

# **SIMUNO**

SISTEMA DE SIMULAÇÃO IMUNOLÓGICA

Alunos: Rodrigo Januário da Silva  
Marcos Tiago Garcia da Silva  
Orientador: Prof. Marcelo Finger  
Colaborador: Prof. Eduardo Finger

# Índice

1	Introdução.....	3
1.1	Objetivos.....	3
2	Conceitos e Tecnologias Estudadas.....	4
2.1	Conceitos Biológicos.....	4
2.1.1	Imunidade Inata.....	4
2.1.1.1	Diferenciando o próprio do patógeno.....	5
2.1.2	Imunidade Adaptativa.....	5
2.1.2.1	Diferenciando o próprio do patógeno.....	5
2.1.3	Funcionamento da resposta imunológica.....	6
2.1.4	O Reconhecimento de um Epítipo.....	9
2.2	Conceitos de Tecnologia e desenvolvimento.....	11
2.2.1	Motivação.....	11
2.2.2	Desafio.....	12
2.2.3	Modelagem e implementação.....	12
2.2.3.1	Diagrama UML.....	14
2.2.3.2	Algoritmo de decisão de reconhecimento.....	15
2.2.4	Ferramentas utilizadas.....	16
2.2.5	Metodologia.....	17
2.2.6	Atividades.....	18
2.2.7	Produto.....	18
2.2.8	Resultados.....	18
3	Conclusões.....	18
4	Bibliografia.....	19
5	Parte Subjetiva.....	20
5.1	Desafios e frustrações.....	20
5.2	Disciplinas Relevantes ao Projeto.....	21
5.3	Futuro do Projeto.....	22

# 1 Introdução

O sistema imunológico é de extrema importância para a sobrevivência e adaptabilidade de qualquer ser vivo. É graças a ele que somos capazes de lidar com o grande número de micro organismos existentes e que insistentemente invadem nosso corpo a todo instante. Não fosse por este complexo, eficiente e adaptável sistema, certamente não mais habitaríamos este planeta, tamanha é a quantidade e variedade de invasões que sofreremos.

Entender nosso sistema imunológico é uma tarefa muito difícil. Existem inúmeras reações bioquímicas e diversas interações celulares acontecendo simultaneamente para garantir que estaremos a salvo das constantes e indesejadas invasões que sofreremos. Além da alta complexidade do sistema, temos um outro agravante que dificulta o entendimento do sistema como um todo: Nem mesmo os especialistas entendem uma grande parte dos acontecimentos que levam a algumas reações diferenciadas do sistema. Como exemplo elucidativo, podemos citar a diferenciação dos linfócitos T4 em linfócitos comuns e linfócitos memória no momento da reprodução linfocítica após o reconhecimento do epítipo.

O sistema imunológico, nos vertebrados, pode ser dividido em três grandes compartimentos:

- *Imunidade inata*: Bloqueia a entrada de microorganismos no meio interno.
- *Imunidade adaptativa*: Elimina agressores do meio interno e acumula uma memória imunológica que previne a recorrência do agente agressor.
- *Imunidade intermediária*: É formado por células tipicamente da imunidade adaptativa que, no entanto, respondem como componentes da imunidade inata (Linfócitos B-1 e NK).

## 1.1 Objetivos

O objetivo do trabalho é fazer uma simulação computacional de uma das partes mais importantes do sistema imunológico: O reconhecimento dos epítomos, pois é a partir daí que os linfócitos T4, T8 e B são ativados para contra-atacar os invasores e livrar-nos deles. No escopo do trabalho está também a simulação da resposta imune inata, que é uma resposta que acontece sempre igual para todos os seres vivos.

## **2 Conceitos e Tecnologias Estudadas**

Nesta seção descreveremos os conceitos de imunologia que estudamos para a realização da simulação, bem como as tecnologias que estamos utilizando para que se torne viável a implementação das partes propostas. Iniciaremos então, explicando detalhadamente os processos imunológicos envolvidos na simulação para então em seguida descrevermos como estes processos foram modelados computacionalmente.

### **2.1 Conceitos Biológicos**

Nesta seção descreveremos os conceitos biológicos que estão relacionados à parte do sistema imunológico que o sistema irá simular.

#### **2.1.1 Imunidade Inata**

A imunidade inata utiliza vários mecanismos físicos, químicos e celulares que foi programada no decorrer da evolução dos seres vivos para reagir aos invasores de uma forma padrão. A resposta imunológica inata é típica e age sempre igual e inclui os seguintes componentes:

- A pele que é uma barreira contra microorganismos;
- Secreções viscosas que bloqueiam o acesso às mucosas;
- Substâncias microbidas, inclusas nas secreções orgânicas;

Suas principais células são os macrófagos e os neutrófilos, ambos capazes de fagocitar e secretar substâncias microbidas contra invasores. Algumas dessas substâncias são as lisozimas, defensinas e radicais livres de oxigênio.

Enquanto os macrófagos residem basicamente nos tecidos, os neutrófilos residem na circulação e migram para o interstício imediatamente após uma agressão, em resposta a estímulos quimiotáticos decorrentes do dano celular.

Sabendo disso, fica uma pergunta: Por que os macrófagos e neutrófilos não atacam o próprio indivíduo à que pertencem? Como são diferenciadas as próprias células de células patógenas?

### **2.1.1.1 Diferenciando o próprio do patógeno**

A imunidade inata promove a diferenciação entre o próprio e o patógeno através de dois métodos.

O primeiro método consiste em equipar as células com receptores sensíveis a substâncias que não existem no organismo. Essas substâncias são chamadas de PAMPs (*Pathogen-Associated Molecular Patterns*). Dentre essas substâncias podemos destacar o lipopolissacarídeo bacteriano, o DNA demetilado, o RNA de dupla hélice viral e a flagelina.

