

Warszawa, 2011-10-31

Dr hab. Jacek Waniewski, prof. nadzw.
Instytut Biocybernetyki i Inżynierii Biomedycznej PAN
Polska Akademia Nauk
Trojdena 4
02-109 Warszawa
jwaniewski@ibib.waw.pl

**Recenzja rozprawy doktorskiej
mgr Andrzeja Mizery
"Metody oraz narzędzia do konstrukcji i analizy modeli obliczeniowych dla
zjawisk biologicznych. Ich zastosowanie do modelowania procesów
odpowiedzi na szok termiczny oraz samoorganizacji filamentów w żywych
komórkach"**

Niniejsza recenzja została wykonana w związku z przewodem doktorskim mgr Andrzeja Mizery na podstawie uchwały Rady Naukowej Instytutu Podstawowych Problemów Techniki PAN z dnia 30 czerwca 2011 roku powołującej mnie na recenzenta jego rozprawy doktorskiej. Recenzja została opracowana na podstawie dokumentów dostarczonych mi przez Sekretarza Rady oraz, na moją prośbę, uzupełnień dostarczonych przez Doktoranta w postaci szczegółowego opisu jego osobistego wkładu w opracowanie artykułów składających się na rozprawę oraz pełnego tekstu artykułu V, który w otrzymanej przeze mnie wersji rozprawy nie zawierał omawianych w nim wykresów.

Przedstawiona mi do recenzji praca ma formę cyklu publikacji poprzedzonych obszernym, liczącym 98 stron, ogólnym omówieniem problematyki rozprawy oraz wyników przedstawionych w publikacjach. Zawiera również czterostronicowe streszczenie w języku polskim. Promotorami rozprawy są dr hab. Barbara Gambin, prof. IPPT PAN, oraz prof. Iona Petre, Abo Akademi University, Turku, Finlandia. Rozprawa jest napisana w języku angielskim i na mocy umowy pomiędzy IPPT PAN a Uniwersytetem w Turku jest podstawą ubiegania się o stopień doktora w obu tych instytucjach.

Tematyka rozprawy dobrze wpisuje się w szybko rozwijający się kierunek badań polegający na zastosowaniu metod obliczeniowych do analizy matematycznych modeli złożonych układów biologicznych, coraz częściej, jak w omawianym przypadku, opisujących procesy na poziomie subkomórkowym, jak m. innymi genetyczne sieci regulacyjne, szlaki sygnałowe, sieci metaboliczne, czy powstawanie dużych struktur biochemicznych. Pojawienie się w ostatnich dziesięcioleciach nowych metod eksperymentalnych umożliwia uzyskanie ilościowych informacji o procesach o wysokiej złożoności, ale ich integracja i oparte na tym przewidywanie zachowania się układu są jednak możliwe wyłącznie przy pomocy modelowania matematycznego. Modele takie, zarówno ze względu na ilość zmiennych jak i na nieliniową strukturę oddziaływań z licznymi pętlami sprzężeń, nie poddają się analizie matematycznej, ale pozostaje do dyspozycji analiza numeryczna. Ta jednak, z tych samych powodów, nastrocza wiele problemów, a jej systematyczne stosowanie wymaga opracowania nowych reguł i technik komputerowych. Mgr Mizera poświęcił swoją rozprawę doktorską właśnie przedstawieniu kilku takich technik, które zostały zastosowane do dwu konkretnych modeli, ale w części wstępnej zwrócił uwagę na ogólne znaczenie tych podejść również w badaniu innych modeli matematycznych o podobnym charakterze.

