



Poznań, 06.09.2023r.

Prof. dr hab. Stanisław Gawiejnowicz  
Wydział Matematyki i Informatyki  
Uniwersytet im. Adama Mickiewicza w Poznaniu  
e-mail: [stgawiej@amu.edu.pl](mailto:stgawiej@amu.edu.pl)  
WWW: <https://algo.wmi.amu.edu.pl>

**RECENZJA ROZPRAWY DOKTORSKIEJ MGR. JAROSŁAWA SYNAKA PT.  
„MODELOWANIE I SYMULOWANIE SYSTEMÓW WIELOAGENTOWYCH  
OPISUJĄCYCH PROCESY W HIPOTEZIE ŚWIATA RNA”**

Niniejszą recenzję sporządzono na mocy uchwały nr 2023-21-159 Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Poznańskiej z dnia 27.06.2023r., powołującej mnie na recenzenta ww. rozprawy. Wersję elektroniczną tej rozprawy otrzymałem 06.07.2023r., natomiast wersję papierową – 11.07.2023r.

**1. Tematyka rozprawy**

Recenzowana rozprawa, napisana pod kierunkiem prof. dr hab. inż. Jacka Błazewicza, czł. rzecz. PAN, jest poświęcona modelowaniu procesów ewolucyjnych związanych z tzw. *hipotezą świata RNA*. Mówi ona, iż życie na Ziemi rozpoczęło się od cząsteczek RNA oraz że są one ewolucyjnie wcześniejsze niż białka i cząsteczki DNA, które z nich powstały.

Hipoteza ta, sformułowana w latach 60-tych XX wieku (Rich, 1962; Woese, 1967), jest ciekawa poznawczo, choć wnioski, jakie można na jej podstawie wysnuć mają głównie charakter spekulatywny. Wynika to z faktu, iż akceptacja tej hipotezy wymaga przyjęcia założenia, iż cząsteczki RNA mogły w przeszłości pełnić rolę zarówno replikatora, jak i enzymu. Ponadto, trudno tę hipotezę zweryfikować, ponieważ nie wiadomo, jakie warunki panowały na Ziemi wtedy, gdy nie było jeszcze białek i DNA, a jedynie cząsteczki lub łańcuchy RNA. To z kolei powoduje, iż w celu przeprowadzenia eksperymentów potwierdzających bądź falsyfikujących hipotezę świata RNA należy przyjąć założenia dotyczące warunków, co do których nie mamy pewności iż były spełnione w rzeczywistości.

Mimo ww. wątpliwości, hipoteza świata RNA jest użyteczna badawczo, ponieważ wyniki eksperymentów związanych z jej badaniem mogą doprowadzić do nowych hipotez związanych z początkami życia na Ziemi. Ponadto, organizacja i przeprowadzanie tego typu eksperymentów to duże wyzwanie badawcze, gdyż wymagają one szerokiej wiedzy z kilku dziedzin oraz stosowania różnych podejść i metod. Z tych powodów uważam, że wyboru tematyki rozprawy dokonano trafnie oraz że jest ona ważna naukowo i istotna poznawczo.

Rozprawa ma charakter interdyscyplinarny, gdyż wykorzystuje się w niej pojęcia oraz metody pochodzące z matematyki, informatyki, biologii i chemii. W związku z tym, że najbliższe mi są zagadnienia związane z dwoma pierwszymi dziedzinami, w niniejszej recenzji skupię się na matematycznych i informatycznych aspektach recenzowanej rozprawy.

## 2. Ogólny układ rozprawy

Rozprawa składa się 10 rozdziałów oraz polskiego i angielskiego streszczenia, wraz ze stroną tytułową liczy 125 stron. Bibliografia rozprawy obejmuje 134 pozycje, w tym 3 pozycje, których współautorem jest Autor rozprawy ([105], [106], [110]). Całość rozprawy można podzielić na dwie części, wstępną i merytoryczną, o następującej zawartości.

