

WNIOSKOWANIE LOGICZNE ZA POMOCĄ DNA. REGUŁA KONTRAPOZYCJI.

Łukasz Rogowski

Uniwersytet Łódzki
Wydział Matematyki i Informatyki
ul. Banacha 22/ p.A227, 90-238 Łódź
e-mail: rogowski@math.uni.lodz.pl

Streszczenie: *Obliczenia DNA (z ang. DNA computing) to nowa dziedzina informatyki, będąca alternatywą dla tradycyjnych systemów komputerowych, polegająca na wykorzystaniu cząsteczek molekularnych do rozwiązywania problemów algorytmicznych, matematycznych i logicznych. Niniejszy artykuł przedstawia istniejące implementacje systemów wnioskowania realizowanych za pomocą DNA oraz nową koncepcję takiego systemu, uwzględniającą nowe elementy: negację i regułę kontrapozycji w implikacji.*

Słowa kluczowe: *Obliczenia DNA, komputer DNA, logika, wnioskowanie, implikacja, implementacja, reguła kontrapozycji*

Logical deduction with DNA. The rule of contraposition.

Abstarct: *DNA computing is one of new computational paradigms which are alternative to traditional computer systems. Biological molecules in special laboratory conditions can be used to solve mathematical, logical and algorithmic purposes. This paper describes some already existing models of logical inference systems and the new proposal – DNA deduction system using new elements: negation and rule of contraposition..*

Keywords: *DNA computing, computer DNA, logic, deduction, implication, implementation, rule of contraposition.*

1. WSTĘP

DNA (kwas deoksyrybonukleinowy) to polimer (substancja chemiczna składająca się z wielokrotnie powtórzonych jednostek), którego elementami są nukleotydy. Są one zbudowane z reszty cukrowej (deoksyrybozy), grupy fosforanowej i zasady azotowej. Dwie pierwsze części są zawsze takie same, natomiast między sobą nukleotydy różnią się zasadami (występują cztery różne: adenina (A), cytozyna (C), guanina (G) i tymina (T), w związku z czym nazwy nukleotydów łączą się z nazwami zasad).

Z punktu widzenia informatyki, chcąc wykorzystać DNA jako nośnik danych, mamy do dyspozycji cztery różne symbole czyli dwa razy tyle co w tradycyjnym zapisie

bitowym (gdzie korzystamy z zer i jedynek). Mniejsze są możliwości manipulowania tymi danymi, a do zbudowania komputera DNA, który będzie mógł wykonywać wszystkie operacje podobnie do komputerów tradycyjnych jeszcze przed nami długa droga.

Obliczenia DNA zostały zapoczątkowane pracą L. Adlemana [1], która przedstawia algorytm rozwiązania problemu drogi Hamiltona w grafie za pomocą DNA. Problem ten polega na rozstrzygnięciu czy istnieje droga w grafie skierowanym, przechodząca przez wszystkie jego wierzchołki tylko raz. Rozwiązanie polega na zakodowaniu wszystkich elementów grafu (wierzchołków i krawędzi) jako jednoniciowych molekuł, w taki sposób, że istnienie owej drogi zostanie zakomunikowane poprzez powstanie długiego dwuniciowego fragmentu DNA. Dalsze badania to rozwiązania kolejnych problemów informatycznych [2], konstrukcja prostych modeli automatów [3][4], a także

symulacji wnioskowań logicznych [5][6] (czego dotyczy poniższa praca).