Część ogólna rozprawy, po wstępnym rozdziale pierwszym, przechodzi do omówienia problemów zastosowania modelowania komputerowego w biologii układów. Przedstawione tu są metody konstrukcji deterministycznych modeli matematycznych opartych na prawie działania mas, metody analizy takich modeli oraz badania ich wrażliwości na parametry, a na zakończenie wersje stochastyczne oparte na łańcuchach Markowa. Kolejny rozdział przedstawia dwa konkretne modele, których konstrukcja i różnorodne analizy są zawarte w cyklu publikacji: 1) model odpowiedzi na szok termiczny w komórkach jądrazystych i 2) model tworzenia się filamentów pośrednich *in vitro*. Te rozdziały zarysowują podstawy do następującego teraz przedstawienia problemów badania takich modeli metodami obliczeniowymi, co, jak zaznacza autor, było jego głównym obiektem zainteresowania w przedstawianych pracach i stanowi jego oryginalny wkład do nauki. Są to: 1) techniki konstruowania i testowania modeli, a w szczególności estymacji parametrów, walidacji modeli, sprawdzania identyfikowalności modeli, oraz porównania modeli deterministycznych i stochastycznych dotyczących tych samych procesów, 2) metody dekompozycji modeli, a w tym „mutanty knockdown”, mody elementarnych przepływów, oraz dekompozycja odnosząca się do układów sterowania, 3) techniki modyfikacji modeli, jak obliczeniowe heurystyki upraszczania modeli biologicznych i problemy formalnego uszczegółowienia modeli (tutaj autor podkreśla swój szczególny wkład), 4) metody porównywania podmodeli (różnych wersji fragmentu większego modelu), określonych jako matematyczne sterowanie porównywaniem modeli i jego rozszerzona wersja oraz lokalne porównywanie podmodeli. Ten rozdział jest zakończony syntetycznym omówieniem zastosowania przedstawionych metod do analizy modelu odpowiedzi na szok termiczny. W części ogólnej zawarte jest jeszcze przedstawienie oryginalnych elementów zawartych w poszczególnych publikacjach oraz wnioski i perspektywy dalszych badań. Bibliografia części ogólnej zawiera 141 pozycji literaturowych.

Część ogólna rozprawy jest napisana ciekawie, poprawnym językiem i stanowi dobre, obszerne wprowadzenie w problematykę cyklu publikacji w jej wszystkich aspektach, a w szczególności w tematykę analizy struktury modeli, która stanowi podstawę ubiegania się mgr Mizery o stopień doktora. Doktorant podjął również próbę nadania ogólnego charakteru metodom, które zostały przedstawione w artykułach w zastosowaniu do konkretnych modeli. Jest to ważny element rozprawy, gdyż istnieje obecnie wyraźna potrzeba skodyfikowania postępowania badawczego przy tworzeniu i analizie rozbudowanych modeli układów biologicznych, które coraz częściej wymykają się analizie intuicyjnej. Warto tu zauważyć, że w przypadku badanych przez autora modeli, aspekty „biologiczne” zostały sprowadzone do modelowania reakcji chemicznych, pomijając specyficzne dla komórki problemy, jak np. rozkład przestrzenny jej struktur wewnętrznych czy rola enzymów. Pozwala to zachować stosunkowo prostą strukturę matematyczną modeli (nieliniowe wyrażenia mają postać wielomianów), a oparcie modelu na prawie działania mas, chociaż możliwe w każdym przypadku reakcji biochemicznych, może stać się nieefektywne przy konieczności uwzględnienia reakcji enzymatycznych. Trzeba też przy okazji wskazać, że definiowanie reakcji odwracalnej jako złożenia dwu reakcji nieodwracalnych (jak to proponuje autor), chociaż wygodne z punktu widzenia informatyka czy matematyka, jest sprzeczne z prawami fizyki, ponieważ reakcje nieodwracalne mogą być rozpatrywane tylko jako pewne przybliżenia przydatne w określonych warunkach.

Cykl oryginalnych publikacji składa się z 9 artykułów, z czego trzy ukazały się w czasopiśmie o międzynarodowym zasięgu, jeden w zbiorowym wydaniu książkowym, jeden w materiałach konferencyjnych wydanych przez Springer, a 4 zostały przedstawione w postaci manuskryptów, z których jeden został przyjęty do druku a trzy zgłoszone do publikacji w czasopiśmie o międzynarodowym zasięgu. Wszystkie artykuły są

współautorskie, a mgr A. Mizera jest pierwszym autorem czterech z tych artykułów. Z przedstawionego przez niego oświadczenia wynika, że w ramach zespołów autorskich był on odpowiedzialny m. innymi za numeryczną implementację modeli oraz metod ich analizy, przeprowadzenie obliczeń, numeryczną estymację parametrów, walidację modeli, analizę wrażliwości, etc. Prace numeryczne były prowadzone w większości przy użyciu dedykowanego oprogramowania (COPASI, AUTO, Abaqus), ale również wymagały programowania w takich środowiskach jak Matlab i R. Obliczenia numeryczne dotyczyły rozwiązywania równań różniczkowych zwyczajnych i cząstkowych oraz równań stochastycznych, metod optymalizacji, walidacji i badania identyfikowalności modelu, analizy wrażliwości oraz bardziej specyficznych procedur związanych z nowymi metodami badań prowadzonych przez autorów. Doktorant udowodnił również twierdzenia matematyczne zawarte w artykułach IV i VIII. Mgr Mizera brał również udział w opracowywaniu koncepcji badań i planowaniu artykułów, współpracował przy opracowywaniu nowych metod badawczych oraz opracowywaniu wyników, pisał fragmenty artykułów oraz napisał artykuły, których jest pierwszym autorem. Podjęcie badań w szerokim zakresie problemów i opanowanie różnorodnych metod matematycznych i numerycznych zastosowanych w omawianych pracach jest znaczącym osiągnięciem Doktoranta. Samodzielne opracowanie czterech artykułów oraz obszernej części ogólnej rozprawy dobrze ilustruje jego możliwości jako autora publikacji naukowych.

