

**Um algoritmo para simplificar  
sistemas de equações diferenciais que  
descrevem a cinética de reações químicas**

Amanda Sayuri Guimarães

TEXTO APRESENTADO AO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DA  
UNIVERSIDADE DE SÃO PAULO  
PARA  
OBTENÇÃO DO TÍTULO  
DE  
MESTRE EM CIÊNCIAS

Programa: Ciência da Computação  
Orientador: Prof. Dr. Junior Barrera

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

17 de agosto de 2016

# Um algoritmo para simplificar sistemas de equações diferenciais que descrevem a cinética das reações químicas

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 10/06/2016. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Junior Barrera (orientador) - IME-USP
- Prof. Dr. Saulo Rabello Maciel de Barros - IME-USP
- Prof. Dr. Marcelo da Silva Reis - Instituto Butantan

# Agradecimentos

Ao meu mestre da vida, Dr. Daisaku Ikeda, o qual me incentiva todos os dias a superar os obstáculos e a fazer a minha revolução humana.

Aos meus mestres, Prof. Dr. Junior Barrera e Prof. Dr. Marcelo Reis por me concederem a oportunidade de trabalhar ao lado de vocês, por confiarem em mim e por nunca negarem um espacinho em suas agendas atribuladas, sempre me auxiliando nas diversas dificuldades que encontramos ao longo desse projeto.

A todos os meus professores, desde a pré-escola até a pós-graduação. Serei eternamente grata a todos vocês.

Aos meus queridos pais, Braga Barbosa Guimarães e Alice Fukuyo Guimarães por não pouparem esforços para me educar, apoiar, aconselhar e incentivar, sempre com muito carinho.

Aos meus queridos irmãos, Gustavo e Thaís, por sempre estarem ao meu lado e completarem os meus dias.

Ao meu companheiro amado, Eduardo Shoiti Sato, o qual torna meus dias ainda mais felizes. Obrigada por todo o amor, apoio, carinho, benevolência e por toda sua grandiosidade!

Às minhas irmãs de faculdade, Ana Livia Rodrigues Queiroz e Lulu Wu, companheiras inseparáveis que me aguentaram desde os tempos de faculdade, sempre divertindo e alegrando meus dias.

A todos os meus amigos, em especial, aos companheiros da Soka Gakkai, por me oferecerem a oportunidade de desenvolver como pessoa junto com vocês.

Aos meus colegas de laboratório e de trabalho, sempre de braços abertos para me ajudarem em qualquer situação.

Agradecimentos também ao Instituto de Matemática e Estatística da Universidade de São Paulo, o qual possui uma ótima infraestrutura, e aos membros da banca examinadora pela disponibilidade para avaliar e discutir sobre este projeto.

Por fim, agradeço à CAPES pela bolsa de estudos concedida.



# Resumo

GUIMARÃES, A. S. **Um algoritmo para simplificar sistemas de equações diferenciais que descrevem a cinética de reações químicas**. 2015. Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2015.

O estudo da evolução da concentração de elementos de uma reação química, conhecida como Cinética Química, é de extrema importância para a compreensão das complexas interações em sistemas biológicos. Uma maneira de descrever a cinética de uma reação química é utilizando um sistema de equações diferenciais ordinárias (EDOs). Uma vez que para resolver um sistema de equações diferenciais ordinárias pode ser uma tarefa difícil (ou mesmo inviável), métodos numéricos são utilizados para realizar simulações, ou seja, para obter concentrações aproximadas das espécies químicas envolvidas durante um determinado período de tempo. No entanto, quanto maior for o sistema simulado de EDOs, mais os métodos numéricos estão sujeitos a erros. Além disso, o aumento do tamanho do sistema muitas vezes resulta em simulações que são mais exigentes do ponto de vista computacional. Assim, o objetivo deste projeto de mestrado é o desenvolvimento de regras para simplificar os sistemas de equações diferenciais ordinárias que modelam a cinética de reações químicas e, portanto, a obtenção de um algoritmo para executar simulações numéricas de um modo mais rápido e menos propenso a erros. Mais do que diminuir o erro e o tempo de execução, esta simplificação possibilita o biólogo escolher a solução mais factível do ponto de vista de medida. Isso porque, a identificação dos sistemas (i.e., inferência dos parâmetros) requer que a concentração de todas as espécies químicas seja conhecida, ao menos em um certo intervalo de tempo. Contudo, em muitos casos, não é possível medir a concentração de todas as espécies químicas consideradas. Esta simplificação gera sistemas equivalentes ao original, mas que dispensa a utilização de certas concentrações de espécies químicas. Um sistema de equações diferenciais ordinárias pode ser simplificado considerando as relações de conservação de massa, que são equações algébricas. Além disso, no caso de reações enzimáticas, o sistema de equações diferenciais ordinárias pode ser simplificado pelo pressuposto de que a concentração do complexo enzima-substrato mantém-se constante, o que permite a utilização da equação de Michaelis-Menten. De todas as combinações possíveis das equações algébricas com as equações diferenciais, uma família de sistemas simplificados de EDOs foi construída, permitindo a escolha do sistema mais simples. Esta escolha segue um critério guloso que favorece a minimização do número de equações diferenciais e do número total de termos. As regras em desenvolvimento de simplificação dos sistemas de equações diferenciais ordinárias foram utilizadas para projetar um algoritmo, que foi implementado usando a linguagem de programação Python. O algoritmo concebido foi testado utilizando instâncias artificiais.

**Palavras-chave** redes de sinalização molecular, cinética química, equação diferencial ordinária.



# Abstract

GUIMARÃES, A. S. **An algorithm to simplify systems of differential equations that describe the kinetics of chemical reactions.** 2015. Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2015.

The study of the evolution of the concentration of species in a chemical reaction, known as Chemical Kinetics, is of paramount importance for the understanding of complex interactions in biological systems. One way to describe the kinetics of a chemical reaction is using a system of ordinary differential equations (ODEs). Once to solve a system of ODEs can be a difficult (or even unfeasible) task, numerical methods are employed to carry out simulations, that is, to obtain approximated concentrations of the involved chemical species for a certain time frame. However, the larger is the simulated system of ODEs, the more numerical methods are subject to error. Moreover, the increase of the system size often results in simulations that are more demanding from the computational point of view. Thus, the objective is the development of rules to simplify systems of ODEs that models the kinetics of chemical reactions, hence obtaining an algorithm to execute numerical simulations in a faster way and less prone to error. More than decrease error and run time, this simplification allows the biologist to choose the most feasible solution from the point of view of measurement. This is because the identification of systems (i.e., inferring parameters) requires that the concentration of all chemical species is known, at least in a certain time interval. However, in many cases it is not possible to measure the concentration of all chemical species considered. This simplification creates systems equivalent to the original, but that does not require the use of certain concentrations of chemical species. A system of ODEs can be simplified considering the relations of mass conservation, which are algebraic equations. Furthermore, in the case of enzymatic reactions, the system of ODEs can be simplified under the assumption that the concentration of enzyme-substrate complex remains constant, which allows us to use the Michaelis-Menten equation. From all possible combinations of the algebraic equations with differential equations, a family of simplified systems of ODEs will be built, allowing the choice of a simplest system. This choice will follow a greedy criterion which favors the minimization of number of differential equations and the total number of terms. The rules under development to simplify systems of ODEs will be used to design an algorithm, which will be implemented using Python programming language. The designed algorithm will be tested using synthetic data.

**Keywords:** molecular signaling networks, chemical kinetics, ordinary differential equation.



# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objetivos . . . . .	3
1.2	Organização do Trabalho . . . . .	3
<b>2</b>	<b>Fundamentos teóricos</b>	<b>5</b>
2.1	Cinética Química . . . . .	5
2.1.1	Taxa de Reação . . . . .	5
2.1.2	Ordem de Reação . . . . .	6
2.1.3	Reações Paralelas . . . . .	8
2.1.4	Reações Sequenciais de Primeira Ordem . . . . .	8
2.1.5	Reações Reversíveis de Primeira Ordem . . . . .	9
2.1.6	Reação Enzimática Simples . . . . .	9
2.2	Simplificação de sistemas de EDOs utilizando relações de conservação de massa . . . . .	10
2.3	Simplificação de sistemas de EDOs utilizando a equação de Michaelis-Menten . . . . .	11
<b>3</b>	<b>Mapeamento de reações químicas para sistemas de EDOs</b>	<b>13</b>
3.1	Geração de listas de taxas de reação (LTs) . . . . .	13
3.2	Geração do sistema de EDOs . . . . .	15
3.3	Comentários . . . . .	15
<b>4</b>	<b>Simplificação de sistemas de EDOs</b>	<b>17</b>
4.1	Simplificação algébrica utilizando relações de conservação de massa . . . . .	17
4.2	Simplificação utilizando hipótese de estado estacionário . . . . .	20
4.3	Comentários . . . . .	21
<b>5</b>	<b>Escolha ótima de um sistema de EADs</b>	<b>23</b>
5.1	Sistemas Lineares Indeterminados . . . . .	24
5.2	Composição de sistemas de EADs . . . . .	27
5.3	Critérios para busca de um sistema de EADs ótimo . . . . .	32
5.4	Algoritmos para a escolha de um sistema ótimo . . . . .	33
5.4.1	Algoritmo Branch and Bound . . . . .	33
5.4.2	Algoritmo Sequential Forward Search . . . . .	34
5.4.3	Percorrimento de uma única cadeia . . . . .	35
5.5	Comentários . . . . .	35

<b>6 Experimentos</b>	<b>41</b>
6.1 Experimentos . . . . .	41
6.1.1 Algoritmo para geração de dados sintéticos . . . . .	41
6.1.2 Simulação . . . . .	42
6.2 Resultados das Simulações e Discussão . . . . .	42
6.2.1 Exemplos de Resultado da Simulação de um Sistema Sintético . . . . .	42
6.2.2 Resultados Gerais . . . . .	48
6.3 Comentários . . . . .	50
<b>7 Conclusões</b>	<b>53</b>
7.1 Revisão e discussão dos resultados apresentados . . . . .	53
7.2 Contribuições . . . . .	54
7.3 Sugestões para Pesquisas Futuras . . . . .	54
<b>Referências Bibliográficas</b>	<b>55</b>

# Capítulo 1

## Introdução

A capacidade de coordenar as atividades com as mudanças ambientais é uma característica comum a todas as células. A comunicação com o ambiente é realizada por meio de vias moleculares que recebem sinais a partir do ambiente externo e também de regiões no interior da célula. A cada dia, um número cada vez maior de vias moleculares estão sendo identificadas e estudadas, evidenciando que elas fazem parte de uma rede molecular ainda mais abrangente, a qual coordena todas as atividades celulares [JLI00]. Por conta do tamanho e da complexidade, recortes dessas redes são, muitas vezes, complicados para se organizar e analisar. Assim, faz-se necessário o desenvolvimento de modelos matemáticos a fim de se entender o comportamento do sistema de redes de sinalização molecular. Com tais modelos é possível, por exemplo, identificar centros relevantes de regulação da propagação de sinal por uma via molecular, o que pode auxiliar na descoberta de possíveis alvos para intervenção farmacológica em doenças [EI04].

Uma abordagem de modelagem é a baseada na cinética química, ou seja, no estudo da evolução ao longo de um determinado tempo da concentração das espécies químicas de uma reação. A equação que descreve a dinâmica de uma espécie química  $C$  ao longo de um determinado período de tempo  $t$  tem a forma:

$$\frac{d[C]}{dt} = (\text{geração}) - (\text{consumo}). \quad (1.1)$$

A equação 1.1 é uma *equação diferencial ordinária* (EDO). Se modelarmos um conjunto de reações químicas (e.g., uma via de sinalização molecular) através de um conjunto de EDOs, cada equação retrataria a dinâmica de formação e consumo de uma espécie química e, ao resolvê-la, obter-se-ia a função que representa a variação da concentração da espécie química em questão com o passar do tempo [CNS10]. Para exemplificar, sejam R um reagente e P um produto. Uma forma de representar uma reação envolvendo estas espécies químicas pode ser descrita como:



na qual,  $k_1$  é a constante de velocidade de transformação de R em P e  $k_{-1}$  é a constante de velocidade de transformação de P em R. A reação 1.2 pode ser expressa pelo seguinte sistema de equações diferenciais:

$$\frac{d[R]}{dt} = k_{-1}[P] - k_1[R] \quad (1.3a)$$

$$\frac{d[P]}{dt} = k_1[R] - k_{-1}[P], \quad (1.3b)$$

no qual  $\frac{d[R]}{dt}$  e  $\frac{d[P]}{dt}$  são as velocidades de mudança da concentração de R e P, respectivamente, para o instante de tempo  $t$ .

A vantagem do uso de modelos baseados em EDOs é que, embora soluções analíticas não sejam garantidas, métodos numéricos para a solução de um sistema com grande número de equações

diferenciais ordinárias não-lineares são bem desenvolvidos [EI04]. No entanto, quanto maior for o sistema de EDOs simulado, mais métodos numéricos estão sujeitos a erros. Além disso, o aumento do tamanho do sistema, muitas vezes, resulta em simulações que são mais exigentes do ponto de vista computacional. Ademais, quanto maior o sistema de EDOs, maior a quantidade de simetrias de termos entre diferentes equações que o mesmo tende a ter. Uma simetria em um sistema de EDOs é a repetição em duas equações de um mesmo termo com sinal invertido [HYD05]; por exemplo, na equação 1.3 ocorrem duas simetrias. A presença de muitas simetrias em um sistema de EDOs acarreta em maior dificuldade em realizar análises de bifurcação do modelo descrito pelo sistema (por exemplo, na avaliação de modelos que possuem histerese em uma ou mais espécies químicas).

Uma tentativa de reduzir o erro numérico, o tempo de execução e o número de simetrias de sistemas de EDOs é simplificá-los através da utilização de relações de conservação de massa. Uma relação de conservação de massa é uma equação algébrica que relaciona uma concentração inicial das espécies químicas ( $t_0 = 0$ ) com os seus respectivos valores em um dado tempo  $t$ ,  $t > 0$ . Trata-se da aplicação da Lei de Conservação de Massas de Lavoisier. Por exemplo, na reação 1.2, para qualquer instante de tempo  $t > 0$ , a soma de  $[R]$  e  $[P]$  tem de ser igual à concentração total de  $R$  em  $t = 0$ . Ou seja:

$$[R_0] = [R] + [P]. \quad (1.4)$$

Dessa forma, podemos simplificar o sistema de equações representado em 1.3 para:

$$\frac{d[R]}{dt} = k_{-1}[P] - k_1[R] \quad (1.5a)$$

$$[P] = [R_0] - [R], \quad (1.5b)$$

que trata-se de um sistema de *equações algébrico-diferenciais* (EAD). Assim, por conter apenas uma equação diferencial, a integração numérica desse sistema é menos sujeita a erros e computacionalmente menos exigente. Além disso, a introdução da EAD faz com que a quantidade de simetrias caia de duas para zero.

O processo de identificação dos parâmetros de um sistema de EDOs a partir de observações experimentais requer que as concentrações de todas as espécies químicas sejam conhecidas, ao menos ao longo de um certo intervalo de tempo. Contudo, em muitos casos, não é possível medir a concentração de todas as espécies químicas consideradas (e.g., proteínas de transmembrana são particularmente difíceis de se obter medidas). Ao simplificarmos um sistema em um sistema de EADs, podemos substituir equações diferenciais por equações algébricas, eliminando assim a necessidade de efetuarmos medidas experimentais para as espécies químicas cujas respectivas equações diferenciais foram substituídas por equações algébricas.

Além disso, diferentes substituições de equações diferenciais por equações algébricas levam a diferentes sistemas de EADs, sendo que alguns podem ser mais interessantes do que outros em termos de factibilidade de produção de medidas experimentais. Por exemplo, suponhamos que possamos realizar medidas apenas para a espécie química  $P$  da reação descrita na equação 1.2. O sistema de EDOs das equações 1.3a–1.3b não serve para ser utilizado em um processo de identificação de parâmetros, pois não é factível produzir medidas experimentais para  $R$ . O sistema de EADs mostrado nas equações 1.5a–1.5b tampouco é adequado, pois a cinética de  $P$  é descrita por uma equação algébrica. Todavia, o sistema de EADS abaixo, equivalente ao das equações 1.5a–1.5b:

$$[R] = [R_0] - [P] \quad (1.6a)$$

$$\frac{d[P]}{dt} = -k_{-1}[P] + k_1[R] \quad (1.6b)$$

é adequado para este fim, pois nele a cinética de  $P$  é descrita por uma equação diferencial, enquanto que a cinética de  $R$  é extraída a partir de uma equação algébrica.

