

Metodologia de Pesquisa para Ciência da Computação

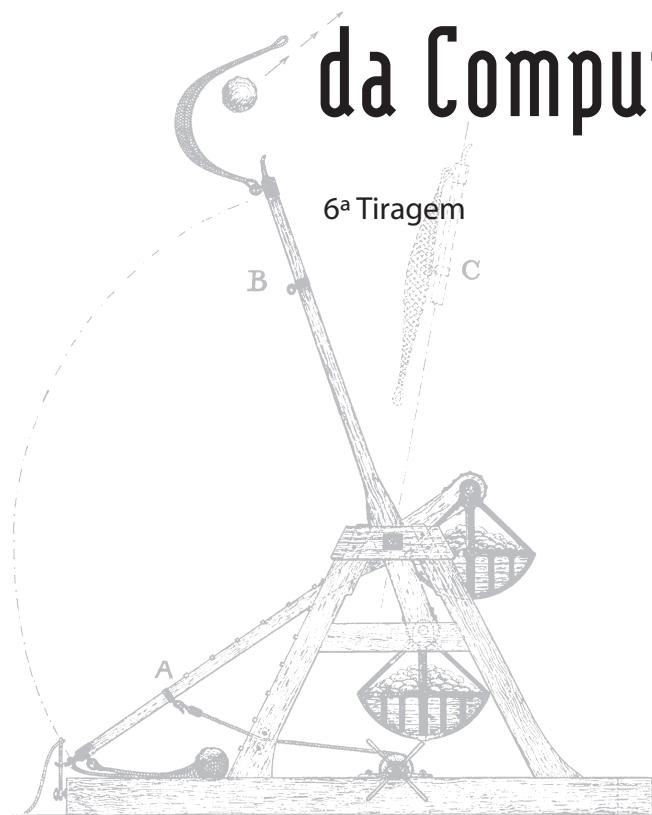


Preencha a **ficha de cadastro** no final deste livro
e receba gratuitamente informações
sobre os lançamentos e as promoções da Elsevier.

Consulte também nosso catálogo completo,
últimos lançamentos e serviços exclusivos no site
www.elsevier.com.br

Raul Sidnei Wazlawick

Metodologia de Pesquisa para Ciência da Computação



© 2009, Elsevier Editora Ltda.

Todos os direitos reservados e protegidos pela Lei nº 9.610, de 19/02/1998.

Nenhuma parte deste livro, sem autorização prévia por escrito da editora, poderá ser reproduzida ou transmitida, sejam quais forem os meios empregados: eletrônicos, mecânicos, fotográficos, gravação ou quaisquer outros.

Copidesque: Adriana Araújo Kramer

Revisão: Marco Antônio Corrêa

Editoração Eletrônica: SBNigri Artes e Textos Ltda.

Elsevier Editora Ltda.

Conhecimento sem Fronteiras

Rua Sete de Setembro, 111 – 16º andar

20050-006 – Centro – Rio de Janeiro – RJ – Brasil

Rua Quintana, 753 – 8º andar

04569-011 – Brooklin – São Paulo – SP – Brasil

Serviço de Atendimento ao Cliente

0800-0265340

sac@elsevier.com.br

ISBN 978-85-352-3522-7

Nota: Muito zelo e técnica foram empregados na edição desta obra. No entanto, podem ocorrer erros de digitação, impressão ou dúvida conceitual. Em qualquer das hipóteses, solicitamos a comunicação ao nosso Serviço de Atendimento ao Cliente, para que possamos esclarecer ou encaminhar a questão.

Nem a editora nem o autor assumem qualquer responsabilidade por eventuais danos ou perdas a pessoas ou bens, originados do uso desta publicação.

CIP-Brasil. Catalogação-na-fonte.
Sindicato Nacional dos Editores de Livros, RJ

W372m Wazlawick, Raul Sidnei, 1967

Metodologia de pesquisa para ciência da computação/Raul
Sidnei Wazlawick. – Rio de Janeiro: Elsevier, 2009. 6ª reimpressão

Inclui bibliografia

ISBN 978-85-352-3522-7

1. Computação – Pesquisa – Metodologia. 2. Pesquisa –
Metodologia. 3. Redação técnica. I. Título.

09-1884.

CDD: 004.072

CDU: 004.001.8

Este livro é dedicado aos meus anjinhos:
Júlia, Amanda, Alice e Carlinha.

Agradecimentos

Agradeço a este livro pela compreensão por tantas horas em que não o escrevi porque estava dedicado à minha família.

Biografia do Autor

Raul Sidnei Wazlawick é professor associado II da Universidade Federal de Santa Catarina (UFSC), lotado no Departamento de Informática e Estatística. Graduado em Ciência da Computação, UFSC, 1988. Mestre em Ciência da Computação, UFRGS, 1991. Doutor em Engenharia de Produção, UFSC, 1993. Pós-doutorado pela Universidade Nova de Lisboa (UNL), 1998.

Conselheiro da Sociedade Brasileira de Computação (SBC). Ex-membro da Comissão de Especialistas de Ensino de Computação e Informática do Ministério da Educação (MEC). Ex-representante do Brasil na International Federation for Information Processing (IFIP). Ex-coordenador do IFIP *Working Group on Higher Education* (WG 3.2). Ex-coordenador do Curso de Bacharelado em Ciência da Computação, UFSC. Ex-coordenador do Programa de Pós-Graduação em Ciência da Computação, UFSC. Ex-coordenador de vários cursos de pós-graduação *lato sensu* da UFSC. Ex-membro da Câmara de Pesquisa da UFSC. Criador e ex-editor da *Revista Brasileira de Informática na Educação* da SBC.

Coordenou vários eventos científicos no Brasil e no exterior, incluindo o Congresso da SBC (2002). Foi presidente de comitês de programa de eventos científicos no Brasil e exterior. Membro de comitê de programa de dezenas de eventos científicos. Membro de comitê de revisores de periódicos científicos. Orientador de dezenas de dissertações de mestrado, teses

de doutorado, monografias de especialização e trabalhos de conclusão de cursos de graduação. Membro de mais de uma centena de bancas de mestrado, doutorado, especialização e graduação em várias universidades brasileiras. Conta com quase uma centena de publicações em periódicos e eventos científicos e é autor do livro *Análise e Projeto de Sistemas de Informação Orientados a Objetos*, de 2004, pertencente à série didática Campus/SBC. Coordenador de vários projetos de pesquisa na UFSC e interinstitucionais, com intensa atividade de consultoria em Engenharia de Software. Não é pesquisador do CNPq. Entre outras disciplinas, leciona Metodologia da Pesquisa em Ciência da Computação, no Programa de Pós-Graduação em Ciência da Computação, trimestralmente, desde 2003, de onde partiu a inspiração para escrever este livro.

Prefácio do Autor

Este livro se destina principalmente aos alunos de cursos de ciência da computação, seja na graduação, especialização, mestrado ou doutorado, que vão realizar um trabalho científico escrito na forma de uma monografia, dissertação ou tese.

Aqueles que, como eu, se sentiram perdidos durante seu curso em relação ao que é o trabalho de pesquisa em ciência da computação poderão encontrar algumas informações valiosas neste livro e, portanto, diminuir um pouco a tensão da tese, ou seja, a sensação de angústia que quase invariavelmente atinge os alunos que precisam escrever um trabalho científico.

O aluno de computação que tenha lido livros de metodologia científica de outras áreas já deve ter percebido que nem sempre os autores falam a linguagem que nós entendemos. Nem sempre os métodos de pesquisa de outras áreas se aplicam à computação, devido às características especiais dessa ciência, que permeia praticamente todas as outras atividades humanas.

Este livro, então, procura apresentar conceitos de ciência que sejam compreensíveis ao aluno de computação. Em especial são apresentadas histórias informativas e dicas baseadas na experiência do autor em cerca de 20 anos de orientação e elaboração de trabalhos científicos.

Fica ainda um agradecimento especial ao meu colega Pedro Alberto Barbetta pelas sugestões dadas à Seção 3.4.

Praia dos Ingleses, Florianópolis, 6 de fevereiro de 2009
(um belo dia de sol).

Capítulo 1

Introdução

Era uma vez um aluno de mestrado que queria fazer uma monografia.¹ Ele pensou um pouco sobre o assunto, olhou ao redor e resolveu que havia um problema relevante em sua cidade que ele poderia resolver durante seu mestrado.

O problema era o seguinte: havia um rio cortando a cidade ao meio e não havia forma segura de atravessá-lo.

Disposto a resolver o problema, o aluno conseguiu convencer seu orientador de que teriam material para uma monografia, e começou a trabalhar. Primeiramente, estudou tudo o que podia sobre rios. À medida que estudava, foi escrevendo um capítulo de revisão bibliográfica. Escreveu sobre água, citou a criação dos oceanos de acordo com o Gênesis, escreveu sobre a molécula de água e seus componentes, o hidrogênio e o oxigênio, sobre as diferentes maneiras como os rios desembocam no oceano (incluindo um estudo detalhado sobre os mais importantes deltas do mundo e sua história) e finalmente concluiu com um pequeno tratado sobre a maneira como a gravidade atrai as moléculas de água para o centro da terra, produzindo assim a correnteza dos rios.

¹ Usualmente no Brasil diz-se “dissertação” de mestrado e “tese” de doutorado. Em outros países usam-se os termos indistintamente ou até com significados diferentes. Neste livro o trabalho individual para obtenção de um grau, seja graduação, especialização, mestrado ou doutorado, será denominado “monografia”, para simplificar o texto.

Encerrada essa parte da pesquisa, o aluno deparou-se com o problema em si, que era a inexistência de um meio para cruzar o rio. Pensando um pouco sobre o assunto, ele se lembrou de um instrumento sobre o qual já tinha ouvido falar e que servia para levar objetos de um ponto *A* para um ponto *B*.

Esse instrumento era a catapulta.

Escolhida a ferramenta de trabalho, o aluno passou a planejar os experimentos. Inicialmente transportou 100 indivíduos de um lado ao outro do rio usando a catapulta. Desses indivíduos, 95 não sobreviveram ao experimento. O aluno concluiu que a eficácia do instrumento era, portanto, de apenas 5%, e que haveria grandes possibilidades de melhoria. Portanto, o tema era promissor.

Como segundo experimento, o aluno entregou um paraquedas a cada uma de suas cobaias, e fez o teste com mais 100 indivíduos. Observou então o seguinte: cerca de 20% dos indivíduos se assustavam durante a travessia e abriam o paraquedas antes da hora, caindo dentro do rio e sendo arrastados pela correnteza; cerca de 30% dos indivíduos se assustavam durante a travessia e esqueciam-se de abrir o paraquedas, sofrendo as consequências da queda na outra margem. Habilmente, o aluno concluiu que houve uma melhora no experimento, pois o índice de sucesso passou de 5% para aproximadamente 50%, o que é uma melhora substancial.

Não satisfeito, o aluno resolveu testar outra abordagem para melhorar o sistema. Eliminou o uso do paraquedas, que causava pelo menos 20% de perda sobre o rio, e instalou um colchão de ar na margem oposta. Realizou um novo teste com mais indivíduos e verificou que 95% das vezes os indivíduos sobreviviam à travessia, sendo que apenas 5% dos casos aterrissaram fora do colchão de ar.

Nesse momento, já satisfeito com os resultados, o aluno encerrou os experimentos (até porque estava ficando difícil encontrar voluntários). Como trabalho futuro ele já havia pensado em propor um algoritmo de cálculo da velocidade da catapulta baseado no peso do passageiro e no seu índice de pânico, com vistas a diminuir ainda

mais a taxa de erro observada. Ele não ia programar esse algoritmo porque não teria tempo. Então, deixou isso como trabalho futuro para outro fazer. Escreveu, portanto, o capítulo do desenvolvimento e entregou o texto para o orientador, sendo que só faltava escrever as conclusões e o resumo.

O aluno foi reprovado.

Por incrível que pareça, exageros à parte, a história contada corresponde à história de muitas dissertações de mestrado em Ciência da Computação.

Ao longo deste livro se tentará mostrar *por que* o aluno foi reprovado.

Apesar de seu trabalho ter ares de ciência, ele pecou em vários aspectos no que concerne ao seu comportamento e à metodologia científica. Apenas para citar alguns:

- a) Desde o momento da definição do tema até a conclusão dos experimentos, ele não voltou a entrar em contato com seu orientador, que poderia tê-lo redirecionado.
- b) Não realizou uma revisão bibliográfica adequada, pois estudou muita coisa sobre rios, mas não sobre as formas que já existiam para cruzá-los, como, por exemplo, pontes, barcos, teleféricos etc. Por ter feito uma revisão bibliográfica inadequada, ele concluiu, erroneamente, que era a primeira pessoa no mundo a tentar resolver esse tipo de problema.
- c) Ele escolheu uma ferramenta *a priori* e começou a trabalhar com ela sem uma justificativa adequada para ter eliminado outras ferramentas candidatas.
- d) Seus resultados consistem na comparação do seu trabalho com o seu próprio trabalho, ou seja, não há comparação com trabalhos correlatos de outros autores.
- e) O aluno escolheu como problema-alvo algo que ele observou apenas na sua cidade. Problemas locais nem sempre são problemas para todo mundo. Soluções locais nem sempre podem ser generalizadas.

Este livro tem como objetivo apresentar aos estudantes de Ciência da Computação e áreas correlatas teoria e técnicas para a execução de bons trabalhos científicos, no nível de exigência adequado aos cursos de graduação, especialização, mestrado e doutorado. Casos semelhantes ao citado já aconteceram muitas vezes, mas podem ser evitados com alguma orientação.

A propósito, nenhuma pessoa ou animal foram feridos pelo autor deste livro enquanto ele escrevia a história da catapulta. Pelo menos não que ele tivesse conhecimento, já que a ciência o leva a não crer em verdades absolutas (ele poderia ter pisado em uma formiga sob a mesa sem perceber).

