

MASARYKOVA UNIVERZITA
FAKULTA INFORMATIKY



Výpočetní analýza nelineárních dynamických systémů

DIPLOMOVÁ PRÁCE

Sven Dražan

Brno, Podzim 2010

Prohlášení

Prohlašuji, že tato diplomová práce je mým původním autorským dílem, které jsem vypracoval samostatně. Všechny zdroje, prameny a literaturu, které jsem při vypracování používal nebo z nich čerpal, v práci řádně cituji s uvedením úplného odkazu na příslušný zdroj.

Vedoucí práce: RNDr. David Šafránek, Ph.D.

Poděkování

Děkuji svému vedoucímu RNDr. Davidu Šafránkovi, Ph.D. za velmi vstřícný přístup při vedení práce, cenné rady a doporučení týkající se struktury práce i diskusi odborných témat.

Děkuji svým rodičům Miladě a Lubošovi za trpělivost, důvěru a veškerou materiální i lidskou podporu, jež mi během mého studia i psaní této práce poskytovali.

Děkuji své přítelkyni Janě za pomoc při formulování myšlenek, korektury práce, za energii, lásku a neustálý optimismus.

Bez Vás by tato práce nevznikla, díky mnohokrát.

Sven

Shrnutí

Dynamika biochemických reakčních systémů může být za podmínek dostatečně vysokých koncentrací látek popsána soustavou diferenciálních rovnic. Řešení této soustavy lze numericky aproximovat užitím metod numerické simulace. Výběrem diskrétních bodů blízkých trajektoriím původního systému lze vytvořit diskretizace, nad nimiž lze ověřovat vlastnosti vyjádřené formulemi lineární temporální logiky.

Specifickým problémem popisu chování dynamických reakčních systémů je problém iniciální podmínky, tedy nalezení takové množiny bodů, v níž začínají trajektorie splňující danou vlastnost.

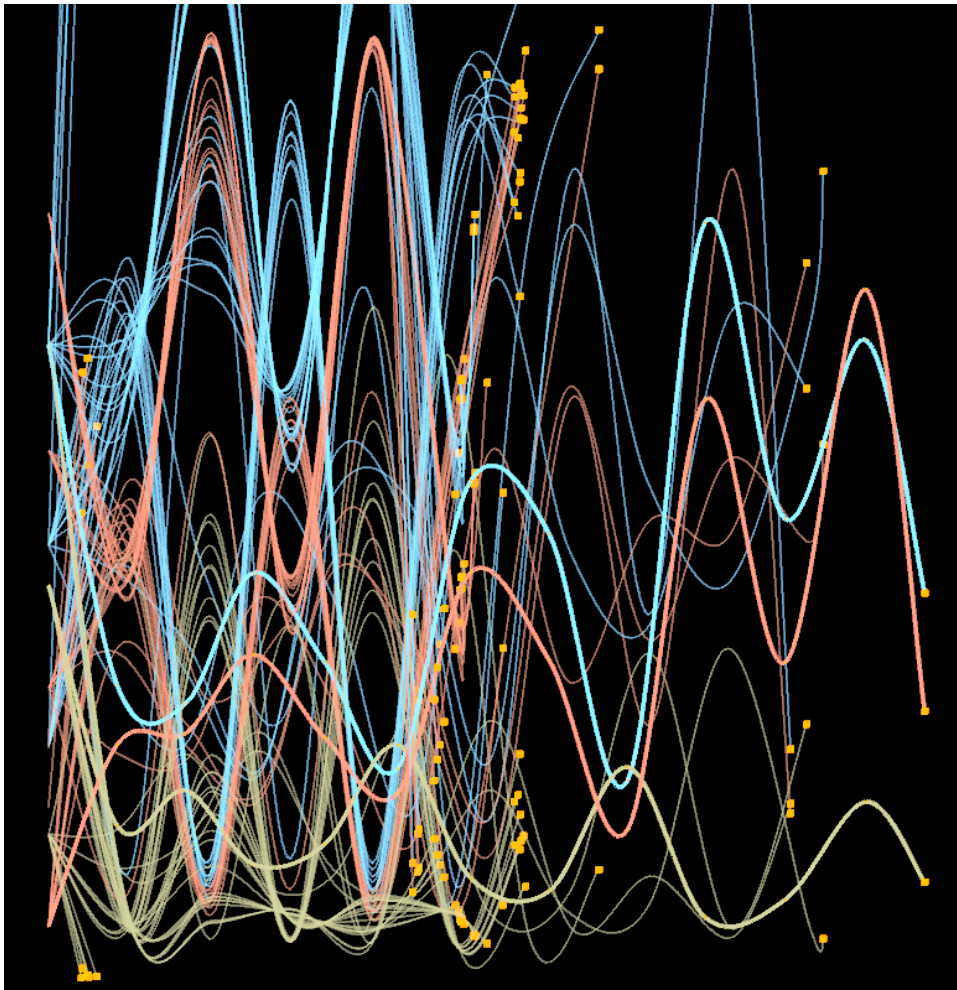
Tato práce představuje přístup jak pomocí dané diskretizace trajektorií řešení systému diferenciálních rovnic a metody počítající relaci vzdálenosti těchto diskretizací lze získat množinu reprezentativních diskretizací. Ověřením požadované vlastnosti nad množinou těchto reprezentantů lze usuzovat o celkovém chování všech trajektorií vycházejících z dané množiny iniciálních podmínek.

Je formálně definována relace vzdálenosti a blízkosti diskretizací trajektorií a abstraktní algoritmus počítající množinu reprezentantů.

Tento algoritmus je implementován a jeho možnosti evaluovány na množině několika malých dynamických systémů s dobře známou dynamikou.

Klíčová slova

Dynamický systém, soustava diferenciálních rovnic, diskretizace, LTL, analýza vlastností.



Obsah

1	Úvod	3
1.1	<i>Dynamické systémy biochemických reakcí</i>	3
1.2	<i>Preludium</i>	4
1.3	<i>Struktura práce</i>	6
2	Pojmy a východiska	7
2.1	<i>Dynamický systém a jeho části</i>	7
2.1.1	Látky	7
2.1.2	Koncentrace	7
2.1.3	Chemická reakce	8
2.1.4	Rychlost chemické reakce	9
2.1.5	Rychlostní rovnice chemické reakce	10
2.1.6	Dynamický systém	12
2.2	<i>Diferenciální rovnice</i>	13
2.3	<i>Dynamika</i>	14
2.4	<i>Vlastnosti dynamických systémů</i>	15
2.4.1	Dosažitelnost	15
2.4.2	Rovnovážné body	16
2.4.3	Oscilace a cykly	17
2.4.4	Bistabilita	17
2.4.5	Popis vlastností - Lineární Temporální Logika	17
2.5	<i>Diskretizace trajektorií numerickou simulací</i>	18
2.6	<i>Existující přístupy</i>	20
3	Nový přístup	23
3.1	<i>Definice problému</i>	23
3.2	<i>Naivní řešení</i>	24
3.3	<i>Blízkost trajektorií jako východisko</i>	25
3.4	<i>Blízkost diskretizací trajektorií</i>	26
3.5	<i>Algoritmus</i>	30
4	Evaluace metody	35
4.1	<i>Linear: jednoduchý model</i>	35
4.2	<i>Lotka-voltera: oscilace bez limitního cyklu</i>	36
4.3	<i>Bayramov: oscilace s limitním cyklem</i>	40

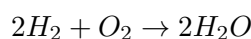
4.4	<i>Wilhelm: minimální bistabilní model</i>	43
5	Návrh a implementace	47
5.1	<i>Nastroje a specifika</i>	47
5.2	<i>Funkcionalita a interaktivita</i>	47
5.3	<i>Popis formátů</i>	48
5.4	<i>Struktura tříd</i>	48
5.5	<i>Algoritmy</i>	50
6	Závěr	51

Kapitola 1

Úvod

1.1 Dynamické systémy biochemických reakcí

Soustavy vzájemně reagujících chemických substancí se vyskytují všude kolem nás. Již na střední škole v hodinách chemie jsme se dozvěděli, že reakcí dvou molekul vodíku $2H_2$ a jedné molekuly kyslíku O_2 vznikají dvě molekuly vody $2H_2O$. Tuto jednoduchou skutečnost lze zapsat výrazem



který možná ještě nezmizel z paměti stejně jako



což je reakce, kterou lze jednak chápat jako anorganický děj hoření dřeva – celulózy – na vzduchu, ale také jako jednu ze základních organických reakcí nutných pro život – dýchání.

Na rozdíl od anorganických reakčních soustav s několika málo reaktanty, reakcemi a produkty se biochemické reakční soustavy odlišují jak velikostí, tedy počtem reagujících součástí a reakcí, tak nutně i svou dynamikou.

Rozdíl mezi kusem hořícího dřeva a člověkem či bakterií je totiž v tom, že ono hořící poleno neodvratně míří do stavu minimální energie. Jakmile se substráty – dřevo nebo kyslík ze vzduchu – vypotřebují a promění v produkty popel, teplo a CO_2 , reakce se zastaví. Naproti tomu člověk nebo bakterie, coby živé soustavy, neustále bojují proti tomuto přirozenému chodu věcí – druhému termodynamickému zákonu a s ním spojenému nárůstu entropie – tím, že svou entropii udržují nebo snižují na úkor okolí, jehož entropii zvyšují, a z něhož neustále čerpají nejrůznější zdroje energie a hmoty.

Život jako hromadná demonstrace proti tomuto zákonu však již během dlouhých stovek miliónů a miliard let pochopil, že příznivé podmínky a nadbytek zdrojů jsou spíše výjimkou, a proto vybavil všechny své soustavy

1. ÚVOD

od nejjednodušších funkčních molekul, přes autonomní jednobuněčné organismy až po složité a rozměrné mnohobuněčné živočichy důmyslnými mechanismy, jak s energií a hmotou hospodařit.

Toto myšlenkové paradigma, tedy, že život je v podstatě velmi důmyslný, protože kdyby nebyl, tak by již dávno přestal existovat, je při analýze biochemických systémů velmi důležité. Pokud analyzujeme soustavu vzájemně reagujících molekul či makromolekul, buněk, tkání nebo orgánů, můžeme se ptát:

- Proč reagují právě tyto molekuly a takto? Jistě je to důležité z pohledu minimální vynaložené energie, pokud by se soustava obešla bez některé ze svých součástí jistě by se tak dávno stalo. Nebo je to z důvodu robustnosti, aby byla funkce zachována i při poškození jedné z částí?
- Proč právě tyto geny a tyto proteiny v tomto reakčním schématu? Umožňuje snad jejich postupné zapínání šetřit zdroje a zahájit dělení buňky jen když je dostatek organel a zásobních látek pro obě dceřiné buňky?
- Jak funguje bičík umožňující bakterii *Escherichia coli* (*E. coli*) vyhledávat oblasti bohaté na zdroje živin, když sama má tak malé rozměry, že není schopna detekovat gradient chemikálií?
- Jak spolu komunikují buňky srdečního svalu skrze změny napětí vyvolané koncentracemi iontů pumpovaných iontovými kanály dovnitř a vně buněčné stěny? Jak jejich vzájemná spolupráce vytváří pravidelné stahy srdečních síní a komor?

Je zřejmé, že nadhozené otázky i systémy jichž se týkají, jsou značně rozdílné svým rozměrem i složitostí a tato práce si neklade za cíl na kteroukoliv z těchto otázek odpovědět. Co si však za cíl klade, je přesvědčit čtenáře, že je nutné v biochemických systémech hledat onu důmyslnost života a pokoušet se ji pochopit jakkoliv to díky vzrůstající složitosti těchto systémů a jejich nelinearitě přestává být jednoduché.

1.2 Preludium

Nasměrování autora této práce do oblasti systémové biologie vzniklo účasti v projektu *Escherichia Coli – Modeling and Analysis* (EC-MOAN [14]).

Cílem tohoto mezinárodního projektu bylo integrovat laboratorní mikrobiologii provádějící *in vitro* experimenty a tvořící modely chování bakterie *E. coli* s matematikou upravujícími tyto modely a navrhuujícími analytické

metody dohromady s informatiky implementujícími tyto metody v algoritmech a využívajícími paralelizace k zrychlení náročných výpočtů.

V centru zájmu bylo integrovat existující modely bakterie *E. coli* na úrovních genetické regulace, metabolismu a mimobuněčných signálů. Hlavními přístupy byly metoda kvalitativní analýzy genetických regulačních sítí s použitím po částech lineárních modelů Hiddeho de Jonga [12] a metoda analýzy dosažitelnosti v multiafinních systémech Calina Belty [7, 20].

Pro první přístup byl upraven existující nástroj DiViNE [5] vyvíjený v Laboratoři paralelních a distribuovaných systémů ParaDiSe [27], byla analyzována složitost tohoto algoritmu a navržen upravený algoritmus počítající podaproximaci s nižší časovou složitostí [4].

Druhý přístup byl implementován do podoby nástroje BioDiVinE [2].

Zahraniční spolupráce v rámci projektu byla nejintenzivnější s prof. Janem H. van Schuppenem [30], jehož výzkumu v oblasti multiafinních systémů, jakožto i celkovému přístupu a nadhledu v mnoha oblastech matematiky si autor této práce velmi váží.

Po vzniku nástroje BioDiVinE a evaluaci jeho možností bylo postupem času zjištěno, že samotná metoda abstrakce multiafinních systémů do podoby přechodového systému, nad nímž se provádí verifikace vlastností metodou ověřování modelu, je až příliš hrubá a nadaproximace, která je takto získána je až příliš velká (množství chování, která jsou navíc oproti původnímu systému přidána je dosti velké). Toto zjištění vedlo ke snaze upravit metodu užitím různých druhů heuristik [3], celkově tak vznikl obtížně uchopitelný objekt podaproximované nadaproximace, jehož vztah k původnímu modelu nebylo možné jednoduše popsat.

Tento vývoj mne postupně vedl k přesvědčení, že zcela striktní, matematicky dokonale definovaná abstrakce poskytující jasné záruky platí za svou rigoróznost až příliš vysokou cenu v podobě značného vzdálení se od reality.

Rozhodl jsem se proto prozkoumat cestu slevující z požadavků zaručeného obsažení všech možných chování daného systému – striktní nadaproximace – a zjistit, zda tím nelze získat metodu produkující výsledky více se s realitou shodující.

Jako základní stavební kámen jsme zvolil numerickou simulaci aproximující trajektorie řešení spojitých systémů obyčejných diferenciálních rovnic (Ordinary Differential Equations - ODE). Tento formalismus je často využíván právě pro popis dynamiky biologických reakčních systémů.

Vlastnosti numerických simulací, jež jsou diskretizací spojitě trajektorie do množiny časoprostorových bodů je možno vyjadřovat formulemi lineární temporální logiky – LTL – které jsou reprezentovatelné Büchiho auto-

1. ÚVOD

maty.

Metoda představená v této práci se zabývá právě množinami numerických simulací daného systému popsaného diferenciálními rovnicemi, které mají splňovat vlastnosti zadané v podobě Büchiho automatů a ze vzájemné polohy jednotlivých simulací se snaží dedukovat vlastnosti celého systému nebo jeho částí.

Metoda ověřování modelu v této práci není použita v klasickém chápání tohoto pojmu [10], neboť běhy dynamického systému jsou zde výhradně lineární řetězce s disjunktními stavy bez prolínání a téměř vždy se v systému vyskytují takové běhy, které definovanou vlastnost porušují. V tomto nastavení tedy není otázkou zda celý systém splňuje danou vlastnost, ale za jakých okolností ji splňuje, resp. jaké musí mít iniciální podmínky, aby ji splnil. Büchiho automaty jsou zde používány spíše jako monitory rozhodující o splnění vlastností nad tím či oním během. Přesná definice problému je uvedena v sekci 3.1.

1.3 Struktura práce

Diplomová práce je členěna do šesti kapitol. Úvodní kapitola se snaží poskytnout motivaci pro hledání nových analytických přístupů dynamických systémů a krátce představuje okolnosti, které k této snaze vedly autora práce.

Druhá kapitola seznamuje čtenáře se základními pojmy oblasti dynamických systémů a nejčastějšími vlastnostmi, které tyto systémy mívají, nebo jsou v nich vyhledávány a s existujícími přístupy analýzy těchto vlastností.

Třetí kapitola je jádrem práce a představuje nový přístup navrhovaný autorem založený na myšlence, že pokud jsou abstraktní objekty při pohybu v prostoru vždy "blízko sebe" mají podobnou dynamiku.

Kapitola čtvrtá evaluuje metodu na konkrétních příkladech několika jednoduchých dynamických systémů vykazujících přítomnost rovnovážných bodů, oscilace a bistability.