Recentemente, um arcabouço foi desenvolvido para auxiliar na construção de modelos cinéticos

de redes de sinalização molecular, permitindo que o interatoma de uma rede seja descrito através de uma lista de reações químicas, que por sua vez é mapeada em um sistema de EDOs. Esse arcabouço, chamado SigNetSim, realiza os seguintes passos para construção de um modelo cinético:

1. descrição da estrutura de rede através de uma lista de reações, suas constantes de velocidade (parâmetros do modelo) e concentrações iniciais;
2. mapeamento das reações em sistema de equações diferenciais ordinárias (EDOs);
3. simulação numérica do sistema de EDOs;
4. avaliação dos resultados por meio de análise de ajuste de curva, que poderá conduzir a uma nova simulação (passo (3)) com parâmetros ajustados.

Dados proteômicos (e.g., quantificações de experimentos cinéticos empregando Western blot) são utilizados para a análise de ajuste de curva (passo (4)). O ciclo de passos (3)–(4) é repetido até que a simulação numérica do último modelo produza resultados satisfatórios [MMF<sup>+</sup>14]. O arcabouço, implementado em Python, mapeia as listas automaticamente para sistemas de EDOs, simula numericamente e avalia os resultados [MMF<sup>+</sup>14]. Todavia, a versão atual do arcabouço SigNetSim não aborda os problemas anteriormente citados: com o aumento do modelo considerado, o aumento do erro do método numérico utilizado para simular o sistema de EDOs pode comprometer a confiabilidade da cinética simulada, além de exigir um tempo computacional que pode ser inafável para os recursos computacionais disponíveis. Ademais, a presença de muitas simetrias no sistema de EDOs pode inviabilizar a execução de uma análise de bifurcação. Portanto, é uma necessidade premente o desenvolvimento de uma metodologia para simplificar sistemas de EDOs que descrevem a cinética reações químicas, em particular a de redes de sinalização molecular.

## 1.1 Objetivos

O objetivo deste projeto de mestrado é o desenvolvimento de regras para simplificar sistemas de equações diferenciais ordinárias (EDOs) que modelam a cinética de reações químicas, utilizando para este fim relações de conservação de massa entre as espécies químicas envolvidas, que por sua vez tratam-se de equações algébricas. Além disso, no caso de sistemas que contêm reações enzimáticas, também utilizaremos, sempre que possível, reduções do sistema original de EDOs que fazem uso de descrições simplificadas da cinética do sistema (e.g. equação de Michaelis-Menten [CNS10]).

O algoritmo a ser desenvolvido, implementando as regras estabelecidas, deverá ser capaz de inspecionar o espaço de EADs possíveis, permitindo a escolha de um sistema de acordo com o critério escolhido. Dentre os critérios possíveis, consideraremos relevantes aqueles que priorizem um sistema de EADs que:

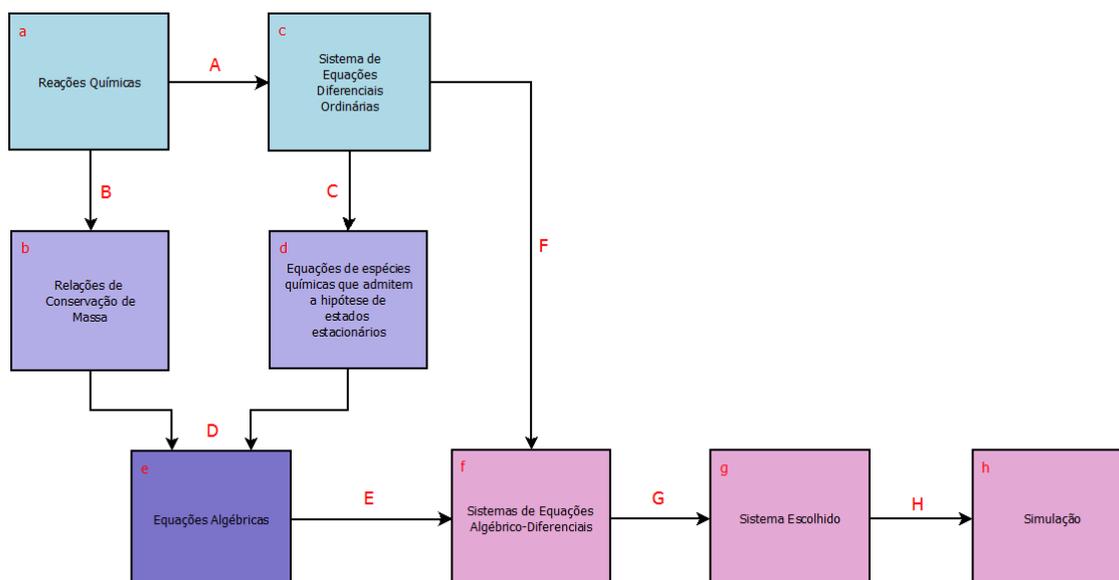
- maximize nas EDOs as espécies químicas mensuráveis;
- diminua o número de simetrias do sistema.

Tal algoritmo será testado experimentalmente através de uma implementação em linguagem de programação Python. Os testes serão feitos utilizando dados sintéticos. Finalmente, pretendemos incorporar os resultados deste trabalho no arcabouço SigNetSim como um de seus componentes.

Este projeto está estruturado em etapas, cuja organização apresentamos em um fluxo de trabalho exibido na figura 1.1.

## 1.2 Organização do Trabalho

Após esta introdução, no capítulo 2, apresentamos fundamentos teóricos nos quais este trabalho foi baseado, incluindo as condições necessárias e suficientes para a utilização da equação de Michaelis–Menten. No capítulo 3, introduzimos um algoritmo para mapear uma lista de reações



**Figura 1.1:** *fluxo de trabalho do projeto. Cada caixa representa um conjunto de dados e cada flecha uma ação sobre eles. “A” representa o algoritmo que mapeia uma lista de reações químicas para um sistema de equações diferenciais. “B” representa o algoritmo que extrai as relações de conservação de massa a partir das listas de reações químicas. “C” refere-se a simplificação empregando a equação de Michaelis-Menten. “D” é a junção das equações algébricas obtidas em “B” e “C”. “E” e “F” representam a combinação entre as equações algébricas e o sistema de equações diferenciais, formando assim, sistemas de equações diferenciais-algébricas. “G” representa um algoritmo que escolhe um sistema simplificado, dentro de um dado critério. “H” refere-se a resolução numérica do sistema.*

químicas para um sistema de equações diferenciais ordinárias (EDOs). No capítulo 4, propomos um algoritmo que extrai relações de conservação de massa entre espécies químicas, que por sua vez podem ser empregadas para simplificar o modelo cinético através do desenho de sistemas de equações algébrico-diferenciais (EADs). No capítulo 5, apresentamos uma forma para construir sistemas de EADs, assim como critérios de escolha de um sistema de EADs ótimo segundo um critério que minimize o número de termos das equações do sistema. No capítulo 6 mostraremos resultados de experimentos realizados para avaliar a metodologia proposta. Finalmente, no capítulo 7 recapitularemos as contribuições deste trabalho e apontaremos possíveis continuações para esta linha de pesquisa.

## Capítulo 2

# Fundamentos teóricos

Neste capítulo, apresentaremos alguns fundamentos teóricos que serão utilizados ao longo desta dissertação. Iniciaremos apresentando alguns conceitos fundamentais de Cinética Química; em particular, mostraremos como descrever a cinética de uma reação química através de um sistema de equações diferenciais ordinárias (EDOs). Na sequência, apresentaremos duas maneiras de simplificar um sistema de EDOs que descreve a cinética de uma reação: a primeira delas através de relações de conservação de massa entre as espécies químicas envolvidas; já a segunda pode ser empregada em algumas situações em que reações enzimáticas estão presentes, e faz uso da equação de Michaelis-Menten.

### 2.1 Cinética Química

Na natureza, reações químicas acontecem a todo momento, e cada uma delas possui uma determinada duração. Por exemplo, enquanto que dissolver uma pastilha efervescente leva alguns segundos, o processo de oxidação de um objeto feito de ferro pode levar anos. Portanto, reações químicas podem ser mais rápidas ou mais lentas, por vezes influenciadas por fatores tais como: temperatura (quanto maior a temperatura do ambiente, maior a velocidade da reação química), pressão (quanto maior a pressão de um sistema gasoso, maior a velocidade da reação), superfície de contato (quanto maior a superfície de contato maior a velocidade) e concentração de reagentes (quanto maior a concentração de substâncias reagentes, mais rapidamente ocorrerá a reação química). Este último fator refere-se à lei de ação de massas, na qual diz que a velocidade de uma reação é diretamente proporcional ao produto das concentrações molares dos reagentes, elevadas a potências determinadas pela estequiometria da reação.

O ramo da química que estuda a velocidade das reações químicas é chamado de Cinética Química. A seguir, apresentaremos alguns conceitos fundamentais dessa disciplina.

#### 2.1.1 Taxa de Reação

Considere a reação abaixo:



A expressão matemática que define a taxa de reação é dada por:

$$taxa = k \prod_{i=1}^n [R_i]^{\gamma_i}, \quad (2.2)$$

em que  $k$  é a constante de velocidade de reação,  $[R_i]$  é a concentração do reagente,  $n$  é a quantidade de reagentes e  $\gamma_i$  é a ordem da reação em relação ao reagente  $R_i$  [SAB05].

Assim, para a reação química da equação 2.1, temos a sua cinética descrita pelo seguinte sistema de EDOs:

$$\frac{d[A]}{dt} = -k[A][B] \quad (2.3a)$$

$$\frac{d[B]}{dt} = -k[A][B] \quad (2.3b)$$

$$\frac{d[C]}{dt} = k[A][B] \quad (2.3c)$$

$$\frac{d[D]}{dt} = k[A][B]. \quad (2.3d)$$

### 2.1.2 Ordem de Reação

A *ordem de uma reação* com respeito a uma dada espécie química é definida como o índice do expoente do termo concentração em uma dada taxa de reação. A seguir, descreveremos as ordens de reação mais comuns.

#### Ordem zero

Uma reação é de ordem zero quando sua velocidade é constante, independente da concentração do reagente. Considere a reação de ordem zero abaixo:



Para se obter a concentração do reagente  $A$ , deve-se partir da equação diferencial que descreve a variação da concentração de  $A$  e resolvê-la, como mostra a sequencia de operações abaixo:

$$\frac{-d[A]}{dt} = k[A]^0 \quad (2.5a)$$

$$\int_{[A]_0}^{[A]} -d[A] = \int_0^t k[A]^0 dt \quad (2.5b)$$

$$-([A] - [A]_0) = kt \quad (2.5c)$$

$$[A] = [A]_0 - kt. \quad (2.5d)$$

#### Primeira ordem

Uma reação é de primeira ordem se a sua velocidade é diretamente proporcional à concentração do reagente. Considere a reação de primeira ordem abaixo:



Os seguintes passos devem ser realizados para se obter a solução da equação diferencial:

$$\frac{-d[A]}{dt} = k[A]^1 \quad (2.7a)$$

$$\int_{[A]_0}^{[A]} \frac{-1}{[A]} d[A] = \int_0^t k dt \quad (2.7b)$$

$$-\ln[A] \Big|_{[A]_0}^{[A]} = kt \quad (2.7c)$$

$$-(\ln[A] - \ln[A]_0) = kt \quad (2.7d)$$

$$\ln\left(\frac{[A]_0}{[A]}\right) = kt \quad (2.7e)$$

$$\frac{[A]_0}{[A]} = e^{kt} \quad (2.7f)$$

$$[A] = [A]_0 e^{-kt}. \quad (2.7g)$$

### Segunda ordem

A reação de segunda ordem pode ser de dois tipos: com um ou dois reagentes. Considere a seguinte reação de segunda ordem com apenas um reagente:



Sua equação da velocidade depende do quadrado da concentração do reagente, e pode ser resolvida da seguinte forma:

$$\frac{-d[A]}{dt} = k[A]^2 \quad (2.9a)$$

$$\int_{[A]_0}^{[A]} \frac{-1}{[A]^2} d[A] = \int_0^t k dt \quad (2.9b)$$

$$\frac{1}{[A]} - \frac{1}{[A]_0} = kt \quad (2.9c)$$

$$\frac{1}{[A]} = \frac{1}{[A]_0} + kt. \quad (2.9d)$$

Agora, considere uma reação de segunda ordem com dois reagentes:



A taxa de reação é dada por:

$$\frac{-d[A]}{dt} = k[A][B] \quad (2.11a)$$

$$\int_{[A]_0}^{[A]} \frac{-1}{[A][B]} d[A] = \int_0^t k dt. \quad (2.11b)$$

Resolvendo a integral, conclui-se que:

$$\frac{1}{[A]_0 - [B]_0} \ln\left(\frac{[A][B]_0}{[B][A]_0}\right) = kt. \quad (2.12)$$

### 2.1.3 Reações Paralelas

Reações paralelas são aquelas que ocorrem ao mesmo tempo no mesmo meio (por exemplo, no caso de reações cujas espécies químicas são dissolvidas em água, aquelas que ocorrem em um mesmo recipiente), compartilhando do mesmo reagente. Considere as seguintes reações paralelas:



no qual  $k_1$  é a constante de velocidade referente à transformação de  $A$  em  $B$ , e  $k_2$  referente à transformação de  $A$  em  $C$ . Se a reação for de primeira ordem em relação ao reagente  $A$ , então a taxa de reação é dada por:

$$\frac{-d[A]}{dt} = k_1[A] + k_2[A] \quad (2.14a)$$

$$[A] = [A]_0 e^{-(k_1+k_2)t}, \quad (2.14b)$$

enquanto que a taxa em relação a  $B$  é dada por:

$$\frac{d[B]}{dt} = k_1[A] \quad (2.15a)$$

$$\frac{d[B]}{dt} = k_1[A]_0 e^{-(k_1+k_2)t} \quad (2.15b)$$

$$[B] = \frac{[A]_0 k_1}{k_1 + k_2} (1 - e^{-(k_1+k_2)t}). \quad (2.15c)$$

Analogamente, para  $C$  temos que:

$$\frac{d[C]}{dt} = k_2[A] \quad (2.16a)$$

$$[C] = \frac{[A]_0 k_2}{k_1 + k_2} (1 - e^{-(k_1+k_2)t}). \quad (2.16b)$$

### 2.1.4 Reações Sequenciais de Primeira Ordem

Duas reações são sequenciais (ou consecutivas) quando o produto de uma delas é utilizado como reagente pela outra. Considere as reações sequenciais abaixo:



nas quais  $k_1$  é a constante de velocidade referente à transformação de  $A$  em  $B$ , e  $k_2$  referente à transformação de  $B$  em  $C$ . O sistema de equações diferenciais é dado por:

$$\frac{-d[A]}{dt} = k_1[A] \quad (2.18a)$$

$$\frac{d[B]}{dt} = k_1[A] - k_2[B] \quad (2.18b)$$

$$\frac{d[C]}{dt} = k_2[B]. \quad (2.18c)$$

Resolvendo o sistema da equação 2.18, temos que a mudança de concentração de  $A$  em um instante de tempo  $t$  é dada por:

$$[A] = [A]_0 e^{-k_1 t}, \quad (2.19)$$

a de  $B$  por:

$$\frac{d[B]}{dt} = k_1[A] - k_2[B] \quad (2.20a)$$

$$\frac{d[B]}{dt} = k_1[A]_0 e^{-k_1 t} - k_2[B] \quad (2.20b)$$

$$[B] = \frac{[A]_0 k_1}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t}), \quad (2.20c)$$

e a de  $C$  por:

$$\frac{d[C]}{dt} = k_2[B] \quad (2.21a)$$

$$\frac{d[B]}{dt} = k_2 \frac{[A]_0 k_1}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t}) \quad (2.21b)$$

$$[C] = [A]_0 \left(1 + \frac{1}{k_1 - k_2} (k_2 e^{-k_1 t} - k_1 e^{-k_2 t})\right). \quad (2.21c)$$

### 2.1.5 Reações Reversíveis de Primeira Ordem

Uma reação reversível é um par de reações químicas no qual o(s) reagente(s) de uma reação é (são) produto(s) da outra e vice-versa. Considere a reação reversível abaixo:



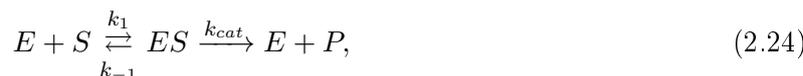
na qual  $k_1$  é a constante de velocidade referente à transformação de A em B, e  $k_{-1}$  referente à transformação de B em A. A cinética dessas espécies químicas é dada pelo seguinte sistema de EDOs:

$$\frac{-d[A]}{dt} = k_1[A] - k_2[B] \quad (2.23a)$$

$$\frac{d[B]}{dt} = k_1[A] - k_2[B]. \quad (2.23b)$$

### 2.1.6 Reação Enzimática Simples

Diversas reações bioquímicas são mediadas por um tipo especial de catalisadores biológicos conhecido como enzima. Adrian Brown, por volta de 1900, propôs que a reação enzimática global é composta por duas reações elementares, na qual o substrato forma um complexo com a enzima, e esse complexo se decompõe para os produtos e enzima [VV11]. Para exemplificar, sejam E uma enzima, S um substrato, ES um complexo enzima-substrato e P o produto. A reação enzimática que envolve essas espécies químicas pode ser descrita como:



na qual  $k_1$  é a constante de velocidade de formação do complexo enzimático,  $k_{-1}$  é a constante de velocidade da dissociação desse complexo e  $k_{cat}$  é a constante de velocidade de transformação do substrato em produto. Para descrever a variação da concentração de E, por exemplo, em um dado instante de tempo  $t$ ,  $t \geq 0$ , é necessário levar em consideração a variação da concentração de E devido à dissociação de uma dada concentração de ES em  $t$ , como também, a variação das concentrações de E e de S devido à formação do complexo em  $t$ . Assim, a velocidade de mudança

da concentração de E no instante de tempo  $t$  é dada pela seguinte equação:

$$\frac{d[E]}{dt} = -k_1[E][S] + (k_{-1} + k_{cat})[ES]. \quad (2.25)$$

O raciocínio é análogo para definir as equações que descrevem as cinéticas das demais espécies químicas:

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (2.26a)$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_{cat})[ES] \quad (2.26b)$$

$$\frac{d[P]}{dt} = k_{cat}[ES]. \quad (2.26c)$$

## 2.2 Simplificação de sistemas de EDOs utilizando relações de conservação de massa

Uma relação de conservação de massa entre reagentes e produtos é uma equação algébrica que relaciona uma concentração inicial dessas espécies químicas ( $t_0 = 0$ ) com os seus respectivos valores em um dado tempo  $t$ ,  $t > 0$ . Trata-se da aplicação da Lei de Conservação de Massas de Lavoisier. Podemos simplificar sistemas de EDOs extraindo relações de conservação de massa referente às espécies químicas.

