

KRAHASIMI I METODAVE TË MODELIMIT TË SISTEMEVE BIOLOGJIKE

*BIMBARI B., SAATÇIU D.

Universiteti i Tiranës, Fakulteti i Shkencave të Natyrës, Departamenti i Informatikës

e-mail: bora.bimbari@fshn.edu.al

Përmbledhje

Sistemet biologjike janë sisteme komplekse, të përbërë nga shumë pjesë që bashkëjetojnë dhe bashkëveprojnë duke përcaktuar sjelljen e të gjithë sistemit. Përbërjet dhe sjelljet e sistemeve janë me interes për t'u studiuar dhe kuptuar. Në mënyrë tradicionale, për këto sisteme ngrihen hipoteza të cilat provohen nëpërmjet eksperimenteve laboratorike. Megjithatë, gjendja aktuale e teknologjisë lejon simulime kompjuterike të sistemeve biologjike dhe hipotezave për to (*in silico*). Kjo do të thotë që ne mund t'i studiojmë sistemet dhe sjelljet e tyre duke i paraqitur në mënyra të kuptueshme për kompjuterat. Sigurisht, përgjigja përfundimtare do të provohet në laboratore, me lëndë reale biologjike (*in vitro*). Studimi i sistemeve nëpërmjet kompjuterave shtron sfidën e ideimit të metodave për paraqitjen e sistemeve me mënyra që kapin të gjitha karakteristikat e sistemeve biologjike dhe që janë të kuptueshme për kompjuterat. Qëllimi i këtij studimi është të përshkruajë dhe krahasojë metodat më efikase të modelimit të sistemeve biologjike.

Abstract

Biological systems are complex and composed of many concurrent parts which interact, and thus define the behavior of the system as a whole. Systems composition and behavior is of study interest. Traditionally, we raise hypothesis on the systems, and these hypothesis need to be proven experimentally, in laboratories. Anyways, technological state of art indicates that we can simulate biological systems and related hypothesis computationally (*in silico*). This means that we can study biological systems by representing them so that they can be read and interpreted by computers. The ultimate results would be needed to be proven experimentally in any case (*in vitro*). Computational study of biological systems demands the elaboration of methods for representing such systems by capturing all biological systems features, and that are understandable by computers. This study aims to explore and compare the most efficient modelling techniques for biological systems.

Fjalëkyçe: modelim, simulim, algjebra procesi, sisteme biologjike.

Hyrje

Modelimi i objekteve apo i proceseve të botës reale (të cilave do t'ju referohemi si sisteme) realizon paraqitjen e tyre në trajta të ndryshme, duke abstraguar në nivele të ndryshme. Për një sistem mund të ndërtohen disa modele në varësi të detajeve që kemi interes të studiojmë në një moment të caktuar, pra në varësi të

këndvështrimit për sistemin. Modelet e ndërtuara lejojnë analizë të sistemeve të cilat ato paraqesin dhe shpesh analizat e modeleve nuk kërkojnë eksperimente reale me sistemin real.

Ndërkohë që modelimi përdoret prej njerëzve prej kohësh, zhvillimi i modeleve që analizohen nëpërmjet kompjuterave dhe programeve kompjuterikë është i lidhur me zhvillimet teknologjike të dekadave të fundit. Pikërisht realizimi i analizave të modeleve nëpërmjet kompjuterave zëvendëson eksperimentet reale dhe në këtë rast themi që simulojmë sistemet nëpërmjet kompjuterave. Psh. nëse konsiderojmë një sistem biologjik (një zinxhir reaksionesh të sintezës së lëndëve të ndryshme në organizëm, aktivitetet e gjeneve, etj), të gjitha hipotezat e ngritura për të do të duhet të provohen nëpërmjet eksperimenteve laboratorike.