Pátá kapitola uvádí podrobnosti návrhu a implementace prototypu algoritmu do podoby nástroje Vizualizátor Verifikace Numerických Simulací – ViVeNuSi, který vznikl ve snaze názorně ukázat průběh výpočtu a umožnit tak jeho lepší pochopení na intuitivní úrovni i zjednodušit jeho další vývoj.

Poslední kapitola shrnuje dosažené výsledky a subjektivně je hodnotí.

Kapitola 2

Pojmy a východiska

Pojmy v této kapitole slouží pro lepší pochopení návazností v dalších částech práce, definice zde obsažené jsou z velké části převzaty z předmětu IV118 – Formální metody v systémové biologii [29] a z Wikipedie [33].

2.1 Dynamický systém a jeho části

2.1.1 Látky

Látkou je v kontextu této práce chápána jakákoliv chemická entita definovaná především svou schopností reagovat s jinými látkami v chemických reakcích, nebo naopak být vůči jiným látkám inertní. Látkami v biochemických či biologických systémech jsou obvykle jak jednoduché molekuly jako H_2O nebo CO_2 tak i složité makromolekuly v podobě nukleotidů, aminokyselin, peptidů až po velké komplexy mnoha makromolekul jako jsou enzymy zajišťující tvorbu a modifikace jiných makromolekul.

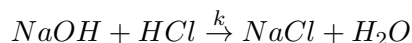
2.1.2 Koncentrace

Koncentrace je dána množstvím dané látky vztažené na jednotku objemu. Koncentrace bývají udávány v $mmol.ml^{-1}$ (milimolech na mililitr) pokud jsou množství látky vzhledem k objemu relativně velká, pokud je naopak počet molekul velmi malý může být udáván i v počtech částic na jednotkový objem buňky. Koncentrace látky A je značena $[A]$.

Příklad pro názornost: Objem krevních buněk se pohybuje v řádech 10^{-14} litrů, počet částic v jednom molu je vyjádřen Avogadrovou konstantou $6,022 \cdot 10^{23} .mol^{-1}$. Pokud by koncentrace např. glukózy v červené krvi s objemem $10\mu m^3$ byla $0,5 mmol.ml^{-1}$, znamenalo by to, že je v ní cca 30 miliónů molekul cukru. Oproti tomuto může být v případě imunitní reakce spouštěcí antigen imunitní reakce v množstvích tisíců či stovek molekul ve stejném objemu.

2.1.3 Chemická reakce

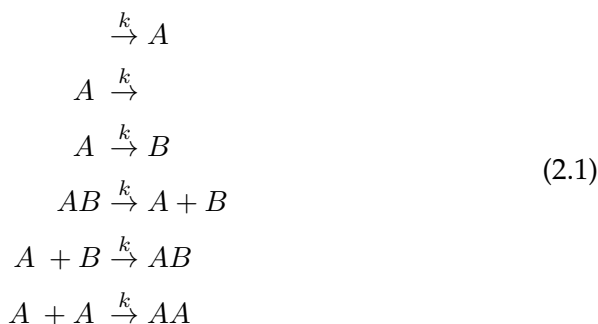
Chemickou reakcí je chápán děj, při kterém dochází k přeměně jedné či více látek v jednu či více jiných látek.



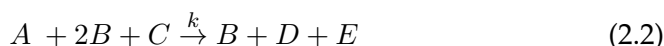
k je specifická kinetická konstanta.

Reakce lze dělit na *elementární* a *obecné*.

Elementární reakce jsou takové, u nichž dochází k přímé jednorázové přeměně reaktantů v produkty, nejsou známy žádné reakční mezistupně nebo je možné je zanedbat z hlediska celkové dynamiky dané reakce. Elementární reakce odpovídají jedinému energetickému maximu, kterým soustava reaktantů prochází při přeměně v produkty. Elementárním reakcím odpovídají reakce s účastí 0, 1 či 2 reaktantů.



Obecné reakce obsahují více reakčních mezistupňů a meziproductů a jejich přesnou kinetiku není možné z podoby dílčích chemických reakcí odvodit.



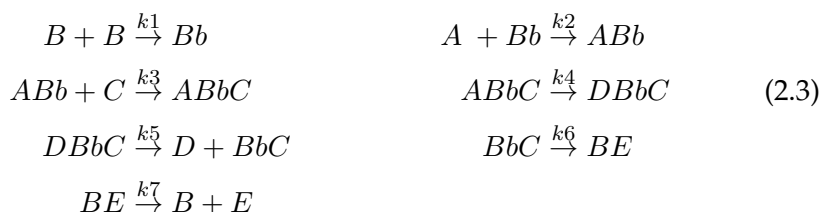
Vlastní doba trvání kontaktu pouhých dvou molekul a jejich přeměny na jinou látku je v realitě považována za velmi krátkou, proto se obecně neuvažuje, že by docházelo k reálné reakci více jak tří molekul naráz (i třímolekulární reakce jsou velmi vzácné a s pokrokem v diagnostických metodách dochází k častému zjištění, že ve skutečnosti jde o rychlý sled dvou elementárních reakcí), jelikož pravděpodobnost jejich setkání ve stejném okamžiku na stejném místě je velmi malá.

Předpokládá se, že skutečně dochází jen k reakci jedné či dvou molekul, které dávají vzniknout komplexu, který nadále funguje jako jeden celek,

nebo zakrátko dojde k jeho opětovnému rozpadu na dvě (postupně možná i více) částí.

Každou obecnou reakci lze tudíž rozložit na posloupnost elementárních reakcí, které probíhají postupně nebo souběžně.

V tomto světle může být reakce (2.2) zjednodušením skutečnosti kdy postupně dojde k reakci nejprve 4 molekul, které vytvoří komplex, jenž projde vnitřní přeměnou, a poté se postupně rozpadne na tři části:



Zjednodušený zápis (2.2) může být rozumnou abstrakcí například pokud jednotlivé dílčí reakce s kinetickými konstantami k_1, k_2, k_3 a k_4, k_5, k_6 probíhají řádově rychleji než reakce s k_4 , během vzniku meziproductů nedochází k jejich dalším reakcím s jinými částmi systému a použitá konstanta k odpovídá celkové dynamice této reakční kaskády s konstantami $k_1 \dots k_7$. Obecně však neexistují pravidla, která by umožňovala z kinetických konstant dílčích reakcí odvodit kinetickou konstantu celé kaskády a v praxi je její hodnota měřena experimentálně.

Také nemusí být jednotlivé dílčí konstanty známy kvůli velké rychlosti dílčích reakcí, jelikož meziproducty rychle reagující mezi sebou může být obtížné izolovat a měřit jejich koncentrace potřebné pro výpočet těchto konstant.

2.1.4 Rychlost chemické reakce

Rychlost r obecné chemické reakce $aA + bB \xrightarrow{k} pP + qQ$ je definována takto:

$$r = -\frac{1}{a} \frac{d[A]}{dt} = -\frac{1}{b} \frac{d[B]}{dt} = \frac{1}{p} \frac{d[P]}{dt} = \frac{1}{q} \frac{d[Q]}{dt} \quad (2.4)$$

za předpokladu uzavřenosti systému (výměny energie nebo hmoty s okolím), stálého objemu (změna objemu vede k změně koncentrací) a bez nárůstu meziproductů reakce.

Mezi rychlostí chemické reakce a rychlostí úbytku reaktantů, resp. nárůstu produktů tedy platí vztah:

$$\frac{d[A]}{dt} = -ra \quad \frac{d[B]}{dt} = -rb \quad \frac{d[P]}{dt} = rp \quad \frac{d[Q]}{dt} = rq \quad (2.5)$$

2.1.5 Rychlostní rovnice chemické reakce

Rychlostní rovnice elementární chemické reakce popisuje závislost mezi rychlostí chemické reakce a koncentracemi všech zúčastněných látek, přičemž platí

$$aA + bB \xrightarrow{k} pP + qQ \quad r = k[A]^\alpha[B]^\beta[P]^\gamma[Q]^\delta \quad k = Ze^{\frac{-E_a}{RT}} \quad (2.6)$$

kde a, b, p, q jsou stechiometrické koeficienty, $\alpha, \beta, \gamma, \delta$ jsou dílčí řády rovnice vzhledem k reaktantům A, B a produktům P, Q . Dílčí řády rovnice se zjišťují experimentálně ze závislosti rychlosti chemické reakce na koncentracích zúčastněných látek. Pro většinu reakcí jsou $\gamma = \delta = 0$ a pro elementární reakce $\alpha, \beta \in \{0, 1, 2\}$ [26].

Řád celé rovnice je součtem dílčích řádů. Pro obecné reakce platí, že dílčí řády ani celkový řád rovnice nemusí být celočíselné hodnoty.

Poslední ze vztahů vyjadřuje empirickou závislost kinetické konstanty k (jenž je ve skutečnosti funkcí) na teplotě T a aktivační energii E_a , Z je konstanta vyjadřující pravděpodobnost účinné srážky a R je molární plynová konstanta, celkově je tento vztah znám jako Arrheniův zákon.

Kinetická konstanta k má různé rozměry pro různé druhy reakcí (pro reakce nultého řádu $\text{mol} \cdot \text{s}^{-1}$, pro reakce prvního řádu s^{-1} , pro reakce druhého řádu $\text{mol}^{-1} \cdot \text{s}^{-1}$, obecně $\text{mol}^{1-n} \cdot \text{s}^{-1}$ pro reakci n -tého řádu).

Reakční kinetické konstanty k jsou tedy spíše empirickými hodnotami zohledňujícími pravděpodobnost setkání reaktantů v prostoru při jejich velikosti, pravděpodobnost reakce vzhledem k průměrné orientaci a kinetické energii molekul za dané teploty. Krom závislosti na teplotě je také netriviální jejich závislost na aktivační energii E_a dané přítomností katalyzátorů, pH prostředí, zdroji energie v podobě elektromagnetického záření či radiace.

Je zřejmé, že vzhledem k množství parametrů je nutné celou řadu z nich fixovat, což je předpoklad, který je mnohdy zcela opodstatněný, avšak jindy k jeho dodržení nemusí docházet:

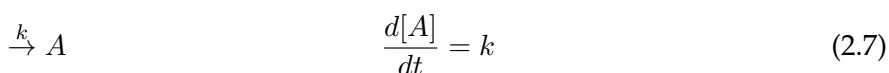
- stálá teplota směsi – reakce mohou být exotermní (uvolňují teplo) do okolí, nebo naopak endotermní (spotřebovávají kinetickou energii okolních molekul a tím celou směs ochlazují),

- stálý tlak směsi – při reakci může docházet k přeměně pevných či kapalných látek v látky, které za dané teploty zaujímají skupenství plynné a zvyšují tak tlak,
- pH prostředí – produkty mohou mít povahu jak neutrální, tak i kyselé či zásaditou,
- rovnoměrnost koncentrace v celém prostoru reakce – pokud jsou reaktanty volně pohyblivé v reakčním prostředí (uvnitř buňky či organely, vně) a teplota (brownův pohyb) či mechanický pohyb zajišťuje neustálé míchání směsi je toto splněno, může však jít o situaci, kdy jsou některé z reaktantů vázány na pevnou strukturu (membrána, vlákno) a jejich volný pohyb není samozřejmý. Také může docházet k postupné změně prostoru, v němž k reakci dochází (růst buňky, difúze do okolí), což opět mění úroveň koncentrací nebo jejich homogenitu.

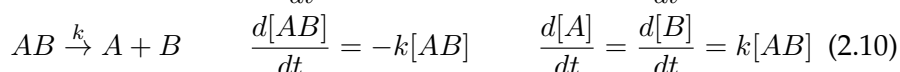
Nejčastěji bývají fixovány všechny parametry krom vlastních koncentrací reagujících látek. V tomto případě se také často předpokládá, že je zajištěn neustálý přísun substrátu a odvod produktů nebo je jinak abstrahováno od faktu, že hromadění produktů reakcí by samovolně snižovalo rychlost průběhu ostatních reakcí, jelikož by při neustále rostoucí koncentraci produktu zakrátko docházelo jen ke srážkám reaktantů s molekulami produktu a nikoliv mezi sebou.

Za těchto předpokladů platí pro *elementární reakce* následující vztahy odvozené z rychlostních rovnic:

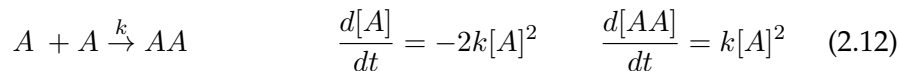
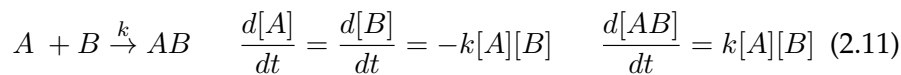
- Reakce nultého řádu (samovolný vznik)



- Reakce prvního řádu (samovolný rozpad nebo přeměna)



- Reakce druhého řádu (reakce dvou molekul)



Pro obecné reakce složené z kaskády mezireakcí není univerzální způsob jak vytvořit rychlostní rovnice a při tvorbě modelu se přihlíží k druhu reakčního mechanismu. Některá opakující se reakční schémata, jako například enzymatické reakce [8] je však možné za určitých předpokladů (rovnováhy probíhajících reakcí a stabilní koncentrace meziproductů, stabilního množství enzymu) zjednodušit a vytvořit tak pro ně rychlostní rovnice s menším počtem proměnných, avšak složitější matematickou podobou.

Obecně se dá říci, že čím více se abstrahuje v reakční kaskádě od dynamiky meziproductů a snahou je vytvořit jediný vztah pro vstupující reaktanty a vystupující produkty, tím složitější vzniká výraz, který má omezenou platnost na doménu pro níž byly vytvořeny abstrakční předpoklady, umožňuje takto ale redukovat množství proměnných v modelu a nutnost znát veškeré vnitřní kinetické konstanty mezireakcí.

2.1.6 Dynamický systém

Pojmem dynamický systém může být označována jakákoliv soustava jedné či více látek dohromady s jednou či více reakcemi, které vedou k časové změně koncentrací alespoň jedné z látek.

Za předpokladu dostatečně velkého množství částic jednotlivých látek soustavy, fixace ostatních parametrů ovlivňující rychlost chemických reakcí (teplota, tlak, pH, přítomnost dalších katalyzátorů) a homogenního rozložení látek v celém objemu soustavy (míchání) lze vyjádřit časovou změnu koncentrací jednotlivých látek pomocí diferenciálních rovnic.

Pokud není splněn předpoklad velkého množství částic, u nichž by bylo možné očekávat deterministické průměrné chování popsatelné empirickými vztahy, může se začít projevovat diskrétní povaha molekul a chování celého systému se stává stochastickým. Metodami stochastické simulace [16] se však tato práce nezabývá, ačkoliv přechod mezi spojitým chováním a stochastickými jevy může být ne zcela zřejmý a existují proto hybridní metody přepínající mezi stochastickou a spojitou simulací v závislosti na množství částic dané látky nebo rychlostmi jednotlivých reakcí [19].

2.2 Diferenciální rovnice

Diferenciální rovnice popisující časový průběh koncentrací látek dynamického systému lze vytvořit ze znalosti všech reakcí a jejich dílčích rychlostních rovnic. Z každé rychlostní rovnice odpovídající jedné reakci se vyjádří změna každého z reaktantů či produktů a tyto dílčí změny se sečtou, čímž vznikne celková změna dané látky v systému.

Formálně mějme dynamickou soustavu l látek $\{A_1, A_2, \dots, A_l\}$ a r *elementárních reakcí*, v nichž má každá látka A_i v reakci s na levé straně reaktantů stechiometrický koeficient n_{si} a na pravé straně produktů stechiometrický koeficient m_{si} (matice levých a pravých stran by byly pro soustavu elementárních reakcí z velké části plné 0 a pro každou reakci by bylo přítomno jen několik 1 či 2). Kinetická konstanta pro reakci s nechť je k_s . Potom časová změna koncentrace látky A_i je dána vztahem [31]:

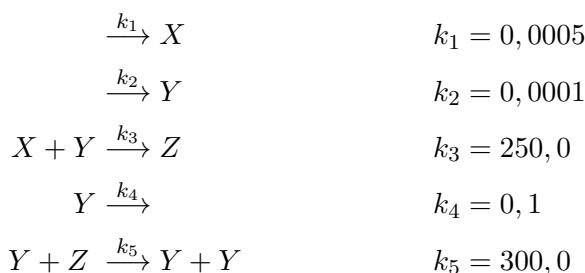
$$\frac{d[A_i]}{dt} = \sum_{s=1}^r \left(k_s (m_{si} - n_{si}) \prod_{j=1}^l [A_j]^{n_j} \right) \quad (2.13)$$

Z podoby diferenciálních rovnic (2.7)–(2.12) vyplývá, že pro každou reakci s platí

$$\left(\sum_{i=1}^l n_{si} \right) \in \{0, 1, 2\}$$

Pokud platí, že tato suma je nejvýše 1, nazývá se systém *multiafinní*. Multiafinní systémy mají tu vlastnost, že při fixaci koncentrací všech proměnných krom jedné je rovnice v této proměnné lineární.