Por exemplo, na reação 2.13, o sistema de equações diferenciais ficaria da seguinte forma:

$$\frac{-d[A]}{dt} = k_1[A] + k_2[A] \quad (2.27a)$$

$$\frac{d[B]}{dt} = k_1[A] \quad (2.27b)$$

$$\frac{d[C]}{dt} = k_2[A]. \quad (2.27c)$$

Para qualquer instante de tempo  $t > 0$ , a soma de [A], [B] e [C] tem que ser igual à concentração total de A em  $t = 0$ . Ou seja:

$$[A_0] = [A] + [B] + [C]. \quad (2.28)$$

Considerando as reações 2.14 e 2.15, temos que:

$$[A_0] = [A] + [B] + [C] \quad (2.29a)$$

$$[A_0] = [A_0]e^{-(k_1+k_2)t} + \frac{[A_0]k_1}{k_1+k_2}(1 - e^{-(k_1+k_2)t}) + [C] \quad (2.29b)$$

$$[C] = [A_0] - ([A_0]e^{-(k_1+k_2)t} + \frac{[A_0]k_1}{k_1+k_2}(1 - e^{-(k_1+k_2)t})) \quad (2.29c)$$

$$[C] = \frac{[A_0]k_2}{k_1+k_2}(1 - e^{-(k_1+k_2)t}), \quad (2.29d)$$

o que condiz com as concentrações obtidas pela resolução da equação diferencial da espécie química C. Dessa forma, podemos simplificar o sistema de equações representado em 2.27 para:

$$\frac{-d[A]}{dt} = k_1[A] + k_2[A] \quad (2.30a)$$

$$\frac{d[B]}{dt} = k_1[A] \quad (2.30b)$$

$$[A_0] = [A] + [B] + [C]. \quad (2.30c)$$

Um outro exemplo é com relação à reação enzimática 2.24. Sabemos que há conservação da quantidade inicial da enzima  $E$  e do substrato  $S$  [RA14][DHZ<sup>+</sup>09]; ou seja:

$$[E_0] = [E] + [ES] \quad (2.31a)$$

$$[S_0] = [S] + [ES] + [P]. \quad (2.31b)$$

Assim, podemos reduzir o sistema original com quatro EDOs para um sistema com duas, simplificando o modelo cinético da reação enzimática:

$$[E] = [E_0] - [ES] \quad (2.32a)$$

$$[P] = [S_0] - ([ES] + [S]) \quad (2.32b)$$

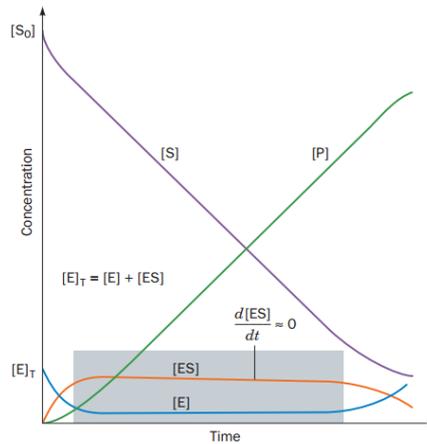
$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (2.32c)$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_{cat})[ES]. \quad (2.32d)$$

## 2.3 Simplificação de sistemas de EDOs utilizando a equação de Michaelis-Menten

Consideremos uma reação enzimática do tipo descrito pela equação 2.24. Ilustramos na figura 2.1 um exemplo da cinética das espécies químicas ao longo de uma reação desse tipo. Tipicamente, com a exceção da fase inicial da reação, a concentração do complexo enzima-substrato,  $[ES]$ , permanece aproximadamente constante até que o substrato esteja quase esgotado, caso  $[S] \gg [E]$  [RA14]. Assim, a taxa de síntese de  $ES$  deve ser igual a de consumo durante a maior parte da reação, mantendo-se em um estado estacionário [VV11]. A hipótese da existência desse estado estacionário implica que:

$$\frac{d[ES]}{dt} = 0. \quad (2.33)$$



**Figura 2.1:** Cinética das espécies químicas de uma reação enzimática. Figura extraída de Voet e Voet [VV11].

Voltando ao sistema de EDOs que descreve a cinética de uma reação enzimática (equações 2.25 e 2.26), considerando a hipótese de estado estacionário, temos que:

$$k_1([E]_0 - [ES])[S] - (k_{-1} + K_{cat})[ES] = 0 \quad (2.34)$$

$$[ES] = \frac{[E_0][S]}{K_m + [S]} \quad (2.35)$$

$$[E] = [E]_0 - [ES] \quad (2.36a)$$

$$[P] = [S]_0 - ([ES] + [S]) \quad (2.36b)$$

$$[ES] = \frac{[E]_0[S]}{K_m + [S]} \quad (2.36c)$$

$$\frac{d[S]}{dt} = \frac{-K_{cat}[E][S]}{K_m + [S]}. \quad (2.36d)$$

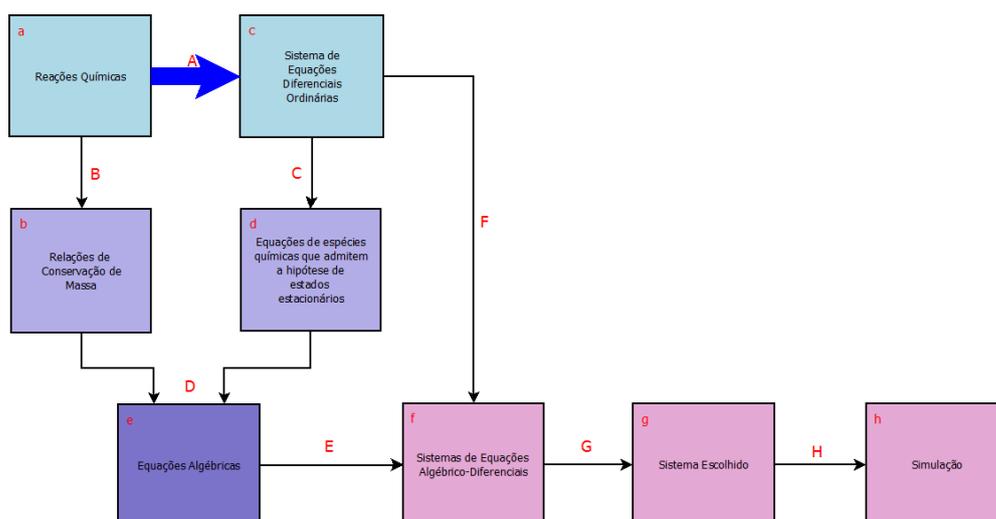
onde  $K_M = \frac{K_{-1} + K_{cat}}{K_1}$  é a constante de Michaelis, e a equação 2.36c é conhecida como equação de Michaelis-Menten.

Observe que, com a aplicação da simplificação por hipótese de estado estacionário, o sistema original com quatro equações diferenciais passa a conter apenas uma. Observe também que o sistema 2.36 usa apenas dois parâmetros ( $K_{cat}$  e  $K_m$ ). Assim, essa simplificação facilita a estimação de parâmetros e mitiga o problema de overfitting, já que diminui a dimensão dos parâmetros.

## Capítulo 3

# Mapeamento de reações químicas para sistemas de EDOs

Como vimos no capítulo anterior, a cinética de um conjunto de reações químicas pode ser descrita através de um sistema de equações diferenciais ordinárias (EDOs). Neste capítulo apresentaremos uma maneira sistemática de fazer o mapeamento de um conjunto de reações químicas em um sistema de EDOs que descreva a sua cinética. Tal sistematização será formalizada em dois algoritmos que serão descritos nas próximas duas seções. O mapeamento descrito neste capítulo faz parte da primeira etapa do fluxo de trabalho do projeto, e é destacado pela flecha azul na figura 3.1.



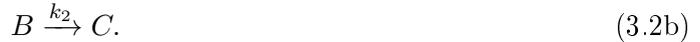
**Figura 3.1:** Mapeamento das reações químicas para sistemas de equações diferenciais ordinárias (EDOs). Essa atividade constitui a atividade “A” do projeto e está destacada pela flecha azul.

### 3.1 Geração de listas de taxas de reação (LTs)

Como vimos anteriormente, algumas reações químicas podem ser decompostas em reações mais elementares de primeira ou segunda ordem. Ao fazermos isso como um pré-processamento desta atividade, menos regras são necessárias no processo de mapeamento propriamente dito. Ou seja, caso apareça alguma reação consecutiva ou reversível, ela será transformada em duas reações. Por exemplo, dada a seguinte reação:



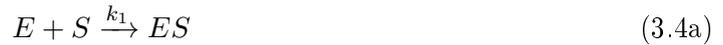
podemos transformá-la em duas:



Da mesma forma, a reação enzimática 3.3:



transforma-se em três reações diferentes, por ter uma parte reversível e outra consecutiva:



Após esse processamento, para cada reação, a taxa propriamente dita é calculada. O algoritmo 1 descreve as etapas para calcular a taxa de cada reação química.

---

**Algorithm 1** Algoritmo para obter as taxas de cada reação

---

```

1: Entrada: Lista de reações químicas (LR)
2: Saída: Lista das taxas de cada reação (LT)
3:
4: //Inicialize LT
5:  $LT \leftarrow \emptyset$ 
6: for cada reação química  $i \in LR$  do
7:   // Realize o produto dos j reagentes da reação i
8:    $rate_i = k \prod_{j=1}^n [R_j]$ 
9:   // Inclua reação na posição i de RT
10:   $LT \leftarrow LT \cup rate_i$ 
11: end for
12: Devolva  $LT$ 
13: FIM

```

---

Para o exemplo da reação enzimática 3.3, temos os seguintes pares de reações decompostas e suas respectivas taxas:



$$taxa_E = taxa_S = k_1[E][S], \quad (3.5b)$$

e também:



$$taxa_{ES_1} = k_{-1}[ES], \quad (3.6b)$$

assim como:



$$taxa_{ES_2} = k_{cat}[ES]. \quad (3.7b)$$

## 3.2 Geração do sistema de EDOs

A partir do cálculo das taxas, equações diferenciais ordinárias podem ser construídas para descrever a cinética de cada espécie química. O algoritmo 2 formaliza o processo de mapeamento, empregando para isso a lista de taxas de reações (LTs) produzida a partir da metodologia apresentada na seção anterior.

---

**Algorithm 2** Algoritmo para obter o sistema de equações diferenciais ordinárias correspondente a uma lista de reações

---

```

1: Entrada: Lista de reações químicas (LR), lista das taxas das reações (LT) e Lista de espécies químicas presentes na lista de reações (LE).
2: Saída: Sistema de equações diferenciais ordinárias (S)
3:
4: //Inicialize do sistema de equações diferenciais ordinárias
5:  $S \leftarrow \emptyset$ 
6: for cada espécie química  $X \in LE$  do
7:   // Inicialize a taxa da espécie química X
8:    $rate_x = 0$ 
9:   for cada reação química  $i \in LR$  do
10:    if X é um dos reagentes da reação i then
11:      // Subtraia a taxa da reação i da taxa da espécie química X
12:       $rate_x = rate_x - rate_i$ 
13:    else
14:      if X é um dos produtos da reação i then
15:        // Adicione a taxa da reação i da taxa da espécie química X
16:         $rate_x = rate_x + rate_i$ 
17:      end if
18:    end if
19:  end for
20:  // Inclua a taxa da reação em S
21:   $\frac{d[X]}{dt} = rate_x$ 
22:   $S \leftarrow S \cup \frac{d[X]}{dt}$ 
23: end for
24: Devolva S
25: FIM

```

---

No exemplo da reação enzimática 3.3, há quatro espécies químicas: E, S, ES e P. Utilizando o algoritmo 2, para produzir uma EDO para a espécie química E, por exemplo, soma-se a taxa da reação em que E é produto e subtrai-se a taxa da reação em que E é reagente. Assim, já que E aparece como reagente em na reação 3.5a e produto nas reações 3.6a e 3.7a, a EDO com relação a E apresenta-se como:

$$\frac{d[E]}{dt} = -k_1[E][S] + k_{-1}[ES] + k_{cat}[ES]. \quad (3.8)$$

Para as outras espécies químicas, o processo é feito de forma análoga: a espécie química S é consumida na reação 3.5a e produzida na reação 3.6a; já o complexo ES é produzido na reação 3.5a e consumido nas reações 3.6a e 3.7a. Por fim, o produto P é produzido na reação 3.7a.

## 3.3 Comentários

Neste capítulo foi apresentada uma maneira sistemática de fazer o mapeamento de um conjunto de reações químicas em um sistema de EDOs. Tal sistematização foi formalizada em dois algoritmos: o algoritmo 1, que descreve as etapas para calcular a taxa de cada reação química e o algoritmo 2,

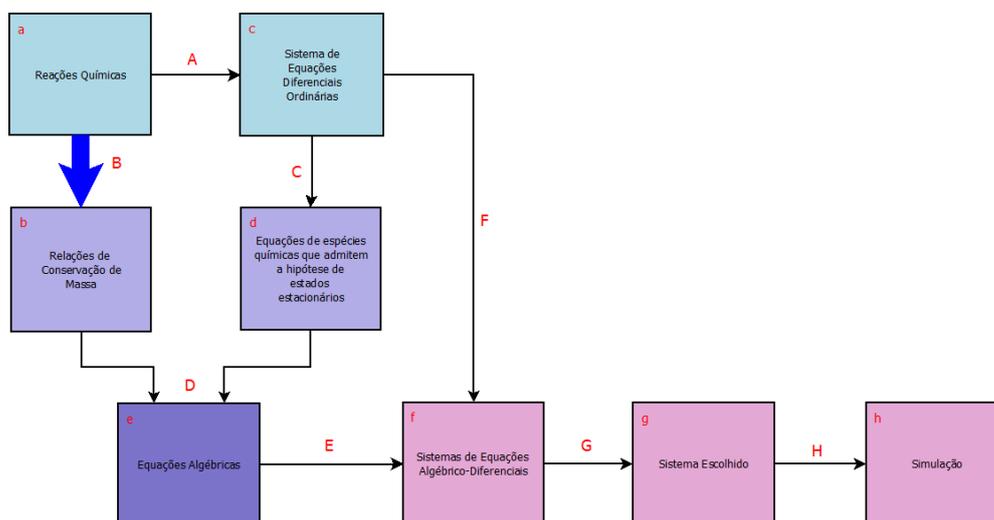
que formaliza o processo de mapeamento. Com isso, para cada lista de reação química, tem-se um sistema de equações diferenciais ordinárias que descreve sua cinética.