Część wstępna rozprawy obejmuje rozdziały 1-4, wprowadzające aparat pojęciowy stosowany w części merytorycznej. Rozdział 1. zawiera opis pochodzących z informatyki metod stosowanych w badaniu początków życia oraz cel i zakres rozprawy. Rozdział 2. przedstawia główne pojęcia z kilku działów matematyki i informatyki, wykorzystywane w dalszych rozdziałach rozprawy, tj. równania różniczkowe zwyczajne i cząstkowe, złożoność obliczeniową, algorytmy wielowątkowe, automaty komórkowe oraz systemy wieloagentowe. Rozdział 3. wprowadza pojęcia związane z podstawami biologicznymi i chemicznymi: dynamiką molekularną, kinetyką reakcji chemicznych, problemem początków życia, hipotezą świata RNA, śladami geologicznymi, strukturą przestrzenną RNA oraz badaniami *in vitro*. Wreszcie rozdział 4. zawiera opis metod analitycznych stosowanych w badaniu hipotezy świata RNA: równanie quasi-gatunku, równanie replikatorowe, równania różniczkowe wykorzystujące pojęcie tzw. *hipercyklu* oraz automaty komórkowe.

Część merytoryczna rozprawy to rozdziały 5-9, zawierające wyniki własne Autora rozprawy. Ta część zawiera dwie grupy wyników. Pierwsza grupa obejmuje wyniki zaprezentowane w rozdziałach 5-7, napisanych w oparciu o prace, odpowiednio, [110], [105] oraz [106]. Rozdziały te dotyczą, odpowiednio, modelu wieloagentowego, modelu wieloagentowego z sekwencjami oraz modelu wykorzystującego równania różniczkowe cząstkowe. W każdym z tych rozdziałów opisano model biologiczny, jego matematyczny odpowiednik, algorytm wykorzystany w symulacji, wyniki symulacji oraz dyskusję tych wyników. Implementacja algorytmu stosowanego w symulacji modelu z rozdziału 5. została napisana w języku C#, pozostałe dwie implementacje – w języku C++.

Druga grupa wyników obejmuje wyniki przedstawione w rozdziałach 8-9. W rozdziale 8. przedstawiono model wykorzystujący ewolucję syntez, a w rozdziale 9. – model wykorzystujący inhibitory kontrolujące populację. Prezentacja wyników w obu tych rozdziałach przebiega wg tego samego schematu co w rozdziałach 5-7, z tą różnicą, iż nie napisano *explicite*, czy przedstawiane wyniki zostały wcześniej opublikowane.

Rozprawę kończy rozdział 10, zawierający podsumowanie.

## 3. Problem badawczy i jego znaczenie

Problemem badawczym rozpatrywanym w rozprawie jest określenie zakresu stosowalności hipotezy świata RNA, natomiast jej celem – konstrukcja modeli służących do weryfikacji ograniczeń i zakresu stosowalności hipotezy świata RNA (rozdz. 1.2, s. 13).

Badany w rozprawie problem ma, moim zdaniem, istotne znaczenie, gdyż wiąże się z fundamentalnym pytaniem o sposób powstania życia na Ziemi. Ponadto, zbadanie ograniczeń oraz zakresu stosowalności hipotezy świata RNA może doprowadzić do odkrycia nowych zależności, dotyczących roli RNA w procesach ewolucyjnych. Nie bez znaczenia

jest także fakt, iż konstrukcja coraz bardziej złożonych modeli, takich jak te opisywane w rozprawie, czyni te modele coraz bliższymi ich odpowiednikom w świecie rzeczywistym, co powoduje, iż symulacje przeprowadzane z wykorzystaniem tych modeli coraz lepiej odzwierciedlają modelowane zjawiska i procesy.

#### 4. Wkład autora

W rozprawie Autor bada 5 modeli. Trzy z nich, przedstawione w rozprawie w rozdziałach 5-7, pochodzą z prac [105], [106] oraz [110], natomiast dwa przedstawione w rozdziałach 8-9 nie zostały jeszcze opublikowane. Pierwsza oraz trzecia z ww. prac zostały opublikowane w *PLOS One* (140pkt wg ministerialnej listy czasopism), natomiast druga – w *Entropy* (100pkt wg tej listy). Moim zdaniem, potwierdza to wagę prezentowanych w tych pracach modeli oraz duże znaczenie wyników symulacji przeprowadzonych z ich użyciem.