Ky është një proces i kushtueshëm, i komplikuar dhe ndonjëherë edhe i ngadaltë. Ky proces mund të përmirësohet duke e mbështetur me simulime kompjutacionale. Pra, duhet të ndërtohen modele sistemesh biologjike, të cilat mund të simulohen nëpërmjet platformave të përshtatshme softuerike dhe harduerike. Simulimet do të kryhen me të dhëna të përfutuara në laborator dhe janë shumë më të shpejta dhe efçente për vërtetimin ose jo të hipotezave. Për më tepër, simulimet lejojnë që të monitorohet njëkohësisht sjellja e disa njësive të sistemit, të ndryshohen parametra sasiorë të eksperimentit dhe në disa raste të analizohet sistemi në shkallë detajimi më të lartë sesa lejon eksperimenti laboratorik. Kjo mund të çojë edhe në krijimin e hipotezave të reja dhe në të kuptuarin e sjelljeve të sistemeve të cilat nuk kapen nga intuita biologjike. Në përfundim të analizës kompjuterike, rezultatet do të duhet të provohen në eksperimente laboratorike, megjithatë procesi i analizës përmirësohet ndjeshëm si nga ana sasimore, ashtu edhe nga ana cilësore.

Për të modeluar sistemet biologjike janë zhvilluar dhe përdorur mënyra të shumta, të cilat kapin disa nga karakteristikat përcaktuese të përbërjes dhe sjelljes së këtyre sistemeve. Sfidat më të madhe mbetet zhvillimi i metodave të modelimit që i kapin të gjithë elementët e sistemit që janë me rëndësi për studimin.

Materiali dhe metodat

Natyra e sistemeve biologjike

Sistemet biologjike janë sisteme komplekse, të përbërë nga shumë pjesë, të cilat janë bazë ndërtimore, por edhe funksionale e tyre. Këto pjesë të veçanta ndërveprojnë me njëra-tjetrën, në kontekstin e të tërë sistemit dhe, ky ndërveprim së bashku me ndikimin e faktorëve të jashtëm mbi sistemin përcakton sjelljet (dinamikën) e sistemit. Pra, sistemet biologjike mund të shikohen si sisteme modulare, të cilat paraqesin ndërveprime dhe evolime.

Sistemet biologjike klasifikohen në klasën e sistemeve hibride. Pra, sistemet biologjike kanë komponente që shfaqin sjellje të vazhdueshme dhe komponente që shfaqin sjellje diskrete gjatë jetëgjatësisë së tyre. Psh. ndryshimi i sjelljes së sistemit pasi është arritur një prag i përqëndrimit të përbërjeve që marrin pjesë në reaksion përbën sjellje diskrete të sistemit. Kjo ngjarje diskrete ndikon tek elementët e vazhdueshëm të sistemit, që në këtë rast do të ishin përqëndrimet e përbërjeve të reaksionit.

Sistemet biologjike janë sisteme me sjellje të rastit. Kjo do të thotë që faktorë të ndryshëm ndikojnë sjelljet e sistemeve në mënyrë jo deterministe.

Tiparet e sipërpërmendura janë përcaktuese dhe themelore për sistemet biologjike, prandaj metodat e modelimit të sistemeve biologjike duhet të ofrojnë mekanizma për paraqitjen e tyre.

Modelimi matematik

Trajta më e hershme e modelimit të proceseve biologjike është shprehja lidhjes mes sasive së elementëve që marrin pjesë në reaksion dhe ndryshimit e tyre me kalimin e kohës nëpërmjet funksioneve matematike. Modele të tilla quhen modelet matematike. Modelet matematike janë të përshtatshme për paraqitjen e sjelljeve të vazhdueshme të sistemeve biologjike. Nëse ato përmbajnë faktorë probabilitarë në përcaktimin e tyre, atëherë arrijnë të modelojnë edhe natyrën jo deterministe të sistemeve.

Modelet matematike nuk janë zhvilluar për t'u analizuar me kompjutera, ato kanë ekzistuar më parë. Për më tepër, këto modele janë një tërësi ekuacionesh të cilat nuk përmbajnë algoritmin për zgjidhjen e tyre. Megjithatë, edhe për këto modele bëhen përpjekje për të ndërtuar programe specifike, të cilët i lexojnë dhe kuptojnë këto modele dhe, simulojnë sistemet e paraqitur prej tyre. Në këtë mënyrë, mund të merren një sërë zgjidhjesh për ekuacionet e modelit (Fisher & Henzinger, 2007).