Capítulo 2

Estilos de Pesquisa Correntes em Computação

Com alguma frequência se ouve dizer que a Computação ou Informática é uma área *nova* no campo das ciências, e que está em *franco desenvolvimento*. Mas isso não justifica que o método científico específico da área de Computação tenha de ser vago e que tantas monografias sejam escritas sem um embasamento metodológico adequado.

Essa discrepância de estilos de pesquisa e baixa conformação à metodologia científica acontecem não só pelo fato de a área ser nova, mas também pelo fato de que a Computação permeia praticamente todas as atividades humanas, e, portanto, se inter-relaciona com muitas outras disciplinas.

A própria observação do surgimento dos primeiros cursos de Ciência da Computação no Brasil é evidência dessa variedade de abordagens e inter-relacionamentos. Alguns cursos surgiram nas faculdades de Engenharia. Em outras universidades, os cursos de Computação foram oriundos das faculdades de Matemática ou de Física. Em alguns casos ainda, cursos de Computação surgiram a partir de departamentos de processamento de dados cuja finalidade era a prestação de serviços e não o ensino.

A variedade de cursos, e mesmo de denominações, causou grande confusão no cenário nacional até aproximadamente o ano 2000. Até essa data, cursos na área poderiam ser denominados “Bacharelado em Ciência da Computação”, “Análise de Sistemas”, “Bacharelado em Informática”, “Engenharia de Computação”, “Engenharia de Infor-

mática”, e assim por diante. Usualmente, não havia nenhum tipo de correspondência entre a denominação do curso e o tipo de formação que era oferecido. Após o ano 2000, os cursos da área foram definidos pela Comissão de Especialistas de Ensino de Computação e Informática, do Ministério da Educação, em apenas quatro denominações:

- a) Bacharelado em Ciência da Computação.
- b) Bacharelado em Sistemas de Informação.
- c) Licenciatura em Informática.
- d) Engenharia de Computação.

Tal classificação, porém, ainda é imprecisa, sendo que em muitos casos um licenciado em Informática poderá perfeitamente exercer as atribuições de um bacharel em Sistemas de Informação. Por outro lado, um engenheiro de Computação poderá dar aulas de Informática, e assim por diante.

Se as ideias já são difusas na nomenclatura dos cursos, quanto mais na pesquisa realizada pelos profissionais da área.

Este capítulo apresenta uma possível classificação para os tipos de pesquisa realizados em Ciência da Computação e áreas correlatas, considerando o grau de amadurecimento da pesquisa na subárea específica, bem como seu inter-relacionamento com outras ciências. Essa classificação é baseada em uma discussão ocorrida na lista sbc-l@sbc.org.br há cerca de 10 anos.

2.1. Estilo “Apresentação de um Produto”

Dentre as áreas emergentes dentro da Computação, ou seja, aquelas que, mesmo para a Computação, são consideradas muito novas, é aceitável uma pesquisa em que simplesmente se procura apresentar algo novo. Nessas áreas, a pesquisa é eminentemente exploratória, sendo difícil comparar um trabalho com trabalhos anteriores, pois estes podem não existir.

Sendo assim, as pesquisas nessas áreas apresentariam resultados da forma “Fiz algo novo. Eis meu produto”. É muito pouco provável que áreas mais maduras reconheçam pesquisas apresentadas assim.

Exemplificando: um artigo do tipo “um novo método para análise de sistemas” dificilmente seria aceito em um evento de Engenharia de Software, a não ser que o autor apresentasse claramente os problemas com os *velhos* métodos e como o seu *novo* método os resolve. Outro exemplo consiste em artigos ou trabalhos que apresentam uma ferramenta ou protótipo. Normalmente existe pouca preocupação em apresentar comparativos com outras ferramentas. Alega-se que não existem ferramentas iguais, o que normalmente é até verdade. Mas se uma ferramenta ou protótipo construído justificasse a concessão do título de mestre, por exemplo, as universidades distribuiriam diplomas de mestre a todos os analistas e programadores que criam ferramentas diariamente nas suas empresas, não é mesmo? Evidentemente existe algo mais.

De qualquer maneira, apresentações desse tipo normalmente são ingênuas e devem ser evitadas. Mesmo que se esteja trabalhando em uma área nova do conhecimento, é interessante que a pesquisa demonstre que se está resolvendo um problema relevante. Se o problema é relevante, então provavelmente já se tentou resolvê-lo, e a partir daí, já é possível traçar um comparativo. O aluno da catapulta, mencionada anteriormente, apresentou uma solução para o problema sem ter tomado conhecimento de outras soluções que já existiam, e por isso falhou em sua pesquisa.

Um tipo de artigo que se encaixa muito bem nessa categoria é aquele em que o aluno desenvolve um sistema e escreve um artigo apresentando-o. Não há comparativos, não se apresenta nenhum conhecimento novo, a não ser o sistema em si, e, portanto, esse tipo de artigo tem poucas chances de ser aceito em um veículo de publicação relevante. Muitas vezes, tais artigos são vistos mais como uma *propaganda* do grupo que desenvolveu o sistema do que como uma contribuição científica. Em outras palavras, artigos do tipo “manual da ferramenta” devem ser evitados.

Esse tipo de publicação poderá ter seu espaço em sessões especiais de apresentação de ferramentas, ou em eventos cujo tema seja a aplicação da informática a alguma outra área, como, por exemplo,

a Medicina, a Educação e outras. Mesmo assim, essas áreas têm, cada vez mais, exigido que os artigos apresentados sejam mais do que uma mera descrição de um sistema, que tragam conhecimento novo para a área e, principalmente, comparem o trabalho apresentado com trabalhos anteriores.

O desenvolvimento de um sistema e sua apresentação podem ser considerados um trabalho relevante em cursos de graduação ou especialização, desde que fique evidente que o aluno aplicou técnicas no sistema ou no processo de desenvolvimento do sistema aprendidas durante o curso. Dificilmente esse tipo de trabalho seria aceito no mestrado e doutorado.

2.2. Estilo “Apresentação de algo Diferente”

Um segundo tipo de pesquisa, um pouco mais amadurecido, consiste na apresentação de uma forma diferente de resolver um problema. Esse tipo de pesquisa também é característico de áreas emergentes, e os trabalhos normalmente são apresentados como uma simples comparação entre técnicas, em que não se exige necessariamente rigor científico na apresentação dos resultados. As comparações normalmente são muito mais qualitativas do que quantitativas.

Um exemplo típico desse tipo de pesquisa seria um trabalho em Engenharia de Software no qual se apresenta uma nova técnica para realizar algo, em que se compara essa técnica com outras técnicas existentes (não necessariamente todas e não necessariamente as melhores, muitas vezes por falta de uma métrica para decidir sobre isso), e em que se apresenta um ou dois estudos de caso para reforçar o argumento.

Os resultados de um artigo deste tipo poderão ser aceitos em algum veículo de publicação, desde que os argumentos utilizados pelo autor sejam convincentes.

Um estudo de caso raramente prova alguma coisa, e a possibilidade de generalizar o resultado é responsabilidade do autor do texto, não do leitor. Se o método funcionou no estudo de caso A ou

no estudo de caso B, isso não quer dizer que funcionará sempre. Não há aqui, portanto, prova com rigor científico, mas uma tentativa de *convencimento* do leitor.

Apesar disso, o estudo de caso pode servir para provar que um método consagrado falha em uma ou outra situação. Esse resultado, sim, poderia ser interessante desde que o motivo da falha fosse claramente identificado e uma solução para o problema fosse proposta e validada.

Esse tipo de pesquisa é típico em áreas novas nas quais não se dispõe de grandes bases de dados para testar teorias empiricamente, ou quando o tempo e os recursos necessários para realizar a pesquisa empiricamente são inviáveis.

Para que esse tipo de pesquisa funcione é necessário que se tenha uma boa hipótese de trabalho, uma boa teoria construída para sustentá-la e uma boa argumentação para fazer com que um eventual leitor se convença da validade da teoria, mesmo sem poder testá-la com métodos estatisticamente aceitos.

Em relação à hipótese, convém mencionar que ela é o coração da monografia. Se a hipótese for mal escolhida, o trabalho pode não alcançar os objetivos. Nesse caso, quem é penalizado? O aluno! Portanto, uma boa hipótese com evidências de efetividade deve ser buscada.

Trabalhos de mestrado e doutorado, em geral, propõem algo: um novo método, uma nova ideia, um novo sistema etc. Porém, “propor” algo é fácil. Difícil é mostrar que a proposta apresenta algum tipo de melhoria em relação a outras propostas semelhantes que existem por aí.

Por exemplo, propor um método de compressão de textos mais eficiente do que os que atualmente estão no mercado é possível e até louvável como objetivo de uma tese. Mas daí o problema é: como criar um método mais eficiente do que os atuais? É necessário ter uma boa hipótese.

Uma hipótese, segundo a Wikipédia (2008), é uma teoria provável, mas ainda não demonstrada, ou uma suposição admissível.

A hipótese norteia o trabalho de pesquisa justamente porque ainda não se sabe se ela é efetivamente verdadeira. Ela será testada ao longo do trabalho. Caso se confirme, o trabalho terá sido um sucesso. Caso não se confirme, será necessário juntar os cacos e tentar outra linha de pesquisa. Por isso é que é necessário ter uma hipótese bem embasada e justificada. O risco é sempre do aluno.

Uma das formas de aumentar a chance de sucesso desse tipo de trabalho é estruturá-lo na forma de uma tabela comparativa. A ideia é que não se vai criar algo simplesmente diferente daquilo que já existe, mas algo que incorpore várias características importantes em um mesmo artefato. Idealmente o trabalho começa com uma boa pesquisa bibliográfica para se descobrir quais são as formas correntes usuais para se resolver o problema em questão. Em seguida, analisam-se diferentes propriedades de cada uma das abordagens, construindo uma tabela como a da Figura 2.1.

	Característica 1	Característica 2	Característica 3	Característica 4
Artefato 1	X	X		
Artefato 2	X			X
Artefato 3		X	X	X

Figura 2.1: Exemplo de tabela comparativa de artefatos e características.

Uma vez identificados os artefatos usados para resolver o problema em questão e as principais características desses artefatos, pode-se proceder à criação ou definição de um novo artefato que abranja todas as características, conforme a Figura 2.2.

	Característica 1	Característica 2	Característica 3	Característica 4
Artefato 1	X	X		
Artefato 2	X			X
Artefato 3		X	X	X
Novo Artefato	X	X	X	X

Figura 2.2: Proposta de um novo artefato que tenha todas as características dos anteriores.

Esse novo artefato será diferente dos demais, pois seu conjunto de características não é possuído por nenhum dos outros artefatos isoladamente. O novo artefato será útil na medida em que as características forem efetivamente relevantes. E a tabela comparativa será uma boa ferramenta para a pesquisa, caso as características possam ser efetiva e independentemente verificadas.

2.3. Estilo “Apresentação de algo Presumivelmente Melhor”

Áreas um pouco mais amadurecidas do que as anteriores exigem que qualquer nova abordagem apresentada seja comparada quantitativamente com outras da literatura. Na falta de bancos dados (*benchmark*) internacionalmente aceitos ou acessíveis, o próprio autor do artigo acaba criando e realizando os testes que demonstram que a sua abordagem é melhor do que outras.

Um problema com esse tipo de pesquisa é que o autor terá de testar a sua abordagem e também as outras, que constam da literatura, resultando em excesso de trabalho, além de, possivelmente, introduzir o risco de erro, visto que não há garantias de que as abordagens apresentadas na literatura estejam testadas nas melhores condições pelo autor do trabalho. Sendo assim, tais comparações muitas vezes são temerárias. Para que uma pesquisa desse tipo seja bem aceita, é necessário que o autor deixe bem claro a forma como aplicou cada uma das técnicas e que isolou todos os fatores que poderiam possivelmente afetar os resultados.

De qualquer maneira, uma abordagem para um determinado problema que se revela melhor do que outras abordagens requer alguns cuidados. Em primeiro lugar, o pesquisador deve se certificar de que está comparando a nova abordagem com alguma outra que seja do estado da arte. Em computação, muitas vezes é inadmissível apresentar um método e compará-lo com outro de uma referência bibliográfica de 15 anos atrás. Mesmo que o novo método seja melhor que o antigo, o artigo terá pouca credibilidade, a não ser que o autor deixe bem claro que nos últimos 15 anos não houve nenhum

avanço nessa área. Um artigo, porém, que apresente melhorias em relação a um processo publicado recentemente, digamos, no máximo um ou dois anos, terá mais credibilidade.

Não é necessário, porém, que o autor de algum método novo demonstre que o seu método é melhor que outro método do estado da arte para toda e qualquer situação. É possível, muitas vezes, apresentar métodos ou abordagens que funcionam melhor em determinadas situações. Nesse caso, o artigo deve deixar bem claro quais são as situações nas quais a nova abordagem funciona melhor e o porquê disso. Experimentos deverão ser feitos para demonstrar tal melhoria.

Aqui entra em foco um aspecto muito importante na pesquisa que leva em consideração o uso de dados comparativos: a *métrica*. Afirmações do tipo “O sistema *x* é mais fácil de usar” não terão fundamento a não ser que se defina claramente o que significa “ser fácil de usar”. Um exemplo de definição nesse caso poderia ser a quantidade de cliques de mouse que o usuário tem de usar para executar uma tarefa dada em um ou outro sistema. Poderia ser questionado se a métrica é boa e eficiente, mas dentro da definição dada é possível confirmar qual sistema é mais fácil de usar.

2.4. Estilo “Apresentação de algo Reconhecidamente Melhor”

O nível mais maduro da pesquisa desta linha, em que a apresentação de dados empíricos é relevante para a aceitação dos resultados, é aquele no qual um trabalho é desenvolvido e seus resultados são apresentados em função de testes padronizados e internacionalmente aceitos. Nesse caso, o autor do trabalho não precisa testar outras abordagens, pois seus resultados já estão publicados. O autor deverá buscar os dados de entrada para testar a sua abordagem em um banco de dados conhecido e apresentar os resultados usando uma métrica aceita pela comunidade. Dessa forma, os experimentos poderão ser reproduzidos por equipes independentes. Se for demonstrado que a nova abordagem é superior às abordagens anteriores, esta passará a ser considerada como estado da arte.

