

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

ALESSANDRO DALLA VECCHIA

**Semáforos adaptativos, uma abordagem
baseada em inteligência de enxames para a
mobilidade urbana**

Trabalho de Conclusão apresentado como
requisito parcial para a obtenção do grau de
Bacharel em Ciência da Computação

Profa. Dra. Ana Lúcia Cetertich Bazzan
Orientador

Porto Alegre, 12 de Dezembro de de 2013

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Alessandro Dalla Vecchia,

Semáforos adaptativos, uma abordagem baseada em inteligência de enxames para a mobilidade urbana /

Alessandro Dalla Vecchia. – Porto Alegre: Graduação em Ciência da Computação da UFRGS, 2013.

98 f.: il.

Trabalho de Conclusão – Universidade Federal do Rio Grande do Sul. Curso de Bacharelado em Ciência da Computação, Porto Alegre, BR–RS, 2013. Orientador: Ana Lúcia Cetertich Bazzan.

1. Comportamento Adaptativo. 2. Aprendizagem por Reforço. 3. Inteligência de Enxame. 4. Controle de Tráfego. 5. Simulação de Tráfego. I. Lúcia Cetertich Bazzan, Ana. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos Alexandre Netto

Vice-Reitor: Prof. Rui Vicente Oppermann

Pró-Reitor de Graduação: Prof. Sérgio Roberto Kieling Franco

Diretor do Instituto de Informática: Prof. Luís da Cunha Lamb

Coordenador do CIC: Prof. Raul Fernando Weber

Bibliotecário-chefe do Instituto de Informática: Alexander Borges Ribeiro

A maioria dos males sociais e políticos com os quais você sofre está sob seu controle, e poderá acabar se você tiver vontade e coragem. Você poderá viver uma outra vida, uma vida mais sábia, se estiver disposto a pensar e a encontrar soluções. Você não tem consciência do poder que possui. (Platão)

AGRADECIMENTOS

Em primeiro lugar gostaria de agradecer àqueles que tem me apoiado constantemente em minha vida (seja na parte pessoal ou acadêmica), e sem os quais eu não seria quem sou hoje, a meus pais, Terezinha e Adilar, e a meu irmão caçula Andrei, muito obrigado por tudo.

Em segundo lugar gostaria de agradecer a minha orientadora, Ana Bazzan, por me aceitar em seu grupo de pesquisa (desejo este que carregava desde os primeiros semestres do curso), por abraçar este trabalho de conclusão que escrevo, pelas cobranças constantes e justas, pelo seu conhecimento e profissionalismo de excelência, muito obrigado.

Não poderia deixar de incluir aqui um agradecimento especial para as professoras Aline Villavicencio, Cleci Bevilacqua e Anna Maciel que me proporcionaram os primeiros passos no mundo da pesquisa científica. E também para todo o pessoal do Departamento de Sistemas da Informação do Centro de Processamento de Dados da UFRGS¹, os quais foram como uma família para mim durante os quase dois anos em que estivemos juntos.

Por último, mas não menos importante, obrigado a todos vocês meus amigos que estiveram juntos nessa longa subida pela graduação, em especial ao André Ferreira por todos os algoritmos que não aprendi, ao Leonardo Moura pelas discussões infundáveis, ao Wagner Schmitt pelos divertidos passeios, e a Júlia Kambara por me fazer manter os pés no chão de vez em quando.

Este trabalho foi parcialmente financiado pelas agências de fomento CNPQ²-CAPES³ e FAPERGS⁴.

¹ <http://www.ufrgs.br/cpd/>

² <http://www.cnpq.br/>

³ <http://www.capes.gov.br/>

⁴ <http://www.fapergs.rs.gov.br/>

SUMÁRIO

SUMÁRIO	5
LISTA DE FIGURAS	8
LISTA DE TABELAS	10
LISTA DE SÍMBOLOS	12
LISTA DE ALGORITMOS	14
SIGLAS	15
RESUMO	15
ABSTRACT	16
1 INTRODUÇÃO	18
1.1 Motivação	18
1.2 Objetivos	19
1.3 Métodos	20
1.4 Contribuições	20
1.5 Estrutura do trabalho	20
2 BASE CONCEITUAL	21
2.1 Engenharia de Tráfego	21
2.1.1 Sistemas Inteligentes de Transporte	26
2.2 Simulação de tráfego	26
2.2.1 Modelo de simulação microscópico	26

2.2.2	Simulation of Urban MObility (SUMO)	27
2.3	Sistemas Multiagente	31
2.4	Inteligência de Enxame	32
2.5	Aprendizagem por Reforço	33
2.5.1	Processo de Decisão de Markov	34
2.5.2	Aprendizagem por reforço passivo	35
2.5.3	Aprendizagem por reforço ativo	37
2.6	Considerações finais	39
3	METODOLOGIA E ALGORITMOS	40
3.1	Ambiente de testes	40
3.1.1	Hardware e software utilizado	40
3.1.2	Simulador e simulação	40
3.2	Modelagem das redes de tráfego urbano e suas demandas	44
3.2.1	Mapa em formato de cruzamento simples	45
3.2.2	Mapa em formato de cruzamento duplo	45
3.2.3	Mapa em formato de bloco	46
3.2.4	Mapa das vias de Sioux Falls	48
3.2.5	Mapa das vias arteriais de Porto Alegre	49
3.2.6	Comparação das características das redes propostas	50
3.3	Abordagem via inteligência de enxame	51
3.3.1	Fluxo de atuação da abordagem	51
3.3.2	Descrição do algoritmo	51
3.4	Abordagem via aprendizagem por reforço	55
3.4.1	Fluxo de atuação da abordagem	55
3.4.2	Descrição do algoritmo	56
3.5	Abordagem híbrida: <i>Smart Ant with Context-Detection (SA-CD)</i>	57
3.5.1	Fluxo de atuação da abordagem	57
3.5.2	Descrição do algoritmo	58
3.6	Considerações finais	60
4	EXPERIMENTOS REALIZADOS E RESULTADOS	61
4.1	Experimentos realizados	61
4.2	Resultados	62
4.2.1	Resultados da rede viária de Cruzamento Simples	62
4.2.2	Resultados da rede viária de Cruzamento duplo	68

4.2.3	Resultados da rede viária de Formato bloco	72
4.2.4	Resultados da rede viária de Sioux Falls	78
4.2.5	Resultados da rede viária Poa Arterials	82
4.3	Considerações finais	88
5	CONCLUSÕES E TRABALHOS FUTUROS	89
5.1	Trabalhos futuros	89
5.2	Dificuldades encontradas	90
	APÊNDICE A MATRIZES ORIGEM-DESTINO	91
A.1	Rede viária da cidade de Sioux Falls	91
A.2	Rede contendo as vias arteriais de Porto Alegre	93
	REFERÊNCIAS	95

LISTA DE FIGURAS

2.1	Fotos de engarrafamentos, do passado até o presente.	22
2.2	Elementos da engenharia de tráfego abordados	24
2.3	Exemplo de controladores semafóricos reais e simulado, plano e fases semafóricos	25
2.4	Comparação das diferentes granularidades de simulação	27
2.5	Modelos de malhas viárias complexas importadas no SUMO.	29
2.6	Modelos de malhas viárias abstratas possíveis de serem criadas no SUMO.	29
2.7	Arquitetura de um sistema de aprendizagem por reforço monoagente .	34
2.8	Exemplo de cadeia de Markov	34
3.1	Exemplo de descrição de via.	41
3.2	Exemplo de descrição de veículo e rota.	42
3.3	Exemplo de descrição de intersecção.	42
3.4	Exemplo de descrição de junção interna.	42
3.5	Exemplo de descrição de semáforo. Inclusão do tipo (<i>type</i>) "traffic_light".	43
3.6	Exemplo de descrição do plano semafórico e das fases.	43
3.7	Modelos de malhas simples criadas no SUMO para testes iniciais. . .	47
3.8	Visão sobre o mapa viário real e simulado da cidade de Sioux Falls, Dakota do Sul, EUA	49
3.9	Visão sobre o mapa viário real e simulado da cidade de Porto Alegre.	50
4.1	Comparação dos resultados de PT na rede viária de Cruzamento simples	64
4.2	Comparação dos melhores resultados de PT na rede viária de Cruzamento simples	65
4.3	Comparação dos resultados de TMV na rede viária de Cruzamento simples	66
4.4	Comparação dos melhores resultados de TMV na rede viária de Cruzamento simples	67

4.5	Comparação dos resultados de PT na rede viária de Cruzamento duplo	69
4.6	Comparação dos melhores resultados de PT na rede viária de Cruzamento duplo	70
4.7	Comparação dos resultados de TMV na rede viária de Cruzamento duplo	71
4.8	Comparação dos melhores resultados de TMV na rede viária de Cruzamento duplo	72
4.9	Comparação dos resultados de PT na rede viária de Formato bloco . .	74
4.10	Comparação dos melhores resultados de PT na rede viária de Formato bloco	75
4.11	Comparação dos resultados de TMV na rede viária de Formato bloco	76
4.12	Comparação dos melhores resultados de TMV na rede viária de Formato bloco	77
4.13	Comparação dos resultados de PT na rede viária de Sioux Falls . . .	79
4.14	Comparação dos melhores resultados de PT na rede viária de Sioux Falls	80
4.15	Comparação dos resultados de TMV na rede viária de Sioux Falls . .	81
4.16	Comparação dos melhores resultados de TMV na rede viária de Sioux Falls	82
4.17	Comparação dos resultados de PT na rede viária de Poa Arterials . . .	84
4.18	Comparação dos melhores resultados de PT na rede viária de Poa Arterials	85
4.19	Comparação dos resultados de TMV na rede viária de Poa Arterials .	86
4.20	Comparação dos melhores resultados de TMV na rede viária de Poa Arterials	87

LISTA DE TABELAS

3.1	Fluxo de veículos da malha viária cruzamento simples	45
3.2	Fluxo de veículos da malha viária cruzamento duplo	46
3.3	Fluxo de veículos da malha viária em formato de bloco	47
3.4	Comparação das redes para execução dos experimentos iniciais	50
3.5	Comparação das redes para execução dos experimentos principais . . .	51
3.6	Colônia de formigas: parâmetros de entrada	53
3.7	Aprendizagem Q: parâmetros de entrada	56
3.8	SA-CD: parâmetros de entrada	59
4.1	Visão geral sobre os experimentos realizados	62
4.2	Resultados de PT e TMV para rede viária de Cruzamento simples, sem uso de algoritmo	63
4.3	Resultados de PT e TMV e desvios padrão para rede viária de Cruza- mento simples, usando Colônia de formigas	63
4.4	Resultados de PT e TMV e desvios padrão para rede viária de Cruza- mento simples, usando Aprendizagem Q	63
4.5	Resultados de PT e TMV e desvios padrão para rede viária de Cruza- mento simples, usando SA-CD	63
4.6	Resultados de PT e TMV para rede viária de Cruzamento duplo, sem uso de algoritmo	68
4.7	Resultados de PT e TMV e desvios padrão para rede viária de Cruza- mento duplo, usando Colônia de formigas	68
4.8	Resultados de PT e TMV e desvios padrão para rede viária de Cruza- mento duplo, usando Aprendizagem Q	68
4.9	Resultados de PT e TMV e desvios padrão para rede viária de Cruza- mento duplo, usando SA-CD	70
4.10	Resultados de PT e TMV para rede viária de Formato bloco, sem uso de algoritmo	73
4.11	Resultados de PT e TMV e desvios padrão para rede viária de For- mato bloco, usando Colônia de formigas	73

4.12	Resultados de PT e TMV e desvios padrão para rede viária de Formato bloco, usando Aprendizagem Q	73
4.13	Resultados de PT e TMV e desvios padrão para rede viária de Formato bloco, usando SA-CD	73
4.14	Resultados de PT e TMV para rede viária de Sioux Falls, sem uso de algoritmo	78
4.15	Resultados de PT e TMV e desvios padrão para rede viária de Sioux Falls, usando Colônia de formigas	78
4.16	Resultados de PT e TMV e desvios padrão para rede viária de Sioux Falls, usando Aprendizagem Q	80
4.17	Resultados de PT e TMV e desvios padrão para rede viária de Sioux Falls, usando SA-CD	80
4.18	Resultados de PT e TMV para rede viária Poa Arterials, sem uso de algoritmo	82
4.19	Resultados de PT e TMV e desvios padrão para rede viária de Poa Arterials, usando Colônia de formigas	83
4.20	Resultados de PT e TMV e desvios padrão para rede viária de Poa Arterials, usando Aprendizagem Q	83
4.21	Resultados de PT e TMV e desvios padrão para rede viária de Poa Arterials, usando SA-CD	83
A.1	Matriz OD da rede viária de Sioux Falls	92
A.2	Matriz OD da rede viária de Porto Alegre	94

LISTA DE SÍMBOLOS

\aleph	Lista de fatores de discretização de estados
α	Taxa de aprendizagem
β	Taxa de dissipação de feromônio
χ	Lista de incrementos e decrementos do tempo de fase em verde
Δt	Intervalo de tempo discreto
Δ_k	Porção de tempo utilizada pela fase k
$\delta_{l,t}$	Densidade de veículos em uma faixa l no tempo t
ϵ	Fator de seleção de ação do agente
γ	Fator de desconto da aprendizagem
λ	Coefficiente de aprendizado/esquecimento
μ	Intervalo de tempo para cálculos de ocupação veicular
Ω	Fator que altera a importância relativa à recompensa e a transição do modelo
$\bar{\epsilon}$	Decaimento do fator de seleção de ação do agente
ϕ	O número máximo de planos aprendidos
π	Política (sequência de ações)
$\pi(s)$	Ação recomendada no estado s
π^*	Política ótima
ρ	Coefficiente de ajuste da qualidade do modelo
σ	Desvio padrão do rastro de feromônio
τ	Expoente de recompensa
θ_{ij}	Limiar do agente i em relação a tarefa j
φ	Um fator que influencia a perda de planos aprendidos

a	Ação
$c_m(s, a)$	Confiança do modelo m baseada na escolha de uma ação a num estado s
$d_{in_{k,t}}$	Densidade de feromônio acumulado nas faixas de entrada da fase k no tempo t
e_m	Qualidade instantânea do modelo m
$E_{m_{cur}}$	Qualidade associada ao modelo corrente
E_{min}	Especifica a qualidade mínima esperada do modelo
e_m^R	Qualidade da predição da recompensa do modelo m
e_m^T	Qualidade da predição da transição do modelo m
k	Possível estado da tabela-Q
L	Tamanho de uma faixa em uma via
MAX_{N_m}	O número máximo de escolha de ações passadas
$N_m(s, a)$	Contador de ações a tomadas no estado s para o modelo m
$N_{l,t}$	Número de veículos em uma faixa l no tempo t
r	Valor de recompensa
$R_{m_{cur}}$	Modelo interno de recompensa para o modelo corrente
s	Estado
s'	Estado futuro
s_j	Estímulo associado a tarefa j
$T_{\theta_{ij}}(s_j)$	Tendência de um agente i executar uma tarefa como uma relação entre o estímulo (s_j) e o limiar de resposta (θ_{ij})
$T_{m_{cur}}$	Modelo interno de transição para o modelo corrente
w	Intervalo de tempo discreto
Z_R	Fator de normalização da recompensa
Z_T	Fator de normalização da transição

LISTA DE ALGORITMOS

1	Colônia de formigas	54
2	Aprendizagem Q	57
3	Smart Ant with Context-Detection (SA-CD)	60

SIGLAS

AM Aprendizagem de Máquina. 29

AR Aprendizagem por Reforço. 29–32, 35, 36

ATIS Sistema Avançado de Informação ao Viajante. 21

ATMS Sistema Avançado de Gestão de Viagem. 21

DT Diferença temporal. 33, 34

IA Inteligência Artificial. 29

ITS Sistemas Inteligentes de Transporte. 21

LM Livre de Modelo. 29

PDA Programação dinâmica adaptativa. 33, 34

PDM Processo de Decisão de Markov. 30, 31, 33, 34

SUMO Simulation of Urban MObility. 22, 23, 37, 41

RESUMO

O problema de mobilidade urbana tem crescido a níveis cada vez mais desgastantes para a população das cidades. Acompanhamos diariamente notícias sobre: os enormes engarrafamentos, a falta de transporte público eficiente e de qualidade, o desrespeito as regras de trânsito, os acidentes e as mortes desnecessárias.

Nem todos os problemas podem ser resolvidos apenas com a melhoria da educação no trânsito, alguns requerem altos investimentos, como na expansão e duplicação de vias, construção de viadutos e pontes, e que permitem suprir a demanda cada vez maior de veículos. Com menor custo de implantação podemos incluir a sinalização de trânsito e semafórica, que ajuda no controle das vias e aumenta a segurança para as pessoas.

Neste trabalho é estudado um sistema de controle de trânsito simulado com a perspectiva de melhorar o sistema semafórico real. Visto que desta forma é possível causar um grande impacto aos sistemas de transporte em geral com um baixo custo de implantação quando comparado as modificações na infraestrutura das vias.

Serão aplicados três algoritmos de inteligência artificial ao sistema semafórico experimental, de forma a torná-lo adaptativo, e assim permitir uma melhoria do fluxo de veículos através de um sistema que não necessita de um controle centralizado.

Os resultados serão apresentados através da comparação dos algoritmos utilizados, levando-se em conta questões como tempo de simulação e tempo de viagem dos veículos.

Palavras-chave: Comportamento Adaptativo, Aprendizagem por Reforço, Inteligência de Enxame, Controle de Tráfego, Simulação de Tráfego.

Adaptive traffic signals, a swarm-based approach for the urban mobility

ABSTRACT

The problem of urban mobility has grown to increasingly stressful levels for the cities population. We see daily reports about: huge traffic jams, the lack of a quality and efficient public transportation, the disregard of the traffic rules, the accidents and the unnecessary deaths.

All problems can not be solved only with the improvement of traffic education, some require high investments, such as the expansion and duplication of streets, construction of viaducts and bridges, which allow to meet the increasing demand of vehicles. With lower cost of deployment we can include traffic signs and traffic lights, which helps in streets control and increases the people safety.

In this work is presented a study about a simulated traffic control system with the prospect of improving the actual traffic system. In this way is possible to generate a major impact on transport systems, in general with a low implementation cost if compared with changes in the infrastructure of roads.

Will be applied three artificial intelligence algorithms to an experimental traffic lights system, in order to make it adaptive, and thus allow an improvement in the traffic flow through a system which does not require a centralized control.

The results are presented through a comparison of the algorithms used, taking into account issues such as simulation time and vehicles travel time.

Keywords: Adaptive Behavior, Reinforcement Learning, Swarm Intelligence, Traffic Control, Traffic Simulation.

1 INTRODUÇÃO

1.1 Motivação

Já vemos como rotina dos cidadãos das grandes cidades do mundo enfrentar longos congestionamentos no trânsito, seja usando um veículo privado ou coletivo, perdendo assim horas em seu deslocamento diário. Como exemplo, a cidade de São Paulo chega a registrar engarrafamentos de mais de 200 quilômetros de veículos parados¹. Nessas condições, muitas pessoas acabam adotando medidas mais extremas, fazendo longas caminhadas, já que se utilizassem veículos acabariam passando mais tempo parados na via do que realmente percorrendo a rota planejada.

Junte aos problemas que já ocorrem, a melhoria econômica nacional, tendo assim um aumento substancial no número de veículos - sejam eles particulares, comerciais ou coletivos - no trânsito das cidades^{2, 3}. Nas palavras do presidente da Associação Nacional dos Fabricantes de Veículos (Anfavea), Luiz Moan Yabiku Júnior:

"O governo teve um papel extremamente importante no desempenho do setor automotivo neste ano, não só quanto ao juro para os caminhões e máquinas agrícolas, mas também na redução da carga tributária dos automóveis e comerciais leves. A carga tributária era muito alta e, com a redução, o consumidor está respondendo."

Mas mesmo com esse fortalecimento na indústria automobilística sendo um bom fator para o país, ainda existe um déficit já conhecido de investimento em mobilidade urbana e são necessários estudos que possam trazer melhorias de uma forma mais barata e sustentável à população que se locomove diariamente⁴.

Das soluções que vem sendo exploradas para reduzir o problema dos congestionamentos podemos citar três:

A solução básica e de mais alto custo, é aumentar a malha viária, alargando vias para comportarem mais veículos, construindo novas vias arteriais e expressas, acessos e viadutos. Esta solução parece simples, mas na prática é de execução complicada em núcleos urbanos. Também possui outros pontos negativos, como o aumento da poluição visual, sonora e do ar.

¹ <http://g1.globo.com/sao-paulo/noticia/2013/06/sp-tem-mais-de-200-km-de-lentidao-na-noite-desta-sexta-feira.html>

² <http://redeglobo.globo.com/globoecologia/noticia/2013/06/crescimento-no-numero-de-carros-e-semelhante-ao-de-onibus-no-brasil.html>

³ <http://agenciabrasil.ebc.com.br/noticia/2013-06-06/crescimento-das-vendas-de-automoveis-mantem-previsao-otimista-para-ano>

⁴ <http://www1.folha.uol.com.br/poder/2013/10/1362010-ao-lado-de-alkmin-dilma-critica-antecessores-do-pt-por-falta-de-investimento-em-metro.shtml>

A segunda é a redução do número de veículos em circulação. Isto pode ser realizado através da cobrança de pedágio em áreas mais congestionadas (como centros urbanos), e incentivo aos meios de transporte públicos e alternativos, como o uso de bicicletas em algumas cidades⁵. Bons resultados já vem sendo alcançados com este último em cidades européias como Paris e Holanda, mas mesmo assim, até elas podem gerar engarrafamentos⁶. Transportes alternativos como a bicicleta também possuem algumas desvantagens, seja devido às intempéries, seja pelo risco associado ao dividir a via com outros veículos quando não existe sinalização adequada ou mesmo quando existe^{7, 8}. Do lado motorizado por sua vez, é necessária maior paciência e atenção do motorista ao dar preferência a estes tipos de transporte, reduzindo assim sua velocidade e aumentando ainda mais o tempo de deslocamento.

Uma terceira solução é a melhor utilização da malha viária existente, ou seja, remover zonas de estacionamento, modificar o sentido do tráfego em algumas ruas, auxiliar os veículos para que tomem uma rota alternativa que consumirá um menor gasto de tempo, e melhorar os sistemas de controle de tráfego, proporcionando maior fluidez do trânsito em pontos mais críticos. Para isso ocorrer, são necessárias técnicas que envolvam a informação histórica do próprio sistema, além de ferramentas computacionais para lidar com toda essa informação.

Neste trabalho o foco estará nesta última solução, sistemas de controle de tráfego. Através de técnicas avançadas de simulação de sociedades será utilizado um sistema multiagente mapeado em uma rede de semáforos simulada de algumas malhas viárias urbanas. Com esta técnica incrementada por uma metáfora inspirada na biologia, e outra baseada na aprendizagem de máquina, proponho a implementação de um semáforo inteligente, ou seja, uma máquina proposta a aprender através do que ela percebe de seu ambiente (nesse caso, o fluxo de veículos) e possa melhorar a fluidez do trânsito urbano que controla. Essa característica de adaptabilidade é muito importante, visto que os sistemas de trânsito são altamente dinâmicos e complexos.

1.2 Objetivos

Este trabalho tem como objetivo principal a experimentação de três algoritmos de Inteligência Artificial sobre um sistema semafórico simulado, de forma a avaliar o comportamento do mesmo.

Espera-se desenvolver uma estratégia de controle descentralizado e em tempo real que possibilite minimizar a necessidade de um especialista em tráfego que atue no controle desse sistema.

Indiretamente, a atuação dessa estratégia em um sistema de trânsito real permitiria uma melhora no fluxo de veículos e gastos com combustível para o usuário final (satisfação do cidadão), causaria um menor impacto financeiro às contas da cidade (menor gasto com aumento de infraestrutura), e um menor impacto ambiental (redução de poluição urbana).

⁵ <http://www.mobilicidade.com.br/bikepoa.asp>

⁶ <http://exame.abril.com.br/meio-ambiente-e-energia/noticias/amsterda-sofre-com-congestionamento-acredite-de-bicicletas>

⁷ <http://g1.globo.com/rs/rio-grande-do-sul/noticia/2013/03/justica-do-rs-mantem-juri-popular-de-motorista-que-atropelou-ciclistas.html>

⁸ <http://noticias.uol.com.br/cotidiano/ultimas-noticias/2013/09/17/um-ciclista-morre-em-acidente-de-transito-por-semana-em-sao-paulo.htm>

1.3 Métodos

O experimento proposto consiste em:

- criar um conjunto de redes viárias simuladas (3 simples e 2 complexas);
- criar uma demanda de veículos que irão transitar por essas redes viárias;
- implementar os algoritmos propostos no sistema semafórico de cada rede;
- realizar diversos testes onde serão mantidos fixos alguns parâmetros e variados alguns fatores de cada algoritmo;
- analisar os resultados obtidos.

1.4 Contribuições

Ao final deste trabalho são deixadas as seguintes contribuições:

- A implementação de um algoritmo descentralizado e de comportamento adaptativo para o controle semafórico simulado experimentado e que funciona em tempo real;
- A implementação de um algoritmo de aprendizagem de máquina para o controle semafórico simulado experimentado e que funciona em tempo real;
- A implementação de um algoritmo bio-inspirado para o controle semafórico simulado experimentado e que funciona em tempo real;

1.5 Estrutura do trabalho

A partir deste ponto o trabalho está estruturado da seguinte forma: serão apresentados no Capítulo 2 os trabalhos relacionados estudados e que formam a base de conhecimento para a evolução dos temas abordados; no Capítulo 3 é apresentada a metodologia aplicada na resolução do problema, contendo as variáveis de ambiente envolvidas na simulação, a modelagem das redes viárias e sua demanda, e os algoritmos utilizados; o Capítulo 4 apresentará os experimentos realizados e os resultados obtidos por ambas as abordagens e delineará uma comparação das mesmas; no Capítulo 5 uma breve conclusão do trabalho e algumas idéias para trabalhos futuros sobre este tema são apresentadas.

2 BASE CONCEITUAL

Neste capítulo serão apresentados os trabalhos relacionados que fazem parte do estado da arte e que são base para o desenvolvimento da abordagem proposta. Este capítulo divide-se em cinco seções: na Seção 2.1 são apresentados alguns tópicos relacionados à engenharia de tráfego; a Seção 2.2 informa sobre as necessidades da simulação de tráfego; na Seção 2.3 serão apresentados alguns trabalhos sobre sistemas multiagente; na Seção 2.4 são abordados os trabalhos sobre inteligência de enxames; ao final, são apresentados na Seção 2.5 trabalhos relacionados ao uso das técnicas de aprendizagem por reforço.

2.1 Engenharia de Tráfego

Os sistemas de transporte atuais são o fruto do desenvolvimento de múltiplos ramos da Engenharia: da Civil temos as construções das obras viárias (infraestrutura) tais como ruas, viadutos, pontes; da Automotiva através do desenvolvimento de veículos mais robustos e seguros; de Tráfego pelos esforços na parte operacional do sistema, buscando um equilíbrio entre a demanda e a capacidade; e da Elétrica pelos esforços em criar mecanismos de controle mais eficientes e automatizados, unida a essa podemos adicionar a da Computação, por prover sistemas computadorizados mais robustos ao controle de todo o sistema.

Neste trabalho estaremos dando um enfoque maior à parte do controle de tráfego através da melhoria de semáforos num ambiente simulado.

Problemas de mobilidade em nossa sociedade são muito antigos. A necessidade de transporte de produtos e de pessoas é necessária, e condições como o aumento da população, produção e urbanização das cidades levaram seus responsáveis a incluir formas de controle a esse sistema complexo.

Questões relativas ao excesso de tráfego e congestionamentos datam desde o Império Romano levando ao banimento da circulação de carroças nos centros urbanos¹. Nos dias atuais esse tipo de atitude também ocorre², a cidade de Porto Alegre também se encaminha neste mesmo sentido³.

Atualmente a complexidade desses sistemas aumentou consideravelmente, e com isso

¹ <http://www.transitocomvida.ufrj.br/HistoriaDoTransitoNoMundo.asp>

² <http://www.estadao.com.br/noticias/cidades,rodizio-de-carros-sera-ampliado-em-240-quilometros-em-sao-paulo,1034988,0.htm>

³ <http://www.cidademaishumana.com.br/site/2013/09/projeto-limita-circulacao-de-veiculos-no-centro-historico/>



(a) *Delancey Street, Nova York, EUA (1900).* (b) *Shirley Highway, Virgínia, EUA (1950).* (c) *Zona Sul, São Paulo, Brasil (2012).*

Figura 2.1: Fotos de engarrafamentos, do passado até o presente.

acarretou diversos problemas, seja de ordem social ^{4, 5}, econômica ⁶ ou ambiental⁷.

Órgãos governamentais responsáveis pelo controle desse sistema tem se esforçado para a criação de normas de convivência no trânsito⁸ e de redução de poluentes⁹, além de aumentar consideravelmente os investimentos para ampliação de vias, como tem-se visto nas obras do PAC da Mobilidade Urbana no Brasil¹⁰. Por parte da população vemos diversas iniciativas para uma mudança na forma como o trânsito é visto e um fortalecimento de alternativas para o transporte tem sido sugeridas, em um movimento chamado de cidade para pessoas¹¹.

A seguir são apresentados alguns termos necessários ao entendimento do resto do texto.

Malha viária

Uma *malha viária* (ver exemplo à Figura 2.2a) é uma formação geralmente complexa que compreende a união de diversas ruas, ligando assim diversos pontos distintos em uma cidade.

Rua

Rua, avenida ou via (ver exemplo à Figura 2.2b) define o espaço público no qual é permitida a movimentação, seja por veículos ou pessoas. Uma rua é dividida em um conjunto de uma ou mais *faixas*, onde cada uma pode ter um sentido de direção do tráfego próprio.

Veículo

É todo e qualquer meio de transporte, neste trabalho estará representado especificamente pelo *carro* (ver exemplo à Figura 2.2c).

Fluxo

Um *fluxo* (ver exemplo à Figura 2.2d) representa um conjunto de veículos que se movimenta em ruas tendo pontos de partida e chegada previamente definidos.