Modelimi kompjutacional

Kohët e fundit, formalizmat e informatikës janë parë si të përshtatshme për modelimin e sistemeve biologjike. Sistemet shikohen si sekuenca veprimesh apo rregullash që përshkruajnë sjelljen e sistemit si një tërësi kalimesh nga një gjendje në tjetrën. Pra, thelbi i modeleve të tilla është paraqitja e të gjitha gjendjeve të mundshme nëpër të cilat mund të ndodhet sistemi dhe kalimin nga njëra gjendje në tjetrën. Një gjendje e sistemit përcaktohet nga kombinimi i vlerave të variablave të përcaktues të sistemit dhe kalimi nga një gjendje në tjetrën përcaktohet nga ndryshimi i vlerave të variablave. Faktorët që ndikojnë në ndryshimin e vlerave quhen shkaktarët e kalimit të sistemit nga një gjendje në tjetrën.

Metoda modelimi që ndjekin formalizmat e informatikës janë:

Rrjetat buleane (boolean networks). Rrjetat buleane janë forma më e hershme e modelimit kompjutacional të sistemeve biologjike. Çdo molekulë (psh. e një gjeni ose proteine) e zinxhirit të reaksioneve që zhvillohen konsiderohet ose si aktive ose si pasive (pra pjesë e reaksionit ose jo), ndërsa gjendjet e ndërmjetme anashkalohen. Gjendja e sistemit përcaktohet nga gjendjet aktive të të gjithë molekulave pjesë e sistemit. Gjendjet e mundshme të sistemit paraqiten si kulme të grafit që paraqet rrjetin, ndërsa kalimet nga njëra gjendje në tjetrën paraqiten nga brinjët e grafit. Rrjetat buleane e thjeshtojnë shumë dinamikën e sistemeve të mëdhenj dhe kjo bën që nëpërmjet tyre të bëhen analiza eficiente. E metë e tyre është paaftësia për të paraqitur struktura hierarkike ose kompozicionale, të cilat lehtësojnë analizat e sistemeve.

Rrjetat Petri (Petri Nets) (Reisig, 1985). Rrjetat Petri janë një formalizëm i përdorur gjerësisht për modelimin e sistemeve të shpërndarë. Ky formalizëm thekson ekzistencën dhe veprimtarinë e njëkohshme të pjesëve të sistemit dhe ky element është shumë i rëndësishëm në trajtimin e sistemeve biologjike. Rrjeta Petri është një graf i orientuar me dy tipe kulmesh: *kulme burime* të cilat paraqesin burimet e sistemit dhe *kulme tranzicione* që paraqesin ngjarjet që mund të shkaktojnë kalime të sistemit nga njëra gjendje në tjetrën. Brinjët lidhin kulmet tranzicione me kulmet burime. Kulmet burime shoqërohen me elementë dallues të cilët u korrespondojnë vlerave të burimeve. Në këtë mënyrë paraqiten gjendjet e mundshme të sistemit. Caktimi i vlerave të ndryshme dalluese përcakton kalimin nga një gjendje tek tjetra. Mqs nga një gjendje mund të kalohet në disa të tjera, rrjetat Petri kapin edhe natyrën jo deterministe të sistemeve.

Rrjetat Petri janë të përshtatshme për modelimin e aspektit të njëkohshmërisë së sistemeve biologjike dhe shpesh përdoren për paraqitjen e zinxhirëve të reaksioneve metabolikë dhe të sintezës së proteinave. Avantazh i rëndësishëm i tyre është aftësia për paraqitje grafike të sistemeve dhe ekzistenca e shumë programeve për analizën e sistemeve të paraqitura nëpërmjet tyre. Megjithatë, edhe rrjetat Petri vuajnë në paraqitjen strukturore të sistemeve të përbërë nga shumë pjesë.

Modelet e kinetikës kimike. Këto modele i paraqesin proceset qelizore si një sistem reaksionesh kimike të dallueshme nga njëri tjetri. Tërësia e gjendjeve në të cilat mund të ndodhet sistemi dhe kalimi nga njëra gjendje në tjetrën përcaktohet nga sasia e molekulave (përqëndrimi) të pranishme në proceset qelizore. Kinetika kimike analizohet duke përdorur ekuacionet diferenciale. Këto ekuacione supozojnë që variablat të cilat ata përshkruajnë kanë gjithmonë vlera të vazhdueshme dhe që proceset që ato përshkruajnë janë deterministe. Këto dy

supozime e thjeshtojnë shumë një sistem biologjik, i cili shpesh ka pjesë diskrete dhe sjellje jo deterministe, por me devijime rastësore (Priami & Quaglia, 2004).