W skład cząsteczki DNA zwykle wchodzi dwa łańcuchy, zbudowane w taki sposób, że znając jeden z nich możemy bez problemu odtworzyć drugi (zasada komplementarności Watsona-Cricka, która mówi, że adenina łączy się wyłącznie z tyminą, a cytozyna z guaniną). Każda z nici ma na jednym końcu (oznaczonym jako 5') wolną grupę fosforanową, a na drugim (oznaczonym jako 3') wolną grupę hydroksylową. Splecione ze sobą dwie nici DNA są względem siebie antyrównoległe. Na ilustracjach górna nić biegnie w kierunku od 5' do 3', dolna zaś od 3' do 5'. Podstawową operacją jest ligacja czyli łączenie się ze sobą pojedynczych łańcuchów komplementarnych (zarówno jednoniciowych jak i dwuniciowych zakończonych lepkiem końcem tj. krótkim jednoniciowym fragmentem). Przeciętną operacją jest cięcie podwójnych łańcuchów przez enzymy restrykcyjne – związki chemiczne, które rozpoznają odpowiedni fragment DNA i rozcinają łańcuch w ściśle określonym miejscu względem rozpoznanego fragmentu. Często pozostawiają wspomniane wcześniej lepkie końce, co umożliwia dalsze łączenie się powstałych cząstek. Komputer DNA oparty na ciągu naprzemiennych łączy i cięć odpowiednich molekuł nazywamy modelem splatania [7][8].

Istnieją inne modele biokomputerów [7][8], jednak wszystkie przedstawione w artykule systemy wnioskowania oparte są na modelu splatania.

2. LOGIKA

Zdaniem logicznym nazywamy dowolne wyrażenie utworzone ze zmiennych zdaniowych (zwanych również termami lub faktami) np. p, q, r, \dots oraz spójników logicznych: koniunkcji \wedge („i”), alternatywy \vee („lub”), negacji \neg („nieprawda, że”), implikacji \Rightarrow („jeśli..., to...”) oraz nawiasów. Na przykład zdaniami logicznymi są: $p, p \Rightarrow q, (p \vee q) \Rightarrow \neg r$. Jeżeli przypiszemy wszystkim zmiennym wartościowanie (prawdy lub fałszu), to posługując się prawami logiki przypisanymi spójnikom, jesteśmy w stanie określić wartościowanie każdego zdania logicznego. Tautologią nazywamy zdanie zawsze prawdziwe, niezależnie od wartościowania użytych w nim zmiennych np. $p \vee \neg p$. Zwyczajowo prawdę oznaczamy symbolem „1”, a fałsz „0”.

p :	q :	$p \wedge q$:	$p \vee q$:	$p \Rightarrow q$:	$\neg p$:
1	1	1	1	1	0
1	0	0	1	0	0
0	1	0	1	1	1

0	0	0	0	1	1
---	---	---	---	---	---

Tabela 1: Wartościowanie spójników logicznych.

W rachunku zdań za prawdziwe uznajemy również te zdania, dla których istnieje możliwość wyprowadzenia ich za pomocą reguł wnioskowania z zestawu zdań, których prawdziwość uznajemy. Jedną z nich (reguła odrywania) mówi, że jeżeli zdania p oraz $p \Rightarrow q$ uznajemy za prawdziwe, to na ich podstawie zdanie q również możemy uznać za prawdziwe. Może ono służyć do wyciągania dalszych wniosków. Oprócz reguły odrywania stosuje się jeszcze regułę podstawiania. Jednak w omawianych systemach opartych na DNA nie jest ona wykorzystywana.

3. IMPLEMENTACJE

Konstrukcje praktyczne, implementujące za pomocą modelu splatania wnioskowanie logiczne za pomocą DNA, aktualnie ograniczają się do zastosowania reguły odrywania. Ich celem jest odpowiadanie na pytania postaci: czy fakt, o który pytamy jest możliwy do wyprowadzenia na podstawie posiadanej wiedzy (faktów, których prawdziwość uznajemy).

Pierwsza implementacja takiego systemu została przedstawiona w roku 2005 przez zespół Olgerda Unolda z Politechniki Wrocławskiej [5]. Wykorzystuje ona: fakty (o wartościowaniu prawdziwym), implikacje, koniunkcję, alternatywę oraz zapytania. Użyty został enzym restrykcyjny *BseXI*. Detekcja odpowiedzi następuje poprzez posortowanie powstałych cząstek wg długości (pojawienie się odpowiednio długiej molekuly oznacza odpowiedź twierdzącą).

Implementację o podobnych możliwościach (omówioną w punkcie 4) przedstawił w roku 2009 zespół Ehuda Shapiro z Izraela [6]. Jest to pierwszy tego typu system przetestowany również laboratoryjnie. Różnicą jest budowa cząstek, użyty enzym restrykcyjny *FokI* oraz detekcja odpowiedzi poprzez umieszczenie na końcach zapytań elementów fluorescencyjnych.