Sześć pierwszych artykułów cyklu jest poświęconych sformułowaniu i analizie modelu matematycznego odpowiedzi na szok termiczny. Model ten powstał w oparciu o schemat oddziaływań wewnątrzkomórkowych oparty na pracach eksperymentalnych Morimoto i współautorów. Dane eksperymentalne z jednej z tych prac posłużyły również do estymacji parametrów modelu. Problem polegał na znalezieniu 17 wartości parametrów (w tym właściwych parametrów modelu i wartości zmiennych w stanie ustalonym, jednocześnie będących wartościami początkowymi zmiennych modelu), a do estymacji skorzystano z 12 punktów eksperymentalnych na krzywej przebiegu czasowego jednej ze zmiennych modelu (Wykres 1A w artykule I, wykres 1 w artykule II). Autorzy nie zwrócili uwagi na fakt, że tak sformułowane zagadnienie jest równoważne (pomijając statystyczny rozrzut wartości pomiarowych) próbie rozwiązania 11 równań na 17 niewiadomych. Wiadomo, że tak sformułowany problem nie ma na ogół jednoznacznych rozwiązań. Numeryczna estymacja parametrów dała dokładnie taki wynik: niejednoznaczność rozwiązań. Autorzy poświęcili wiele wysiłku i wykazali dużą pomysłowość, aby można było uznać jeden z otrzymanych zestawów parametrów jako właściwy. Trzeba tu było wykorzystać dodatkowe informacje i ciekawym pomysłem było skorzystanie z danych jakościowych lub półilościowych pochodzących z innych eksperymentów. Przeszukując wielowymiarową przestrzeń parametrów metodą *latin hypercube sampling*, wybierając ich zestawy odpowiadające dobremu dopasowaniu do krzywej eksperymentalnej, a następnie porównując przewidywania modelu z danym zestawem parametrów w innych sytuacjach lub dla innych zmiennych, autorzy mogą odrzucić większość z tych zestawów jako nie odpowiadających posiadanym informacjom o badanym procesie biologicznym. Metoda ma charakter heurystyczny, ale w tym konkretnym przypadku okazała się przekonującą, chociaż oczywiście można tu zgłaszać dalsze pytania. Ostateczny model wraz z wyestymowanymi parametrami dobrze opisuje (przynajmniej jakościowo) dosyć szeroki zestaw informacji o układzie odpowiedzi na szok termiczny, tym niemniej nie najlepsze dopasowanie rozszerzenia modelu przedstawione na wykresie 1B w artykule I oraz zbyt szybka odpowiedź układu na oscylacyjne zmiany temperatury, omówiona w artykule IX, sugerują, że w przyszłości może okazać się konieczne uwzględnienie innych aspektów problemu, jak np. procesy transportowe w komórce i rozbudowany opis procesu ekspresji genu.

Podstawowy model odpowiedzi na szok termiczny jest sformułowany i dyskutowany w artykułach I i II. Ważnym aspektem tej analizy jest ściśle powiązanie stanu ustalonego układu i dynamicznej odpowiedzi na szok termiczny. Wszystkie wartości zmiennych w stanie ustalonym oraz wartości parametrów modelu są traktowane jako nieznane i wyznaczone w oparciu o dane eksperymentalne (patrz wyżej). Artykuł III przynosi uproszczoną wersję tego modelu (redukcja z 14 do 10 zmiennych) uzyskaną poprzez analizę kinetyki poszczególnych zmiennych i wyeliminowanie tych, które mają zaniedbywalny wpływ na przebieg reakcji lub charakteryzują się szybkim dochodzeniem do stanu ustalonego, ale przy zachowaniu dobrego dopasowania do danych eksperymentalnych. Ta problematyka, o czym autorzy nie wspominają, ma swoje ściśle matematyczne metody, jak np. metoda stanów quasi-ustalonych czy metoda małego parametru, pozwalające jednocześnie na określenie warunków stosowania danego przybliżenia. Podobnie problem (nie)zależności stanu ustalonego od wartości początkowych postawiony w tej pracy jest dobrze znanym problemem matematycznym stabilności punktu stacjonarnego i obszaru jego przyciągania. Uproszczony model został poddany kilku testom numerycznym, aby sprawdzić czy jego zachowanie po modyfikacjach zgadza się z przewidywaniami pełnego modelu. Artykuł IV poświęcony jest stochastycznej wersji dynamiki układu i pokazuje, że typowe zachowania się układu stochastycznego mają podobny przebieg jak przewidywania wcześniej sformułowanego modelu deterministycznego. Ta obserwacja pozwala uzasadnić stosowanie modelu deterministycznego w sytuacji, kiedy ilość danych białek w komórce nie jest zbyt duża.