Pesquisas que apresentam resultados desse tipo são típicas de boas teses de doutorado. Supõe-se que após a publicação dos resultados ninguém mais possa ignorar essa nova abordagem em função das vantagens que ela oferece em relação às anteriores. Isso é o que se entende por “avançar o estado da arte”.

Por incrível que pareça, essa é a pesquisa mais fácil de executar, desde que o autor tenha uma boa hipótese de trabalho. Por que isso? Porque os testes-padrão já estão definidos e os dados já estão disponíveis. Basta implementar a abordagem e realizar os testes. O grande problema e dificuldade inerente, então, consistem em encontrar uma boa hipótese de trabalho, que faça sentido e que seja promissora (isso, infelizmente, em geral não é trivial).

Portanto, esse tipo de pesquisa exigirá, por parte do autor, amplo estudo sobre o estado da arte em uma determinada área, e muita reflexão sobre a forma como as técnicas são desenvolvidas para resolver os problemas dessa área. Problemas em aberto serão excelentes focos de atenção para pesquisa.

Além disso, poderá ser de muita valia se o autor possuir conhecimentos em outras áreas, que muitas vezes nem estejam relacionadas ao problema em questão. Algumas vezes técnicas de áreas distintas aplicadas a um problema produzem resultados muito interessantes. Porém, deve-se lembrar que apenas aplicar uma técnica diferente para um problema remete o trabalho para “apresentar algo diferente”. Quando se opta por utilizar uma técnica alienígena em um problema conhecido, é necessário que se tenha bons motivos para acreditar que a técnica possa produzir resultados melhores do que as técnicas correntes.

2.5. Estilo: “Apresentação de uma Prova”

Os diferentes tipos de pesquisa apresentados anteriormente se enquadram nas subáreas da computação em que normalmente os resultados são apresentados a partir de evidências empíricas ou pelo menos de argumentações ou estudos de caso que sugerem provas.

Outro tipo de pesquisa exige provas matemáticas, de acordo com as regras da lógica. A área de métodos formais ou compiladores, por exemplo, dificilmente aceitará trabalhos que não apresentem demonstrações claras de correção ou eficiência.

Deve ser construída uma teoria, afirmando claramente quais são os conceitos utilizados e mostrando que a aplicação desses conceitos leva, logicamente, a determinados resultados. Esses resultados podem ser a demonstração de que um determinado algoritmo é o melhor algoritmo possível para resolver um determinado tipo de problema, ou que um algoritmo para resolver um determinado tipo de problema não existe, ou ainda que a complexidade de qualquer algoritmo que resolve um determinado tipo de problema não pode ser menor do que um determinado polinômio.

2.6. Discussão

A partir das observações feitas, pode-se verificar que diferentes subáreas da Computação podem ser caracterizadas por diversos estilos de pesquisa. É possível classificar esses estilos, então, em três tipos básicos:

- a) Pesquisas formais, em que é exigida a elaboração de uma teoria e uma prova formal de que essa teoria é correta. A lógica formal será a grande ferramenta de trabalho do pesquisador que optar por essa linha.
- b) Pesquisas empíricas, em que uma nova abordagem apresentada é comparada com outras através de testes aceitos pela comunidade. Os métodos estatísticos serão a grande ferramenta de trabalho do pesquisador que optar por essa linha.
- c) Pesquisas exploratórias, em que não se consegue provar uma teoria nem apresentar resultados estatisticamente aceitos. Mas entram aqui os estudos de caso, as análises qualitativas e as pesquisas exploratórias em áreas emergentes. A argumentação e o convencimento são as principais ferramentas do pesquisador.

Embora a pesquisa formal aparentemente seja mais difícil de realizar, seus resultados, quando obtidos, são mais difíceis de refutar.

A pesquisa empírica, mesmo baseada em métodos estatísticos, poderá ser refutável se não estiver também embasada em uma boa teoria, isso porque a estatística não explica causas. É bem conhecida a anedota do pesquisador que mandava uma aranha pular e em seguida lhe arrancava uma das pernas. Após arrancar sete pernas, a aranha ainda pulava com a perna que restava. Após arrancar a última perna, o cientista percebeu que a aranha não atendia mais à ordem de pular. A conclusão do pesquisador foi de que a aranha sem pernas fica surda, pois não escutava mais a ordem para pular.

Em último lugar, a pesquisa exploratória parece ser mais fácil de realizar, porque não é necessário utilizar os métodos da lógica formal e nem realizar experimentos exaustivos. Porém, em termos de pesquisa, é a abordagem mais arriscada, pois a aceitação dos argumentos não é universal, e artigos que não se fundamentam em uma boa teoria e/ou em um bom conjunto de testes têm menor chance de serem publicados em bons veículos do que os demais tipos. A apresentação de estudos de caso e exemplos, no caso de pesquisas exploratórias, poderá ajudar o pesquisador a convencer o leitor do seu ponto de vista, mas não constituem provas.

Capítulo 3

Preparação de um Trabalho de Pesquisa

A preparação de um trabalho de pesquisa é uma etapa que deve ser realizada antes que se comece a escrever sobre a pesquisa.

Parece estranho, mas muitas vezes esse é um ponto que deve ser ressaltado. Alunos, ansiosos por escrever a monografia, começam a escrever páginas e páginas sem terem feito nenhuma pesquisa. Uma ressalva: é necessário diferenciar o termo “pesquisa”, que pode ser entendido como revisão bibliográfica, da sua significação como pesquisa científica, ou seja, produção de conhecimento novo. A revisão bibliográfica, diga-se de passagem, não produz conhecimento novo, mas apenas supre ao estudante conhecimentos que lhe faltavam.

Portanto, não se recomenda que o aluno comece a escrever sua monografia ou trabalho de conclusão de curso sem ter realizado alguma pesquisa que tenha produzido conhecimento novo. Explicando melhor: não há necessidade de sair escrevendo o capítulo de revisão bibliográfica antes de saber o que efetivamente será feito em termos de produção de conhecimento.

A revisão bibliográfica de um trabalho de pesquisa em Computação, em geral, não deve ser um tratado sobre a área de pesquisa. Muitas vezes, um aluno que começa a escrever o capítulo da revisão antes de decidir o objetivo da pesquisa acabará escrevendo demais e desnecessariamente. Esse capítulo será cansativo para o leitor, e muitas vezes ele não compreenderá por que determinados assuntos são ali colocados se não são abordados na monografia em si, que vai iniciar mais adiante.

O segredo de um trabalho de pesquisa de sucesso consiste em ter um bom objetivo. Uma vez definido o objetivo do trabalho, tudo o mais gravita em redor dele. A justificativa vai dizer por que vale a pena buscar esse objetivo; o método informa como o objetivo pode ser alcançado; os resultados esperados mostram o que muda no mundo após o objetivo ser atingido; o capítulo de revisão bibliográfica vai apresentar os conceitos necessários para a compreensão do objetivo e os trabalhos relacionados ao objetivo.

3.1. Escolhendo o Objetivo de Pesquisa

A escolha de um objetivo de pesquisa é frequentemente a tarefa mais difícil em um trabalho de mestrado ou doutorado.

Muitas vezes, o objetivo é confundido com o tema da pesquisa. O *tema* da pesquisa pode ser a influência da batata inglesa na língua portuguesa. Mas nesse caso qual é o objetivo? Ou seja, o que se quer provar?

O objetivo normalmente comporta uma *hipótese de trabalho*. Um bom objetivo de pesquisa normalmente terá a forma “demonstrar que a hipótese x é verdadeira”.

Nem todo objetivo pode ser considerado um bom objetivo de pesquisa. Por exemplo, algo do tipo “o objetivo deste trabalho é aumentar os meus conhecimentos na área de estudo” pode até ser muito sincero, mas não convence ninguém de que algum conhecimento novo para a humanidade será produzido. Portanto, isso deve ser evitado.

Outro objetivo algumas vezes encontrado é a forma “propor...”. Alguma coisa é proposta, normalmente um método, uma abordagem, uma técnica, um algoritmo, uma comparação, ou qualquer outra coisa. A questão é: se o autor fizer a proposta, então o objetivo estará atingido? Se o aluno se propõe a propor e propôs, então está proposto! O que for proposto não é necessariamente melhor ou diferente daquilo que já existia antes. Então, o estágio da pesquisa neste caso ainda é dos mais ingênuos.

É necessário que o objetivo diga que aquilo que está sendo proposto é melhor do que alguma outra coisa ou que resolve algum problema que antes não podia ser resolvido.

Segundo Chinneck (1988), a descrição de um problema de pesquisa tem três partes:

- a) Um enunciado preciso da questão ou problema de que trata a monografia.
- b) Uma explicação por referência direta à bibliografia de que tal questão de pesquisa ainda não foi tratada.
- c) Uma discussão sobre por que é importante tratar essa questão de pesquisa.

O item (b) falhará se o aluno não conseguir deixar claro que a questão de pesquisa nunca foi tratada. Uma boa revisão bibliográfica é necessária para apresentar tal justificativa com suficiente autoridade. Devem ser evitadas afirmações do tipo “não encontrei nada parecido”. O aluno deve sempre mostrar o que encontrou nas fontes relevantes que examinou e comparar aquilo que foi encontrado com aquilo que ele pretende fazer. Se ele disser que não encontrou nada, a banca examinadora provavelmente pensará que ou ele não pesquisou direito ou está tratando um problema de pouco interesse. Em alguns casos pode ocorrer que efetivamente nada de muito semelhante seja encontrado, mas em todo o caso, sempre existe algum problema que possa ser considerado o mais próximo possível. Haverá situações em que serão encontradas abordagens quase idênticas, variando em poucos detalhes; e, em outros casos, a abordagem mais próxima será tão distinta que será necessária uma boa dose de explicação para que se possa entender por que ela é relevante. Leonardo da Vinci não conhecia a tecnologia utilizada pelos modernos aviões, então baseou seus estudos sobre máquinas voadoras no modelo mais próximo que estava disponível na sua época, a estrutura de voo dos pássaros.

Por outro lado, é importante que um trabalho de pesquisa se relacione principalmente com conhecimentos que sejam inerentes ao estado da arte, ou seja, que sejam recentes. Não faria muito sen-

tido hoje, com todos os conhecimentos de engenharia aeronáutica disponíveis, ignorar essas informações e tentar construir máquinas voadoras baseadas nos modelos de Leonardo.

Apesar disso, os conhecimentos antigos não devem ser de todo ignorados. Muitas vezes o conhecimento antigo aliado ao estado da arte pode produzir resultados muito interessantes. No caso da indústria aeronáutica, por exemplo, tenta-se hoje produzir aviões flexíveis que, de certa forma, imitam as estruturas de voo dos pássaros. Mas isso é feito com base em conhecimentos atuais, não se está reinventando a roda, mas aperfeiçoando conceitos já existentes.

3.1.1. O Caminho para a Escolha de um Objetivo de Pesquisa

Para que alguém seja capaz de pensar em um objetivo de pesquisa relevante, essa pessoa deve conhecer a área de pesquisa na qual está trabalhando. Portanto, o caminho lógico consiste de três passos:

- a) Escolher um *tema* de pesquisa, ou seja, uma área de conhecimento na qual se vai trabalhar.
- b) Realizar a *revisão bibliográfica*. A não ser que o autor já seja especialista na área escolhida, ele vai precisar ler muitos trabalhos já publicados nessa área para saber o que está sendo feito (estado da arte) e o que ainda precisa ser feito (problemas em aberto).
- c) Definir o *objetivo* de pesquisa. Uma vez feita a revisão bibliográfica, o objetivo de pesquisa possivelmente será fortemente relacionado com um dos problemas em aberto verificados no passo anterior.

Em poucos casos, a revisão bibliográfica será feita depois da definição do objetivo de pesquisa. Não é razoável conceber que alguém decida por um caminho de pesquisa sem conhecer exatamente a área onde vai atuar. Essa atitude possivelmente levará a objetivos mal definidos e que precisarão ser revistos quando o pesquisador perceber que está reinventando a roda.

Mas é possível ainda que os passos (b) e (c) devam ser repetidos algumas vezes de forma a refinar o objetivo (Figura 3.1). De

fato, ao realizar a revisão bibliográfica sobre um determinado tema, o pesquisador terá ideias sobre aspectos do tema que ainda não foram explorados, e esses aspectos darão origem a um objetivo de pesquisa. Mas antes de começar a gastar energia tentando atingir esse objetivo, o pesquisador deve ainda refinar sua pesquisa bibliográfica tentando verificar se tal objetivo já não foi perseguido em trabalhos anteriores e que tipo de resultados foi obtido.

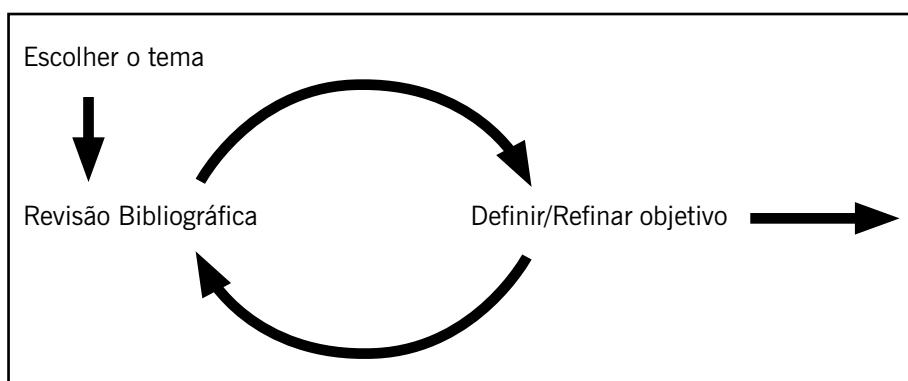


Figura 3.1: O caminho lógico para a definição de um objetivo de pesquisa.