Jako příklad je zvolen model 3 látek a 5 reakcí [6]:



Lze jej také chápat jako model 5 proměnných, kdy X a Y vznikají ze substrátu S a Y se mění v produkt P , přičemž je zajištěna stacionární koncentrace substrátu a stálý odběr produktu, čímž nedochází k jeho hromadění. Poslední z těchto reakcí navíc neodpovídá přímo žádnému z navržených

2. POJMY A VÝCHODISKA

schémat, lze ji však chápat jako zjednodušení dvou reakcí $Y + Z \xrightarrow{k_5} YY$ a $YY \xrightarrow{k_6} Y + Y$. Meziprodukt YY je inertní vůči ostatním látkám a rozpadá se mnohem rychleji než vzniká ($k_6 \gg k_5$).

Pro tuto soustavu platí podle (2.13) tyto rovnice:

$$\begin{aligned}\frac{d[X]}{dt} &= 0,0005 + (-250,0)[X][Y] \\ \frac{d[Y]}{dt} &= 0,0001 + (-0,1)[Y] + (-250,0)[X][Y] + 300,0[Y][Z] \\ \frac{d[Z]}{dt} &= 250,0[X][Y] + (-300,0)[Y][Z]\end{aligned}\quad (2.14)$$

Definice 1. Necht' $n \in \mathbb{N}^+$ je počet látek $X_0, X_1, \dots, X_n$ jejíž dynamika je vyjádřena diferenciálními rovnicemi tvaru 2.13. Tyto rovnice lze chápat jako funkce nad koncentracemi látek, potom **dynamický systém** $\mathcal{DS}$ je trojice s první složkou určující počet látek, druhou určující tyto látky a třetí určující časovou změnu těchto látek, formálně:

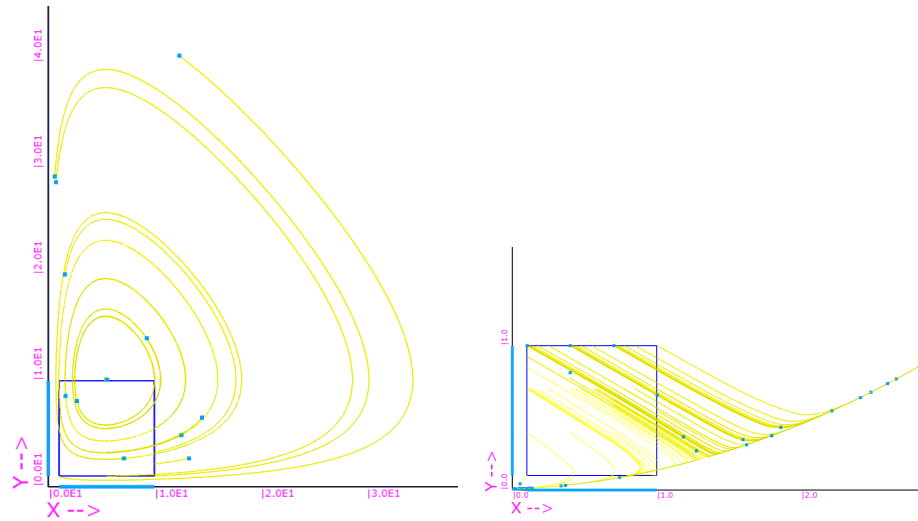
$$\mathcal{DS} = \langle n, \bar{X}, f \rangle \quad n \in \mathbb{N} \quad \bar{X} = (X_1, X_2, \dots, X_n) \quad \frac{d[\bar{X}]}{dt} = f([\bar{X}])$$

2.3 Dynamika

Řešením soustavy diferenciálních rovnic pro daný dynamický systém je funkce $\bar{Y}(t, t_0, [\bar{X}]_0)$ vyjadřující závislost koncentrace všech látek v čase t na počátečních koncentracích látek $[\bar{X}]_0$ v čase t_0 .

Řešení v podobě těchto funkcí jsou pro jednodušší systémy s několika málo látkami a jednoduchými rovnicemi přímo nalezitelné pomocí integrace a symbolických úprav, pro složitější systémy se však spíše používají numerické metody.

Pokud má dynamický systém n látek, lze chápat množinu všech možných hodnot koncentrací těchto látek jako n -dimenzionální prostor $\mathbb{R}^n$ respektive jeho pozitivní hyperkvadrant $\mathbb{R}_0^{+n}$. Pokud je uvažován čas jako další dimenze, půjde o $n + 1$ dimenzionální prostor $\mathbb{R}_0^{+n+1}$. Počáteční stav dynamického systému s koncentracemi $([X_0]_0, [X_1]_0, \dots, [X_n]_0) = [\bar{X}]_0$ je v tomto chápání jediný bod se souřadnicemi $[[X_0]_0, [X_1]_0, \dots, [X_n]_0]$ na osách koncentrací a t_0 na ose časové. Řešením soustavy pro čas $t \in (t_0, t_a)$ je trajektorie začínající v bodě $[\bar{X}]_0$ a končící v $[\bar{X}]_a$, přičemž pokud je původní systém diferenciálních rovnic spojitý, je spojitý i jeho řešení a tato trajektorie tvoří spojitou křivku.



Obrázek 2.1: Ukázka dvou dynamických systémů a jejich trajektorií

Definice 2. *Dynamikou dynamického systému $\mathcal{DS}$ chápeme množinu všech řešení začínajících a pohybujících se v kladném hyperkvadrantu $\mathbb{R}_0^{+n}$.*

2.4 Vlastnosti dynamických systémů

2.4.1 Dosažitelnost

Prostor, v němž se pohybují trajektorie reálných dynamických systémů je omezený na každé ze svých dimenzí. Nejedná se tedy o celý pozitivní hyperkvadrant $\mathbb{R}_0^{+n}$, ale jen o jeho konečnou část.

Důvodem této omezenosti je to, že koncentrace každé látky je zdola i shora omezená $X_i \in [\theta_i^{min}, \theta_i^{max}]$. Toto vychází z biologické podstaty modelovaných systémů, neboť koncentrace každé látky je nezáporná tedy zdola omezená, navíc může být pro fyziologické funkce daného systému nějaká minimální koncentrace přímo nutná. Vzhledem k omezenému objemu modelovaného systému (buňková organela, buňka, orgán, soustava) a fyzikálními vlastnostmi molekul dané látky (objem, stlačitelnost, přitažlivé a odpuzivé síly) je koncentrace omezená také shora, přičemž fyziologické stavy živého systému většinou obsahují koncentrace řádově nižší, než jaké jsou fyzikálně možné. Takto je definován n -dimensionální hyperkvádr v níž se budou pohybovat trajektorie řešení.

2. POJMY A VÝCHODISKA

Definice 3. *Nechť je dán dynamický systém $\mathcal{DS} = \langle n, \overline{X}, f \rangle$ a pro každou látku X_i interval hodnot její koncentrace $[\theta_i^{min}, \theta_i^{max}]$, v nichž se může v systému vyskytovat.*

Globální prostor koncentrací je kartézský součin intervalů hodnot koncentrací všech látek, formálně:

$$\mathcal{G}^n = [\theta_0^{min}, \theta_0^{max}] \times [\theta_1^{min}, \theta_1^{max}] \times \dots \times [\theta_n^{min}, \theta_n^{max}]$$

O dosažitelnosti nějaké oblasti má smysl uvažovat jen pokud se blíže definuje bod nebo množina bodů, odkud má být dosažitelná.

Definice 4. *Nechť jsou dány pro každou látku X_i její počáteční podmínky v podobě intervalu $[\iota_i^{min}, \iota_i^{max}]$. Prostor iniciálních podmínek je kartézský součin počátečních intervalů všech látek, formálně:*

$$\mathcal{I}^n = [\iota_0^{min}, \iota_0^{max}] \times [\iota_1^{min}, \iota_1^{max}] \times \dots \times [\iota_n^{min}, \iota_n^{max}]$$

Vlastnosti dosažitelnosti dané oblasti z daných iniciálních podmínek jsou zajímavé z pohledu otázek, na které je skrze tyto vlastnosti možné se ptát:

- Je daná oblast dosažitelná a pokud ano z jak velké části?
- Jaká je oblast dosažitelná z daných podmínek, je dosažitelná odkudkoliv?
- Je nějaká oblast prostoru nedosažitelná z jiné části?
- Jaká je celková množina bodů z níž je daná oblast dosažitelná? Této množině se říká atrakční množina dané množiny bodů.

2.4.2 Rovnovážné body

Rovnovážné body a jejich stabilita je rozhodující pro dynamiku daného systému. Pokud v něm existuje jeden či více rovnovážných bodů, trajektorie se kolem nich stáčí. Pokud je daný bod labilní, leží na hranici oblastí s různou dynamikou, pokud je stabilní, potom k sobě trajektorie ze svého okolí přitahuje.

Rovnovážné body a jejich existence v daném systému je oblast jíž se extenzivně zabývá lineární analýza.

2.4.3 Oscilace a cykly

Oscilace a cykly umožňují ohraničenost soustav i jejich stabilitu.

V živých systémech je mnoho dějů cyklické povahy: cyklus buněčného dělení, biorytmy a cyklus střídání spánku a bdění, rozmnožovací cykly organismů. Všechny tyto děje jsou povětšinou emergentní vlastnosti daných systémů a nelze tudíž jednoduše nalézt konkrétní látku nebo příčinu přesné frekvence či amplitudy daných oscilací.

2.4.4 Bistabilita

Bistabilita je vlastnost umožňující systému zaujímat vícero stabilních stavů. Umístění více stabilních stavů a tvar jejich atrakčních množin výrazným způsobem ovlivňuje dynamiku daného systému.

V biologii je bistabilita považovaná za důvod diferenciaci buněk při vývoji organismu, tvorbě vzorů a jiným narušením jinak homogenních podmínek.

Poznámka 5. *Všechny výše zmíněné vlastnosti, tedy dosažitelnost, existence rovnovážných bodů, oscilací i bistability lze vyjádřit pomocí formulí lineární temporální logiky [24].*

2.4.5 Popis vlastností - Lineární Temporální Logika

Lineární temporální logika umožňuje formulovat vlastnosti systémů proměnlivé v čase. Přesná definice tohoto pojmu a mimo rámec této práce a lze proto čtenáře odkázat na jejího autora Amira Pnueliho [24].

Základní formulace této logiky byla zamýšlena pro výpočetní systémy s jejich explicitní diskretizovaností. Dostatečně přesná diskretizace ale umožňuje formulovat vlastnosti i jinak spojitých systémů jako jsou například trajektorie systémů diferenciálních rovnic, jak bude nastíněno v následující části práce.

Lineární temporální logika – LTL – umožňuje formulovat vlastnosti lineárních běhu systémů, toto je také hlavní důvod, proč může být použita pro popis chování deterministických systémů, jimiž dynamické systémy popsané soustavou diferenciálních rovnic jsou. V případě, kdy daný systém vykazuje z důvodu nízkých koncentrací některých látek stochastické vlastnosti, mohou být použity některé z logik větvícího se času jako například CTL (Computational Tree Logic).

Interpretace formulí LTL nad posloupnostmi stavů diskretizace numerické aproximace řešení soustavy diferenciálních rovnic použitá v této práci

je zcela převzatá ze článku *On a Continuous Degree of Satisfaction of Temporal Logic Formulae with Applications to Systems Biology* [25], autory jsou Rizk, Batt, Fages a Soliman, strany 3 a 4.

Pro účely implementace byla zvolena reprezentace LTL formulí Büchiho automaty. Příklady použité v kapitole 4 byly vypočítané s pomocí formulí přeložených online nástrojem LTL 2 BA [23] do podoby Büchiho automatů, jež jsou součástí příloh této práce (soubory s příponou `*.ltl`).

2.5 Diskretizace trajektorií numerickou simulací

Trajektorie řešení dynamického systému jsou křivky začínající v bodech počátečních podmínek. Jako takové jsou to matematické objekty, se kterými se dá pro jednodušší systémy manipulovat symbolicky, avšak pro složitější systémy v současnosti taková řešení nejsou dostupná a využívají se numerické integrační metody, označované někdy též jako numerické simulace.

Tyto numerické metody počítají aproximace s konečnou mírou přesnosti jinak nekonečně přesných formálních objektů. U některých numerických metod existuje také formální odhad chyby, které se pro dané nastavení dopouští. Tuto chybu lze chápat jako velikost okolí nebo maximální vzdálenost, v níž se vyskytuje skutečná absolutně přesná křivka od její numerické aproximace. Chyba numerické simulace může být zadána v podobě absolutní vzdálenosti ϵ , relativní odchylky η nebo procentní odchylky δ [32]:

$$\epsilon = |x_{real} - x_{aprox}| \quad \eta = \left| \frac{x_{real} - x_{aprox}}{x_{real}} \right| \quad \delta = \left| \frac{x_{real} - x_{aprox}}{x_{real}} \right| \times 100$$

Numerickou aproximaci křivky lze chápat jako množinu diskrétních bodů v čase a prostoru, které se od skutečné křivky vyskytují do vzdálenosti dané chybou numerické metody v použitém nastavení. Mnoho numerických metod je schopno vytvořit tuto aproximaci s tak velkou přesností, že by explicitní uložení všech těchto bodů bylo paměťově velmi náročné, proto jsou vybírány jen některé body reprezentující celou aproximaci opět s určitou přesností vhodnou pro daný problém. Nejčastějším parametrem výběru těchto reprezentativních bodů je jejich dostatečná vzdálenost, většinou v čase, nebo v prostoru. Tento výběr bodů lze chápat jako diskretizaci původní numerické aproximace matematicky přesného objektu trajektorie.

Definice 6. *Metoda numerické simulace* $\mathcal{M}_\eta \left(f, [\overline{X}]_{t_0}, \Delta t \right) = [\overline{X}]_t$ s parametrem relativní chyby η je funkce, jež má na vstupu soustavu diferenciálních rovnic f , vektor koncentrací $[\overline{X}]_{t_0}$ v čase t_0 a časový úsek délky Δt . Tato funkce počítá numerickou aproximaci trajektorie začínající v bodě $[\overline{X}]_{t_0}$ na intervalu $[t_0, t_0 + \Delta t]$.

Výstupem této funkce je vektor koncentrací látek $\overline{[X]}_t$ v čase $t = t_0 + \Delta t$, případně také vektor časových změn složek jednotlivých látek $\overline{d[X]}_t$ v bodě $\overline{[X]}_t$, přitom platí pro každou látku X_i , že skutečná poloha trajektorie řešení začínající v bodě $\overline{[X]}_{t_0}$ v čase $t_0 + \Delta t$ má od vypočítané polohy relativní odchylku nejvýše η .

Definice 7. Diskretizace numerické aproximace řešení soustavy diferenciálních rovnic $\mathcal{DS} = \langle n, \overline{X}, f \rangle$ provedenou pomocí numerické metody $\mathcal{M}_\eta$ s časovým krokem nejvýše Δt , počátečními podmínkami v podobě bodu $\overline{[X]}_a$, pro časový interval $[t_0, t_0 + \tau]$ délky τ necht' je posloupnost trojic tvaru $\langle t, \overline{[X]}, \overline{d[X]} \rangle$. První složka trojice je čas, druhá složka je vektor koncentrací látek – tedy bod prostoru koncentrací a třetí složkou je vektor časových změn koncentrací látek v bodě.

$$\begin{aligned} \overline{\mathcal{X}}(\mathcal{DS}, \mathcal{M}_\eta, \overline{[X]}_a, t_0, \tau, \Delta t) &= \left(\langle t_0, \overline{[X]}_0, \overline{d[X]}_0 \rangle, \langle t_1, \overline{[X]}_1, \overline{d[X]}_1 \rangle, \dots \right) \\ &= \overline{\mathcal{X}}_a^\tau \end{aligned}$$

Zkrácené označení $\overline{\mathcal{X}}_a^\tau$ necht' je používáno v případech, kdy jsou ostatní parametry jasné z kontextu.