Najciekawsze rozważania zawierają artykuły V i VI, gdzie poddana została analizie struktura dynamiczna (sterowania) modelu (w wersji uproszczonej, artykuł III) i rola zawartych w nim trzech pętli sprzężenia zwrotnego. W tym celu sformułowane są „zmutowane” modele pozbawione wszystkich, dwu lub tylko jednej pętli zwrotnej we wszystkich możliwych kombinacjach (nazywane przez autorów „*knockdown mutant models*”). Tak zmutowane modele w liczbie siedmiu są dopasowane do tych samych danych co pełny model na zasadach opracowanych we wcześniejszych pracach (Artykuł V). Następnie modele są testowane względem spełniania dodatkowych, jakościowych, ale sformułowanych „empirycznie” jako ilościowe, kryteriów na dwie ze zmiennych modelu w dwu różnych temperaturach. Celem tej części pracy jest ustalenie, które z mutantów spełniają dane kryterium. Ciekawym pomysłem jest tu zapis logiczny (Boolowski) istnienia lub nieistnienia danej pętli sprzężenia w danym mutancie, a następnie połączony zapis dla tych mutantów, które spełniają dany warunek. Przeprowadzenie operacji logicznych pozwala z kolei na znalezienie tych pętli sprzężenia, które są odpowiedzialne za spełnienie danego warunku. W kolejnym artykule VI, mutanty są badane w całej przestrzeni parametrów modelu (wykorzystano tu metodę *latin hypercube sampling*), a przewidywania każdego z mutantów są porównywane z przewidywaniami pełnego modelu dla tych samych parametrów przy pomocy wskaźników takich jak maksymalne wartości wybranych zmiennych oraz powierzchnie pod krzywymi opisującymi odpowiedź układu na szok termiczny. Pomysłowe ujęcie wyników w postaci wykresów uzyskanych w oparciu o wszystkie zestawy użytych parametrów pozwala na scharakteryzowanie zachowania się mutantów niezależnie od przyjętych wartości parametrów. Metody analizy poszczególnych elementów modelu przedstawione w obu tych artykułach są nowe, oryginalne i mogą być wykorzystane przy badaniu innych modeli o podobnej strukturze i charakterze danych eksperymentalnych.

W ostatnim artykule związanym z modelem odpowiedzi na szok termiczny, mgr Mizera połączył model lokalnych reakcji chemicznych tej odpowiedzi z przestrzennym rozkładem temperatury w tkance ogrzewanej cyklicznie promieniowaniem naddźwiękowym, mając na uwadze potencjalne metody hipertermicznej terapii niektórych schorzeń (artykuł IX). Czasoprzestrzenne rozkłady temperatury modelowano przy pomocy równania Pennesa

rozchodzenie się ciepła w ośrodku biologicznym. Praca ta dobrze ilustruje możliwe kliniczne aplikacje opracowanego modelu, ale jednocześnie dostarcza informacji o zachowaniu się układu przy periodycznym krótkim pobudzeniu temperaturowym, które wskazują na konieczność jego rozbudowania dla tego typu zastosowań.