Na história da catapulta, o tema de pesquisa foi o rio que cortava a cidade. Quando foi estabelecido o objetivo de pesquisa “encontrar uma forma segura de atravessar o rio”, o aluno deveria ter verificado na literatura os principais trabalhos já publicados sobre travessia de rios. Mas ele se limitou apenas a informações sobre a composição dos rios e, portanto, perdeu informações importantíssimas para sua pesquisa, ou seja, que já existiam métodos para atravessar rios e que ele deveria primeiro tentar melhorar esses métodos já existentes, ou pelo menos conhecê-los e seus defeitos, antes de enveredar pela pesquisa de algo totalmente novo. No extremo desse raciocínio fica a possibilidade de que alguns anos antes outros pesquisadores já tivessem tentado usar a catapulta para cruzar rios, sem saber um do outro. A cada vez que o estudo é repetido, esforços são despendidos, medidas tomadas e conclusões (possivelmente as mesmas) são obtidas, sem que nenhum avanço ocorra para a ciência.

3.1.2. O Tema

O tema da pesquisa frequentemente depende do interesse do aluno e do orientador. Não se recomenda, em hipótese alguma, uma pesquisa cujo tema não seja compatível com os conhecimentos do orientador. No caso do aluno, recomenda-se que ao passar da graduação ao mestrado e do mestrado ao doutorado procure trabalhar no mesmo tema, embora, buscando objetivos distintos. Por que isso? Porque, se o aluno se mantiver no mesmo tema, o passo de revisão bibliográfica será concluído mais rapidamente ao passar de um estágio para outro em sua formação. Bastará que ele se atualize nos últimos desenvolvimentos da área para poder decidir por um bom objetivo de pesquisa. Se o aluno mudar de tema, então terá de fazer toda a revisão bibliográfica sobre outro tema, o que lhe tomará muito tempo.

Não é impossível que uma pessoa com formação em uma área faça mestrado ou doutorado em uma área totalmente diferente, mas com certeza essa pessoa terá muito mais trabalho e levará muito mais tempo para amadurecer os conceitos da nova área do que uma pessoa que já tem a vivência na área.

O tema pode ser especializado a partir de uma grande área em subáreas cada vez mais específicas. Por exemplo:

1. Ciência da Computação.

1.1. Inteligência Artificial.

1.1.1. Métodos de busca.

1.1.1.1. Busca heurística.

1.1.1.1.1. Algoritmo A*.

Nessa lista, cada item é uma especialização do item anterior, mas cada um é apenas um tema de pesquisa, embora cada vez mais específico.

Indo para outra direção, pode-se combinar um tema de pesquisa com uma área de aplicação. O tema, possivelmente, será mais específico do que geral. Não faz muito sentido, por exemplo, falar de “aplicação da Ciência da Computação no problema da pavimentação

das estradas”. Faria mais sentido um tema de pesquisa como “aplicação de busca heurística no problema do transporte de máquinas para pavimentação de estradas”.

Mesmo assim, em alguns casos encontram-se trabalhos cujo tema é algo do tipo “aplicação de técnicas de Inteligência Artificial para resolver o problema *x*”. O pesquisador deve ter a noção de proporção para saber se o nível de especificidade do tema de pesquisa é adequado ou não. No exemplo anterior, ao se falar de Inteligência Artificial, abre-se um leque enorme de possibilidades (pode-se conferir como a área é extensa no livro de Russel e Norvig, 1995). Portanto, um tema tão extenso quanto este não é adequado. Ao se falar de busca heurística, porém, o leque é reduzido a alguns poucos algoritmos conhecidos, e o tema fica mais viável. Quanto mais amplo o tema, maior a quantidade de livros e artigos que terão de ser lidos. Portanto, recomenda-se buscar temas cada vez mais específicos antes de se propor um objetivo de pesquisa. Quando se escolhe um tema de pesquisa que tem aplicação em outra área, deve-se tomar cuidado. Quando se está realizando um mestrado ou doutorado em Ciência da Computação, deve-se observar que a principal contribuição do trabalho deve ser para a área da Computação. Ou seja, o problema a ser resolvido deve estar ligado à não-existência ou inadequação das ferramentas de computação existentes, e não aos aspectos ou técnicas da área de aplicação (pelo menos não predominantemente). O aluno, nesse caso, deve evitar a tentação de *contribuir* para uma área que ele não conhece ou que conhece muito pouco. Por exemplo, uma monografia sobre Informática aplicada à Medicina deve apresentar contribuições em relação às ferramentas de Computação aplicada à Medicina e às melhorias que podem ser obtidas. Dificilmente o aluno de Computação irá propor e defender uma nova técnica cirúrgica. Não quer dizer que seja impossível que isso aconteça, nem quer dizer que o aluno não seja execrado na banca por algum médico que esteja ali presente, mas é um risco a ser evitado.

3.1.3. O Problema

Uma monografia deve apresentar uma solução para um problema. Inicialmente, portanto, um problema deve ser identificado. Seria errado iniciar a monografia simplesmente resolvendo criar um novo método para isso ou aquilo.

No caso da catapulta, o aluno propôs e testou um novo método para cruzar rios. Aqui vem uma questão interessante. Segundo o aluno, o problema consistia em cruzar o rio. Mas esse problema identificado já não é mais um problema sem solução, porque existem diferentes abordagens para cruzar um rio: ponte, balsa, teleférico etc. Então, caso o aluno quisesse insistir neste tema, teria de indicar quais problemas as soluções existentes apresentam. Ou seja, quais os problemas encontrados quando se tenta cruzar um rio com uma ponte ou com uma balsa? Ele poderia descobrir, por exemplo, que pontes são muito caras e balsas são muito lentas. Se todas as soluções existentes apresentarem algum tipo de problema, então é possível que se esteja abrindo caminho para uma nova abordagem. Caso contrário, as pessoas continuarão a cruzar rios com os meios usuais e não com a *nova* abordagem.

Algumas propostas de pesquisa são apresentadas inicialmente sem ter um problema claramente identificado. Por exemplo: “este trabalho propõe usar a metáfora de formigueiro para modelar pacotes em uma rede”. Esse tema até pode vir a ser um trabalho interessante, mas qual é o problema que essa modelagem vai resolver? O que há de errado com outras formas de modelagem, sejam elas quais forem, que essa metáfora do formigueiro vai possivelmente resolver?

Segundo Griffiths (2008), se o autor não consegue estabelecer claramente qual é o problema tratado em sua monografia, será muito difícil para outras pessoas especularem sobre os possíveis usos dela. Também será difícil avaliar se ela obteve sucesso.

3.1.4. Perspectiva Profissional

O tema de pesquisa a ser escolhido, em qualquer nível de formação, além de ser do agrado do aluno, deve estar relacionado com a sua perspectiva de desenvolvimento profissional.

Não faz sentido gastar vários anos investindo em uma pesquisa na área X, para depois trabalhar o resto da vida na área Y. O ideal seria que cada pesquisa gerasse algum tipo de produto ou uma melhoria a um produto existente que pudesse, após o final do curso, gerar algum tipo de perspectiva de trabalho e de renda para o pesquisador.

Isso é mais difícil quando se trata de pesquisa básica, mas mesmo assim esse tipo de pesquisa pode gerar patentes que em algum momento poderão se tornar produtos e gerar riquezas para o país. Não é admissível que todo um esforço despendido durante o mestrado ou mesmo durante um doutorado acabe sendo depois simplesmente guardado em uma gaveta.

3.2. A Revisão Bibliográfica

A revisão bibliográfica, conforme já comentado, não produz conhecimento novo, mas apenas supre as deficiências de conhecimento que o pesquisador tem em uma determinada área. Portanto, ela deve ser muito bem planejada e conduzida.

Supondo que o pesquisador seja praticamente iniciante em uma determinada área, ele deveria iniciar a pesquisa lendo algum tipo de *survey* sobre o assunto. Não é difícil encontrar artigos ou mesmo livros que abordem toda uma área de conhecimento na forma de *survey*. Os livros, aliás, servem exatamente a esse propósito.

Pode-se iniciar a pesquisa com uma leitura de trabalhos mais abrangentes que deem uma visão do todo para depois ir se aprofundando cada vez mais em temas cada vez mais específicos.

Quando se faz uma pesquisa em que alguma técnica de computação é aplicada a alguma outra área do conhecimento, é necessário que se faça a revisão bibliográfica sobre a técnica em si, sobre a área de aplicação e, mais do que tudo, sobre as aplicações que já foram tentadas com essa técnica ou com técnicas semelhantes na mesma área ou em áreas equivalentes. Exemplificando, um aluno pretende desenvolver um sistema multiagentes para auxiliar controladores de voo. Esse aluno deve conhecer profundamente os sistemas multiagentes e deverá conhecer também os problemas que os controla-

dores de voo enfrentam para exercer sua profissão. Porém, ele não deve pensar, como algumas vezes acontece, que essa é a primeira vez que alguém vai tentar desenvolver um sistema multiagentes para esse tipo de aplicação. O aluno da catapulta, mencionado anteriormente, estudou rios e estudou catapultas, mas não procurou saber se alguém já tinha tentado atravessar um rio usando uma catapulta. Se essa pesquisa existisse e ele tivesse acesso a ela, teria visto que os resultados não eram animadores e talvez tivesse escolhido outro tema de pesquisa antes de ter dedicado a maior parte do seu tempo de mestrado a algo infrutífero.

Uma monografia sobre pesquisa aplicada que apresenta uma boa revisão bibliográfica da ferramenta de computação e da área de aplicação, mas que não menciona nenhuma tentativa anterior de aplicação dessa ferramenta na área sofre da “Síndrome da Intersecção Esquecida”. Uma monografia com esse problema estará possivelmente “reinventando a roda”.

A Síndrome da Intersecção Esquecida em geral é justificada pelos alunos com frases do tipo “mas não encontrei nada parecido com o que estou fazendo”. Esse raciocínio negativo deve ser evitado. Nunca se deve dizer que não se achou nada semelhante. Algo sempre deve ser apresentado como referência. Essa referência poderá ser mais semelhante ou menos semelhante à abordagem usada de um ponto de vista relativo. Mas sempre a abordagem mais semelhante de todas (por menos semelhante que seja) deve ser buscada.

Pode ser interessante pensar assim: “Ninguém fez algo parecido com o que estou fazendo, *mas* muitas coisas já foram feitas pelos seres humanos ao longo da sua história. Então, eu poderia classificar as coisas que já foram feitas em termos de grau de semelhança com aquilo que estou fazendo. As coisas mais parecidas com o meu trabalho serão minha referência, mesmo que a semelhança seja pequena”.

Assim, evita-se o *fundamento vazio*, ou seja, dizer que seu trabalho é original porque ninguém nunca fez nada parecido. Não se deve fundamentar todo um trabalho de pesquisa em uma negação.

Deve-se mostrar o que outros fizeram, e depois mostrar que o trabalho feito é diferente ou melhor do que essas referências.

3.2.1. Fichas de Leitura

Durante todo o processo de leitura, é fundamental que sejam feitas anotações. Conceitos-chave e ideias novas devem ser anotados sempre que forem detectados na leitura. É necessário que se saiba de onde tais ideias e conceitos saíram. Em geral, inicia-se uma ficha de leitura, seja em papel, seja no computador, escrevendo a referência bibliográfica da obra sendo consultada. Em seguida são feitas as anotações relevantes.

Essas fichas serão extremamente importantes no futuro para saber de onde as ideias saíram. Depois de ler algumas dezenas de artigos sobre um determinado assunto, será difícil lembrar-se de onde saíram determinadas ideias.

Porém, deve-se ter em mente que o conjunto das fichas de leitura não é o mesmo que o capítulo de revisão bibliográfica. As fichas são apenas um registro, com memória, de leituras feitas, que é organizado por fonte bibliográfica. Esse trabalho normalmente será desenvolvido antes da definição do objetivo de pesquisa, pois, como foi comentado, consiste exatamente na busca de informações para que esse objetivo possa ser definido. Já o capítulo de revisão bibliográfica será organizado após a execução da pesquisa. Ou seja, após definir o objetivo, definir o método de trabalho, executar os experimentos, coletar os resultados e esboçar as conclusões, é que o pesquisador vai organizar o capítulo de revisão bibliográfica onde ele vai inserir os principais conceitos e trabalhos correlatos relevantes para a compreensão da sua pesquisa.

Se as fichas de leitura são organizadas por fonte bibliográfica, o capítulo de revisão bibliográfica não deve ser organizado dessa forma, mas por conceitos. Ou seja, em vez de dizer tudo o que um autor pensa sobre vários conceitos, deve-se dizer o que vários autores pensam sobre um conceito de cada vez.

3.2.2. Tipos de Fontes Bibliográficas

Há vários tipos de fontes bibliográficas. Cada uma terá sua utilidade em determinados momentos da pesquisa. Os *livros* normalmente contêm informação mais completa, didática e bem amadurecida. O objetivo do livro é justamente apresentar uma determinada área da ciência de forma didática e bem fundamentada.

Raramente serão encontradas em livros informações sobre trabalhos futuros, que conduzem a ideias de pesquisa.

Alguns livros se dedicam especificamente à apresentação de problemas em aberto em determinadas áreas, mas estes não são maioria. Em geral, tais livros são produto de eventos científicos. A maioria dos livros didáticos procura apresentar apenas a informação que já está consolidada.

Ideias de pesquisa serão encontradas mais facilmente em artigos curtos que normalmente são publicados em *eventos ou periódicos*. A maioria das ciências exatas preza mais a publicação em periódicos. A área de Computação, porém, tem características distintas em relação a esse aspecto, visto que considera publicações em eventos como sendo tão importantes, e muitas vezes até mais importantes, que publicações em periódicos.

Isso significa que, na área de Computação, bons artigos poderão ser encontrados tanto em eventos quanto em periódicos. Que diferenças podem então ser esperadas?

O processo de submissão e publicação em eventos e periódicos é diferente. Por isso, pode-se esperar diferentes tipos de artigos nesses dois veículos.