⁴ <http://veja.abril.com.br/noticia/saude/poluicao-do-transito-eleva-risco-de-hiperatividade-infantil>

⁵ <http://noticias.terra.com.br/brasil/com-295-km-sao-paulo-registra-maior-congestionamento-da-historia,0e1ae7948c1da310VgnCLD200000bbcc0aRCRD.html>

⁶ <http://www.mobilize.org.br/estudos/78/transito-de-sp-ja-causa-perdas-de-r-50-bi-por-ano.html>

⁷ <http://noticias.r7.com/sao-paulo/poluicao-do-ar-mata-mais-do-que-acidentes-de-transito-em-sao-paulo-diz-pesquisa-20092013>

⁸ <http://www.denatran.gov.br/>

⁹ <http://www.mma.gov.br/port/conama/>

¹⁰ <http://www.pac.gov.br/cidade-melhor/mobilidade-urbana>

¹¹ <http://www2.camara.leg.br/camaranoticias/noticias/TRANSPORTE-E-TRANSITO/452610-CAMARA-CELEBRA-DIA-MUNDIAL-SEM-CARRO-COM-DEBATES-E-INCENTIVO-AO-USO-DA-BICICLETA.html>

Escolha de rota

É a capacidade de um agente escolher uma *rota* (ver exemplo à Figura 2.2e) que será traçada de seu ponto de partida até a chegada. Sua escolha afeta o fluxos dos outros veículos, em geral fatores como tempo, custos ou poluição tentam ser minimizados na escolha de rotas.

Controlador semafórico

Este termo possui vários sinônimos, como sinal de trânsito, cruzamento, junção e intersecção, nomeado assim devido a apenas um plano semafórico estar em funcionamento em um conjunto de semáforos que estejam em uma via, representando assim uma única entidade. Semáforos são controladores de tráfego que combinam hardware e software de modo a implementar um temporizador. A parte fundamental no projeto de um semáforo com temporizador é seu ciclo [Roess et al., 2004]. Um ciclo é definido como uma rotação completa por todas as indicações de verde das ruas adjacentes à junção. Geralmente todos os movimentos legais para um veículo seguir serão indicados em verde durante cada ciclo. Intervalos são os períodos de tempo durante o qual não ocorrem mudanças no sinal indicado. Os intervalos mais comuns são: intervalo de mudança (sinal amarelo), intervalo de liberação (sinal vermelho para todos), e intervalos de verde e de vermelho. Além do ciclo e dos intervalos, outro componente do temporizador é a fase do sinal. Ela consiste de um intervalo verde seguido de um intervalo de mudança e de um de liberação.

Na Figura 2.3c pode-se notar a sequência dos intervalos do ciclo para cada um dos semáforos de uma mesma intersecção.

Podemos classificar os semáforos de acordo com sua forma de operação de várias formas: baseado na sua dimensão, de tempo fixo (pré determinado) versus sensível ao tráfego (atuado); no controle isolado de intersecção versus controle coordenado; ou no controle local por microprocessador versus controle remoto por central de controle.

Quando operado como tempo fixo, o tamanho do ciclo, a sequência das fases, e a temporização de cada intervalo são constantes e seguem um plano pré definido, isso é definido desta forma pois o volume de tráfego já foi previamente computado através de dados históricos. Em uma operação parcialmente atuada, também é necessária a aquisição de dados, seja através de detectores enterrados (laços de indução) ou outros dispositivos, como câmeras de vigilância. Estes detectores são colocados apenas em lugares próximos às intersecções (não nas ruas principais). Em uma operação totalmente atuada, todas as vias possuem um detector e o tempo de sinal verde é alocado de acordo com regras específicas, dessa forma o tamanho do ciclo, a sequência de fases e as fatias de tempo do sinal verde podem variar de ciclo para ciclo. Um controlador de tempo fixo é a escolha mais acessível e lógica quando se tem uma malha viária com um comportamento do tráfego estável ou previsível.

Em sistemas gerenciados por uma central de controle, o computador é peça principal do sistema, fazendo a coordenação do tempo dos controladores de tráfego. Através de dados dos detectores ou da hora do dia, ele consegue fazer a seleção ou calcular planos de coordenação ótimos para cada intersecção. Entretanto, para que essa coordenação seja efetiva, todos os sinais precisam ter o mesmo ciclo (ou múltiplos dele), e isso acaba tornando difícil de ser mantido quando é permitido

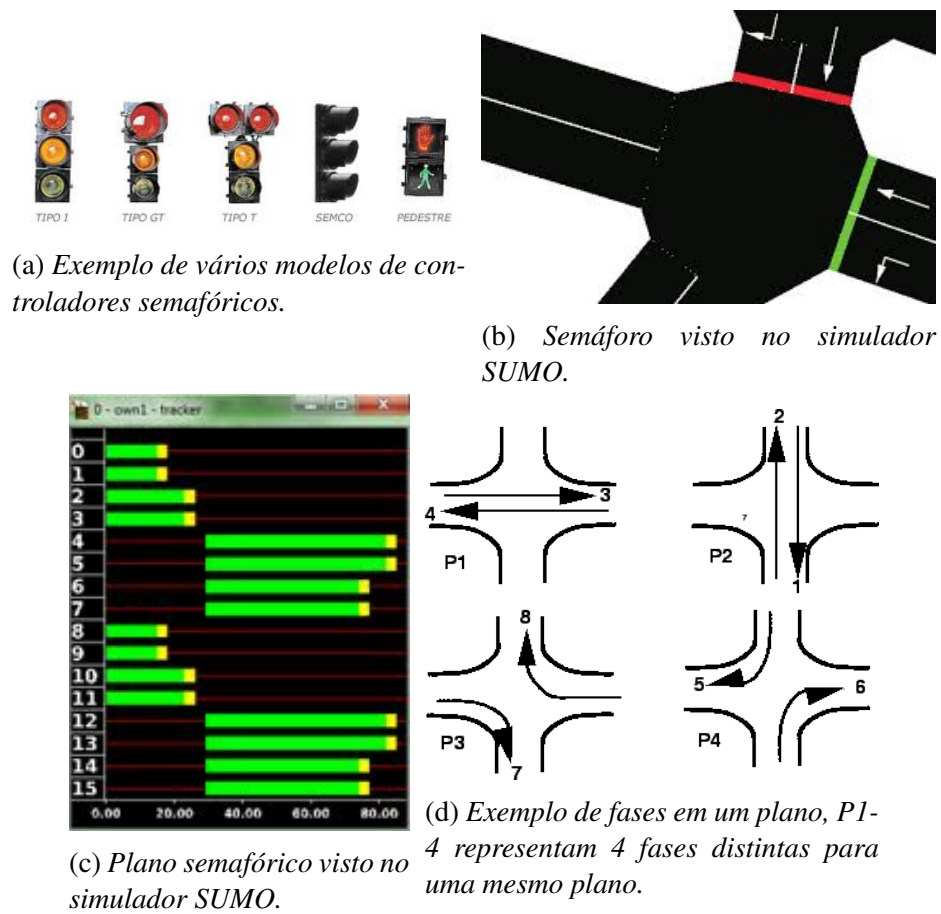


Figura 2.3: Da esquerda para a direita, exemplo de controladores semafóricos reais, um semáforo no simulador SUMO, um diagrama de um possível plano semafórico no mesmo simulador e um conjunto de fases.

2.1.1 Sistemas Inteligentes de Transporte

Esses sistemas são um novo esforço para tornar os sistemas de transporte mais eficientes, seguros e independentes através do uso da tecnologia da informação [Roess et al., 2004, Bazzan, 2012]. Algumas linhas nessa área trabalham no desenvolvimento de formas de comunicação inter-veicular¹² e autonomia dos mesmos¹³, centros de controle especializados, semáforos inteligentes^{14, 15}, entre outros. Neste trabalho estaremos focando no último caso.

Sistemas Inteligentes de Transporte (ITSs) voltados à infraestrutura de transporte e à engenharia de tráfego são conhecidos como Sistemas Avançados de Gestão de Viagens (ATMSs), e compreendem aplicações como simuladores de trânsito, semáforos inteligentes, ou ainda dispositivos de fiscalização eletrônica [Bazzan, 2012]. Já os sistemas usados para distribuir informações aos motoristas são conhecidos como Sistemas Avançados de Informação ao Viajante (ATISs), e englobam serviços como mapas online, boletins informativos de rádio, dispositivos embarcados como celulares do tipo smartphone e sistemas de navegação veicular.

Como pôde ser visto nesta seção, os sistemas de tráfego são uma questão complexa e desafiadora para muitas das áreas envolvidas, muitos dos temas relacionados e problemas em aberto não cabem aqui neste texto, mas podem ser revistos em outras fontes da literatura [Roess et al., 2004, Bazzan and Klügl, 2007, Bazzan, 2012].

2.2 Simulação de tráfego

Uma ferramenta computacional muito útil para estudos de sistemas complexos são os simuladores. Com eles é possível uma emulação da realidade através da construção do ambiente desejado, dos agentes envolvidos e das interações dos mesmos.

Os modelos de simulação podem ser divididos em três tipos primários: Microscópico, Mesoscópico e Macroscópico, e a diferença entre os mesmos está na granularidade dos detalhes. Abordaremos apenas o primeiro tipo neste trabalho e recomendamos a leitura de material elucidativo sobre os outros tipos de modelos de simulação e outros simuladores de trânsito [Kotushevski and Hawick, 2009, Hallmann, 2011].

Adotou-se neste trabalho o simulador SUMO [Krajzewicz et al., 2012], pois o mesmo se adequa as necessidades do experimento. Mais informações sobre ele serão apresentadas nas seções seguintes.

2.2.1 Modelo de simulação microscópico

Este é o modelo que permite a modelagem mais próxima do motorista no mundo real, pois permite um maior detalhamento das características do movimento de cada veículo a ser simulado. Dessa forma, ele permite, modelar por exemplo, a aceleração, desaceleração, mudanças de rota e mudanças de faixa na via para cada veículo.

Na prática, esse tipo de simulação microscópica é implementada de forma síncrona,

¹²<http://www.its.dot.gov/research/v2v.htm>

¹³<http://www.wired.com/autopia/2013/08/nissan-autonomous-drive/>

¹⁴http://www.newscientist.com/article/dn13306-could-smart-traffic-lights-stop-motorists-fuming.html#.UjysD9_LL3w

¹⁵http://www.thestar.com/news/insight/2013/06/14/can_smart_traffic_lights_ease_torontos_road_congestion.html

ou seja, cada passo de tempo explora todas as entidades no modelo e atualiza os estados das entidades de uma só vez.

A forma de modelagem mais comum desse tipo é chamada de car-following, onde um veículo que está à frente é seguido pelo que vem atrás. É um dos modelos mais antigos surgindo na década de 50 e possuindo diversas variantes [Olstam and Tapani, 2004]. Outros tipos de modelagens microscópicas são as baseadas em autômatos celulares [Nagel and Schreckenberg, 1992] e em agentes [Bazzan et al., 2010]. Daremos maior ênfase ao modelo Car-following aqui por ser o adotado no simulador de tráfego utilizado, mais informações sobre esse modelo serão descritas a seguir na Seção 2.2.2.1.

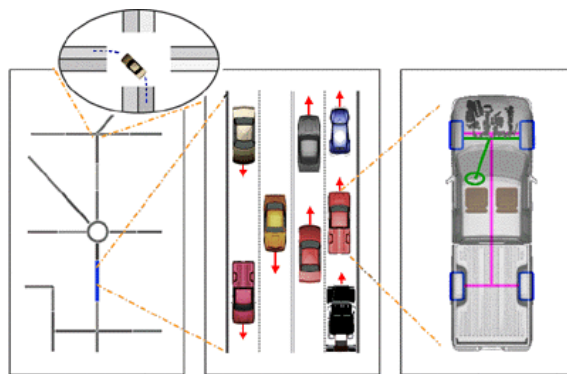


Figura 2.4: *Comparação das diferentes granularidades de simulação^a*. Da esquerda pra direita: macroscópico, microscópico, sub-microscópico (no círculo detalhe mesoscópico)

^ahttp://sumo-sim.org/userdoc/Theory/Traffic_Simulations.html

2.2.2 Simulation of Urban MObility (SUMO)

SUMO é um simulador de tráfego desenvolvido dentro do Instituto de Sistemas de Transporte da Agência Aeroespacial Alemã¹⁶. Ele possui código aberto, e portabilidade para várias plataformas, é de fácil extensão, reduzido consumo de memória e tem uma considerável velocidade de simulação.

Sua utilização pode ser feita através de uma interface gráfica ou da linha de comando. Recebe como entrada e devolve como saída arquivos no formato XML¹⁷. Possui diversas ferramentas associadas para a geração das malhas viárias, geração de demandas e controle semafórico. Possui também uma interface de controle (TraCI¹⁸) que permite acessar a simulação em tempo real, e manipular o comportamento dos objetos envolvidos.

2.2.2.1 Características de movimento

O movimento dos veículos na simulação segue o modelo microscópico car following proposto por Stefan Krauß [Krauß, 1998], onde o tempo é discreto e o espaço é contínuo. Uma inovação deste modelo é assumir que um motorista não é perfeito em realizar a velocidade desejada, pelo contrário, a velocidade realmente determinada é menor. Com este diferencial, ocorrem congestionamentos espontâneos e o efeito chamado de início lento (*slow-to-start*), características dos motoristas reais. O modelo é rápido durante a

¹⁶<http://www.dlr.de/dlr/en/desktopdefault.aspx/tabid-10002/>

¹⁷<http://www.w3.org/XML/>

¹⁸<http://sumo-sim.org/userdoc/TraCI.html>

execução, principalmente devido ao pequeno número de computações, em comparações com outros modelos. Durante o movimento o veículo deve observar a infraestrutura a sua frente e possíveis alterações na velocidade permitida, para assegurar um sistema livre de colisões.

Algumas dúvidas ainda pairam sobre esse modelo em dois pontos, o modelo de imperfeição do motorista é relativamente simples, não sendo validado contra trajetórias reais e o uso do tempo de reação do motorista de apenas 1 segundo é uma problemática conhecida em cenários densos [Barceló, 2010, capítulo 7].

O comportamento de troca de faixas (*lane-change*) pelos veículos está baseado no modelo de [Krajzewicz, 2009], o qual computa um caminho válido através da malha viária. A avaliação de mudança se dá baseada na faixa atual do veículo e nas faixas mais próximas, se ele quiser fazer a troca é necessário que ela ocorra até um limite onde a rota não seja afetada.

Para simular o comportamento de aumentar a velocidade do movimento, o motorista mantém informações sobre o benefício de trocar de faixa a cada intervalo de tempo discreto. Esse fator é a diferença entre a velocidade segura na faixa de interesse e a corrente, dividida pela velocidade máxima que o veículo pode utilizar em condições de fluxo livre. A troca de faixa então só poderá ser iniciada se este benefício alcançar um limiar discreto e claro, se houver espaço na faixa desejada. Dessa forma a simulação tem um comportamento bom para cenários urbano e de vias rápidas, assumindo que os veículos escolhem antecipadamente para onde se mover e tendem a ocupar todo o espaço disponível, entretanto ela não leva em conta o comportamento dos demais veículos, podendo causar problemas, como acidentes [Barceló, 2010, capítulo 7].

2.2.2.2 Geração da malha viária

SUMO possui diversas formas de utilização de malhas viárias, podendo importar modelos prontos de outros simuladores¹⁹, tais como VISUM²⁰, VISSIM²¹, Robocup Rescue²² e OpenStreetMaps²³ ou por geração manual que são convertidos pela ferramenta OSM²⁴ e depois pela ferramenta NETCONVERT²⁵. Além disso, permite a criação de malhas abstratas com a ferramenta NETGENERATE²⁶, nesse caso consegue construir malhas em três formatos: tipo Grade (ou *Manhattan*), Teia de aranha, e Aleatória. A Figura 2.5 apresenta alguns modelos complexos que podem ser importados para o uso no simulador, a Figura 2.6 apresenta algumas malhas que podem ser criadas pelo próprio simulador.

2.2.2.3 Controladores de tráfego

Para a inclusão de controladores semafóricos existem algumas opções:

1. durante a importação das malhas viárias, nesse caso as junções que originalmente

¹⁹ <http://sumo-sim.org/userdoc/Networks/Import.html>

²⁰ <http://vision-traffic.ptvgroup.com/en-uk/products/ptv-visum/>

²¹ <http://vision-traffic.ptvgroup.com/en-uk/products/ptv-vissim/>

²² <http://www.robocuprescue.org/>

²³ <http://www.openstreetmap.org/>

²⁴ http://sumo-sim.org/userdoc/Networks/Import/OpenStreetMap.html#Editing_OSM_networks

²⁵ <http://sumo-sim.org/userdoc/NETCONVERT.html>

²⁶ <http://sumo-sim.org/userdoc/NETGENERATE.html>

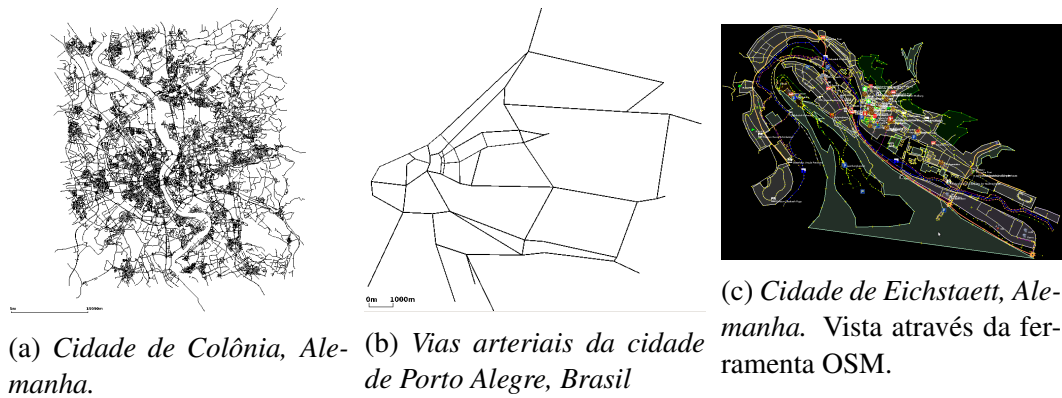


Figura 2.5: Modelos de malhas viárias complexas importadas no SUMO.

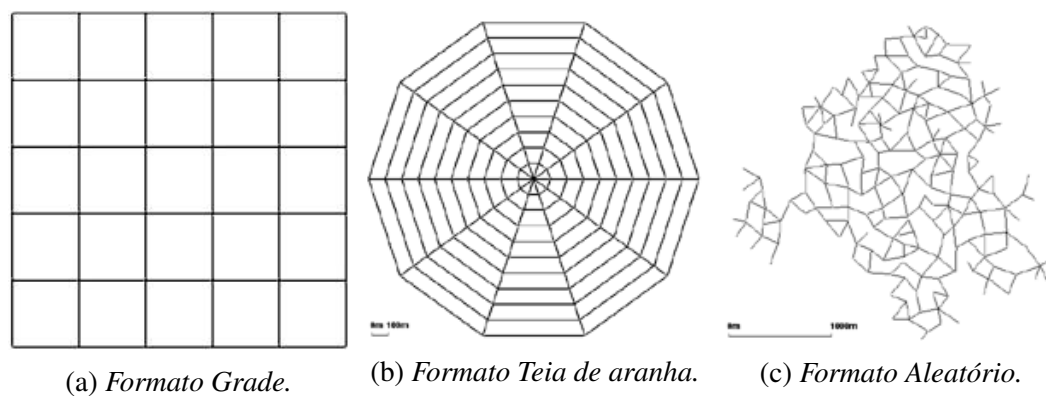


Figura 2.6: Modelos de malhas viárias abstratas possíveis de serem criadas no SUMO.

possuíam controladores terão os mesmos importados e convertidos;

2. durante a criação de malhas viárias, ao definir uma junção é possível adicionar uma referência de que a mesma possui um controlador de tráfego;
3. durante a execução da simulação, é possível incluir arquivos adicionais que contêm informações de ativação de semáforos em determinadas intersecções da rede viária utilizada.

Os semáforos definidos dessa forma, em geral não irão representar toda a complexidade de comportamento percebida em um semáforo real, principalmente se forem adotados os comportamentos padrão na importação. Duas outras opções para melhorar o comportamento dos mesmos existem, a primeira consiste na criação de programas regulares diários²⁷, permitindo assim que um semáforo alterne seus planos semaforicos durante o dia. Note que neste caso isso deve ser feito por um especialista em tráfego que irá levar em consideração o conhecimento histórico do sistema na região desse semáforo, e claro adotará nesse caso planos estáticos. Uma segunda opção seria o uso da ferramenta TraCi, pois a mesma permite a interação com qualquer elemento da simulação, incluindo os semáforos, permitindo assim trocas de fase e até de planos semaforicos completos do mesmo ou sua desativação em tempo real.

2.2.2.4 Atribuição de demanda

O simulador SUMO possui um grande número de ferramentas acopladas que permitem a atribuição de demanda veicular. Abaixo uma lista das possíveis formas de atribuição é apresentada:

Matrizes Origem-Destino (OD)

Utilizadas por algumas das ferramentas, as matrizes OD são representações na forma de tabela onde cada veículo possui um ponto de saída e outro de chegada (podendo ser uma via ou mesmo uma intersecção) e um tempo de saída.

DUAROUTER

Utilizando-se das matrizes OD a ferramenta DUAROUTER²⁸ gera uma lista de rotas (trajeto a ser percorrido) para cada veículo através de um algoritmo de caminho mínimo.

OD2TRIPS

Com OD2TRIPS²⁹ é possível converter uma matriz OD em viagens. Com ela podemos distribuir os veículos uniformemente durante o tempo definido na matriz OD ou definir um intervalo padrão para a distribuição dos mesmos (por exemplo, incluir um novo carro na simulação a cada 5 segundos). Além disso, com essa ferramenta podemos fatiar o tempo de demandas muito grandes (por exemplo escolher somente a demanda entre 6h e 8h).

²⁷ http://sumo-sim.org/userdoc/Simulation/Traffic_Lights.html#Defining_Program_switch_Times_and_Procedure

²⁸ <http://sumo-sim.org/userdoc/DUAROUTER.html>

²⁹ <http://sumo-sim.org/userdoc/OD2TRIPS.html>

Fluxos

Pensando num modelo de interação de partículas, é possível atribuir fluxos de veículos através de rotas fixas, agregando assim um número maior de veículos em determinada região de interesse com uma demanda uniforme³⁰.

DUA

Atribuição Dinâmica de Usuário/Demanda (DUA³¹) é uma ferramenta do simulador que busca encontrar um equilíbrio da simulação, ou seja, um momento onde nenhuma modificação na rota de cada um dos veículos envolvidos melhore o seu desempenho (reduza seu tempo de viagem em geral).

JTRROUTER

Através da ferramenta JTRROUTER³² é possível adicionar valores de probabilidade (variados) que permitem ao veículo fazer um desvio em cada intersecção da malha viária. Com isso, um veículo que faça sua viagem através da malha poderá alterar por quais vias fará seu trajeto a cada nova intersecção. Como essa forma estocástica não é muito precisa, pode ocorrer de determinados veículos ficarem presos em laços (andem em círculos).

Demanda Baseada em Atividade

Através da descrição de uma população em determinada área. A atribuição de demanda baseada em atividade³³, permite uma simulação ainda mais próxima da realidade, podendo incluir populações e interesses, como atividades de trabalho, escolas, bicicletas, ônibus e pessoas caminhando.

DFROUTER

Com a ferramenta DFROUTER³⁴, são definidas fontes e destinos de veículos, dessa forma os veículos fazem rotas aleatórias dentro da malha viária e saem da mesma ao encontrar um destino. Simulações com esse tipo de demanda podem não ter fim, ou mesmo ser incompatíveis com a realidade.

Roteamento Automático por Dispositivo

Roteamento automático por dispositivo³⁵, nessa forma de demanda alguns veículos (ou mesmo todos, de acordo com a probabilidade escolhida) possuirão um aparelho que permite o re-roteamento do veículo em tempo de simulação. A idéia é simular um veículo que possui um aparelho de GPS [Hofmann-Wellenhof et al., 1997] e um motorista que busca melhorar seu desempenho (chegar ao destino mais rápido).

2.3 Sistemas Multiagente

Estudos envolvendo sistemas multiagente vem sendo usados a algum tempo, mesmo assim, alguns pesquisadores da área da computação ainda não entraram em consenso sobre o que exatamente define um agente [Wooldridge, 2009]. Neste trabalho considero

³⁰ http://sumo-sim.org/userdoc/Demand/Shortest_or_Optimal_Path_Routing.html

³¹ http://sumo-sim.org/userdoc/Demand/Dynamic_User_Assignment.html

³² <http://sumo-sim.org/userdoc/JTRROUTER.html>

³³ http://sumo-sim.org/userdoc/Demand/Activity-based_Demand_Generation.html

³⁴ <http://sumo-sim.org/userdoc/DFROUTER.html>

³⁵ http://sumo-sim.org/userdoc/Demand/Automatic_Routing.html

que: um sistema multiagente pode ser percebido como um grupo de agentes (atores) que interagem uns com os outros, percebendo e agindo no ambiente em que se situam.

Ao criar um agente é possível definir um comportamento próprio para o mesmo, entretanto é uma tarefa difícil e indesejada pois na maioria das vezes o ambiente em que eles se encontram terá um comportamento dinâmico (outros agentes estarão alterando-o), dessa forma uma solução é permitir que os agentes possam adaptar-se ao meio (capacidade de aprender)[Busoniu et al., 2008, Bazzan, 2009].

A tarefa de aprendizagem para apenas um agente já vem sendo estudada a bastante tempo, vários algoritmos já foram propostos, porém para um sistema multiagente diversos desafios ainda estão em aberto, pois para um agente é necessário a adaptação tanto ao ambiente, quanto aos outros agentes que agem simultaneamente. Uma opção nesse caso é ignorar as ações tomadas pelos outros agentes na sua própria aprendizagem, porém isso não é realístico quando tratamos de sociedades, além disso, os agentes podem acabar se limitando a obter resultados não ótimos se comparados com um resultado global do sistema que poderia ser alcançado.

Sistemas de trânsito encaixam-se assim como uma ótima opção de testes de metodologias sobre sistemas multiagentes, pois possuem uma alta complexidade e interação de seus agentes no ambiente. Vários trabalhos vem sendo desenvolvidos com esse foco, utilizando-se de por exemplo, técnicas de aprendizagem [Tavares and Bazzan, 2012], mediação [Oliveira et al., 2005b, Oliveira, 2005], comunicação inter-veicular[Tavares and Bazzan, 2013] e abordagens bio-inspiradas [Oliveira et al., 2004, 2005a].

2.4 Inteligência de Enxame

Inteligência de enxame é uma técnica que se baseia no comportamento coletivo de grupos de agentes (geralmente insetos), e que nos permite contemplar comportamentos emergentes na ausência de um controle central. Originou-se em trabalhos da área de robótica [Beni and Wang, 1993] onde foram notados padrões de ordenação não previstos.

Uma das grandes vantagens encontradas em espécies que vivem em enxames é a sua capacidade de dividir e alocar as tarefas de forma dinâmica entre seus membros, visto que diversos fatores afetam esse tipo de sociedade. Tais como: variações no tamanho da colônia, instabilidade das fontes alimento, ação de predadores, mudanças climáticas.

Por consequência disso, as colônias alteram seu comportamento de forma a ajustar-se as condições do ambiente, aumentando ou diminuindo a quantidade de servidores envolvidos em cada tarefa.

Diversos pesquisadores vem trabalhando sobre este tema, seja na ordem biológica[Gordon, 2002] ou computacional [Bonabeau et al., 1999], através da geração de modelos heurísticos.

Cada indivíduo dentro da colônia possui uma tendência a fazer determinadas tarefas, e em geral essa tendência muda conforme a idade do mesmo, permitindo que ocorra uma melhor resposta às necessidades da colônia.

O modelo adotado neste trabalho [Oliveira et al., 2005a, Oliveira, 2005] pode ser visto na Seção 3.3 onde trato da metodologia empregada na resolução do problema.

2.5 Aprendizagem por Reforço³⁶

Estudos voltados para a Aprendizagem por Reforço (AR) surgiram a muito tempo, porém com vieses nas áreas de estatística, psicologia e neurociências. Na última década, essa área tem atraído um interesse crescente da comunidade de Inteligência Artificial (IA), principalmente na de Aprendizagem de Máquina (AM). Isso se deve a esta abordagem oferecer a oportunidade de um agente resolver tarefas sem nenhuma especificação de como fazê-lo, sendo levado aos resultados através de pequenas recompensas ou punições. Podemos dizer que, a AR é o processo pelo qual um agente aprende um comportamento através de interações do tipo "tentativa e erro" em seu ambiente.

Em geral os problemas de busca e planejamento em IA estão fortemente ligados a capacidade do agente aprender em um ambiente desconhecido. Algoritmos de busca descobrem trajetórias satisfatórias em grafos de estados, enquanto que técnicas de planejamento operam tipicamente através de construção de estruturas de dados mais complexas, utilizando, por exemplo, composição de expressões lógicas para representar os estados. Esses tipos de algoritmos são de mais alto nível, se comparados com a AR pois requerem um modelo pré-definido de mundo (conhecimento total do ambiente), em geral determinístico. A AR, por outro lado, é dita Livre de Modelo (LM), pois não assume a existência de um modelo explícito de funcionamento do mundo no qual atua.

As principais estratégias para a resolução de problemas de AR são fazer uma busca no espaço de comportamentos possíveis de maneira a encontrar o que melhor se adapta ao ambiente (semelhante ao usado nos algoritmos genéticos e na computação evolutiva), usar técnicas estatísticas e de programação dinâmica para estimar a utilidade de se tomar ações em determinados estados do ambiente. A maior parte dos trabalhos que vêm sendo desenvolvidos utilizam essa segunda abordagem.

Em um modelo padrão de AR o agente está conectado ao seu ambiente através de percepções e ações. A cada passo de interação, este agente recebe como entrada uma indicação do estado atual do ambiente, com base na qual decide e efetua uma ação. Cada ação modifica o estado atual do ambiente. A transição para o novo estado gera um valor escalar que é comunicado ao agente na forma de um sinal de reforço (recompensa/punição). O agente baseia assim seu comportamento na escolha de ações que maximizem a soma esperada das recompensas ao longo do tempo.