O segundo método são as opsoninas, que são proteínas que se ligam simultaneamente a receptores de leucócitos e de patógenos mediando, desta maneira, a fagocitose. Essa ligação pode ser específica, através dos anticorpos ou inespecífica, através do complemento.

## **2.1.2 Imunidade Adaptativa**

São parte constituinte da imunidade inata os linfócitos T e B. Ambos os linfócitos são caracterizados por possuírem receptores, TCR e BCR respectivamente, em sua superfície que são encarregados de reconhecer peptídeos contendo entre 5 e 17 aminoácidos chamados epítomos. O reconhecimento de um epítomo é que torna o patógeno visível para a imunidade adaptativa.

Cada TCR / BCR é único e por serem gerados por recombinação aleatória de seus genes, podemos dizer que sua especificidade é totalmente obra do acaso. Por serem receptores gerados aleatoriamente o potencial destrutor da imunidade adaptativa é muito maior que o da imunidade inata. Tanto o linfócito T quanto o linfócito B destroem tudo que encontram pela frente uma vez que fizeram o reconhecimento do epítomo apresentado. Entretanto por passarem por um rígido processo de seleção os linfócitos T e B que são selecionados são capazes de reconhecer apenas epítomos de patógenos.

### **2.1.2.1 Diferenciando o próprio do patógeno**

Graças a imunidade adaptativa, o foco das atividades do sistema imunológico mudou de um pequeno grupo existente de PAMPs para um substrato extremamente variado, existente em todos os seres vivos: as proteínas. O sistema

imunológico dos vertebrados foi reestruturado, fazendo com que a imunidade inata passasse a ser responsável pela resistência local imediata em zonas de interface entre o meio interno e externo, deixando para a imunidade adaptativa a responsabilidade de neutralizar os agressores em nível sistêmico. Integrando a resposta inata a a resposta adaptativa estão as células apresentadoras de antígenos, conhecidas como APC ou DC.

As APCs ou DCs são células que internalizam através da fagocitose os antígenos nas áreas de agressão, processam-nos em epítopos e então migra para o baço e para os linfonodos afim de apresentá-los para os linfócitos T através do Complexo Maior de Histocompatibilidade conhecido como MHC, do inglês *Major Histocompatibility Complex*. O MHC é um conjunto de 6 genes extremamente polimórficos que define uma assinatura celular única para cada indivíduo denominado haplótipo. Além disso, também codifica proteínas especializadas em selecionar epítopos para apresentá-los em sua superfície. Como os linfócitos T só são capazes de reconhecer epítopos que são apresentados através do MHC, podemos dizer que o haplótipo define o que o sistema imunológico tolera ou ataca.

O MHC pode ser dividido em dois tipos distintos: o MHC classe I (MHC I) que é expresso por todas as células do organismo e apresenta epítopos produzidos no citosol (auto-antígenos) e o MHC classe II (MHC II) expresso apenas por APCs e que apresenta antígenos extracelulares fagocitados.

É graças à existência dos auto-antígenos que a imunidade adaptativa é capaz de diferenciar o próprio dos patógenos, pois uma vez que um linfócito T reconhece um auto-antígeno apresentado através do MHC classe I ele é eliminado antes mesmo de ser liberado para o ataque. Isso garante que os linfócitos que atacariam as células do próprio organismo sejam eliminados para que não haja uma destruição em massa.

### 2.1.3 Funcionamento da resposta imunológica

A resposta imunológica é iniciada quando o dano celular desencadeia cascatas inflamatória, que ocorre devido a ativação de uma série de substâncias e reações entre as mesmas. Essas cascatas produzem ativação celular e fatores quimiotáticos, que faz com que ocorra um fluxo maciço de neutrófilos e monócitos, que como os macrófagos e as células dendríticas (também conhecidas como APCs), respondem aos processos inflamatórios e aos PAMPs através da fagocitose dos patógenos e da secreção de substâncias bactericidas e citocinas que são proteínas essenciais para a comunicação intercelular e podem ser divididas em interleucinas, quimiocinas e hematopoietinas.

Após serem ativadas pelo dano celular e pelos PAMPs, as células dendríticas iniciam um intenso período de fagocitose e pinocitose quando, então, adquirem uma grande quantidade de antígenos locais. Esses antígenos então são processados em epítopos durante o caminho de migração das DCs (ou APCs) para os linfonodos, onde serão então apresentados sobre o MHC para

reconhecimento dos linfócitos. As DCs ativadas possuem uma importante diferença das não ativadas: Apenas as ativadas expressão a molécula de co-estimulação B7 em sua superfície.

É então que os linfócitos T virgens testam seus TCRs contra os epítomos apresentados nos MHCs das DCs. Para que os linfócitos virgens possam ser ativados, é necessário, sem exceção, que o sinal enviado pelo TCR seja acompanhado de um sinal paralelo produzido pela ligação entre o ligante CD28 do linfócito com o ligante B7 da DC ativada. Uma vez que este sinal paralelo não é enviado o linfócito reconhece o antígeno testado como sendo um auto-antígeno e, portanto, ele é descartado. Se o linfócito não reconhecer nenhum dos epítomos apresentados ele migra para outros linfonodos a fim de continuar tentando reconhecer outros epítomos. Cabe aqui ressaltar que apenas os linfócitos T virgens necessitam do co-estímulo B7 para sua ativação, já que os linfócitos T memória respondem à sinalização exclusiva do TCR com ativação plena.