Eventos normalmente têm uma data-limite para entrega de trabalhos. De um conjunto de trabalhos enviados e avaliados, os melhores são encaminhados para publicação. Em geral, são sugeridas algumas modificações, mas dificilmente uma segunda rodada de avaliação é realizada.

Já no caso de periódicos, não existem datas-limite, a não ser no caso de edições especiais. Os artigos são submetidos, avaliados, e revisões são sugeridas. Posteriormente, os artigos são enviados de

novo, avaliados e assim por diante. Esse processo de ida e volta pode acontecer várias vezes e em geral pode levar alguns anos até que o artigo seja publicado.

Então, a diferença que se pode esperar é a seguinte: os artigos em eventos terão informações mais atuais, mas poderão variar bastante em termos de qualidade. Já os artigos em periódicos terão sido arduamente revisados e lapidados ao longo de iterações entre autores e revisores, mas quando publicados talvez já não sejam mais tão atuais quanto os artigos em eventos.

Fica a ressalva de que toda regra tem exceção, porque existem periódicos que publicam artigos rapidamente e eventos cujo índice de exigência é tão alto que publicará apenas artigos do mais alto grau de excelência.

3.2.3. O que deve ser Necessariamente Lido

O pesquisador iniciante em uma determinada área deverá começar sua revisão bibliográfica pelos *surveys*. Livros sobre o assunto também seriam uma boa escolha nesse momento. Algumas ferramentas de pesquisa bibliográfica, como *citeSeer* (<http://citeSeer.nj.nec.com/impact.html>), permitem uma busca específica por artigos do tipo *survey*. Além disso, existem publicações especializadas nesse tipo de artigo.

Essas coletâneas apresentam ao pesquisador o estado da arte da área de pesquisa e sua evolução histórica, indicando diferentes desdobramentos e as principais realizações.

Na sequência, alguns trabalhos clássicos devem ser buscados, normalmente destacados nos *surveys*. Além disso, quando se usa uma ferramenta de pesquisa bibliográfica eletrônica, é possível, muitas vezes, solicitar que os trabalhos publicados em uma determinada área sejam ordenados a partir de sua importância. Um trabalho será mais importante na medida em que for citado por outros trabalhos. Os trabalhos clássicos são aqueles que já receberam ao longo do tempo o maior número de citações.

Continuando a pesquisa, deverão ser buscadas as fontes mais recentes sobre o assunto da pesquisa. Artigos muito atuais dificilmente serão clássicos, pois ainda não terá havido tempo para que sejam citados em outras publicações. Porém, é fundamental que um trabalho de pesquisa tenha como referência também os desenvolvimentos mais recentes na área.

3.2.4. Leitura Crítica

A leitura de trabalhos científicos não deve ser encarada apenas como um aprendizado. O pesquisador deve exercer, antes de tudo, o espírito crítico, para questionar a validade de todas as informações registradas nos textos que estão sendo lidos. A aceitação passiva de tudo aquilo o que é lido não gera no pesquisador o espírito de busca por novas informações.

Para que o tema de pesquisa se transforme em um objetivo, é necessário que a cada instante a leitura produza questionamentos. Sem perguntas não há respostas. Muitas perguntas que o pesquisador fizer a si mesmo ao longo da leitura de um texto possivelmente ainda não terão respostas e serão, portanto, excelentes candidatas a objetivo de pesquisa.

Algumas perguntas-chave poderão ajudar o leitor a transformar uma leitura passiva em uma leitura rica e geradora de ideias para pesquisa. Entre elas:

- a) De onde o autor parece tirar suas ideias?
- b) O que foi obtido como resultado deste trabalho?
- c) Como este trabalho se relaciona com outros na mesma área?
- d) Qual seria um próximo passo razoável para dar continuidade a essa pesquisa?
- e) Que ideias de áreas próximas poderiam ser aproveitadas neste trabalho?

As perguntas geradoras mencionadas também poderiam ser usadas para avaliar a qualidade do trabalho que está sendo lido. No caso da primeira pergunta, se não for possível descobrir de onde o autor tira suas ideias, provavelmente se estará diante de um trabalho

fraco, pois as ideias devem vir de referências bibliográficas, ou da observação de fenômenos, ou então são hipóteses criadas pelo autor, as quais serão comprovadas ao longo do trabalho.

Em geral, os autores não podem simplesmente escrever frases como “o interesse pela Internet tem crescido muito ao longo dos últimos anos”. Uma afirmação como esta, embora, à primeira vista, seja consensual, não pode deixar de ter uma base. Essa base pode ser uma referência a outro trabalho, que tenha realizado uma pesquisa sobre o assunto. A base também pode ser um levantamento estatístico realizado pelo próprio autor que demonstra a validade da afirmação. Mais adiante será visto que essa frase específica ainda comporta vários outros problemas, por exemplo:

- a) Como se define e se mede “interesse”?
- b) Como se conceitua “crescer muito”?
- c) Que período de tempo compreende os “últimos anos”?

Em relação à segunda pergunta geradora, “o que exatamente foi obtido como resultado deste trabalho em particular?”, se não houver possibilidade de resumir em poucas palavras a contribuição real do trabalho, então, possivelmente, o texto será confuso e mal organizado, não deixando clara a efetiva contribuição do artigo para a comunidade científica.

Em relação à terceira pergunta geradora, “como este trabalho se relaciona com outros na mesma área?”, espera-se, a princípio, que o próprio artigo deixe bem claro, citando adequadamente os trabalhos correlatos. Caso isso não seja feito, o leitor poderá tentar estabelecer as relações entre o trabalho lido e outras obras. Muitas vezes, aspectos importantes sobre o trabalho (falhas) são descobertos através dessas comparações.

Em relação à quarta pergunta geradora, “qual seria um próximo passo razoável para dar continuidade a essa pesquisa?”, a resposta poderá ser um excelente objetivo de pesquisa. Muitas vezes as questões de pesquisa já estão colocadas no trabalho pelos autores na esperança de que outros grupos deem continuidade à pesquisa.

A quinta pergunta, “que ideias de áreas próximas poderiam ser aproveitadas neste trabalho?”, traz ao pensamento do pesquisador possíveis melhoramentos ao trabalho sendo estudado em função de conceitos correlatos de outras áreas de pesquisa, que possivelmente não eram conhecidos dos autores do trabalho. Dessa forma, o sucesso da aplicação de algum desses conceitos correlatos no trabalho em questão poderá dar origem a uma interessante hipótese de pesquisa, que, se tiver uma justificativa plausível, poderá ser um excelente objetivo de pesquisa.

3.2.5. Exposição à Pesquisa

Além da leitura, o pesquisador, na fase de geração de ideias, deverá estar constantemente exposto a um ambiente científico. No caso de alunos de mestrado e doutorado, na fase de elaboração do objetivo de pesquisa, é fundamental que se tente, pelo menos uma vez por semana, gerar uma ideia de pesquisa para ser discutida com o orientador.

Regularmente, o pesquisador deverá ler pelo menos os resumos dos artigos publicados nos principais periódicos e eventos na sua área de pesquisa. Além dos resumos, deve-se procurar ler pelo menos um ou dois artigos de maior relevância para a área de pesquisa.

Além disso, sempre que possível, o pesquisador deverá participar de palestras e seminários nos quais poderá trocar ideias com outros pesquisadores, além de observar a forma de trabalho de outros grupos de pesquisa. No caso de alunos de mestrado e doutorado, isso implica também participar, como ouvinte, do maior número possível de defesas de teses e dissertações, mesmo que não sejam referentes à sua área de pesquisa específica.

3.2.6. A Ideia de Pesquisa

A ideia de pesquisa, ou problema, surge a partir da leitura e observação. Pode-se comparar a ideia de pesquisa a uma semente que germina no solo da revisão bibliográfica. Com uma revisão bi-

bliográfica superficial ou inexistente até se pode ter ideias, mas possivelmente serão ideias fracas, que dificilmente se desenvolverão de forma saudável. Já com uma boa revisão bibliográfica, tem-se um solo forte para germinação de ideias consistentes.

Portanto, antes de tentar ter ideias, o aluno deve ter experimentado uma boa carga de leitura relacionada ao tema.

3.2.7. Como Sistematizar a Pesquisa Bibliográfica

Um dos problemas que os alunos eventualmente encontram, conforme já mencionado, é justificar que não encontraram nada efetivamente ligado ao assunto que desejam pesquisar. Buscas desorganizadas dificilmente levarão a bons resultados e dificilmente darão segurança ao aluno quando lhe for perguntado “mas será que alguém já não fez isso?”.

Fica aqui então uma sugestão sobre como proceder à pesquisa bibliográfica de maneira sistemática. Cada aluno poderá mudar essas sugestões de acordo com suas necessidades ou disponibilidade. Sugere-se seguir estes passos:

- a) Listar os títulos de periódicos e eventos relevantes para o tema de pesquisa e os títulos de periódicos gerais em computação que eventualmente possam ter algum artigo na área do tema de pesquisa.
- b) Obter a lista de todos os artigos publicados nos últimos cinco (ou mais) anos nesses veículos.
- c) Selecionar dessa lista aqueles títulos que tenham relação com o tema de pesquisa.
- d) Ler o *abstract* desses artigos e, em função da leitura, classificá-los como relevância “alta”, “média” ou “baixa”.
- e) Ler os artigos de alta relevância e fazer fichas de leitura anotando os principais conceitos e ideias aprendidos. Anotar também títulos de outros artigos possivelmente mencionados na biografia de cada artigo (mesmo que com mais de cinco anos) e que pareçam relevantes para o trabalho de pesquisa. Incluir

esses artigos na lista dos que devem ser lidos (inicialmente o *abstract* e, se for relevante, o artigo todo).

- f) Dependendo do caso, ler também os artigos de relevância média e baixa, mas iniciando sempre pelos de alta relevância.

O conjunto de referências mencionadas no terceiro passo deve ser efetivamente produzido na forma de uma listagem que possa ser posteriormente consultada ou até indicada na bibliografia do trabalho como fonte de pesquisa. Portanto, essa fase do processo deve ser bem documentada.

Depois do último passo, o aluno poderá decidir se:

- a) Já tem material suficiente para elaborar uma ideia de pesquisa consistente.
- b) Precisa expandir a pesquisa examinando artigos mais antigos (expandindo o passo “b”) ou periódicos menos relevantes (expandindo o passo “a”).

A decisão sobre se já há material suficiente para finalizar a revisão bibliográfica usualmente será tomada em conjunto com o orientador.

Uma última técnica ainda para expandir a revisão bibliográfica, quando necessário, é consultar as referências bibliográficas citadas nos trabalhos mais importantes consultados.

3.2.8. Como Terminar a Revisão Bibliográfica

Em geral, existe sempre conhecimento novo sendo produzido em todas as áreas da ciência. A revisão bibliográfica, portanto, praticamente não termina nunca.

Mas o trabalho de pesquisa deve ter um término claro. E é necessário saber que, durante o seu desenvolvimento, apenas a revisão bibliográfica não produz nenhum conhecimento novo.

É necessário, portanto, passar gradualmente da fase de revisão bibliográfica para a fase de realização propriamente dita do trabalho de pesquisa, ou seja, a construção de teorias e realização de experimentos para testar hipóteses.

3.3. O Objetivo

O objetivo da pesquisa deve ser diretamente verificável ao final do trabalho. Um bom objetivo de pesquisa possivelmente irá demonstrar que alguma hipótese sendo testada é ou não verdadeira.

Portanto, o objetivo geral e os objetivos específicos do trabalho devem ser expressos na forma de uma condição não trivial cujo sucesso possa vir a ser verificado ao final do trabalho. Um objetivo bem expresso em geral terá verbos como “demonstrar”, “provar”, “melhorar” (de acordo com alguma métrica definida) etc.

Deve-se tomar cuidado com certos verbos que determinam objetivos cuja verificação é trivial e, portanto, inadequada. Entre eles pode-se citar “propor”, “estudar”, “apresentar” etc. Se o objetivo do trabalho é *propor* algo, basta que a coisa seja proposta para que o objetivo seja atingido e, portanto, essa forma é trivial e inadequada, pois a definição do objetivo não menciona a qualidade daquilo que será proposto.

Se o objetivo do trabalho é *estudar* algo, então ele terá sido alcançado se aquilo foi estudado, não importando se alguma nova informação foi aprendida ou não, sendo, portanto, inadequado como objetivo de pesquisa. Estudar, normalmente, é o objetivo do *aluno* e não do trabalho.

Se o objetivo do trabalho consiste em *apresentar* algo, novamente ele é trivial e inadequado. Uma simples apresentação não produz necessariamente conhecimento novo. Por exemplo, “o objetivo deste trabalho é apresentar os operadores da lógica booleana”; tal objetivo pode ser alcançado com um pequeno texto explicando os operadores conhecidos, mas, como não traz informação nova, não é um objetivo de pesquisa.

A proposta, o estudo e a apresentação podem ser justificáveis como objetivo de pesquisa desde que o objeto da proposta, estudo ou da apresentação seja algo original.

Segundo Chinneck (1988), uma monografia deve apresentar uma contribuição original ao conhecimento. Dessa forma, ao final do trabalho, o estudante deverá ser capaz de mostrar que identifi-

cou um problema que valia a pena ser resolvido, mas que ainda não havia sido. Além disso, o estudante deverá mostrar que ele resolveu o problema que propôs e apresentar a solução.

Em função disso, Chinneck conclui que um avaliador, ao ler o texto de uma monografia, vai procurar responder às seguintes questões:

- a) Qual é a questão de pesquisa que o aluno propôs?
- b) É uma boa questão? (Já foi respondida alguma vez? Vale a pena respondê-la?)
- c) O aluno conseguiu convencer que a questão foi respondida adequadamente?
- d) O aluno fez uma contribuição adequada ao conhecimento?

A falha em encontrar respostas para alguma dessas questões poderá colocar o aluno em apuros, sendo que a banca avaliadora provavelmente exigirá revisões extensas no trabalho ou poderá até reprovar o candidato.