Ww. prace są współautorskie i zostały napisane przez, odpowiednio, trzech, trzech oraz pięciu autorów. W materiałach dołączonych do rozprawy ani w samej rozprawie nie znalazłem informacji nt. wkładu poszczególnych autorów w te prace. Brak ten, w pewnym stopniu, rekompensują opisy tego wkładu (ang. author contributions), podane w tych pracach. Wynika z nich, iż w przypadku pracy [105] Autor rozprawy miał wkład w 7 spośród 9 wymienionych na końcu tej pracy aspektów. Wkład Autora rozprawy w przypadku prac [106] oraz [110] wynosi, odpowiednio, 8 spośród 10 oraz 4 spośród 14 ww. aspektów. Oznacza to, moim zdaniem, że Autor rozprawy miał dominujący udział w pierwszych dwu pracach, natomiast mniejszy – w ostatniej z nich.

#### 5. Poprawność

Problem badany w rozprawie Autor rozwiązuje za pomocą symulacji komputerowych rozwoju populacji złożonych z cząsteczek RNA oraz pasożytów, wykorzystując w tym celu następujące modele, których jest współautorem:

- *wieloagentowy*, w którym cząsteczki obu ww. rodzajów są reprezentowane przez agentów, przy czym parametry agentów są wyliczane na podstawie tej samej zależności, wspólnej dla wszystkich agentów tego samego rodzaju;
- *wieloagentowy z sekwencjami*, w którym parametry poszczególnych agentów są różne, a oblicza się je na podstawie sekwencji nukleotydów tych agentów;
- *różniczkowy*, w którym populacja cząsteczek RNA jest opisana za pomocą układu równań różniczkowych, przy czym rozpatruje się dwa przypadki: bez możliwości mutacji oraz z możliwością mutacji;
- *model funkcyjny*, w którym pasożyty mogą nabywać nowe funkcje, które mimo mutacji pozwalają efektywniej zapamiętywać informacje genetyczną przekazywaną kolejnym pokoleniom cząsteczek i pełnią rolę stabilizatora chroniącego populację przed wymarciem;
- *model z inhibitorami*, w którym krótkie sekwencje RNA pełnią rolę inhibitorów, kontrolujących stężenia cząsteczek poszczególnych typów RNA w danej populacji.

Dla niektórych z tych modeli, np. tych z rozdziału 5., 7. oraz 9., Autor przedstawił modele matematyczne opisujące pewne procesy specyficzne dla danego modelu (np. proces dyfuzji dla modelu z rozdziału 5.) bądź parametry go opisujące (np. liczebności replikaz, pasożytów bądź zasobów w modelu z rozdziału 7.). Rozumowania oraz algebraiczne przekształcenia

formuł matematycznych, podane w tych rozdziałach w celu konstrukcji tych modeli, są moim zdaniem prawidłowe.

Symulacje przeprowadzone w oparciu o ww. modele matematyczne zostały dokonane bardzo starannie, uruchamiając implementacje badanych modeli dla bardzo dużej liczby kroków (rzędu dziesiątek tysięcy) oraz dla różnych ich wariantów, zależnych od wartości parametrów opisujących badaną populację. Stąd też, moim zdaniem, wyniki tych symulacji są reprezentatywne, tzn. prawidłowo odzwierciedlają ewolucję badanych populacji przy przyjętych założeniach początkowych i wartościach parametrów kontrolujących liczebności cząsteczek RNA oraz pasożytów w tych populacjach.