W historii DNA obliczeń pojawiało się więcej pomysłów częściowej implementacji operacji logicznych, jednak te dwa posiadają największe możliwości rozbudowy.

4. KONCEPCJA SHAPIRO

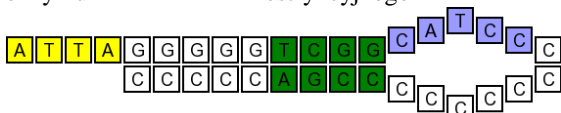
Molekuly reprezentujące fakty (p, q, r, \dots) składają się z trzech elementów: dystansownika (siedmiu dowolnych nukleotydów), startera enzymu restrykcyjnego *FokI* oraz jednoniciowego łańcucha identyfikującego (każda zmienna

zdaniowa posiada przyporządkowany unikalny czteronukleotydowy fragment DNA).



Rys.1: Molekuła reprezentująca fakt p .

Molekuły reprezentujące implikacje ($p \Rightarrow q, \dots$) składają się z czterech elementów: łańcucha identyfikującego poprzednik implikacji, dystansownika (pięciu dowolnych nukleotydów), łańcucha identyfikującego następnik implikacji oraz kolistego fragmentu, zawierającego starter enzymu restrykcyjnego *FokI*.



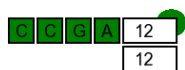
Rys.2: Molekuła reprezentująca implikację $p \Rightarrow q$.

Dla zakończenia wnioskowania za pomocą reguły odrywania oraz dołączenia wniosku do zbioru faktów, niezbędna jest jeszcze molekuła terminalna.



Rys.3: molekuła terminalna wnioskowania.

Molekuły reprezentujące zapytania składają się z dwóch elementów: jednonuciowego łańcucha zmiennej, o którą pytamy oraz dystansownika (dwunastu dowolnych nukleotydów), zawierającego substancję fluorescencyjną.

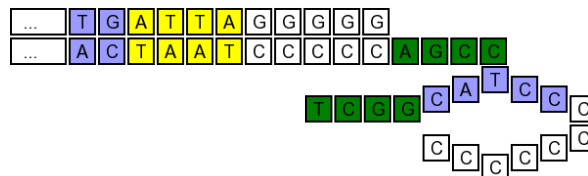


Rys.4: Molekuła reprezentująca zapytanie q ?

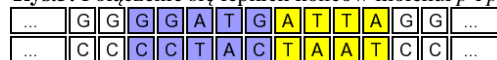
5. PRZYKŁADOWA REAKCJA

Jeżeli w próbówce umieścimy cząsteczki pokazane jako przykładowe w części 4 artykułu: fakt p , implikację $p \Rightarrow q$, cząsteczkę terminalną oraz zapytanie q ?, to reakcja powinna wyglądać następująco: na mocy reguły odrywania uznajemy fakt q za prawdziwy, zatem system zwraca odpowiedź twierdzącą.

W reakcji biomolekularnej wygląda to następująco:



Rys.5: Połączenie się lepkich końców molekuł p i $p \Rightarrow q$.



Rys.6: Rozcięcie molekuły przez enzym *FokI*.



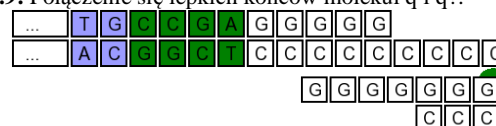
Rys.7: Rozwinięcie powstałej molekuły jednonuciowej.



Rys.8: Przyłączenie się molekuły terminalnej, powstaje q .



Rys.9: Połączenie się lepkich końców molekuł q i q ?



Rys.10: Rozcięcie molekuły przez enzym *FokI*.



Rys.11: Rozpad i uaktywnienie fluorescencji.

Reakcja może być dowolnie rozbudowywana, a do końcowego wyniku może prowadzić dowolnej długości skończony ciąg wnioskowań. Nie wszystkie cząsteczki znajdujące się w bazowej próbówce (czyli których prawdziwość uznajemy) muszą zostać wykorzystane. Niepojawienie się świecenia oznacza, że fakt o który pytamy nie jest możliwy do wyprowadzenia z informacji wejściowych (oznacza to, że system zwraca odpowiedź przeczącą).