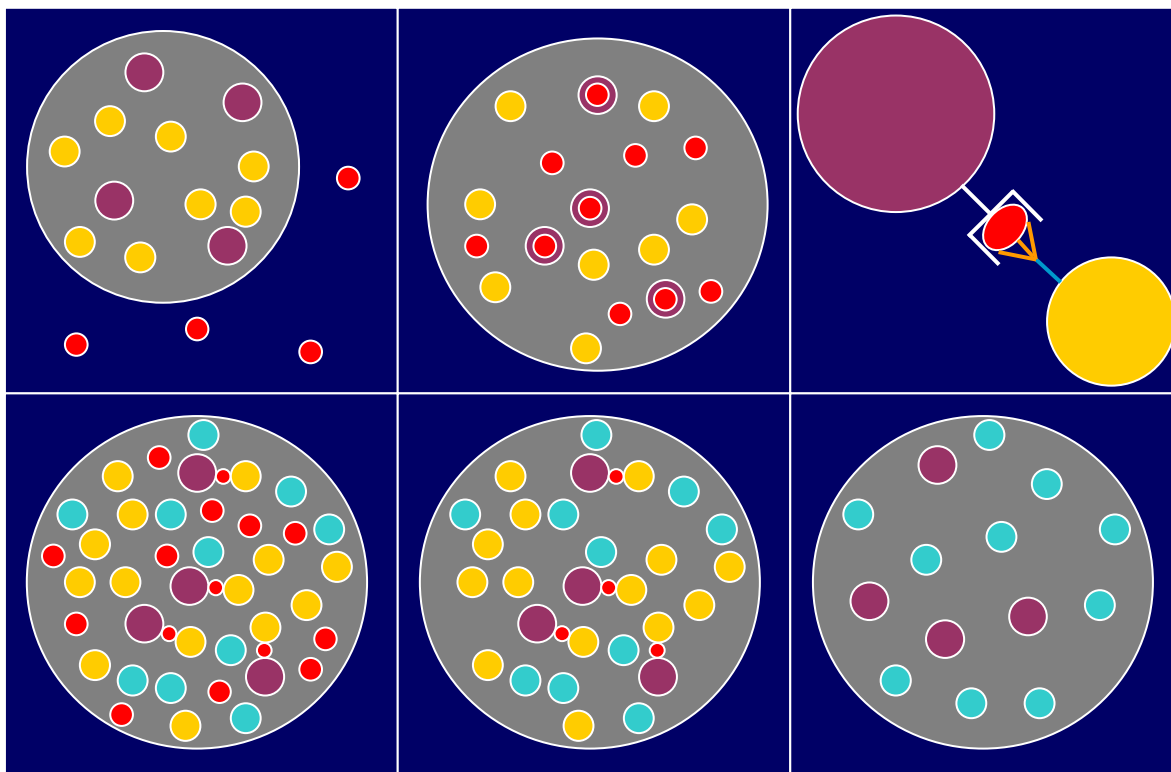
Uma vez ativados, os linfócitos T4 auxiliam as DCs na produção do co-estimulante B7 afim de ativar os linfócitos T8. A ativação de linfócitos deste tipo requer uma co-estimulação significativa maior do que aquela que as DCs podem produzir antes de receber o auxílio dos linfócitos T4. Já a ativação de linfócitos B depende da co-estimulação através do ligante CD40L que também é expresso por linfócitos T4 ativados, ou seja, os linfócitos T4 organizam e comandam a resposta adaptativa.

Além do ligante CD40L, linfócitos T4 ativados secretam altas quantidades do receptor IL-2, que é uma citocina fortemente mitogênica que promove a expansão exponencial da população de linfócitos T4 ativados. Durante essa fase de reprodução são produzidas tanto linfócitos efetores da resposta imunológica quanto linfócitos memória.

O contato do linfócito T4 com a APC, avalia, também, uma série de parâmetros para definir o tipo de resposta que será utilizada contra o agressor. Essa resposta, chamada de resposta auxiliadora (T helper ou Th) poderá ser predominantemente celular (Th1) ou humoral (Th2) sendo que respostas do tipo Th1 são mais efetivas contra vírus e bactérias e respostas do tipo Th2 são mais efetivas contra infecções por helmintos e neutralização de toxinas.

Interessantemente, os mesmos estímulos que induzem a ativação dos linfócitos induzem, também a expressão de moléculas inibidoras (CTLA-4). Enquanto houver invasores nos organismos o sinal ativador dos linfócitos supera o sinal inibidor e na medida que os invasores vão sendo eliminados o sinal inibidor vai, gradativamente, superando o sinal ativador. A falta do sinal ativador induz à morte (apoptose) em massa dos linfócitos efetores (não dos linfócitos memória).

Abaixo vemos uma figura representando o ciclo do sistema imunológico:



No primeiro quadro vemos um organismo saudável. Os antígenos em vermelho ainda não efetuaram a invasão.

No quadro 2 os antígenos já efetuaram a invasão, provocando a produção de substâncias que ativam os macrófagos e as APCs a iniciarem a fagocitose que, após efetuada, faz com que estas células comecem sua migração para os linfonodos. Durante o caminho os antígenos são processados em epítopos para a apresentação através do MHC.

No quadro 3 temos uma representação de um epítipo sendo apresentado no MHC com um linfócito T4 conectado à ele, tentando o reconhecimento.

No quadro 4 vemos a reprodução em massa dos linfócitos T4 em células efetoras (amarelo) e células memória (azul). Essas células então iniciam o ataque contra os agressores.

No quadro 5 vemos os agressores eliminados. Uma vez eliminados inicia-se o processo de apoptose das células efetoras.

No quadro 6 vemos apenas as células efetoras que restaram.



## 2.1.4 O Reconhecimento de um Epítopo

Um epítopo como já foi dito anteriormente é uma seqüência que contém de 5 a 17 aminoácidos. Entretanto para o propósito deste trabalho consideraremos apenas seqüências entre 13 e 17 aminoácidos que é o tamanho da seqüência que pode ser acomodada no MHC classe II.

