

Katedra Grafiki, Wizji Komputerowej i Systemów Cyfrowych

Politechniki Śląskiej

Recenzja rozprawy doktorskiej

Autor: mgr inż. Michał Kadlof

Tytuł: Komputerowe modelowanie trójwymiarowej struktury chromatyny

Promotor: prof. dr hab. Dariusz Plewczyński

Ogólna charakterystyka rozprawy

Przedłożona do recenzji rozprawa doktorska liczy 170 stron tekstu, składa się z dziesięciu rozdziałów, dodatków A, B, spisu rysunków, tabel, bibliografii oraz skorowidza. Dziesięć rozdziałów tekstu pracy jest podzielonych na trzy części, teoretyczną, praktyczną i dodatki. Praca przedstawia bardzo szeroki obraz zagadnień związanych z problematyką modelowania struktury chromatyny. Obraz ten obejmuje zarówno aspekty biologiczne jak też eksperymentalne, modelowania matematycznego, bioinformatyki, budowy oprogramowania, technik obliczeniowych łącznie z aspektami konstrukcyjnymi klastra obliczeniowego. Oryginalne, naukowe elementy rozprawy leżą głównie w obszarach modelowania matematycznego oraz bioinformatyki, są udokumentowane recenzowanymi publikacjami naukowymi, których współautorem jest Doktorant.

Praca ma charakter monografii. Uzyskane wyniki naukowe, opublikowane w wieloautorskich, recenzowanych artykułach naukowych są w niezależny sposób opisane, z dodatkowymi wyprowadzeniami i wyjaśnieniami.

Część pierwsza pracy (rozdziały 1-5) przedstawia motywację do podjęcia badań, cele naukowe, przedstawienie i pewna formalizacji pojęć i metod wykorzystywanych w pracy.

Rozdział 1, wstęp stanowi bardzo popularne i poglądowe omówienie obecnej wiedzy biologicznej dotyczącej struktury i własności chromatyny. Doktorant przytacza wiele faktów, które uzasadniają złożoność problematyki badania struktury i funkcjonalności chromatyny. Długość łańcucha DNA upakowanego w chromosomach, w jądrze komórkowym, aspekty

funkcji/zadań biologicznych realizowanych przez DNA, które muszą być precyzyjnie realizowane dla funkcjonowania komórek. Na końcu rozdziału wstępnego autor charakteryzuje obszar badań prowadzonych w ramach pracy. Jest nim powiązanie metod modelowania matematycznego, biologii obliczeniowej i strukturalnej oraz bioinformatyki dla budowy i analizy przestrzennych modeli chromatyny. Autor słusznie podkreśla tu interdyscyplinarny charakter problematyki poruszanej w ramach pracy.

W rozdziale 2, na wstępie przytacza się fakty dotyczące budowy fizycznej i chemicznej łańcucha DNA. Następnie Doktorant przechodzi do wyższych poziomów organizacji chromatyny. Przedstawia koncepcje pętli, domen oraz kompartmentów. Wspomina o mechanizmach i czynnikach formowania pętli. Przykładową strukturę kontaktów i kompartmentów ilustruje na (zapożyczonym z literatury) rysunku 2.9. Cytuje za publikacjami z literatury metodę wykorzystania pierwszej składowej głównej macierzy kontaktów do oceny położenia/kształtu kompartmentów chromatyny.

Rozdział 3 poświęcony jest przedstawieniu szeregu technik eksperymentalnych biologii molekularnej użytecznych do badania struktury chromatyny, 3C, 4C, 5C, Hi-C, ChIA-PET. Spośród omówionych metod dwie ostatnie mają obecnie bardzo duże znaczenie poznawcze, pozwalają na tworzenie map kontaktów w łańcuchu DNA, co pozwala na budowę i rozwój algorytmów do badania trójwymiarowej geometrii chromatyny. Doktorant w rozdziale 3 omawia także jeszcze inne metody eksperymentalne możliwe do zastosowania w badaniu chromatyny, ChIA-Drop, FISH, EM-ISH, iPALM. Ostatnie z wymienionych metod są technikami związanymi z mikroskopią. Doktorant dyskutuje charakterystyczne cechy, zalety i wady dostępnych technik eksperymentalnych.