## Capítulo 4

# Simplificação de sistemas de EDOs

Resolver um sistema de equações diferenciais ordinárias (EDOs) pode ser uma tarefa difícil, ou mesmo inactivável. Assim, utilizamos métodos numéricos para obter, via integração numérica das EDOs do sistema, valores aproximados das concentrações das espécies químicas em um determinado tempo. Cada método numérico está associado a um erro intrínseco ao mesmo, o qual pode ser reduzido a partir de simplificações utilizando propriedades que o problema admite. Neste capítulo, apresentaremos duas propriedades que proporcionam simplificações em um sistema de EDOs que descreve a cinética de um conjunto de reações químicas. A primeira propriedade envolve a exploração de relações de conservação de massa entre as espécies químicas, cujos princípios vimos no capítulo 2.2. Já a segunda propriedade está relacionada à aplicação da hipótese de estado estacionário em reações enzimáticas, o que permite o uso da equação de Michaelis-Menten, conforme apresentado no capítulo 2.3.

### 4.1 Simplificação algébrica utilizando relações de conservação de massa



**Figura 4.1:** Extração de relações de conservação de massa a partir de um conjunto de reações químicas. A flecha azul destaca a atividade “B”, que é formalizada nesta seção utilizando dois algoritmos.

Consideremos o sistema de EDOs referentes à reação 3.3. Esse sistema pode ser simplificado com a utilização das seguintes relações de conservação de massa, como descrito no capítulo 1:

$$[E_0] = [E] + [ES] \quad (4.1a)$$

$$[S_0] = [S] + [ES] + [P], \quad (4.1b)$$

na qual a primeira relação representa a conservação da quantidade inicial de enzimas e a segunda representa a conservação da quantidade inicial de substrato. A extração de relações de conservação de massa entre espécies químicas a partir de um conjunto de reações químicas é formalizada em dois algoritmos que serão apresentados a seguir; tais algoritmos constituem a atividade destacada com uma flecha azul no fluxo de trabalho desde projeto (figura 4.1).

O algoritmo 3 formaliza uma maneira de extração de relações de conservação de massa a partir de uma lista de reações químicas. Assim, para cada espécie química que possui concentração inicial diferente de zero é produzida uma ou mais relações de conservação de massa, as quais serão adicionadas à lista total de relações.

Para extrair uma relação de conservação de massa a partir de uma espécie química, o algoritmo 3 utiliza a sub-rotina representada em 4. Esta sub-rotina requer a lista de reações, uma espécie química e a equação que representa a sua relação de conservação de massa. Esta última é inicializada atribuindo-se a concentração da espécie química à sua concentração inicial. Com esses três parâmetros de entrada, a sub-rotina verifica apenas as reações em que a espécie química participa como reagente. Se uma dessas reações contiver apenas um produto, automaticamente sua concentração é somada à relação. Caso contrário, soma-se apenas as concentrações de produtos derivados da espécie química em questão. E para cada um desses produtos, o algoritmo 4 é chamado recursivamente em busca de outros produtos derivados indiretamente da espécie química, até que não haja mais produtos para serem inseridos na relação de conservação de massa.

---

**Algorithm 3** Algoritmo para extrair as relações de conservação de massa

---

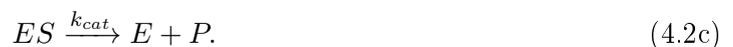
```

1: Entrada: Lista das reações químicas (LR) e Lista de espécies químicas presentes na lista de
   reações (LE).
2: Saída: Lista de relações de conservação de massa (LCM).
3:
4: //Inicialize LCM
5:  $LCM \leftarrow \emptyset$ 
6: for cada espécie química  $X \in LE$  que possui concentração inicial diferente de zero do
7:   // Modifique a variável global igualando a X
8:    $VG \leftarrow X$ 
9:   // Inicialize equação de conservação de massa
10:   $Eq \leftarrow [X_0] = [X]$ 
11:   $R \leftarrow Algoritmo4(LR, X, Eq)$ 
12:  // Inclua R em LCM
13:  if  $R \notin LCM$  then
14:     $LCM \leftarrow LCM \cup R$ ;
15:  end if
16: end for
17: Retorne LCM
18: FIM

```

---

Retornando ao exemplo da reação enzimática 3.3, consideremos a sua decomposição em reações elementares:



Suponhamos que apenas E e S possuem concentrações diferentes de zero. Assim, os mesmos terão apenas duas relações de conservação de massa a ser consideradas: primeiramente, a relação referente à E se iniciará com  $[E_0] = [E]$ . Como na reação 4.2a E aparece como reagente e há apenas um

---

**Algorithm 4** Algoritmo para extrair relações de conservação de massa para uma espécie química

---

```

1: Entrada: Lista das reações químicas (LR), Espécie química (X) , Equação de X (Eq).
2: Saída: Lista contendo as relações de conservação de massa para a espécie X (Eq).
3:
4: for cada reação  $i \in LR$  em que X estiver presente como reagente do
5:   if Reação i contiver apenas um produto P then
6:     // Acrescente em Eq a concentração de P
7:      $Eq = Eq + [P]$ 
8:     // Chame recursivamente o algoritmo para a espécie P
9:      $Eq \leftarrow Algoritmo4(LR, P, Eq)$ 
10:  else
11:    // Reação do tipo  $R \rightarrow P_1 + P_2$ 
12:    if  $P_1$  for derivado da variável global VG,  $P_2$  não for e  $P_1 \notin Eq$  then
13:      // Acrescente em Eq a concentração de  $P_1$ 
14:       $Eq = Eq + [P_1]$ 
15:      // Chame recursivamente o algoritmo para  $P_1$ 
16:       $Eq \leftarrow Algoritmo4(LR, P_1, Eq)$ 
17:      // Modifique a variável global igualando a  $P_1$ 
18:       $VG \leftarrow P_1$ 
19:    else
20:      if  $P_2$  for derivado da variável global VG,  $P_1$  não for e  $P_2 \notin Eq$  then
21:        // Acrescente em Eq a concentração de  $P_2$ 
22:         $Eq = Eq + [P_2]$ 
23:        // Modifique a variável global igualando a  $P_2$ 
24:         $VG \leftarrow P_2$ 
25:        // Chame recursivamente o algoritmo para  $P_2$ 
26:         $Eq \leftarrow Algoritmo4(LR, P_2, Eq)$ 
27:      else
28:        // Tanto  $P_1$  como  $P_2$  são derivados
29:         $Eq[1] = Eq + [P_1]$ 
30:         $Eq[2] = Eq + [P_2]$ 
31:        // Modifique a variável global igualando a  $P_1$ 
32:         $VG \leftarrow P_1$ 
33:        // Chame recursivamente o algoritmo
34:         $Eq[1] \leftarrow Algoritmo4(LR, P_1, Eq[1])$ 
35:        // Modifique a variável global igualando a  $P_2$ 
36:         $VG \leftarrow P_2$ 
37:         $Eq[2] \leftarrow Algoritmo4(LR, P_2, Eq[2])$ 
38:      end if
39:    end if
40:  end if
41: end for
42: Retorne Eq
43: FIM

```

---

produto ES, este é adicionado na relação. O algoritmo fará uma chamada recursiva para ES. Este último aparece como reagente nas reações 4.2b e 4.2c; não obstante, mas em ambas reações, E aparece como produto. Assim, nada é feito e a primeira relação obtida será:

$$[E_0] = [E] + [ES]. \quad (4.3)$$

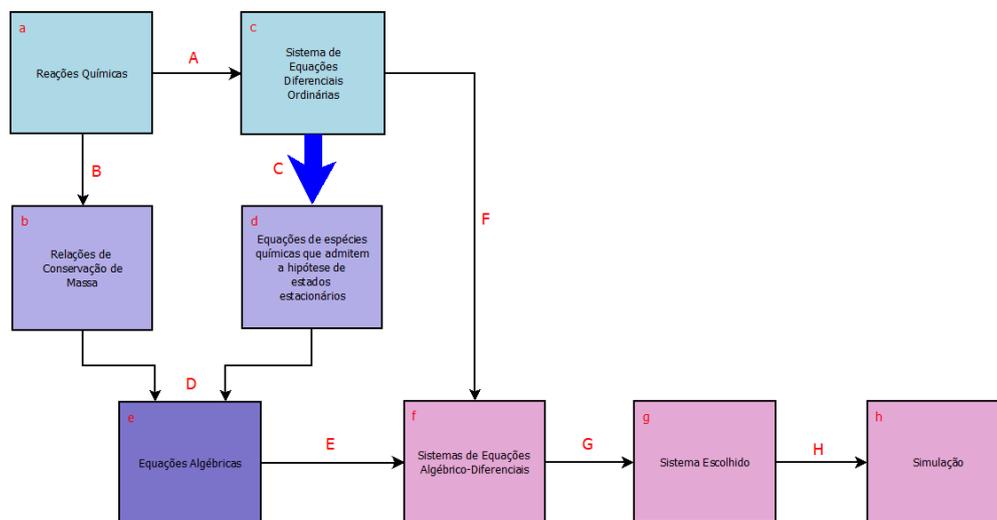
Para S, a relação se iniciará com  $[S_0] = [S]$ . Como na reação 4.2a S aparece como reagente e há apenas um produto ES, este é adicionado na relação. O algoritmo chamará recursivamente para ES. Na reação 4.2b, S aparece como produto e, portanto, nada é feito. Entretanto, na reação 4.2c, P é produto de S e, portanto, sua concentração é adicionado na relação. Assim, a segunda relação obtida será:

$$[S_0] = [S] + [ES] + [P]. \quad (4.4)$$

5

## 4.2 Simplificação utilizando hipótese de estado estacionário

Como vimos no capítulo 2.3, outro modo de simplificar os sistemas de equações diferenciais que descrevem reações químicas é, dada uma reação enzimática, admitir hipótese de estado estacionário para o complexo substrato-enzima, caso  $[S] \gg [E]$ . Dessa forma, a velocidade de mudança de concentração desse complexo é zero, o que nos permite igualar a zero a equação diferencial do mesmo. Dado um conjunto de reações químicas, o processo de determinação das reações enzimáticas que admitem estado estacionário é destacado em nosso fluxo de trabalho com uma flecha azul na figura 4.2.



**Figura 4.2:** Identificação das equações de espécies químicas que admitem estado estacionário.

No caso da reação enzimática da equação 3.3, admitir a hipótese de estado estacionário levaria à seguinte simplificação:

$$0 = -k_1[E][S] + (k_{-1} + k_{cat})[ES] \quad (4.5a)$$

$$0 = k_1[E][S] - (k_{-1} + k_{cat})[ES] \quad (4.5b)$$

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (4.5c)$$

$$\frac{d[P]}{dt} = k_{cat}[ES]. \quad (4.5d)$$

Ao juntarmos essa simplificação com a simplificação a partir de relações de conservação de massa

entre as espécies químicas, apresentada na seção anterior, teremos o seguinte sistema:

$$[E_0] = [E] + [ES] \quad (4.6a)$$

$$0 = k_1[E][S] - (k_{-1} + k_{cat})[ES] \quad (4.6b)$$

$$[S_0] = [S] + [ES] + [P]. \quad (4.6c)$$

Dessa forma, completamos o quadro de simplificações possíveis para a reação 3.3 ao empregarmos duas propriedades conhecidas do problema.

### 4.3 Comentários

Neste capítulo, foi apresentado duas propriedades que proporcionam simplificações em um sistema de EDOs. A primeira propriedade envolve a exploração de relações de conservação de massa entre as espécies químicas e a segunda propriedade está relacionada à aplicação da hipótese de estado estacionário em reações enzimáticas, o que permite o uso da equação de Michaelis-Menten. Assim, completamos o quadro de simplificações.



## Capítulo 5

# Escolha ótima de um sistema de EADs

No capítulo anterior foi apresentado um algoritmo para extrair relações de conservação de massa a partir de uma lista de reações químicas, permitindo a simplificação de sistemas de equações diferenciais ordinárias (EDOs) a partir da composição de sistemas de equações que combinam EDOs e equações algébricas; tais sistemas híbridos são conhecidos como sistemas de equações algébrico-diferenciais (EADs).

Voltemos ao exemplo da reação enzimática 3.3, que é descrita pelo seguinte sistema de EDOs:

$$\frac{d[E]}{dt} = -k_1[E][S] + (k_{-1} + k_{cat})[ES] \quad (5.1a)$$

$$\frac{d[S]}{dt} = -k_1[E][S] + k_{-1}[ES] \quad (5.1b)$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_{cat})[ES] \quad (5.1c)$$

$$\frac{d[P]}{dt} = k_{cat}[ES]. \quad (5.1d)$$

O sistema acima pode ser simplificado utilizando as seguintes relações de conservação de massa entre as espécies químicas envolvidas (utilizando, por exemplo o algoritmo 3):

$$[E_0] = [E] + [ES] \quad (5.2a)$$

$$[S_0] = [S] + [ES] + [P], \quad (5.2b)$$

Admitindo-se que  $[ES]$  mantenha um estado estacionário, temos a seguinte simplificação adicional:

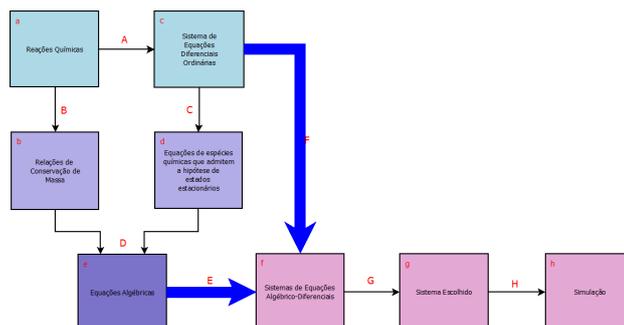
$$\frac{d[ES]}{dt} = 0 \quad (5.3a)$$

$$k_1[E][S] - (k_{-1} + k_{cat})[ES] = 0. \quad (5.3b)$$

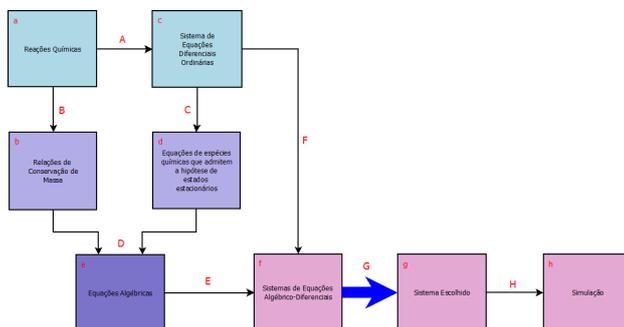
$$(5.3c)$$

Todavia, dependendo da escolha das relações de conservação de massa para simplificação do sistema de EDOs, diferentes sistemas de EADs são obtidos, e podem ter diferentes níveis de otimalidade conforme o critério adotado (e.g., número de termos das equações). Assim, neste capítulo será apresentada uma metodologia para, dado um sistema de EDOs que descreve a cinética de um conjunto de reações e também um conjunto de relações de conservação de massa entre as espécies químicas envolvidas, produzir diferentes sistemas de EADs que descrevam a cinética modelada e escolher um sistema ótimo. Em nosso fluxo de trabalho, esta metodologia cobre a atividade destacada com flechas azuis na figura 5.1.

Inicialmente, mostraremos na seção 5.1 um método para adquirir um conjunto solução do sistema linear formado pelas relações de conservação de massa entre as espécies químicas. Após isso, na seção 5.2 apresentaremos uma forma para obter diferentes sistemas de EADs. Na seção 5.3



(a) obtenção de um sistema de equações algébrico-diferenciais (EADs) para descrever a cinética de um conjunto de reações químicas; esta atividade parte de um sistema de EDOs e de relações de conservação de massa entre espécies químicas.



(b) escolha de um sistema de EADs ótimo, dentro de um critério estabelecido.

**Figura 5.1:** *Etapas deste capítulo*

apresentaremos alguns possíveis critérios de escolha de um sistema de EADs ótimo. Finalmente, na seção 5.4 será apresentado alguns algoritmos conhecidos da literatura que resolvem o problema de encontrar o sistema ótimo

## 5.1 Sistemas Lineares Indeterminados

Um sistema linear que possui menos equações que incógnitas pode não ter soluções, como por exemplo:

$$x + y + z = 1 \quad (5.4a)$$

$$2x + 2y + 2z = 3. \quad (5.4b)$$

Alternativamente, tal sistema pode apresentar infinitas soluções como:

$$x + y + z = 1 \quad (5.5a)$$

$$w + y = 3, \quad (5.5b)$$

já que seu conjunto solução pode ser representado pela tupla  $(x, y, 3 - y, 1 - y - x)$  [CDC83]. Esse tipo de sistema linear é chamado de sistema linear indeterminado, em que as incógnitas  $x$  e  $y$  são variáveis independentes e  $w$  e  $z$ , variáveis dependentes.

Nota-se que as relações de conservação de massa formam um sistema linear indeterminado com

infinitas soluções. Por exemplo, o sistema:

$$[E_0] = [E] + [ES] \quad (5.6a)$$

$$[S_0] = [S] + [ES] + [P] \quad (5.6b)$$

pode ser representado da seguinte forma:

$$([S], [ES], [E], [P]) = ([S], [ES], [E_0] - [ES], [S_0] - [ES] - [S]), \quad (5.7)$$

na qual os valores de  $[E]$  e  $[P]$  dependem das variáveis independentes  $[S]$  e  $[ES]$ .