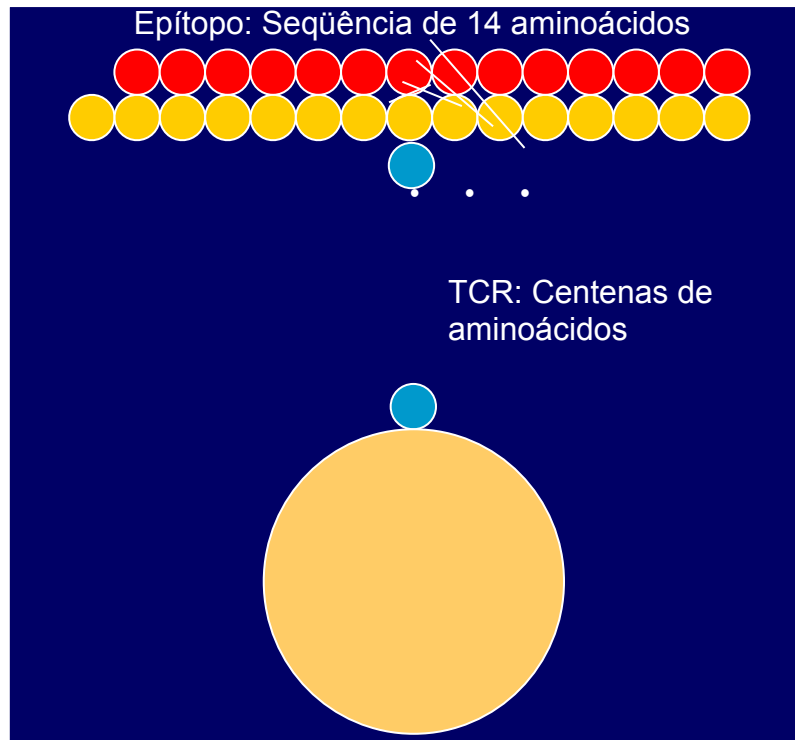
O TCR ou CDR dos linfócitos T4 constituído por mais de 100 aminoácidos, mas para a função do reconhecimento de epítopos são designados apenas 15 aminoácidos que estão em três partes distintas de contato do CDR, conhecidos como CDR1, CDR2 e CDR3. Cada um dos CDRs possui uma seqüência de 5 aminoácidos que são testados de forma pareada com os aminoácidos do epítopo em questão.

Existe uma série de forças que são levadas em consideração no reconhecimento de um epítopo. Abaixo temos uma tabela contendo as principais forças que serão levadas em consideração no projeto:

Forças não covalentes	Origem
Forças Eletrostáticas	Ação entre cargas opostas
Pontes de hidrogênio	Hidrogênios compartilhados entre átomos eletronegativos.
Forças de Van der Waals	Flutuações em nuvens de elétrons ao redor das moléculas polarizam os átomos de maneira oposta.
Forças Hidrofóbicas	Grupos hidrofóbicos interagem desfavoravelmente com a água e tendem a se juntar para expulsar moléculas de água. Esta atração também envolve Forças de Van der Waals.

É a resultante entre essas diversas interações dos aminoácidos que nos diz que se um epítopo é reconhecido ou não. Quando temos uma resultante repulsiva, dizemos que um epítopo não é reconhecido, pois o linfócito T4 não consegue estabelecer uma ligação entre seu TCR e o epítopo em questão. Entretanto quando a resultante é atrativa o reconhecimento é efetuado.

É importante ressaltar que como a força resultante é variável o reconhecimento do epítopo desencadeia respostas imunológicas de maneira diferente dependendo da força de reconhecimento. Quanto maior a força resultante mais rápida e efetiva será a resposta imunológica contra o agente invasor. Abaixo temos um exemplo ilustrativo de como um epítopo é pareado pelo TCR a fim de ser reconhecido pelo mesmo.



Na figura acima vemos o pareamento entre um epítipo com 14 aminoácidos e o TCR do linfócito T4. Vale ressaltar que a interação não é apenas entre os Aminoácidos pareados, mas também há interação entre os aminoácidos que estão nas diagonais. A força de interação é proporcional ao distância entre os aminoácidos, em angstroms sendo que 1 angstrom = 0,1 nanômetro.

## **2.2 Conceitos de Tecnologia e desenvolvimento**

Este trabalho visou estudar a área de simulação de fenômenos naturais através de modelagem computacional. Neste caso, o fenômeno estudado pertence à área biológica. Nosso trabalho consistiu em modelar o funcionamento do sistema imunológico humano e deste contexto vem o maior desafio que encontramos.

### **2.2.1 Motivação**

No século vinte houve um avanço considerável nas áreas das ciências em geral. A área médica não ficou de fora dessa evolução. Assim, melhoraram-se os equipamentos de pesquisa por meio de tecnologia, descobriram-se novos compostos químicos e biológicos que auxiliam na pesquisa e tratamento de doenças, aumentou-se o conhecimento geral sobre a saúde e muitos outros avanços. Mas, com tudo isso, houve também um olhar mais crítico da sociedade quanto aos métodos empregados pelos pesquisadores. Discussões sobre ética, como aconteceu ao ser inventada a clonagem por exemplo, tornaram-se públicas e relevantes. Grupos ativistas cada vez mais surgem na mídia com seus protestos contra os testes de medicamentos em animais e a favor de um método menos cruel ou agressivo com a natureza.