## 6. Wiedza kandydata

Rozprawa ma charakter interdyscyplinarny, a przedstawione w niej wyniki wymagają znajomości wybranych pojęć i metod matematyki (analiza matematyczna, równania różniczkowe), informatyki (teoria algorytmów, programowanie równoległe), biologii (procesy ewolucyjne, struktury cząsteczek i łańcuchów RNA oraz DNA) oraz chemii (równania reakcji chemicznych). Mimo pewnych niedociągnięć w prezentacji tych pojęć i metod, wymienionych w punkcie 7., w szczególności tych związanych z matematyką i informatyką, nie mam wątpliwości, iż Autor rozprawy opanował odpowiednią wiedzę z ww. dziedzin oraz że potrafi stosować ją w praktyce.

Podobna ocena dotyczy aparatu teoretycznego z informatyki. Konstrukcja algorytmów wykorzystanych w symulacjach opisanych w rozdziałach 5-9 wymagała bardzo dobrej znajomości algorytmiki, zaawansowanych technik programowania (implementacji dokonano w obiektowych językach C# i C++) oraz programowania równoległego (standard OpenMPI). Także w tym przypadku, mimo iż Autor rozprawy nie ustrzegł się kilku usterek opisanych w punkcie 7., nie mam wątpliwości, iż w pełni opanował ww. techniki programowania oraz że potrafi z nich prawidłowo korzystać.

## 7. Inne uwagi

W trakcie czytania rozprawy nasunął mi się szereg uwag, które można podzielić na uwagi ogólne oraz uwagi szczegółowe. Dotyczą one głównie sposobu prezentacji pewnych zagadnień i nie wpływają istotnie na moją bardzo pozytywną ocenę recenzowanej rozprawy.

### 7.1 Uwagi ogólne

- W rozprawie formuły matematyczne są numerowane, ale nie do wszystkich z nich istnieją odwołania w tekście. Przykładowo, brak odwołań do numerowanych formuł w rozdziałach 2-4, a pierwsze odwołanie ma miejsce w punkcie 5.2.5.1 (40:22 od góry). Zauważmy także, iż odwołując się do numerowanych formuł należy stosować nawiasy okrągłe w celu odróżnienia numerów tych formuł od liczb.
- Autor traktuje formuły matematyczne jako byty niezależne od tekstu, pomijając znaki przestankowe takie jako przecinki, średniki czy kropki w przypadkach, gdy formuły są częściami zdań. Przykładowo, po formule (2.1) winien być przecinek, a po formule (2.2) należało postawić kropkę.

- Podany w rozprawie opis pewnych pojęć wprowadzonych w rozdziale 2 może prowadzić do nieporozumień. Przykładowo, pisząc o funkcji różniczkowalnej w punkcie (15:8-5 od dołu) Autor używa tego samego symbolu „ $x$ ” na oznaczenie zmiennej niezależnej funkcji  $f$  oraz punktu w którym obliczana jest pochodna tej funkcji; by uniknąć nieporozumień, w tym drugim przypadku zamiast „ $x$ ” lepiej użyć „ $x_0$ ”.
- Rozdziały 4-7 mają układ podobny do układu prac na podstawie których je napisano. Moim zdaniem, brak w tych rozdziałach części zawierającej wnioski (obecnie są one ukryte w podrozdziałach zawierających dyskusję wyników poszczególnych symulacji) oraz sugestii co do dalszych badań związanych z określonymi modelami.
- Pseudokody algorytmów prezentowanych w rozdziałach 5-9 mogłyby być przedstawione w bardziej formalny sposób. Po pierwsze, poszczególne wiersze tych pseudokodów mogłyby nosić numery, co umożliwiłoby odwoływanie się do nich w tekście rozprawy. Po drugie, brak w nich opisu wejścia oraz wyjścia danych algorytmów. Po trzecie, w pseudokodzie nie powinno być komentarzy bądź wyjaśnień (por. „patrz równanie (5.2)” w pseudokodzie Algorytmu 1, zob. 42:19 od góry), a jedynie opis operacji wykonywanych przez dany algorytm. Po czwarte wreszcie, pseudokody algorytmów powinny być, tak dalece jak to możliwe, precyzyjne, co oznacza iż należy unikać opisów typu „zrób to lub tamto” bądź „zrób to i to” (por. pseudokod Algorytmu 1, 42:2-4 od góry).
- Opisując implementacje badanych modeli, Autor nie podał dokładniejszych danych na temat środowiska sprzętowego (np. zegar, ilość pamięci RAM itp.) ani programowego (np. numer wersji kompilatora, nazwa stosowanej odmiany Linuxa itp.).
- Moim zdaniem, w rozprawie doktorskiej należy unikać korzystania ze źródeł internetowych, jeśli znane są źródła drukowane. Spośród 134 pozycji bibliografii recenzowanej rozprawy 12 to odwołania do stron z internetowej wersji encyklopedii matematycznej oraz *Encyclopaedia Britannica*. Korzystanie z nich w kontekście definicji pojęć matematycznych uważam za nietrafne – lepiej było odwołać się do odpowiednich pozycji książkowych, szczególnie, że Autor rozprawy tak czyni w przypadku pojęć związanych z bioinformatyką.