Um modelo formal de AR consiste de três elementos, um conjunto discreto de estados do ambiente; um conjunto discreto de ações do agente; e um conjunto de sinais escalares de reforço, tipicamente o conjunto dos números reais.

O agente pode perceber o ambiente de duas formas, através de uma função identidade (o valor do ambiente é retornado corretamente), nesse caso sendo chamado de ambiente completamente observável ou através de uma função que retorne resultados imprecisos, sendo chamado assim de ambiente parcialmente observável.

Através das percepções do seu estado e das recompensas imediatamente recebidas, o agente busca encontrar uma política que mapeie estados para ações de forma a maximizar a medida de recompensa a longo prazo. Na maior parte das vezes considera-se que o ambiente é não determinístico, dessa forma uma mesma ação pode gerar resultados diferentes em ocasiões diferentes. Porém, assumimos que o ambiente é estacionário, o que equivale a dizer que, embora seja probabilístico, o conjunto de probabilidades que guia as

³⁶ Devido a clareza e concisão do texto, esta seção foi fortemente baseada no Trabalho Individual I de Bruno Silva [Silva, 2005].

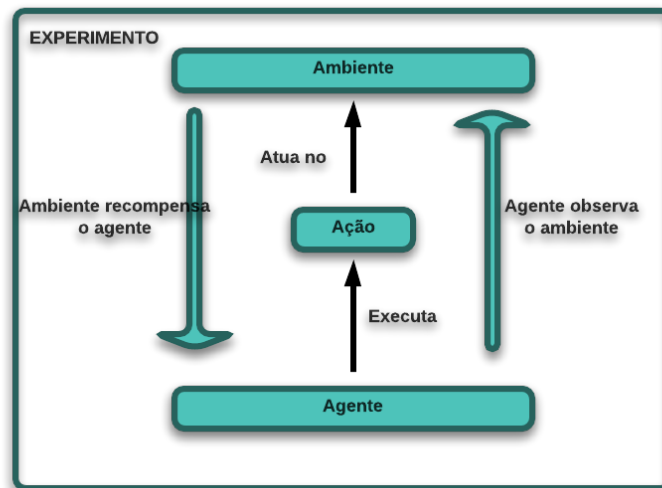


Figura 2.7: Arquitetura de um sistema de aprendizagem por reforço monoagente.

transições de estado não é alterado ao longo do tempo.

2.5.1 Processo de Decisão de Markov

Os Processo de Decisão de Markovs (PDMs) são particularmente importantes para a AR pois seu ambiente é modelado como um PDM.

Um PDM é um processo estocástico caracterizado por um conjunto de estados, ações e matrizes de probabilidade de transição entre os estados. Essa transição entre os estados ocorre de maneira discreta no tempo e depende de uma ação. Além disso, a cada estado é associada uma função de recompensa, a qual define o ganho instantâneo que o sistema recebe por estar situado naquele estado.

Os PDMs em geral assumem que a Hipótese de Markov é válida, isto é o mesmo que dizer que um sistema é markoviano caso a "distribuição de probabilidades condicionais para os estados futuros do processo dependa apenas do estado atual". Isto significa que os estados futuros do sistema são condicionalmente independentes dos estados passados.

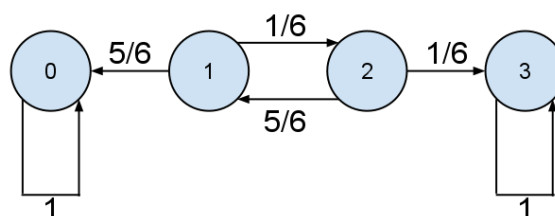


Figura 2.8: Exemplo de cadeia de Markov. Número em cada nodo representa o estado, número em cada aresta representa a probabilidade de mudar para o estado apontado.

Para resolver um PDM basta encontrar uma sequência de ações que garanta o maior ganho esperado para o sistema, dado um determinado estado inicial. Uma política π é a especificação de tal sequência de ações na forma de um mapeamento *estado* \rightarrow *ação*. Denotamos $\pi(s)$ como a ação recomendada dado que o sistema esteja no estado s e π^* como a política ótima, ou seja, a política que gera o maior ganho esperado.

A dificuldade para encontrar uma política ótima piora quanto mais vezes for utilizado o fator de desconto γ , o qual faz com que os ganhos futuros do sistema tenham seu valor diminuído de forma que o sistema tenda a preferir recompensas imediatas (torna-se guloso).

Existem dois algoritmos tradicionalmente usados para resolver um PDM: o algoritmo de *iteração de valor* e o de *iteração de política*. Ambos fazem uso da chamada Equação de Bellman possibilitando assim obter uma estimativa do possível ganho que o sistema receberá das ações feitas a partir de um determinado estado. Esse ganho é chamado de *utilidade de estado* e é usualmente denotado por U .

A Equação de Bellman consiste na formalização matemática da seguinte afirmação: "*a utilidade de um estado é a recompensa imediata correspondente a este estado somada à utilidade descontada esperada do próximo estado, supondo-se que o agente escolha a ação ótima*". E é definida por:

$$U(s) = R(s) + \gamma \max_a \sum_{s'} T(s, a, s') U(s') \quad (2.1)$$

onde $U(s)$ é a utilidade esperada no estado s , $R(s)$ é a recompensa imediata em s , γ é o fator de desconto e $T(s, a, s')$ é a probabilidade de ir para o estado s' , dado que se efetuou a ação a estando no estado s . Note que o cálculo de $U(s)$ através da Equação de Bellman exige a resolução de um sistema de equações cujas variáveis são os valores de utilidade de cada um dos estados.

O algoritmo de iteração de valor resolve iterativamente um sistema de equações de forma a estimar progressivamente o valor de utilidade para um conjunto de estados. Por sua vez, o algoritmo de iteração de política consiste em um processo que alterna o cálculo das utilidades dos estados sob a política atual com o aperfeiçoamento da política atual em relação às utilidades atuais [Russell and Norvig, 2004].

Embora as técnicas de AR possam se valer do formalismo dos PDMs, há uma diferença essencial: tipicamente, no último conhece-se tanto o modelo completo do ambiente quanto a função de recompensa. Porém na AR, não se supõe o conhecimento de nenhum desses dois componentes. Essa suposição pode ser útil em situações onde é difícil fornecer avaliações precisas e consistentes para um grande número de situações.

2.5.2 Aprendizagem por reforço passivo

Nesse modelo a política π que o agente irá seguir é fixa: em determinado estado, o agente sempre efetuará a mesma ação. Com essa abordagem, é possível aprender o quão boa a política é, ou seja, aprender um valor de utilidade para cada estado. Essa tarefa é semelhante ao passo de *avaliação de política* do algoritmo de iteração de política sobre Cadeias de Markov. O diferencial é que na AR, o agente não conhece *a priori* o modelo de transição entre estados. Além disso, a função de recompensa também não é conhecida.

O agente poderia executar um conjunto de explorações no ambiente usando sua política fixa π . Em cada uma destas explorações, o agente percebe uma sequência de recompensas imediatas R_i relacionadas a cada um dos estados E_j que são alcançados.

A partir desse conjunto de observações, o agente deve ser capaz de aprender os valores de utilidade para cada um dos estados do sistema. Existem três métodos principais

para aprendizagem por reforço passivo: estimativa de utilidade direta, programação dinâmica adaptativa e aprendizagem de diferença temporal. Estes métodos serão descritos nas seções a seguir.

2.5.2.1 Estimativa de utilidade direta

Este método [Widrow, 1962] calcula sucessivas aproximações para as utilidades a partir de uma sequência de explorações no espaço de estados. A idéia é que a utilidade de um estado é a recompensa total esperada desse estado em diante, e portanto cada teste (exploração) fornece uma amostra desse valor para cada estado visitado.

Essa estimativa nada mais é do que um processo de aprendizagem supervisionada, onde cada exploração pode ser vista como um pareamento que relaciona estados visitados com recompensas. Com um número grande o suficiente de explorações a amostra média convergirá para a verdadeira função de utilidade, dada pela seguinte equação:

$$U^\pi(s) = \sum_{t=0}^{\infty} \gamma^t R(s_t) | \pi, s_0 = s \quad (2.2)$$

Nessa equação, $U^\pi(s)$ é a utilidade esperada para o estado s , $R(s_t)$ é a recompensa obtida pelo estado s no tempo t e γ é um fator de desconto.

A maior limitação desse método está no fato que ele não leva em conta que os vários estados do sistema não são independentes. Isso significa que a utilidade, conforme definida pela equação de Bellman para uma política fixa, é igual à recompensa daquele estado somada à utilidade esperada de seus estados sucessores:

$$U^\pi(s) = R(s) + \gamma \sum_{s'} T(s, \pi(s), s') U^\pi(s') \quad (2.3)$$

onde $U^\pi(s)$ é a utilidade do estado s dada a política π .

Isso significa que o método da estimativa direta não tira proveito das conexões entre os estados: se sabemos que um estado s leva a estados de alta utilidade, então deveríamos ser levados a crer que o próprio estado s também tende a ter alto grau de utilidade. Ao ignorar este tipo de conexão, o método de estimativa direta acaba fazendo uma procura em um espaço de hipóteses muito maior do que o necessário, uma vez que considera hipóteses de valores de utilidade que violam as equações de Bellman. Por essa razão, este método frequentemente converge de maneira muito lenta.

2.5.2.2 Programação dinâmica adaptativa

A Programação dinâmica adaptativa (PDA) é uma técnica capaz de aprender modelos de transição do ambiente a medida em que as interações ocorram a partir da resolução do PDM correspondente.

Num ambiente totalmente observável essa aprendizagem dos modelos é relativamente simples, podendo-se utilizar uma tabela que mantenha as probabilidades de transição de um estado para o outro, por determinada ação. Essa forma de mapeamento é a de um problema de aprendizagem supervisionada onde uma entrada é um pareamento de estado-ação e o resultado esperado é um novo estado. Caso seja mantido um contador para a

frequência com que cada estado é alcançado é possível estimar a probabilidade de transição correspondente.

Em se tratando de um modelo de aprendizagem passiva, a política π é fixa. Iterativamente o agente aprende uma estimativa da função de transição de estados e , a partir desse conhecimento e da política pré-definida, calcula uma nova aproximação para a função de utilidade. Essa aproximação é implementada através da resolução de um sistema de equações de Bellman, que por ter uma política fixa, é linear.

2.5.2.3 Diferença temporal

A aprendizagem de Diferença temporal (DT) tenta utilizar o melhor da PDA e da estimativa de utilidade direta. Basicamente, busca uma aproximação das equações de Bellman sem ter que resolvê-las para todos os estados possíveis. Isso é feito através do ajuste dos valores de utilidade dos estados observados, após cada transição, de forma que eles concordem com as equações de restrição de Bellman. O valor da utilidade de cada estado é então atualizado por uma fração entre sua utilidade atual e o valor que seria recebido do próximo estado, o método então está construindo progressivamente uma estimativa da capacidade que cada estado possui em levar a situações com uma boa recompensa esperada.

Dado que a restrição $U^\pi(s) = R(s) + U^\pi(s')$ deve ser obedecida, podemos observar que da transição $s \xrightarrow{a} s'$, $U^\pi(s)$ é atualizado na direção de $U^\pi(s) + R(s)$. Tornando essa atualização proporcional à diferença de utilidades entre estados sucessivos, obtemos a seguinte regra:

$$U^\pi(s) \leftarrow U^\pi(s) + \alpha(R(s) + \gamma U^\pi(s') - U^\pi(s)), \quad (2.4)$$

onde α é uma taxa de aprendizagem. Essa equação move $U^\pi(s)$ na direção do valor dado pela equação de Bellman para políticas fixas, mas leva em conta apenas o estado sucessor observado (s'), enquanto que a equação original prevê a contribuição de todos os estados seguintes possíveis.

Mesmo que seja um método aproximativo, o uso dessa regra de atualização alcança um valor médio de $U^\pi(s)$ que tende ao valor correto conforme o estado s é visitado várias vezes. Entretanto essa convergência é mais lenta do que a obtida com o algoritmo de PDA, a vantagem da DT está em ser muito mais simples, uma vez que não exige um modelo de transições e consegue aproximar os vetores de utilidade conforme os estados vão sendo visitados.

2.5.3 Aprendizagem por reforço ativo

O problema da aprendizagem por reforço ativo não consiste apenas da descoberta de uma função de utilidade, mas também na escolha de uma política de ações que maximize a recompensa esperada no futuro.

2.5.3.1 Agente guloso e a falta de exploração

Se desejarmos projetar um agente que aprenda como agir, podemos partir de um modelo baseado na PDA. O agente irá aprender um modelo completo de probabilidades para

em seguida escolher a ação que leve ao estado de maior utilidade. As utilidades podem ser calculadas pela equação de Bellman para o caso genérico (2.1).

Nesse caso podemos notar que o operador de maximização constitui uma não-linearidade, o que dificulta a resolução deste sistema de equações. Geralmente, a obtenção dos valores de utilidade se dá através do uso de um algoritmo padrão para o PDM, como o algoritmo de iteração de valor ou de iteração de política.

Um agente que executa imediatamente a política que leva ao estado de maior utilidade é chamado de agente guloso, pois ele age de acordo com o modelo de ambiente aprendido a cada etapa. O problema é que o modelo de ambiente de cada etapa pode não corresponder ao ambiente real, e portanto, a política ótima em cada iteração pode também não ser ótima no mundo real. Por isso, um agente guloso muito raramente consegue agir de acordo com a política ideal, e pode até mesmo ficar preso em políticas péssimas. Em outras palavras, um agente guloso é conservador no sentido em que não explora o ambiente além das ações que considera boas naquele momento.

Existe então a necessidade de se encontrar uma política ótima de exploração que represente o quanto o agente deve se arriscar em novos estados. Esse problema decisório é de difícil resolução exata. Porém, podemos modificar a equação de Bellman a fim de definir a utilidade otimista. Uma utilidade otimista consiste na adição de um termo que pondera positivamente a expectativa de estados poucos explorados. O agente então terá mais curiosidade em visitar novos estados que podem ser fontes de uma boa utilidade. Essa simples modificação na equação de Bellman já garante uma rápida convergência para políticas bastante próximas da ótima.

2.5.3.2 Aprendizagem Q

Essa técnica consiste em uma adaptação da DT para o caso de não possuímos uma política de ação fixa. Através dela o agente aprende uma representação ação-valor, chamada de *valor Q* , ao invés de aprender diretamente as utilidades. Podemos dizer que existe uma relação entre o valor de Q e o valor de U da seguinte forma:

$$U(s) = \max_a Q(a, s) \quad (2.5)$$

onde $Q(a, s)$ representa o valor da execução da ação a no estado s . Q também é conhecida como tabela- Q . A grande vantagem de se armazenar informações de utilidade na forma de valores Q é que o aprendizado passa a não depender da existência de um modelo do mundo. Por essa razão, costuma-se dizer que a aprendizagem Q é *livre de modelo*.

De forma a aprender as ações a serem tomadas, sem ter conhecimento do modelo, é feita uma relação entre o valor de Q para determinado estado com o valor para os estados vizinhos, como visto nas técnicas de Diferença temporal (DT). A equação resultante é:

$$Q(a, s) \leftarrow Q(a, s) + \alpha(R(s) + \gamma \max_{a'} Q(a', s') - Q(a, s)) \quad (2.6)$$

Essa equação é calculada sempre que a ação a for executada no estado s e levar ao estado s' . Um fator que deve ser levado em conta é que esse método necessita de um maior tempo para convergir por não possuir um modelo explícito de probabilidades de transição, principalmente em ambientes muito complexos.

2.5.3.3 *Detecção de contexto*

Em ambiente dinâmicos e não estacionários, como o tráfego, uma solução possível mas custosa para o sistema, seria manter múltiplos modelos de ambientes e suas respectivas políticas. Mesmo o uso de modelos parciais não é recomendado, pois estas abordagens requerem um número fixo de modelos, ou seja, assumem implicitamente que o número aproximado das diferentes dinâmicas do ambiente seja conhecida à priori.

Uma alternativa é a construção incremental de novos modelos [Silva et al., 2006a,b,c]. Nesse tipo de abordagem, assume-se algumas coisas como: as mudanças no ambiente são restritas a um pequeno número de contextos (padrões de tráfego) que são ambientes estacionários e que possuem dinâmicas distintas; o contexto atual não pode ser diretamente observado, mas pode ser estimado de acordo com os tipos de transição e recompensas observadas; as mudanças no contexto do ambiente são independentes da ação do agente; e que mudanças no contexto são geralmente infrequentes.

Num cenário de tráfego, as suposições anteriores significam que padrões de fluxo são não estacionários mas eles podem ser divididos em dinâmicas estacionárias que não precisam ser conhecidas a priori. É fato que um aspecto interessante do método é exatamente sua capacidade de particionar automaticamente as dinâmicas do ambiente em modelos parciais relevantes.

Para cada modelo é atribuída uma política ótima (que é mapeada dos padrões de tráfego para os planos semaforicos), e serve para traçar uma predição do erro das transições e recompensas, objetivando estimar a qualidade de dado modelo parcial. A criação de novos modelos é controlada por uma avaliação contínua da predição de erro gerada por cada modelo parcial. Um modelo parcial contém transições estimadas e funções de recompensa estimadas.

Para cada modelo parcial, métodos clássicos baseados em AR podem ser usados para computar localmente uma política ótima. Se o ambiente muda e uma política se torna sub-ótima (engarrafamento aumenta além do limite), então o sistema cria um novo modelo. Sempre que possível, o sistema irá reusar modelos existentes ao invés de criar novos. Novos modelos são criados apenas quando não existem modelos com um valor de erro menor que um limiar definido. Resultados dessa abordagem mostram que o mecanismo de AR com detecção de contexto é mais eficiente que uma estratégia gulosa e que outras abordagens de AR baseadas em modelo.

2.6 **Considerações finais**

Este capítulo apresentou uma visão geral sobre diversos tópicos interessantes. Através deles é possível perceber a necessidade da adoção de sistemas mais inteligentes para o controle do sistema de tráfego, evitando assim uma sobrecarga aos controladores, aumentar o desempenho e a segurança do mesmo. Também foram comentadas as vantagens em utilizar simuladores, principalmente quando se trata de um sistema complexo. As potencialidades do uso de técnicas de sistemas multiagente para a representação sociedades de agentes. A capacidade de autocontrole de sistemas emergentes e as dificuldades em aprender sem modelos de mundo previamente definidos.

3 METODOLOGIA E ALGORITMOS

Este capítulo descreve a metodologia empregada na resolução do problema de trânsito. Serão apresentadas na Seção 3.1 as características relacionadas à máquina utilizada na execução dos experimentos e ao ambiente de simulação. Na Seção 3.2 são apresentados os passos para a geração dos mapas viários utilizados no simulador e suas demandas. Na Seção 3.3 a forma como a abordagem bio-inspirada resolve o problema, os parâmetros e o algoritmo utilizados nos experimentos, na Seção 3.4 a forma como a técnica da Aprendizagem Q resolve o problema, os parâmetros e o algoritmo utilizados nos experimentos, e na Seção 3.5 a forma como a abordagem híbrida proposta (SA-CD) resolve o problema, os parâmetros e o algoritmo utilizados nos experimentos.

3.1 Ambiente de testes

3.1.1 Hardware e software utilizado

Para todos os testes executados foi utilizado o mesmo computador, ele possui as seguintes características de hardware: Processador Intel®Core™i7-3770 de 64 bits com velocidade de 3.40GHz, 16GB de memória RAM com velocidade de 1600 MHz.

O sistema operacional utilizado foi o Ubuntu 12.04 LTS de 64 bits (Precise Pangolin)¹ com Linux kernel na versão 3.5.0-41. Alguns scripts foram desenvolvidos para automatizar determinadas tarefas, eles foram desenvolvidos com a linguagem de programação Python²[Rossum, 1995] na versão 2.7.3.

3.1.2 Simulador e simulação

O simulador utilizado foi o SUMO descrito na Seção 2.2. Foi utilizada a versão 0.18 por possuir diversas melhorias de desempenho em relação as versões anteriores.

3.1.2.1 Parâmetros de entrada

Cada simulação recebe os seguintes parâmetros ao ser executada:

- um arquivo contendo o mapa viário
- um arquivo contendo as rotas dos veículos

¹ <http://www.ubuntu.com/>

² <http://www.python.org/>

- uma porta de conexão remota (utilizada pela ferramenta TraCi), de tipo inteiro
- o nome de um arquivo de saída para o sumário da simulação, de tipo string
- o nome de um arquivo de saída para as informações das viagens dos veículos, de tipo string

3.1.2.2 Número de execuções e iterações

Como lidamos neste trabalho com algoritmos estocásticos é necessária a execução de várias vezes a mesma simulação, de forma a garantir uma aproximação média dos valores reais esperados. Por tanto, cada teste é executado 10 vezes com os mesmos parâmetros para os algoritmos propostos (os parâmetros estão descritos no Capítulo 3). Foi executada uma única simulação utilizando-se plano fixo para cada rede, servindo assim como um comparador para os outros experimentos. A simulação que utiliza intersecções sem controle semaforico (irreal) foi incluída apenas para observação do comportamento do sistema, pois sua comparação com redes totalmente controladas não seria factível.

O número de iterações (passos de tempo) de cada simulação são variáveis de acordo com a demanda inserida no mapa viário e com suas condições de fluxo, por tanto não são impostos limites ao número de iterações, e sendo assim, cada simulação termina quando todos os veículos que foram inseridos na simulação terminam suas viagens.

3.1.2.3 Descrição das características das vias

Cada via em uma malha viária, possui uma descrição em formato XML das suas características. Ela pode conter uma ou mais faixas e ter ou não ligação a um semáforo. Abaixo é apresentado um exemplo fictício de como definir uma via (*edge*) com uma faixa (*lane*). Mais informações podem ser obtidas em http://sumo-sim.org/userdoc/Networks/SUMO_Road_Networks.html

```
<edge id="<ID>" from="<FROM_NODE_ID>" to="<TO_NODE_ID>" priority="<PRIORITY>">
  <lane id="<ID>" index="0" speed="<SPEED>" length="<LENGTH>"
    shape="0.00,495.05 248.50,495.05"/>
</edge>
```

Figura 3.1: Exemplo de descrição de via.

3.1.2.4 Descrição das características dos veículos e rotas

Cada veículo na simulação é uma entidade independente, visto que a simulação é de nível microscópico. Em geral ele possui um tipo (*Vtype*) que define suas características, e uma rota (*route*) que define seu trajeto na malha viária. Abaixo é apresentado um exemplo fictício de como definir um veículo (*vehicle*) padrão. Mais informações podem ser obtidas em http://sumo-sim.org/userdoc/Definition_of_Vehicles,_Vehicle_Types,_and_Routes.html

```

<routes>
  <vType id="type1" accel="0.8" decel="4.5" sigma="0.5" length="5" maxSpeed="70"/>
  <vehicle id="0" type="type1" depart="0" color="1,0,0">
    <route edges="beg middle end rend"/>
  </vehicle>
</routes>

```

Figura 3.2: Exemplo de descrição de veículo e rota.

3.1.2.5 Descrição das características das intersecções e junções internas

Uma malha viária possui além de suas vias, as conexões entre as mesmas, chamadas intersecções, e entre cada faixa essas conexões são chamadas de junções internas, as quais informam os movimentos válidos permitidos para os veículos que transitam de uma faixa para outra. Abaixo é apresentado um exemplo fictício de como definir uma intersecção (*junction*) e uma junção interna (*connection*) padrão. Mais informações podem ser obtidas em http://sumo-sim.org/userdoc/Networks/SUMO_Road_Networks.html

```

<junction id="<ID>" type="<JUNCTION_TYPE>" x="<X-POSITION>" y="<Y-POSITION>"
  incLanes="<INCOMING_LANES>" intLanes="<INTERNAL_LANES>"
  shape="<SHAPE>">
</junction>

```

Figura 3.3: Exemplo de descrição de intersecção.

```

<connection from="<FROM_EDGE_ID>" to="<TO_EDGE_ID>" fromLane="<FROM_LANE_INDEX>"
toLane="<TO_LANE_INDEX>" via="<VIA_LANE_ID>" tl="<FROM_EDGE_ID>"
linkIndex="12" dir="r" state="o"/>

```

Figura 3.4: Exemplo de descrição de junção interna.

3.1.2.6 Descrição das características dos semáforos

Um semáforo é uma intersecção especial da malha viária com comportamento que afeta o trânsito dos veículos. Ele afeta o sistema através de uma lógica própria, o plano semafórico, o qual possui diversas fases formadoras do ciclo semafórico. Abaixo é apresentado um exemplo fictício de como definir uma junção do tipo semáforo (*type traffic_light*) e um plano semafórico (*tlLogic*) padrão. Mais informações podem ser obtidas em http://sumo-sim.org/userdoc/Simulation/Traffic_Lights.html

```

<junction id="<ID>" type="traffic_light" x="<X-POSITION>" y="<Y-POSITION>"

```

```

        incLanes="<INCOMING_LANES>" intLanes="<INTERNAL_LANES>"
        shape="<SHAPE>">
</junction>

```

Figura 3.5: Exemplo de descrição de semáforo. Inclusão do tipo (*type*) "traffic_light".

```

<tlLogic id="0" programID="my_program" offset="0" type="static">
  <phase duration="31" state="GGgrrrrGGgrrrr"/>
  <phase duration="5" state="yygrrrryygrrrr"/>
  <phase duration="6" state="rrGGrrrrrrGGrrrr"/>
  <phase duration="5" state="rryyrrrrrryyrrrr"/>
  <phase duration="31" state="rrrrGGgrrrrGGgg"/>
  <phase duration="5" state="rrrryygrrrryygg"/>
  <phase duration="6" state="rrrrrrGGrrrrrrGG"/>
  <phase duration="5" state="rrrrrryyrrrrrryy"/>
</tlLogic>

```

Figura 3.6: Exemplo de descrição do plano semafórico e das fases.

3.1.2.7 Modelos de controle de intersecção abordados

Sem controle: neste modelo, cada intersecção da malha viária não possui nenhum controle específico de sinalização, dessa forma os veículos que estiverem tentando cruzar determinada intersecção, acabarão tendo que parar para evitar acidentes. Adicionado a isso, o simulador define certas prioridades de transposição que podem ocasionar a formação de congestionamentos em determinados pontos, ou até mesmo acidentes, já que o modelo de motorista possui um fator de impaciência associado que faz com que ele tente seguir sua viagem.

Plano semafórico fixo: neste modelo, cada intersecção da malha viária possuirá um controle semafórico simples, onde um plano semafórico padrão pré-determinado será instalado. A definição desse plano não prioriza nenhuma via, ou seja, ele divide igualmente o tempo de ciclo entre todas as vias que o circundam. Esse tempo de ciclo permite a cada fase verde um tempo de 45 segundos.

Esse mesmo plano também é escolhido como plano inicial por todos os semáforos nas abordagens subsequentes deste trabalho que aplicam algoritmos de mudança no comportamento dos semáforos.

Plano semafórico dinâmico: neste plano, é adotado o uso do algoritmo bio-inspirado, o qual será descrito a seguir na Seção 3.3. A principal característica deste tipo de abordagem se deve a adicionar ao agente a capacidade de mudar de tarefas (mudar seu plano atual por outro pré-determinado), e dessa forma permitir que emergja uma coordenação entre os semáforos de forma a aumentar a vazão em determinadas vias. Essa coordenação acontece devido ao efeito de uma comunicação implícita que ocorre entre os agentes (estigmergia).

Plano semafórico aprendido: neste plano, será adotado o uso da Aprendizagem Q, a qual será descrita a seguir na Seção 3.4. A principal característica no uso desta técnica está em permitir que cada semáforo seja induzido a modificar seu comportamento através

de pequenos incentivos (reforços), de acordo com o que ele percebe do tráfego que controla. Dessa maneira, ao perceber que determinadas vias costumam ter um maior fluxo de veículos, ele irá corresponder aumentando a permanência da fase de verde do plano semafórico para aquela via, possibilitando assim que a fluidez do tráfego seja maior.

Plano semafórico híbrido com memória: neste plano, será adotando o uso do algoritmo híbrido, o qual será descrito a seguir na Seção 3.5. Com as informações dos dois tipos de plano semafórico anteriores, podemos notar que eles possuem certas capacidades de adaptação ao seu meio, entretanto ela ocorre de forma pontual no tempo, necessitando assim de re-adaptações contínuas, e podendo nunca mais alcançar um plano ótimo que possa ter sido aprendido e depois esquecido.

Por esse motivo, a proposta de criação de um algoritmo híbrido com capacidade de memorização de planos permite, além de uma independência maior de especialistas de trânsito que manipulem esses planos, tornar os semáforos capazes de guardar os planos mais adaptados às demandas de sua região e reduzir a curva de aprendizagem para a troca entre os planos já aprendidos.

3.2 Modelagem das redes de tráfego urbano e suas demandas

Como foi apresentado anteriormente (Seção 2.2.2.2), existem muitas formas para a geração de malhas viárias ao utilizar o SUMO. Abaixo são adicionados mais alguns termos necessários ao entendimento desta seção.

Distrito

É uma região pré-determinada de uma malha viária, onde viagens serão originadas ou destinadas. Ele pode ser representado por uma área contendo diversas vias e intersecções, apenas vias ou apenas intersecções.

Variantes da rede

Para todas as redes estudadas foram geradas três variantes a partir do seu modelo de estrutura básico, e que servem para comparação com os algoritmos propostos. Assim, foram desenvolvidas as seguintes variantes de rede:

- a primeira variante possui as junções sem controle (ver Seção 3.1.2.7);
- na segunda foram incluídas condições de controle de tráfego (controle semafórico fixo, ver Seção 3.1.2.7).

Geração das demandas

Para evitar duplicações no texto que segue, note que:

1. para todas as viagens dos veículos definidas por um arquivo de fluxo ou matriz OD (Seção 2.2.2.4) é realizado um roteamento com o algoritmo de caminho mínimo Dijkstra, que é padrão da ferramenta DUAROUTER;
2. se a demanda for feita por arquivo de fluxo, cada veículo é enviado para a simulação em um período de tempo distribuído uniformemente (por exemplo a cada 5 segundos)

3. se a demanda for feita por uma matriz OD, cada veículo é enviado para a simulação em um período de tempo sorteado aleatoriamente dentro do período máximo definido no arquivo da matriz OD;
4. a duração das demandas é de 3 horas, permitindo assim a observação da convergência dos algoritmos experimentados.