Tato posloupnost necht' je definována induktivně:

$$\begin{aligned} \langle t_0, \overline{[X]}_0, \overline{d[X]}_0 \rangle &= \langle t_0, \overline{[X]}_a, f(\overline{[X]}_a) \rangle \\ \langle t_{n+1}, \overline{[X]}_{n+1}, \overline{d[X]}_{n+1} \rangle &= \langle t_n + \Delta_n, \mathcal{M}_\eta(f, \overline{[X]}_n, \Delta t), f(\overline{[X]}_{n+1}) \rangle \end{aligned} \quad (2.15)$$

Kde Δ_n je skutečná délka časového kroku mezi n -tým a $(n+1)$ -ním prokem posloupnosti. Platí $\Delta_n \leq \Delta t$, jelikož krok mohl být numerickou metodou záměrně zkrácen, aby byla splněna možná další prostorová kritéria vzdálenosti.

Je-li současně definována při diskretizaci také podmínka ohraničenosti globálního prostoru $\mathcal{G}^n$, nemusí diskretizace pokrývat celý interval $[t_0, t_0 + \tau]$, jelikož v jistém okamžiku může být další prvek posloupnosti již mimo $\mathcal{G}^n$. V tomto případě bude mít poslední prvek časovou složku menší než $t_0 + \tau$.

Možnost, že by diskretizace nepokrývala celý specifikovaný časový interval, existuje také je-li vybavena vnitřním mechanismem detekujícím případy, kdy se prvky posloupnosti začínají opakovat. Z důvodů úspory paměti tuto situaci může signalizovat a na výstupu vrátit pouze unikátní část posloupnosti. Tato situace může z pohledu vlastností dynamického systému nastat pokud daný $\mathcal{DS}$ obsahuje ekvilibrium nebo cyklickou trajektorii a tyto jsou dosažitelné z oblastí iniciálních podmínek.

Pro jednoznačnou interpretaci výstupů diskretizace proto zaved' me interní stavovou proměnnou diskretizace $\mathcal{DS}.end_status$, která po vrácení

výstupu obsahuje některou z konečně mnoha konstant popisujících koncové podmínky (ekvilibrium, cyklus, dosažení hranice prostoru, úspěšné pokrytí celého intervalu, vnitřní chybu).

Poznámka 8. *Jelikož přesný výraz **diskretizace numerické aproximace řešení soustavy diferenciálních rovnic** je poněkud zdlouhavý, zaved' me jeho zkrácenou, formálně nepřesnou, avšak intuitivní podobu **diskretizace trajektorie**.*

Jak již bylo naznačeno v komentáři k diskretizaci trajektorie výše, není tento objekt abstraktním matematickým konstruktem, ale reálnou metodou s fyzickými omezeními. Není tedy možné libovolný dynamický systém aproximovat s libovolnou přesností a chybou a tuto aproximaci poté diskretizovat. I velmi pokročilé numerické metody [18] mají svá omezení a jsou použitelné jen pro specifické druhy soustav diferenciálních rovnic.

Hlavním důvodem je způsob výpočtu numerické integrace založený na dělení časového intervalu na mnoho menších intervalů. Pro některé velmi rychle rostoucí funkce by totiž metoda musela zvolit velmi malý krok, což naráží buď na hranice přesnosti numerických operací, nebo na časové možnosti výpočtu.

Díky těmto nepřesnostem mohou mít diskretizace trajektorií odlišné vlastnosti od skutečných abstraktních objektů. Platí například, že pokud existuje v dynamickém systému ekvilibrium, tedy bod, kde jsou časové změny všech složek nulové, je tento bod z kteréhokoli jiného bodu nedosažitelný, neboť by trajektorie do něj vedoucí musela obsáhnout nekonečně dlouhý časový interval (jak se trajektorie blíží ekvilibriu, blíží se nule všechny derivace, tudíž se přibližování zpomaluje). Stejně tak, pokud v systému existuje uzavřená cyklická trajektorie, opět platí, že je ze všech bodů přímo neležících na ní nedosažitelná (jiné trajektorie se k ní mohou libovolně blízko přiblížit, avšak splynout s ní by mohly až v nekonečno, kde však již narážíme na hranice definic cyklického objektu). Při skutečném výpočtu však po jisté době narazíme na numerickou chybu vlastní metody i přesnost reprezentace čísel a tudíž může k dosažení ekvilibríí nebo cyklických trajektorií dojít.

2.6 Existující přístupy

Existující přístupy k analýze dynamických systémů lze rozdělit podle oblasti, do které kladou největší nejistotu:

- nízké koncentrace látek – V tom případě jsou aplikovány metody stochastické simulace [16] s různým stupněm pokročilosti, optimalizace

či paralelizovatelnosti. Výstupem jsou množiny trajektorií vycházejících z daných iniciálních podmínek, které lze hodnotit užitím statistických metod [9]. Výstupem může být pravděpodobnost s níž daný systém resp. množina jeho trajektorií splňuje danou vlastnost. Pro nedeterministické běhy bývají občas místo logiky LTL [24] používány logiky větvícího se času - CTL a podobné.

- nejisté hodnoty kinetických konstant nebo dalších numerických konstant v pravých stranách rovnic. V tomto případě jde o problémy tzv. *parameter estimation*, tedy odhadů parametrů, nebo tzv. *parameter fitting, curve fitting*, tedy "pasování" dané množině chování. Je známa množina chování skutečného systému a hledají se jeho numerické parametry tak, aby model chování co nejvíce odpovídal. Způsoby řešení tohoto problému jsou různé, vesměs však jde o různé inteligentní prohledávání prostoru všech možných parametrů. Čím více parametrů je neznámých, a čím větší jsou intervaly jejich nejistoty, tím větší je prostor, který je nutné prohledávat. Pokud se podaří zadefinovat metriku v prostoru parametrů vyjadřující se o tom, jak moc daný systém odpovídá specifikaci, lze využít některou z heuristik prohledávání jako je např. simulované žíhání. Pokud je specifikace chování dána formulí je možné definovat i spojitou metriku splnění formule nad během systému – diskretizací trajektorie [25].
- nejistota v počátečních podmínkách jinak přesně daného dynamického systému – o řešení tohoto problému se snaží tato práce. Současné existující postupy (pokud je mi známo) nejdou dále než masivní navzorkování počátečních podmínek a paralelní vyhodnocení.
- Je důležité také zmínit analytické přístupy z oblasti lineární algebry, které pro jednodušší systémy, ale i leckteré složitější dokáží přesně lokalizovat významné body určující dynamiku systému jako jsou např. ekvilibria nebo hranice, kde se mění znaménka derivací.

Kapitola 3

Nový přístup

3.1 Definice problému

Nechť je dán dynamický systém $\mathcal{DS} = \langle n, \overline{X}, f \rangle$ n látek, jejichž změna v čase je vyjádřena n diferenciálními rovnicemi $\frac{d[X_i]}{dt} = f_i(\overline{X})$ se spojitými multiafinními pravými stranami. Nechť jsou známy přesné numerické hodnoty všech kinetických konstant vyskytujících se v těchto rovnicích. Soubor těchto diferenciálních rovnic je souhrnně označen $\frac{d[\overline{X}]}{dt} = f(\overline{X})$.

Nechť je pro každou látku X_i dán interval hodnot její koncentrace

$$[\theta_i^{min}, \theta_i^{max}]$$

v nichž se může v systému vyskytovat. Tyto intervaly určují globální prostor koncentrací

$$\mathcal{G}^n = [\theta_0^{min}, \theta_0^{max}] \times [\theta_1^{min}, \theta_1^{max}] \times \dots \times [\theta_n^{min}, \theta_n^{max}]$$

Nechť jsou dány pro každou látku X_i její počáteční podmínky v podobě intervalu $[\iota_i^{min}, \iota_i^{max}]$. Tyto určují prostor iničiálních podmínek

$$\mathcal{I}^n = [\iota_0^{min}, \iota_0^{max}] \times [\iota_1^{min}, \iota_1^{max}] \times \dots \times [\iota_n^{min}, \iota_n^{max}]$$

Nechť je dána metoda numerické simulace $\mathcal{M}_\eta \left(f, \overline{X}_{t_0}, \Delta t \right) = \overline{X}_{t_0 + \Delta t}$ s relativní chybou η , která má na vstupu soustavu diferenciálních rovnic f , vektor koncentrací $\overline{X}_{t_0}$ v čase t_0 a časový úsek délky Δt . Tato metoda vypočte nový vektor koncentrací látek $\overline{X}_{t_0 + \Delta t}$ v čase $t_0 + \Delta t$ numerickou integrační metodou s relativní chybou η .

Je dána metoda diskretizace řešení soustavy diferenciálních rovnic pomocí numerické metody $\mathcal{M}_\eta$ s časovým krokem nejvýše Δt , pro počáteční podmínky v podobě bodu $\overline{X}_a$, časový interval $[t_0, t_0 + \tau]$ délky τ

$$\overline{\mathcal{X}}_a^\tau = (\langle t_0, \overline{X}_0, \overline{d[X]}_0 \rangle, \langle t_1, \overline{X}_1, \overline{d[X]}_1 \rangle, \dots)$$

3. NOVÝ PŘÍSTUP

Nechť je dána vlastnost P vyjádřitelná formulí LTL a reprezentovaná Büchiho automatem B .

Definice 9. *Problém iniciálních podmínek* je zjistit v kterých částech prostoru iniciálních podmínek $\mathcal{I}^n$ začínají trajektorie řešení systému, jejichž diskretizace splní (nebo naopak poruší) podmínku P .

$$\begin{array}{lll} \mathcal{A} \subseteq \mathcal{I}^n & \forall [\overline{X}]_a \in \mathcal{A}. \overline{\mathcal{X}}_a^T \models P & \mathcal{A} = ? \\ \mathcal{B} \subseteq \mathcal{I}^n & \forall [\overline{X}]_b \in \mathcal{B}. \overline{\mathcal{X}}_b^T \not\models P & \mathcal{B} = ? \end{array}$$

Definice 10. Pokud množina iniciálních podmínek $\mathcal{I}^n$ pokrývá celý prostor globálních koncentrací $\mathcal{G}^n$ je problém iniciálních podmínek nazýván **problémem globální analýzy**.

3.2 Naivní řešení

Naivním řešením problému iniciálních podmínek by mohlo být diskretizovat prostor iniciálních podmínek do husté sítě mřížových bodů, podél každé dimenze s takovou hustotou, s jakou je požadován výsledek. Z těchto mřížových bodů spustit nezávislé numerické simulace diskretizující trajektorie a u těchto diskretizací ověřit platnost vlastnosti P (buď po vypočtení celé simulace, nebo v průběhu metodami *on-the-fly*).

Výhodami tohoto naivního řešení by byla jednoduchá implementace, snadná distribuce i paralelizovatelnost výpočtů díky jejich nezávislosti a jistá záruka nalezení hledaných trajektorií odvíjející se od hustoty iniciálních mřížových bodů.

Nevýhodou tohoto přístupu by byla velká náročnost na čas i paměť, neboť oba dva parametry jsou lineárně závislé na počtu mřížových bodů, jejichž počet by při rozdělení iniciálního intervalu dimenze i na d_i částí byl $\prod_{i=1}^n (d_i + 1) \sim d^n$ pro $d_0 = d_1 = \dots = d_n$. Pro reálné dynamické systémy s dimenzionalitou n často se pohybující v počtech látek spíše v řádech desítek než jednotek by při možném rozpětí koncentrací určité látky v rozmezí několika řádů byl tento přístup i při logaritmickém dělení iniciálních intervalů velmi náročný.

Nabízí se otázka, zda je možné toto naivní řešení upravit tak, aby zůstaly jeho podstatné výhody a přitom bylo možné snížit alespoň parametricky jeho výpočetní náročnost.

3.3 Blízkost trajektorií jako východisko

Východiskem navrhovaného přístupu je na intuitivní rovině empirické zjištění

Hypotéza 11. *Většina řešení začínajících v iniciálních bodech **blízko** sebe zůstávají **blízko** i v průběhu času.*

Formálněji toto pozorování lze formulovat jako větu o spojitě závislosti řešení systémů diferenciálních rovnic se spojitými pravými stranami na počátečních podmínkách a času [17]. Tato věta má jako svůj předpoklad unikátnost řešení. Unikátnost řešení znamená, že každým bodem prostoru prochází právě jedno řešení. Toto je pro dynamické systémy, tak jak jsou uvažovány v této práci splněno, jelikož jsou to systémy autonomní (derivate jsou nezávislé na čase).

Jelikož platnost vlastnosti P na trajektoriích procházejících prostorem lze chápat také jako spojitou závislost [25] na tomto prostoru, lze úvodní hypotézu reformulovat do podoby:

Hypotéza 12. *Většina diskretizací trajektorií začínajících v iniciálních bodech **blízko** sebe mají **podobnou** míru platnosti vlastností závislých na prostoru, v němž se pohybují.*

Toto východisko umožňuje formulovat empirickou hypotézu:

Hypotéza 13. *Pro to, aby bylo zjištěno chování diskretizací trajektorií v určité oblasti prostoru s danou přesností dělení, není třeba počítat diskretizace trajektorií všech bodů v tomto prostoru s danou přesností dělení. Stačí vyzkoušet jen množinu **reprezentantů** a to za předpokladů:*

- *diskretizace trajektorií v jejich okolí jim zůstanou **blízko** na celém časovém intervalu τ , na němž je chování systému – vyjádřené vlastností P – ověřováno,*
- *pokrývají celou množinu iniciálních podmínek.*

Podmínku **blízkosti** diskretizací trajektorií během průběhu prostorem (resp. časoprostorem) jejich reprezentantovi lze dále nahradit podmínkou vzájemné **blízkosti** reprezentantů a to na základě stejné úvahy o spojitě závislosti řešení systému na počátečních podmínkách.

Pokud se tedy podaří nalézt množinu reprezentantů diskretizací trajektorií, kteří zůstávají po celou dobu τ ve vzájemné **blízkosti**, lze očekávat, že jim zůstanou blízko i diskretizace trajektorií v okolí mezi těmito reprezentanty. Na základě úvahy o spojitě závislosti míry platnosti P budou poté

3. NOVÝ PŘÍSTUP

zjištěné hodnoty platnosti vlastnosti P na těchto reprezentantech odpovídat hodnotám platnosti v jejich okolí.

Problém iniciálních podmínek lze při tomto pohledu rozložit na dva podproblémy:

- Nalezení reprezentantů, tedy množiny diskretizací trajektorií, které začínají v iniciálních podmínkách a po celý průběh sledovaného intervalu $[t_0, t_0 + \tau]$ si budou vzájemně **blízko**.
- Určení pravdivostních hodnot vlastnosti P na množině reprezentantů.

Řešením problému by v této konstrukci byla množina počátečních bodů reprezentantů splňujících (resp. porušujících) vlastnost P a okolí těchto bodů.

Ústředním pojmem, s nímž tyto úvahy operují je pojem **blízkosti** diskretizací trajektorií, následuje proto návrh jeho možné definice.

3.4 Blízkost diskretizací trajektorií

Ještě před vlastní definicí označme $\mathcal{A} \leftrightarrow \mathcal{B}$ relaci blízkosti dvou diskretizací trajektorií a $\mathcal{A} \nleftrightarrow \mathcal{B}$ relaci vzdálenosti, přičemž $\mathcal{A} \leftrightarrow \mathcal{B} \Leftrightarrow \neg \mathcal{A} \nleftrightarrow \mathcal{B}$.

Důvodem proč je zavedena relace a nikoliv rovnou vzdálenost je ten, že je třeba mít diskrétní vztah definující, kdy jsou si diskretizace trajektorií blízké a kdy vzdálené, přičemž zavedení metriky mezi diskretizacemi a prosté porovnání dvou čísel není nad takto složitými objekty jednoduché. Po postupném zadefinování pomocných pojmů je o to učiněn pokus v definici 18.

Pro definici blízkosti je důležitá zvolená metrika. Při její volbě byla uvažována následující kritéria:

- Nezávislost na škálách jednotlivých dimenzí – ve skutečných modelech dynamických systémů mohou mít různé látky koncentrace různých se o mnoho řádů.
- Invariantnost vůči modularitě – biologické dynamické systémy jsou považovány za částečně modulární. To znamená, že je možné je komponovat a jejich dynamika zůstává do určité míry zachována. Proto je vhodné, aby byla relace blízkosti zachována i při kompozici dvou dynamických systémů.