## 7.2 Uwagi szczegółowe

- Rachunek różniczkowy i całkowy są ze sobą nierozdzielnie związane, ponieważ całkowanie i różniczkowanie to operacje odwrotne względem siebie. Nie można więc pisać o pierwszym, pomijając drugi. Z tego powodu na początku podrozdziału 2.1 (15:14 od góry) należy dodać słowa „i całkowy” po słowach „rachunek różniczkowy”.
- Tytuł rys. 2.1 (s. 16) powinien być krótszy, a wszelkie wyjaśnienia go dotyczące należało umieścić w tekście. Podobna uwaga dotyczy innych rysunków w rozprawie.
- Z kontekstu wynika, iż  $F$  jest funkcją  $n$  zmiennych, więc po słowach „dowolną funkcją” (17:7 od góry) należy dodać frazę „ $n$  zmiennych”.
- Zamiast pisać o pochodnej „po konkretnym argumentie” (17:8 od dołu) lepiej pisać o pochodnej „względem argumentu”.
- Używając określonych nazw (terminów) należy konsekwentnie stosować tę samą pisownię. W rozprawie nie zawsze tak jest, np. w punkcie 2.2 Autor pisze zarówno o „przestrzeni euklidesowej” (18:1-2 od góry), jak i o przestrzeni Euklidesowej” (18:5-6 od góry).

- Definicja symbolu  $O(f)$ , podana u dołu s. 18 jest niepoprawna, gdyż sugeruje iż symbol „ $N_0$ ” oznacza stałą rzeczywistą, podczas gdy oznacza on indeks (=liczbę naturalną); by skorygować tę definicję wystarczy przed „ $N_0$ ” wstawić słowo „indeks”. Podobna uwaga dotyczy definicji symboli  $\Omega(f)$  oraz  $\Theta(f)$ , podanych u góry s. 19.
- W zdaniu dotyczącym klas złożoności (19:11-13 od góry) należy usunąć frazę „dla poszczególnych”, słowo „asymptotycznie” oraz wstawić słowo „je” po „dzieli się”.
- By uniknąć niejasności należało dodać „XX w.” po „60” (25:16 od dołu).
- Po równaniu (5.3) na s. 39 nie podano zakresu zmienności  $a$ .
- Jako że w matematyce nie ma terminu „kompletnie niezależne” (41:16 od dołu), słowo „kompletnie” należy usunąć.
- Rysunki 5.3-5.6 oraz 5.9 (ss. 46-47 oraz 49, odpowiednio) zawierają angielskie słowo „Mean”, podczas gdy rozprawa jest w języku polskim.
- Rozdział 6 oparto o pracę [105] (por. 51:5 od góry), więc zamiast „w [110]” (63:3 od góry) winno być „w [105]”.
- W opisie wartości brzegowych (67:19 od góry) zamiast „funkcje reprezentujące” winno być „wartości funkcji reprezentujących”.
- Potoczne określenie „metoda delty” nie powinno być używane w tekście rozprawy doktorskiej – stąd zamiast „metodą delty” (99:19 od góry) lepiej napisać „stosując wzory na pierwiastki równania kwadratowego”.
- Pozycje [30] i [31] bibliografii dotyczą tej samej książki (podręcznik Cormena i innych z teorii algorytmów).
- Zbyteczne jest podawanie w opisie bibliograficznym książki nazwy kraju, w którym ją opublikowano. Ponadto, Autor nie jest konsekwentny: podaje tę nazwę w opisie kilku pozycji (np. [28], [79], [98], [124]), a nie podaje w innych (np. [8], [30-31]).
- W przypadku pozycji [102] Autor podaje nazwę kraju w języku polskim, podczas gdy opis dotyczy pozycji w języku angielskim.
- Pewne pozycje bibliografii można było pominąć, podczas gdy inne – dodać. Przykładowo, cytowana w rozprawie książeczka [124] (zob. 18:2 od góry) dotycząca elementarnych podstaw rachunku na wektorach, została napisana w 1901 roku (Autor podał wersję z 1913, choć było jeszcze kilka innych) i jest przestarzała – należało odesłać Czytelnika rozprawy do innych, bardziej aktualnych, źródeł (w języku polskim bądź angielskim). Podobna uwaga dotyczy doboru literatury nt. analizy matematycznej oraz równań różniczkowych.
- W języku polskim przyjęto rozpoczynać nazwiska oraz pisać skróty dużymi literami, stąd też nazwiska (Arrhenius, Weierstrass) czy skróty (RNA, MPI), występujące np. w pozycjach [23], [56], [58], [93], winny być napisane zgodnie z tą konwencją.