A seguir estão descritas cada uma das malhas viárias criadas para este trabalho, em cada uma é apresentada a motivação por trás de sua criação e a demanda utilizada, e ao final desta seção uma tabela apresenta uma comparação das características de todas estas redes.

3.2.1 Mapa em formato de cruzamento simples

Neste mapa simples (ver Figura 3.7a) são realizados os testes iniciais utilizando o algoritmo de Aprendizagem Q (Seção 3.4). O mesmo foi escolhido devido a simplicidade na observação da atuação do algoritmo no agente. O mapa consiste de um conjunto de 5 nós (mas apenas 1 cruzamento central) e 4 arestas, cada aresta comporta um fluxo unidirecional de veículos e possui comprimento de aproximadamente 200 metros. O comprimento total da rede, dado pela soma do comprimento das arestas multiplicado pelo número de pistas, é de aproximadamente 800 metros. Seriam necessários cerca de 160 veículos para ocupar a rede completamente. A velocidade máxima permitida para cada aresta em condições de fluxo livre é de 13.89 m/s (50 km/h) para todas as arestas. O tempo de viagem em fluxo livre em cada aresta é de 14.4 segundos.

3.2.1.1 Demanda produzida

A demanda é formada a partir de um arquivo que descreve o fluxo de viagens dos veículos e está descrito na Tabela 3.1. Nela a primeira coluna indica nodos de origem da demanda e a primeira linha indica nodos de destino. Volume na origem e no destino indicam respectivamente o número de viagens que partem e que chegam em determinado nodo.

Tabela 3.1: Fluxo de veículos da malha viária cruzamento simples. Volume total de 3600 veículos para 3 horas.

	S	E	Volume na origem
W	1200	1200	2400
N	600	600	1200
Volume no destino	1800	1800	3600

3.2.2 Mapa em formato de cruzamento duplo

É semelhante ao mapa em formato de cruzamento simples, possuindo agora dois semáforos (ver Figura 3.7b). O interesse nessa rede está em verificar mudanças de comportamento de um dos agentes ocasionada pela cadeia de ações tomadas pelo outro. O mapa consiste de um conjunto de 8 nós (mas com apenas 2 cruzamentos centrais) e 7 arestas, cada aresta comporta um fluxo unidirecional de veículos e possui comprimento de aproximadamente 200 metros. O comprimento total da rede, dado pela soma do comprimento

das arestas multiplicado pelo número de pistas, é de aproximadamente 1.4 quilômetro. Seriam necessários cerca de 280 veículos para ocupar a rede completamente. A velocidade máxima permitida para cada aresta em condições de fluxo livre é de 13.89 m/s (50 km/h) para todas as arestas. O tempo de viagem em fluxo livre em cada aresta é de 14.4 segundos.

3.2.2.1 Demanda produzida

A demanda é formada a partir de um arquivo que descreve o fluxo de viagens dos veículos e está descrito na Tabela 3.2. Nela a primeira coluna indica nodos de origem da demanda e a primeira linha indica nodos de destino. Volume na origem e no destino indicam respectivamente o número de viagens que partem e que chegam em determinado nodo.

Tabela 3.2: Fluxo de veículos da malha viária cruzamento duplo. Volume total de 3600 veículos para 3 horas.

	S1	S2	E	Volume na origem
W	800	800	800	2400
N1	200	200	200	600
N2	0	300	300	600
Volume no destino	1000	1300	1300	3600

3.2.3 Mapa em formato de bloco

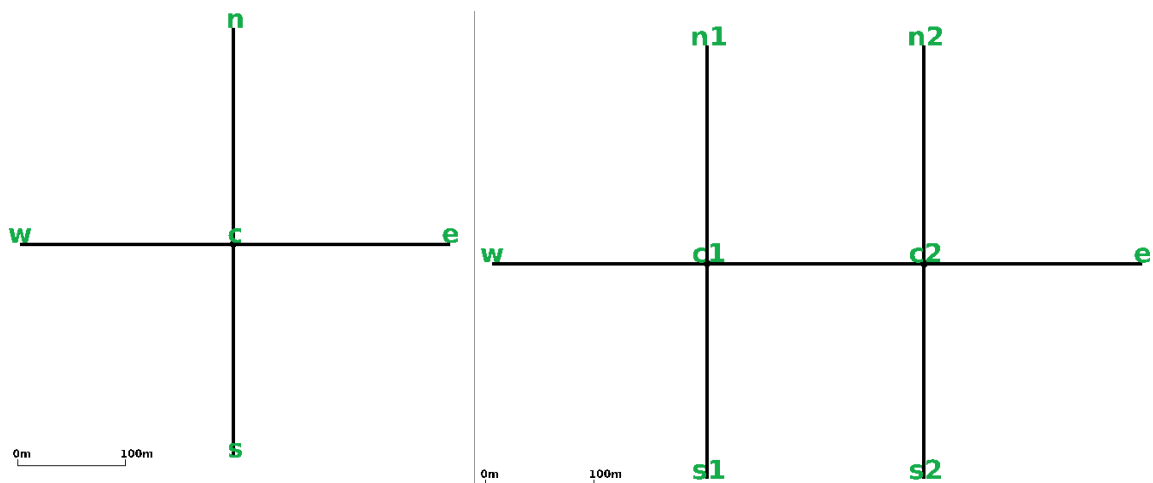
Este mapa possui um formato de grade simples com algumas vias de entrada (ver Figura 3.7c), o objetivo aqui está em experimentar o comportamento do algoritmo bio-inspirado (Seção 3.3), percebendo a ocorrência da coordenação implícita entre alguns dos semáforos pela troca de tarefas. O mapa consiste de um conjunto de 12 nós (com 4 cruzamentos centrais) e 12 arestas, cada aresta comporta um fluxo unidirecional de veículos e possui comprimento de aproximadamente 200 metros. O comprimento total da rede, dado pela soma do comprimento das arestas multiplicado pelo número de pistas, é de aproximadamente 2.4 quilômetros. Seriam necessários cerca de 480 veículos para ocupar a rede completamente. A velocidade máxima permitida para cada aresta em condições de fluxo livre é de 13.89 m/s (50 km/h) para todas as arestas. O tempo de viagem em fluxo livre em cada aresta é de 14.4 segundos.

3.2.3.1 Demanda produzida

A demanda é formada a partir de um arquivo que descreve o fluxo de viagens dos veículos e está descrito na Tabela 3.3. Nela a primeira coluna indica nodos de origem da demanda e a primeira linha indica nodos de destino. Volume na origem e no destino indicam respectivamente o número de viagens que partem e que chegam em determinado nodo.

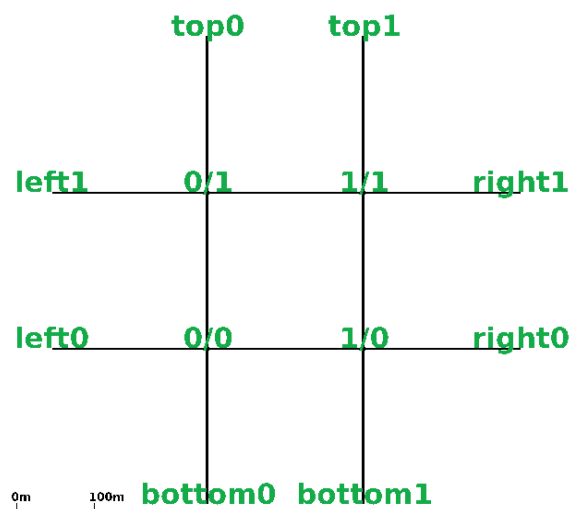
Tabela 3.3: Fluxo de veículos da malha viária em formato de bloco. Volume total de 3600 veículos para 3 horas.

	right0	right1	bottom0	bottom1	Volume na origem
left0	100	0	150	150	400
left1	600	600	600	600	2400
top0	100	100	100	100	400
top1	150	100	0	150	400
Volume no destino	950	800	850	1000	3600



(a) Mapa em formato de cruzamento simples.

(b) Mapa em formato de cruzamento duplo.



(c) Mapa em formato de bloco.

Figura 3.7: Modelos de malhas simples criadas no SUMO para testes iniciais.

3.2.4 Mapa das vias de Sioux Falls

A rede de tráfego chamada Sioux Falls³(ver Figura 3.8b) utilizada nesse trabalho consiste de um conjunto de 24 nós e 76 arestas, cada nó representa uma intersecção de vias e ao mesmo tempo um distrito para origem e destino de veículos, e cada aresta comporta um fluxo bidirecional de veículos e possui comprimento variado. O comprimento total da rede, dado pela soma do comprimento das arestas multiplicado pelo número de pistas, é de aproximadamente 196.8 km. Seriam necessários aproximadamente 39.4 mil veículos para ocupar a rede completamente. A velocidade máxima permitida para cada aresta em condições de fluxo livre é de 13.89 m/s (50 km/h) para todas as arestas. O tempo de viagem em fluxo livre depende do comprimento de cada aresta. Essa rede é uma representação⁴ de algumas das vias da cidade de Sioux Falls, Dakota do Sul, EUA⁵.

A origem de seu uso em experimentos ocorreu através do trabalho de LeBlanc et al. [1975], do mesmo trabalho foram extraídas as informações relativas ao fluxo de veículos demandado para nossos experimentos.

A criação da rede foi feita através dos passos descritos abaixo.

- Captura de informações: das coordenadas cartesianas dos nodos, arestas que os conectam, e viagens dos veículos entre os nodos (demanda).
- Conversão da escala de coordenadas: devido as originais possuírem um formato em polegadas foi necessária uma manipulação de seus valores (redução de 40 vezes), e que gerou uma escala praticamente idêntica a do mapa real (menos de 4% de diferença).
- Conversão da escala de coordenadas: devido as originais possuírem um formato desconhecido foi feita uma manipulação de seus valores (redução de 40 vezes) para a geração de uma escala praticamente idêntica a do mapa real (menos de 4% de diferença).
- Utilizando-se das partes básicas da rede citadas anteriormente, foram geradas três variantes do modelo de rede para o simulador microscópico e contínuo SUMO⁶ através da ferramenta NETCONVERT.

3.2.4.1 Demanda produzida

A demanda é formada a partir de uma matriz OD das viagens dos veículos (ver Apêndice A.1), ela possui um total de 3606 viagens. No trabalho original ela é sofre uma escala em um fator de 1000 vezes (3.606.000 veículos) para uma distribuição com duração de 24 horas. Devido as alterações feitas na estrutura da rede para este trabalho, optou-se por manter o modelo de demanda, mas utilizou-se um fator de escala equivalente a 8.4 vezes para ampliar essa demanda (cerca de 30290 viagens), que foi então distribuída durante 3 horas de simulação. Mesmo que essa demanda pareça pequena ela é suficiente para a geração de filas de veículos em alguns pontos. Como citado no começo desse capítulo, a duração de 3 horas na simulação é necessária para a observação do comportamento dos algoritmos.

³ <http://www.bgu.ac.il/~bargera/tntp/>

⁴ <http://www.bgu.ac.il/~bargera/tntp/SiouxFalls/Sioux-Falls-Network.doc>

⁵ http://www.bgu.ac.il/~bargera/tntp/SiouxFalls/SiouxFallsMap_AAA1998.jpg

⁶ <http://sumo-sim.org/>

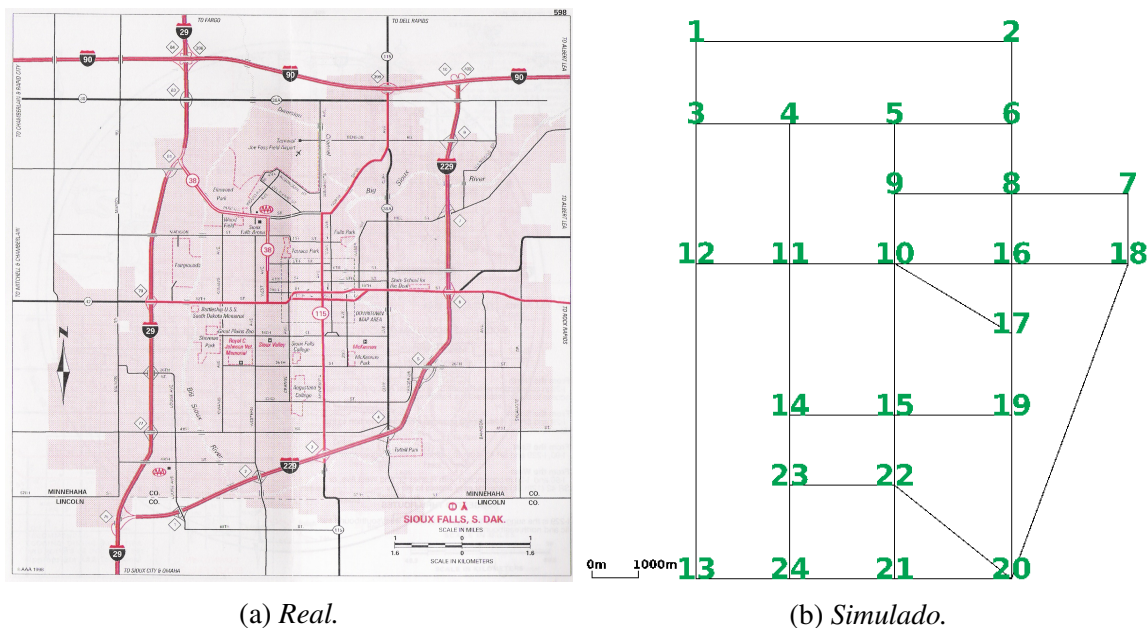


Figura 3.8: Visão sobre o mapa viário real e simulado da cidade de Sioux Falls, Dakota do Sul, EUA.

3.2.5 Mapa das vias arteriais de Porto Alegre

A rede de tráfego das vias arteriais de Porto Alegre (ver Figura 3.9b) utilizada nesse trabalho consiste de um conjunto de 61 nós e 152 arestas, cada nó representa uma intersecção de vias e cada aresta uma via de comprimento e sentido de fluxo variados. O comprimento total da rede, dado pela soma do comprimento das arestas vezes o número de pistas, é de aproximadamente 636.6 km. Seriam necessários aproximadamente 125 mil veículos para ocupar a rede completamente. A velocidade máxima permitida para cada aresta em condições de fluxo livre é de 13.89 m/s (50 km/h) para todas as arestas. O tempo de viagem em fluxo livre depende do comprimento de cada aresta.

Essa rede já vem sendo utilizada dentro do projeto MASLAB⁷. Foi necessário então apenas a geração das três variantes do modelo de rede para o simulador através da ferramenta NETCONVERT.

3.2.5.1 Demanda produzida

A demanda foi produzida a partir de uma matriz OD das viagens dos veículos (ver Apêndice A.2), ela possui um total de 210 viagens. Utilizou-se um fator de escala equivalente a 95 vezes para ampliar essa demanda (cerca de 20000 viagens), que foi então distribuída durante 3 horas de simulação. Mesmo que essa demanda pareça pequena ela é suficiente para a geração de filas de veículos em alguns pontos. Como citado no começo desse capítulo, a duração de 3 horas na simulação é necessária para a observação do comportamento dos algoritmos.

⁷ <http://www.inf.ufrgs.br/maslab/maslab/>

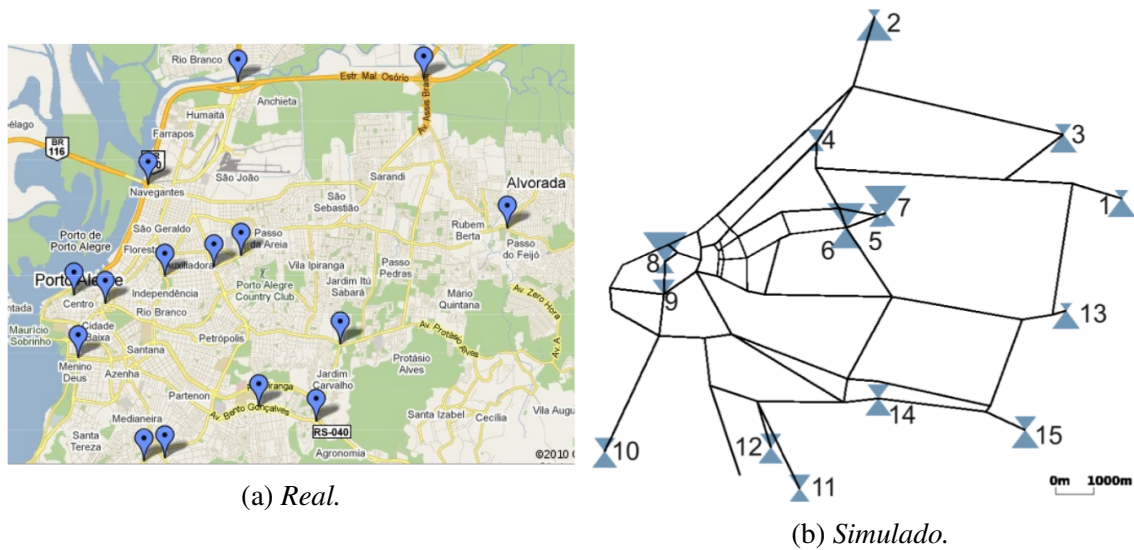


Figura 3.9: Mapa real (a) retirado de [Amarante, 2012] e simulado (b) retirado de [Tavares, 2013] das principais vias arteriais de Porto Alegre. O tamanho de um triângulo mostra qualitativamente a proporção de viagens que se iniciam (ponta para cima) ou terminam (ponta para baixo) em um distrito. Os números identificam os distritos.

3.2.6 Comparação das características das redes propostas

Na Tabela 3.4 são apresentadas as informações gerais sobre as redes mais simples. Pode-se notar que ocorre uma duplicação do número de agentes (semáforos) da esquerda para a direita. Mesmo sendo simples, essas redes serviram como forma de validar os resultados empíricos através de observação visual do comportamento dos agentes. Quando as redes se tornam mais complexas, torna-se difícil esse tipo de observação, sendo necessária uma análise mais profunda sobre os dados de saída da simulação.

Na Tabela 3.5 são apresentadas as informações gerais sobre as redes mais complexas. Como nas redes anteriores, também existe uma duplicação no número de semáforos da esquerda para a direita.

Os resultados obtidos através da avaliação das redes propostas nos experimentos podem ser visto no Capítulo 4.

Tabela 3.4: Comparação das redes para execução dos experimentos iniciais.

	Cruzamento simples	Cruzamento duplo	Formato Bloco
Numero de semáforos	1	2	4
Numero de nós	5	8	12
Número de vias	4	7	12
Capacidade máxima* (veículos)	80	280	480
Comprimento total (m)	400	1400	2400
Velocidade máxima (m/s)	13.89	13.89	13.89
Tempo em fluxo livre (s)	14.4	14.4	14.4
Duração da demanda (h)	3	3	3
Volume total (viagens)	3.600	3.600	3.600

*Comprimento / tamanho do veículo. Por padrão um veículo possui 5 metros de comprimento.

Tabela 3.5: Comparação das redes para execução dos experimentos principais.

	Sioux Falls	Poa Arterials
Numero de semáforos	20	43
Numero de nós	24	61
Número de vias	76	152
Capacidade máxima* (veículos)	39.400	125.000
Comprimento total (km)	196.8	636.6
Velocidade máxima (m/s)	13.89	13.89
Tempo em fluxo livre (s)	variado	variado
Duração da demanda (h)	3	3
Volume total (viagens)	30.300	20.000

*Comprimento / tamanho do veículo. Por padrão um veículo possui 5 metros de comprimento.

3.3 Abordagem via inteligência de enxame

3.3.1 Fluxo de atuação da abordagem

Como apresentado na Seção 2.4, colônias de insetos possuem uma capacidade interessante relacionada a sua adaptação ao meio ambiente e na resolução de tarefas através de uma comunicação implícita com o uso de feromônios.

Da observação de uma malha viária através da metáfora de insetos sociais, os semáforos são representados pelos insetos que necessitam alocar tarefas. Nesse contexto as tarefas serão representadas pela priorização do fluxo de veículos em determinadas vias (mudanças no plano semafórico) e por conseguinte, os veículos serão assimilados como o rastro hormonal usado na avaliação da tarefa. Quanto maior esse rastro, maior a tendência do semáforo priorizar aquela via. Essa formulação gera uma espécie de comunicação indireta entre os semáforos adjacentes (no contexto biológico chamada de stigmergia).

Inicialmente a cada semáforo é atribuído um conjunto de possíveis planos semafóricos, ele decide então qual deles deve executar de acordo com a potência percebida desse hormônio contra seu limiar de troca de tarefas.

Os planos semafóricos iniciais são pré-calculados através de uma lista de intervalos de ocupação de vias descrito junto aos parâmetros na Seção 3.3.2.

Após a fase de troca de plano, existe um período de espera (um tempo de 3 ciclos semafóricos, que é algo próximo de 5 minutos), onde o semáforo fará novamente uma avaliação sobre as mudanças no seu meio para então passar por uma nova fase de trocas, esse comportamento segue então indefinidamente.

Casos de falha semafórica não são abordados, mas supõe-se um sistema onde os planos sejam salvos e recarregados ou em caso contrário o comportamento padrão de equilíbrio é novamente iniciado e o algoritmo passa a agir desde o início.

3.3.2 Descrição do algoritmo

Seja s_j a intensidade de um estímulo associado a uma tarefa j . O limiar de resposta θ_{ij} , expresso em unidades de intensidade de estímulo, é uma variável interna que determina a tendência de um indivíduo i , respondendo ao estímulo s_j , realizar a tarefa j .

O estímulo do plano semafórico é calculado a partir dos rastros de feromônio acumulados em cada faixa controlada por ele. Enquanto os veículos estão parados eles permanecem liberando feromônio. Assim, uma via com maior quantidade de veículos parados possuirá maior densidade de feromônio. A variação da densidade de feromônio em uma faixa l no tempo t é dada por:

$$d_{l,t} = \frac{\sum_{t=0}^w \beta^{-t} (\delta_{l,t})}{\sum_{t=0}^w \beta^{-t}}, \quad (3.1)$$

onde w é o tamanho do intervalo de tempo, β é a taxa de dissipação do feromônio e $\delta_{l,t}$ é a densidade de veículos em uma faixa l no tempo t . Esta é dada por:

$$\delta_{l,t} = \frac{\sum_{l=1}^L N_{l,t}}{L}, \quad (3.2)$$

onde L é o tamanho da faixa em número de células (cada célula tem tamanho equivalente a 1 veículo) e $N_{l,t}$ é o número de veículos em uma faixa l no tempo t .

O estímulo s de um plano j é calculado utilizando-se uma soma ponderada do feromônio acumulado em cada uma de suas fases:

$$s_j = \sum_{k=0}^n (d_{in_k,t}) \Delta_k, \quad (3.3)$$

onde n é o número de fases do plano j , $d_{in_k,t}$ é a densidade de feromônio acumulado nas faixas de entrada da fase k no tempo t e Δ_k a porção de tempo utilizada pela fase k . Cada fase tem uma porção de tempo

$$\Delta_k = (\text{tempo}_{fim} - \text{tempo}_{inicio}) / \text{tempo}_{ciclo}, \quad (3.4)$$

que indica o tempo de sinal verde que o plano dedica para aquela fase. Fases com maiores porções de tempo indicam que o plano prioriza determinados movimentos.

A flexibilidade na troca de planos pelo agente é consequência das mudanças induzidas pelo ambiente no estímulo e nos limiares do agente. Cada plano semafórico possui um estímulo associado de acordo com a direção que ele prioriza. Os agentes tem uma tendência maior de trocar de tarefa se essa direção tem um estímulo que ultrapassa seu limiar de resposta. A tendência de um indivíduo executar uma tarefa como uma relação entre o estímulo e o limiar de resposta é dado por:

$$T_{\theta_{ij}}(s_j) = \frac{s_j^2}{s_j^2 + \theta_{ij}^2}, \quad (3.5)$$

onde θ_{ij} é o limiar do indivíduo i em relação a tarefa j e s_j é o estímulo associado à tarefa j .

Utiliza-se o modelo de especialização no qual o limiar de resposta do indivíduo para a execução de uma tarefa é atualizado utilizando um mecanismo de aprendizado por reforço. Insetos reais são diretamente influenciados pelo seu grau de sucesso na execução de uma determinada tarefa [Gordon, 2002]. Formigas bem sucedidas são motivadas a

continuar na execução da mesma tarefa, enquanto as mal sucedidas tendem a parar de executar a tarefa e buscar outra. Para adicionar essa propriedade que faz o inseto buscar ser bem sucedido, o modelo de aprendizado de [Bonabeau et al., 1999] foi estendido para incluir na parte da mudança da tendência um coeficiente que representa tanto o aprendizado quanto o esquecimento (quando o coeficiente é negativo o agente está esquecendo, quando é positivo ele está aprendendo). Isto se dá através da fórmula:

$$\theta_{ij} = \theta_{ij} - \lambda \Delta t, \quad (3.6)$$

onde λ é o coeficiente de aprendizado/esquecimento e Δt é um intervalo de tempo discreto, por exemplo a cada novo ciclo do semáforo. O grau de sucesso de um indivíduo pode ser calculado de forma linear

$$\lambda = 1 - 2\sigma \quad (3.7)$$

ou exponencial

$$\lambda = 2e^{(-5\sigma)} - 1. \quad (3.8)$$

Em ambos os casos σ é o desvio padrão do rastro de feromônio nas seções de rua controladas pelo semáforo.

Nos experimentos, o cálculo de λ foi utilizado apenas em sua forma linear.

Na Tabela 3.6 são apresentados os parâmetros que podem ser passados ao algoritmo. O parâmetro *fatores_de_tempo* é uma lista de fatores de mudança de tempo utilizada na criação dos planos iniciais do algoritmo, nela cada valor representa uma porcentagem que afeta o tempo de cada fase em verde do plano semafórico para mais ou para menos.

Tabela 3.6: Colônia de formigas: visão geral sobre os parâmetros de entrada.

	Valor inicial	Intervalo	Descrição
μ	100	$[0, \infty)$	Intervalo de tempo para cálculos de ocupação.
<i>fatores_de_tempo</i>	[0.10, 0.25, 0.5]	[[0,1]. . .]	Lista de fatores de mudança de tempo.
β	0.5	[0,1]	Taxa de dissipação de feromônio.
s_j	0.5	[0,1]	É o estímulo associado à tarefa j .
θ_{ij}	0.5	[0,1]	É o limiar do indivíduo i em relação a tarefa j .

Algoritmo 1 Colônia de formigas

```

// Inicialização:
i ← 1;
para cada semáforo  $S_i$  na rede faça
   $S_i(\beta) \leftarrow \beta_{entrada}$ ;
   $P_{inicial} \leftarrow cria\_plano\_padrao(s_{entrada})$ ; // Cria plano inicial com tempo das fases em
  verde de 45 segundos e estímulo inicial  $s_{entrada}$ .
   $S_i \leftarrow P_{inicial}$ ; // Atribui o plano inicial ao semáforo atual.
  j ← 1;
  para cada  $fv$  em  $fases\_em\_verde$  no  $P_{inicial}$  faça
    para cada fator em  $fatores\_de\_tempo$  faça
       $P_j \leftarrow copia(P_{inicial})$ ; // Cria um novo plano com base na cópia do plano inicial.
      // Ver observação ao fim do algoritmo.
       $atualiza\_fases(P_j, fv, fator)$ ; // Altera o tempo da fase do plano atual.
      // Atribui ao agente  $S_i$  o estímulo em relação ao plano  $j$ .
       $S_i(\theta_{ij}) \leftarrow \theta_{entrada}$ ;
      j ← j + 1;
    fim
  fim
  i ← i + 1;
fim
// Atualização:
para cada novo passo de tempo  $t$  faça
  i ← 1;
  para cada semáforo  $S_i$  na rede faça
     $observa\_vias(S_i)$ ; // Captura dados de ocupação nas vias que controla
     $atualiza \delta_{i,t}$  para cada via  $l \in S_i$  de acordo com 3.2;
     $atualiza d_{i,t}$  para cada via  $l \in S_i$  de acordo com 3.1;
    para cada fase  $k$  no plano atual ( $P_{atual}$ ) faça
       $atualiza \Delta_k$  de acordo com 3.4;
    fim
     $atualiza s_{atual}$  de acordo com 3.3;
    se momento de atuar então
       $atualiza \lambda$  de acordo com 3.7;
       $atualiza \theta_{i,atual}$  de acordo com 3.6;
       $atualiza T_{\theta_{i,atual}}(s_{atual})$  de acordo com 3.5;
      para todo plano  $P_j$  de  $S_i$  faça
        se  $limiar(P_j) > T_{\theta_{i,atual}}$  então
           $planos\_escolhidos \leftarrow P_j$ ;
        fim
      fim
      se  $planos\_escolhidos \neq \emptyset$  então
        // Escolhe aleatoriamente um plano que teve bom desempenho.
         $S_i(P_{atual}) \leftarrow escolha(planos\_escolhidos)$ ;
      fim
    fim
    i ← i + 1;
  fim
fim

```

Observação: O método $atualiza_fases(P_j)$ consiste na atualização de uma das fases em verde do plano P_j , isso é feito através do incremento do tempo da fase de acordo com a redução das outras fases em verde. Por exemplo, supondo que o *fator* atual (de *fatores_de_tempo*) seja 0.1 (o mesmo que 10%) isso equivale a dizer que cada outra fase em verde do plano sofrerá uma redução de 10% no seu tempo de verde, esse valor é então

acumulado e após somado ao tempo de verde da fase atual. Alguns ajustes são feitos para que o tempo de ciclo após as alterações permaneça o mesmo. Para um caso simples onde haja um cruzamento de duas vias com duas fases em verde (uma para cada via), o efeito causado pela atualização acima será a criação de dois planos semafóricos, onde cada um deles irá priorizar uma das vias em 10% do tempo de verde.