Omówiony model ma jednak wady. Dobrze funkcjonuje wyłącznie dla prostych wnioskowań za pomocą reguły odrywania. Brakuje w nim ujęcia np. operacji negacji, tak istotnej w rozumowaniach (np. dla dowodów nie wprost) i związanych z nią aksjomatów (np. reguły kontrapozycji).

6. NOWA KONCEPCJA

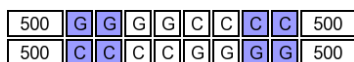
Celem wprowadzenia nowej koncepcji systemu wnioskującego za pomocą DNA jest wyeliminowanie trudności powstałych w modelach Shapiro i Unolda, a także rozszerzenie ich możliwości [9]. Główną zmianą jest wprowadzenie operacji negacji (czyli możliwości używania zarówno faktów np. p jak i ich zaprzeczeń np. $\neg p$).



Rys.22: Przyłączenie się detektora sprzeczności.



Rys.23: Rozcięcie molekuly przez enzym *AcuI*.



Rys.24: Ponowne przyłączenie się detektora sprzeczności.

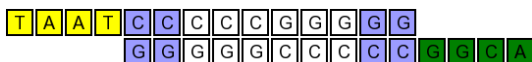
Pojawienie się cząstki długości rzędu 1000 oznacza, że w systemie znalazła się sprzeczność. Oczywiście nie musi ona wystąpić na etapie projektowania, ale może w wyniku skończonego ciągu wnioskowań np. p , $p \Rightarrow q$, $q \Rightarrow \neg r$, $\neg r \Rightarrow \neg p$.

9. REGUŁA KONTRAPOZYCJI

Jednym z aksjomatów klasycznego rachunku zdań jest reguła kontrapozycji. Mówi ona, że jeżeli prawdziwa jest implikacja $p \Rightarrow q$ (jeśli p to q), to zawsze prawdziwe jest również rozumowanie „w drugą stronę” $\neg q \Rightarrow \neg p$ (jeśli nieprawda, że q to nieprawda, że p). Wprowadzenie negacji do systemu wnioskującego wymaga również uwzględnienia tego aksjomatu.

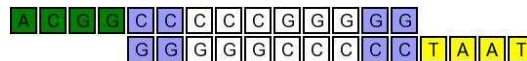
Wykorzystując wcześniejszą obserwację dotyczącą możliwości „obracania” cząsteczek oraz odpowiednio projektując cząsteczkę reprezentującą implikację tak by była ona symetryczna względem punktu środkowego (co umożliwia oderwanie wniosku zarówno poprzednika i następnika implikacji) możemy wprowadzić regułę kontrapozycji we wnioskowaniu za pomocą DNA.

Molekuły reprezentujące implikację uwzględniającą regułę kontrapozycji składają się z pięciu elementów umieszczonych w sposób symetryczny: dystansownika (sześciu nukleotydów) w środku cząsteczki, dwóch dwójek nukleotydów odpowiedzialnych za wykrywanie sprzeczności i dwóch czwórek jednoznacznie identyfikujących zmienne zdaniowe poprzednika i następnika implikacji.



Rys.25: Molekuła reprezentująca implikację $p \Rightarrow q$.

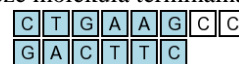
Warto zwrócić uwagę, że ta sama molekula zapisana jako odwrócona o 180° będzie przedstawiać implikację $\neg q \Rightarrow \neg p$.



Rys.26: Molekuła reprezentująca implikację $\neg q \Rightarrow \neg p$.

Oczywiście w rzeczywistości jest to ta sama cząsteczka ponieważ możliwość obracania jej jest naturalna w próbówce.

Dla zakończenia wnioskowania za pomocą reguły odrywania oraz dołączenia wniosku do zbioru faktów, niezbędna jest jeszcze molekula terminalna.



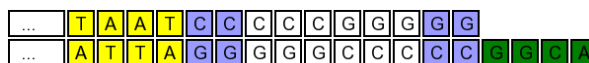
Rys.27: Molekuła terminalna wnioskowania.