Tekst rozprawy został zredagowany starannie, z wykorzystaniem LaTeX-a, co powoduje, iż wzory matematyczne są bardzo czytelne, a tekst złożony zgodnie z zasadami typografii. Jedynie w kilkunastu miejscach dostrzegłem drobne błędy:

- zamiast „zainspirowane” winno być „zainspirowanych” (7:15-16 od dołu);
- po „niemniej” brak słowa „jednak” (12:22 od góry);
- zamiast „zależności” (13:5 od góry) lepiej napisać „własności”;
- zamiast „bilologicznych” (23:9 od dołu) winno być „biologicznych”;
- zamiast „ilości” (29:9 od dołu, 113:21 od góry) winno być „liczby”;
- po słowie „Wygodnie” (35:6 od góry) brak słowa „jest”;
- zamiast „bilologicznym” (8 od góry) winno być „biologicznym”;

- zamiast „sekwencje” (56:2 od dołu) winno być „sekwencje”;
- zamiast „ilość” (66:14 od góry, 97:16 od góry) winno być „liczba”;
- zamiast „funckję” (99:9 od góry) winno być „funkcję”;
- zamiast „rozkładem” (106:6 od dołu) winno być „rozkładem”.

## 8. Konkluzje

Biorąc pod uwagę opinie wyrażone w poprzednich punktach mojej recenzji oraz wymagania zdefiniowane w art. 187 ustawy z dnia 20 lipca 2018 r. (Prawo o Szkolnictwie Wyższym, Dz. U. 2018, poz. 1668, z późniejszymi zmianami), moja ocena rozprawy pod względem trzech podstawowych kryteriów jest następująca (symbol „☒” oznacza wybór, natomiast symbol „☐” oznacza brak wyboru danej opcji odpowiedzi):

A. Czy rozprawa zawiera oryginalne rozwiązanie problemu naukowego?

- zdecydowanie tak
- raczej tak
- trudno powiedzieć
- raczej nie
- zdecydowanie nie

B. Czy po przeczytaniu rozprawy zgadzasz się, że kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie informatyka techniczna i telekomunikacja?

- zdecydowanie tak
- raczej tak
- trudno powiedzieć
- raczej nie
- zdecydowanie nie

C. Czy kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej?

- zdecydowanie tak
- raczej tak
- trudno powiedzieć
- raczej nie
- zdecydowanie nie

prof. dr hab. Stanisław Gawiejnowicz

Wydział Matematyki i Informatyki  
Uniwersytet im. Adama Mickiewicza w Poznaniu