Pro příklad zvolme dva dynamické systémy s disjunktními množinami látek, které jsou vzájemně intertní. Diskretizace trajektorií v těchto

systémech jsou posloupnosti obsahující body v čase a prostoru, přičemž kompozicí obou systémů by vznikl nový systém, jehož dynamika na každé z proměnných by odpovídala původnímu systému (neuvažujeme nyní zkracování posloupností v důsledku detekce ekvilibrií či cyklů). Diskretizace trajektorií tohoto nového systému by byly opět posloupnosti obsahující body, přičemž jednotlivé složky ve vektorech koncentrací a jejich změn by odpovídaly těm v původních diskretizacích (neuvažujeme nyní proměnlivost časových kroků Δ_n).

Bylo by nyní *žádoucí*, aby platilo, že pokud byly dvě dvojice trajektorií ($\mathcal{A}, \mathcal{B}$ v systému 1 a $\mathcal{C}, \mathcal{D}$ v systému 2) v daném vztahu před kompozicí, měl by zůstat zachován i po kompozici:

$$\mathcal{A} \leftrightarrow \mathcal{B} \Rightarrow \mathcal{A}|\mathcal{C} \leftrightarrow \mathcal{B}|\mathcal{D} \quad \mathcal{A} \leftrightarrow \mathcal{B} \wedge \mathcal{C} \leftrightarrow \mathcal{D} \Rightarrow \mathcal{A}|\mathcal{C} \leftrightarrow \mathcal{B}|\mathcal{D}$$

Je dobré uvést, že tento postup není z modelovacího hlediska správný, jelikož skládat modely, které spolu nesouvisí dohromady není smysluplné. Z hlediska logického je však tento požadavek *vhodný*, neboť je to jistá limita kompozice dvou systémů při níž lze zachovávání relace blízkosti testovat.

- Nezávislost na čase – toto kritérium je diskutabilní a vychází z praktické podoby sledovaných vlastností dynamických systémů. Pokud tyto vlastnosti závisí čistě na prostorových souřadnicích (dosažitelnost, existence rovnovážných bodů, cyklů) tedy nezahrnují časovou složku, zjednoduší se výpočet blízkosti a to nikoliv na elementární úrovni (porovnání dvou čísel), ale na úrovni vyšší rozhodující o tom, zda budou dvě trajektorie uznány za reprezentanty, nebo zda bude třeba dále hledat. Pokud však jsou sledované vlastnosti přímo závislé na časové složce (porovnávání časového rozměru s konkrétní hodnotou) nebo se v ní projevují (porovnávání hodnot derivací), je možné toto kritérium relaxovat a závislost na čase v relaci blízkosti také uvažovat.

Na základě těchto kritérií byla zvolena upravená metrika součinu vážených absolutních rozdílů jednotlivých prostorových složek.

Pro dva body $\overline{[A]}, \overline{[B]}$ je jejich vzdálenost definována takto:

$$|\overline{[A]}, \overline{[B]}| = \prod_{i=1}^n \max\left(1, \frac{|A_i - B_i|}{\Delta_i^{\max}}\right) \quad (3.1)$$

n je počet látek, tedy dimenze prostoru, $\Delta_i^{\max}$ je parametr určující váhu dimenze i a lze jej interpretovat jako maximální povolenou vzdálenost dvou

3. NOVÝ PŘÍSTUP

bodů na dimenzi i , pro kterou budou ještě považovány za **blízké**. Tento parametr je součástí vstupu pro daný model dynamického systému v podobě n -tice $\Delta^{max} = (\Delta_0^{max}, \Delta_1^{max}, \dots, \Delta_n^{max})$ a je odvozený buď uživatelem ze znalosti rozsahu koncentrací nebo je možné jej automaticky odhadnout z velikosti iniciálního intervalu $[\iota_0^{min}, \iota_0^{max}]$ a počtu částí, na než byl rozdělen při prvotním nastavení reprezentantů (viz následující sekce popisující algoritmus).

S touto metrikou je možné definovat relaci blízkosti dvou bodů takto

Definice 14. *Nechť $\overline{[A]}$ a $\overline{[B]}$ jsou dva body n -dimenzionálního prostoru, nechť platí*

$$\overline{[A]} \leftrightarrow \overline{[B]} \Leftrightarrow |\overline{[A][B]}| = 1 \quad \overline{[A]} \leftrightarrow \overline{[B]} \Leftrightarrow |\overline{[A][B]}| > 1$$

Tato metrika má všechny výše uvedené požadované vlastnosti a má tu vlastnost, že pokud jsou body $\overline{[A]}$ a $\overline{[B]}$ na každé dimenzi i blíže než Δ_i^{max} je jejich vzdálenost 1, pokud v kterékoliv dimenzi jejich vzdálenost překročí Δ_i^{max} je jejich celková vzdálenost větší než 1.

Důkaz tohoto tvrzení:

Nezávislost na škálách jednotlivých dimenzí je dána samostatným vážením rozdílů vzdáleností na každé prostorové složce, invariantnost vůči kompozici vychází z monotónnosti operátoru součinu pro čísla větší nebo rovná jedné a vztahu

$$|\overline{[A]}|\overline{[C][B]}|\overline{[D]}| = |\overline{[A][B]}| \cdot |\overline{[C][D]}|$$

nezávislost na čase je zřejmá z důvodu nezahrnutí časové složky do váženého součinu.

Blízkost dvou bodů nyní umožňuje zadefinovat pomocí indukce relaci blízkosti bodu $\overline{[A]}_i$ a diskretizované trajektorie $\mathcal{B}$, nejprve však zavedme pojmy *podposloupnost blízká bodu* a *množina podposloupností blízkých bodu*.

Definice 15. *Podposloupnost diskretizace $\mathcal{B}$ blízká bodu $\overline{[A]}_0$ od indexu j nechť je nepřerušená podposloupnost všech bodů diskretizace $\mathcal{B}$ od j -tého počínaje takových, že jsou **blízké** bodu $\overline{[A]}_0$, formálně:*

$$P(\overline{[A]}_0, \mathcal{B}, j) = \bigcup_{k \in \mathbb{N}_0} \left\{ \overline{[B]}_i \mid \forall i. j \leq i \leq j + k : \overline{[A]}_0 \leftrightarrow \overline{[B]}_i \right\}$$

V této definici není podposloupnost blízká bodu skutečnou posloupností v klasickém matematickém chápání tohoto pojmu jako zobrazení, jedná

se o množinu. Pojmenování podposloupnost je použito pro zdůraznění toho, že všechny body množiny jsou v původní diskretizaci $\mathcal{B}$ za sebou a jsou její podčásti. Tato definice by mohla být rozšířena do podoby zobrazení přiřazujícího číslu i prvek $\overline{B}_i$ tak, aby v následujících definicích zůstávala smysluplná, v tom případě by však docházelo k změnám definičních oborů těchto zobrazení.

Definice 16. *Množina podposloupností diskretizace $\mathcal{B}$ blízkých bodu $\overline{A}_i$ označená $\mathbf{M}(\overline{A}_i, \mathcal{B})$ je definována induktivně:*

$$\begin{aligned}\mathbf{M}(\overline{A}_0, \mathcal{B}) &= \{P(\overline{A}_0, \mathcal{B}, 0)\} \\ \mathbf{M}(\overline{A}_{k+1}, \mathcal{B}) &= \bigcup_{\overline{B}_j \in \mathbf{M}(\overline{A}_k, \mathcal{B})} P(\overline{A}_{k+1}, \mathcal{B}, j)\end{aligned}$$

Množina podposloupností blízkých bodu je ve své podstatě opět množinou bodů diskretizace $\mathcal{B}$ ovšem s tou strukturou, že se v ní vyskytují sekvence po sobě jdoucích bodů s mezerami jako například

$$\{\overline{B}_a, \overline{B}_{a+1}, \dots, \overline{B}_{a+m}, \overline{B}_b, \overline{B}_{b+1}, \dots, \overline{B}_{b+n}, \dots\}$$

Definice 17. *Relace blízkosti bodu $\overline{A}_i$ a diskretizované trajektorie $\mathcal{B}$ je definována pomocí množiny podposloupností diskretizace $\mathcal{B}$ blízkých bodu $\overline{A}_i$:*

$$\overline{A}_i \leftrightarrow \mathcal{B} \Leftrightarrow \mathbf{M}(\overline{A}_i, \mathcal{B}) \neq \emptyset$$

Definice 18. *Relace blízkosti dvou diskretizací $\mathcal{A}$ a $\mathcal{B}$ je definována pomocí blízkosti všech bodů jedné z nich vůči druhé:*

$$\mathcal{A} \leftrightarrow \mathcal{B} \Leftrightarrow \left(\forall \overline{A}_i \in \mathcal{A}. \overline{A}_i \leftrightarrow \mathcal{B} \right) \wedge \left(\forall \overline{B}_i \in \mathcal{B}. \overline{B}_i \leftrightarrow \mathcal{A} \right)$$

Zcela formálně vzato je každá diskretizace tvořena posloupností trojic $\langle t_i, \overline{X}_i, d[\overline{X}_i] \rangle$, z nichž bodem je pouze druhá složka, výraz $\forall \overline{A}_i \in \mathcal{A}$ je tedy chápán jako množina všech bodů, jež jsou druhé složky trojic v diskretizaci.

V následující části popisující algoritmus bude využito skutečnosti, že ve výše uvedené definici blízkosti dvou diskretizací jsou obě požadované podmínky ve většině případů ekvivalentní a pokud se ošetří jisté specifické hraniční případy na jejich koncích, lze testovat jen jednu z nich.

Tato poněkud složitá definice blízkosti dvou diskretizací trajektorií oproti možné jednodušší definici:

$$\mathcal{A} \leftrightarrow \mathcal{B} \Leftrightarrow \left(\forall \overline{A}_i \in \mathcal{A}, \forall \overline{B}_j \in \mathcal{B}. \overline{A}_i \leftrightarrow \overline{B}_j \right)$$

má tu výhodu, že tvoří jistý kompromis v posledním požadavku na metriku, tedy v nezávislosti na čase. V této složitější podobě totiž neobsahuje přímou závislost na čase, umožňuje však formulovat vlastnosti obsahující informace o čase tím, že neumožňuje množiny posloupností blízkých bodům rozšiřovat směrem zpět v čase. Implementace s použitím této složité podoby je také efektivnější.

3.5 Algoritmus

Algoritmus zde popsaný se soustředí především na řešení prvního podproblému, tedy nalezení množiny reprezentativních diskretizací trajektorií, které budou po celou dobu τ na niž je vlastnost P ověřována **blízko** sebe. Druhý podproblém, tedy výpočet pravdivosti P nad reprezentanty je již přímočarý a nikterak inovativní.

Vstupem algoritmu nechť je dynamický systém $\mathcal{DS} = \langle n, \overline{X}, f \rangle$, definice globálního prostoru koncentrací $\mathcal{G}^n$ a prostoru iniciálních podmínek $\mathcal{I}^n$, počáteční dělení iniciálních podmínek $\mathcal{D}^n$ určující na kolik částí mají být rozděleny iniciální intervaly v každé dimenzi, časový úsek τ , metoda numerické simulace $\mathcal{M}_\eta$ a maximální časový krok Δt , ověřovaná vlastnost P a množina váhových parametrů Δ^{max} .

Výstupem bude množina bodů iniciálních podmínek splňujících vlastnost P .

Pro následující popis algoritmu budiž základní datovou a výpočetní jednotkou reprezentant F_a sestávající ze čtveřice

$$F_a = \langle \overline{[X]}_a, \overline{\mathcal{X}}_a^\tau, N, S \rangle$$

$\overline{[X]}_a$ je počáteční bod, $\overline{\mathcal{X}}_a^\tau$ je diskretizace trajektorie vycházející z počátečního bodu, N je množina indexů sousedních reprezentantů, S je koncový status diskretizace (jedna z hodnot bound, equilibrium, cycle, timeout, null).

V algoritmu se vyskytuje několik výrazů a funkcí, jež je vhodné definovat:

- *Division* ($\mathcal{I}^n, \mathcal{D}^n$) (řádek 2)
funkce generující dvojici $(\overline{[X]}_\alpha, N_\alpha)$. První složka je mřížový bod, druhá je množina indexů okolních mřížových bodů. Sít' iniciálních mřížových bodů vznikne rozdělením prostoru iniciálních podmínek v každé dimenzi na $\mathcal{D}_i^n$ částí, tudíž celkem takto vznikne $\prod_{i=1}^n (\mathcal{D}_i^n + 1)$ mřížových bodů.
Způsob jakým jsou tvořeny množiny indexů okolních mřížových bodů výrazným způsobem ovlivní další průběh výpočtu i citlivost detekce

Algoritmus 1 Obecná struktura

```

1: init  $R_{old} := \emptyset$ 
2: init  $R_{new} := \left\{ \langle [\overline{X}]_\alpha, \emptyset, N_\alpha, \text{null} \rangle \mid ([\overline{X}]_\alpha, N_\alpha) \in \text{Division}(\mathcal{I}^n, \mathcal{D}^n) \right\}$ 
3: while  $R_{new} \neq \emptyset$  do
4:   for all  $F \equiv \langle [\overline{X}], \overline{\mathcal{X}}, N, S \rangle \in R_{new}$  do
5:      $F.\overline{\mathcal{X}} := \mathcal{X}\left(\mathcal{DS}, \mathcal{M}_\eta\left(f, [\overline{X}], \Delta t\right), [\overline{X}], t_0, \tau, \Delta t\right)$ 
6:      $F.S := \mathcal{M}_\eta.\text{end\_status}$ 
7:   end for
8:    $R_{tmp} = R_{new}$ 
9:    $R_{new} = \emptyset$ 
10:  for all  $F_a \in R_{tmp}$  do
11:    for all  $F_b \in \{F_i \mid i \in F_a.N\}$  do
12:       $\text{proximity} := \text{check\_distance}(F_a.\overline{\mathcal{X}}, F_a.S, F_b.\overline{\mathcal{X}}, F_b.S)$ 
13:      if  $\text{proximity} == \text{false}$  then
14:         $R_{new} := R_{new} \cup \left\langle \frac{F_a.\overline{[X]} + F_b.\overline{[X]}}{2}, \emptyset, \{a, b\}, \text{null} \right\rangle$ 
15:      end if
16:    end for
17:  end for
18:   $R_{old} := R_{old} \cup R_{tmp}$ 
19: end while
20: init  $\text{result} := \emptyset$ 
21: for all  $F \in R_{old}$  do
22:    $\text{satisfied} := \text{check\_property}(F.\overline{\mathcal{X}}, P)$ 
23:   if  $\text{satisfied} == \text{true}$  then
24:      $\text{result} := \text{result} \cup \text{neighbourhood}\left(F.\overline{[X]}_0^\alpha\right)$ 
25:   end if
26: end for
27: return  $\text{result}$ 

```

3. NOVÝ PŘÍSTUP

blízkosti trajektorií. V současné implementaci je každému mřížovému bodu přiřazeno 0 až n sousedů se souřadnicemi vždy o jedna nižší v každé dimenzi (body na hranicích $\mathcal{I}^n$ mají méně sousedů, většina bodů uvnitř má n sousedů).

- $\mathcal{M}_\epsilon.end_status$ (řádek 6)
Tento výraz určuje stav s jakým skončila diskretizační metoda počítající numerickou simulaci. Mohlo dojít k těmto 4 situacím
 1. equilibrium - výpočet dosáhl stavu, kdy se všechny změny látek blíží nule (s danou přesností, jež může být nastavitelným parametrem metody), další generování bodů až do dosažení času τ je zbytečné,
 2. cycle - výpočet dosáhl stavu, kdy se body začínají opakovat, stačí jen zaznamenat začátek cyklu,
 3. bound - numerická simulace dosáhla na některé z dimenzí hranice globálního prostoru $\mathcal{G}^n$, buď minimální nebo maximální, také v tomto případě již nemá smysl generovat další body,
 4. timeout - numerická simulace vygenerovala diskretizované body na celém intervalu délky τ , aniž by dosáhla ekvilibria, cyklu nebo hranice prostoru.
- $check_distance(F_a.\overline{\mathcal{X}}, F_a.S, F_b.\overline{\mathcal{X}}, F_b.S)$ (řádek 12)
funkce počítající relaci blízkosti mezi diskretizovanými trajektoriemi $F_a.\overline{\mathcal{X}}$ a $F_b.\overline{\mathcal{X}}$ beroucí v potaz jejich koncový status. Více o implementaci této funkce v kapitole 5.
- $\frac{F_a.[X] + F_b.[X]}{2}$ (řádek 14)
výraz vyjadřující umístění iniciálního bodu nově vkládaného reprezentanta v případě, kdy bylo detekováno vzdálení dvou reprezentantů. Pokud došlo k oddálení, je vložen nový reprezentant s iniciálním bodem na geometrickém středu mezi iniciálními body původních reprezentantů. Současně s tímto jsou tomuto reprezentantovi vloženy do množiny sousedních reprezentantů indexy obou oddálivších se reprezentantů. Tímto postupem je zajištěno, že původně nahrazená relace blízkosti není porušena ale je transformována do podoby dvou nových relací.