W punktach 10 i 11 przeanalizujemy wnioskowanie w obu możliwych kierunkach z wykorzystaniem reguły kontrapozycji.

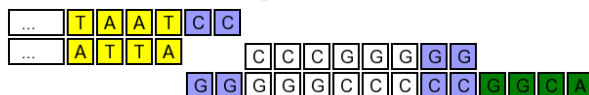
10. WNISKOWANIE $P \Rightarrow Q$

Jeżeli w próbówce umieści się cząsteczki pokazane w części 6 i 9 artykułu: fakt p , implikację $p \Rightarrow q$ oraz cząsteczkę terminalną, to reakcja wygląda następująco: na mocy reguły odrywania uznajemy fakt q za prawdziwy.

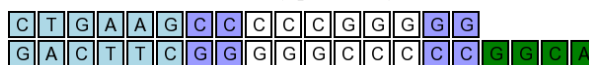
W reakcji biomolekularnej wygląda to następująco:



Rys.28: Połączenie się lepkich końców molekul p i $p \Rightarrow q$.



Rys.29: Rozcięcie molekuly przez enzym *AcuI*.



Rys.30: Przyłączenie się molekuly terminalnej, powstaje q .

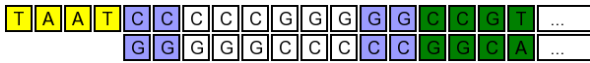
Powstała cząsteczka q o wartościowaniu prawdziwym może być wykorzystywana do dalszej reakcji i przeprowadzania dalszych wnioskowań.

11. WNISKOWANIE $\neg Q \Rightarrow \neg P$

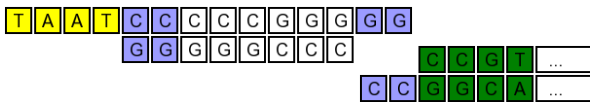
Jeżeli w próbówce umieści się cząsteczki pokazane w części 6 i 9 artykułu: zaprzeczenie faktu $\neg q$, implikację

$p \Rightarrow q$ oraz cząsteczkę terminalną, to reakcja wygląda następująco: na mocy reguły odrywania uznajemy fakt p za fałszywy (zaprzeczenie $\neg p$ za prawdziwe).

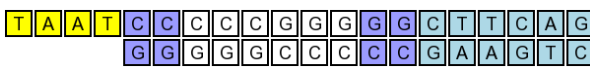
W reakcji biomolekularnej wygląda to następująco:



Rys.31: Połączenie się lepkich końców molekuł $\neg q$ i $p \Rightarrow q$.



Rys.32: Rozcięcie molekuły przez enzym *AclI*.



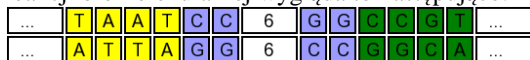
Rys.33: Przyłączenie się molekuły terminaln., powstaje $\neg p$.

Powstała cząsteczka $\neg p$ o wartościowaniu fałszywym może być wykorzystywana do dalszej reakcji i przeprowadzania dalszych wnioskowań.

12. WYKRYCIE SPRZECZNOŚCI $P, P \Rightarrow Q, \neg Q$

Jeżeli w próbówce umieści się cząsteczki pokazane w części 6 i 9 artykułu: fakt p , implikację $p \Rightarrow q$, zaprzeczenie faktu $\neg q$, to powinna wystąpić sprzeczność. Może zająć wnioskowanie $p \Rightarrow q$ lub $\neg q \Rightarrow \neg p$ – w obu tych sytuacjach sprzeczność zostanie wykryta tak jak opisane zostało to w części 8. Może też zająć sytuacja, w której molekuły faktów jednocześnie łączą się z molekułą implikacji.

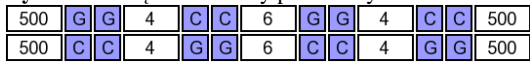
W reakcji biomolekularnej wygląda to następująco:



Rys.34: Połączenie się lepkich końców $p, p \Rightarrow q$ i $\neg q$.