A seguir, apresentaremos um algoritmo que devolve um conjunto solução para um sistema linear indeterminado (algoritmo 5).

Para exemplificarmos a dinâmica do algoritmo 5, considere a matriz  $S$  do sistema linear 5.6, seu vetor de incógnitas  $V$  e o seu vetor de constantes  $C$ :

$$S = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad V = \begin{bmatrix} S \\ ES \\ P \\ E \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} S_0 \\ E_0 \end{bmatrix}.$$

A matriz  $M_1$  conterá apenas '1' nas posições referentes às variáveis dependentes. Vamos supor que  $[E]$  e  $[P]$  são as variáveis dependentes como em 5.7. Assim, a matriz  $M_1$  será:

$$M_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (5.8)$$

Como  $M_2 = S - M_1$ , então temos que:

$$M_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad (5.9)$$

$$M_1 * V = C - M_2 * V \quad (5.10a)$$

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} S \\ ES \\ P \\ E \end{bmatrix} = \begin{bmatrix} S_0 \\ E_0 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} * \begin{bmatrix} S \\ ES \\ P \\ E \end{bmatrix}. \quad (5.10b)$$

O algoritmo 5 calcula então  $V_2$  e  $V_3$  para resolver o sistema linear:

$$V_2 = M_1 * V \quad (5.11)$$

$$= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} S \\ ES \\ P \\ E \end{bmatrix} \quad (5.12)$$

$$= \begin{bmatrix} P \\ E \end{bmatrix}, \quad (5.13)$$

**Algorithm 5** Algoritmo para obter um conjunto solução para um sistema linear indeterminado

---

```

1: Entrada: Matriz do sistema linear indeterminado ( $S$ ), Número de equações ( $n$ ), Número de incógnitas ( $m$ ), Vetor de incógnitas ( $V$ ) e Vetor de Constantes ( $C$ ), Lista de variáveis dependentes ( $L$ ).
2: Saída: Tupla que representa o conjunto solução.
3:
4: //Considere  $T$  uma tupla vazia de tamanho  $m$ 
5:  $T \leftarrow \emptyset$ 
6: //Considere  $M_1$  uma matriz vazia de tamanho  $n \times m$ .
7:  $M_1 \leftarrow \emptyset$ 
8: //Considere  $M_2$  uma matriz vazia de tamanho  $n \times m$ .
9:  $M_2 \leftarrow \emptyset$ 
10:
11: for  $i = 0$  até  $n - 1$  do
12:   for  $j = 0$  até  $m - 1$  do
13:     // Se a variável  $V[j]$  for uma variável dependente
14:     if  $S[i][j] == 1$  e  $V[j] == L[i]$  then
15:        $M_1[i][j] = 1$ 
16:     else
17:        $M_1[i][j] = 0$ 
18:     end if
19:   end for
20: end for
21:
22: //Resolva a operação com matrizes
23:  $M_2 = S - M_1$ 
24:  $V_2 = M_1 * V$ 
25:  $V_3 = C - M_2 * V$ 
26:
27: Obtenha o sistema linear SL, a partir de  $V_2$  e  $V_3$ , com as variáveis dependentes em função das variáveis independentes.
28:
29: //  $T$  recebe as variáveis independentes
30: for  $i = 0$  até  $m - 1$  do
31:   if  $V[i]$  não pertencer a  $L$  then
32:      $T \leftarrow T + V[i]$ 
33:   end if
34: end for
35: // $T$  recebe as variáveis dependentes
36:  $T \leftarrow T + SL$ 
37: Devolva  $T$ 
38: FIM

```

---

$$V_3 = C - M_2 * V \quad (5.14)$$

$$= \begin{bmatrix} S_0 \\ E_0 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} * \begin{bmatrix} S \\ ES \\ P \\ E \end{bmatrix} \quad (5.15)$$

$$= \begin{bmatrix} S_0 - S - ES \\ E_0 - ES \end{bmatrix}, \quad (5.16)$$

$$V_2 = V_3 \quad (5.17)$$

$$\begin{bmatrix} P \\ E \end{bmatrix} = \begin{bmatrix} S_0 - S - ES \\ E_0 - ES \end{bmatrix}. \quad (5.18)$$

E, a partir de  $V_2$  e  $V_3$ , obtém-se o sistema linear com as variáveis dependentes em função das variáveis independentes.

$$[P] = [S_0] - [S] - [ES] \quad (5.19a)$$

$$[E] = [E_0] - [ES], \quad (5.19b)$$

e  $T = ([S], [ES], [E]_0 - [ES], [S]_0 - [ES] - [S])$  será devolvido.

Agora que o conjunto solução do sistema linear indeterminado foi obtido, o próximo passo é combiná-lo com as equações diferenciais ordinárias.

## 5.2 Composição de sistemas de EADs

A combinação das equações para formar um sistema de equações algébrico-diferenciais será feita substituindo as variáveis independentes do conjunto solução dos sistema indeterminado pelas suas respectivas EDOs. Para exemplificar, consideremos a reação enzimática 3.3; suponhamos que as variáveis  $[E]$  e  $[P]$  são as variáveis dependentes do sistema linear. Assim, o conjunto solução será:

$$T = ([S], [ES], [E_0] - [ES], [S_0] - [ES] - [S]). \quad (5.20)$$

Ao substituirmos as variáveis independentes pelas suas respectivas EDOs, forma-se um sistema de equações algébrico-diferenciais (EADs), o qual é mais simples que o sistema original de quatro EDOs referente à reação 3.3:

$$[E] = [E_0] - [ES] \quad (5.21a)$$

$$[P] = [S_0] - ([ES] + [S]) \quad (5.21b)$$

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (5.21c)$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_{cat})[ES]. \quad (5.21d)$$

Há uma outra simplificação que pode ser realizada no sistema 5.21, que é quando o complexo  $[ES]$  admitir a hipótese de estado estacionário. Essa simplificação adicional está representada no sistema

5.22.

$$[E] = [E_0] - [ES] \quad (5.22a)$$

$$[P] = [S_0] - ([ES] + [S]) \quad (5.22b)$$

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (5.22c)$$

$$0 = k_1[E][S] - (k_{-1} + k_{cat})[ES]. \quad (5.22d)$$

$$(5.22e)$$

Realizando operações algébricas, chega-se no sistema 5.23:

$$[E] = [E_0] - [ES] \quad (5.23a)$$

$$[P] = [S_0] - ([ES] + [S]) \quad (5.23b)$$

$$[ES] = \frac{[E]_0[S]}{K_m + [S]} \quad (5.23c)$$

$$\frac{d[S]}{dt} = k_1([E]_0 - [ES])[S] + k_{-1}[ES], \quad (5.23d)$$

no qual  $K_M = \frac{K_{-1} + K_{cat}}{K_1}$ .

Observe que existem várias combinações possíveis na escolha das variáveis dependentes nas equações algébricas. Por exemplo, para o sistema 5.6, as possíveis escolhas de variáveis na primeira e na segunda equação, respectivamente, são:  $([ES],[S]);([ES],[P]);([E],[ES]);([E],[S])$  e  $([E],[P])$ . Dessa forma, para se obter todas as possíveis combinações de equações algébricas com as equações diferenciais ordinárias, além de identificar as espécies químicas que admitem a hipótese de estado estacionário, construímos o algoritmo 6.

---

**Algorithm 6** Algoritmo para obter os sistemas de equações diferenciais-algébricas

---

- 1: **Entrada:** Quantidade de espécies químicas(n), Lista de espécies químicas que admitem estado estacionário(LEE).
  - 2: **Saída:** Lista de tuplas, os quais representam sistemas de equações diferenciais-algébricas (LS)
  - 3:
  - 4: //Inicialize LS
  - 5:  $LS \leftarrow \emptyset$
  - 6: //Inicialize LTEE (Listas de tuplas as quais admitem estado estacionário)
  - 7:  $LS \leftarrow \emptyset$
  - 8: //Inicialize X = tupla de tamanho n contendo apenas "1"
  - 9:  $X \leftarrow [1, 1, \dots, 1]$
  - 10: //Inicialize Y = tupla de tamanho n contendo apenas "0"
  - 11:  $Y \leftarrow [0, 0, \dots, 0]$
  - 12: *Algoritmo7*(n, X, Y, LS, LEE, LTEE)
  - 13: Devolva LS
  - 14: FIM
- 

A ideia desse algoritmo é utilizar uma árvore T como espaço de busca, a qual é um sub-grafo do diagrama de Hasse do Reticulado Booleano. Cada nó dessa árvore representa um sistema e é formado por uma tupla, sendo que cada posição representa uma espécie química. A figura 5.2 representa um nó da árvore de busca referente ao exemplo 3.3.

Neste caso, a primeira posição representa o substrato, a segunda, a enzima, a terceira, o complexo enzima-substrato e a última o produto. O número zero na posição significa que, para o elemento correspondente, está sendo utilizado sua EDO no sistema. Assim, no exemplo, como todas as posições contem zero, o sistema representado pelo nó não possui nenhuma simplificação.

Já a figura 5.3 representa um outro sistema, para o mesmo exemplo, contendo uma equação

---

**Algorithm 7** Algoritmo para obter os sistemas de equações diferenciais-algébricas

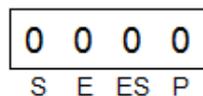
---

```

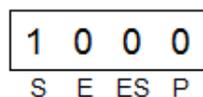
1: Entrada:  $n, X, Y, LS, LEE, LTEE$ .
2: Saída: Lista de tuplas, os quais representam sistemas de equações diferenciais-algébricas (LS)
3: if  $Y$  não pertence à  $LS$  then
4:   //Adicione  $Y$  a  $LS$ 
5:    $LS \leftarrow LS + Y$ 
6: end if
7: if  $X$  é diferente de vazio then
8:   //Retire um elemento de  $X$ 
9:    $X \leftarrow X - 1$ 
10: for  $i$  de índice do último elemento 1 de  $Y$  até  $n$  do
11:    $Y[i + 1] \leftarrow 1$ 
12:   // Se a tupla  $Y$  representa um sistema possível determinado, gera outros tuplas derivadas
   de  $Y$ 
13:   if  $Y$  forma sistema possível determinado then
14:      $Algoritmo7(n, X, Y, LS, LEE)$ 
15:   end if
16:   for  $j$  de 0 à  $n$  do
17:     if  $Y[j] = 0$  a espécie química que  $Y[j]$  representa está contido em  $LEE$  then
18:       if  $Y$  com  $\frac{d[ES]}{dt} = 0$  forma sistema possível determinado then
19:         // Adiciona  $Y$  à listas de tuplas as quais admitem estado estacionário
20:          $LTEE \leftarrow LTEE + Y$ 
21:       end if
22:     end if
23:   end for
24: end for
25: end if

```

---

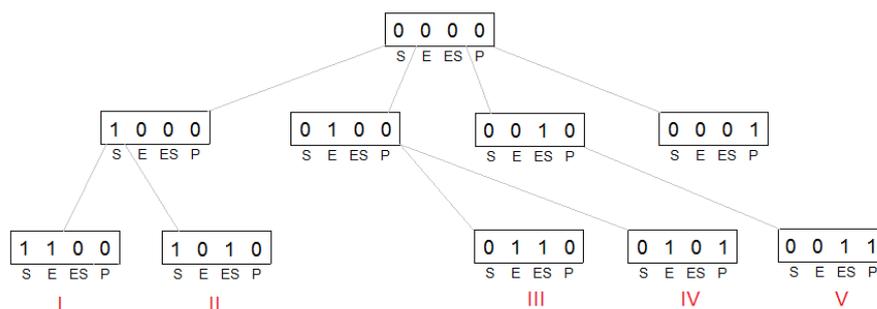


**Figura 5.2:** Nó, referente ao exemplo 3.3, que representa o sistema sem nenhuma simplificação.



**Figura 5.3:** Nó, referente ao exemplo 3.3, que representa o sistema que utiliza uma equação algébrica no lugar da EDO do elemento "S".

algébrica para a espécie química 'S'. Assim, o sistema é uma EAD, já que possui três EDOs e uma EA. Este sistema, representado pelo array "Y" no algoritmo, é gerado a partir do nó 5.2 com inserção de "uns" em determinadas posições. Portanto, este algoritmo gera todas as combinações de sistemas que possuem solução viável. A figura 5.4 apresenta todos os sistemas possíveis para a reação 3.3.



**Figura 5.4:** Árvore contendo todos os sistemas possíveis para a reação 3.3.

Este algoritmo também identifica sistemas que permitem a simplificação utilizando a equação de Michaelis-Menten. Caso um nó possua elementos que são complexos de enzima-substrato e que permitem estado estacionário, o algoritmo gera essa tupla, armazenando em uma lista.

Para exemplificar o mapeamento dos nós gerados para os sistemas, consideremos a reação enzimática 3.3. Ao aplicarmos o algoritmo 6 utilizando essa reação como entrada obtemos os seguintes sistemas de EADs:

$$[ES] = [E_0] - [E] \quad (5.24a)$$

$$[S] = [S_0] - ([ES] + [P]) \quad (5.24b)$$

$$\frac{d[E]}{dt} = -k_1[E][S] + (k_{-1} + k_{cat})[ES] \quad (5.24c)$$

$$\frac{d[P]}{dt} = k_{cat}[ES], \quad (5.24d)$$

$$[ES] = [E_0] - [E] \quad (5.25a)$$

$$[P] = [S_0] - ([ES] + [S]) \quad (5.25b)$$

$$\frac{d[E]}{dt} = -k_1[E][S] + (k_{-1} + k_{cat})[ES] \quad (5.25c)$$

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES], \quad (5.25d)$$

$$[E] = [E_0] - [ES] \quad (5.26a)$$

$$[ES] = [S_0] - ([S] + [P]) \quad (5.26b)$$

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (5.26c)$$

$$\frac{d[P]}{dt} = k_{cat}[ES], \quad (5.26d)$$

$$[E] = [E_0] - [ES] \quad (5.27a)$$

$$[S] = [S_0] - ([ES] + [P]) \quad (5.27b)$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_{cat})[ES]. \quad (5.27c)$$

$$\frac{d[P]}{dt} = k_{cat}[ES] \quad e \quad (5.27d)$$

$$[E] = [E_0] - [ES] \quad (5.28a)$$

$$[S] = [S_0] - ([ES] + [P]) \quad (5.28b)$$

$$[ES] = \frac{[E]_0[S]}{K_m + [S]} \quad (5.28c)$$

$$\frac{d[P]}{dt} = k_{cat}[ES] \quad e \quad (5.28d)$$

$$[E] = [E_0] - [ES] \quad (5.29a)$$

$$[P] = [S_0] - ([ES] + [S]) \quad (5.29b)$$

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (5.29c)$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_{cat})[ES]. \quad (5.29d)$$

$$[E] = [E_0] - [ES] \quad (5.30a)$$

$$[P] = [S_0] - ([ES] + [S]) \quad (5.30b)$$

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (5.30c)$$

$$[ES] = \frac{[E]_0[S]}{K_m + [S]}. \quad (5.30d)$$

Assim, o sistema 5.24 é representado pela nó II, o sistema 5.25 por V, o sistema 5.26 por III, o sistema 5.27 por I e o sistema 5.29 por IV. Os sistemas 5.28 e 5.30 também são gerados e são armazenados na lista de tuplas que admitem estado estacionário, já que permitem que a  $\frac{d[ES]}{dt}$  seja igualada à zero.

Dentre os sistemas produzidos pelo algoritmo 6 e que foram apresentados nas equações 5.24–5.29, alguns podem ser melhores do que outros, dependendo do critério escolhido. Na próxima seção, discutiremos possíveis critérios que poderiam ser adotados para escolher um sistema de EADs que seja “ótimo”.

### 5.3 Critérios para busca de um sistema de EADs ótimo

Como visto na seção anterior, há várias combinações possíveis para a formação de sistemas de EADs. Por exemplo, para a reação 3.3 existem cinco possíveis sistemas de EADs para descrever a sua cinética.

Dentre os vários critérios possíveis para definir a otimalidade de um sistema de EADs, inicialmente consideraremos dois deles:

1. número de equações diferenciais;
2. número de termos das equações diferenciais.

Com esses critérios, esperamos minimizar o número de simetrias existentes no sistema de EADs a ser escolhido.

#### Número de EDOs como critério de otimização

Este critério refere-se à quantidade de equações diferenciais ordinárias presentes no sistema de equações algébrico-diferenciais. Por exemplo, utilizando este critério para avaliar os cinco sistemas possíveis da reação 3.3, apenas os sistemas 5.27 e 5.29 seriam escolhidos, já que, diferentemente dos outros, possuem apenas uma equação diferencial.