Rozdział 4 poświęcony jest modelowaniu chromatyny na poziomie mezoskopowym. Doktorant przedstawia przegląd mezoskopowych modeli mechanicznych budowy łańcucha DNA, poczynawszy od łańcucha idealnego, jednowymiarowego, o doskonałej wiotkości, poprzez kolejne modele proponowane w literaturze, Kratky-Poroda, WLC, model OSF oraz model fraktalny. Dyskutuje zasady ich budowania oraz własności powstałych modeli.

Wreszcie ostatni w części pierwszej, rozdział 5 poświęcony jest zasadom modelowania struktur wewnętrzządrowych oraz możliwości ich weryfikacji. Dla obrazowego sformułowania swojej oceny problematyki modelowania chromatyny, Doktorant przeciwstawia problematykę modelowania chromatyny oraz problematykę modelowania struktur białek. Dyskutując różnice w tych zasadach modelowania formułuje słuszną tezę, że zarówno sam proces modelowania chromatyny jak też możliwości weryfikacji modeli struktury i funkcji chromatyny stanowią większe wyzwanie naukowe niż procesy modelowania struktur białkowych. W aspekcie faktu, że przeprowadzenie projektu modelowania struktury chromatyny wymaga podjęcia decyzji o wielu parametrach algorytmu modelowania oraz o jego krokach i kolejności ich wykonywania, Doktorant w ostatnich podpunktach rozdziału 5 stara się sformułować zasady określania strategii modelarskiej dla chromatyny. Decyzje, które należy podjąć dotyczą takich parametrów jak

skali modelowania, własności symulowanego łańcucha DNA, własności otoczenia badanego polimeru, lub jego części.

Część druga pracy obejmująca rozdziały od 6 do 10 nazwana jest przez Doktoranta częścią praktyczną. Zawiera ona opisy algorytmów/narzędzi modelowania opracowanych w ramach pracy oraz opisy ich zastosowania.

Rozdział 6 stanowi pewnego rodzaju wstęp do części drugiej. Doktorant wymienia tradycyjnie definiowane/wykorzystywane w genomice struktury wewnątrzchromosomowe, geny, elementy regulacyjne, centromery, telomery oraz przedstawia ich odpowiedniki w algorytmach modelowania kształtu i struktury chromatyny, domeny i ich definicje TAD, CCD. W dalszej części oraz na ilustracjach przedstawia struktury wykorzystywane w modelowaniu geometrii trójwymiarowej łańcucha DNA, kontakty, mapy kontaktów, pętle, diagramy łuków. Opisuje możliwe metody detekcji zdefiniowanych struktur wewnątrzchromosomowych. W podpunktach 6.1 oraz 6.2 doktoratu autor przedstawia dwa opracowane przez siebie narzędzia do modelowania struktury chromatyny, CCD Caller oraz MetaTAD Caller. Dwa opisane narzędzia mogą być dopasowane do rodzaju dostępnych danych eksperymentalnych stanowiących informacje wejściowe dla modelowania, Hi-C lub ChIA-PET. Dla obu narzędzi dostępne są otwarte kody źródłowe oprogramowania.

W rozdziale 7 opisany jest problem skalowania wielowymiarowego (MDS, multidimensional scaling) oraz proponowane dla jego rozwiązania algorytmy jako narzędzia modelowania trójwymiarowej struktury chromatyny. W podrozdziale 7.1 Doktorant proponuje schemat wykorzystania podejścia do konstrukcji trójwymiarowego modelu struktury chromatyny z wykorzystaniem algorytmów dystansu grafowego.