V krajním případě dvou bodů, jejichž trajektorie divergují na opačné strany může dojít k několika iteracím tohoto *zahušťovacího* procesu. Z podmínky spojitě závislosti řešení vyplývá, že po konečném počtu

kroků (s přihlédnutím na metodu půlení intervalu a konečnou přesnost numerické simulace) musí být dosažen takový bod, jehož trajektorie skončí v blízkém okolí ekvilibria, jenž se muselo (opět z podmínky spojitosti řešení) někde mezi body, kam původní dvě trajektorie divergovaly, vyskytovat.

- *check_property* ($F.\overline{\mathcal{X}}, P$) (řádek 22)
funkce počítající nad diskretizací trajektorie $F.\overline{\mathcal{X}}$ pravdivost vlastnosti P . Vzhledem k tomu, že se jedná o lineární posloupnost stavů je tento postup celkem přímočarý, je však nutné ošetřit několik možných okrajových podmínek daných různými způsoby, kterak mohla metoda provádějící diskretizaci skončit (viz druhý bod výše), detailněji je toto rozvedeno v kapitole 5.
- *neighbourhood* ($F.\overline{[X]}_0^\alpha$) (řádek 24)
funkce vracející okolí počátečního bodu $F.\overline{[X]}_0^\alpha$.

Kapitola 4

Evaluace metody

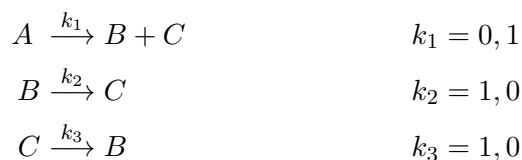
Vzhledem k způsobu implementace, který je podrobněji představen v následující kapitole je vhodnější představit spíše možnosti této metody než její rychlost a optimalizovanost. K tomuto účelu byly vybrány čtyři jednodušší modely.

Pro rámcovou představu jsou uvedeny časy výpočtů u jednotlivých příkladů. Měření proběhlo na notebooku s procesorem Intel Core 2 Duo, T7300, 2.0 GHz, 2.0 GB RAM, operační systém Windows Vista, JRE 6. Jelikož nebylo možné zamezit spouštění některých náročných procesů na pozadí, nejsou měření zcela přesná, odchylka z několika měření s pohybuje v řádu 0,1 s. Paměťové nároky aplikace byly v rozmezí 50-700 MB (nejvíce pro jednoduchý model Linear s velkým počtem iniciálních bodů).

Byl také měřen čas běhu metody při ověřování vlastností jednak uvedených v následujících sekcích, druhak ale také obecné vždy platné vlastnosti $G \models true$ vyjádřené jednoduchým jednostavovým Büchiho automatem se vždy umožněným přechodem. Rozdíly v naměřených časech byly v řádu chyby měření (0,1s) a tudíž ukázaly, že náročnou částí celého výpočtu je výpočet diskretizací a relace blízkosti ověření vlastností dané Büchiho automatem.

4.1 Linear: jednoduchý model

Tento model obsahující pouze jednomolekulárních reakce rozpadu a dvou přeměn, jedná se o umělý multiafinní systém, ve kterém se látka A rozpadá na B a C . Koncentrace látek B, C se díky rychlé vratné reakci vyrovnávají. V závislosti na počáteční koncentraci látky A se systém ustálí v některém z nekonečně mnoha stabilních ekvilibrií.



Dynamika látek A, B, C je popsána třemi rovnicemi

$$\begin{aligned}\frac{d[A]}{dt} &= -0.1A \\ \frac{d[B]}{dt} &= 0.1A + (-1.0)B + 1.0C \\ \frac{d[C]}{dt} &= 0.1A + (-1.0)C + 1.0B\end{aligned}$$

V tomto modelu byla hledána část množiny iniciálních podmínek, z níž se systém ustálí ve stavu, kdy bude koncentrace B i C větší než 6.

Formule LTL odpovídající této vlastnosti je $G(B > 6 \wedge C > 6)$.

Globálních prostor koncentrací byl omezen na

$$\mathcal{G}^n = [0, 0; 10, 0] \times [0, 0; 10, 0] \times [0, 0; 0, 10]$$

prostor iniciálních podmínek byl zvolen

$$\mathcal{I}^n = [4, 0; 6, 0] \times [0, 0; 2, 0] \times [2, 0; 4, 0]$$

počáteční podmínky byly rozděleny na nejprve na 4 části a poté na 9 částí na každé z os, čímž vzniklo nejprve 125 bodů a poté 1000. Délka časového intervalu τ byla zvolena 240s a maximální krok Δt na 0,1s. Parametry metriky Δ_i^{max} byly zvoleny 0,5 z těch důvodů, že při 1000 bodech docházelo zprvu při procesu zahušťování množiny reprezentantů k nedostatkům paměti.

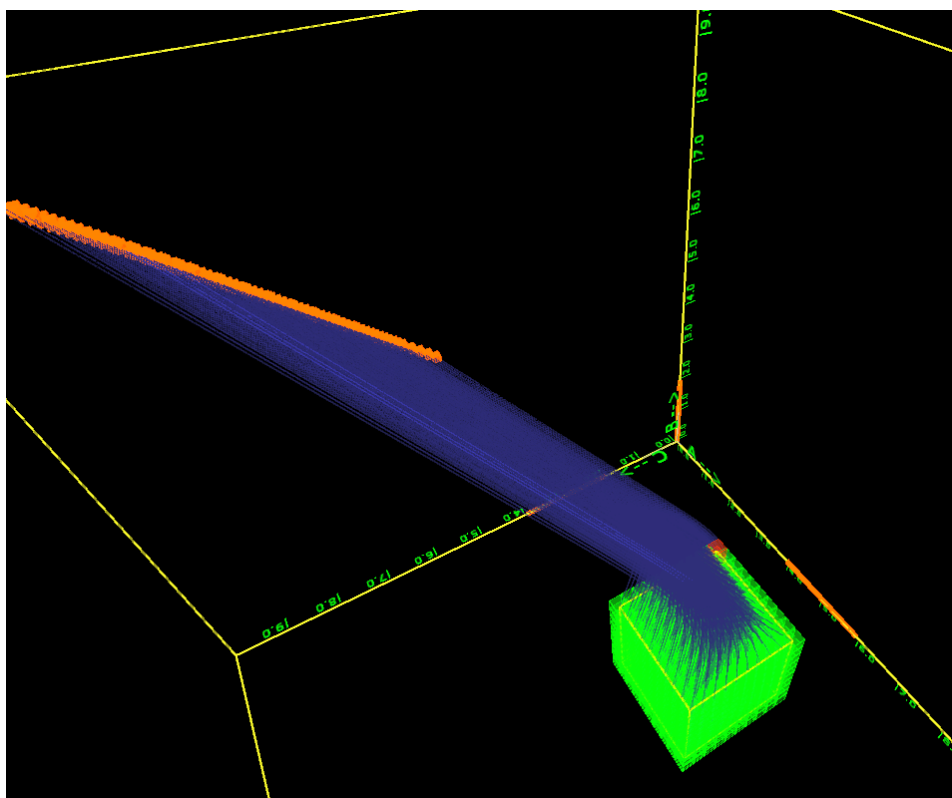
Výpočet s 125 iniciálními body skončil za 16 sekund s 274 reprezentanty, druhý výpočet s 1000 iniciálními body skončil za 52 sekund s 1237 reprezentanty.

Byla identifikována podmnožina prostoru iniciálních podmínek, ze které jsou dosažitelné vysoké hodnoty B a C , zároveň byla zjištěna i ta část, ze které jsou dosažitelné jen nízké koncentrace.

Tento jednoduchý model ověřil, že metoda je schopna detekovat množiny z níž jsou jiné definované množiny dosažitelné. Také na tomto modelu byla ověřena schopnost detekovat rovnovážné body dané nulovou změnou složek v čase.

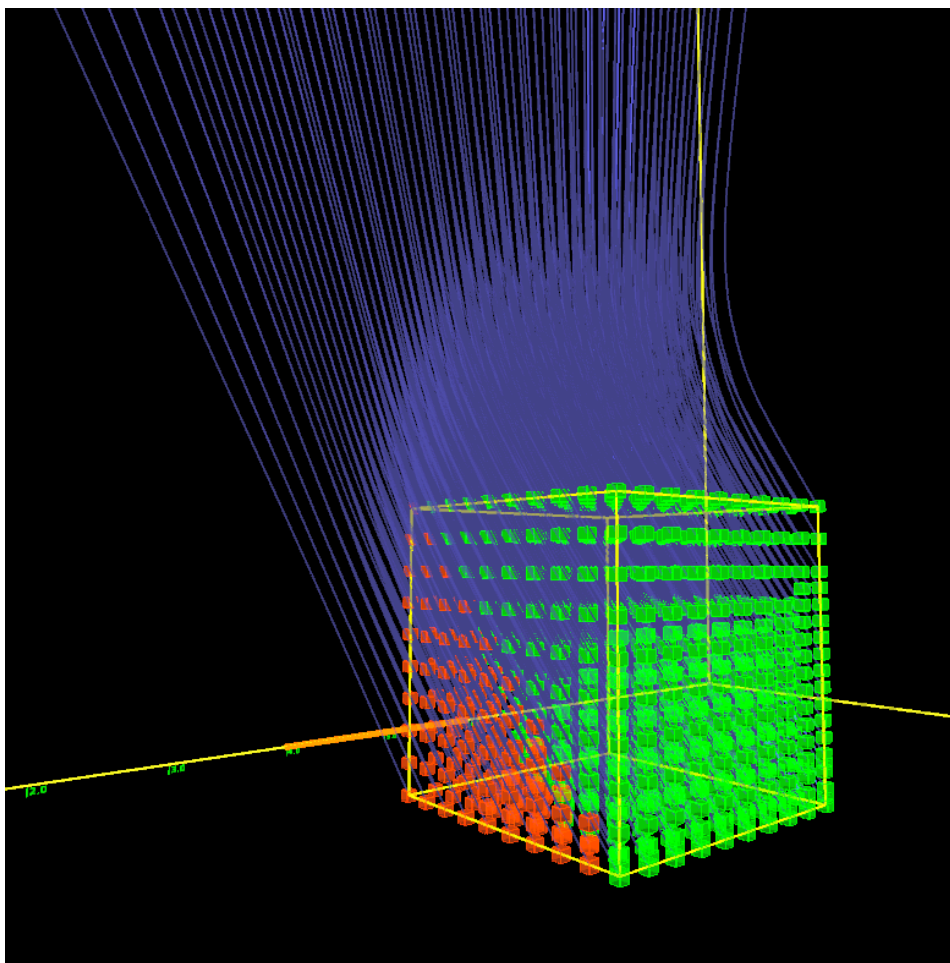
4.2 Lotka-voltera: oscilace bez limitního cyklu

Model Lotka-Volterra je velmi dobře prostudovaný dvourozměrný systém diferenciálních rovnic popisující dynamiku chování dvou druhů entit, jedna



Obrázek 4.1: Linear: Pohled na množinu koncových bodů diskretizací i počátečních podmínek, zeleně jsou označené ty části iniciálních podmínek, ze kterých jsou dosažitelné vysoké koncentrace B, C .

(označovaná jako kořist) se živý na zdrojích pocházejících z okolí v stabilním množství, druhá (označovaná jako lovec) se živí na první entitě. Tento vztah je často používán pro popis vývoje populací určitých druhů organismů a umožňuje zjistit jak různé parametry prostředí ovlivňují periodické oscilace v počtech jedinců jednoho a druhého druhu a jak je celý systém robustní vůči impulsům (velikost populací v depresní fázi). Ačkoliv byl tento model zaveden a popularizován na makroskopických systémech, mohou



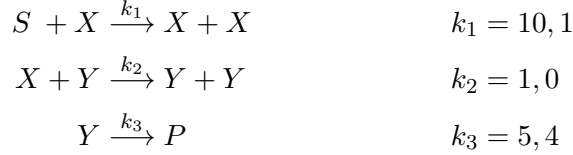
Obrázek 4.2: Linear: Bližší pohled na rozdělení intervalu iniciálních podmínek na část splňující danou podmínku (zeleně) a nesplňující (červeně).

vzniknout systémy se stejnou dynamikou i na mikroskopické úrovni v podobě populací bakterií [1].

Tento model obsahující dvě dvoumolekulární a jednu jednomolekulární reakci a jedná se o multiafinní systém.

Sestává ze čtyř látek, přičemž dvě z nich – substrát a produkt, nebo též potrava a odpadní látky – mají stabilní koncentrace, tudíž pro ně nejsou

uvažovány žádné rovnice.



kinetické koeficienty byly zvoleny náhodně.

Dynamika proměnlivých látek X, Y, Z je popsána třemi rovnicemi

$$\begin{aligned}
 \frac{d[X]}{dt} &= 10.1X + (-1.0)XY \\
 \frac{d[Y]}{dt} &= 1.0XY + (-5.4)Y
 \end{aligned}$$

Veškeré trajektorie tohoto modelu by měly být uzavřené cykly pohybující se soustředně kolem nestabilního rovnovážného bodu.

Je tedy zajímavé se ptát, zda jak budou vypadat cyklické trajektorie a zda budou zdetekovány.

Cyklus lze definovat formulí LTL $G((prop \Rightarrow F(\neg prop)) \wedge (\neg prop \Rightarrow F(prop)))$ kde $prop$ je výraz definující prahovou hodnotu kolem které určitá látka osciluje. Zvolen byl výraz $(Y < 1)$.

Globálních prostor koncentrací byl omezen na

$$\mathcal{G}^n = [0, 0; 100, 0] \times [0, 0; 100, 0]$$

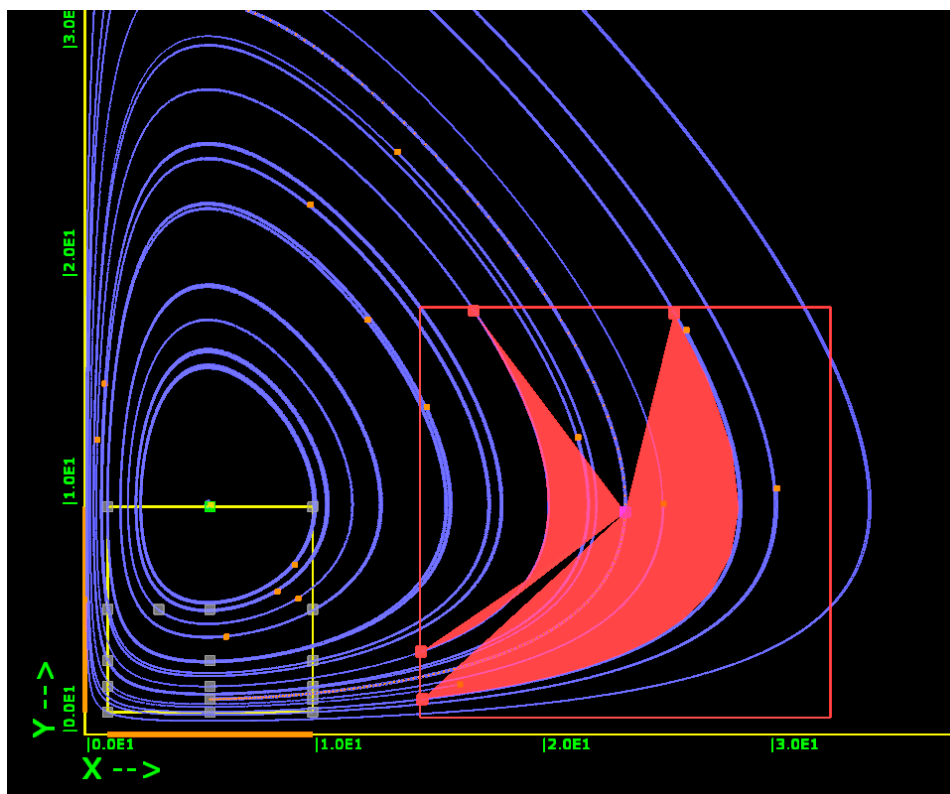
prostor iniciálních podmínek byl zvolen

$$\mathcal{I}^n = [1, 0; 10, 0] \times [1, 0; 10, 0]$$

počáteční podmínky byly rozděleny na dva intervaly v obou osách intervalu τ byla zvolena 120s a maximální krok Δt na 0,01s.