Rys.35: Rozcięcie molekuły przez enzym *AclI*.



Rys.36: Przyłączenie się dwóch detektorów sprzeczności.

Pojawienie się cząstki długości rzędu 1000 oznacza, że w systemie znalazła się sprzeczność.

13. PODSUMOWANIE

Celem artykułu było przedstawienie koncepcji komputera DNA, który potrafi logicznie wnioskować.

W pierwszej części przedstawiony został model przedstawiony przez zespół Ehuda Shapiro, który potrafi

wysnuwać wnioski na podstawie informacji o prawdziwości pewnych faktów. Został on przetestowany laboratoryjnie, a uzyskane wyniki były identyczne jak w rozważaniach teoretycznych [3]. Umożliwia on jedynie proste wnioskowanie, wyłącznie w jednym kierunku, nieuwzględniające operacji negacji i związanych z nią aksjomatów, dlatego stanowi wyłącznie punkt wyjścia do dalszych rozważań i stworzenia pełniejszego systemu wnioskującego.

W dalszej części zaprezentowana została nowa koncepcja systemu wnioskującego za pomocą DNA, rozszerzająca możliwości modelu omówionego wcześniej. Przedstawione zostały wprowadzone zmiany: inna budowa cząsteczek, inny użyty enzym restrykcyjny, wprowadzenie operacji negacji, kontrola spójności systemu poprzez wykrywanie mogących pojawić się w wyniku obliczeń sprzeczności.

Nowym elementem, przedstawianym tutaj po raz pierwszy, jest implementacja reguły kontrapozycji w systemie wnioskującym za pomocą DNA. Wprowadzenie jej było możliwe dzięki obserwacji, że „odwracanie” cząsteczek o 180° nie zmienia ich faktycznej orientacji w próbówce (odbywa się w naturalny sposób). Istnieje dzięki temu więcej możliwości ligacji (reakcji łączenia się ze sobą komplementarnych końców). Poprzez stworzenie cząsteczki symetrycznej względem swojego środka (różniacej się jedynie fragmentem identyfikującym zdania atomowe) możliwe stało się oderwanie wniosku zarówno poprzednika i następnika implikacji, zależnie od posiadanej wcześniej wiedzy, aktywującej ową reakcję.

Używając tych samych obserwacji (dotyczących „odwracania” i symetrii cząsteczek) molekuły faktów (zdań atomowych) zostały tak zaprojektowane, aby przy użyciu jednego (wspólnego dla wszystkich zdań atomowych) detektora mogła odbywać się kontrola spójności systemu [5].

W planach znajduje się dalsze rozszerzanie możliwości omawianego systemu wnioskującego oraz badania laboratoryjne potwierdzające jego poprawność.

Literatura

- ADLEMAN L. Molecular computation of solutions to combinatorial problems. Science 1994; 226: 1021-1024.
- LIPTON R. DNA solution of hard computational problems. Science 1995; 268: 542-545.
- BENENSON Y, PAZ-ELIZUR T, ADAR R, KEINAN E, LIVNEH Z, SHAPIRO E. Programmable and autonomous computing machine made of biomolecules. Nature 2001; 414: 430-434.

4. BENENSON Y, ADAR R, PAZ-ELIZUR T, LIVNEH Z, SHAPIRO E. DNA molecule provides a computing machine with both data and fuel. PNAS 2003; 100: 2191-2196.
5. UNOLD O, TROĆ M. Biomolekularne modele obliczeniowe. Inżynieria Komputerowa 2005; 73-85.
6. RAN T, KAPLAN S, SHAPIRO E. Molecular implementation of simple logic program. Nature Nanotechnology 2009; 203(2): 1-7.
7. PAUN G, ROZENBERG G, SALOMAA A. DNA Computing. New Computing Paradigms. Springer-Verlag 1999.
8. KRASIŃSKI T, SAKOWSKI S. Przegląd modeli i praktycznych implementacji DNA obliczeń. Studia Informatica 2008; 29: 5-31.
9. ROGOWSKI Ł. Implementacje systemów logicznych za pomocą DNA – koncepcje i pomysły rozbudowy. WMiI Preprint (Wydział Matematyki i Informatyki)