Bazat e të dhënave funksionale me objekte. Këto sisteme ruajnë informacione për sistemin e reaksioneve molekulare (psh. EcoCyc, MPW, etj) dhe janë mekanizma për studimin e sistemeve biologjike. Ato përdorin objekte për të realizuar paraqitje hierarike të njërive molekulare. Shumica e bazave të të dhënave ofrojnë paraqitje grafike dhe mjete për kërkimin e të dhënave që janë ruajtur në to. Bazat e të dhënave funksionale rezultojnë të përshtatshme për ruajtjen, organizimin dhe vizualizimin e të dhënave të sistemeve të reaksioneve molekulare. Si të tilla, ato nuk u përgjigjen të gjitha nevojave për analizën e sistemeve, veçanërisht atyre nevojave që kanë të bëjnë me analizën e dinamikave të tyre. Rrjedhimisht, ato ofrojnë pak ose aspak mundësi për simulime të reaksioneve (Priami & Quaglia, 2004).

Mënyrat e modelimit të sipërpërmendura kanë secila attribute të veçanta, prandaj janë më të përshtatshme për të paraqitur aspekte të caktuara të sistemeve, por jo për t'i paraqitur ato në mënyrë të tillë që të mund të kryhet një studim i të gjitha aspekteve të tyre. Asnjëra prej tyre nuk ofron njëkohësisht mundësi për modelimin e natyrës kompozicionale të sistemeve biologjike, sjelljes së tyre diskrete dhe të vazhdueshme. Nga ana tjetër, algjebrat e proceseve janë metoda modelimi të ideuara pikërisht për paraqitje të sistemeve me komponentë që ekzistojnë njëkohësisht, që ndërveprojnë dhe që konkurrojnë vazhdimisht për të fituar të njëjtat burime të kufizuara. Mqs të gjitha këto tipare përputhen me karakteristikat e sistemeve biologjike, atëherë algjebrat e proceseve mund të konsiderohen si mënyra potenciale modelimi të tyre.

Algjebrat e proceseve

Algjebrat e proceseve janë zhvilluar fillimisht për të kapur semantikën e sistemeve të njëkohshëm dhe janë përdorur për të modeluar rrjetet kompjuterikë ose komunikimet në distancë (Milner, 1989). Algjebrat përcaktohen nga sintaksa dhe semantika e tyre. Semantika përcaktohet nga aksiomat e algjebërës dhe nga rregulla që rrjedhin prej aksiomave apo njëri-tjetrit.

Elementët kryesorë përbërës të algjebrave të proceseve janë proceset e njëkohshëm, të cilët ndërveprojnë nëpërmjet komunikimit dhe shkëmbejnë informacion. Sjellja e sistemit kur ai merr input përcaktohet nga gjendja e tij e momentit dhe nga natyra (përmbajtja) e inputit. Rrjedhojë direkte e ndërveprimit mund të jetë ndryshimi i gjendjes së brendshme dhe i aftësive të procesit për ndërveprim. Pjesët komplekse mund të paraqiten në trajtë hierarkike, si kompozim pjesësh më të thjeshta.

Algjebrat e proceseve mund të përdoren për paraqitjen e sistemeve biologjike duke modeluar molekulat e tyre si procese kompjutacionale të cilat

ndërveprojnë. Ndërveprimi mund të paraqitet duke përdorur primitiva të algjebrave, të përcaktuara posaçërisht për to. Një reaksion paraqitet si ndërveprimi i proceseve që paraqesin pjesëmarrësit në reaksion dhe produktet e tij. Efekti përfundimtar është ndryshimi i gjendjes së proceseve në mënyrë të tillë që të reflektojnë transformimin e reaktantëve në produkte.