3.4 Abordagem via aprendizagem por reforço

3.4.1 Fluxo de atuação da abordagem

Como visto na Seção 2.5, a aprendizagem em um sistema dinâmico possui certas dificuldades inerentes a cada problema, mas no geral o espaço de estados do sistema será muito grande. Aprender dinamicamente como este espaço de estados funciona torna-se uma habilidade muito interessante ao agente, mais ainda, aprender quais são os estados mais valiosos torna-se indispensável para evitar uma exploração desnecessária.

O uso da Aprendizagem Q (Seção 2.5.3.2) permite então que possamos construir essa visão de mundo de forma iterativa, através da escolha de ações que maximizem a utilidade do novo estado para o agente.

Além dos elementos presentes nas equações da Seção 2.5, como α, γ , os seguintes parâmetros $\aleph, \chi, \epsilon, \bar{\epsilon}, \eta, \tau$ são incluídos e estão descritos brevemente na Tabela 3.7.

Em especial, \aleph representa uma lista de fatores de discretização de estados, que é utilizada para particionar o espaço de estados de transição possíveis para o agente. Os valores representam a ocupação de determinada aresta, estes valores estão entre 0 (aresta livre) e 1 (aresta totalmente ocupada) e foram escolhidos de forma empírica. O processo consiste em gerar um produto cartesiano entre os intervalos de discretização e as fases do plano semafórico. Por exemplo, se \aleph for a lista [0.5] (ele é equivalente a dois intervalos discretos distintos: [0, 0.5] e [0.5, 1], chamados de A e B respectivamente, onde A possui ocupação da aresta inferior a 50% e B ocupação igual ou superior a 50%) e o plano atual tiver duas fases (chamadas de $fase - 1$ e $fase - 2$), teremos uma lista com o seguinte conjunto de estados:

$$[[A, fase - 1], [B, fase - 1], [A, fase - 2], [B, fase - 2]].$$

Com a definição dos estados do parágrafo anterior, é utilizada a lista de fatores de ação χ para a geração da tabela-Q (que contém as informações de transição de estados para o agente). Essa lista corresponde a um conjunto de valores que afetam o tempo da fase do plano semafórico atual para mais ou para menos. Por exemplo, seja χ a lista [-10%, +10%] (que alteraria o tempo da fase em -10% ou +10%) e utilizando os estados descritos anteriormente, teríamos o seguinte conjunto de estados-ação na tabela-Q de exemplo:

$$[[A, fase - 1, -10\%], [A, fase - 1, +10\%], \dots, [B, fase - 2, -10\%], [B, fase - 2, +10\%]].$$

Ou seja, 8 estados-ação. O número de elementos na tabela-Q cresce então como o produto dos conjuntos de estados, fases e ações, respectivamente $SxFxA$. Note como a quantidade de elementos contidos nesta tabela cresce rapidamente.

Passada essa inicialização da tabela-Q, para cada novo passo de tempo, o agente irá capturar informações do seu ambiente (ocupação das vias) até que chegue o momento de atuar, a fase de ação.

A **fase de ação** para o agente consiste na seleção de uma ação baseada nas informações adquiridas do meio, essa seleção consiste na avaliação do estado de ocupação atual da via controlada e uma discretização do estado, semelhante ao que foi exemplificado anteriormente. É então escolhida através da tabela-Q o conjunto de possíveis ações a tomar (é válido também não tomar nenhuma ação). As ações podem ser escolhidas através de uma busca gulosa (ação mais valiosa dentre os valores-Q) ou seleção aleatória, ambas dependem do fator de seleção de ações do agente (ϵ). Caso o agente já tenha agido anteriormente, nesse momento ele irá atualizar a tabela-Q com a recompensa obtida pela ação da rodada anterior através da função de utilidade (Equação 2.6). A partir disso, o agente fará a atualização do tempo semafórico.

Após a fase em que ocorre uma ação, existe um período de espera (um tempo de 3 ciclos semafóricos, que é algo próximo de 5 minutos), onde o semáforo fará novamente uma avaliação sobre as mudanças no seu meio para então passar por uma nova fase de ação, esse comportamento segue então indefinidamente.

3.4.2 Descrição do algoritmo

Dentre os elementos constantes na função de atualização de utilidade da tabela-Q (Equação 2.6), a recompensa $R(s)$ é calculado com base na seguinte equação:

$$R(s) \leftarrow (2 * (1 - ocupação)^\tau) - 1 \quad (3.9)$$

onde *ocupação* é calculada com base na média das ocupações médias de cada via que chega ao semáforo. A ocupação média é baseada em um período de tempo discreto de 100 passos. τ é um fator que afeta o valor da recompensa, quanto menor seu valor, maior será acumulado para a recompensa, quanto maior, menor a recompensa. A recompensa em si será devolvida como um valor no intervalo $[-1, 1]$ quando $\tau = 2$.

Tabela 3.7: Aprendizagem Q: visão geral sobre os parâmetros de entrada.

	Valor padrão	Intervalo	Descrição
μ	100	$[0, \infty)$	Intervalo de tempo para cálculos de ocupação
\aleph	[0.1, 0.21, 0.33, 0.5, 0.6, 0.65, 0.7]	$[[0,1]. \dots]$	Lista de fatores de discretização de estados
χ	[-12.5%, -8.5%, -5%, +10%, +17%, +25%]	$[[0,1]. \dots]$	Lista de incrementos e decrementos do tempo de fase em verde
α	0.5	$[0,1]$	Taxa de aprendizagem
γ	0.5	$[0,1]$	Fator de desconto na aprendizagem
ϵ	0.5	$[0,1]$	Fator de seleção de ação do agente
$\bar{\epsilon}$	0.5	$[0,1]$	Decaimento do fator de seleção de ação do agente
<i>valor - Q</i>	0	$[0,1]$	Valor inicial da recompensa para tupla estado-ação)
τ	2	$[0, \infty)$	Expoente de recompensa

Algoritmo 2 Aprendizagem Q

```

// Inicialização:
i ← 1;
para cada semáforo  $S_i$  na rede faça
   $S_i \leftarrow \text{atribui\_parametros}(\alpha, \gamma, \epsilon, \kappa, \eta, \text{valor} - Q, \tau)$ ;
  // Cria tabela-Q com base nos discretizadores
  // de estado e fatores de ação.
   $S_i \leftarrow \text{tabelaQ}(\mathfrak{X}, \chi)$ 
  // Cria plano inicial com tempo das fases em verde de
  // 45 segundos e estímulo inicial  $s_{\text{entrada}}$ .
   $P_{\text{inicial}} \leftarrow \text{cria\_plano\_padrao}(s_{\text{entrada}})$ ;
   $S_i \leftarrow P_{\text{inicial}}$ ; // Atribui o plano inicial ao semáforo atual.
fim
// Atualização:
para cada novo passo de tempo t faça
  i ← 1;
  para cada semáforo  $S_i$  na rede faça
     $\text{observa\_vias}(S_i)$ ; // Captura dados de ocupação nas vias que controla
    se momento de atuar então
      Seja  $a$  a ação escolhida por  $\pi_{S_i}(s)$ 
      Observe o próximo estado  $s'$  e a recompensa  $r$  de acordo com 3.9
      se já atuou antes então
        atualiza tabela-Q de acordo com 2.6;
      fim
      // Atuação:
      atualiza tempos do plano  $P_{\text{inicial}}$  de acordo com ação  $a$  escolhida
      // Atualiza o estado atual:
       $s \leftarrow s'$ ;
    fim
  i ← i + 1;
fim
fim

```

3.5 Abordagem híbrida: *Smart Ant with Context-Detection (SA-CD)*

3.5.1 Fluxo de atuação da abordagem

Utilizando a abordagem da colônia de insetos, associada a técnica da Aprendizagem Q e a técnica da Detecção de Contexto (Seção 2.5.3.3) formalizada a seguir, foi desenvolvida a aplicação chamada *Smart Ant with Context-Detection (SA-CD)*, que une a capacidade de escolher planos dinamicamente (insetos), alterar os tempos de verde de suas fases (aprendizado) e criar novos planos semafóricos de acordo com a percepção do ambiente (detecção de contexto).

A atuação consiste na inicialização e criação dos planos iniciais, como no Algoritmo 1, seguida da inicialização e criação da tabela-Q como no Algoritmo 2. Após, ocorre a captura dos parâmetros da técnica proposta, atualizações e avaliações dos modelos.

As atualizações seguem também da mesma forma das ocorridas nos algoritmos anteriores, o diferencial é que aqui ocorre a atualização da qualidade associada ao modelo pelo algoritmo de detecção de contexto. Escolhe-se então o melhor modelo, baseado na sua qualidade ($E_{m_{\text{cur}}}$) e na tendência do agente em relação ao plano (θ_{ij}). Esse plano é ativado como o plano corrente do agente.

Depois das atualizações, ocorre a avaliação do plano semafórico corrente (modelo corrente). Em caso de uma má avaliação a criação de um novo modelo será necessária, esse modelo segue os valores padrão de semáforo de plano fixo, já comentadas na Seção 3.1.2.7. Após essa fase ocorre a atualização do modelo interno de transição ($T_{m_{cur}}$), do modelo interno de recompensa ($R_{m_{cur}}$) e do contador de ações tomadas ($N_m(s, a)$).

Ocorre então a tomada de ação pelo agente, que consiste na atualização dos tempos de verde do semáforo com base no que é feito na tomada de ação do Algoritmo 2.

Após a fase em que ocorre uma ação, existe um período de espera (um tempo de 3 ciclos semafóricos, que é algo próximo de 5 minutos), onde o semáforo fará novamente uma avaliação sobre as mudanças no seu meio para então passar por uma nova fase de ação, esse comportamento segue então indefinidamente.

3.5.2 Descrição do algoritmo

As Equações 3.10 a 3.22 foram incluídas aqui apenas para auxiliar no entendimento do algoritmo proposto (Algoritmo 3), para maiores detalhes sugere-se verificar no trabalho original [Silva et al., 2006a].

Para todas as equações em que se façam presentes: s e k representam um dos possíveis estados da tabela-Q, s' representa um possível estado futuro, r representa um valor de recompensa obtido, e a representa uma ação válida a ser tomada pelo agente.

Atualização do modelo interno de transição:

$$\Delta T_m(k) = \begin{cases} \frac{1-T_m(s,a,k)}{N_m(s,a)+1} & k = s' \\ \frac{0-T_m(s,a,k)}{N_m(s,a)+1} & k \neq s' \end{cases} \quad \forall k \in S \quad (3.10)$$

$$T_m(s, a, k) = T_m(s, a, k) + \Delta T_m(k), \forall k \in S \quad (3.11)$$

Atualização do modelo interno de recompensa:

$$\Delta R_m = \frac{r - R_m(s, a)}{N_m(s, a) + 1} \quad (3.12)$$

$$R_m(s, a) = R_m(s, a) + \Delta R_m \quad (3.13)$$

Atualização do contador de ações tomadas:

$$N_m(s, a) = \min(N_m(s, a) + 1, MAX_{N_m}) \quad (3.14)$$

Atualização da confiança baseada na escolha de uma ação a num estado s :

$$c_m(s, a) = \frac{N_m(s, a)}{MAX_{N_m}} \quad (3.15)$$

Atualização da qualidade instantânea baseada na combinação linear da qualidade da predição da recompensa (e_m^R) e da qualidade da predição da transição (e_m^T):

$$e_m = c_m(s, a) (\Omega e_m^R + (1 - \Omega) e_m^T) \quad (3.16)$$

Atualização da qualidade da predição da recompensa e de seu fator de normalização (Z_R):

$$e_m^R = 1 - 2(Z_R(\Delta R_m)^2) \quad (3.17)$$

$$Z_R = (R_{max} - R_{min})^{-1} \quad (3.18)$$

Atualização da qualidade da predição da transição e de seu fator de normalização (Z_T):

$$e_m^T = 1 - 2(Z_T \sum_{k \in S} \Delta T_m(k)^2) \quad (3.19)$$

$$Z_T = \frac{1}{2}(N_m(s, a) + 1)^2 \quad (3.20)$$

Atualização da qualidade do modelo:

$$E_m = E_m + \rho(e_m - E_m) \quad (3.21)$$

Inicialização do modelo interno de transição baseado no conjunto de estados:

$$T_m(s, a, k) \leftarrow \frac{1}{|S|} \quad \forall s \in S, \forall a \in A, \forall k \in S \quad (3.22)$$

Tabela 3.8: SA-CD: visão geral sobre os parâmetros de entrada.

	Valor padrão	Intervalo válido	Descrição
μ	100	$[0, \infty)$	Intervalo de tempo para cálculos de ocupação.
ϕ	12	$[1, \infty)$	O número máximo de planos aprendidos.
φ	0.5	$[0, 1]$	Um fator que influencia a perda de planos aprendidos.
MAX_{N_m}	10	$[0, \infty)$	O número máximo de escolha de ações passadas.
E_{min}	0.5	$[-1, 1]$	Especifica a qualidade mínima esperada do modelo.
Ω	0.5	$[0, 1]$	Altera a importância relativa à recompensa e a transição do modelo.
ρ	0.5	$[0, 1]$	Coefficiente de ajuste da qualidade do modelo.

Algoritmo 3 Smart Ant with Context-Detection (SA-CD)

```

// Inicialização:
Inicializa atributos e planos como no algoritmo 1
Inicializa atributos e cria tabela-Q como no algoritmo 2
Seja novo_modelo() uma função que cria e inicializa um novo modelo;
Seja  $m_{cur}$  o modelo corrente;
Seja  $M$  o conjunto de todos os modelos;
 $m_{cur} \leftarrow novo\_modelo();$ 
 $M \leftarrow m_{cur};$ 
 $s \leftarrow s_{inicial};$ 
// Atualização:
para cada novo passo de tempo  $t$  faça
   $i \leftarrow 1;$ 
  para cada semáforo  $S_i$  na rede faça
    Atualiza atributos como no Algoritmo 1
    se momento de atuar então
      // Escolhe a ação e avalia o modelo:
      Seja  $a$  a ação escolhida por  $\pi_{m_{cur}}(s)$ 
      Observe o próximo estado  $s'$  e a recompensa  $r$ 
      para todo  $m \in M$  faça
        Atualiza  $E_m$  de acordo com 3.21
      fim
      // Gera a lista de planos escolhidos de acordo com o Algoritmo 1
       $planos\_escolhidos \leftarrow planos\_escolhidos \cup arg \max_m(E_m);$ 
      // Escolhe um dentre os possíveis planos com bom desempenho:
       $m_{cur} \leftarrow escolha(planos\_escolhidos);$ 
      se  $E_{m_{cur}} < E_{min}$  então
         $m_{cur} \leftarrow novo\_modelo();$ 
         $M \leftarrow M \cup m_{cur};$ 
      fim
      Atualiza  $T_{m_{cur}}$  de acordo com 3.11
      Atualiza  $R_{m_{cur}}$  de acordo com 3.13
      Atualiza  $N_m(s, a)$  de acordo com 3.14
      Atua como no Algoritmo 2
    fim
     $i \leftarrow i + 1;$ 
  fim
fim

```

3.6 Considerações finais

Foram apresentados neste capítulo os diversos elementos que foram necessários para a criação dos experimentos executados, a definição do número de execuções, os tipos de experimentos para cada variante das redes viárias, e as demandas para essas redes, além disso, foram apresentados os três algoritmos utilizados.

4 EXPERIMENTOS REALIZADOS E RESULTADOS

Neste capítulo, na Seção 4.1 são descritos os experimentos realizados para as redes propostas e o porquê de sua escolha. Na Seção 4.2 são apresentados os resultados obtidos e as conclusões sobre os mesmos.

4.1 Experimentos realizados

Antes da apresentação dos experimentos, é necessário salientar que os diversos algoritmos utilizados possuem uma grande quantidade de parâmetros que podem ser configurados, e devido a falta de tempo seria inviável testar todos eles e suas combinações. Portanto foram realizados experimentos com uma gama reduzida de parâmetros, e quem tiver interesse poderá executar outros experimentos apartir dos códigos fonte disponibilizados¹.

A Tabela 4.1 apresenta uma visão geral sobre todos os experimentos realizados para cada rede viária simulada (outros detalhes das redes viárias se encontram nas Tabelas 3.4 e 3.5) utilizando-se uma demanda com duração de 3 horas (10800 passos de tempo de simulação), e terminaram apenas quando todos os veículos deixaram a simulação. Foram realizadas 10 execuções de cada simulação de onde foram extraídas as médias aritméticas e desvios padrão apresentados na Seção 4.2. Para os parâmetros não citados nos experimentos foram utilizados os valores padrão, descritos nas Tabelas 3.6, 3.7, e 3.8.

Nos Experimentos 1-3 optou-se apenas pela variação no valor $\theta_{entrada}$, ou seja, na capacidade de resposta do agente em relação aos planos conhecidos. Um valor menor indica que o agente é contido e prefere manter o plano semaforico atual, um valor maior indica que ele é mais suscetível à mudanças de plano.

No Experimento 4 utilizou-se os valores padrão, ou seja, uma distribuição igual nas taxas de aprendizagem e no seu desconto, assim como no fator de seleção de ações e seu decaimento. O Experimento 5 busca incentivar o agente na aprendizagem de novos planos, tornando-o mais curioso. O Experimento 6 torna o agente menos suscetível à aprendizagem, fazendo com que ele demore mais para aprender novos estados. O Experimento 7 faz com que o agente perca rapidamente seu interesse por mudanças (torna-o guloso).

No Experimento 8 utiliza-se um fator padrão para a qualidade do modelo e a taxa de perda de memória, o Experimento 9 tenta apresentar o efeito da falta de confiança na

¹ Projeto swarm em: <https://code.google.com/p/maslab/>

Tabela 4.1: Visão geral sobre os experimentos realizados.

Sem uso de algoritmos				
Experimento U	Intersecção não controlada.			
Experimento F	Intersecção com plano fixo.			
Colônia de formigas				
	$\theta_{entrada}$ (limiar do agente)			
Experimento 1	0.25			
Experimento 2	0.50			
Experimento 3	0.75			
Aprendizagem Q				
	α (taxa de aprendizagem)	γ (desconto na aprendizagem)	ϵ (fator de seleção de ações)	$\bar{\epsilon}$ (decaimento do fator de seleção de ações)
Experimento 4	0.5	0.5	0.5	0.5
Experimento 5	0.75	0.5	0.75	0.5
Experimento 6	0.25	0.25	0.5	0.5
Experimento 7	0.5	0.75	0.5	0.75
SA-CD				
	ρ (ajuste de qualidade do modelo)		φ (fator da perda de planos)	
Experimento 8	0.5		0.5	
Experimento 9	0.05		0.95	
Experimento 10	0.95		0.05	

qualidade do modelo em conjunto com a alta capacidade de memorização. O Experimento 10 busca demonstrar o efeito da alta confiança na qualidade do modelo unida à baixa capacidade de memorização.

4.2 Resultados

Nas tabelas que seguem, o item **PT** (Passos de Tempo (s)) indica o número de passos de tempo em segundos de uma simulação, **TMV** (Tempo médio de viagem (s)) indica quanto tempo em segundos, na média, os veículos que entraram nessa rede levaram para chegar ao seu destino, o item **Desvio** representa o desvio padrão para aquele experimento. Para os Experimentos 1-10 esses valores são uma média de 10 simulações, para os Experimentos U e F os valores são resultado de uma única simulação, por isso, o desvio padrão dos mesmos é zero. Foi adicionada uma barra de erro ao resultado principal de cada gráfico (PT ou TMV) que representa o desvio padrão daquele conjunto de experimentos apresentado.

4.2.1 Resultados da rede viária de Cruzamento Simples

Em relação ao tempo de simulação, na comparação dos experimentos do Gráfico 4.1a podemos notar um leve ganho do Experimento 1 em relação aos demais e ao plano Fixo, possivelmente devido a seu baixo interesse em trocar de plano semafórico. Entretanto os Experimentos 2 e 3 possuem um desvio considerável e que demonstra que é possível uma melhora ainda maior sobre o plano Fixo após alguns ajustes de seus parâmetros. No Gráfico 4.1b os resultados foram muito parecidos não permitindo nenhuma conclusão. No Gráfico 4.1c apenas o Experimento 10 consegue um melhor desempenho em relação aos outros, além de um baixo desvio, isso demonstra certa estabilidade através da manu-

tenção do plano semafórico atual. No Gráfico 4.2 é possível notar de uma forma mais fina a diferença no resultado dos experimentos com melhor desempenho e do plano Fixo, mesmo assim, ela é muito pequena.

Em relação ao tempo médio de viagem, na comparação dos experimentos do Gráfico 4.3a nota-se um desempenho levemente pior de todos os experimentos em relação ao plano Fixo. O Gráfico 4.3b apresenta um desempenho melhor nos Experimentos 4 e 5, entretanto essa diferença é ínfima. Nos Gráficos 4.3c e 4.4 nada pôde ser afirmado.

Tabela 4.2: Resultados de PT e TMV para rede viária de Cruzamento simples, sem uso de algoritmo.

Cruzamento simples	Experimento U	Experimento F
PT	10824	10876
TMV	33.05	54.12

Tabela 4.3: Resultados de PT e TMV e desvios padrão para rede viária de Cruzamento simples, usando Colônia de formigas.

Cruzamento simples	Experimento 1	Experimento 2	Experimento 3
PT	10855.1	10933.8	11070
Desvio (PT)	4.86	265.74	248.43
TMV	54.934	57.604	69.354
Desvio (TMV)	3.03	7.12	5.05

Tabela 4.4: Resultados de PT e TMV e desvios padrão para rede viária de Cruzamento simples, usando Aprendizagem Q.

Cruzamento simples	Experimento 4	Experimento 5	Experimento 6	Experimento 7
PT	10855.8	10850.8	10888	11059.5
Desvio (PT)	12.80	4.24	93.75	537.71
TMV	52.727	51.949	54.998	58.876
Desvio (TMV)	1.14	0.28	7.17	11.06

Tabela 4.5: Resultados de PT e TMV e desvios padrão para rede viária de Cruzamento simples, usando SA-CD.

Cruzamento simples	Experimento 8	Experimento 9	Experimento 10
PT	10923.1	10874.4	10854.4
Desvio (PT)	127.44	55.22	20.05
TMV	56.26	53.573	55.473
Desvio (TMV)	3.65	2.61	3.45

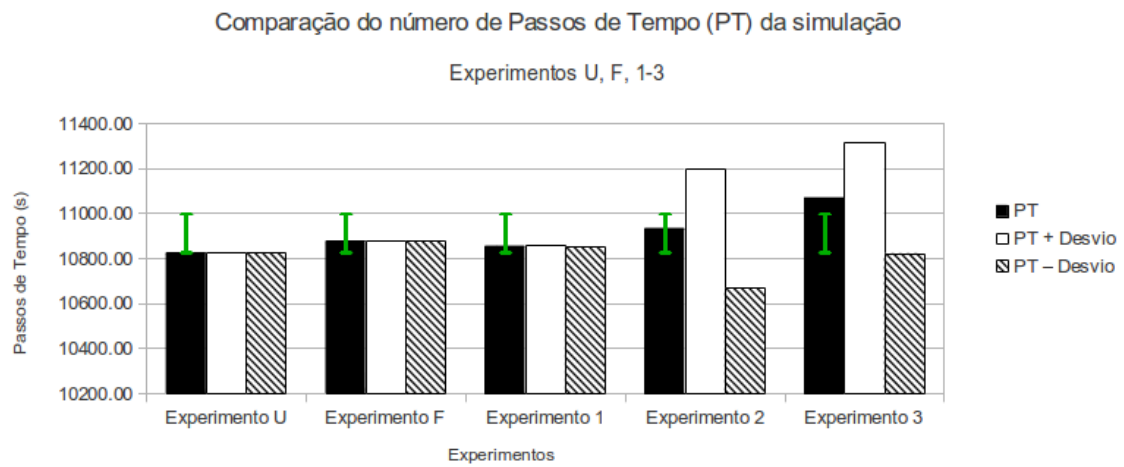
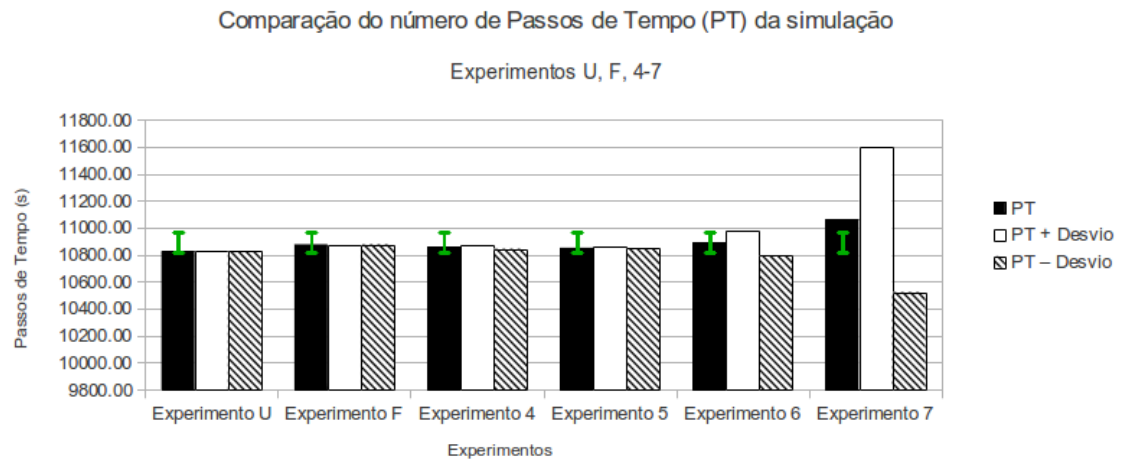
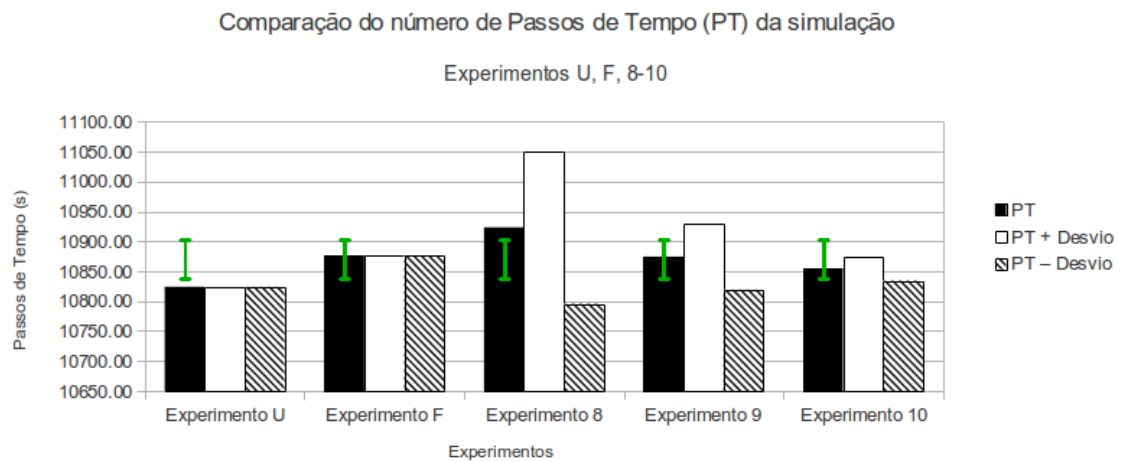
(a) *Colônia de Formigas.*(b) *Aprendizagem Q.*(c) *SA-CD.*

Figura 4.1: Comparação dos resultados de PT na rede viária de Cruzamento simples.