#### Número de termos das EDOs como critério de otimização

Este critério refere-se ao número de termos presentes nas equações diferenciais do sistema de equações algébrico-diferenciais. Ele será utilizado em caso de empate, como ocorreu no exemplo da reação 3.3. Assim, por esse critério, o sistema escolhido para a simulação seria o sistema 5.27, pois  $\frac{d[P]}{dt}$  possui menos termos que  $\frac{d[S]}{dt}$ , presente no sistema 5.29.

Além dos dois critérios citados acima, também é possível utilizar outro critério, o qual retorna um sistema que possui a solução mais factível do ponto de vista de medida. Isso é importante porque, a identificação dos sistemas, ou inferência dos parâmetros, requer que a concentração de todas as espécies químicas seja conhecida, ao menos em um certo intervalo de tempo. Entretanto, em muitos casos, não é possível medir a concentração de todas as espécies químicas consideradas. Como foi apresentado um número combinatório de possibilidades de representar equivalentemente o sistema de equações diferenciais completos em termos de um conjunto menor de equações diferenciais e equações de conservação de massa, basta escolher o sistema que maximize a abrangência das concentrações conhecidas nas equações diferenciais, sabendo que a concentração das demais componentes pode ser calculada a partir das equações de conservação de massa. Assim, utilizando esse novo critério, dá-se a possibilidade do biólogo escolher a solução mais factível do ponto de vista de medida.

## 5.4 Algoritmos para a escolha de um sistema ótimo

### 5.4.1 Algoritmo Branch and Bound

O Algoritmo Branch and Bound é uma abordagem que garante solução ótima, o qual ramifica o nó até que o custo do elemento comece a subir. Caso isso ocorra, a árvore é podada, descartando o ramo derivado do nó em questão. Para o caso da busca de um sistema ótimo, cada nó da árvore é uma tupla que representa um sistema de EADs e o custo é calculado pelo número de EAs no sistema. Caso não seja possível resolver um sistema, este é descartado juntamente com suas tuplas "filhas". Ao final, o algoritmo devolve uma lista contendo os sistemas ótimos locais e estes serão analisados quanto à quantidade de termos. O sistema que contiver o menor número de termos é o sistema ótimo. Este algoritmo está descrito em 8.

---

#### Algorithm 8 Branch and Bound

---

```

1: Saída: Tupla R representando sistema ótimo
2:
3: //Inicialize X = tupla de tamanho n contendo apenas "1"
4:  $X \leftarrow [1, 1, \dots, 1]$ 
5: //Inicialize Y = tupla de tamanho n contendo apenas "0"
6:  $Y \leftarrow [0, 0, \dots, 0]$ 
7: //Inicialize L: lista com tuplas sub-ótimas
8:  $L \leftarrow \emptyset$ 
9:
10: //Lista contendo os sistemas mínimos globais
11: Algoritmo9(X, Y, L)
12:  $R \leftarrow x \in L : c(x) \text{mínimo}$ 
13: Devolva R
14: FIM

```

---



---

#### Algorithm 9 Algoritmo Branch and Bound

---

```

1: Entrada: Tupla X com n posições, Tupla Y com n posições, L.
2: Saída: Lista L contendo tuplas que representam sistemas sub-ótimo
3:
4:  $L \leftarrow L \cup Y$ 
5:
6: while  $X \neq \emptyset$  do
7:   //Remova um elemento de X
8:    $X \leftarrow X - x$ 
9:   //Adicione em Y'
10:   $Y' \leftarrow Y \cup x$ 
11:   $cY' \leftarrow cost(Y')$ 
12:
13:  if  $cY < cOptimalY$  and optimalY é um sistema possível then
14:     $N \leftarrow \text{Algoritmo9}(X, \text{optimalY})$ 
15:  end if
16: end while
17:
18: Algoritmo9(X, optimalY, L)
19: FIM

```

---

A figura 5.5 e 5.6 apresentam a simulação do algoritmo descrito acima para a reação 3.3 desconsiderando, a princípio, a simplificação de Michaelis-Menten.

### 5.4.2 Algoritmo Sequential Forward Search

SFS ou Busca Sequencial para Frente é uma heurística (algoritmo subótimo) que percorre o espaço de busca de forma incremental: isto, é, ele começa com um sistema com zero EAs, e inicialmente escolhe um sistema com uma EA que minimiza o número de termos; em seguida, escolhe outro sistema derivado do anterior, o qual possui outra EA que minimiza o número de termos e assim sucessivamente até o ponto em que novas sistemas não podem mais ser escolhidos. Este raciocínio é apresentado pelo algoritmo 10.

Note que o a função custo utilizado neste algoritmo é calculado pela quantidade de termos que o sistema possui, além da quantidade de EADs.

---

#### Algorithm 10 SFS

---

```

1: Saída: Tupla R representando sistema sub-ótimo
2:
3: //Inicialize X = tupla de tamanho n contendo apenas "1"
4:  $X \leftarrow [1, 1, \dots, 1]$ 
5: //Inicialize Y = tupla de tamanho n contendo apenas "0"
6:  $Y \leftarrow [0, 0, \dots, 0]$ 
7:
8:  $R \leftarrow \text{Algoritmo11}(X, Y)$ 
9: Devolva R
10: FIM

```

---



---

#### Algorithm 11 Algoritmo SFS

---

```

1: Entrada: Tupla X com n posições, Tupla Y com n posições.
2: Saída: Tupla R representando sistema sub-ótimo
3:
4:  $cY \leftarrow \text{cost}(Y)$ 
5:  $\text{minCost} \leftarrow cY$ 
6:  $\text{optimalY} \leftarrow Y$ 
7:
8: while  $X \neq \emptyset$  do
9:   //Remova um elemento de X
10:   $X \leftarrow X - x$ 
11:  //Adicione em Y'
12:   $Y' \leftarrow Y \cup x$ 
13:   $cY' \leftarrow \text{cost}(Y')$ 
14:
15:  if  $cY < c\text{optimalY}$  and  $\text{optimalY}$  é um sistema possível then
16:     $\text{minCost} \leftarrow cY'$ 
17:     $\text{optimalY} \leftarrow Y'$ 
18:  end if
19: end while
20:
21:  $R \leftarrow \text{Algoritmo11}(X, \text{optimalY})$ 
22: Devolva R
23: FIM

```

---

A figura 5.7 apresenta a simulação do algoritmo descrito acima para a reação 3.3. Coincidentemente, o algoritmo retorna o mesmo sistema ótimo que o algoritmo Branch and Bound. Entretanto, isso nem sempre pode acontecer. A figura 5.8 apresenta outro resultado possível caso o algoritmo ramificasse outro nó que possui o mesmo custo da simulação apresentada em 5.7.

### 5.4.3 Percorrimento de uma única cadeia

Foi considerado também um algoritmo que utiliza apenas um ramo da árvore de combinações. Para isso, substituí a EDO pela relação de conservação de massa do elemento que gerou a relação. Por exemplo:

Considere as seguintes relações de conservação de massa para a reação 3.3:

$$[S_0] = [S] + [ES] + [P] \quad (5.31a)$$

$$[E_0] = [E] + [ES] \quad (5.31b)$$

Este algoritmo substituí as EDOs dos elementos 'S' e 'E', formando o seguinte sistema representado pela tupla  $T = [1100]$ :

$$[E] = [E_0] - [ES] \quad (5.32a)$$

$$[S] = [S_0] - ([ES] + [P]) \quad (5.32b)$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_{cat})[ES]. \quad (5.32c)$$

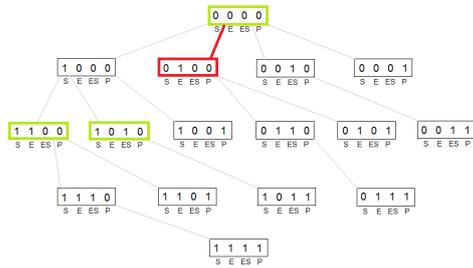
$$\frac{d[P]}{dt} = k_{cat}[ES] \quad (5.32d)$$

## 5.5 Comentários

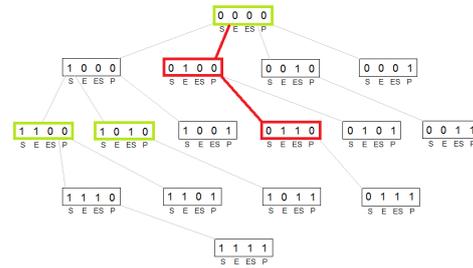
Neste capítulo, foi apresentado uma metodologia para, dado um sistema de EDOs e um conjunto de relações de conservação de massa entre as espécies químicas envolvidas, produzir diferentes sistemas de EADs que descrevam a cinética modelada. Foram apresentados também critérios pra escolha de apenas um sistema de EADs e algoritmos da literatura para que o faça.

No próximo capítulo resolvemos o problema de otimização aqui apresentado em experimentos que utilizaram os algoritmos citados.

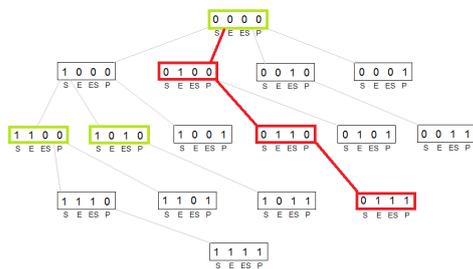




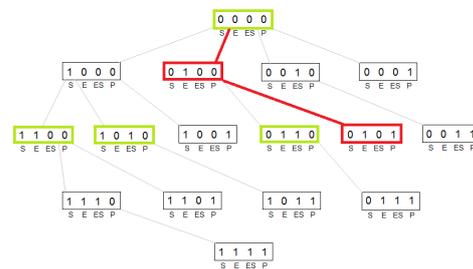
(a) Mínimo local é definido e a busca continua nos próximos ramos.



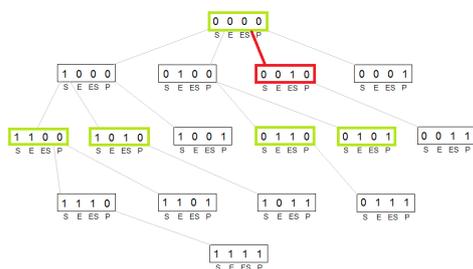
(b) Tupla visitada possui custo menor que o seu pai, já que contém duas equação algébrica.



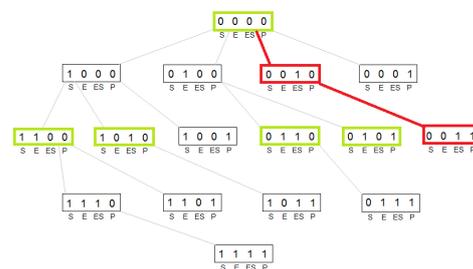
(c) Tupla visitada representa um sistema indeterminado. Ramo é descartado.



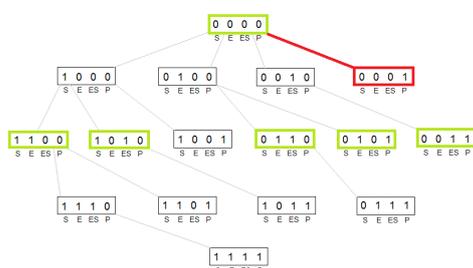
(d) [Mínimo local é definido e a busca continua nos próximos ramos.



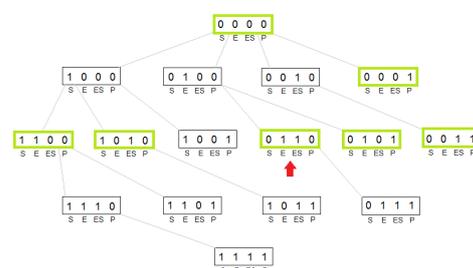
(e) Mínimo local é definido e a busca continua nos próximos ramos.



(f) Tupla visitada possui custo menor que o seu pai, já que contém duas equação algébrica.

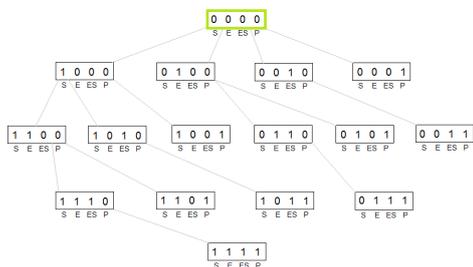


(g) Mínimo local é definido e a busca continua nos próximos ramos.

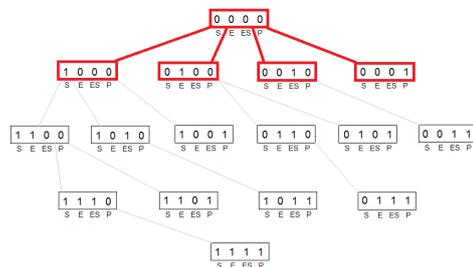


(h) Árvore de busca é esgotada. As tuplas em verde representam o mínimos locais. A tupla indicada representa o sistema ótimo, já que possui menor quantidade de termos.

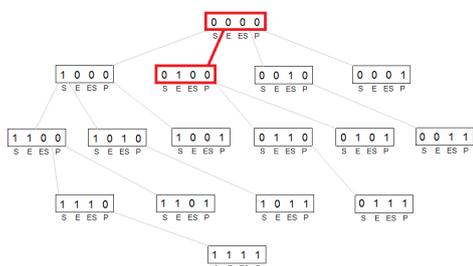
Figura 5.6: Segunda parte da simulação do algoritmo Branch and Bound para a reação 3.3



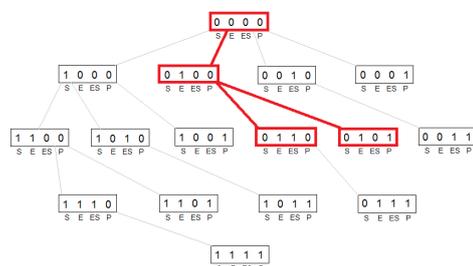
(a) Algoritmo inicia com um sistema sem nenhuma simplificação.



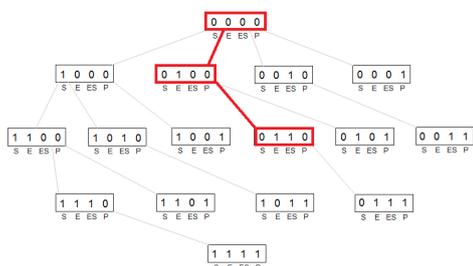
(b) O custo é calculado para todas as tuplas filhas.



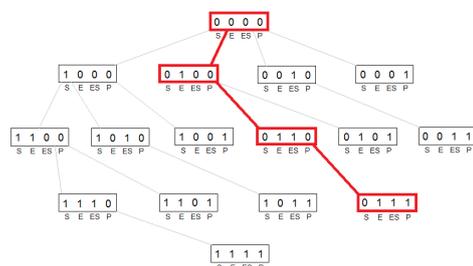
(c) A tupla que minimiza o número de termos do sistema é escolhido.



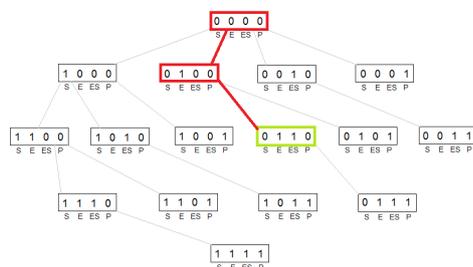
(d) A tupla que minimiza o número de termos do sistema é escolhido.



(e) A tupla que minimiza o número de termos do sistema é escolhido.

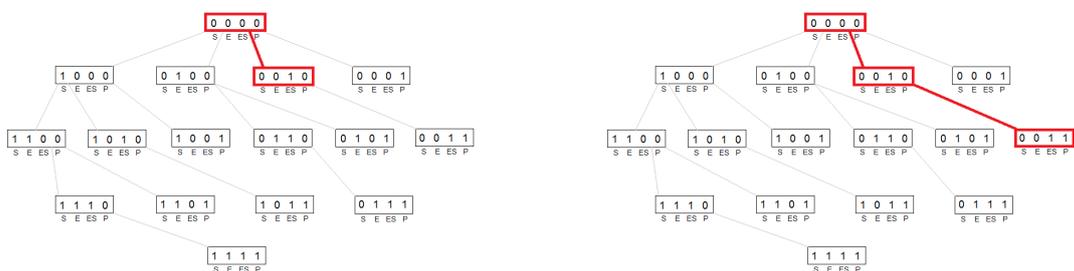


(f) [Tupla visitada representa um sistema indeterminado. Ramo é descartado.



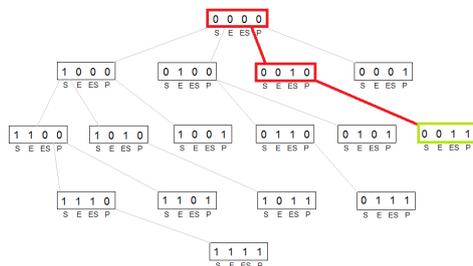
(g) Tupla escolhida.