Paralelo a tudo isso, a área da computação também evoluiu muito dos anos 50 até os dias de hoje, quando já se é possível ter computadores com altíssimas capacidades dentro de casa e a um custo razoavelmente baixo. Outra característica importante existente na computação e que relaciona-se de modo ortogonal à biologia é o fato de ela ser “pervasiva”. Ou seja, atualmente a computação tem a capacidade de se infiltrar nas mais diversas áreas e às vezes de maneira imperceptível aos usuários em geral. Hoje em dia milhares de tipos diferentes de aparelhos e dispositivos possuem chips e funcionam por meio de software e contribuíram e muito para essa grande evolução que houve na ciência, em especial no nosso caso de estudo, a área biológica.

Assim, considerando-se o surgimento do problema da ética e da necessidade inerente da área biológica de fazer testes em seres vivos para pesquisar novos tratamentos e verificar a eficácia deles unido à tecnologia existente no mundo da computação moderna, surgiu a idéia de se reproduzir o mundo real em que vivemos de uma maneira virtual, num ambiente em que seria possível realizar testes à vontade, e em grande escala, sem necessariamente criar problemas éticos ou fazer mal ao meio ambiente, por exemplo. Essa reprodução da vida real em um ambiente virtual seria feita através de uma modelagem abstrata do universo em que vivemos, tentando-se extrair da realidade os conceitos e características essenciais para seu funcionamento. No caso da biologia, a idéia seria compreender exatamente como funciona o ser vivo em estudo e reproduzir de modo abstrato o seu funcionamento, a fim de poder utilizá-lo à vontade. Neste trabalho a idéia foi a de simular o sistema imunológico

humano, para que o pesquisador pudesse testar métodos preventivos de doenças sem ter que efetivamente contaminar indivíduos ou testar remédios e/ou vacinas que potencialmente poderiam fazer mal à pessoa.

## 2.2.2 Desafio

Diferentemente da computação, que é uma área exata e tem todo o seu domínio bem definido e consistente, a biologia é altamente imprecisa e muitas vezes desconhecida. Para se criar um modelo de um processo natural, como a imunologia, deveríamos saber o máximo de detalhes possível – senão todos eles – para reproduzir fielmente no modelo o funcionamento observado na natureza. Foi então que encontramos nosso primeiro grande problema: os próprios imunologistas não sabem como todas as partes do sistema funcionam.

Existe uma quantidade enorme de informação, como visto na primeira parte, mas ainda assim não se sabe tudo. E por esse motivo principalmente, nosso projeto teve que reduzir seu escopo de trabalho a uma porção menor do sistema imunológico. Esta parte, pelas necessidades que um modelo computacional exige, teria de ter seu funcionamento completamente conhecido e compreendido pelo homem. Assim, ficaria possível fazer um modelo abstrato de seu funcionamento e, portanto, reproduzi-lo dentro de um computador. Depois de muita conversa com um especialista no assunto, chegamos à conclusão que faríamos a modelagem de uma parte do processo de reconhecimento de um antígeno pelo sistema imunológico chamada Reconhecimento de Epítomos.

## 2.2.3 Modelagem e implementação

Inicialmente procuramos entender como o processo acontece, e isolamos seus elementos, tentando deixá-los bem definidos quanto suas características e funcionalidades. Assim, modelamos todos os elementos que são importantes nesse processo, como por exemplo células (linfócito, APC), conectores (MHC, T4, TCR), aminoácidos e sequência de aminoácidos (epítopo), antígeno e o ambiente. Dessa maneira, implementamos cada um desses itens representados em classes, com seus atributos e métodos de manipulação dos seus dados e de interação com outros itens no sistema.

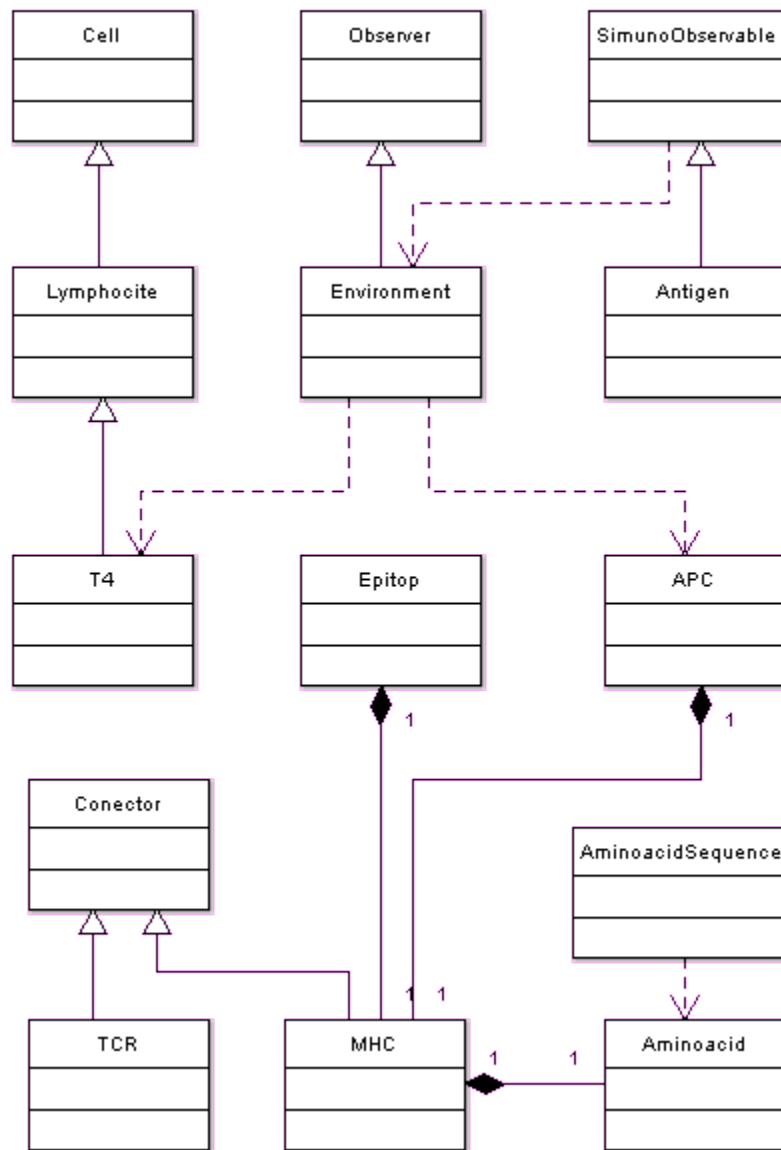
Para fazer esta modelagem, decidimos utilizar uma técnica de desenvolvimento de software chamada Orientação a Objetos.