Dwie publikacje (artykuły VII i VIII) są poświęcone modelowaniu procesu powstawania filamentów pośrednich z tetramerów wimentyny w procesie polimeryzacji poprzez kolejne stadia oktamerów, heksadekamerów, najprostszych filamentów o jednostkowej długości (dwa heksadekamery) i powiększania się filamentów poprzez dołączanie tetrameru, filamentu jednostkowego lub filamentu o większej długości. Punktem odniesienia do walidacji modelu są wyniki eksperymentów *in vitro* zawierające informacje o zależności średniej długości filamentu od czasu dla dwu różnych początkowych stężeń tetrameru. Autorzy stanęli tu przed podobnym problemem ustalenia stopnia złożoności modelu koniecznego do opisanie danych eksperymentalnych. W pracy VII porównano dwie wersje modelu: prostą (4 zmienne) i złożoną (5 zmiennych), i badano również podmodele z pominiętymi niektórymi mechanizmami powstawania filamentów. Autorzy opracowali własną formułę obliczania średniej długości filamentów i estymowali, w oparciu o dane pomiarowe o kinetyce średniej długości filamentów dla jednej z początkowych wartości stężeń tetramerów, parametry wszystkich rozpatrywanych wersji modeli, wykorzystując dane dla drugiej wartości początkowej do walidacji modelu. Pozwoliło to na wyróżnienie trzech z ośmiu rozpatrywanych wersji jako pozwalających na poprawny opis danych. Autorzy opracowali również rozszerzony model dla filamentów o różnej długości aż do pewnej wartości granicznej, poza którą wszystkie filamenty są już opisane przez jedną zmienną. Opracowali też metodę doboru parametrów przy zwiększaniu granicznego rozmiaru filamentu tak, aby zachować informacje o parametrach posiadane dla modelu o niższej rozdzielczości. Pozwoliło to na przewidzenie stężeń filamentów o różnych długościach w funkcji czasu trwania eksperymentu; nie można jednak było zweryfikować tego wyniku z braku danych eksperymentalnych. W artykule VIII badano dokładniej zarysowaną już metodę zmiany rozdzielczości modelu nadając jej ogólną postać matematyczną i wykazując kilka twierdzeń dowodzących jej poprawności. Te ogólne rozważania zostały zilustrowane na modelu z wcześniejszego artykułu VII. Autorzy poruszają przy okazji sprawę pominięcia sztucznego założenia o granicznej wielkości uwzględnianej w modelu, warto więc przypomnieć, że takie teorie matematyczne zostały już sformułowane.

Badane w rozprawie modele matematyczne mają prostą dynamikę, nie wykraczającą poza dochodzenie do stanu ustalonego, ich rozmiar (mierzony ilością zmiennych i parametrów) jest nieduży a własności tych układów możliwe do uchwycenia przez specjalistę w danej dziedzinie. Tym niemniej ilość problemów, jakie pojawiają się przy analizie tych modeli i ich porównaniu z danymi empirycznymi jest znaczna, zwłaszcza przy ambitnym założeniu, że wszystkie parametry modelu powinny być wyznaczone na podstawie tych danych. Autorzy, oprócz prowadzenia standardowych analiz, jak np. wrażliwość modelu na zmiany parametrów i identyfikowalność modelu, opracowali szereg oryginalnych metod postępowania związanych przede wszystkim z problemem wyodrębniania i porównywania różnych wersji modelu między sobą i z danymi empirycznymi oraz analizą wewnętrzną struktury oddziaływań pomiędzy zmiennymi modelu. Metody te, chociaż niejednokrotnie związane ze specyficznymi cechami rozpatrywanych modeli, mają duże znaczenie heurystyczne i mogą być stosowane do znacznie większych modeli o podobnej strukturze jak i stanowić podstawę i zachętę do poszukiwania analogicznych metod dla modeli o innej strukturze.

Wkład mgr Andrzeja Mizery w przedstawione w rozprawie prace zespołowe jest znaczący. Doktorant aktywnie uczestniczył we wszystkich etapach rozwiązywania poszczególnych zagadnień i opracowywania wszystkich artykułów; jest on również pomysłodawcą wielu oryginalnych rozwiązań szczegółowych oraz autorem tekstu czterech artykułów. W napisanych przez siebie artykułach, których jest pierwszym autorem (artykuły IV, VI, VIII, IX), samodzielnie zaproponował nowe metody i podejścia do rozważanych zagadnień. Jego ważnym osiągnięciem jest opracowanie i napisanie obszernej części ogólnej rozprawy. Trzeba również podkreślić znaczny zasięg problemów i metod matematycznych i numerycznych, które podejmował i twórczo stosował w swojej pracy badawczej. Uważam, że mgr A. Mizera przekonująco wykazał swoje zdolności do prowadzenia samodzielnej pracy badawczej.

Podsumowując moje opinie o rozprawie mgr Andrzeja Mizery stwierdzam, że spełnia ona warunki do nadania stopnia doktora określone w ustawie o stopniach i tytule naukowym z dnia 14 marca 2003 r. ze zmianami wprowadzonymi ustawami z dnia 27 lipca 2005 r. i z dnia 18 marca 2011 r., i stawiam wniosek o dopuszczenie Doktoranta do dalszych etapów przewodu doktorskiego.

J. Wanicewski