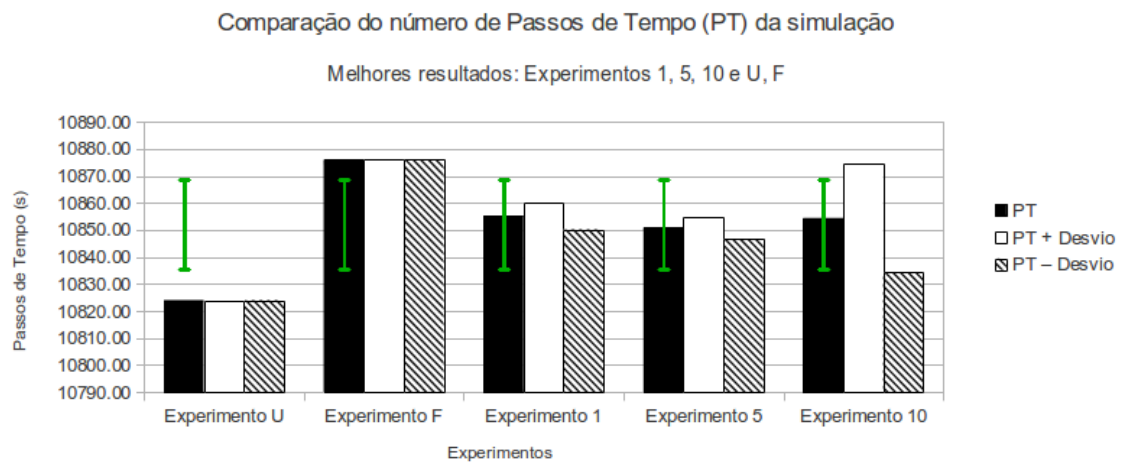
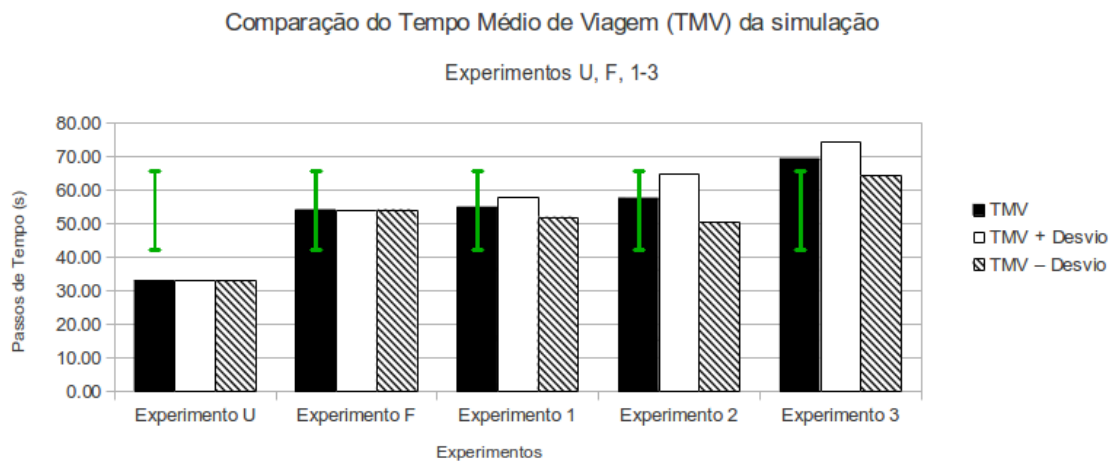
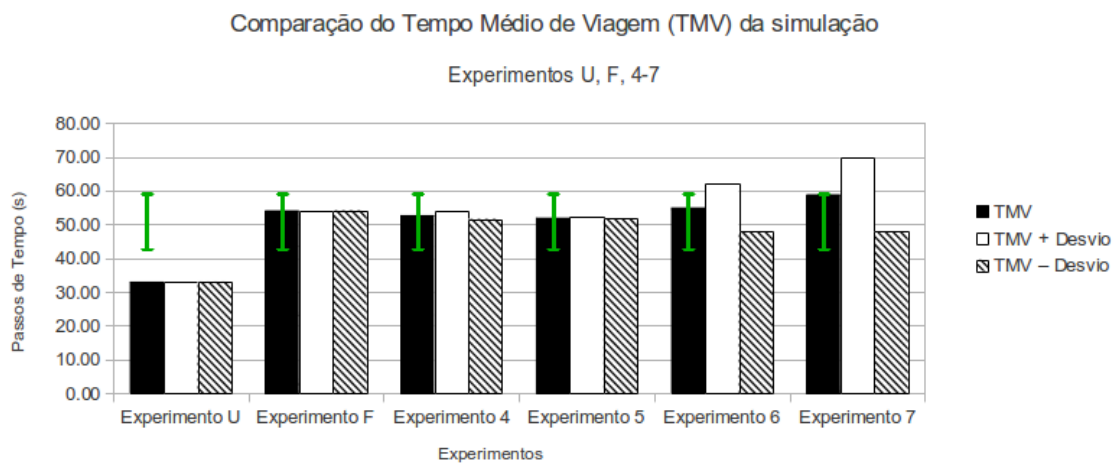


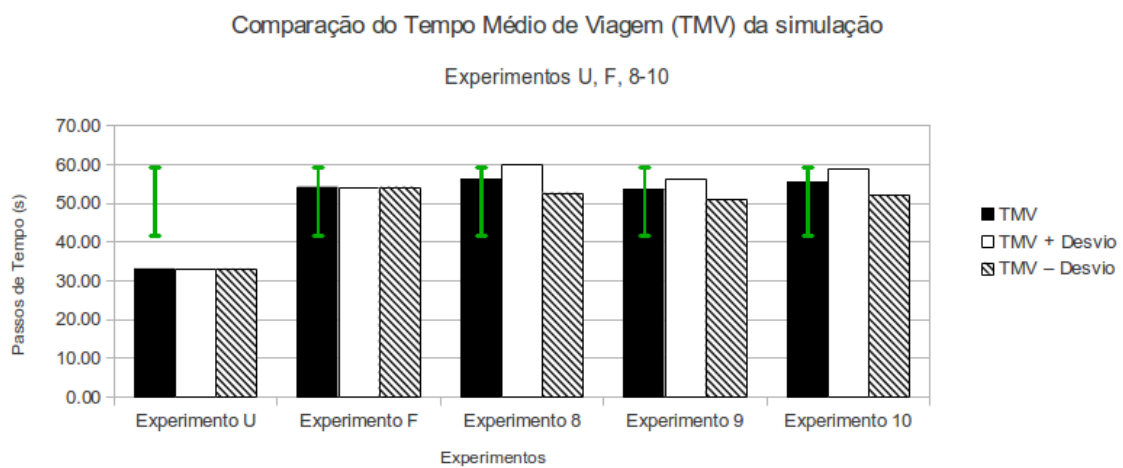
Figura 4.2: Comparação dos melhores resultados de PT na rede viária de Cruzamento simples. Experimentos 1, 5 e 10 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.



(a) *Colônia de Formigas.*



(b) *Aprendizagem Q.*



(c) *SA-CD.*

Figura 4.3: Comparação dos resultados de TMV na rede viária de Cruzamento simples.

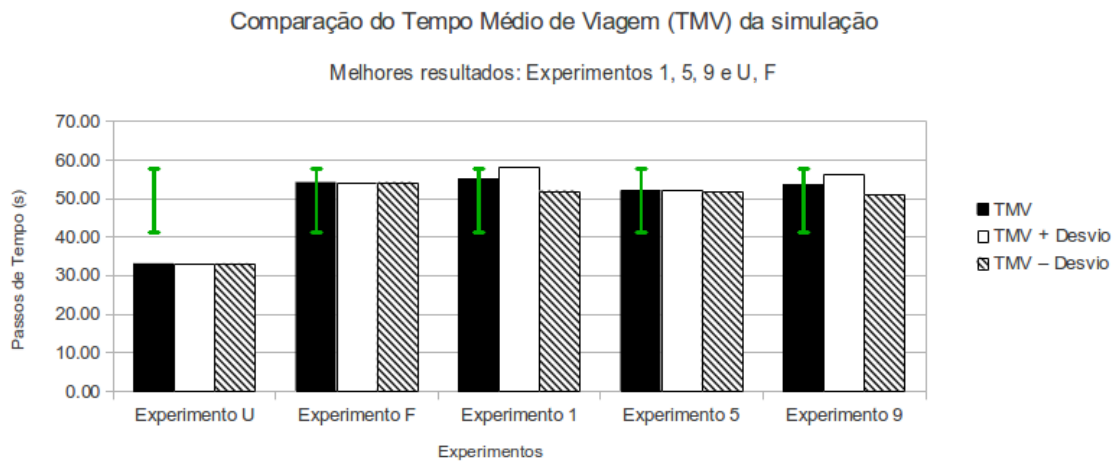


Figura 4.4: Comparação dos melhores resultados de TMV na rede viária de Cruzamento simples. Experimentos 1, 5 e 9 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

4.2.2 Resultados da rede viária de Cruzamento duplo

Em relação ao tempo de simulação, na comparação dos experimentos do Gráfico 4.5a todas as variações de experimentos foram piores que o plano Fixo. No Gráfico 4.5b os resultados estão muito próximos, o Experimento 7 possuindo um resultado pior, mas um desvio que daria margem para um melhor desempenho através de uma correção de parâmetros. No Gráfico 4.5c os Experimentos 8 e 10 possuem um melhor desempenho em relação ao plano Fixo. O Gráfico 4.6 apresenta resultados semelhantes para todos os experimentos à exceção do Experimento 2 que foi pior, porém, o mesmo apresenta um desvio padrão que indica a possibilidade de melhoria de seu resultado.

Em relação ao tempo médio de viagem, na comparação dos experimentos do Gráfico 4.7a todos os experimentos tiveram um desempenho pior que o plano Fixo. No Gráfico 4.7b os Experimentos 4-6 tiveram um resultado levemente melhor que o plano Fixo, porém, a expectativa de que os resultados entre o Experimento 5 (maior aprendizagem) e o Experimento 6 (menor aprendizagem) tivessem um maior contraste não pôde ser observada. O Gráfico 4.7c apresenta resultados melhores nos Experimentos 8-10 em relação ao plano Fixo, porém, essa melhoria é bem pequena. Pelo Gráfico 4.8 podemos notar que o Experimento 5 (Aprendizagem Q) possui melhor desempenho que os outros, se devendo a sua capacidade de exploração maior.

Tabela 4.6: Resultados de PT e TMV para rede viária de Cruzamento duplo, sem uso de algoritmo.

Cruzamento duplo	Experimento U	Experimento F
PT	10841	10889
TMV	42.43	72.64

Tabela 4.7: Resultados de PT e TMV e desvios padrão para rede viária de Cruzamento duplo, usando Colônia de formigas.

Cruzamento duplo	Experimento 1	Experimento 2	Experimento 3
PT	11138.2	11007.4	11267.5
Desvio (PT)	430.83	402.46	320.70
TMV	101.562	80.192	97.572
Desvio (TMV)	39.34	8.97	6.06

Tabela 4.8: Resultados de PT e TMV e desvios padrão para rede viária de Cruzamento duplo, usando Aprendizagem Q.

Cruzamento duplo	Experimento 4	Experimento 5	Experimento 6	Experimento 7
PT	10867.3	10871.5	10873.7	11186.5
Desvio (PT)	3.80	20.26	10.07	770.86
TMV	70.415	67.466	68.055	75.919
Desvio (TMV)	2.98	2.46	2.08	13.74

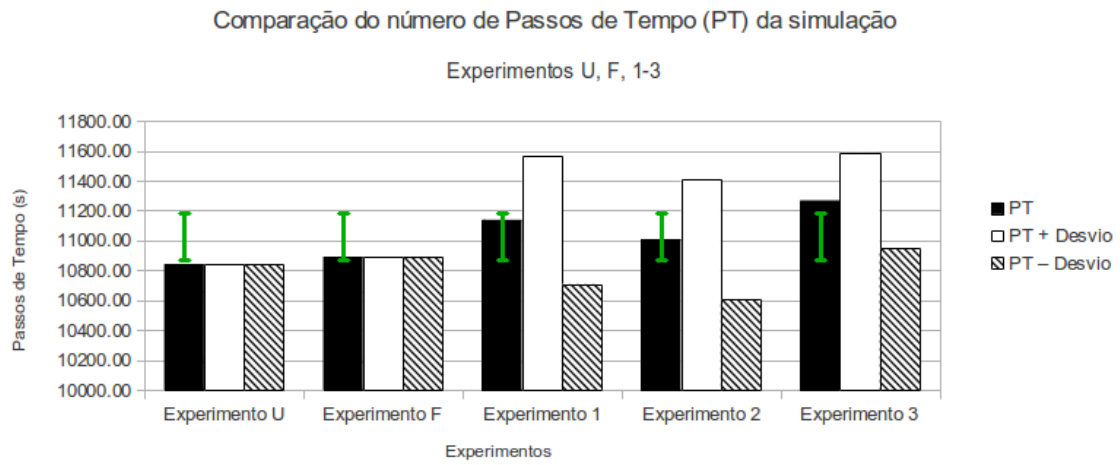
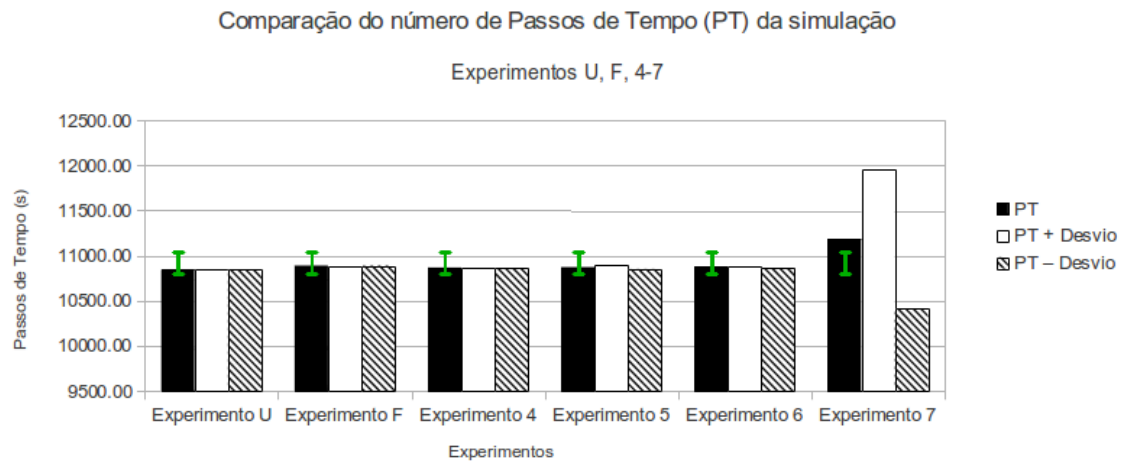
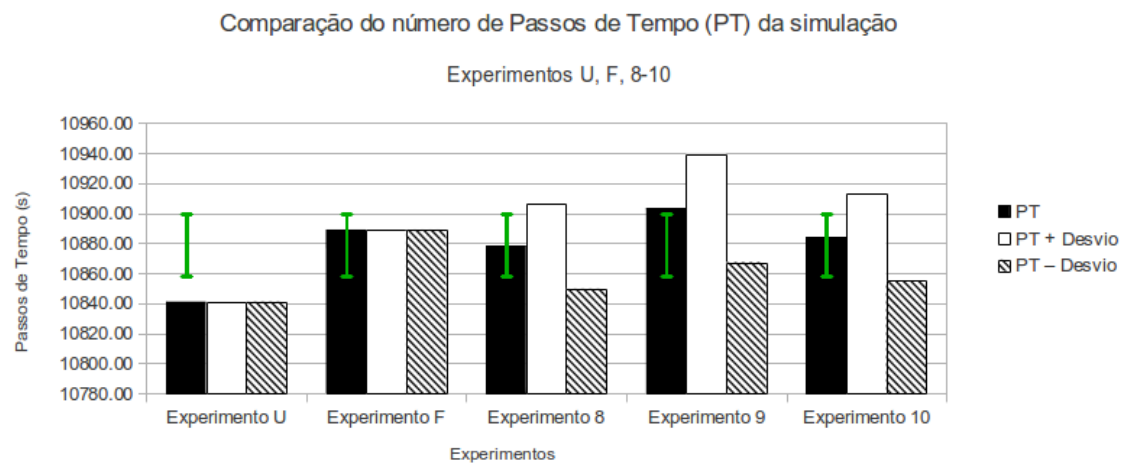
(a) *Colônia de Formigas.*(b) *Aprendizagem Q.*(c) *SA-CD.*

Figura 4.5: Comparação dos resultados de PT na rede viária de Cruzamento duplo.

Tabela 4.9: Resultados de PT e TMV e desvios padrão para rede viária de Cruzamento duplo, usando SA-CD.

Cruzamento duplo	Experimento 8	Experimento 9	Experimento 10
PT	10878.2	10903.4	10883.8
Desvio (PT)	28.60	36.15	29.03
TMV	71.587	69.498	72.236
Desvio (TMV)	2.80	2.77	2.20

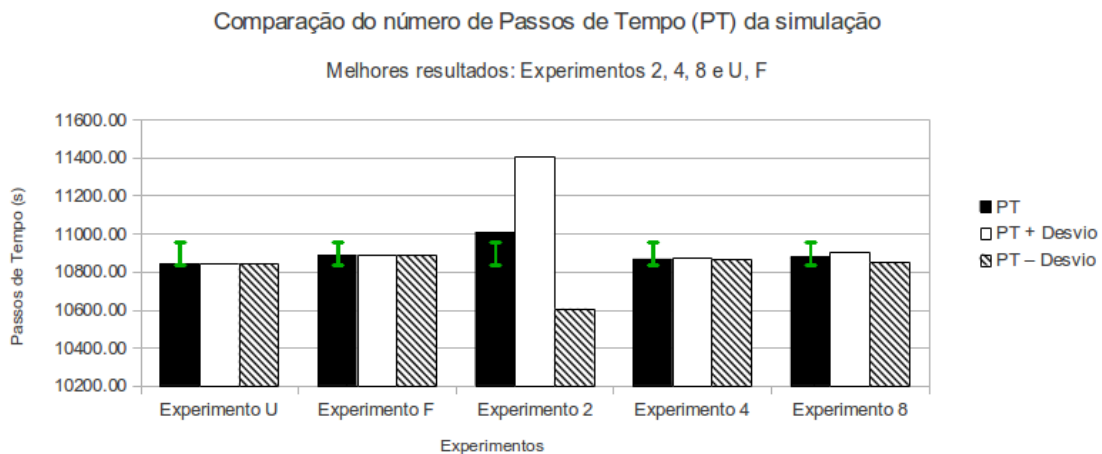
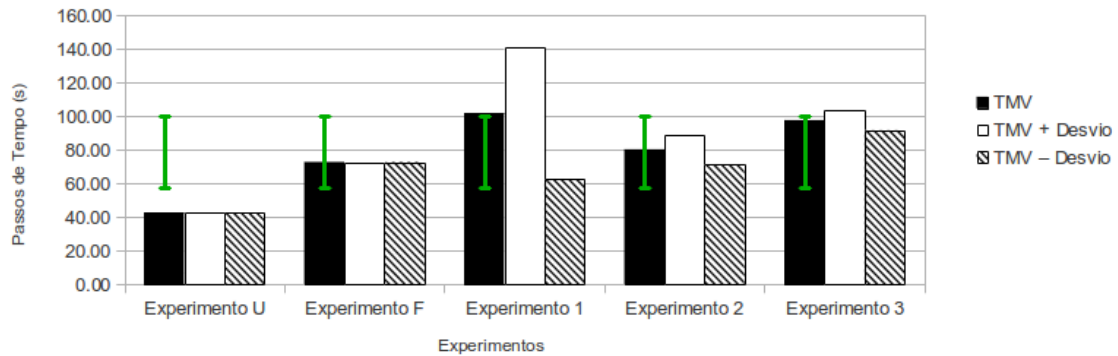


Figura 4.6: Comparação dos melhores resultados de PT na rede viária de Cruzamento duplo. Experimentos 2, 4 e 8 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

Comparação do Tempo Médio de Viagem (TMV) da simulação

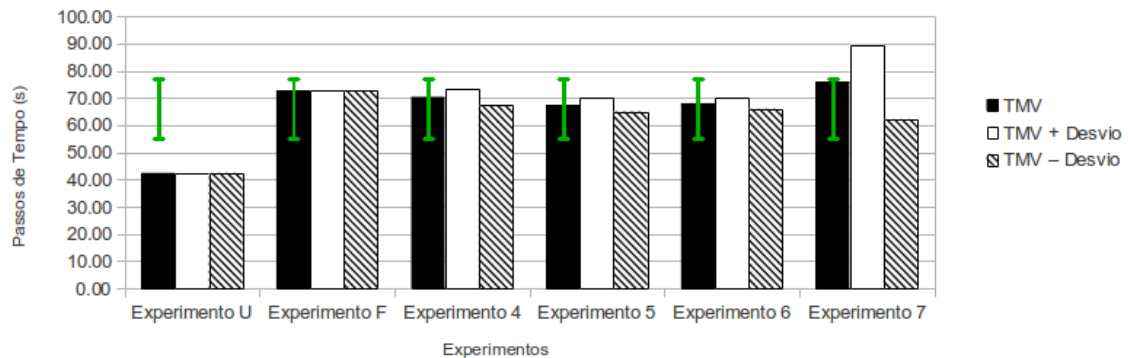
Experimentos U, F, 1-3



(a) *Colônia de Formigas.*

Comparação do Tempo Médio de Viagem (TMV) da simulação

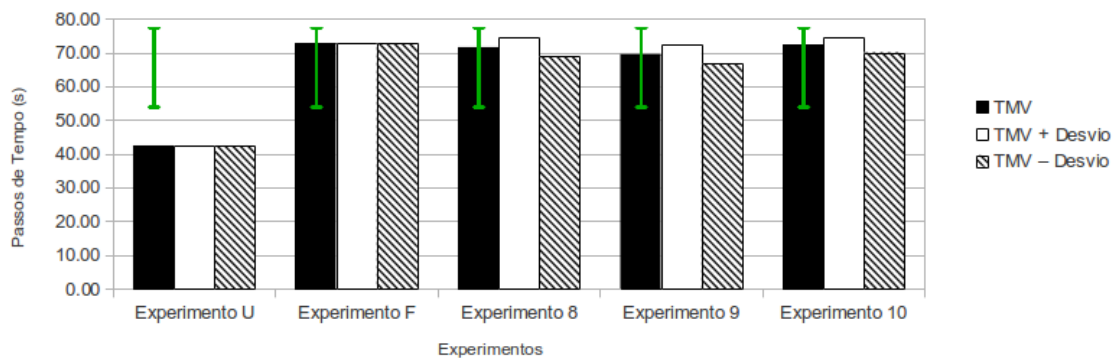
Experimentos U, F, 4-7



(b) *Aprendizagem Q.*

Comparação do Tempo Médio de Viagem (TMV) da simulação

Experimentos U, F, 8-10



(c) *SA-CD.*

Figura 4.7: Comparação dos resultados de TMV na rede viária de Cruzamento duplo.

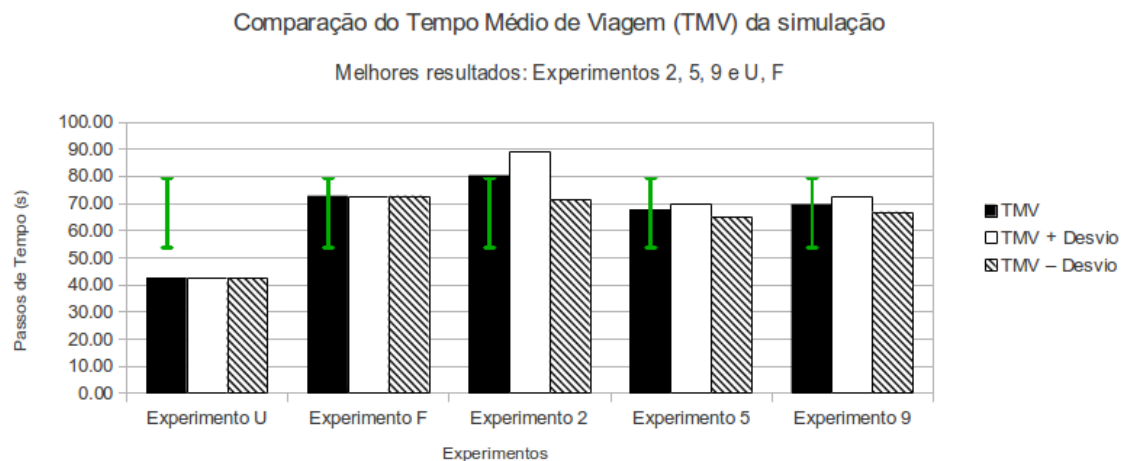


Figura 4.8: Comparação dos melhores resultados de TMV na rede viária de Cruzamento duplo. Experimentos 2, 5 e 9 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

4.2.3 Resultados da rede viária de Formato bloco

Em relação ao tempo de simulação, na comparação dos experimentos do Gráfico 4.9a é possível notar que apenas o Experimento 2 teve um resultado aproximado ao do plano Fixo, também nota-se que o Experimento U (sem semáforo) possui um resultado equivalente ao do plano Fixo o que pode indicar que a longo prazo não seja possível melhorar o desempenho obtido nessa rede. Isso pode ser um problema que está relacionado a manutenção de uma demanda fixa nos experimentos realizados. No Gráfico 4.9b os resultados apresentados são muito parecidos, não havendo um consenso. Os resultados dos experimentos no Gráfico 4.9c apresentam um desempenho ruim em comparação com o plano Fixo. Ao comparar os melhores resultados do Gráfico 4.10 notamos que o Experimento 2 possui uma pequena melhoria em relação ao plano Fixo, isso indica que ao utilizar o algoritmo de Colônia de Formigas é possível alcançar valores melhores, porém, o problema é que este experimento não é direcionado (possui um valor neutro).

Em relação ao tempo médio de viagem, na comparação dos experimentos do Gráfico 4.11a, temo novamente apenas o Experimento 2 com melhor desempenho. O Gráfico 4.11b, apresenta os Experimentos 4-6 com melhores resultados que o plano Fixo, o que indica que existe a possibilidade de melhorar esse resultado com alguns ajustes. No Gráfico 4.11c todos os experimentos tiveram um resultado melhor que o plano Fixo, por um lado isso demonstra que o método pode obter resultados interessantes, por outro existe a necessidade de mais testes, visto que a expectativa era de que os resultados obtidos com os Experimentos 8-10 apresentassem algum contraste perceptível entre eles, não sendo possível afirmar o que está causando melhoria nos resultados. O Gráfico 4.12 mostra todos os resultados com valor próximo ou menor que o plano Fixo, mais uma vez demonstrando que é possível alcançar melhorias, mesmo que pequenas.

Tabela 4.10: Resultados de PT e TMV para rede viária de Formato bloco, sem uso de algoritmo.

Formato bloco	Experimento U	Experimento F
PT	10853	10895
TMV	57.09	103.81

Tabela 4.11: Resultados de PT e TMV e desvios padrão para rede viária de Formato bloco, usando Colônia de formigas.

Formato bloco	Experimento 1	Experimento 2	Experimento 3
PT	11287.6	10888.5	11750.5
Desvio (PT)	734.96	21.09	527.93
TMV	125.201	105.882	130.081
Desvio (TMV)	36.16	3.85	8.56

Tabela 4.12: Resultados de PT e TMV e desvios padrão para rede viária de Formato bloco, usando Aprendizagem Q.

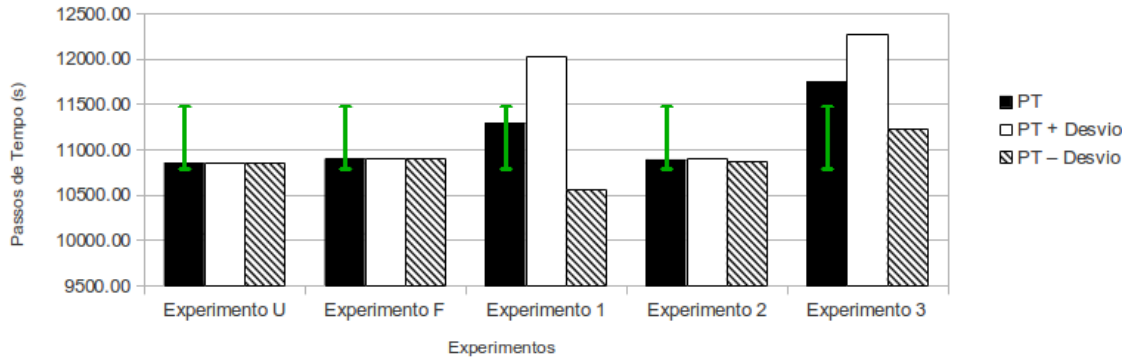
Formato bloco	Experimento 4	Experimento 5	Experimento 6	Experimento 7
PT	10923.1	10916.8	10927.4	11539.9
Desvio (PT)	19.52	20.15	17.83	1001.07
TMV	96.666	94.617	95.661	111.495
Desvio (TMV)	2.86	0.99	2.33	15.19

Tabela 4.13: Resultados de PT e TMV e desvios padrão para rede viária de Formato bloco, usando SA-CD.

Formato bloco	Experimento 8	Experimento 9	Experimento 10
PT	10923.8	10924.3	10918.7
Desvio (PT)	21.22	14.87	13.43
TMV	99.861	99.692	99.9
Desvio (TMV)	2.03	1.89	1.77

Comparação do número de Passos de Tempo (PT) da simulação

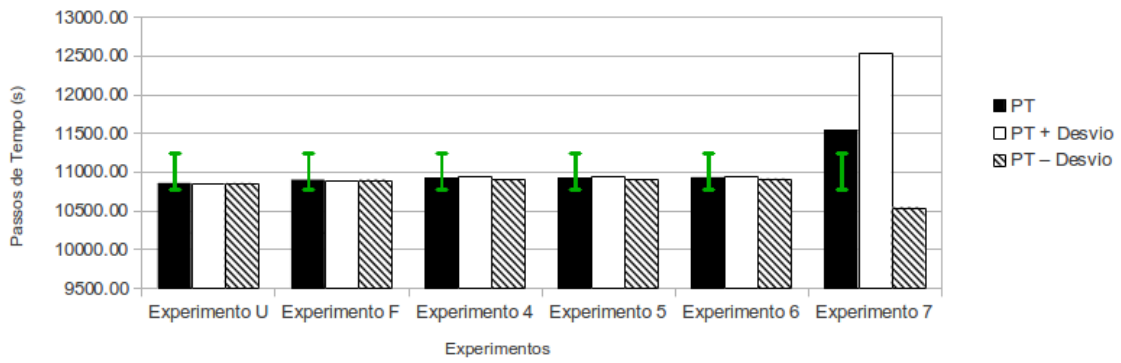
Experimentos U, F, 1-3



(a) *Colônia de Formigas.*

Comparação do número de Passos de Tempo (PT) da simulação

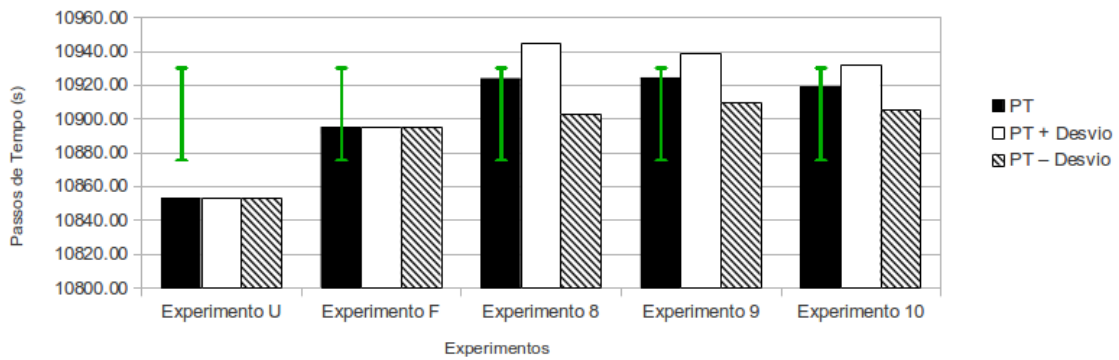
Experimentos U, F, 4-7



(b) *Aprendizagem Q.*

Comparação do número de Passos de Tempo (PT) da simulação

Experimentos U, F, 8-10



(c) *SA-CD.*

Figura 4.9: Comparação dos resultados de PT na rede viária de Formato bloco.