Orientação a Objetos consiste em dividir o problema em questão em pequenas partes individuais, que têm uma certa autonomia. Essas partes são chamadas abstratamente de classes, e quando cria-se um item de uma classe, este passa a chamar-se objeto. Ou seja, um objeto é uma realização de uma classe. A Orientação a Objetos veio de uma tentativa de tornar o software fácil de ser desenvolvido por se assemelhar muito ao mundo real, onde temos objetos de

verdade. Por exemplo, poderíamos criar uma classe chamada Carro, que possui atributos como nome, cor, velocidade e preço, e possui métodos (ações) como ligar, desligar, acelerar e frear. E com esta classe, poderíamos criar um objeto dela (uma instância), e atribuir valores aos atributos deste objeto, que é do tipo Carro, como “Ferrari”, “vermelho”, “200 km/h”, “R\$ 500.000”.

Por essa praticidade e semelhança com o mundo real, a Orientação a Objetos mostrou-se muito adequada para o nosso problema de modelar o sistema imunológico. Assim, criamos classes para todos os elementos citados, e também algumas classes que não são representações de itens reais, mas são necessárias para poder realizar a nossa simulação. Por exemplo, utilizamos o padrão Observer para criar uma classe que é responsável por monitorar o ambiente e perceber a entrada de novos antígenos. Também utilizamos muito o conceito de herança, já que temos vários tipos de células, de conectores e de aminoácidos. A herança permite que uma classe seja subclasse de uma outra, aproveitando todos os métodos e atributos da classe mãe. No exemplo do Carro, esta classe poderia ser subclasse de uma classe chamada Veículo, por exemplo. Esta poderia ter outras subclasses, como Moto ou Barco. No nosso trabalho, as classes Linfocito e APC herdam da classe Celula, por exemplo. Diz-se que elas são do mesmo tipo. Por fim, também utilizamos programação concorrente, que consiste em rodar vários pequenos programas simultaneamente e que interagem entre si. Isto serviu para simular as células “vivas” interagindo entre si e agindo e no ambiente.

### 2.2.3.1 Diagrama UML



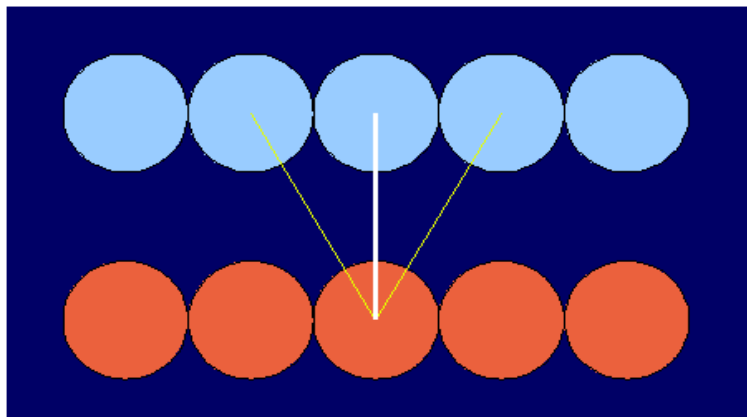
Modelagem UML dos Objetos

Em relação aos algoritmos, nos concentramos em entender como os componentes atuam e interagem entre si, para poder codificar de maneira mais a se assemelhar à realidade. Em especial no caso da ligação entre os aminoácidos, tivemos que desenvolver um algoritmo que leva em consideração a força de atração/repulsão e a distância entre eles e entre os vizinhos. Uma combinação de todas essas forças é o que decide se a ligação vai ser feita ou não.

### 2.2.3.2 Algoritmo de decisão de reconhecimento

- Propósito:
  - Dado duas sequências de aminoácidos, decidir se elas se atraem com força suficiente ou não.
- Entrada:
  - Duas sequência de aminoácidos a serem testadas;
  - Valores de forças eletrostáticas, hidrofóbicas, Van der Waals e pontes de hidrogênio para cada aminoácido;
  - Distância entre um aminoácido e seus vizinhos;
- Funcionamento do Algoritmo:

O algoritmo consiste em calcular, para cada par de aminoácidos associado, um de cada sequência, a força de atração entre eles. O algoritmo leva em consideração os valores informados para cada um dos aminoácidos, e também faz o mesmo cálculo considerando os vizinhos imediatos dos aminoácidos, mas usando um fator de redução da força de atração devido à distância e à posição menos privilegiada. Veja a figura:



Forças de atração entre aminoácidos de duas sequências

O traço branco indica a principal força entre os aminoácidos sendo analisados em uma iteração do algoritmo. Os traços amarelos indicam as forças secundárias que também influenciam na atração entre os aminoácidos. Nesta imagem vemos apenas a atração da sequência azul sobre um dos aminoácidos da sequência abóbora. O algoritmo considera também, para cada par, as forças secundárias análogas, para o aminoácido do lado oposto ao que vemos na imagem.

Assim, podemos fazer uma iteração por todos os aminoácidos da sequência calculando a força de atração entre cada par e ao final calcular uma força resultante da atração entre as duas sequências.