Výpočet s tímto nastavením byl ukončen po 15 minutách uživatelem, původních 9 reprezentantů bylo *zahuštěno* na 22. Přičemž u žádné z diskretizací nebyl detekován cyklus a místo spojitých křivek byly vytvořeny dlouhé spirály pomalu se vytácející směrem ven, které dospěly v čase až na hranici τ .

Tento model ukázal, že jednoduchá Eulerova metoda numerické simulace není s použitým nastavením dostatečně přesná pro detekci cyklů heuristikou extrémních bodů (viz kapitola 5).



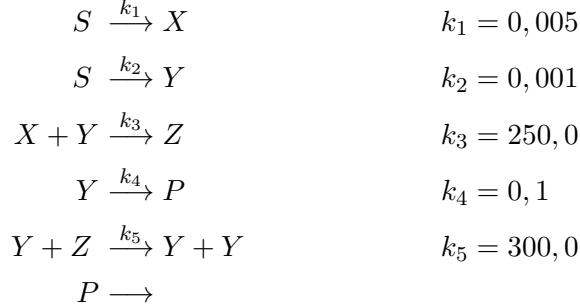
Obrázek 4.3: Lotka-Volterra: nepřesná numerická simulace způsobila vznik spirál místo detekce uzavřených křivek, červené trojúhelníky jsou vizualizací metody počítající relaci blízkosti mezi diskretizacemi trajektorií.

4.3 Bayramov: oscilace s limitním cyklem

Tento model obsahující pouze jedno a dvoumolekulárních reakce, jedná se o multiafinní systém a vykazuje ustálené oscilující chování [6].

Sestává z pěti látek, přičemž dvě z nich – substrát a produkt – mají sta-

bilní koncentrace, tudíž pro ně nejsou uvažovány žádné rovnice.



Dynamika proměnlivých látek X, Y, Z je popsána třemi rovnicemi

$$\begin{aligned}
 \frac{d[X]}{dt} &= 0.0005 + (-250.0)XY \\
 \frac{d[Y]}{dt} &= 0.0001 + (-0.1)Y + (-250.0)XY + 300.0YZ \\
 \frac{d[Z]}{dt} &= 250.0XY + (-300.0)YZ
 \end{aligned}$$

V tomto modelu by měla existovat jediná stabilní cyklická trajektorie, s poměrně velkou atrakční množinou. Toto znamená, že trajektorie uvnitř i vně cyklu by k němu měly konvergovat. Uvnitř cyklu je labilní rovnovážný bod.

Je tedy zajímavé se ptát, ze kterých částí prostoru iniciálních podmínek je dosažitelný limitní cyklus?

Cyklus lze definovat formulí LTL $G((prop \Rightarrow F(\neg prop)) \wedge (\neg prop \Rightarrow F(prop)))$ kde $prop$ je výraz definující prahovou hodnotu kolem které určitá látka osciluje. Na základě prvotního prozkoumání dynamiky systému byla zvolena podoba výrazu $(Y > 0.001)$.

Globálních prostor koncentrací byl omezen na

$$\mathcal{G}^n = [0, 0; 0, 01] \times [0, 0; 0, 01] \times [0, 0; 0, 01]$$

prostor iniciálních podmínek byl zvolen

$$\mathcal{I}^n = [0, 0015; 0, 004] \times [0, 00095; 0, 00405] \times [0, 00038; 0, 00482]$$

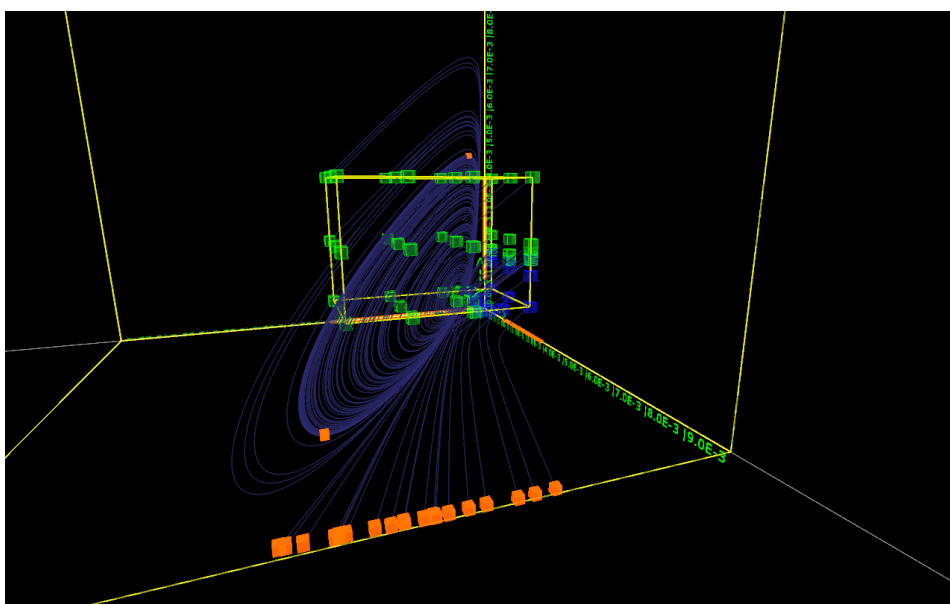
počáteční podmínky byly rozděleny na dva intervaly v osách X, Y a na tři v ose Z . Délka časového intervalu τ byla zvolena 960s a maximální krok Δt na 0,1s.

Výpočet s tímto nastavením skončil za 6 minut 34 sekund, původních 36 reprezentantů bylo *zahuštěno* na 74.

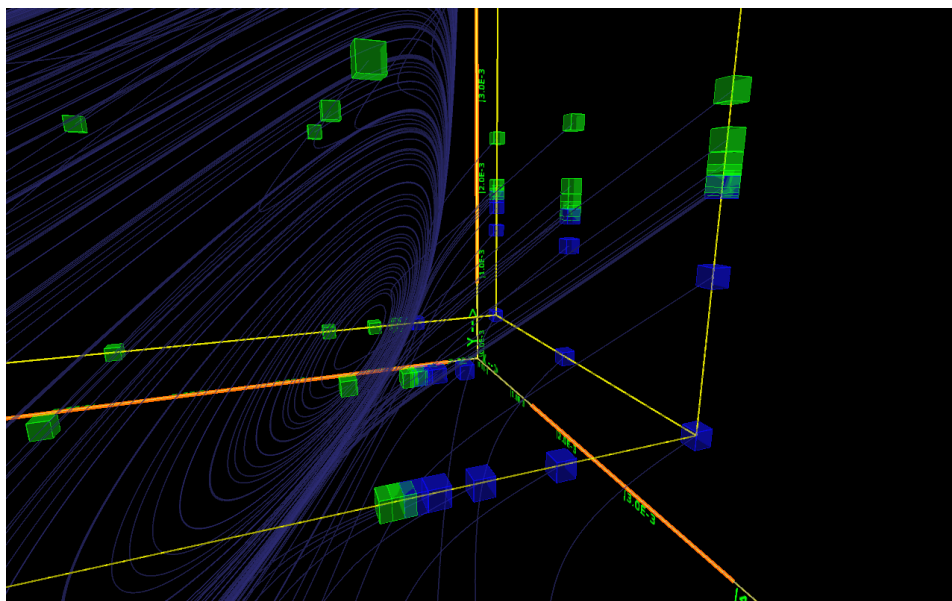
4. EVALUACE METODY

Bylo zjištěno, že velká část prostoru iniciálních podmínek je skutečně atrakční množinou limitního cyklu. Pro část prostoru iniciálních podmínek blízko počátku soustavy však trajektorie divergují od cyklu až narazí na hranici globálního prostoru koncentrací a vlastnost P pro ně zůstane nerozhodnuta, jelikož je možné, že by se ještě otočily zpět a dokonvergovaly k limitnímu cyklu, ale stejně tak mohou divergovat do nekonečna, to však nelze uvnitř prostoru $\mathcal{G}^n$ touto metodou rozhodnout.

Tento model tedy prokázal, že metoda je schopna detekovat atrakční množiny limitních cyklů a jejich hranice. Zároveň také prokázal, že je schopen detekovat vlastní cyklus, tedy uzavřenou trajektorii, resp. její diskretizaci s konečnou přesností, pomocí heuristiky extrémních bodů cyklu (viz kapitola 5).



Obrázek 4.4: Bayramov: Globální prostor koncentrací $\mathcal{G}^n$ s výsledky analýzy: zelená = P platí tedy část atrakční množiny cyklu uvnitř iniciálních podmínek, modrá = nerozhodnuto – trajektorie odtud vycházející dosáhly hranice $\mathcal{G}^n$.

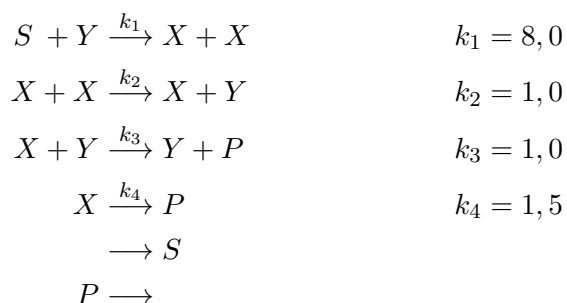


Obrázek 4.5: Bayramov: Bližší pohled na hranici atrakční množiny limitního cyklu.

4.4 Wilhelm: minimální bistabilní model

Tento model složený jen z jedno a dvoumolekulárních reakcí je minimální v počtu látek, reakcí a počtu výrazů na pravých stranách diferenciálních rovnic [34].

Sestává ze čtyř látek, přičemž dvě z nich – substrát a produkt – mají stabilní koncentrace, tudíž pro ně není nutné uvažovat změny v čase.



Dynamika látek je popsána dvěma rovnicemi

$$\begin{aligned}\frac{d[X]}{dt} &= 16,0Y + (-1,0)X^2 + (-1,0)XY + (-1,5) \\ \frac{d[Y]}{dt} &= 1,0X^2 + (-8,0)Y\end{aligned}$$

V tomto modelu by měly existovat dva rovnovážné body, jeden v oblasti vyšších koncentrací a druhý u počátku soustavy. V oblasti kolem počátku se také vyskytuje dělicí hranice mezi atrakčními množinami těchto dvou stabilních ekvilibrií, toto také znamená, že by se zde měl vyskytovat také labilní rovnovážný bod.

Položené otázky tedy byly, kde se vyskytují ony rovnovážné body, kudy prochází hranice atrakčních množin a zda systém vykazuje bistabilitu.

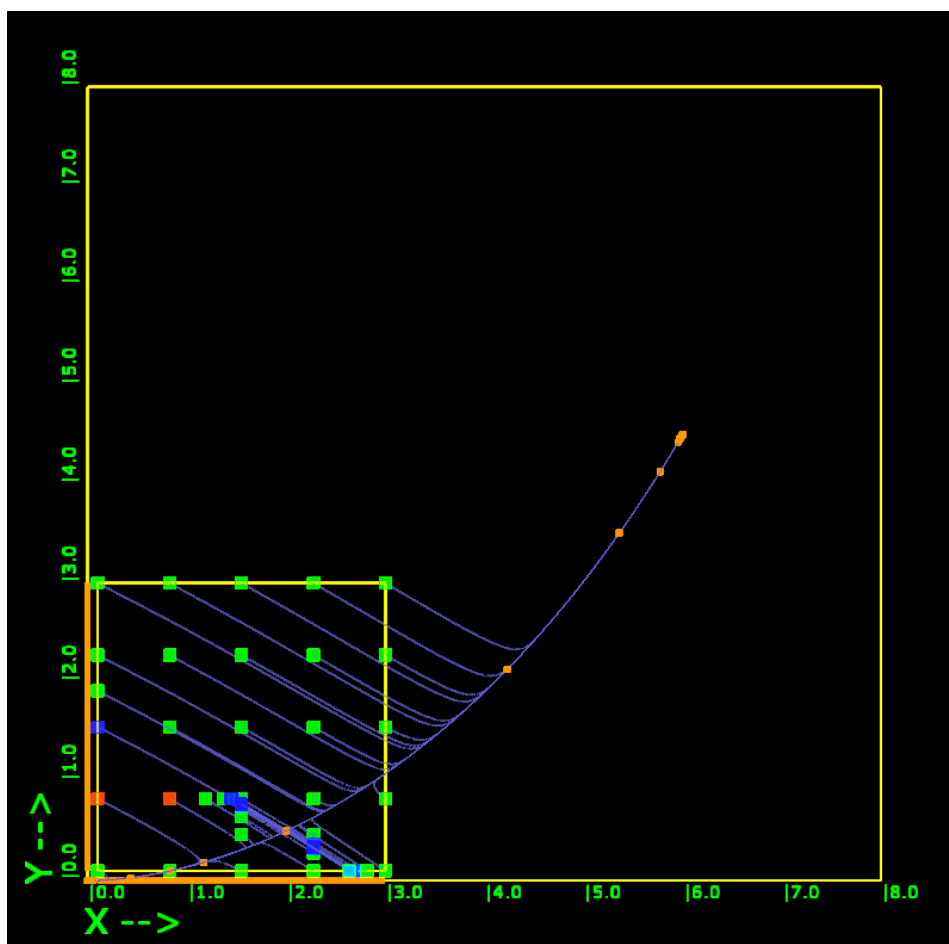
Bistabilita může být definována formulí LTL $G((prop1 \Rightarrow Gprop1) \wedge (prop2 \Rightarrow Gprop2))$ kde $prop1$ a $prop2$ jsou výrazy definující jednu a druhou oblast stability. Oblast ekvilibria s vysokou koncentrací byla definována výrazem $(X > 5,0 \wedge Y > 5,0)$, oblast u počátku souřadnic $(X < 0,2 \wedge Y < 0,2)$.

Globálních prostor koncentrací byl omezen na $\mathcal{G}^n = [0, 0; 8, 0] \times [0, 0; 8, 0]$, prostor iniciálních podmínek byl zvolen $\mathcal{I}^n = [0, 1; 3, 0] \times [0, 1; 3, 0]$, počáteční podmínky rozděleny na tři části na každé dimenzi. Délka časového intervalu τ byla zvolena 20s a maximální krok Δt na 0,01s.

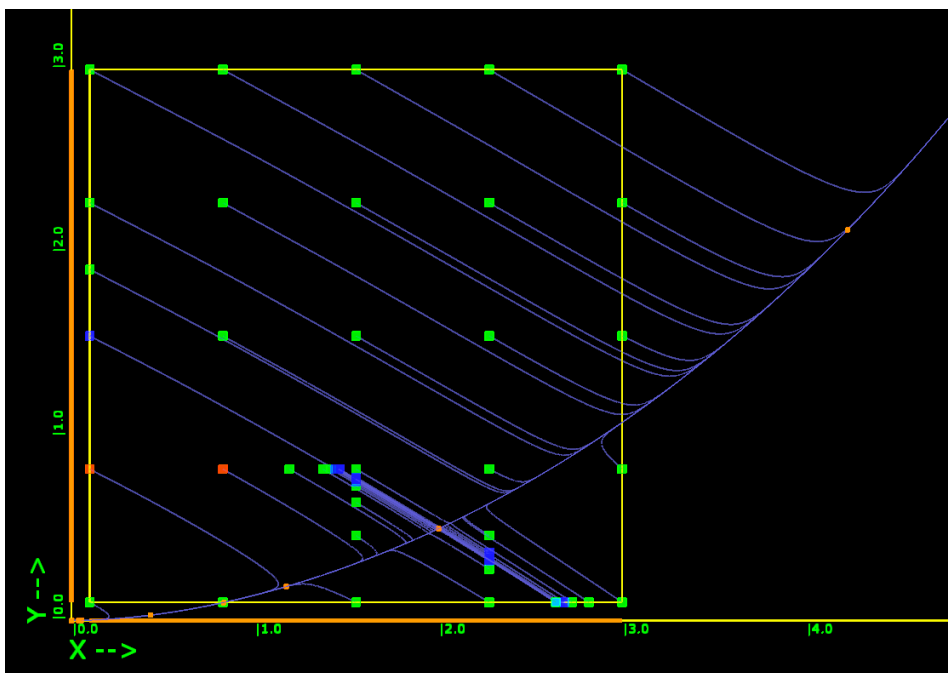
Výpočet s tímto nastavením skončil za 2 minuty 47 sekund, původních 16 reprezentantů bylo *zahuštěno* na 97. Horní stabilní ekvilibrium bylo lokalizováno na souřadnicích $[6, 0; 4, 5]$, dolní v $[0; 0]$ a labilní v bodě $[2, 0; 0, 5]$. Tyto hodnoty jsou v souladu s výsledky publikovanými v článku, z něhož byl model převzat.