3.3.1. A Extensão do Objetivo de Pesquisa

Um objetivo de pesquisa, dependendo do nível de formação pretendido (graduação, especialização, mestrado ou doutorado), não pode ser demasiadamente trivial nem demasiadamente complexo.

Um objetivo demasiadamente trivial será atingido rapidamente, mas dificilmente será defensável perante uma banca, pois a cada grau acadêmico espera-se do aluno não apenas inspiração, mas também transpiração, ou seja, não basta ter uma boa ideia, é necessário trabalhar sobre ela demonstrando todos os seus diferentes aspectos dentro do nível de complexidade exigido pelo grau almejado.

Um objetivo demasiadamente complexo dificilmente será atingido no tempo disponível para o cumprimento dos requisitos do grau almejado. Objetivos complexos podem ser colocados como projetos de longo prazo na carreira de pesquisadores já formados, os quais terão, muitas vezes, décadas à frente e equipes de pesquisa para buscar esses objetivos.

Alunos de graduação e pós-graduação devem atingir os objetivos colocados dentro do tempo regulamentar que seus cursos estabelecem, e, portanto, a complexidade desses objetivos deve ser consistente com esse tempo. Não basta, para a obtenção de um grau, apresentar uma ideia complexa que não pode ser demonstrada ou concluída no tempo disponível. Alegar para a banca que não houve tempo para concluir o trabalho dificilmente é uma boa desculpa.

Sendo assim, o aluno deverá sempre buscar apoio no seu orientador e a definição de um objetivo de pesquisa que possa ser buscado no tempo disponível, de acordo com o curso realizado. Quanto mais experiente for o orientador, ou seja, quanto mais trabalhos ele tiver orientado com sucesso em sua vida, mais se pode esperar que seja capaz de apoiar o aluno na escolha de um objetivo adequado.

3.3.2. Objetivo de Pesquisa *versus* Objetivo Técnico

É aceitável que um trabalho de graduação e mesmo de especialização tenham objetivos técnicos, ou seja, espera-se nesses graus que os alunos sejam capazes de demonstrar que aprenderam determinados conceitos e conseguem colocá-los em prática. Assim, é aceitável que um aluno de graduação, ao final de seu curso, desenvolva um sistema usando conceitos aprendidos durante o curso e que apresente o sistema como trabalho final. Porém, esse tipo de abordagem não é aceitável nos níveis de mestrado e doutorado. Espera-se que o mestre e o doutor tenham de alguma forma contribuído para o avanço do conhecimento. Assim, o desenvolvimento de um sistema, embora possa ser necessário para provar alguma hipótese previamente estabelecida, não é em si suficiente para a concessão de um título de mestre ou doutor.

Se o simples desenvolvimento de um sistema ou de um protótipo fosse suficiente para outorgar um grau acadêmico, então as universidades deveriam estar conferindo título de mestre a todos os programadores ou analistas que diariamente desenvolvem sistemas, muitas vezes complexos, nas suas empresas. Se isso não acontece é porque existe algo mais nas dissertações e teses que precisa ser

buscado. Esse algo mais é o *conhecimento*. Uma monografia é um documento que apresenta de forma organizada uma contribuição para o estado da arte, apresentando, portanto, informações que não eram conhecidas e que a partir do momento em que são publicadas passam a integrar o corpo de conhecimento relevante para quem for atuar em uma determinada área.

Os trabalhos técnicos, que são aceitáveis na graduação e na especialização, apenas usam o conhecimento já disponível.

Os trabalhos científicos, que devem obrigatoriamente ser desenvolvidos no mestrado e no doutorado, devem, além de usar o conhecimento já disponível, criar novos conhecimentos, associando-os dentro de uma estrutura coerente àqueles que já são conhecidos. Sendo assim, o desenvolvimento de sistemas ou protótipos servirá apenas como apoio para demonstrar a aplicabilidade desses novos conhecimentos, se assim for necessário.

3.3.3. Os Objetivos Específicos

Os objetivos específicos devem ser escolhidos da mesma forma que o objetivo geral, ou seja, devem ser não triviais e verificáveis ao final do trabalho. Normalmente, os objetivos específicos não são etapas do trabalho, mas subprodutos. Deve-se tomar cuidado para não confundir os objetivos específicos com os passos do método de pesquisa.

A implementação de um protótipo ou a coleta de dados empíricos possivelmente serão etapas dentro de um trabalho e, portanto, parte do método de trabalho. Esses passos não são, assim, objetivos específicos.

Deve-se entender, portanto, que os objetivos específicos são detalhamentos ou subprodutos do objetivo geral. Se o objetivo geral consiste em provar uma determinada hipótese, os objetivos específicos podem estabelecer a prova de uma série de condições associadas a tal hipótese.

3.4. O Método de Pesquisa

Em geral as monografias têm um capítulo ou seção designados como “metodologia”. Entretanto, *metodologia* seria o estudo dos métodos. Apesar do uso corrente, linguisticamente seria mais correto afirmar que um trabalho científico individualmente tem um *método* de pesquisa e não uma metodologia.

Segundo a Wikipédia (2009a), “A Metodologia é o estudo dos métodos. (...) Tem como finalidade captar e analisar as características dos vários métodos disponíveis, avaliar suas capacidades, potencialidades, limitações ou distorções e criticar os pressupostos ou as implicações de sua utilização.”

Dificilmente um aluno escrevendo uma monografia apresenta uma *metodologia* com essas características. O usual é que, dependendo do tipo de problema, se escolha um método de trabalho e siga utilizando-o até comprovar ou refutar as hipóteses. Assim, os termos *método* e *metodologia* serão aqui claramente diferenciados.

Ainda segundo a Wikipédia (2009b), “O método científico é um conjunto de regras básicas para desenvolver uma experiência a fim de produzir novo conhecimento, bem como corrigir e integrar conhecimentos preexistentes. Na maioria das disciplinas científicas consiste em juntar evidências observáveis, empíricas (ou seja, baseadas apenas na experiência) e mensuráveis e as analisar com o uso da lógica.”

Este livro trata de metodologia porque apresenta um estudo de métodos. Mas o livro em si teve um método para ser concebido e escrito, o qual não é aqui descrito. Porém, uma monografia deverá sempre descrever o método usado para chegar às conclusões.

O método propriamente dito de um trabalho científico só pode ser estabelecido depois que o objetivo tiver sido definido. Por esse motivo, no caso da Computação, normalmente a revisão bibliográfica não deveria nem fazer parte do método. A revisão bibliográfica consiste em um passo do trabalho no qual o aluno vai iniciar ou aprofundar seus conhecimentos em um campo do saber para que possa então propor um objetivo que seja coerente com o grau que deseja obter. Ou seja, a etapa de revisão bibliográfica não seria parte

do método, mas um *prerrequisito* para realização do trabalho de pesquisa, pois quem não estudou o assunto não tem como propor um objetivo válido.

O método consiste na sequência de passos necessários para demonstrar que o objetivo proposto foi atingido, ou seja, se os passos definidos no método forem executados, os resultados obtidos deverão ser convincentes.

O método deve então indicar se protótipos serão desenvolvidos, se modelos teóricos serão construídos, quais experimentos eventualmente serão realizados, como os dados serão organizados e comparados, e assim por diante, dependendo do tipo de trabalho.

A definição do método de pesquisa é um passo fundamental a ser executado logo após a definição do objetivo. Dado o objetivo, o método descreve o caminho para atingi-lo. Assim, deverá ser suficiente trilhar o caminho descrito pelo método para se alcançar o objetivo. Se o objetivo e o método foram coerentes, então grande parte do trabalho de pesquisa já terá sido executada, restando apenas a execução dos passos descritos no método.

Porém, descrever um conjunto de passos que constitua um método de trabalho científico aceitável exige alguns conhecimentos sobre o método científico que serão detalhados nesta seção. A não observância do método científico pode levar a conclusões erradas ou forçadas.

Propostas metodologicamente ingênuas não são totalmente incomuns em computação. Coisas como “trabalhar com dois grupos, um com a ferramenta e outro sem a ferramenta”, até poderia ser parte de um método, mas não é suficiente. Se a diferença entre as médias dos dois grupos for de 0,5 ponto percentual, pode-se concluir que um grupo foi melhor que o outro? Ou pode ter sido obra do acaso? E se a diferença for de cinco pontos percentuais? Como saber? Existem algumas informações trazidas pela estatística que devem ser do conhecimento de qualquer pessoa que se aventure a desenvolver pesquisa científica.

Além disso, existe ainda a possibilidade do mau uso da lógica. Os antigos sofistas eram bastante requisitados entre os gregos para argumentação. Entretanto, nem sempre a lógica usada por eles era exatamente a que poderia ser aceita cientificamente. Um exemplo hilário de uso tortuoso da lógica aparece no filme *Monty Python em Busca do Cálice Sagrado*, quando Bedevere propõe um método infalível para determinar a identidade de uma bruxa. Ao final de uma série de argumentações pseudológicas, ele conclui que, se a mulher pesar tanto quanto um pato, será uma bruxa. Basicamente, segundo Bedevere, bruxas queimam e madeira também queima. Madeira flutua na água assim como os patos. Logo, se a mulher pesar tanto quanto um pato é feita de madeira e, portanto, é uma bruxa.

3.4.1. Dados *versus* Conceitos

O método de pesquisa não consiste apenas em coletar dados para suportar a hipótese de trabalho. É necessário elaborar um discurso ponderado e esclarecedor a partir desses dados. O aspecto mais importante de uma monografia é o pensamento crítico e não apenas a coleta de informações. Trabalhos acadêmicos que se restringem à realização de pesquisas de opinião através de questionários com a consequente tabulação dos dados e apresentação de gráficos não terão validade se não trouxeram consigo alguma informação nova.

Lakatos e Marconi (2006) identificam o questionário como um instrumento de pesquisa que dispensa a presença do pesquisador. Porém, deve-se evitar a distribuição de questionários para serem respondidos se não se sabe com antecedência qual a informação que se está buscando, ou seja, qual a hipótese que se está tentando comprovar.

Além disso, em geral respostas diretas e interpretações simplistas não são as mais interessantes. Por exemplo, há algum tempo pesquisadores realizaram uma pesquisa na Inglaterra, onde entrevistaram homens e mulheres perguntando quantos parceiros sexuais haviam tido ao longo da vida. O resultado médio para as mulheres foi três e o resultado médio para os homens foi 10. Uma interpre-

tação ingênuo e simplista desse fato diria que o homem em média tem mais parceiros do que as mulheres. Porém, considerando-se que existe aproximadamente o mesmo número de homens e mulheres na sociedade, isso é impossível, visto que cada vez que um homem tem uma nova parceira, uma mulher (a parceira) também tem automaticamente um novo parceiro. Portanto, as duas médias deveriam ser praticamente iguais. A conclusão da pesquisa foi, portanto, que ou os homens mentem para mais, ou as mulheres mentem para menos, ou ambos mentem quando se trata de quantificar o número de parceiros.

3.4.2. Pesquisa Experimental e Não-Experimental

Uma das caracterizações clássicas das formas de pesquisa consiste na classificação da pesquisa experimental e não-experimental.

Basicamente, a pesquisa não-experimental consiste no estudo de fenômenos sem a intervenção sistemática do pesquisador. Por exemplo, um pesquisador que estude o dia a dia de uma empresa de desenvolvimento de software para detectar ali determinadas práticas previamente catalogadas está fazendo uma pesquisa não-experimental, pois ele age apenas observando e tirando conclusões a partir de um arcabouço teórico preconcebido.

Já a pesquisa experimental implica que o pesquisador sistematicamente provocará alterações no ambiente a ser pesquisado de forma a observar se cada intervenção produz os resultados esperados.

No exemplo anterior, se o pesquisador resolver criar artificialmente situações de estresse dentro da empresa para verificar como os funcionários reagem, então ele estará realizando pesquisa experimental.

3.4.3. Objetividade

Uma característica bastante importante da pesquisa científica é a objetividade. O pensamento humano permite a tirada de conclusões que nem sempre são objetivas, como no caso das opiniões.

Por exemplo, um desenvolvedor pode considerar que a programação orientada a objetos é mais fácil de usar do que a programação estruturada. Isso é apenas uma questão de opinião. Mesmo que o desenvolvedor consiga estabelecer todo um conjunto de argumentos para justificar essa opinião, ela ainda carece de objetividade científica. Por quê? Porque a facilidade de uso não é uma conclusão à qual chegaria qualquer observador independente. Um programador experimentado em orientação a objetos pode considerar muito fácil trabalhar com objetos e mensagens, mas um iniciante ou ainda um programador experimentado em Cobol poderia ter dificuldade com tais conceitos. Esse último ainda poderia sugerir que Cobol é que é muito mais fácil de usar.

Segundo Kerlinger (1980), “a condição principal para satisfazer o critério de objetividade é, idealmente, que quaisquer observadores com um mínimo de competência concordem com seus resultados”.

Os experimentos e as observações no trabalho científico devem, então, ser tratados de forma objetiva. Qualquer observação ou grandeza que se queira avaliar deve ser definida de forma a que leituras possam ser feitas independentemente do observador que as toma. Por exemplo, o pesquisador poderia dizer que um sistema é “fácil de usar” se um determinado conjunto de tarefas predefinido puder ser executado por um usuário com um determinado grau de treinamento dentro de um período de tempo predeterminado. A objetividade do conceito passa a ser então a pertinência a um conjunto discreto.

Ainda é possível definir um fenômeno objetivamente, mas de forma difusa, isto é, em vez de trabalhar com apenas duas categorias como “fácil” ou “difícil”, trabalha-se com um grau de dificuldade inherente. Novamente, a definição objetiva deve considerar uma medição que possa ser realizada de forma independente por diferentes observadores e ainda assim chegar aos mesmos resultados. Elaborando sobre o exemplo anterior, então, seria possível definir o grau de dificuldade como o tempo médio que usuários com determinado grau de treinamento levam para realizar um conjunto de tarefas pre-

viamente definido usando a ferramenta. Nesse caso, diversas medições chegarão a valores objetivos e comparações poderão ser feitas entre sistemas.