W rozdziale 8 dokładnie opisany jest model oraz związane z nim algorytmy obliczeniowe oznaczane w pracy doktorskiej najczęściej skrótem SM, (spring model, model sprężynkowy, model kulkowo - sprężynkowy). Model ten jest związany z kilkoma publikacjami Doktoranta. Główna idea konstrukcji modelu SM przedstawiona jest w punkcie 8.1. W algorytmie oceny struktury chromatyny dokonuje się iteracyjnej modyfikacji jej konformacji tak aby minimalizować energię polimeru DNA. W punkcie 8.2 opisana jest funkcja energii polimeru DNA składająca się z energii wiązania pomiędzy kolejnymi kulami, energii zgięcia związanej z kątem płaskim pomiędzy trójkami kolejnymi kul, energii wyłączonej objętości oraz energii interakcji (odległych wiązań) pomiędzy kulami tworzącymi oddziałującą parę. Dla energii wiązania kolejnych kul oraz dla energii zgięcia przyjęte są modele harmoniczne. Energia wyłączonej objętości wyliczana jest na podstawie modelu Lennarda – Jonesa. Energia wiązań odległych jest wyznaczana na bazie modelu harmonicznego z obcięciem. Kilka podpunktów rozdziału 8 poświęconych jest możliwym sposobom ustalania początkowej struktury polimeru chromatyny do dalszych obliczeń metodą SM. Struktury polimerów DNA analizowane w trakcie obliczeń, a także struktury wejściowe i wyjściowe są przechowywane/zapisywane w formacie PDB. Podawane są przykłady zastosowania algorytmu i oprogramowania SM, badanie wpływu delecji regionu

CTCF na przykładzie locusa genu TAL1 (za pracami w literaturze), konstrukcja globuli fraktalnej zgodnie ze scenariuszem literaturowym polegającym na ściągnięciu fragmentów globuli do środka sfery. Rozdział 8 kończy opis struktury i funkcjonalności opracowanego oprogramowania oraz aplikacji internetowej. Przedstawione są możliwości obliczeniowe i możliwe warianty obliczeń.

W rozdziale 9 przedstawione jest jeszcze jedno podejście do modelowania struktury chromatyny, związane z dążeniem do dopasowania wyników modelowania do eksperymentów. Polega ono na konstrukcji algorytmów modyfikacji modelowanej struktury polimeru DNA sterowanych obrazami. Obrazy uzyskiwane technikami mikroskopowymi (EM-ISH, IPALM, STED) służą do zdefiniowania potencjałów lub pól sił powodujących odpowiednią modyfikację struktury modelowanej chromatyny. Wspomina się dwie metody implementacji eksperymentalnych danych obrazowych w algorytmach modelowania chromatyny, metodę siłową i metodę korelacyjną. Obie metody prowadzą do modyfikacji pola sił oddziałujących na polimer DNA. Kolejno obie te metody są dokładniej omówione. Następnie przedstawione są przykłady obliczeniowe. Pierwszy z nich dotyczy obrazu wygenerowanego sztucznie, na podstawie modelu fragmentu polimeru DNA, a drugi bazuje na rzeczywistym obrazie mikroskopowym jądra komórkowego rzodkiewnika.

Rozdział 10 jest opisem infrastruktury obliczeniowej używanej w pracach związanych z pracą, której Doktorant jest autorem i projektantem.

Ocena rozprawy

Recenzowaną rozprawę oceniam bardzo pozytywnie. Bardzo zwracając uwagę cecha pracy, to szeroki zakres tematyczny obejmujący podstawy biologiczne zasady modelowania matematycznego, obszerne bioinformatyki polimerów DNA, wyczerpujące omówienie technik eksperymentalnych orientowanych na badanie budowy polimerów DNA, budowy oprogramowania, technik obliczeniowych łącznie z aspektami konstrukcyjnymi klastra obliczeniowego. Doktorant wykazuje się tym samym szeroką wiedzą i opanowaniem aspektów modelowania, obliczeń, konstrukcji oprogramowania na bardzo szczegółowym i zaawansowanym poziomie. Uzyskiwane wyniki naukowe mają walory interdyscyplinarne. Motorem wielu przeprowadzonych w ramach doktoratu prac jest zrozumienie podstaw biologicznych zjawisk. Opanowany przez Doktoranta warsztat naukowy ma potencjalnie bardzo wiele zastosowań w medycynie, farmakologii, klinicyście.

Warte podkreślenia jest, że praca jest napisana bardzo ładnym, żywym i jasnym językiem. Zawiera formalizmy matematyczne sformułowane poprawnie i odpowiednio charakteryzujące przeprowadzone badania z zakresu modelowania matematycznego.