Algjebtrat e proceseve kanë natyrë diskrete, që rrjedh nga përcaktimi i tyre. Kjo do të thotë që do të paraqitet numri i molekulave dhe jo përqëndrimi i vazhdueshëm i tyre. Semantika e tyre përfshin edhe elementë për të përshkruar sjellje rastësore. Në këtë kontekst, reaksionet shoqërohen me probabilitet që të ndodhin dhe jo me ritme të ndodhjes së tyre. Algjebtrat janë të pajisura edhe me relacione ekuivalence të cilat janë një mjet shumë i fuqishëm në modelimin e sistemeve biologjike. Psh. ekuivalenca e së njëjtës njësi funksionale në organizma të ndryshëm mund të përdoret si masë e matjes së ngjashmërisë së sjelljeve apo strukturës së organizmave.

Algjebtrat më të rëndësishme të ideuara dhe përdorura për modelimin e sistemeve biologjike janë:

π -calculus biokimik. Kjo algjebër (Priami *et al.*, 2001) rrjedh nga π -calculus (Milner, 1999) i ideuar posaçërisht për modelime biokimike, në të cilat ndërveprimet paraqiten si komunikime input/ output ndërmjet çifte procesesh dhe kështu, zbatohet drejtpërdrejt abstragimi *molekulat si procese*. Pavarësisht se kjo algjebër ka baza të mira teorike, modelet e paraqitura me të janë shpesh shumë abstrakte dhe të vështira për t'u kuptuar. Psh. përdorimi i komunikimeve binare e bën të pamundur paraqitjen në mënyrë direkte të reaksioneve që përfshijnë tre ose më shumë procese.

Algjebra beta-binders (Priami & Quaglia 2005). Kjo algjebër është një tjetër zgjerim i π -calculus dhe që ka në themel përshkrimin e ndërveprimit dhe formimit të molekulave nëpërmjet komunikimit mes pjesëve të përputhshme të proteinave. Proceset që modelojnë proteinat paraqiten si të mbylluara nëpër kuti me kapacitete ndërveprimi (të cilat paraqesin aftësinë e veprimit, lidhjes).

Paraqitja e reaksioneve me struktura që kanë aftësi të lidhen me njëra-tjetrën, në varësi të kapaciteteve vepruese (lidhëse) që kanë dhe përcaktimi i primitivave në algjebër që i bëjnë ndërveprimet të mundura ose jo, është intuitiv dhe përshkruar. Megjithatë, reaksionet mbeten binare, prandaj reaksione të përgjithshme që përfshijnë më shumë sesa dy lloje molekulash nuk mund të përcaktohen në mënyrë eksplicite.

Kjo algjebër është realizuar nëpërmjet gjuhës BlenX dhe për të ndërtuar dhe simuluar e analizuar me metoda kompjutacionale modelet e paraqitura me këtë algjebër është ndërtuar platforma BetaWB (Dematte *et al.*, 2008).

Bio-pepa (Ciocchetta & Hillston, 2009). Abstragimi që përdoret tek Bio-Pepa është *lënda si proces*: një proces paraqet një lloj molekule dhe ndodhja e një reaksioni paraqitet si ndryshim i sasisë së molekulave në dispozicion për llojet e molekulave të përfshira në reaksion. Proceset ndërveprojnë nëpërmjet veprimesh të paracaktuara dhe të emërtuara. Secili veprim luan nga një rol në reaksion. Proceseve u përcaktohen edhe koeficientët përkatëse në reaksion. Efekti i reaksionit është zvogëlimi i numrit të molekulave të lëndëve dhe rrijta e numrit të molekulave të produktit të reaksionit. Bio-Pepa lejon modelimin e reaksioneve të përgjithshme, cilido qoftë numri i proceseve që marrin pjesë në to.

Për nga sintaksa, Bio-Pepa është e ngjashme me modelet e bazuara tek ekuacionet diferenciale. Sasia e llojeve të molekulave mund të shprehet ose si numër molekulash, ose si përqëndrim i tyre. Kjo lejon analizë me metoda të proceseve të rastit dhe me metoda numerike për ekuacionet diferenciale.

Për ndërtimin dhe analizimin e sistemeve të modeluara me Bio-Pepa është ndërtuar platforma Bio-Pepa Eclipse Plug-in.

HYPE (Galpin *et. al*, 2008). Është një algjebër e ideuar për modelimin e sistemeve hibride. HYPE përcakton rrjedhat si ndikuesit në ndryshimin e variablave të vazhdueshëm. Për sistemet biologjike kjo do të thotë identifikimi i proceseve të ndryshme biologjike që çojnë në ndryshime të përqëndrimit të lëndëve. Rrjedha të zakonshme do të ishin prodhimi, degradimi, transformimi dhe transportimi i lëndëve.