- Saída:
  - Reconhecido (a resultante das forças foi forte o suficiente para as sequências se ligarem) ou não reconhecido (caso contrário).

Um problema que também ficou evidente foi que o algoritmo se utiliza de um valor ajustável, que é a força suficiente de atração necessária entre dois aminoácidos. Necessitamos desta informação para poder afirmar que a conexão entre as cadeias acontece e é estável, ou não. Esse parâmetro vem da observação biológica no mundo real.

## 2.2.4 Ferramentas utilizadas

Para codificar tudo isso, utilizamos a linguagem de programação Java, da Sun Microsystems. Foi criada nos anos 90 e ela diferencia-se das outras linguagens de programação convencionais por utilizar uma máquina virtual. Ao se compilar um código Java, gera-se um código binário chamado de bytecode, que é interpretado pela máquina virtual. Assim, o mesmo bytecode pode funcionar em qualquer máquina virtual Java, não importa em que sistema operacional esta máquina virtual está instalada. O Java é a linguagem orientada a objetos mais utilizada no mundo, com extensa documentação e suporte. Ela também é robusta e confiável, além de ser escalável, e dessa maneira adequa-se muito bem ao trabalho desenvolvido por nós.

Também utilizamos um ambiente gráfico de desenvolvimento, chamado Eclipse. O Eclipse é um projeto de código aberto gratuito que foi originalmente criado para se desenvolver aplicações em Java, mas que é extensível por meio de plugins. Assim, hoje em dia pode-se desenvolver no Eclipse praticamente em qualquer linguagem utilizando os seus recursos. Ele é escrito em uma versão especial de Java que é compilado diretamente para a plataforma. Assim, existe versões de Eclipse para Windows, Mac e Linux/Unix. Isso foi feito para que a interação com o usuário fosse mais rápida, já que quando o acesso às bibliotecas gráficas do Sistema Operacional é feito através da máquina virtual Java, ele acaba ficando lento e assim o tempo de resposta alto pode tornar a experiência



do usuário um pouco desagradável. No caso do nosso trabalho, entretanto, não necessitamos de interface gráfica nessa primeira fase do projeto, pois a simulação é feita toda internamente com o software apenas apresentando resultados finais em modo texto. Portanto este problema está descartado, por enquanto. Mas mesmo para o futuro, o projeto Eclipse fornece sua biblioteca gráfica gratuitamente, assim pode-se desenvolver aplicações que rodam interfaces gráficas com rapidez e robustez. O único problema é que o software desenvolvido deixará de ser multiplataforma. Porém se isso não for uma restrição impactante, pode ser uma boa saída.

Como nosso projeto foi feito em duas pessoas e como as duas pessoas utilizam diferentes computadores, nós achamos necessário utilizar um repositório de código para concentrar nosso projeto em um único lugar e assim facilitar a organização. Para isso utilizamos um repositório chamado Subversion (SVN). O SVN é um dos repositórios mais utilizados hoje e possui algumas vantagens sobre o CVS, como possibilidade de reestruturar a árvore de diretórios, remover arquivos, etc. O SVN também é um controlador de versões. Assim, podemos evoluir o nosso código sem perder o rastro do que foi feito anteriormente, podendo assim desfazer grandes mudanças ou corrigir bugs em versões antigas do software. Nós utilizamos o SVN do site Sourceforge.net, que é um repositório gratuito online, podendo ser acessado de qualquer computador conectado à internet.

## **2.2.5 Metodologia**

A Metodologia que utilizamos para desenvolver este projeto foi uma combinação de alguns métodos de engenharia de software. Basicamente nós utilizamos conceitos de Prototipação e de Programação eXtrema (XP). Do primeiro, nós utilizamos a idéia de se desenvolver protótipos do programa final e mostrar ao cliente (no caso, o imunologista). Assim, a cada vez que o cliente tem contato com o software (mesmo que uma versão mais simples e menos funcional), ele tem novas idéias e corrige possíveis erros que nós estamos cometendo. Isso ajuda muito quando o cliente conhece muito pouco do universo de desenvolvimento de software, como foi o caso. Do XP nós utilizamos o conceito de programação pareada, que ajuda a evitar erros simples e acelera o desenvolvimento e o refactoring, que consiste em rever constantemente o código a fim de melhorar a implementação, refazendo algumas partes.

### **2.2.6 Atividades**

As principais atividades que fizemos neste projeto foi inicialmente uma pesquisa intensa na área biológica, lendo papers, livros e consultando o pesquisador. Depois tivemos algumas reuniões tanto em nosso ambiente (faculdade) como no dele (hospital), para refinar o problema e chegar no escopo em que trabalhamos, que foi o Reconhecimento de Epítomos. Finalmente passamos para a implementação, fase em que geramos protótipos para mostrar e refinamos os algoritmos, tentando fazer nosso simulador se comportar o mais próximo possível com o mundo real.

### **2.2.7 Produto**

Como produto de nosso trabalho, produzimos um software que simula, embora sem muita precisão, a parte do sistema imunológico que faz o reconhecimento do antígeno. O nosso software foi desenvolvido de maneira a permitir sua extensão com facilidade, ampliando assim o seu uso. No futuro a idéia é que poderemos simular não só o sistema imunológico, mas sim muitos outros processos biológicos humanos.

### **2.2.8 Resultados**

Os resultados obtidos não foram satisfatórios para o imunologista, pois ainda necessitamos de muito tempo para criar um software de tamanha complexidade. Mas para nós serviu como aprendizado na área biológica, em lidar com clientes que não tem noção de como se desenvolve um software e também aprendemos a modelar problemas que têm muitas variáveis e não são completamente conhecidos pelo homem.