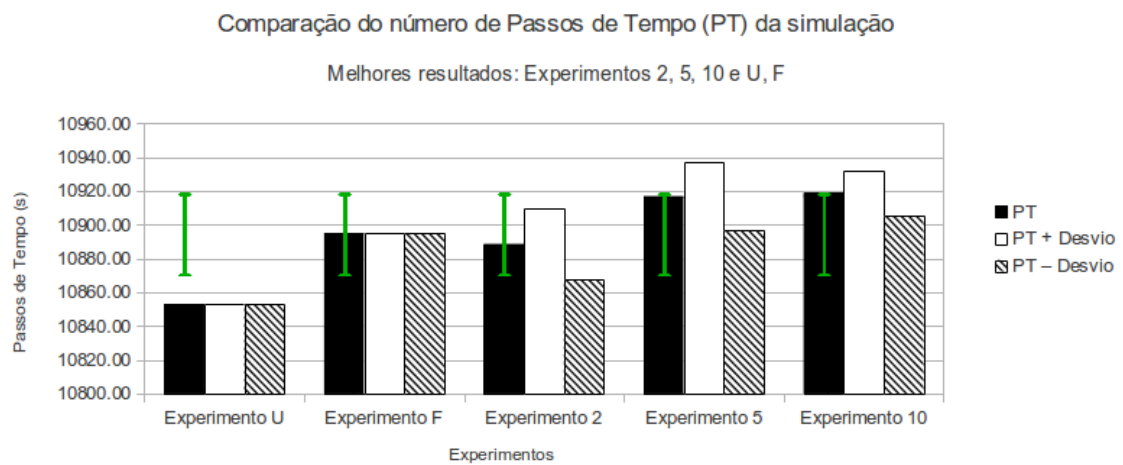


Figura 4.10: Comparação dos melhores resultados de PT na rede viária de Formato bloco. Experimentos 2, 5 e 10 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

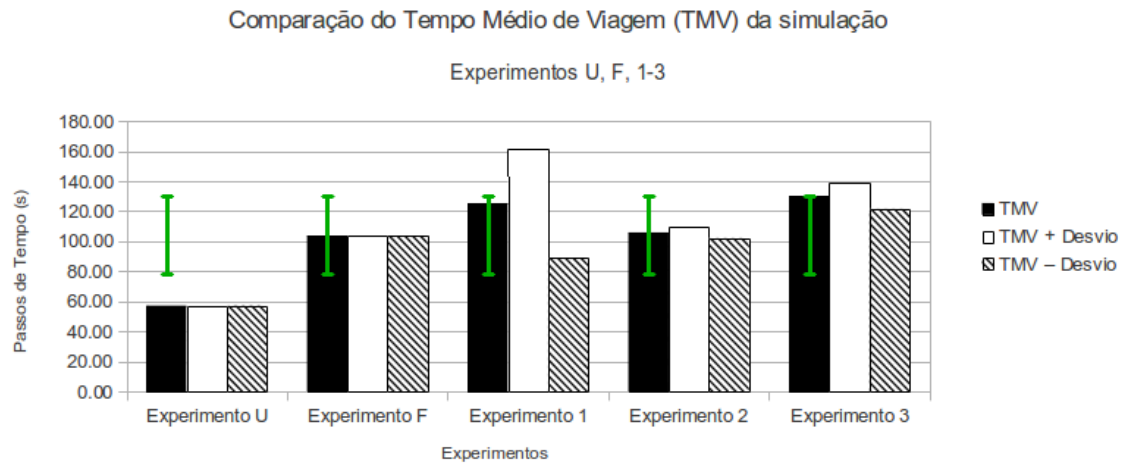
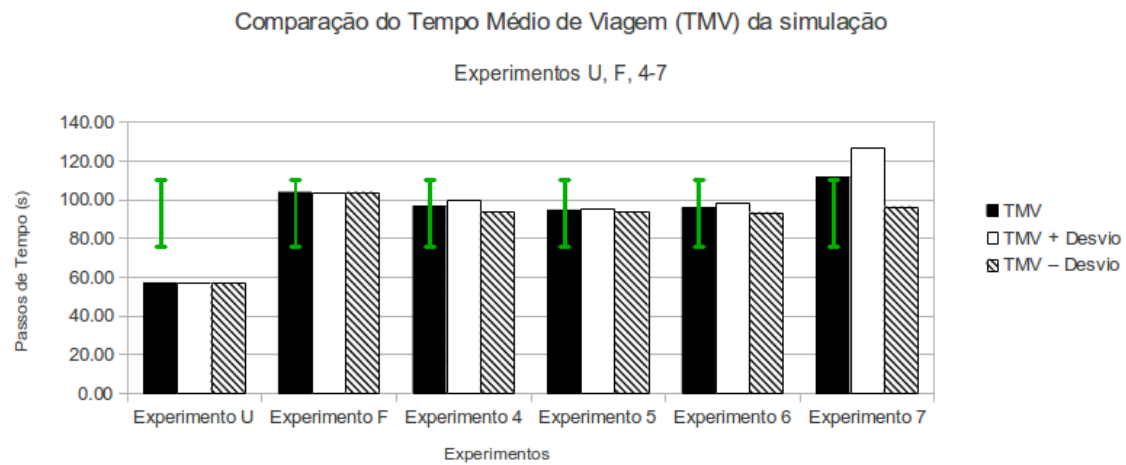
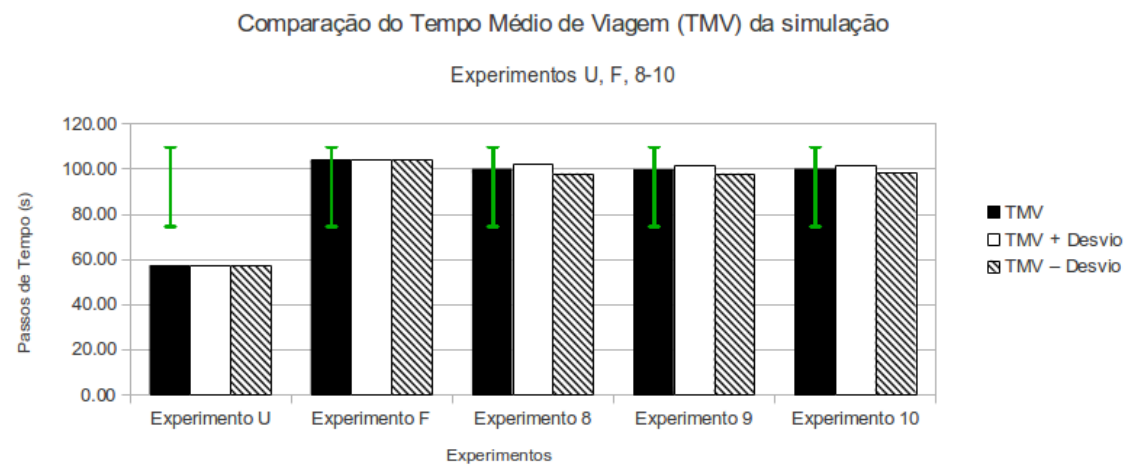
(a) *Colônia de Formigas.*(b) *Aprendizagem Q.*(c) *SA-CD.*

Figura 4.11: Comparação dos resultados de TMV na rede viária de Formato bloco.

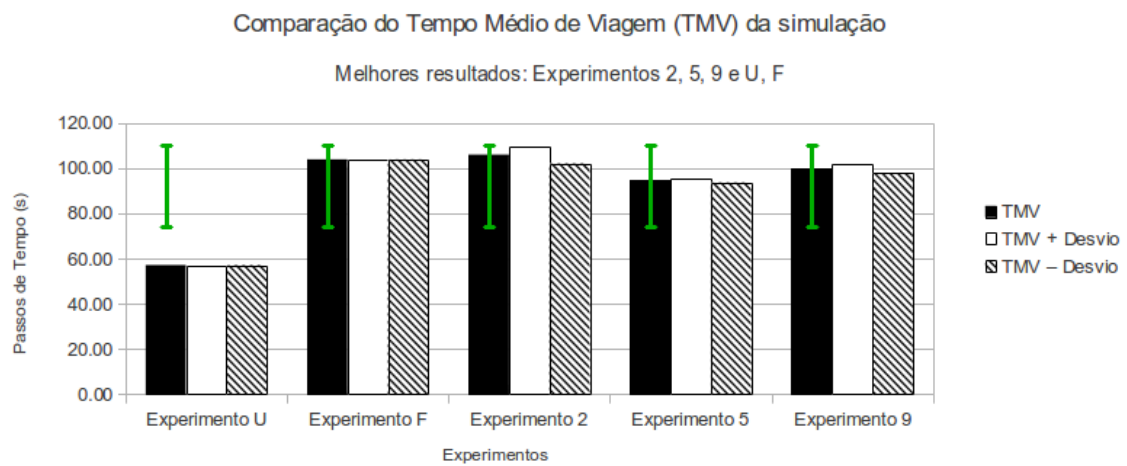


Figura 4.12: Comparação dos melhores resultados de TMV na rede viária de Formato bloco. Experimentos 2, 5 e 9 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

4.2.4 Resultados da rede viária de Sioux Falls

Em relação ao tempo de simulação, na comparação dos experimentos do Gráfico 4.13a nota-se que existe uma redução crescente nas barras dos Experimentos 1-3 (indicando melhoria dos resultados), o que demonstra que ao aumentar a suscetibilidade do agente à troca de planos seu desempenho tende a melhorar. No Gráfico 4.13b como esperado temos um pior desempenho do Experimento 7 (devido a redução de sua capacidade exploratória) e bons resultados em relação aos Experimentos 4-6, reduzindo em quase 10% o tempo de simulação. No Gráfico 4.13c os resultados foram semelhantes aos obtidos pela Aprendizagem Q, porém, eles são muito próximos uns dos outros, não havendo conclusões mais profundas sobre o que aconteceu para que isso ocorresse. O Gráfico 4.14 apresenta como melhor resultado o Experimento 5 (agente mais curioso), mas com os Experimentos 3 e 9 também apresentando resultados melhores que o plano Fixo.

Em relação ao tempo médio de viagem, na comparação dos experimentos do Gráfico 4.15a nota-se que existe uma redução crescente nas barras dos Experimentos 1-3 (indicando melhoria dos resultados), o que demonstra que ao aumentar a suscetibilidade do agente à troca de planos seu desempenho tende a melhorar. No Gráfico 4.15b como esperado temos um pior desempenho do Experimento 7 (devido a redução de sua capacidade exploratória), além disso, os resultados em relação aos Experimentos 4-6 não tiveram um bom desempenho em relação ao plano Fixo. No Gráfico 4.15c os resultados foram semelhantes aos obtidos pela Aprendizagem Q, com nenhum experimento alcançando um resultado satisfatório. O Gráfico 4.16 apresenta os melhores resultados, entretanto, todos foram piores que o plano Fixo, o que indica que essa rede viária foi um pouco mais problemática para a resolução que utiliza de algum dos algoritmos. Mais do que isso, todos os resultados estão muito próximos, não permitindo uma melhor análise do porquê desses resultados.

Tabela 4.14: Resultados de PT e TMV para rede viária de Sioux Falls, sem uso de algoritmo.

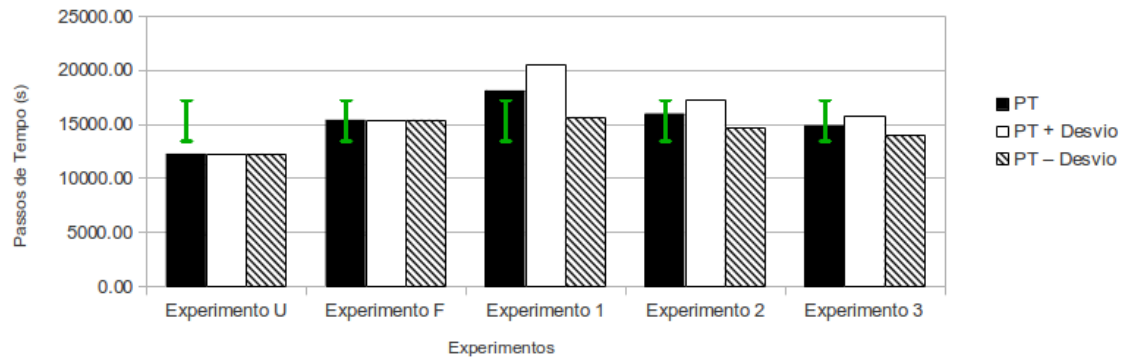
Sioux Falls	Experimento U	Experimento F
PT	12247	15382
TMV	412.78	704.7

Tabela 4.15: Resultados de PT e TMV e desvios padrão para rede viária de Sioux Falls, usando Colônia de formigas.

Sioux Falls	Experimento 1	Experimento 2	Experimento 3
PT	18098.2	15957.4	14862.5
Desvio (PT)	2453.19	1265.71	915.09
TMV	842.824	741.303	728.56
Desvio (TMV)	91.50	53.64	22.49

Comparação do número de Passos de Tempo (PT) da simulação

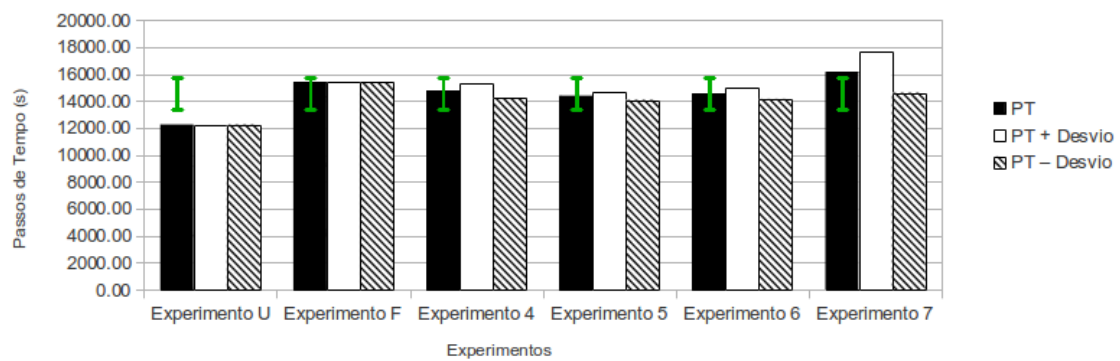
Experimentos U, F, 1-3



(a) *Colônia de Formigas.*

Comparação do número de Passos de Tempo (PT) da simulação

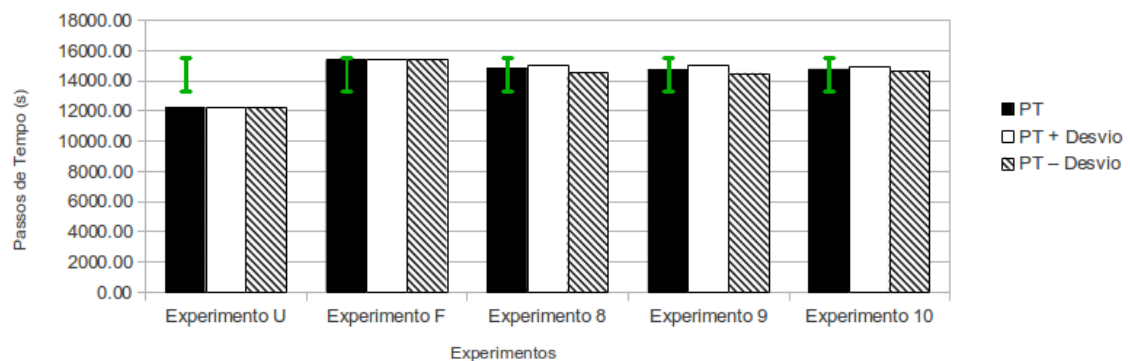
Experimentos U, F, 4-7



(b) *Aprendizagem Q.*

Comparação do número de Passos de Tempo (PT) da simulação

Experimentos U, F, 8-10



(c) *SA-CD.*

Figura 4.13: Comparação dos resultados de PT na rede viária de Sioux Falls.

Tabela 4.16: Resultados de PT e TMV e desvios padrão para rede viária de Sioux Falls, usando Aprendizagem Q.

Sioux Falls	Experimento 4	Experimento 5	Experimento 6	Experimento 7
PT	14766.7	14388.3	14556.2	16144.3
Desvio (PT)	552.39	305.46	403.63	1527.59
TMV	756.979	727.965	736.573	799.229
Desvio (TMV)	27.83	11.66	23.56	49.55

Tabela 4.17: Resultados de PT e TMV e desvios padrão para rede viária de Sioux Falls, usando SA-CD.

Sioux Falls	Experimento 8	Experimento 9	Experimento 10
PT	14799.5	14744.3	14764.3
Desvio (PT)	277.58	265.40	136.52
TMV	734.118	715.797	722.884
Desvio (TMV)	25.65	27.23	22.18

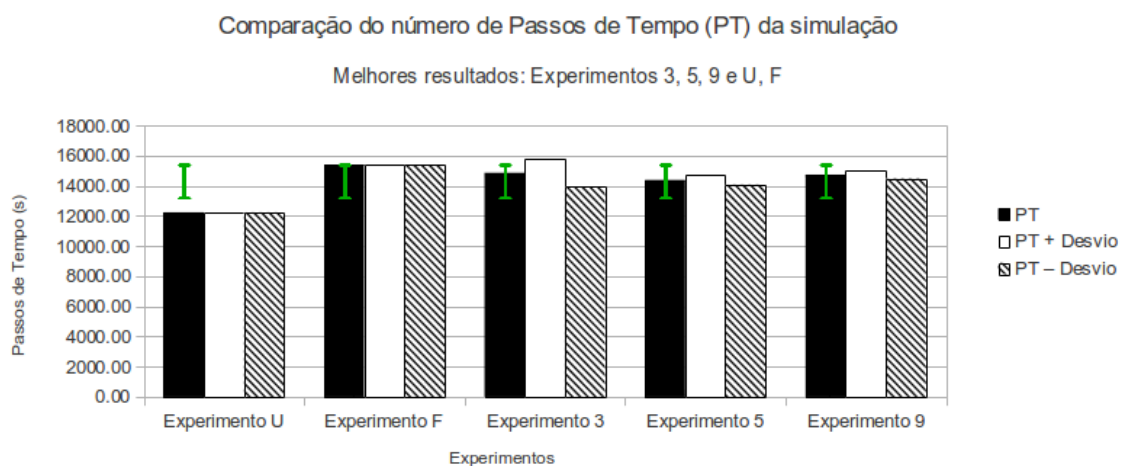


Figura 4.14: Comparação dos melhores resultados de PT na rede viária de Sioux Falls. Experimentos 3, 5 e 9 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

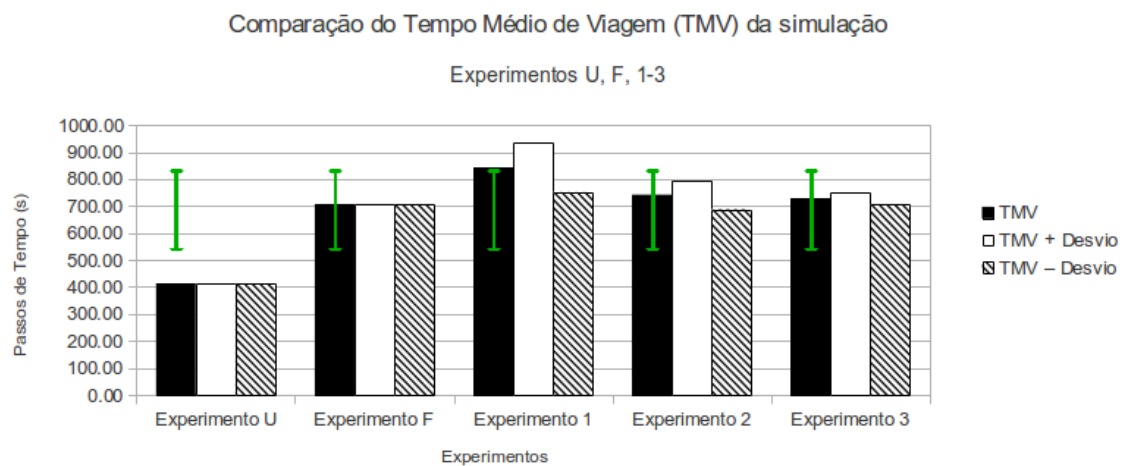
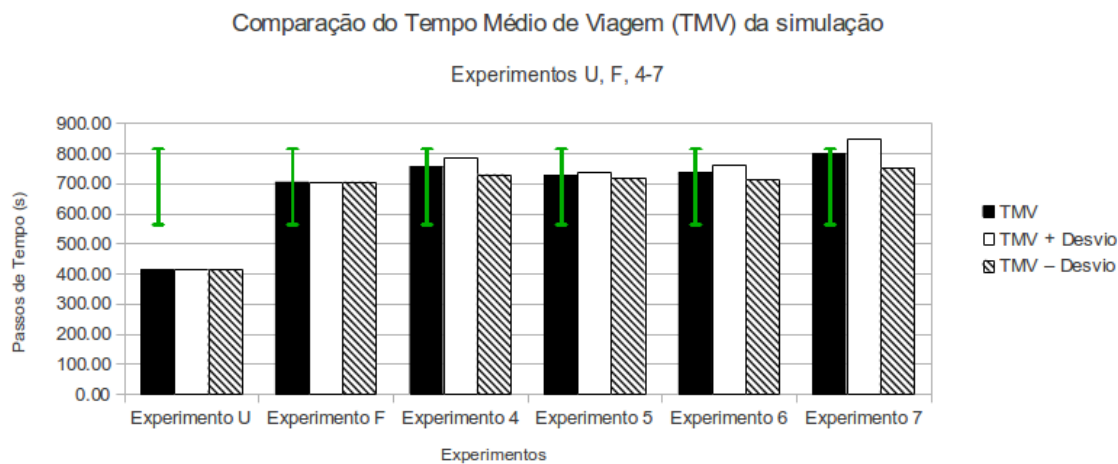
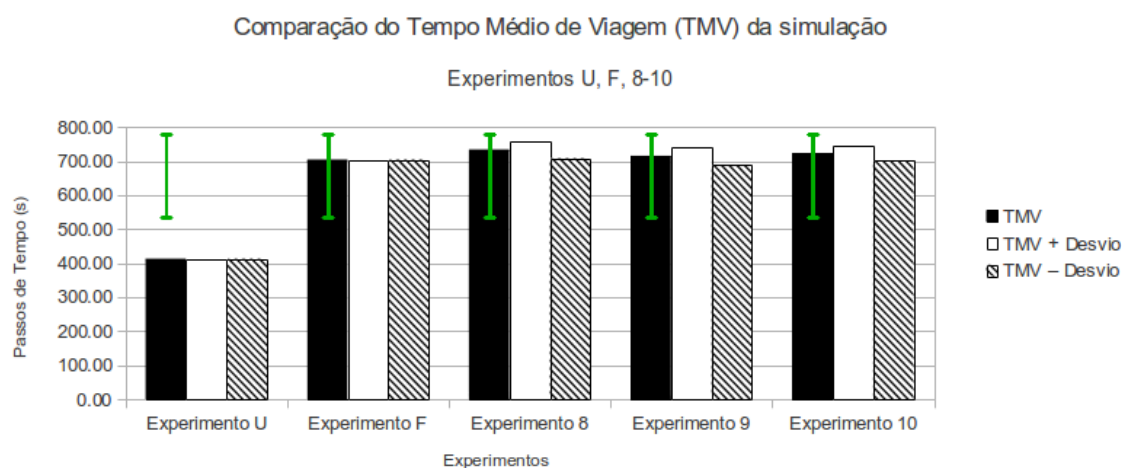
(a) *Colônia de Formigas.*(b) *Aprendizagem Q.*(c) *SA-CD.*

Figura 4.15: Comparação dos resultados de TMV na rede viária de Sioux Falls.

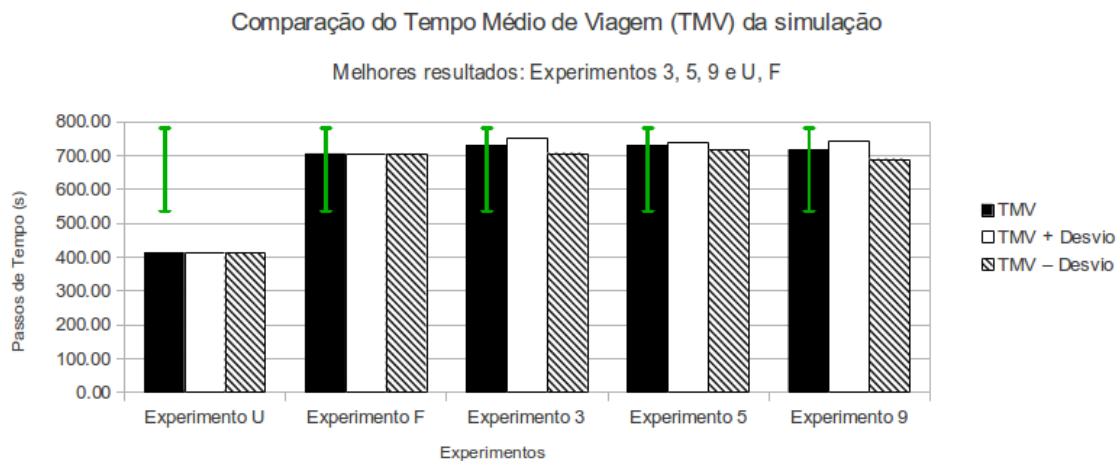


Figura 4.16: Comparação dos melhores resultados de TMV na rede viária de Sioux Falls. Experimentos 3, 5 e 9 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

4.2.5 Resultados da rede viária Poa Arterials

Em relação ao tempo de simulação, na comparação dos experimentos do Gráfico 4.17a podemos notar o decréscimo das barras quando ocorre o aumento no limiar de atuação do agente (Experimentos 1-3 respectivamente), isso é um ótimo sinal, pois indica aqui que bons resultados podem ser alcançados para o uso do algoritmo de Colônia de Formigas, apresentando no Experimento 3 uma melhoria de quase 15%. No Gráfico 4.17b os Experimentos 4 e 5, que possuem maior capacidade exploratória, apresentam bons resultados em relação ao plano Fixo (aproximadamente 10% de melhoria), isso demonstra como o formato das redes viárias influencia nos resultados para um mesmo algoritmo. O Gráfico 4.17c apresenta para todos os algoritmos, resultados inferiores ao do plano Fixo. O Gráfico 4.18 mostra que o desempenho obtido pelo Experimento 3 foi superior a todos os outros.

Em relação ao tempo médio de viagem, na comparação dos experimentos do Gráfico 4.19a podemos notar uma melhoria grande (quase 25%) em relação ao plano Fixo para os Experimentos 2 e 3. No Gráfico 4.19b são apresentados resultados interessantes nos Experimentos 4 e 5, indicando que valores maiores de aprendizagem tendem a levar o modelo a um melhor desempenho global, foram alcançados quase 25% de melhoria nestes 2 casos. O Gráfico 4.19c também apresenta bons resultados em relação ao plano Fixo (cerca de 10%) para todos os 3 experimentos, entretanto, entre si esse desempenho é muito próximo, sendo difícil uma avaliação factível sobre o que os tornou melhores que o plano Fixo. O Gráfico 4.20 apresenta como as 3 abordagens implementadas alcançaram resultados melhores que o plano Fixo num nível de desempenho global do sistema.

Tabela 4.18: Resultados de PT e TMV para rede viária Poa Arterials, sem uso de algoritmo.

Poa Arterials	Experimento U	Experimento F
PT	15531	15045
TMV	1528.63	1266.08

Tabela 4.19: Resultados de PT e TMV e desvios padrão para rede viária de Poa Arterials, usando Colônia de formigas.

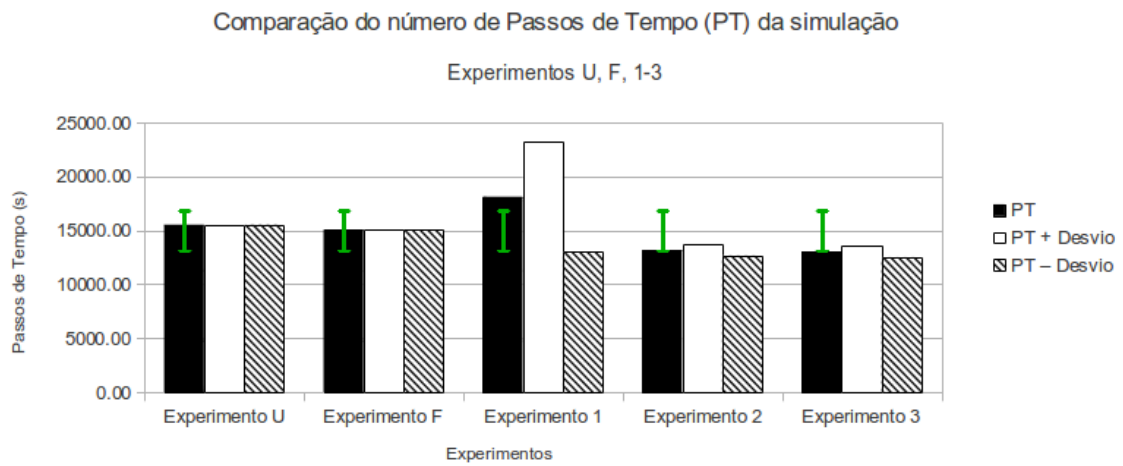
Poa Arterials	Experimento 1	Experimento 2	Experimento 3
PT	18117.5	13163.1	13042.7
Desvio (PT)	5094.03	503.43	522.76
TMV	1395.92	984.054	986.776
Desvio (TMV)	446.72	54.65	50.68

Tabela 4.20: Resultados de PT e TMV e desvios padrão para rede viária de Poa Arterials, usando Aprendizagem Q.

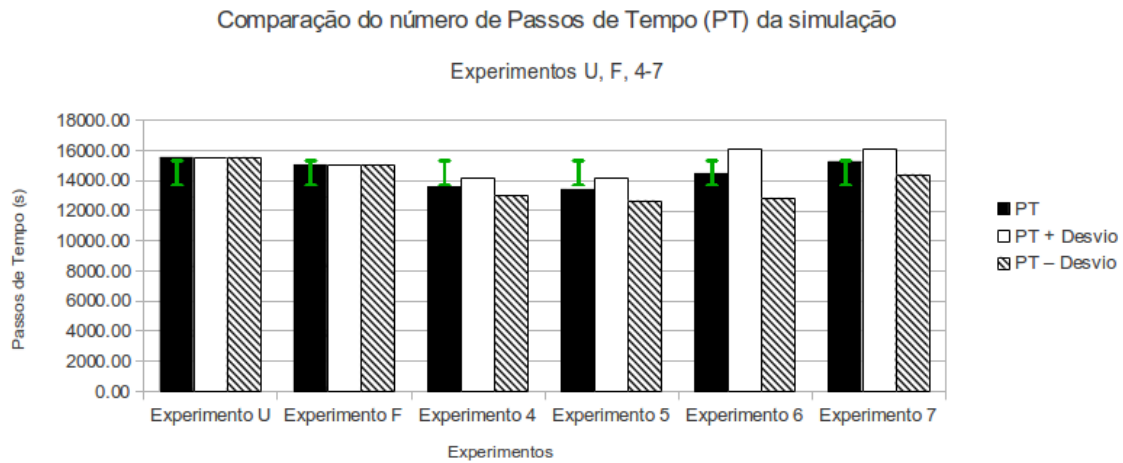
Poa Arterials	Experimento 4	Experimento 5	Experimento 6	Experimento 7
PT	13571.6	13411.5	14467.9	15251.3
Desvio (PT)	559.58	764.04	1634.61	864.05
TMV	948.831	958.625	1065.925	1145.993
Desvio (TMV)	44.81	79.20	178.34	107.94

Tabela 4.21: Resultados de PT e TMV e desvios padrão para rede viária de Poa Arterials, usando SA-CD.