Za najważniejsze, oryginalne osiągnięcie naukowe pracy uważam opisany w rozdziale 8 algorytm modelowania bazujący na strukturze koralikowo – sprężynkowej, a także w

publikacjach współautorskich Doktoranta. Jest to system o złożonej strukturze, skomplikowanej postaci funkcji energii oraz o potencjalnie złożonej dynamice molekularnej. Uzyskanie poprawnych wyników modelowania, weryfikowanych porównaniem uzyskanych struktur z wynikami eksperymentów jest dużym osiągnięciem naukowym.

Także inne opisywane wyniki naukowe, jak metody wyszukiwania domen chromatynowych, dopasowanie struktur sterowane obrazem są bardzo wartościowe.

Publikacje naukowe, których współautorem jest Doktorant (dodatek A), tematycznie związane z modelowaniem genomu ukazały się w bardzo renomowanych, wysoko punktowanych czasopismach naukowych.

Uwagi krytyczne i dyskusyjne, pytania do doktoratu

Szerokość tematyki badawczej podjętej w pracy powoduje, że wiele zagadnień przedstawionych jest pobieżnie. Przydałoby się więcej przykładów użycia opisywanych w pracy algorytmów do rzeczywistych danych, a także szersze porównania opracowanych metod z wynikami algorytmów z literatury.

Wyniki rozdziału 5 o zasadach modelowania chromatyny mają zbyt mało odniesień w innych częściach pracy.

W pracy brak jest rozdziału, który stanowiłby podsumowanie przeprowadzonych prac oraz uzyskanych wyników.

W pracy jest dość dużo usterek edycyjnych. Przykłady są przedstawione poniżej

Str 104:

Okazuje się, taka częstość w globuli fraktalnej jest zbliżona do częstości obserwowanej w kontaktach chromatyny dlatego. Z tego powodu może być wykorzystywana jako model jądra komórkowego. (urwane zdanie).

str 92:

... kąty pomiędzy każdym trzem sąsiadującymi punktami (błąd gramatyczny)

str 91, 96, 97, 98:

Okręg (powinno być okrąg)

Str 111:

Minimalizacją energii z opcjonalną relaksacją (błąd gramatyczny)

Str 111:

Pytania:

1. Na rysunku 8.14 przedstawiono schemat generowania oraz rysunek wygenerowanego modelu fraktalnego jądra komórkowego. Skąd wiadomo, że wygenerowana struktura spełnia prawo potęgowe. Ile wynosi wartość alfa? Czy ta wartość zależy od jakichś parametrów modelowania?
2. W pracy w rozdziale 8 wspomina się o termostacie Andersena. Użycie tej idei wymagałoby dodania członu energii kinetycznej do bilansu energii. Czy taka modyfikacja jest dokonywana?
3. Czy bada się wpływ wartości temperatury na konformację chromatyny?
4. Wspominana praca Andersena zawiera nie tylko koncepcję termostatu ale także barostatu. Czy w eksperymentach kontroluje się ciśnienie? Czy ma ono jakiś wpływ na wyniki?
5. W pracy cytuje się artykuł [41] o modelowaniu polimerów. W artykule tym stosuje się inną koncepcję kontroli temperatury, przez odpowiednie zdefiniowanie współczynnika tłumienia w równaniach Langevina oraz dodanie do każdego bilansu sił dodatkowej losowej siły. Czy ta idea jest stosowana w opracowanych algorytmach? Czy porównywano ją z koncepcją termostatu?

Konkluzja

Praca stanowi podsumowanie oryginalnych, interdyscyplinarnych badań naukowych, w których sformułowano i weryfikowano oryginalne hipotezy badawcze. Dokumentuje wiedzę i kompetencje Doktoranta. Osiągnięcia i oryginalne elementy pracy są na pewno wystarczające do jej ogólnej pozytywnej oceny. Stwierdzam, że rozprawa spełnia warunki stawiane pracom doktorskim i wnioskuję o jej dopuszczenie do publicznej obrony.

Ponadto, pomimo uwag krytycznych, biorąc pod uwagę poziom naukowy pracy, opanowania przez Doktoranta bardzo szerokiego zakresu tematycznego oraz dorobek publikacyjny Doktoranta wnioskuję o wyróżnienie rozprawy.