HYPE ka semantikën e saj operacionale e cila përcakton sistemin e kalimeve të mundshme nëpër gjendjet e mundshme të sistemit. Gjendjet përcaktohen nga lidhja mes rrjedhave të mundshme që janë aktive dhe vlerave të tyre. Për më tepër, HYPE përcakton edhe ekuacione diferenciale për modelimin e pjesëve të vazhdueshme të sistemeve biologjike.

HYPE përcakton edhe rregulla për shndërrimin e sistemit të kalimeve nëpër gjendjet e mundshme në një automat hibrid. Automatët hibridë janë mjete të fuqishme për modelimin formal të sistemeve hibride (Henzinger, 1996). Për to janë ndërtuar shumë programe të cilat verifikojnë formalisht saktësinë e modeleve të paraqitura prej tyre. Mqs HYPE përcakton rregulla për shndërrimin e modeleve të saj në automate hibride, mund të përdoren programet ekzistuese të verifikimit të saktësisë së modeleve.

Përfundime

Në këtë punim u përshkruan tiparet përcaktuese të disa metodave popullore të modelimit të sistemeve biologjike të cilat mundësojnë analiza kompjutacionale të tyre. Sfida e madhe e tyre është paraqitja e sistemeve me shumë saktësi dhe të gjitha varësitë logjike që ato mbartin. Paraqitja me rigorozitet e sistemeve hap

hapësira të reja, që analiza ose intuita nuk i kap. Kjo rrit cilësinë e studimeve të hipotezave për sistemet biologjike.

Metodat e modelimit shpesh kapin aspekte të ndryshme të sistemeve (ndërtimit dhe sjelljeve të tyre). Prandaj, ato janë të përshtatshme për studimin e anëve të caktuara të sistemit. Metodat e modelimit që lejojnë paraqitje të plotë të natyrës së sistemeve biologjike janë algjebra të proceseve të ideuara kryesisht për sistemet biologjike. Fuqia e tyre shtohet edhe nga platforma të shumta softuerike që lexojnë, interpretojnë dhe analizojnë modelet e ndërtuara mbi këto algjebra.

Literatura

Priami C. & Quaglia P. (2004): Modelling the dynamics of biosystems. In: Briefings in Bioinformatics. VOL 5. NO 3. 259-269. Henry Stewart Publications 1467-5463

Milner R. (1989): Communications and Concurrency. International Series in Computer Science. Prentice Hall

Milner R. (1999): Communicating and mobile systems: the π -calculus. Cambridge University Press

Priami, C., Regev, A., Silverman, W., and Shapiro, E. (2001): Application of a stochastic name-passing calculus to representation and simulation of molecular processes. Information Processing Letters 80, 25–31

Priami C. & Quaglia P. (2005): Beta binders for biological interactions. CMSB' 04 Proceedings of the 20 international conference on Computational Methods in Systems Biology, 20-33. Springer

Priami C. & Quaglia P.(2005): Operational patterns in Beta-Binders. Transactions on Computational Systems Biology, LNBI 3380, 50-65. Springer

Dematte L., Priami C., Romanel A. (2008): Modelling and simulation of biological processes in BlenX. SIGMETRICS Performance Evaluation Review 35(4): 32-39

Ciocchetta, F. & Hillston, J. (2009): Bio-PEPA: A framework for the modelling and analysis of biological systems. Theoretical. Computer. Science. 410, 3065–3084

Galpin V., Hillston J., Bortolussi L. (2008): HYPE Applied to the Modelling of Hybrid Biological Systems. Electronic Notes in Theoretical Computer Science 218 (2008) 33-51. ELSEVIER

Henzinger T.A. (1996): The Theory of hybrid automata. Logic in Computer Science, 1996, 278-292

Fisher J. & Henzinger T.A. (2007): Executable cell biology. Nature Biotechnology, volume 25, no 11

Reisig, W. (1985): Petri Nets: An Introduction. Springer-Verlag, New York

Bio-PEPA Home Page. <http://www.biopepa.org/>