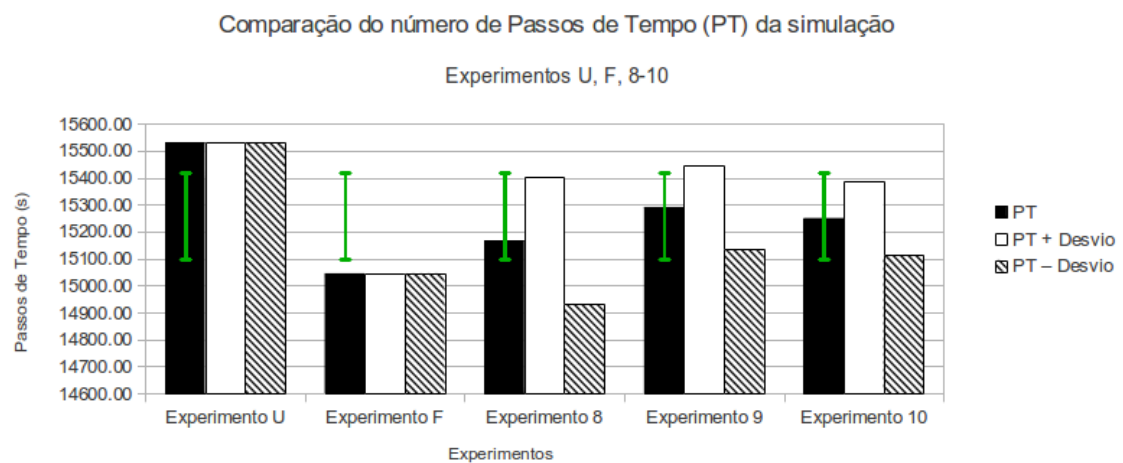
Poa Arterials	Experimento 8	Experimento 9	Experimento 10
PT	15166.8	15291.3	15249.9
Desvio (PT)	234.50	156.59	138.10
TMV	1137.212	1162.829	1144.263
Desvio (TMV)	34.41	32.09	21.92



(a) Colônia de Formigas.



(b) Aprendizagem Q.



(c) SA-CD.

Figura 4.17: Comparação dos resultados de PT na rede viária de Poa Arterials.

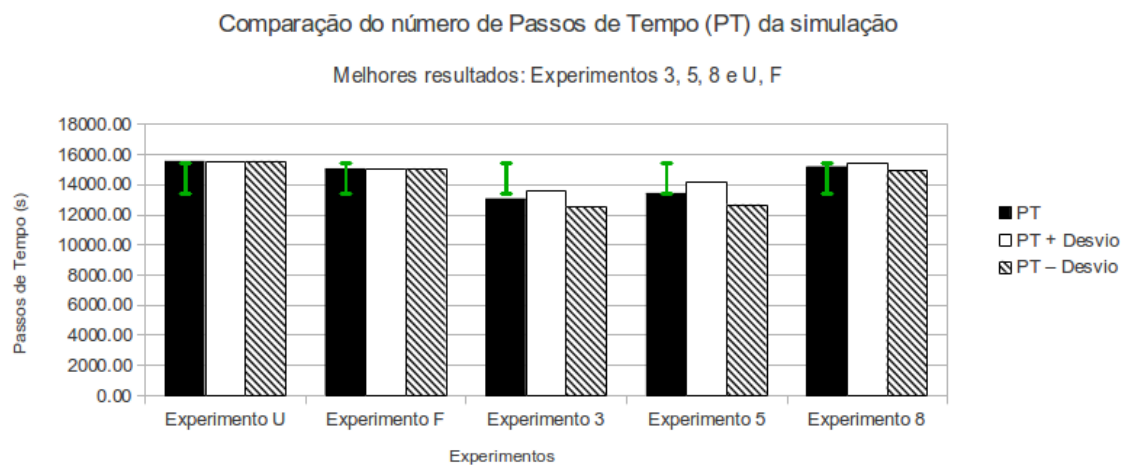
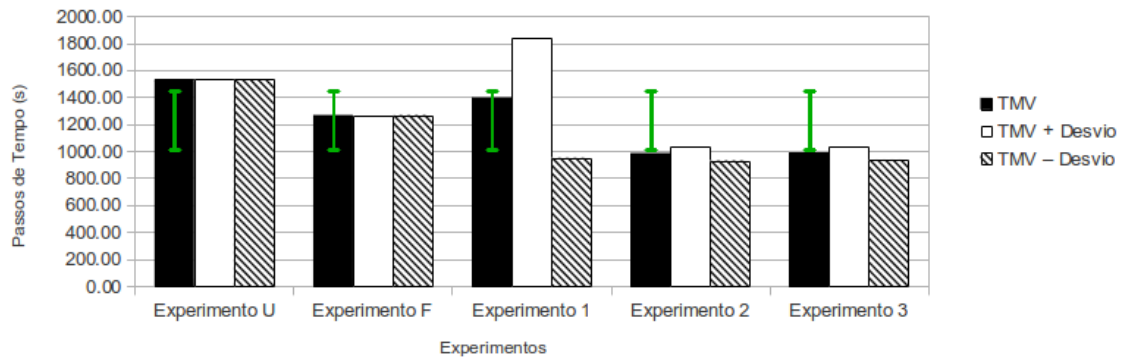


Figura 4.18: Comparação dos melhores resultados de PT na rede viária de Poa Arterials. Experimentos 3, 5 e 8 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

Comparação do Tempo Médio de Viagem (TMV) da simulação

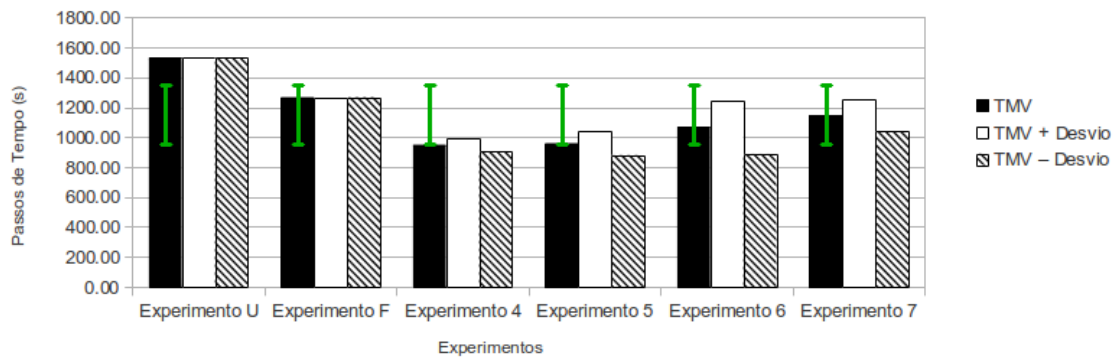
Experimentos U, F, 1-3



(a) *Colônia de Formigas.*

Comparação do Tempo Médio de Viagem (TMV) da simulação

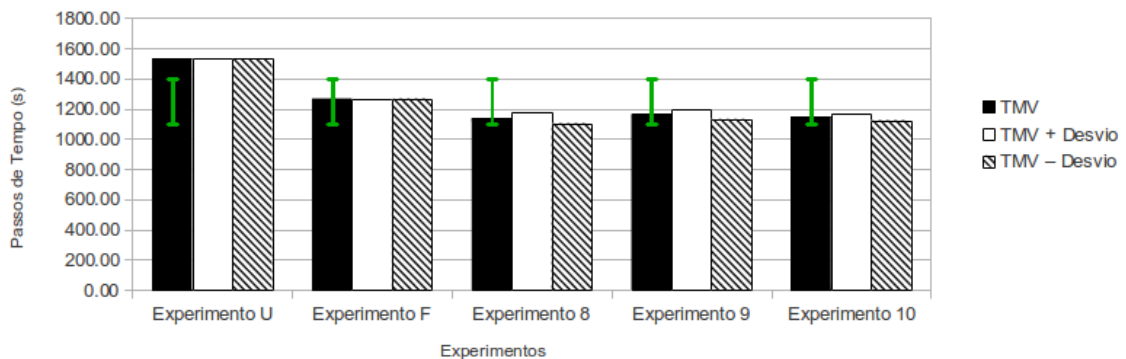
Experimentos U, F, 4-7



(b) *Aprendizagem Q.*

Comparação do Tempo Médio de Viagem (TMV) da simulação

Experimentos U, F, 8-10



(c) *SA-CD.*

Figura 4.19: Comparação dos resultados de TMV na rede viária de Poa Arterials.

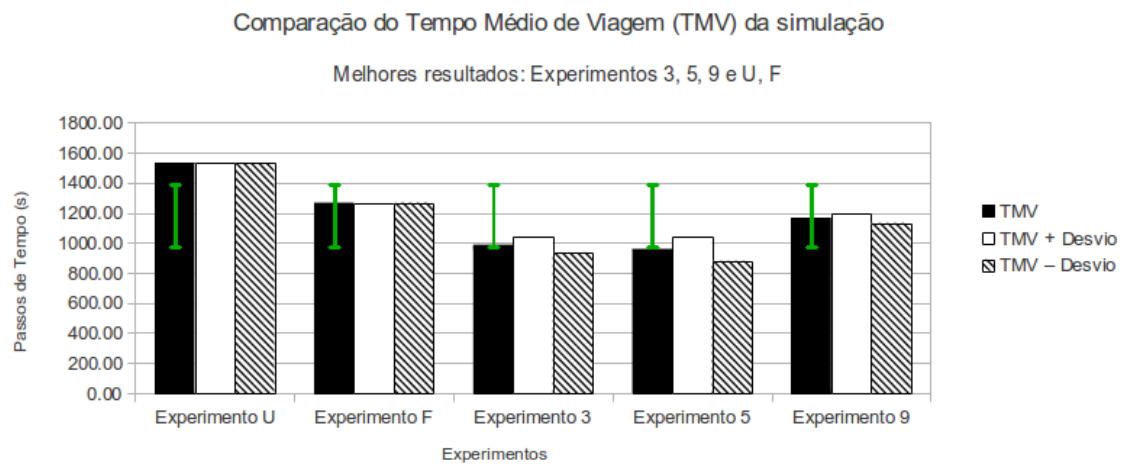


Figura 4.20: Comparação dos melhores resultados de TMV na rede viária de Poa Arterials. Experimentos 3, 5 e 8 correspondem respectivamente a: Colônia de Formigas, Aprendizagem Q e SA-CD.

4.3 Considerações finais

Foram apresentados os resultados obtidos para as 5 redes viárias propostas e para os 3 algoritmos abordados. Em geral os resultados não são os melhores possíveis, mas são motivadores, indicando que existe uma margem de melhoria possível.

O formato da rede viária e o modelo de demanda utilizados são partes muito influentes no resultado dos algoritmos, existe a necessidade de mais testes e ajustes de forma a alcançar um melhor desempenho global.

Na rede Poa Arterials o desempenho dos Experimentos se tornou mais perceptível, indicando que é possível alcançar melhorias em redes mais complexas.

5 CONCLUSÕES E TRABALHOS FUTUROS

Foram apresentadas neste trabalho três técnicas de inteligência artificial, objetivando-se a comparação e combinação das mesmas em um ambiente simulado. Os estudos de sistemas multiagentes permitem ao cientista a capacidade de criar novas sociedades e analisar seu comportamento. A abordagem de trânsito simulado auxilia de forma satisfatória na experimentação dessas sociedades, além de ser um ambiente seguro, não prejudicando os envolvidos que estariam em um sistema real.

Findo os experimentos, analisados os dados, nota-se que a abordagem baseada na Aprendizagem Q obteve melhor desempenho, a técnica SA-CD teve um desempenho inicial satisfatório e possivelmente se mostre de grande valia com um melhor ajuste dos parâmetros. A técnica de Colônia de formigas é muito interessante e possivelmente também teria melhores resultados após uma fase de aprimoramento.

Foram utilizadas demandas fixas nas redes iniciais experimentadas, as redes principais possuíam matrizes-OD, entretanto existe a necessidade de fazer experimentos com outras formas de demandas, de forma a dar validade a modelagem apresentada, de forma a verificar que não existem vieses nos resultados obtidos.

5.1 Trabalhos futuros

Dentre os possíveis estudos que poderiam ser feitos no futuro relacionados a este trabalho posso citar os seguintes:

- a avaliação dos mesmos sistemas supondo que os planos semaforicos pudessem ter um tempo de ciclo variável;
- a análise de desempenho das redes viárias experimentadas na condição de falha de partes do sistema semaforico;
- estudo de impacto: ambiental (poluição), de aumento de demandas em determinados distritos (vias mais populosas), na substituição de veículos pessoais por coletivos, bicicletas ou alternativos;
- colaboração entre múltiplas colônias de insetos (agentes semaforos), fugindo um pouco da realidade biológica, qual seria o impacto de grupos isolados de semaforos trabalhando juntos, supondo que nem todas as vias possuam controle semaforico mas algumas compartilhem determinadas conexões viárias.

5.2 Dificuldades encontradas

Existe uma grande dificuldade na obtenção de informações abertas e de fácil manipulação relacionadas ao trânsito urbano, e em especial sobre ocupação e congestionamentos das vias urbanas. Sem esses dados, é necessária uma grande carga de testes apenas na modelagem das redes viárias simuladas de forma a perceber os seus limites. Se os mesmos dados estivessem ao alcance dos pesquisadores, haveria uma melhor percepção sobre o problema a ser tratado e conseqüentemente os resultados de experimentos sobre esses dados obteriam uma melhor qualidade condizente com a situação real do trânsito.

Creio que acarretada pela falta de dados citada anteriormente, também existe a dificuldade em encontrar modelos de rede viária previamente testados e validados com suas demandas, pode não parecer, mas com existência de tais redes, poderíamos passar a comparar os algoritmos de uma forma muito mais simples, além de não ser mais necessário um dispêndio de tempo na modelagem.

Mesmo com diversas ferramentas voltadas para simulação de sistemas complexos, as de trânsito são escassas, e ainda boa parte delas é paga. Ferramentas de simulação gratuitas, a exemplo do SUMO, são muito importantes e deveriam ter uma maior incentivo de produção, principalmente para prova de conceitos na academia.

APÊNDICE A MATRIZES ORIGEM-DESTINO

A.1 Rede viária da cidade de Sioux Falls

Na rede viária da cidade de Sioux Falls, foram definidos 24 distritos, um para cada intersecção, considerados os pontos de origem e destino de viagens, os dados são apresentados na Tabela A.1 onde a primeira coluna contém cada distrito de origem de viagens e a primeira linha contém cada distrito de destino dessas viagens. O valor contido nas células indica a quantidade proporcional de veículos que irá de um distrito para outro, por exemplo, existem 44 viagens partindo do distrito 10 para o distrito 16, de um total de 3606 viagens demandadas, isso equivale a aproximadamente 1,2% de todas as viagens nessa rede.

Tabela A.1: Matriz OD da rede viária de Sioux Falls. Cada valor é a proporção de viagens geradas em um total de 3606 (soma de todos os valores)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0	1	1	5	2	3	5	8	5	13	5	2	5	3	5	5	4	1	3	3	1	4	3	1
2	1	0	1	2	1	4	2	4	2	6	2	1	3	1	1	4	2	0	1	1	0	1	0	0
3	1	1	0	2	1	3	1	2	1	3	3	2	1	1	1	2	1	0	0	0	0	1	1	0
4	5	2	2	0	5	4	4	7	7	12	14	6	6	5	5	8	5	1	2	3	2	4	5	2
5	2	1	1	5	0	2	2	5	8	10	5	2	2	1	2	5	2	0	1	1	1	2	1	0
6	3	4	3	4	2	0	4	8	4	8	4	2	2	1	2	9	5	1	2	3	1	2	1	1
7	5	2	1	4	2	4	0	10	6	19	5	7	4	2	5	14	10	2	4	5	2	5	2	1
8	8	4	2	7	5	8	10	0	8	16	8	6	6	4	6	22	14	3	7	9	4	5	3	2
9	5	2	1	7	8	4	6	8	0	28	14	6	6	6	9	14	9	2	4	6	3	7	5	2
10	13	6	3	12	10	8	19	16	28	0	40	20	19	21	40	44	39	7	18	25	12	26	18	8
11	5	2	3	15	5	4	5	8	14	39	0	14	10	16	14	14	10	1	4	6	4	11	13	6
12	2	1	2	6	2	2	7	6	6	20	14	0	13	7	7	7	6	2	3	4	3	7	7	5
13	5	3	1	6	2	2	4	6	6	19	10	13	0	6	7	6	5	1	3	6	6	13	8	8
14	3	1	1	5	1	1	2	4	6	21	16	7	6	0	13	7	7	1	3	5	4	12	11	4
15	5	1	1	5	2	2	5	6	10	40	14	7	7	13	0	12	15	2	8	11	8	26	10	4
16	5	4	2	8	5	9	14	22	14	44	14	7	6	7	12	0	28	5	13	16	6	12	5	3
17	4	2	1	5	2	5	10	14	9	39	10	6	5	7	15	28	0	6	17	17	6	17	6	3
18	1	0	0	1	0	1	2	3	2	7	2	2	1	1	2	5	6	0	3	4	1	3	1	0
19	3	1	0	2	1	2	4	7	4	18	4	3	3	3	8	13	17	3	0	12	4	12	3	1
20	3	1	0	3	1	3	5	9	6	25	6	5	6	5	11	16	17	4	12	0	12	24	7	4
21	1	0	0	2	1	1	2	4	3	12	4	3	6	4	8	6	6	1	4	12	0	18	7	5
22	4	1	1	4	2	2	5	5	7	26	11	7	13	12	26	12	17	3	12	24	18	0	21	11
23	3	0	1	5	1	1	2	3	5	18	13	7	8	11	10	5	6	1	3	7	7	21	0	7
24	1	0	0	2	0	1	1	2	2	8	6	5	7	4	4	3	3	0	1	4	5	11	7	0

A.2 Rede contendo as vias arteriais de Porto Alegre

Na rede das vias arteriais de Porto Alegre, foram definidos 15 distritos considerados os principais pontos de origem e destino de viagens, os dados são apresentados na Tabela A.2 onde a primeira coluna contém cada distrito de origem de viagens e a primeira linha contém cada distrito de destino dessas viagens. O valor contido nas células indica a quantidade proporcional de veículos que irá de um distrito para outro, por exemplo, existem 4 viagens partindo do distrito 6 para o distrito 8, de um total de 210 viagens demandadas, isso equivale a aproximadamente 1,9% de todas as viagens nessa rede.

Tabela A.2: Matriz OD da rede viária de Porto Alegre. Cada valor é a proporção de viagens geradas em um total de 210 (soma de todos os valores).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	1	1	1	1	2	2	2	1	1	1	1	1	1	1
2	1	0	1	1	1	4	4	4	1	1	1	1	0	1	1
3	1	1	0	1	1	2	2	2	1	1	1	1	1	1	1
4	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0
5	0	0	0	1	0	1	1	1	1	1	1	1	0	1	1
6	1	1	1	1	1	0	4	4	1	1	1	1	0	1	1
7	1	0	1	1	1	1	0	1	1	1	1	1	0	1	1
8	1	1	0	0	0	1	1	0	1	1	1	1	1	1	1
9	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
10	0	0	1	1	1	2	2	2	1	0	1	1	1	1	1
11	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1
12	0	0	1	1	1	2	2	2	1	1	1	0	1	1	1
13	1	1	1	1	1	2	2	2	1	1	1	1	0	1	1
14	1	1	1	1	1	2	2	2	1	1	1	1	0	0	1
15	1	1	1	1	1	2	2	2	1	1	1	1	1	1	0

REFERÊNCIAS

Maicon de Brito do Amarante. Uma abordagem baseada em agentes para avaliação do balanceamento de carga em redes veiculares: dois estudos de caso. Dissertação de mestrado, Instituto de Informática, UFRGS, Porto Alegre, Brazil, Abril 2012. URL <http://www.lume.ufrgs.br/handle/10183/49744>.

J. Barceló. *Fundamentals of Traffic Simulation*. International series in operations research & management science. Springer, 2010. ISBN 9781441961426.

Ana L. C. Bazzan. Opportunities for multiagent systems and multiagent reinforcement learning in traffic control. *Autonomous Agents and Multiagent Systems*, 18(3):342–375, June 2009. doi: 10.1007/s10458-008-9062-9. URL <http://www.springerlink.com/content/j1j0817117r8j18r/>.

Ana L. C. Bazzan. Lessons learned from one decade of developing agent-based tools for traffic modeling, simulation, and control: how to make cities smarter. In *Braz. Symp. on Information Systems (SBSI)*, pages 67–72. SBC, May 2012. URL www.inf.ufrgs.br/maslab/pergamus/pubs/Bazzan2012sbsi.pdf.

Ana L. C. Bazzan and Franziska Klügl. Sistemas inteligentes de transporte e tráfego: uma abordagem de tecnologia da informação. In Tomasz Kowaltowski and Karin Koogan Breitman, editors, *Anais das Jornadas de Atualização em Informática*. SBC, Rio de Janeiro, Brasil, Julho 2007. ISBN 978-85-7669-116-7. URL <http://www.inf.ufrgs.br/maslab/pergamus/pubs/jai07BazzanKluegl.zip>.

Ana L. C. Bazzan, Maicon de Brito do Amarante, Tiago Sommer, and Alexander J. Benavides. ITSUMO: an agent-based simulator for ITS applications. In Rosaldo Rossetti, Henry Liu, and Shuming Tang, editors, *Proc. of the 4th Workshop on Artificial Transportation Systems and Simulation*. IEEE, September 2010. URL www.inf.ufrgs.br/maslab/pergamus/pubs/ATSS_Bazzan.pdf.

Gerardo Beni and Jing Wang. Swarm intelligence in cellular robotic systems. In Paolo Dario, Giulio Sandini, and Patrick Aebischer, editors, *Robots and Biological Systems: Towards a New Bionics?*, volume 102 of *NATO ASI Series*, pages 703–712. Springer-Verlag Berlin Heidelberg, 1993. ISBN 978-3-642-63461-1. doi: 10.1007/978-3-642-58069-7_38. URL http://dx.doi.org/10.1007/978-3-642-58069-7_38.

Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Proceedings volume in the Santa Fe Institute studies in the sciences of complexity. Oxford University Press, USA, 1999. ISBN 9780195131598.

L. Busoniu, R. Babuska, and B. De Schutter. A comprehensive survey of multiagent reinforcement learning. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(2):156–172, 2008. ISSN 1094-6977. doi: 10.1109/TSMCC.2007.913919.

Debora M. Gordon. *Formigas em Ação*. Ciência e Cultura. Jorge Zahar Editor Ltda, 2002. ISBN 9788571106642.

Henrique Vieira Hallmann. Comparação entre softwares simuladores de trânsito. Trabalho de conclusão (graduação), UFRGS, 2011. URL <http://www.lume.ufrgs.br/handle/10183/36891>.

Bernhard Hofmann-Wellenhof, Herbert Lichtenegger, and James Collins. *Global Positioning System: Theory and Practice*. Springer-Verlag, 1997.

G. Kotushevski and K. A. Hawick. A review of traffic simulation software. Technical Report CSTN-095, Computer Science, Massey University, Albany, North Shore 102-904, Auckland, New Zealand, 2009. URL <http://www.massey.ac.nz/~kahawick/cstn/095/cstn-095.pdf>.

D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker. Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3&4):128–138, December 2012. URL <http://elib.dlr.de/80483/>.

Daniel Krajzewicz. Kombination von taktischen und strategischen einflüssen in einer mikroskopischen verkehrsflussimulation. *Fahrermodellierung in Wissenschaft und Wirtschaft*, 2. Berliner Fachtagung für Fahrermodellierung, pages 104–115, 2009. ISSN 1439-958X.

Stefan Krauß. *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*. PhD thesis, 1998. URL <http://e-archive.informatik.uni-koeln.de/319/>.

Larry J. LeBlanc, Edward K. Morlok, and William P. Pierskalla. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research*, 9(5):309–318, 1975. ISSN 0041-1647. doi: [http://dx.doi.org/10.1016/0041-1647\(75\)90030-1](http://dx.doi.org/10.1016/0041-1647(75)90030-1). URL <http://www.sciencedirect.com/science/article/pii/0041164775900301>.

Kai Nagel and Michael Schreckenberg. A cellular automaton model for freeway traffic. *Journal de Physique I*, 2(12):2221–2229, December 1992. ISSN 1155-4304. doi: 10.1051/jp1:1992277. URL <http://dx.doi.org/10.1051/jp1:1992277>.

Denise de Oliveira. Um estudo de coordenação dinâmica de agentes aplicado ao gerenciamento de tráfego veicular urbano. Dissertação de mestrado, Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil, Abril 2005. URL <http://hdl.handle.net/10183/6224>.

Denise de Oliveira, Paulo R. Ferreira, Jr., Ana L. C. Bazzan, and Franziska Klügl. A swarm-based approach for selection of signal plans in urban scenarios. In *Proceedings of Fourth International Workshop on Ant Colony Optimization and Swarm Intelligence* -

ANTS 2004, volume 3172 of *Lecture Notes in Computer Science*, pages 416–417, Berlin, Germany, 2004. URL <http://inf.ufrgs.br/maslab/pergamus/pubs/2-antsTrafficCR.pdf.tar.gz>.

Denise de Oliveira, Ana L. C. Bazzan, Paulo R. Ferreira, Jr., and Bruno C. Silva. Coordenação dinâmica de semáforos: dois casos de estudo. In *V Encontro Nacional de Inteligência Artificial (ENIA05)*, pages 761–770, 2005a. URL <http://inf.ufrgs.br/maslab/pergamus/pubs/1-denise1.pdf.tar.gz>.

Denise de Oliveira, Ana L. C. Bazzan, and V. Lesser. Using cooperative mediation to coordinate traffic lights: a case study. In Frank Dignum, Virginia Dignum, Sven Koenig, Sarit Kraus, Munindar P. Singh, and Michael Wooldridge, editors, *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS)*, pages 463–470. New York, IEEE Computer Society, July 2005b. URL http://inf.ufrgs.br/maslab/pergamus/pubs/pa4b3_463.pdf.

Johan Janson Olstam and Andreas Tapani. Comparison of car-following models. Technical report, Swedish National Road Administration, June 2004.

R.P. Roess, E.S. Prassas, and W.R. McShane. *Traffic engineering*. Pearson/Prentice Hall, 2004. ISBN 9780131424715.

Guido van Rossum. *The Python Language Reference*, 1995. URL <http://docs.python.org/release/2.7/reference/index.html>.

S. Russell and P. Norvig. *Inteligência Artificial*. Editora Campus, 2004.

Bruno C. Da Silva, Eduardo W. Basso, Ana L. C. Bazzan, and Paulo M. Engel. Dealing with non-stationary environments using context detection. In *In Proceedings of the 23rd International Conference on Machine Learning (ICML 2006)*, pages 217–224. ACM Press, 2006a.

Bruno C. da Silva, Denise de Oliveira, Ana L. C. Bazzan, and E. W. Basso. Adaptive traffic control with reinforcement learning. In Ana L. C. Bazzan, Brahim Chaib-Draa, Franziska Klügl, and Sascha Ossowski, editors, *Proceedings of the 4th Workshop on Agents in Traffic and Transportation (at AAMAS 2006)*, pages 80–86, May 2006b. URL www.inf.ufrgs.br/~bazzan/downloads/WS28ATT.pdf.

Bruno Castro da Silva. Um estudo sobre métodos de aprendizagem de conceitos e de modelos de mundo. Trabalho individual i, PPGC/UFRGS, 2005.

Bruno Castro da Silva, Eduardo Wisnieski Basso, Ana Lucia Cetertich Bazzan, and Paulo Martins Engel. Improving reinforcement learning with context detection. In *Proceedings of the 5th International Joint Conference On Autonomous Agents And Multiagent Systems, AAMAS, 2006*, pages 811–812, Hakodate, Japan, May 2006c. New York, ACM Press. ISBN 1-59593-303-4. URL www.inf.ufrgs.br/maslab/pergamus/pubs/Silva+2006.pdf.

Anderson R. Tavares. Uma abordagem baseada em agentes para simulação de tarifação viária e comunicação inter-veicular. Dissertação de mestrado, Instituto de Informática, UFRGS, Porto Alegre, Brazil, Julho 2013. URL <http://www.lume.ufrgs.br/handle/10183/79589>.

Anderson R. Tavares and Ana L. C. Bazzan. Reinforcement learning for route choice in an abstract traffic scenario. In *VI Workshop-Escola de Sistemas de Agentes, seus Ambientes e aplicações (WESAAC)*, pages 141–153, 2012. URL www.inf.ufrgs.br/maslab/pergamus/pubs/TavaresBazzan2012-wesaac.pdf.

Anderson R. Tavares and Ana L. C. Bazzan. Networked malicious agents in transportation systems: a simulation study. In *The Fifth International Workshop on Emergent Intelligence on Networked Agents (WEIN)*, 2013. URL www.inf.ufrgs.br/maslab/pergamus/pubs/TavaresBazzan2013-wein.pdf.

B. Widrow. Generalization and Information Storage in Networks of Adaline 'Neurons. *Self-Organizing Systems, symposium proceedings*, pages 435–461, 1962.

Michael J. Wooldridge. *An Introduction to MultiAgent Systems*. Wiley Publishing, 2nd edition, 2009. ISBN 0470519460, 9780470519462.