Figura 5.7: Simulação do algoritmo SFS para a reação 3.3



(a) Outro nó com o mesmo custo da figura 5.7 é escolhido.

(b) O custo é calculado a única tupla filha.



(c) Tupla escolhida.

**Figura 5.8:** Outro resultado possível para a simulação do algoritmo SFS para a reação 3.3.



# Capítulo 6

## Experimentos

No capítulo anterior, foi apresentada uma metodologia para, dado um sistema de EDOs e um conjunto de relações de conservação de massa entre as espécies químicas envolvidas, produzir diferentes sistemas de EADs que descrevam a cinética modelada. Foram apresentados também dois critérios pra escolha de um sistema de EADs para ser simulado: quantidade de EDOS e quantidade de termos. Nesse capítulo, serão detalhados os experimentos realizados, juntamente com os respectivos resultados das simulações. Será apresentado também um algoritmo de geração de reações químicas e seus parâmetros, o qual foi utilizado para as simulações.

Assim, este capítulo compreende a fase final do fluxo de trabalho, apresentado na figura 6.1

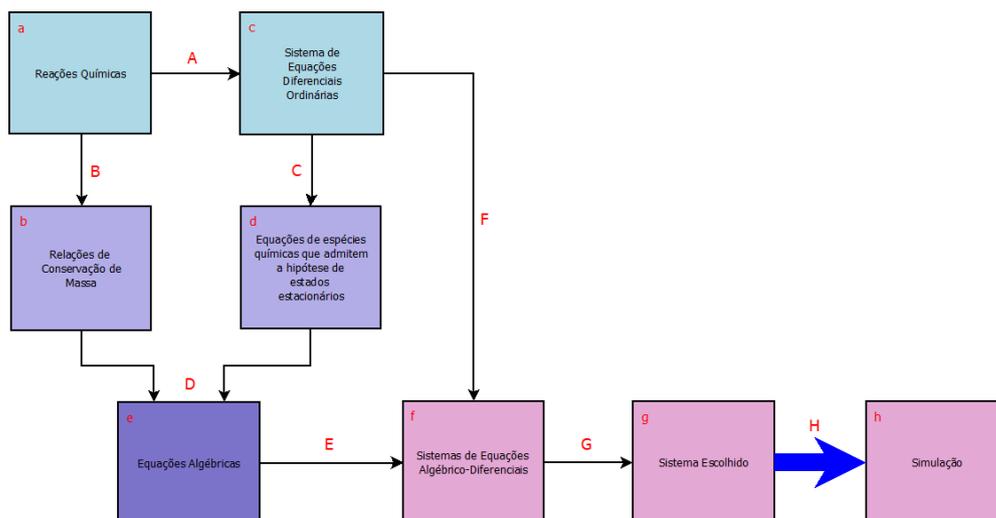


Figura 6.1: Simulação do sistema escolhido.

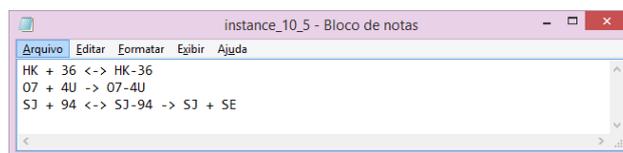
## 6.1 Experimentos

### 6.1.1 Algoritmo para geração de dados sintéticos

Na seção anterior, foram apresentados algoritmos com o objetivo de encontrar um sistema ótimo para uma determinada reação. Para que eles sejam testados, é necessário que haja uma base de dados de reações, com suas respectivas velocidades e concentrações iniciais. Para isso, foi implementado um algoritmo que gera dados sintéticos possibilitando a realização das simulações. Os dados gerados estão organizados em três tipos de arquivos: Reações, Taxas de Reações e Concentrações Iniciais.

O arquivo de reações contém as reações descritas no capítulo 2. Para a sua construção, o algoritmo recebe como parâmetro de entrada a quantidade de espécies químicas desejada e sorteia várias reações de diferentes tipos até que o número de espécies químicas utilizadas ultrapasse o valor do parâmetro. O nome de cada espécie química é gerado com dois caracteres alfanuméricos sorteados.

Caso a espécie química seja um reagente, o algoritmo permite que esta tenha o mesmo nome de um produto de reações anteriores em vez de de sortear um novo nome para ela. A figura 6.2 exemplifica um arquivo de reações com 10 espécies químicas.



**Figura 6.2:** Exemplo de um arquivo contendo reações que utilizam 10 espécies químicas no total.

Da mesma forma, são sorteadas as velocidades para cada reação e as concentrações iniciais para cada espécie química.

Para a realização das simulações, foram geradas 10 instâncias de cada tipo de arquivo para cada valor de 'n', o qual começa com valor 10 e varia de 10 em 10 até chegar a 100. Ou seja, para 'n' igual à 10, há 10 instâncias de reações juntamente com as respectivas 10 instâncias de taxas e 10 instâncias de concentrações iniciais. O mesmo vale para 'n' igual à 20, 'n' igual à 30 e assim por diante, até que 'n' seja igual à 100.

### 6.1.2 Simulação

Todos os sistemas gerados foram resolvidos numericamente e comparados com a solução dos respectivos sistemas resultantes dos três algoritmos citados na seção 5.4. Para cada rodada de cada algoritmo foram guardados:

- Tempo
- Número de simetrias eliminadas
- Valores das espécies químicas nos instantes de tempo
- Erro

O cálculo do erro foi realizado utilizando o Erro Quadrático Médio (MSE) das curvas simuladas do sistema não-simplificado contra as do simplificado. Ou seja,

$$Erro = \sqrt{(x_0 - y_0)^2 + (x_1 - y_1)^2 + \dots + (x_n - y_n)^2} \quad (6.1)$$

no qual  $x_t$  é o valor da espécie química no instante de tempo  $t$  no sistema sem simplificação,  $y_t$  é o valor da espécie química no instante de tempo  $t$  no sistema com simplificação e  $n$  é a quantidade de passos utilizado pelo Solver de sistemas de EADs.

Essas simulações foram realizadas em um dos servidores do laboratório de visão computacional do Instituto de Matemática e Estatística da Universidade de São Paulo. Este servidor possui 62 GB e 24 núcleos. Todos os algoritmos citados foram implementados na linguagem Python, ao qual é utilizada na implementação do arcabouço SigNetSim, e utilizaram a biblioteca Sundials para a solução numérica dos sistemas.

## 6.2 Resultados das Simulações e Discussão

### 6.2.1 Exemplos de Resultado da Simulação de um Sistema Sintético

Considere a reação sintética 6.2 representada pela figura 6.2.



O algoritmo 2 constrói as equações diferenciais ordinárias 6.3 para descrever a cinética de cada espécie química e o algoritmo 3 extrai as relações de conservação de massa representadas por 6.4. Note que este sistema não utiliza todas as constantes de velocidade, ou seja, não utiliza  $k_3$  e nem o  $k_{-3}$ . Portanto, é possível verificar a redução de parâmetros por meio da utilização da equação de Michaelis-Menten.

$$\frac{d[HK]}{dt} = -k_1[HK][36] + k_{-1}[HK-36] \quad (6.3a)$$

$$\frac{d[36]}{dt} = -k_1[HK][36] + k_{-1}[HK-36] \quad (6.3b)$$

$$\frac{d[HK-36]}{dt} = k_1[HK][36] - k_{-1}[HK-36] \quad (6.3c)$$

$$\frac{d[O7]}{dt} = -k_2[O7][4U] \quad (6.3d)$$

$$\frac{d[4U]}{dt} = -k_2[O7][4U] \quad (6.3e)$$

$$\frac{d[O7-4U]}{dt} = k_2[O7][4U] \quad (6.3f)$$

$$\frac{d[SJ]}{dt} = 0 \quad (6.3g)$$

$$\frac{d[94]}{dt} = \frac{-k_{cat}[SJ][94]}{km + [94]} \quad (6.3h)$$

$$\frac{d[SJ-94]}{dt} = 0 \quad (6.3i)$$

$$\frac{d[SE]}{dt} = \frac{+k_{cat}[SJ][94]}{km + [94]}. \quad (6.3j)$$

onde  $km = \frac{k_{-3} + k_{cat}}{k_3}$  e  $[94] \gg [SJ]$ .

$$[HK_0] = [HK] + [HK-36] \quad (6.4a)$$

$$[36_0] = [36] + [HK-36] \quad (6.4b)$$

$$[O7_0] = [O7] + [O7-4U] \quad (6.4c)$$

$$[4U_0] = [4U] + [O7-4U] \quad (6.4d)$$

$$[SJ_0] = [SJ] + [SJ-94] \quad (6.4e)$$

$$[94_0] = [94] + [SJ-94] + [SE]. \quad (6.4f)$$

Já que o algoritmo de percorrimento de uma única cadeia substitui as EDOs pela relação de conservação de massa do elemento que gerou a relação, o sistema a ser simulado gerado por ele é o 6.5, o qual está representado pela tupla 6.1.

**Tabela 6.1:** *Tupla que representa sistema escolhido pelo algoritmo de percorrimento de uma única cadeia.*

HK	36	HK-36	O7	4U	O7-4U	SJ	94	SE	SJ-94
1	1	0	1	1	0	1	1	0	0

$$[HK] = [HK_0] - [HK-36] \quad (6.5a)$$

$$[36] = [36_0] - [HK-36] \quad (6.5b)$$

$$\frac{d[HK-36]}{dt} = k_1[HK][36] - k_{-1}[HK-36] \quad (6.5c)$$

$$[O7] = [O7_0] - [O7-4U] \quad (6.5d)$$

$$[4U] = [4U_0] - [O7-4U] \quad (6.5e)$$

$$\frac{d[O7-4U]}{dt} = k_2[O7][4U] \quad (6.5f)$$

$$[SJ] = [SJ_0] - [SJ-94] \quad (6.5g)$$

$$[94] = [94_0] - [SJ-94] - [SE] \quad (6.5h)$$

$$\frac{d[SJ-94]}{dt} = 0 \quad (6.5i)$$

$$\frac{d[SE]}{dt} = \frac{+k_{cat}[SJ][94]}{km + [94]}. \quad (6.5j)$$

Para gerar este sistema e resolvê-lo, foram necessários 1,58 segundos. O sistema eliminou todas os termos simétricos e em comparação com o sistema 6.3, o qual não possui nenhuma simplificação, o erro calculado pela fórmula 6.1 é de 2,195 e-10 mols/(um)\*\*3.

Já para o algoritmo SFS, o sistema a ser simulado gerado por ele é o 6.6, o qual está representado pela tupla 6.2.

**Tabela 6.2:** *Tupla que representa sistema escolhido pelo algoritmo SFS.*

HK	36	HK-36	O7	4U	O7-4U	SJ	94	SE	SJ-94
1	1	0	1	1	0	0	1	0	0

$$[HK] = [HK_0] - [HK-36] \quad (6.6a)$$

$$[36] = [36_0] - [HK-36] \quad (6.6b)$$

$$\frac{d[HK-36]}{dt} = k_1[HK][36] - k_{-1}[HK-36] \quad (6.6c)$$

$$[O7] = [O7_0] - [O7-4U] \quad (6.6d)$$

$$[4U] = [4U_0] - [O7-4U] \quad (6.6e)$$

$$\frac{d[O7-4U]}{dt} = k_2[O7][4U] \quad (6.6f)$$

$$\frac{d[SJ]}{dt} = 0 \quad (6.6g)$$

$$[94] = [94_0] - [SJ-94] - [SE] \quad (6.6h)$$

$$\frac{d[SJ-94]}{dt} = 0 \quad (6.6i)$$

$$\frac{d[SE]}{dt} = \frac{+k_{cat}[SJ][94]}{km + [94]}. \quad (6.6j)$$

Isso porque, esse algoritmo inicialmente escolhe um sistema com uma EA que minimiza o número de termos; em seguida, escolhe outro sistema derivado do anterior, o qual possui outra EA que minimiza o número de termos e assim sucessivamente até o ponto em que novas sistemas não podem mais ser escolhidos.

Para gerar este sistema e resolvê-lo, foram necessários 1,72 segundos. O sistema também eliminou todos os termos simétricos e em comparação com o sistema 6.3, o erro calculado é de 8,312 e-10 mols/(um)\*\*3.

Por fim, para o algoritmo Branch and Bound, o sistema a ser simulado é o 6.7, o qual está representado pela tupla 6.3. Este algoritmo é uma abordagem que garante solução ótima, e portanto, escolhe este sistema pois é mínimo tanto na quantidade de EDOs quanto na quantidade de termos.

Para gerar este sistema e resolvê-lo, foram necessários 2,38 segundos. O sistema eliminou todos os termos simétricos e em comparação com o sistema 6.3, o erro calculado é de 8,312 e-10 mols/(um)\*\*3.

**Tabela 6.3:** Tupla que representa sistema escolhido pelo algoritmo Branch and Bound.

HK	36	HK-36	O7	4U	O7-4U	SJ	94	SE	SJ-94
0	1	1	0	1	1	0	0	1	0

$$\frac{d[HK]}{dt} = -k_1[HK][36] + k_{-1}[HK-36] \quad (6.7a)$$

$$[36_0] = [36] + [HK-36] \quad (6.7b)$$

$$[HK-36] = [HK_0] - [HK] \quad (6.7c)$$

$$\frac{d[O7]}{dt} = -k_2[O7][4U] \quad (6.7d)$$

$$[4U_0] = [4U] + [O7-4U] \quad (6.7e)$$

$$[O7-4U] = [O7_0] - [O7] \quad (6.7f)$$

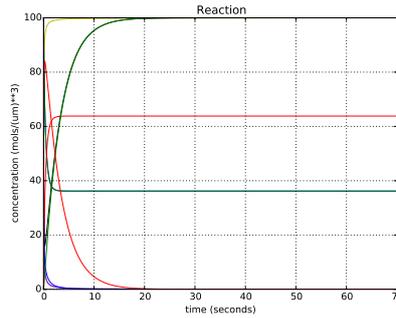
$$\frac{d[SJ]}{dt} = 0 \quad (6.7g)$$

$$\frac{d[94]}{dt} = \frac{-k_{cat}[SJ][94]}{km + [94]} \quad (6.7h)$$

$$\frac{d[SJ-94]}{dt} = 0 \quad (6.7i)$$

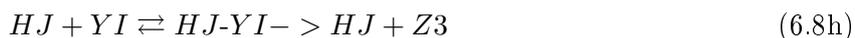
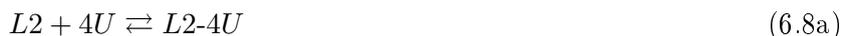
$$[SE] = [94_0] - [SJ-94] - [94]. \quad (6.7j)$$

A figura 6.3 apresenta o gráfico das concentrações de cada espécie química envolvida na reação nos instantes de tempo, as quais foram obtidas pelos sistemas gerados pelos algoritmos citados.



**Figura 6.3:** Gráfico das concentrações obtidas pela solução do sistema sem nenhuma simplificação. Os gráficos das concentrações obtidas pela solução do sistema gerado pelo algoritmo de percorrimento de uma única cadeia, SFS e BB mostram-se iguais a este.

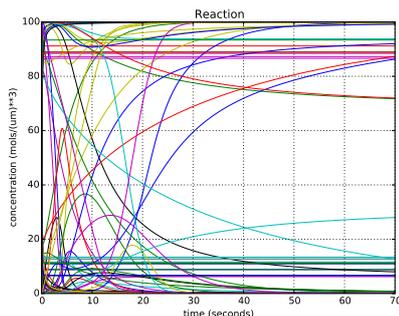
Serão apresentados também os resultados do sistema sintético que obteve maior erro, cuja reação está representado pelas equações 6.8 e pela sua continuação em 6.9.



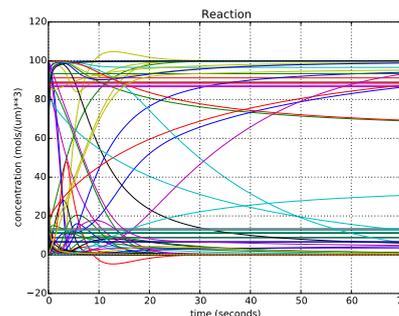
Para gerá-lo pelo algoritmo de percorrimento de uma única cadeia e resolvê-lo, foram necessários 33,38 segundos. Em comparação com o sistema sem nenhuma simplificação, o erro foi de  $773,99 \text{ e-}10 \text{ mols}/(\text{um})^{**3}$ .

Já para o algoritmo SFS, foram necessários 27,9 segundos. Em comparação com o sistema 6.3, o erro calculado foi de  $862,81 \text{ e-}10 \text{ mols}/(\text{um})^{**3}$ .

