrycia sprzeczności, tak w zbiorze aksjomatów początkowych, jak i jako wynik pośrednich operacji odrywania. W krótkim podrozdziale opisana jest ogólna zasada działania systemu indukcji (szkoda, że w tym miejscu nie podano również ogólnych zasad projektowania reprezentacji molekularnej,

które wyszczególniono dopiero w kolejnym rozdziale), po którym podano dwa przykłady „złożonego” wnioskowania. W ostatnim podrozdziale (nie licząc dwuakapitowego podsumowania) Doktorant wskazuje (a może raczej rekapitułuje) w jaki sposób jego system indukcji logicznej realizuje wybrane prawa klasycznego rachunku zdań, a precyzyjniej prawo podwójnego przeczenia, regułę odrywania, kontrapozycji, przechodności implikacji i prawa de Morgana.

Rozdział 5 zawiera analizę zaproponowanego systemu wnioskowania realizującego obliczenia za pomocą ciągów DNA. Rozdział rozpoczyna opis ograniczeń systemu wynikających z przyjętych założeń. W pierwszej sekcji podrozdziału Doktorant precyzyjnie wskazuje liczbę zmiennych możliwych do zakodowania w założonej metodzie reprezentacji molekularnej. Obliczenia pokazują, że system ograniczony jest do 65 zmiennych (wskazanych w Dodatku do pracy – tu pytanie, dlaczego w spisie jest 120 ciągów, a nie 130), choć zdaniem Recenzenta realna liczba (równoczesnych) zmiennych, możliwa do implementacji w środowisku laboratoryjnym, zapewne jest o rząd mniejsza. Podobna analiza przeprowadzona jest dla teoretycznej liczby zapytań w systemie, również ze wskazaniem Dodatku, w którym wymieniono *explicite* 16 ciągów DNA. Dalej, wskazane są ograniczenia dla przypadku, gdy pytamy jednocześnie o fakt i jego negację (ograniczenia dosyć naturalne praktycznie w każdym systemie wnioskującym), a także, gdy pytanie dotyczy bardziej złożonych zapytań, w których występują komplementarne wartości. Wskazanie zastosowania większej liczby próbek zdaje się tutaj zasadne. Kolejne ograniczenia dotyczą zapytań o klauzule złożone oraz wnioskowania faktu ze zbioru alternatyw. Wskazane jest zastosowanie dowodu nie wprost, jak pokazują kolejne sekcje, możliwego teoretycznie w realizacji. Następny podrozdział zawiera ciekawą dyskusję możliwych dodatkowych zastosowań wprowadzanego systemu indukcji molekularnej. Pokazane są możliwości wartościowania funkcji logicznych, w tym również, gdy tych funkcji jest więcej. Naturalną konsekwencją możliwości obliczania wartości logicznych dla zadanych argumentów wykorzystana jest w tzw. częściowym rozwiązaniu problemu SAT. Ta częściowość rozwiązania polega na możliwości nie pojawienia się w procesie wnioskowania sprzeczności, informującej o braku spełnialności reguły dla jakichkolwiek wartości literałów. W takim wypadku obliczenia należy prowadzić innymi metodami. Ostatni podrozdział dotyczy możliwych rozszerzeń, w tym przeprowadzania dowodu nie wprost, z takim jednak ograniczeniem, że nie możemy zadawać w jednej próbce więcej niż jednego zapytania. Ciekawym podrozdziałem jest analiza możliwości zastosowania innych enzymów restrykcyjnych we wnioskowaniu. Dopiero w tej sekcji pojawia się opis zasad projektowania operacji molekularnych w

zaproponowanym systemie wnioskowania. Ostatnim wskazanym rozszerzeniem jest zastąpienie elektroforezy na żelu fluorescencją.

Rozprawę zamyka Podsumowanie punktujące przewagę koncepcji Doktoranta nad pracami Recenzenta i Shapiry, a także podkreślające zastosowane innowacje. Całość kończy wyjaśnienie dlaczego nie przeprowadzono badań laboratoryjnych.

**Kompozycja i struktura rozprawy nie budzą większych zastrzeżeń.** Wątpliwość budzić może przyjęty tytuł rozprawy. Implementacja systemu wnioskującego (i nie tylko) dosyć jednoznacznie kojarzy się (przynajmniej w inżynierii oprogramowania) z jego uruchomieniem i przetestowaniem. Tego jednak w pracy nie znajdziemy, co najwyżej odesłanie w kilku miejscach rozprawy do podobnych (tylko!) badań laboratoryjnych innych autorów. Zdecydowanie bardziej przekonująco brzmi wyjaśnienie wskazujące na czasochłonność i koszty badań laboratoryjnych, a podobieństwo – przynajmniej zdaniem Recenzenta – nie ma tutaj zastosowania. Z tego też powodu może mniej obiecującym, ale na pewno bardziej precyzyjnym tytułem rozprawy, co więcej użytym przez Autora chociażby na str. 29 czy 91, byłby zapis o koncepcji rozszerzonego molekularnego systemu wnioskującego.