Porém, ter uma definição objetiva de um fenômeno não é suficiente para se produzir um trabalho de qualidade. É necessário observar ainda a qualidade da própria definição. Ou seja, poderá haver pesquisadores que não concordem que o tempo médio para realizar um conjunto de tarefas seja uma boa definição para “facilidade de uso”. Levando o caso a um extremo para melhor compreensão, se um pesquisador definisse que um sistema é fácil de usar apenas se puder ser utilizado via Internet, essa definição dificilmente seria aceita, pois se afasta muito daquilo que em geral se considera, mesmo intuitivamente, como facilidade de uso.

Então, para trabalhar com observações de fenômenos, o pesquisador deve definir de maneira objetiva o fenômeno que vai observar e convencer os demais de que essa definição é razoavelmente intuitiva, mas em especial útil, para chegar a algum resultado.

3.4.4. Empirismo

Há um ditado que já vem se tornando clássico em computação: “Teoria é quando o fenômeno é compreendido, mas não funciona. Prática é quando funciona, mas não se sabe por quê. Na Computação coexistem a teoria e a prática: nada funciona e não se sabe por quê.”

Essa *prática* tem sido um dos significados adotados para o termo “empírico” por desenvolvedores na área de computação. Não importa por que, basta que funcione. Se funcionar está certo! Esse ponto de vista pragmático, porém, não corresponde ao significado de *empírico* dado pela ciência. Segundo Kerlinger (1980), “empírico significa guiado pela evidência obtida em pesquisa científica sistemática e controlada”.

A computação, enquanto ciência, fundamenta suas pesquisas no empirismo e não no princípio da autoridade. Em computação, na maioria das vezes, pouco importa a opinião deste ou daquele expoente, mas as conclusões objetivas obtidas empiricamente.

A falta de empirismo pode levar a conclusões erradas. Na idade média acreditava-se que o homem tinha uma costela a menos do que a mulher, já que a Bíblia relatava que a mulher havia sido criada a partir de uma costela do homem. Essa crença poderia ser refutada sem dificuldades caso alguém usasse de empirismo para simplesmente contar a quantidade de costelas em um e outro sexo.

Já Descartes (2004) acreditava que o coração batia porque era o lugar mais quente do corpo, e seu calor inflava o sangue que entrava nele. Essa conclusão obtida com alguma reflexão mas pouca evidência empírica não serve para mais nada hoje, pelo menos na área da Medicina.

O empirismo é importante para a ciência, portanto, porque é uma maneira sensata de olhar o mundo. Não basta acreditar em sua intuição ou nas palavras dos mestres. É preciso verificar objetivamente se o fenômeno descrito realmente é verdadeiro.

Antes foi mencionado o fato de que o interesse pela Internet vem crescendo muito nos últimos anos. Será verdade? Quem disse? E de onde observou esse fato? É um sentimento comum e intuitivo. Possivelmente vários livros e artigos sobre Internet têm falado isso ao longo dos anos. Mas será essa afirmação realmente verdadeira hoje? Não estará o interesse pela Internet estável ou diminuindo? O leitor provavelmente responderá “claro que não” a essa última pergunta. Mas como pode ter tanta certeza? Fez alguma observação empírica? Tem dados? Afinal o que é *interesse*? Como se mede?

Essas dúvidas são importantes para o cientista que quer realmente entender os fenômenos e descobrir novos conhecimentos. Se os médicos atualmente continuassem acreditando que o coração bate porque aquece o sangue, baseados na afirmação de Descartes, pouquíssimos avanços da Medicina teriam sido conseguidos. **Duvidar das conclusões de outros cientistas e duvidar do próprio senso comum, muitas vezes, é a chave para grandes descobertas.**

3.4.5. Variáveis

Foi mencionada anteriormente a existência de fenômenos para os quais interessa ao cientista realizar medições objetivas. Em geral, **tais fenômenos que podem ser medidos objetivamente são identificados como variáveis.** Uma variável é um nome que se dá a um fenômeno que pode ser medido e que varia conforme a medição. Se não variasse seria uma *constante* e não teria maior interesse para a pesquisa.

Variáveis em experimentos, assim como as variáveis em programas computacionais possuem um domínio, ou seja, um conjunto de valores dentro do qual a variável se altera. Por exemplo, uma temperatura em geral será um número racional limitado inferiormente pelo zero absoluto (o valor literal dependerá da escala que se use). Uma variável relacionada à temperatura não poderá assumir o valor “Z” ou “falso”, já que esses não pertencem ao seu domínio.

O domínio de uma variável pode ser discreto ou contínuo. Variáveis contínuas assumem valores reais, a ideia de contínuo vem do fato de que entre dois valores sempre existe um terceiro. Já as variáveis discretas assumem seus valores em conjuntos cujos elementos podem ser ordenados ou em conjuntos finitos. O número de participantes de uma videoconferência é um fenômeno representado em uma variável cujo domínio é discreto e corresponde ao conjunto dos números naturais. Os valores desse conjunto podem ser ordenados, e entre o n -ésimo valor e o $(n+1)$ -ésimo valor não existe uma terceira possibilidade para qualquer n que seja um número natural.

Algumas variáveis discretas assumem seus valores em conjuntos finitos. Tais variáveis são denominadas *categóricas*. Por exemplo, as notas que um estudante de mestrado na UFSC pode obter em uma disciplina variam no conjunto {A, B, C, E}. Em ciência, muitas vezes fenômenos observados são classificados com variáveis categóricas, pois estas são em geral mais fáceis de compreender do que valores contínuos. Caso se avaliasse alunos com valores contínuos, seria razável concluir que um aluno com nota 6,7812 é melhor do que um aluno com nota 6,7811?

Em alguns casos é possível estabelecer regras de transformação de valores contínuos para discretos. Esse processo, chamado *discretização*, consiste em atribuir um valor discreto diferente para variados intervalos de valores contínuos. Por exemplo, considerando-se notas arredondadas para uma casa decimal após a vírgula, pode-se estabelecer uma correspondência entre o domínio contínuo e a variável categórica da seguinte maneira: notas de 0,0 a 4,9 poderiam ser consideradas como o conceito E, notas de 5,0 a 6,9, como conceito C, notas de 7,0 a 8,9, como conceito B, e notas de 9,0 a 10,0, como conceito A. Note-se, porém, que quando se aplica o conceito de arredondamento para uma casa decimal depois da vírgula o domínio da variável já foi transformado de contínuo em discreto, pois, dessa forma, as notas possíveis limitam-se a 101 valores: {0,0, 0,1, 0,2, ..., 9,8, 9,9, 10,0}.

Na pesquisa científica é importante classificar ainda as variáveis como *medidas* ou *manipuladas*. Uma variável *medida* é aquela cujo fenômeno será observado pelo pesquisador. Por exemplo, quantas vezes um usuário de uma ferramenta vai olhar no manual para obter informações para desempenhar a tarefa que lhe foi proposta. Essa variável tem como domínio o conjunto dos números naturais, e seus valores não são determinados pelo observador, mas simplesmente medidos.

Já a variável *manipulada* é aquela que o experimentador vai deliberadamente modificar para realizar seu experimento. Por esse motivo, tal variável também é chamado de variável *experimental*. Um exemplo de variável manipulada poderia ser o número de passos da tarefa repassada aos usuários. O pesquisador poderia passar a alguns usuários, por exemplo, tarefas com cinco passos, a outros com 10 passos, a outros com 15 passos etc. Assim, ao fazer a experimentação, o pesquisador *manipula* a variável referente ao número de passos da tarefa e *observa* o comportamento da variável medida que consiste em contar quantas vezes o usuário vai olhar o manual da ferramenta.

Mas por que pesquisadores manipulam uma ou mais variáveis enquanto observam outras? É porque eles querem encontrar *dependências* entre essas variáveis. No exemplo anterior, possivelmente o pesquisador estaria tentando descobrir se tarefas mais longas implicam ou não o usuário consultar o manual do sistema mais vezes.

A princípio pode-se testar a dependência entre quaisquer variáveis manipuladas e observadas. Mas nem sempre esse teste fará sentido. Antes de analisar uma dependência experimentalmente, o pesquisador em geral desenvolve uma teoria ou hipótese. No caso anterior, a hipótese poderia ser que quanto maior a tarefa, maior a consulta ao manual feita pelo usuário. Outra hipótese possível seria que não importa o tamanho da tarefa, isso não influenciará no número de vezes que o usuário consulta o manual. É uma dependência que tanto em um caso quanto no outro parecem valer a pena testar.

Mas nem sempre dependências que podem ser testadas fazem sentido enquanto hipótese ou teoria. Por exemplo, manipular a variável “número de botões na tela do aplicativo” e medir a variável “número de vezes que o usuário espirra enquanto está usando a ferramenta”, pode até ser divertido, mas dificilmente haverá uma conexão entre os dois fenômenos. Outro exemplo seria observar o dia da semana e tentar verificar se programas rodam mais rapidamente na sexta-feira. O pesquisador até poderia criar experimentos controlados para mensurar isso. Mas dificilmente encontraria uma conexão entre o dia da semana e a velocidade dos programas. Até porque a teoria corrente sobre programas não estabelece nenhuma conexão entre sua velocidade e o dia da semana. Tal hipótese então se revelaria falha.

Daí vem a importância de se trabalhar com uma boa teoria em pesquisa. Não basta realizar experimentos e encontrar relações entre variáveis. É preciso ter uma teoria que procure explicar o porquê dessas relações.

Há aqui também a caracterização, bastante comum em pesquisa, dos termos *variável dependente* e *variável independente*. A variável independente é aquela que, se supõe, influencia outra. A variável dependente é a influenciada.

Uma hipótese de pesquisa científica, então, em geral terá uma associação de implicação, ou seja, antecedente/consequente, entre uma ou mais variáveis independentes e uma ou mais variáveis dependentes. Por exemplo, o número de passos em uma tarefa implica o aumento do número de vezes que o usuário consulta o manual? Essa seria uma hipótese de pesquisa em que a variável dependente é o número de consultas ao manual, e a variável independente é o número de passos da tarefa. Em geral, o pesquisador manipula a variável independente e mede a dependente. Nesse caso, ainda, como ambas as variáveis são numéricas, pode-se afirmar que se busca encontrar uma dependência *direta*, ou seja, quanto maior o número de passos, maior o número de consultas. Em outros casos, se poderia pesquisar dependências inversas, em que o aumento no valor da variável independente causa uma diminuição na variável dependente.

Ainda existe a dependência linear e não-linear. A linear pode ser aproximada por um polinômio de primeiro grau. Por exemplo, poderia ser observado que, para cada passo na tarefa, o usuário faz mais duas consultas ao manual. Nesse caso, a relação entre o número de passos x e o número de consultas y poderia ser expressa pela função $y = 2x$.

Dependências não-lineares usualmente são representadas por polinômios de grau superior a um ou ainda por equações com exponenciais, raízes ou logaritmos. Este livro apresentará apenas os conceitos relacionados à dependência linear. Para o leitor que quiser se aprofundar em outras formas de dependência recomenda-se a leitura de um bom livro de estatística, como o de Barbetta, Reis e Bornia (2008).

3.4.6. Variância e Desvio-Padrão

Na pesquisa científica, frequentemente o pesquisador defronta-se com o problema de analisar conjuntos de dados. Por exemplo, ao avaliar um determinado sistema, o pesquisador contabiliza o tempo de interação de cada pessoa dentre um conjunto previamente

definido. Em geral, a *média* é considerada uma medida importante na avaliação de conjuntos de valores. Por exemplo, se quatro pessoas foram analisadas e os tempos medidos em minutos foram 10, 12, 14, 9, então se pode dizer que o tempo médio observado foi de 11,25 minutos.

A média aritmética simples consiste em somar os valores e dividir o resultado pela quantidade de valores. Por isso, a média acima resultou do cálculo $(10+12+14+9)/4 = 11,25$.

Porém, na pesquisa científica há outra medida importante, que é a *variância* do conjunto de valores. Considerem-se os três conjuntos¹ de valores a seguir:

- {10, 12, 14, 9}
- {1, 20, 2, 22}
- {11, 11, 11, 12}

É possível notar certa semelhança entre eles? Aparentemente são conjuntos bem diferentes. Mas todos têm a mesma média: 11,25. A média, então, passa alguma informação sobre a natureza dos conjuntos de valores, mas ela sozinha não é a única informação importante.

Qual a diferença notável entre esses conjuntos? O terceiro apresenta apenas valores muito próximos da média (ou muito próximos entre si). Já o segundo apresenta valores bastante distanciados da média. O primeiro conjunto é um caso intermediário.

Essa observação do distanciamento dos elementos em relação à média é chamada de *variância*. Então, além da média, o pesquisador deve ficar atento também à *variância* do conjunto de valores, já que esta complementa a caracterização do conjunto.

A variância de um conjunto pode ser definida de forma numérica. Basicamente deseja-se uma variância alta quando os valores se afastam muito da média, seja para cima, seja para baixo. Então, a

¹ O conjunto, na matemática, não admite repetição de elementos. No entanto, aqui a palavra “conjunto” é usada em sentido mais amplo, significando “coleção” ou “multiconjunto” e, portanto, admitindo repetição de elementos.

variância pode ser o cálculo da diferença de cada valor do conjunto em relação à média deste. No caso dos três exemplos anteriores, subtraindo 11,25 de cada valor o resultado seria:

$$\{10-11,25, 12-11,25, 14-11,25, 9-11,25\} = \{-1,25, 0,75, 2,75, -2,25\}$$

$$\{1-11,25, 20-11,25, 2-11,25, 22-11,25\} = \{-10,25, 8,75, -9,25, 10,75\}$$

$$\{11-11,25, 11-11,25, 11-11,25, 12-11,25\} = \{-0,25, -0,25, -0,25, 0,75\}$$

O resultado corresponde à intuição: o terceiro conjunto, de baixa variância, tem valores absolutos pequenos referentes à diferença entre os valores originais e a média. Já o segundo conjunto tem os valores absolutos mais altos, e o primeiro tem valores intermediários.