## **3 Conclusões**

Implementar um sistema que tenta copiar, mesmo que de maneira simplista, um fenômeno da área biológica é extremamente difícil. As informações disponíveis hoje na medicina imunológica são úteis para os médicos mas precisam passar por um processo intenso de padronização e modelagem para que se possa fazer um software que as utilize de maneira adequada. Também pudemos perceber que este programa requer um hardware de capacidade relativamente alta, pois ele deve simular milhares de componentes do sistema imunológico para chegar perto do que acontece no corpo humano.

## 4 Bibliografia

Cabaniols JP, Cibotti R, Kourilsky P, Kosmatopoulos K, Kanellopoulos JM.  
Dose-dependent T cell tolerance to an immunodominant self peptide.  
Eur J Immunol. 1994 Aug;24(8):1743-9.  
PMID: 8056033 [PubMed - indexed for MEDLINE]

Dubey C, Croft M, Swain SL.  
Naive and effector CD4 T cells differ in their requirements for T cell receptor versus costimulatory signals.  
J Immunol. 1996 Oct 15;157(8):3280-9.  
PMID: 8871622 [PubMed - indexed for MEDLINE]

Wulfing C, Rabinowitz JD, Beeson C, Sjaastad MD, McConnell HM, Davis MM.  
Kinetics and extent of T cell activation as measured with the calcium signal.  
J Exp Med. 1997 May 19;185(10):1815-25.  
PMID: 9151707 [PubMed - indexed for MEDLINE]

Kim S, Patrick SM, Braunstein NS, Thomas JL, Leonard EF.  
Modeling of early events in T cell signal transduction after controlled T cell activation by peptide major histocompatibility complex.  
Ann Biomed Eng. 2001 May;29(5):373-83.  
PMID: 11400719 [PubMed - indexed for MEDLINE]

Altan-Bonnet G, Germain RN.  
Modeling T cell antigen discrimination based on feedback control of digital ERK responses.  
PLoS Biol. 2005 Nov;3(11):e356. Epub 2005 Oct 25.  
PMID: 16231973 [PubMed - indexed for MEDLINE]

Imunobiologia (Janeway, Travers, Walpont, Shlomchik).

## **5 Parte Subjetiva**

Nesta parte da monografia, descreveremos os principais desafios e frustrações encontradas no desenvolvimento do projeto, as disciplinas mais relevantes do bacharelado para o projeto bem como uma breve descrição sobre a aplicação dessas disciplinas no projeto. Finalizaremos falando um pouco do futuro do projeto.

### **5.1 Desafios e frustrações**

O principal desafio encontrado neste projeto foi o fato de estarmos trabalhando com uma área, até então, totalmente desconhecida. Como já foi dito anteriormente, o sistema imunológico é extremamente complexo e entender seu funcionamento é uma tarefa difícil, pois envolve vários outros conceitos de biologia que invariavelmente precisam ser estudados.

Conversar com os biólogos também não é uma tarefa fácil, e foi um grande desafio, pois a linguagem que falamos e o tipo de raciocínio são completamente diferentes. Tirar informações que são realmente relevantes para a modelagem muitas vezes foi uma tarefa árdua, já que praticamente a todo o momento as informações passadas eram tão detalhadas que chegava a ponto de não conseguirmos entender o que estava sendo pedido. Como estávamos trabalhando com uma modelagem top-down, queríamos inicialmente uma explicação superficial, para modelarmos o problema em alto-nível para só então fazermos refinamentos sucessivos e entrarmos nos detalhes.

A maior frustração dentro do projeto foi descobrir sua complexidade. Inicialmente pensávamos em modelar o sistema imunológico completo, mas ao vermos o quão complexo esse sistema é, frustramo-nos, pois vimos que apenas uma pequena parte poderia ser modelada.

Além disso, no decorrer do desenvolvimento do projeto fomos descobrindo que existem vários aspectos do sistema imunológico que ainda não são conhecidos nem pelos especialistas e modelar estas partes é extremamente difícil, pois temos que assumir muitas coisas que não sabemos se vai funcionar computacionalmente ou não. Por isso deixamos estas partes desconhecidas para o futuro, na continuação do projeto.

## 5.2 Disciplinas Relevantes ao Projeto.

Abaixo listaremos um conjunto de disciplinas que nos ajudaram no desenvolvimento do projeto:

Disciplinas que são relevantes do ponto de vista algorítmico.

- MAC0122 (Princípios de Desenvolvimento de Algoritmos)
- MAC0323 (Estruturas de Dados)
- MAC0328 (Algoritmos em Grafos)

Disciplinas relacionadas ao desenvolvimento do sistema.

- MAC0211 (Laboratório de Programação I)
- MAC0239 (Métodos Formais em Programação )
- MAC0242 (Laboratório de Programação II)
- MAC0332 (Engenharia de Software)
- MAC0338 (Análise de Algoritmos)
- MAC0441 (Programação Orientada a Objetos)
- MAC0413 (Tópicos de Programação Orientada a Objetos)
- MAC0438 (Programação Concorrente)

Disciplinas matemáticas que ajudaram na modelagem do sistema.

- MAE0121 (Introdução à Probabilidade e Estatística I)
- MAE0212 (Introdução à Probabilidade e Estatística II)

### **5.3 Futuro do Projeto**

Como a simulação do sistema imunológico é algo muito complexo e que ajudaria muito os pesquisadores da área a tirar conclusões importantes sobre alguns aspectos da imunologia e sobre algumas doenças, creio que o projeto poderá ser continuado por nós mesmos ou por algum outro interessado na área, tanto em um novo trabalho de formatura quanto em um mestrado.

O sistema que está sendo entregue é apenas uma pequena parte de um todo muito grande e complexo. Tenho, da minha parte, a pretensão de continuar com o projeto, seguir um mestrado nesta área de biologia computacional, pois foi uma área que me despertou muito interesse.