Definovaná vlastnost bistability byla shledána pravdivou (zelená barva) pro většinu prostoru iniciálních podmínek, přičemž výjimkou je část prostoru $[0; 1] \times [0, 5; 1]$ jelikož trajektorie z ní vycházející opustí námi definovanou atrakční množinu $(X < 1 \wedge Y < 1)$ a poté se do ní opět vrátí (vyznačeno červeně). Vlastnost také zůstává nerozhodnuta (modrá barva) u několika trajektorií na hranici obou atrakčních množin, které buď uvízly v labilním rovnovážném bodě, nebo se průchodem v jeho blízkosti natolik *opozdily*, že dosáhly konce časového intervalu, pro něž byla vlastnost rozhodována.

Tento model tedy prokázal, že metoda je schopna detekovat bistabilitu, hranice atrakčních množin, ekvilibria stabilní i labilní.



Obrázek 4.6: Wilhelm: Globální prostor koncentrací $\mathcal{G}^n$ s výsledky analýzy: zelená = P platí, červená = P neplatí, modrá = nerozhodnuto.



Obrázek 4.7: Wilhelm: Bližší pohled na hranice atrakčních množin, labilní ekvilíbrium a *opožděné* body.

Kapitola 5

Návrh a implementace

5.1 Nastroje a specifika

Jako prostředí pro vývoj implementace navrhované metody jsem zvolil nástroj Processing 1.0.9. Jedná se o třídu jazyka Java a přidružené vývojové prostředí, které umožňuje velmi rychle vizualizovat myšlenkové koncepty díky soustavě příkazů notně zjednodušujících práci s grafikou a základními datovými objekty, avšak umožňující stále využívat plných možností jazyka Java.

Toto je odklon od původního zadání práce, které vyžadovalo paralelní implementaci v MPI. Považuji to však za rozhodnutí nejen výhodné, ale také správné, neboť teprve potom co je metoda vyvinuta a evaluována její schopnost a korektnost má smysl se pouštět do její optimalizace a paralelizace.

Je však potřeba dodat, že při tvorbě implementace bylo pamatováno na možnou budoucí paralelizaci a datový model i schéma výpočtu jí umožňují.

V rámci Processingu je využita grafická knihovna OpenGL. Při implementaci metody numerické simulace jsem se také snažil inspirovat existujícími metodami jako například Runge-Kutta-Fehlberg avšak nakonec jsem z časových důvodů použil jen jednoduchou Eulerovu integrační metodu s velkou chybou, tento aspekt však nebyl jádrem práce a bude moci být v budoucnosti případně modulárně vyměněn.

5.2 Funkcionalita a interaktivita

Aplikaci je možné z Processingu exportovat do webového appletu nebo aplikace pro Windows, Mac či Linux.

Po spuštění aplikace je možné načíst model ve formátu *.ode stiskem klávesy 'o' a soubor s vlastností v podobě Büchiho automatu stiskem klávesy 'p'.

Po načtení modelu (pokud není dodána explicitní vlastnost, použije se implicitní *Gtrue*) může uživatel s prostorem koncentrací rotovat (levé tla-

čítka myši a pohyb kurzorem), přibližovat a oddalovat (klávesy plus a mínus) či posouvat (shift + rotace).

Výpočet se spouští a pozastavuje mezerníkem. Vizualizaci výpočtu lze zrychlit tím, že se sníží preciznost vykreslování křivek (D) nebo zvýší (E).

Klepnutím na objekty, jsou tyto vybírány a jsou vizualizovány jejich vlastnosti.

Již vypočítané diskretizace jsou tmavší a průsvitnější, u vybrané trajektorie jsou zobrazeny dílčí body a průběh výpočtu relace blízkosti.

Množina iniciálních bodů je barevně označena podle platnosti definované vlastnosti nad diskretizací vycházející z onoho bodu (šedá - výpočet nedospěl k rozhodnutí, zelená - splněno, červená - porušeno, modrá - nelze rozhodnout).

Během chodu aplikace je do konzole jazyka Java směrován textový výstup, z něhož lze získat další informace o průběhu výpočtu (vkládání nových reprezentantů, koncové stavy dílčích výpočtů atp.)

Na co je aplikace nachystána, avšak nebylo zatím realizováno, je přepínání zobrazení látek na různých souřadnicích u více-dimenzionálních modelů. Stejně tak možnost kontrakce či dilatace každé z os nezávisle, by byla velmi vhodná.

5.3 Popis formátů

Popis vstupních formátů je v souborech *ode_format.txt* a *ltl_format.txt* v adresáři *examples*.

Formát *.ode je převzatý a modifikovaný formát *.bio nástroje BioDiVinE, *.ltl je restringovaný formát používaný nástrojem DiVinE pro popis vlastnostních automatů. Oba jsou to textové formáty s intuitivním významem.

5.4 Struktura tříd

Jednotliví dílčí třídy jsou seskupeny do samostatných souborů. Ve skutečnosti jsou to všechno vnitřní třídy výsledné aplikace, takže vzájemné volání vnitřních metod je jednoduché, avšak méně nutí do separace funkcionality a zapouzdřenosti.

Následuje krátký popis tříd v jednotlivých souborech, u velké části z tříd jsou rozsáhlé komentáře přímo ve zdrojovém kódu, u velké části však také chybí z důvodů velmi překotného vývoje některých částí aplikace v určitých obdobích:

- AtomicPropositions.pde – definuje atomické propozice a jejich evaluaci v prostorovém bodě, navíc i kompozici propozic do výrazů
- Computation.pde – hlavní třída řídící běh výpočtu, tři nejdůležitější metody jsou `extend_simulations` – vypočítá další kus diskretizace každé trajektorie a `check_simulation_distances` – počítá průběžně relaci blízkosti
- CycleDetectionDataBlock.pde – obsahuje datové struktury a metody pro detekci cyklu na diskretizaci pomocí heuristiky extrémních bodů
- DistanceCheckDataBlock.pde – struktura udržující reprezentaci množiny podposloupností blízkých bodů, a výpočet relace blízkosti mezi dvěma diskretizacemi
- GUI.pde – řídí uživatelské rozhraní a interaktivitu
- ODE.pde – reprezentuje dynamický systém a jeho soustavu diferenciálních rovnic dohromady s prostorem globálních koncentrací, iniciálních podmínek a metriky vzdáleností bodů
- ODEParser.pde – umožňuje parsovat a načítat *.ode soubory do třídy ODE
- PointsDataBlock.pde – datová struktura umožňující držet body diskretizace v zřetěžených blocích, aby bylo možné je po kusech přenášet nebo dealokovat, když už nejsou potřeba
- Property.pde – reprezentuje vlastnostní automat
- PropertyCheckDataBlock.pde – struktura provádějící ověřování vlastnosti nad diskretizací
- PropertyParser.pde – parsuje *.ltl soubory do třídy Property
- SimulationData.pde – datová struktura sdružující všechny potřebné objekty obhospodařující výpočet jednoho reprezentanta
- SimulationStorage.pde – kontejner na reprezentanty
- utility.pde – pomocné metody a konstanty
- ViewPort.pde – třída definující uživatelův pohled na právě probíhající výpočet (rotace, přiblížení, způsoby promítání...)
- vizualizace.pde – hlavní zastřešující třída kombinující vše do jediné aplikace

5.5 Algoritmy

Algoritmy v metodě použité nejsou nikterak nové ani objevné a krom dvou (výpočet množiny podposloupností blízkých bodu a detekce cyklu) jsou to základní principy ovšem vzájemně dosti provázané.

Detekce cyklů vychází z pozorování, že každý cyklus musí mít pro každou dimenzi extrémní bod, tedy takový bod, jenž má buď minimální, nebo maximální hodnotu v této dimenzi. Stačí si tedy pro každou dimenzi pamatovat několik posledních extrémních bodů a v okamžiku, kdy se změny derivace z kladné na zápornou, vložit tento bod jako další, resp. porovnat ho s těmi, které už v seznamu jsou.

V principu je tento postup heuristika, neboť lze nezávislou kompozicí vytvořit ze dvou oscilujících dynamických systémů s nesoudělnou periodou nový, s libovolně mnoha lokálními extrémy, takže pro detekci takového umělého pseudocyklu by nestačil libovolně velký seznam. Naštěstí se takové případy v reálných dynamických systémech nevyskytují a tudíž pro ně tato heuristika funguje celkem dobře.

Výpočet množiny podposloupností blízkých bodu přímo kopíruje svou formální definici a využívá jen faktu, že opakované vkládání téhož prvku do množiny není zapotřebí.

Kapitola 6

Závěr

Cílem práce bylo navrhnout a implementovat metodu analýzy dynamiky systémů popsaných soustavou diferenciálních rovnic využívající tradiční techniky numerické simulace a kombinující je s metodami ověřování vlastností diskretních systémů vycházejících z metod ověřování modelu.

V této práci byla formálně definována relace blízkosti diskretizací numerické aproximace trajektorií řešení dynamických systémů, která umožňuje konstrukci množiny reprezentativních diskretizací. Ověřením vlastnosti vyjádřitelné formulí lineární temporální logiky a reprezentované Büchiho automatem nad množinou takto získaných reprezentantů lze usuzovat o chování celé množiny řešení původního dynamického systému.

Metoda byla implementována v nástroji Processing, využívající jazyka Java do podoby prototypu, který vizualizuje průběh vnitřních výpočtů metody. Toto je odlišnost od původního zadání diplomové práce, která požadovala paralelní implementaci v MPI. Důvodem je hlavně možnost lepšího chápání vnitřního fungování metody za účelem dalšího vývoje, jež by v prostředí MPI nebylo tak snadné.

Metoda byla evaluována na příkladech malých dynamických systémů a byly prezentovány její schopnosti detekce vlastností vyjadřujících se o dosažitelnosti, existenci rovnovážných bodů, existenci cyklických trajektorií a bistability.

Z empirických vlastností dynamických systémů byla využita hlavně spojitost závislosti řešení na počátečních podmínkách, ne však příliš vlastnost multiafinity. Metoda by tudíž mohla být použitelná pro obecné dynamické systémy popsané diferenciálními rovnicemi se spojitými pravými stranami, avšak některá silná projevy multiafinity tak nadále zůstávají nevyužity a čekají na možné vytěžení v budoucnosti.

Literatura

- [1] Frederick K. Balagadde, Hao Song, Jun Ozaki, Cynthia Collins, Matthew Barnett, Frances H Arnold, Stephen R Quake, and Lingchong You.
- [2] J. Barnat, L. Brim, I. Černá, S. Dražan, J. Fabriková, J. Láník, D. Šafránek, and Hongwu Ma. BioDiVinE: A Framework for Parallel Analysis of Biological Models. *Electronic Proceedings in Theoretical Computer Science (COMPMOD 2009)*, 6:31–45, 2009.
- [3] J. Barnat, L. Brim, I. Černá, S. Dražan, J. Fabriková, and D. Šafránek. Computational Analysis of Large-Scale Multi-Affine ODE Models. In *2009 International Workshop on High Performance Computational Systems Biology (HiBi 2009)*, pages 81–90. IEEE Computer Society Press, 2009.
- [4] J. Barnat, L. Brim, I. Černá, S. Dražan, J. Fabriková, and D. Šafránek. On algorithmic analysis of transcriptional regulation by LTL model checking. *Theor. Comput. Sci.*, 410(33-34):3128–3148, 2009.
- [5] J. Barnat, L. Brim, I. Černá, S. Dražan, and D. Šafránek. Parallel Model Checking Large-Scale Genetic Regulatory Networks with DiVinE. *ENTCS*, 194(3):35–50, 2008.
- [6] Sh.K. Bayramov. Concentration oscillations in three-component reaction systems. *Biochemistry (Moscow)*, 68:349–353, 2003. 10.1023/A:1023070604117.
- [7] C. Belta and L.C.G.J.M. Habets. Controlling a class of nonlinear systems on rectangles. *IEEE Transactions on Automatic Control*, 51(11):1749–1759, 2006.
- [8] George Edward Briggs and John Burdon Sanderson Haldane. A note on the kinetics of enzyme action. *Biochemical Journal*, 19(2):338—339, 1925.

- [9] Edmund Clarke, James Faeder, Christopher Langmead, Leonard Harris, Sumit Jha, and Axel Legay. Statistical model checking in biolab: Applications to the automated analysis of t-cell receptor signaling pathway. In Monika Heiner and Adelinde Uhrmacher, editors, *Computational Methods in Systems Biology*, volume 5307 of *Lecture Notes in Computer Science*, pages 231–250. Springer Berlin / Heidelberg, 2008.
- [10] Edmund M. Clarke, Orna Grumberg, and Doron A. Peled. *Model Checking*. MIT Press, 2000.
- [11] Copasi.
URL: http://www.copasi.org/tiki-view_articles.php.
- [12] Hidde de Jong, Jean-Luc Gouzé, Céline Hernandez, Tewfik Sari, Michel Page, and Johannes Geiselmann. Qualitative simulation of genetic regulatory networks using piecewise-linear models, 2002.
- [13] Divine homepage.
URL: <http://anna.fi.muni.cz/divine>.
- [14] Ec-moan escherichia coli modeling and analysis.
URL: <http://www.ec-moan.org>
URL: http://cordis.europa.eu/fetch?CALLER=FP6_PROJ&ACTION=D&RCN=85153.
- [15] Gna: Genetic network analyzer. URL:
<http://www-helix.inrialpes.fr/article122.html>.
- [16] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(No. 25):2340–2381, 1977.
- [17] Philip Hartman. *Ordinary Differential Equations*. Cambridge University Press, 2002.
- [18] Alan C. Hindmarsh. Serial fortran solvers for ode initial value problems.
URL: http://computation.llnl.gov/casc/odepack/odepack_home.html
URL:
<http://computation.llnl.gov/casc/nsde/pubs/u88007.pdf>.
- [19] Thomas R. Kiehl, Robert M. Mattheyses, and Melvin K. Simmons. Hybrid simulation of cellular behavior. *Bioinformatics*, 20(3):316–322, 2004.

-
- [20] M. Kloetzer and C. Belta. Reachability analysis of multi-affine systems. *Transactions of the Institute of Measurement and Control*, 2008.
- [21] A.J. Lotka. Contribution to the theory of periodic reaction. *J. Phys. Chem.*, 14(3):271–274, 1910.
- [22] Denis Noble. *Music of life*. Oxford University Press, 2006.
- [23] Denis Oddoux and Paul Gastin. Ltl 2 ba: fast translation from ltl formulae to büchi automata.
URL: <http://www.lsv.ens-cachan.fr/~gastin/ltl2ba/index.php>.
- [24] Amir Pnueli. The temporal logic of programs. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 46–57, 1977.
- [25] Aurélien Rizk, Grégory Batt, François Fages, and Sylvain Soliman. On a continuous degree of satisfaction of temporal logic formulae with applications to systems biology. In *CMSB*, volume 5307 of *LNCS*, pages 251–268. Springer, 2008.
- [26] Donald R. Sadoway. *Introduction to Solid State Chemistry*. MIT Open Courseware.
URL: <http://ocw.mit.edu/courses/materials-science-and-engineering/3-091-introduction-to-solid-state-chemistry-fall-2004/video-lectures/lecture-22>.
- [27] Stránky laboratoře paradise.
URL: <http://paradise.fi.muni.cz>.
- [28] Stránky laboratoře systémové biologie.
URL: <http://sybila.fi.muni.cz>.
- [29] Iv118 - formální metody v systémové biologii. URL:
<http://is.muni.cz/predmety/predmet.pl?id=540092>.
- [30] Jan H. van Schuppen.
URL: <http://homepages.cwi.nl/~schuppen>.
- [31] Jan H. van Schuppen. System reduction and abstraction of biochemical reaction systems for a genetic network of escherichia coli. EC-MOAN Deliverable, 2008.

- [32] Eric W. Weisstein. Relative error.
URL: <http://mathworld.wolfram.com/RelativeError.html>.
- [33] Wikipedia the free encyclopedia.
URL: <http://en.wikipedia.org/wiki/Abiogenesis>
URL: <http://en.wikipedia.org/wiki/Entropy>
URL: http://en.wikipedia.org/wiki/Red_blood_cell#Human_erythrocytes
URL: http://en.wikipedia.org/wiki/Rate_equation
URL: http://en.wikipedia.org/wiki/Reaction_rate
URL: <http://en.wikipedia.org/wiki/Stoichiometry>
URL: http://en.wikipedia.org/wiki/Order_of_reaction.
- [34] Thomas Wilhelm. The smallest chemical reaction system with bistability. *BMC Systems Biology*, 3(1):90, 2009.