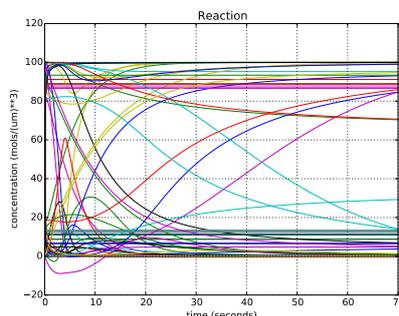
A figura 6.4 apresenta o gráfico das concentrações de cada espécie química envolvida nessa reação.



(a) Gráfico das concentrações obtidas pela solução do sistema sem nenhuma simplificação.



(b) Gráfico das concentrações obtidas pela solução do sistema gerado pelo algoritmo de percorrimento de uma única cadeia.



(c) Gráfico das concentrações obtidas pela solução do sistema gerado pelo algoritmo SFS.

**Figura 6.4:** Gráficos das concentrações de cada espécie química envolvida na reação sintética de maior erro nos instantes de tempo

### 6.2.2 Resultados Gerais

Nesta subsecção, serão apresentados os resultados gerais das simulações. A tabela 6.4 apresenta as médias dos tempos de execução dos sistemas sem nenhuma simplificação, juntamente com seu desvio padrão. Esses valores foram calculados utilizando os resultados das 10 instâncias para cada valor de  $n$ , o qual representa o número de espécies químicas envolvidas na reação e varia de 10 a 100. O resultado apresentado é intuitivo já que, quanto maior o sistema, maior o tempo para resolvê-lo.

A tabela 6.5 apresenta os resultados dos sistemas de EADs gerados pelo algoritmo de percorrimento de uma única cadeia. A primeira coluna apresenta a média de tempo de execução dos sistemas, a terceira apresenta o erro calculado pela equação 6.1, o qual é o Erro Quadrático Médio (MSE) das curvas simuladas do sistema não-simplificado contra as do simplificado. A penúltima coluna apresenta a média de equações simétricas retiradas do sistema.

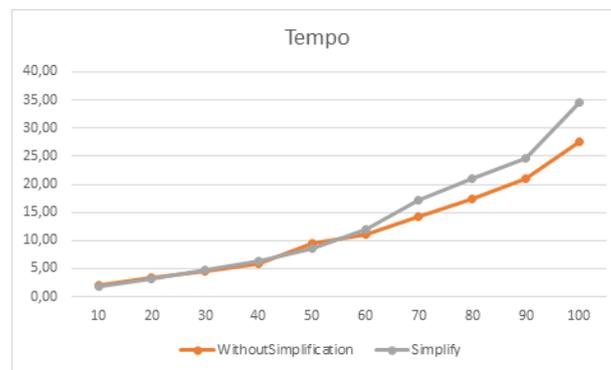
Em comparação com os sistemas sem nenhuma simplificação, o algoritmo de percorrimento de uma única cadeia apresentou resultados de tempo e erro bem próximos, além de eliminar uma quantidade considerável de simetrias. A diferença entre os tempos da tabela 6.4 e da tabela 6.5 está apresentado na figura 6.5

A tabela 6.6 apresenta os resultados dos sistemas de EADs gerados pelo algoritmo *Sequential Forward Search*. A primeira coluna apresenta a média de tempo de execução dos sistemas, a terceira apresenta o erro calculado pela equação 6.1 e a penúltima coluna apresenta a média de equações simétricas retiradas do sistema.

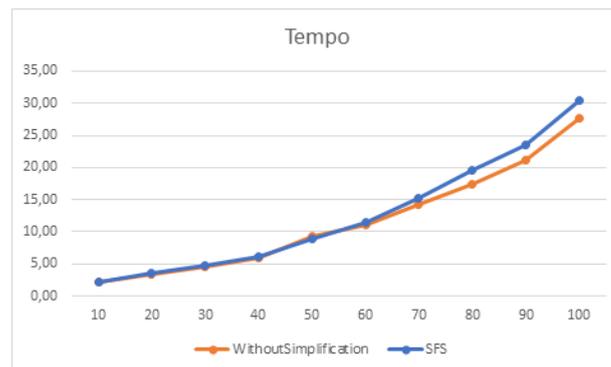
Em comparação com os sistemas sem nenhuma simplificação, o algoritmo SFS também apresentou resultados de tempo bem próximos. Entretanto, comparado com os resultados do algoritmo de percorrimento de uma única cadeia, apresentou erros maiores além de eliminar uma quantidade

**Tabela 6.4:** Tabela com valores de tempo de execução dos sistemas sem nenhuma simplificação.

N	Tempo Médio	Desvio Padrão	Simetrias
10	2.14	0.58	8.2
20	3.33	1.08	24.6
30	4.52	0.81	36.0
40	5.93	1.98	42.1
50	9.40	9.16	50.5
60	11.10	4.25	61.9
70	14.27	3.40	75.5
80	17.42	2.94	82.3
90	21.08	2.96	90.6
100	27.66	4.94	102.4

**Figura 6.5:** Gráfico que apresenta a diferença das médias de tempos de execução dos sistemas sem nenhuma simplificação contra os sistemas gerados pelo algoritmo de percorrimento de uma única cadeia.

reduzida de simetrias. A diferença entre os tempos da tabela 6.4 e da tabela 6.6 está apresentado na figura 6.6

**Figura 6.6:** Gráfico que apresenta a diferença das médias de tempos de execução dos sistemas sem nenhuma simplificação contra os sistemas gerados pelo algoritmo SFS.

Por fim, a tabela 6.7 apresenta os resultados dos sistemas de EADs gerados pelo algoritmo *Branch and Bound*. A primeira coluna apresenta a média de tempo de execução dos sistemas, a terceira apresenta o erro calculado pela equação 6.1 e a penúltima coluna apresenta a média de equações simétricas retiradas do sistema.

Embora este algoritmo apresente uma solução ótima do problema, o tempo para a execução torna a sua utilização inviável.

Note que, pelas figuras 6.5 e 6.6, a diferença entre os tempos ficam mais evidentes conforme

**Tabela 6.5:** Tabela com os resultados da execução dos sistemas com a simplificação resultante do percurso de uma única cadeia.

N	Tempo Médio	Desvio Padrão	Erro Médio	Desvio Padrão	Média Simetria
10	1.90	0.35	3.06	1.74	0
20	3.18	0.66	6.76	45.4	1.3
30	4.68	0.54	1.50 e-09	3.93 e-09	3.2
40	6.24	1.11	0.00028	0.0019	5.4
50	8.64	1.73	0.57	3.46	5.5
60	11.93	2.23	1.15	1.47e-09	6.6
70	17.24	10.24	2.13e-09	6.35e-09	8.0
80	20.95	3.92	0.00041	0.0027	8.2
90	24.60	3.08	2.09	14.02	9.3
100	34.60	12.98	76.08	490.69	10.4

**Tabela 6.6:** Tabela com os resultados da execução dos sistemas com a simplificação resultante do algoritmo SFS.

N	Tempo Busca	Tempo Médio	Desvio Padrão	Erro Médio	Desvio Padrão	Média Simetria
10	0.076	2.02	0.81	10.59	4.58	1.4
20	0.089	3.45	1.02	14.34	29.47	1.9
30	0.136	4.71	0.98	7.35	4.39	3.6
40	0.156	6.15	1.19	6.77	2.44	6.0
50	0.192	8.73	2.13	8.37	3.99	6.6
60	0.217	11.11	2.71	7.18	2.89	7.1
70	0.229	15.03	2.64	7.19	2.5	8.7
80	0.254	19.38	3.51	7.31	3.42	9.2
90	0.278	23.18	2.26	9.51	11.64	10.8
100	0.311	30.12	6.46	95.19	572.33	11.5

aumenta a quantidade de espécies químicas participantes das reações.

Note também que os resultados de erro apresentados na tabela 6.6, a qual refere-se ao algoritmo SFS e na tabela 6.7, referente ao algoritmo BB, estão iguais. Na verdade, estes valores são tão próximos que ao arredondá-los, apresentam-se de forma igual.

Para saber qual a melhor metodologia, é necessário aplicá-las à sistemas que possuem soluções conhecidas, já que um erro maior, como no caso de SFS e BB, podem indicar que estão mais próximos dos valores reais, ou seja, menos sujeitos aos erros numéricos devido às simplificações.

### 6.3 Comentários

Neste capítulo, foram apresentados os detalhes das simulações juntamente com resultados gerais. No próximo capítulo serão apresentadas as contribuições e algumas sugestões para pesquisas futuras.

**Tabela 6.7:** Tabela com os resultados da execução dos sistemas com a simplificação resultante do algoritmo *Branch and Bound*.

N	Tempo Busca	Tempo Médio	Desvio Padrão	Erro Médio	Desvio Padrão	Média Simetria
10	2.88	2.89	0.58	10.60	4.58	0
20	+10h	4.45	1.61	14.34	29.47	1.3



# Capítulo 7

## Conclusões

No capítulo anterior, foram apresentados os detalhes de como foram realizadas as simulações, bem como seus resultados. Neste capítulo, serão lembradas as contribuições deste trabalho e apontados possíveis alternativas para a continuidade desta linha de pesquisa.

### 7.1 Revisão e discussão dos resultados apresentados

No início deste trabalho foi discutida a necessidade do desenvolvimento de modelos matemáticos a fim de se entender o comportamento do sistema de redes de sinalização molecular. A modelagem utilizada foi baseada na cinética química, ou seja, no estudo da evolução da concentração das espécies químicas de uma reação ao longo de um determinado tempo. A equação que descreve a dinâmica de uma determinada espécie química ao longo do tempo tem a forma de uma Equação Diferencial Ordinária (EDO). A vantagem do uso de modelos baseados em EDOs é que, embora soluções analíticas não sejam garantidas, métodos numéricos para a solução de um sistema com grande número de equações diferenciais ordinárias não-lineares são bem desenvolvidos. No entanto, quanto maior for o sistema de EDOs simulado, mais métodos numéricos estão sujeitos a erros. Outro ponto é que o aumento do tamanho do sistema, muitas vezes, resulta em simulações que são mais exigentes do ponto de vista computacional. Além disso, quanto maior o sistema de EDOs, maior a quantidade de simetrias de termos entre diferentes equações. Assim, o objetivo deste projeto de mestrado foi o desenvolvimento de regras para simplificar sistemas de equações diferenciais ordinárias utilizando para este fim relações de conservação de massa entre as espécies químicas envolvidas, que por sua vez tratam-se de equações algébricas. Além disso, no caso de sistemas que contêm reações enzimáticas, também foi utilizado reduções do sistema original de EDOs por meio da equação de Michaelis-Menten.

Para isso, no capítulo 2, foi apresentado os fundamentos teóricos nos quais este trabalho foi baseado, incluindo as condições necessárias e suficientes para a utilização da equação de Michaelis-Menten. No capítulo 3, foi introduzido um algoritmo para mapear uma lista de reações químicas para um sistema de equações diferenciais ordinárias (EDOs). No capítulo 4, foi proposto um algoritmo que extrai relações de conservação de massa entre espécies químicas a fim de serem empregadas para simplificar o modelo cinético através do desenho de sistemas de equações algébrico-diferenciais (EADs). No capítulo 5, foram apresentadas três metodologias para, dado um sistema de EDOs e um conjunto de relações de conservação de massa entre as espécies químicas envolvidas, produzir diferentes sistemas de EADs que descrevam a cinética modelada. Foram apresentados também critérios de escolha de um sistema de EADs ótimo. No capítulo 6 foi apresentado os resultados dos experimentos realizados para avaliar a metodologia proposta, além de um algoritmo de geração de reações químicas e seus parâmetros, o qual foi utilizado para as simulações. Neste capítulo, vimos que pelo critério de minimização do número de equações diferenciais e do número de termos das equações diferenciais, o algoritmo de percorrimento de uma única cadeia apresentou resultados de tempo e erro bem próximos, além de eliminar uma quantidade considerável de simetrias em comparação aos sistemas sem nenhuma simplificação. O algoritmo SFS também apresentou resultados de tempo

bem próximos. Entretanto, comparado com os resultados do algoritmo de percorrimento de uma única cadeia, apresentou erros maiores além de eliminar uma quantidade reduzida de simetrias. Por fim, embora o algoritmo Branch and Bound apresente uma solução ótima do problema, vimos que o tempo para a execução torna a sua utilização inviável.

Entretanto, por mais que estes resultados apresentaram pouca vantagem em relação ao tempo de execução, eles foram capazes de inspecionar o espaço de EADs possíveis, permitindo a escolha de um sistema de acordo com o critério escolhido. Assim, como vimos no capítulo 1, ao simplificarmos um sistema em um sistema de EADs, podemos substituir equações diferenciais por equações algébricas, eliminando a necessidade de efetuarmos medidas experimentais para as espécies químicas cujas equações diferenciais foram substituídas por equações algébricas. Outro ponto é que a simplificação utilizando a equação de Michaelis-Menten facilitou a estimação de parâmetros e mitigou o problema de overfitting, já que diminuiu a dimensão dos parâmetros, conforme descrito na seção 2.3.

## 7.2 Contribuições

Resumindo, as contribuições deste trabalho são:

- Um algoritmo que formaliza uma maneira de extração de relações de conservação de massa a partir de uma lista de reações químicas, o qual foi apresentado em 3
- Um algoritmo para simplificar sistemas de equações diferenciais que descrevem a cinética de reações químicas, o qual gera sistemas alternativos que eliminam simetrias, fato importante para uma melhor análise de bifurcação e que possibilita os biólogos a resolverem um sistema sem conhecer todos os valores de concentrações iniciais das espécies químicas.
- Geração de sistemas alternativos que possuem parâmetros reduzidos, facilitando a estimação de parâmetros e suavizando o problema de overfitting.
- Contribuição tecnológica: Construção de um componente do arcabouço SigNetSim. Ou seja, foi desenvolvido um programa em Python com base nos algoritmos definidos, de forma que o mesmo pudesse ser aproveitado de maneira imediata como um componente do arcabouço SigNetSim.

## 7.3 Sugestões para Pesquisas Futuras

Para concluir este trabalho, listamos algumas sugestões para a continuidade desse trabalho:

- Teste da metodologia em problemas biológicos reais.
- Verificação da hipótese de que um sistema simplificado acarreta em menor erro numérico: para isso, será necessário utilizar como *benchmarking* sistemas de EDOs que possuam solução de fórmula fechada, e que esta última seja conhecida.
- Estudo e aplicação de outros algoritmos de busca possivelmente mais robustos para encontrar o sistema ótimo de forma mais eficaz e eficiente.
- Sugestão de melhoramento tecnológico: Estudo e aplicação de estruturas de dados mais robustas, o que provavelmente redundaria em menor tempo para resolver o problema de otimização.

# Referências Bibliográficas

- [CDC83] Carlos A. Callioli, Hygino H. Domingues e Roberto C.F. Costa. *Álgebra Linear e Aplicações*. Atual Editora, 4ª edição, 1983. 24
- [CNS10] William W. Chen, Mario Niepel e Peter K. Sorger. Classic and contemporary approaches to modeling biochemical reactions. *Genes & Development*, 2010. 1, 3
- [DHZ<sup>+</sup>09] Jayajit Das, Mary Ho, Julie Zikherman, Christopher Govern, Ming Yang, Arthur Weiss, Arup K. Chakraborty e Jeroen P. Roose. Digital signaling and hysteresis characterize ras activation in lymphoid cells. *Cell*, 2009. 11
- [EI04] Narat J Eungdamrong e Ravi Iyengar. Modeling cell signaling networks. *Biology of the Cell*, 2004. 1, 2
- [HYD05] PETER E. HYDON. Introduction to symmetry methods in the solution of differential equations that occur in chemistry and chemical biology. *Wiley InterScience*, 2005. 2
- [JLI00] J.Dedrick Jordan, Emmanuel M Landau e Ravi Iyengar. Signaling networks: The origins of cellular multitasking. *Cell*, 2000. 1
- [MMF<sup>+</sup>14] Reis MS, Dias MHS, Nakano F, Barrera J e Armelin HA. Mathematical modeling of ras/mapk signaling network kinetics provides insights into mechanisms of fgf2-induced cell cycle arrest in the k-ras-driven mouse y1 adrenocortical tumor cells. *RECOMB*, 2014. Web page:<http://f1000.com/posters/browse/summary/1095383>. 3
- [RA14] Marcelo S. Reis e Hugo A. Armelin. *Introduction to Modeling of the Kinetics of Signaling Networks*, 2014. 11
- [SAB05] Robert J. Silbey, Robert A. Alberty e Mounqi G. Bawendi. *Physical Chemistry*. John Wiley & Sons, Inc, 2005. 5
- [VV11] Dinald Voet e Judith G. Voet. *Biochemistry*. John Wiley & Sons, 2011. 9, 11