Kolejną uwagą jest brak jasno i klarownie podanego celu, zakresu i tez(y) rozprawy. Jeżeli przyjmiemy, że celem rozprawy jest udowodnienie postawionej tezy, to cel i tezę pracy można znaleźć w dwóch miejscach Wstępu, natomiast zakres pracy, to jest czego się Autor podejmuje, a czego nie w celu udowodnienia tezy, ze Wstępu do dysertacji już nie wynika i jest do znalezienia dopiero w rozdziale 4.

Język rozprawy jest poprawny, choć Autor nie ustrzegł się stosunkowo licznych błędów interpunkcyjnych. Pojawiają się też błędy gramatyczne polegające na niewłaściwym używaniu słowa ilość z rzeczownikami policzalnymi (wszystkie błędy zaznaczono w otrzymanym egzemplarzu).

Bibliografia zawiera 40 numerowanych pozycji, z których Doktorant jest autorem lub współautorem 5 publikacji, w tym 1 preprintu, 1 pracy konferencyjnej, 2 artykułów ulokowanych w polskich czasopismach oraz 1 opublikowanego w czasopiśmie zagranicznym (z zagranicznym współautorem). Zastanawiać może niewielka w spisie liczba publikacji z ostatnich 2-3 lat. Owszem, Doktorant bazuje przede wszystkim na dwóch koncepcjach opublikowanych w 2005 i 2009 roku, ale dziedzina, w której się porusza dynamicznie się rozwija, przynosząc co raz to nowe propozycje tak teoretyczne, jak i laboratoryjne. Być może wynika to faktu skupienia się na mechanizmie splatania, który ostatnio cieszy się jakby mniejszą "popularnością" wśród badaczy, na korzyść chociażby obliczeń bez użycia restryktaz, na przykład „DNA strand displacement” (w samym roku 2006 google scholar

notuje prawie 300 publikacji na ten temat). Należy jednak oddać Autorowi, że w spisie artykułów są publikacje na ten temat, niestety jednak, na wspomnienie mechanizmu DSD nie znalazło się miejsce w rozprawie.

### 3. Oryginalne osiągnięcia Autora

Teza rozprawy, którą można za Autora ująć w sformułowaniu, że jest możliwy od strony teoretycznej rozszerzony w stosunku do doniesień literaturowych system wnioskujący oparty o mechanizm splatania, wymagał od Doktoranta wykonania nietrywialnych badań teoretycznych.

Do najważniejszych i oryginalnych wyników przedstawionych w recenzowanej rozprawie należy zaliczyć:

- Opracowanie koncepcji w pełni autonomicznego molekularnego systemu wnioskującego, istotnie rozszerzającego możliwości analogicznych systemów opisywanych w literaturze. Rozszerzenie wprowadzone jest zasadniczo przez spójnik negacji i związany z nim cały zestaw praw rachunku zdań. Dodatkowo, proponowany system m.in. umożliwia zadawanie pytań o fakt, alternatywę lub koniunkcję faktów, dowolne zdanie logiczne zapisane w formie dysjunkcyjnej oraz zbiór zdań. Zdolność systemu do wykrywania sprzeczności prowadzi do kolejnego, poza regułą odrywania, mechanizmu wnioskowania, czyli dowodu nie wprost.
- Zastosowanie opracowanego systemu wnioskowania do zadania wartościowania funkcji logicznych i częściowego rozwiązywania problemu  $k$ -SAT.
- Opracowanie reguł konstrukcji molekuł dla potencjalnych, nowych restryktaz.
- Wykazanie możliwości zastąpienia elektroforezy na żelu zjawiskiem fluorescencji.

Dodatkowo, w rozprawie uzyskano następujące oryginalne rezultaty badawcze, jednak o mniejszej randze i funkcji pomocniczej w stosunku do dowodzonej tezy:

- Krytyczna analiza literaturowych systemów wnioskujących, stanowiących podstawę wprowadzonego systemu wnioskowania.
- Porównanie krytyczne modeli wnioskujących zespołów Recenzenta i Shapiry.
- Zastosowanie cząstek symetrycznych, umożliwiających różne przekształcenie tej samej klauzuli.

Zgodnie z powyższym można sformułować wniosek, że uzyskane wyniki teoretyczne potwierdzają tezę rozprawy.

#### 4. Uwagi krytyczne i dyskusyjne

Podczas lektury rozprawy nasunęły się pewne uwagi, niżej wymienione i w większości już wcześniej poruszone:

- Zasadniczą uwagą krytyczną, wzmiankowaną już wcześniej, jest brak eksperymentalnego potwierdzenia opisanej koncepcji. Dopiero wyniki laboratoryjne mogą w pełni udowodnić poprawność stawianych założeń i wskazać rzeczywiste ograniczenia modelu, szczególnie w dziedzinie obliczeń molekularnych. Weryfikacja eksperymentalna otwiera też możliwości umieszczenia pracy w prestiżowych czasopismach z dziedziny.
- Kolejna uwaga, związana z poprzednią i również już wyartykułowana, dotyczy zbyt szerokiego tytułu rozprawy, obiecującego więcej, niż tylko koncepcję.
- Brak właściwej postawionej tezy dysertacji, jak i jej zakresu również został już w recenzji poruszony.
- Przegląd literatury z dziedziny jest ograniczony głównie do tych pozycji, które stosują podobne do prezentowanej koncepcji mechanizmy. Szkoda też, że na liście Bibliografii nie znalazły się także prace polskiego zespołu prof. Muławki.
- W dysertacji jest wzmianka o ciekawym, opartym na innych założeniach, systemie wnioskowania rozwijanym na Politechnice w Madrycie. Interesującym byłoby porównanie możliwości obydwu systemów (patrz dostępny w Internecie doktorat Muriety z 2013 r.).
- Zasadna wydaje się też budowa symulatora obliczeń, który mógłby ułatwić projektowanie wnioskowania i jednocześnie usunąłby możliwe przy ręcznym wykonywaniu obliczeń błędy. W połączeniu z generalnymi zasadami dotyczącymi stosowania innych restryktaz (patrz rozdz. 5.3.2), narzędzie to mogłoby stanowić ciekawą alternatywą do istniejących narzędzi i formalizmów dla mechanizmu DSD.

Poniżej wymieniono uwagi natury szczegółowej:

- Błąd edycyjny – zbędny nawias w 6 linii od dołu na str. 8.
- Na str. 11 wymieniane są dwie reguły wnioskowania stosowane w rachunku zdań. Dlaczego nie wymienia się tam reguły *modus tollens*?

- Linki do źródeł internetowych na str. 13 nie mają podanych dat dostępu.
- Schemat obliczeń, które wykonał Adleman, a opisany na str. 21 jest nieprecyzyjny, gdyż nie wystarczy sprawdzenie długości ścieżki, w celu wykazania, że wszystkie kodowane wierzchołki na ścieżce się nie powtarzają.
- Błąd edycyjny w legendzie do rys. 49b.
- Błąd edycyjny (braku kursywy) w zapisie zmiennej  $n$  na str. 45.
- Błędne odwołanie do rysunku na str. 51 – jest podany rys. 73b, powinno być 74b.
- Wnioskowanie przedstawione na str. 54 ma z założenia dowodzić poprawności alternatywy. Przeprowadzone wnioski dowodzi w istocie jedynie poprawności tylko jednego z członów alternatywy.
- Błąd w legendzie do rys. 91g – jest  $x?$ , powinno być  $b?$ .
- Błąd w legendzie do rys. 92a – jest  $a$ , powinno być  $\neg d$ .
- Niepotrzebne wielkie litery w tytule podrozdziału 4.6.
- Błąd na rys. 95a – w pierwszej linii jest (q), powinno być (q).
- Błąd na rys. 95b – w pierwszej linii jest (q), powinno być (q).
- Błędne odwołania do podrozdziału na str. 71 – jest 5.2.1, powinno być 5.3.1.
- Błąd w legendzie do rys. 129a – jest  $\neg r$ , powinno być  $\neg p$ .
- Błąd edycyjny na str. 85, legenda do rys. 130b rozdzielona od rysunku.
- Błąd zapisu na rys. 132 i następnych. Druga linia powinna mieć, posługując się notacją zaproponowaną przez Doktoranta, sekwencje STARTER, ENZ, PM, PYT, zapisywane jako negacja komplementarnej sekwencji z pierwszej linii.
- Błąd edycyjny - pomiędzy wymienianymi referencjami literaturowymi brakuje koniecznej spacji.
- Doktorant konsekwentnie nie stosuje w tekście pracy odsyłaczy do rysunków.

**Wszystkie wymienione wyżej krytyczne uwagi i spostrzeżenia nie podważają pozytywnej oceny pracy.**

## **5. Wniosek końcowy**

Zawarte w recenzji mają na ogół charakter dyskusyjny, a te krytyczne nie wpływają na pozytywną ocenę recenzowanej dysertacji. Uważam, że rozprawa mgra Łukasza Rogowskiego zdecydowanie spełnia wymogi stawiane rozprawom doktorskim. Recenzowana dysertacja niewątpliwie stanowi oryginalne rozwiązanie postawionego problemu naukowego,



P 1 0 0 0

wprowadzając bardzo obiecującą koncepcję molekularnego systemu wnioskowania. Doktorant wykazał się właściwą wiedzą w dziedzinie nauk matematycznych, w dyscyplinie naukowej Informatyka, dowiódł też opanowania umiejętności samodzielnego prowadzenia pracy naukowej.

Biorąc pod uwagę powyższe fakty stwierdzam, że recenzowana rozprawa w pełni spełnia wymagania odpowiedniej Ustawy i **wnioskuję o dopuszczenie mgra Łukasza Rogowskiego do dalszych etapów postępowania o nadanie stopnia doktora.**

Olgierech Unala