Agora, para obter um valor escalar para a variância se poderia somar os valores obtidos. Porém, alguns deles são negativos. Não seria intuitivo que um valor negativo pudesse anular a influência de um valor positivo no cálculo da variância. Inclusive, pela própria definição até aqui usada, essas somas invariavelmente resultariam em zero (o leitor pode conferir somando os valores dos três conjuntos anteriores). Então, a simples soma dos valores não é uma solução. É preciso somar apenas valores de mesmo sinal para que se tenha efetivamente um escalar que corresponda à medida da distância dos valores em relação à média do conjunto.

Uma opção seria trabalhar com o valor absoluto de cada diferença. Mas a definição oficial de variância eleva os valores das diferenças ao quadrado como forma de obter apenas valores positivos. Elevar esses valores ao quadrado também permite aumentar a influência dos valores mais afastados da média. Essa escolha tem origem em questões cuja explicação foge ao escopo deste livro.

Então, elevando ao quadrado cada um dos valores dos conjuntos de diferenças, obtém-se:

$$\{(-1,25)^2, 0,75^2, 2,75^2, (-2,25)^2\} = \{1,5625, 0,5625, 7,5625, 5,0625\}$$

$$\{(-10,25)^2, 8,75^2, (-9,25)^2, 10,75^2\} = \{105,0625, 76,5625, 85,5625, 115,5625\}$$

$$\{(-0,25)^2, (-0,25)^2, (-0,25)^2, 0,75^2\} = \{0,0625, 0,0625, 0,0625, 0,5625\}$$

Observa-se que os valores do terceiro conjunto são muito baixos (próximos de zero), enquanto os do segundo conjunto são altos e os do primeiro conjunto são intermediários, como se poderia esperar.

O próximo passo é somar os valores obtidos para os elementos de cada conjunto, ficando-se respectivamente com:

$$\sum \{1,5625, 0,5625, 7,5625, 5,0625\} = 14,75$$

$$\sum \{105,0625, 76,5625, 85,5625, 115,5625\} = 382,75$$

$$\sum \{0,0625, 0,0625, 0,0625, 0,5625\} = 0,75$$

Novamente como esperado, o maior valor, 382,75, fica com o segundo conjunto, o menor com o terceiro e o intermediário com o primeiro.

Porém, somando-se simplesmente as diferenças, como foi feito antes, fica-se com um valor de variância que aumenta conforme o tamanho do conjunto. Isso não é intuitivo. A variância de um conjunto seria a dispersão *média* dos elementos em relação à média do conjunto. Então ela não deveria crescer com o tamanho do conjunto. É necessário então, para eliminar esse efeito, dividir o valor escalar encontrado pelo número de elementos do conjunto.

Porém, não é intuitivo dizer que um conjunto com um único elemento tenha variância. No mínimo dois elementos são necessários para que faça sentido dizer que eles se afastam da média. Assim, a divisão não será feita pelo número de elementos no conjunto, mas pelo número de elementos menos um. Assim, no caso de um conjunto com um elemento, a variância será o resultado de uma divisão por zero e, portanto, indefinida.

No exemplo anterior, aplicando-se a divisão por três aos valores de variância obtidos, fica-se com os valores finais de:

$$14,75/3 = 4,9166\dots$$

$$382,75/3 = 127,5833\dots$$

$$0,75/3 = 0,25$$

O que se pode concluir aqui é que o segundo conjunto é o que mais varia, ou seja, seus elementos mais se afastam da média, enquanto o terceiro conjunto é o que menos varia, ou seja, ele é o que tem os elementos relativamente mais próximos da média.

A fórmula da variância pode ser então assim definida:

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})}{n - 1}$$

Onde s_x^2 é o símbolo comumente usado para representar a variância de X , n é o número de elementos do conjunto, \bar{x} representa a média aritmética do conjunto e x_i representa cada um dos elementos do conjunto no somatório.

O *desvio-padrão* é uma medida também bastante utilizada para analisar conjuntos e é definido simplesmente como a raiz quadrada da variância, ou seja:

$$s_x = \sqrt[2]{s_x^2}$$

onde s_x é o símbolo comumente usado para representar o desvio-padrão.

No exemplo anterior, os valores de desvio-padrão dos três conjuntos são respectivamente e aproximadamente:

$$\sqrt[2]{4,9166...} = 2,217355783$$

$$\sqrt[2]{127,5833...} = 11,29527925$$

$$\sqrt[2]{0,25\ 0,75 / 3} = 0,5$$

3.4.7. Covariância

A variância é uma medida muito importante para a pesquisa científica. Basicamente pode-se dizer que não haveria muito o que fazer se os fenômenos não variassem. Se cada variável estudada tivesse variância zero, então os valores seriam previsíveis e pouco se poderia aprender sobre a natureza das coisas.

Entretanto, os fenômenos variam, e quanto maior a variância, mais interessante pode ser o fenômeno. A pesquisa em engenharia de software, por exemplo, tenta descobrir formas de estimar quanto tempo um programa levaria para ser desenvolvido. É uma medida difícil porque mesmo que se tenha uma descrição detalhada de cada função ou caso de uso a ser desenvolvido, o tempo que o desenvolvedor levaria para programar cada um desses elementos poderia variar de alguns minutos a várias semanas.

Saber o tempo esperado (ou o tempo médio) ao criar os programas necessários para realizar um caso de uso² pode até ser útil ao fazer previsões para conjuntos de casos de uso, mas médias aritméticas de um conjunto de valores só costumam ser boas estimativas quando uma quantidade significativa de valores está em jogo. Saber o tempo médio que se gasta para programar um caso de uso não permite prever quanto tempo vai-se levar para programar um dado caso de uso, tomado individualmente. Similarmente, mesmo sabendo que em 1.000 jogadas de uma moeda, aproximadamente 500 serão cara e aproximadamente 500 serão coroa, não há meios de saber qual resultado será obtido em uma jogada específica.

No caso de conjuntos que variam muito, como do tempo que se leva para programar casos de uso, será que não se trataria de considerar que não se tem um único conjunto mas sim um certo número de subconjuntos, cada um dos quais com características distintas? Cada um com sua própria média e variância?

Voltando ao conjunto de maior variância do exemplo anterior, {1, 20, 2, 22}. Observando-se esse conjunto, não seria possível concluir que talvez se trate de dois subconjuntos distintos? Ou seja, {1, 2} e {20, 22}. Nesse caso, o que se tem são dois subconjuntos com médias distintas, e cada um dos quais com variância bem menor do que a do conjunto original.

Seria necessário, então, construir uma teoria para determinar quais as causas de certos elementos apresentarem medidas associa-

2 O conceito de caso de uso é descrito de forma comprehensível por Wazlawick (2004).

das ao primeiro subconjunto ou ao segundo subconjunto, ou seja, qual a causa da existência desses dois subconjuntos.

Voltando ao exemplo dos casos de uso. Não seria mais útil imaginar que seria possível classificar esses casos de uso em, por exemplo, simples, médios e complexos, para obter conjuntos de medidas com variância menor e, portanto, mais previsíveis? E se, por exemplo, os casos de uso simples pudessem ser programados em menos de um dia, os médios entre um dia e uma semana e os complexos em mais de uma semana? Seria mais fácil fazer previsões sobre o tempo que se levaria para desenvolver um sistema. Bastaria contar a quantidade de casos de uso simples, médios e complexos e multiplicar cada quantidade pela média de tempo associada a cada subconjunto. A informação seria mais relevante, possivelmente, do que uma única média aplicada ao conjunto inteiro dos casos de uso, pois com essa abordagem seriam usadas características dos elementos do conjunto para reduzir a incerteza sobre eles.

Porém, essa subclassificação introduz outro fator de incerteza: não se sabe, *a priori*, se a forma de determinar que um caso de uso é simples, médio ou complexo *realmente* classifica os casos de uso em subconjuntos nos quais os valores de tempo de desenvolvimento tenham variância mais baixa.

Nesse ponto poderia ser aplicada a experimentação científica para validar a hipótese de que uma determinada técnica de classificação dos casos de uso efetivamente classifica a complexidade destes adequadamente, ou seja, que esse sistema de classificação colocará no conjunto “simples” os casos de uso que efetivamente sejam mais rápidos de programar, e no conjunto “complexos” os casos de uso mais difíceis, ficando os demais no conjunto “médios”.

Para testar essa hipótese é necessário comparar dois conjuntos de valores: o valor dado a um caso de uso pelo método de classificação e o valor do tempo que efetivamente se leva para programar o caso de uso. Para efetuar essa comparação é necessário usar o conceito de *covariância*, ou seja, determinar em que grau os dois conjuntos variam conjuntamente.

Espera-se, então, que ao se ter as medidas dos tempos, os maiores tempos estejam no conjunto “complexos”, os menores tempos no conjunto “simples” e os demais tempos no conjunto “médios”. Para dar um tratamento totalmente numérico aos conjuntos é possível denotar os conjuntos simples, médios e complexos por números. Alguns trabalhos de referência na área de pontos de caso de uso sugere aplicar valores numéricos 1, 2 e 3 para casos de uso simples, médios e complexos, respectivamente. Nesse caso, o que se busca é uma covariância entre esses valores de pontos de casos de uso e os valores das medidas de tempo obtidas. Pode ser questionado se a escala é correta. Por exemplo, não seria mais adequado ter valores como 1, 2 e 5 em vez de 1, 2 e 3? É uma dúvida válida, mas como a literatura propõe os valores 1, 2 e 3 e não é objetivo deste exemplo questionar esse aspecto da teoria, então trabalha-se com esses valores mesmo e não com outros (se esse exemplo fosse um projeto de pesquisa, o fato de que não será considerada a hipótese de a proporção ser diferente poderia ser uma *limitação* do trabalho).

Supondo que o pesquisador queira saber se um determinado método de estimativa realmente classifica os casos de uso adequadamente, ele poderia fazer experimentos da seguinte forma: (1) tomar um conjunto de casos de uso aleatoriamente escolhidos cujo tempo de programação já seja conhecido (neste caso, trata-se de um *benchmark*; se o tempo não for conhecido, o pesquisador terá de efetuar ele mesmo os cálculos, possivelmente solicitando a um ou mais programadores que desenvolvam os casos de uso para verificar quanto tempo levam); (2) aplicar o método para classificar cada caso de uso com os valores 1, 2 ou 3 e verificar se existe covariância entre os dois conjuntos de valores obtidos. Supondo que os valores obtidos pelo método de classificação para um conjunto de 10 casos de uso sejam os que estão apresentados na Tabela 3.1, pode-se verificar intuitivamente que para os valores mais altos de tempo tem-se a classificação com 3 pontos e os valores mais baixos

de tempo estão classificados com 1 ponto, portanto, intuitivamente parece haver covariância.

Tabela 3.1: Exemplo de covariância alta e direta

Caso de Uso	Tempo conhecido (horas)	Pontos de caso de uso
UC1	1	1
UC2	18	2
UC3	4	1
UC4	67	3
UC5	22	2
UC6	12	2
UC7	2	1
UC8	7	1
UC9	18	2
UC10	55	3

A covariância é considerada alta porque efetivamente os menores valores da coluna “tempo” coincidem com os menores valores da coluna “Pontos de caso de uso”, e os maiores valores da coluna “tempo” coincidem com os maiores valores da coluna “Pontos de caso de uso”. A variância também é considerada direta ou positiva porque quanto maior o valor da coluna “tempo”, maior o valor da coluna “Pontos de caso de uso”.

Em outra situação, poderia ser observada uma relação inversa, ou seja, quanto maior o valor numa coluna, menor na outra. Se fosse assim, então a covariância seria negativa.

Também seria possível que não houvesse nenhuma ou pouca relação entre os tempos conhecidos e o resultado do método de classificação. Um método de classificação totalmente arbitrário não poderia, a princípio, gerar alta covariância, como, por exemplo, um método que atribua aos casos de uso valores sequenciais 1, 2 e 3 em função da ordem em que estes aparecem. A Tabela 3.2 apresenta uma comparação entre os tempos conhecidos e os resultados desse método arbitrário.

Tabela 3.2: Exemplo de baixa covariância

Caso de uso	Tempo conhecido (horas)	Pontos de caso de uso
UC1	1	1
UC2	18	2
UC3	4	3
UC4	67	1
UC5	22	2
UC6	12	3
UC7	2	1
UC8	7	2
UC9	18	3
UC10	55	1

Intuitivamente é possível perceber que não existe uma relação bem definida entre os valores altos da coluna “tempo” e da coluna “pontos” nem entre os valores baixos da coluna “tempo” e os da coluna “pontos”.

A questão agora consiste em como tratar numericamente a covariância para que se possa, por exemplo, comparar diferentes métodos de classificação de casos de uso, decidindo quais são os melhores.

Numericamente, então, a covariância pode ser calculada a partir dos desvios da média em cada conjunto. Se no primeiro conjunto um valor se desvia muito para cima da sua média, espera-se que no segundo conjunto o valor correspondente também se desvie muito para cima da sua própria média. Por outro lado, se ocorre o inverso, ou seja, se o desvio no segundo conjunto ocorre para baixo da média, então há uma relação negativa.

No final, se todos os desvios para cima ou para baixo são semelhantes nos dois conjuntos, haverá covariância alta e positiva. Se os desvios forem sempre invertidos, ou seja, para cima em um conjunto e para baixo no segundo e vice-versa, então a covariância também será alta, mas negativa. Finalmente, se em alguns casos os desvios

coincidem e em outros não, têm-se valores positivos e negativos, levando a uma covariância próxima de zero e, portanto, baixa.

Assim, para calcular a covariância costuma-se usar o produto das diferenças de cada elemento em relação à média do conjunto ao qual ele pertence. Seja $X = \{x_1, x_2, \dots, x_n\}$ o primeiro conjunto e $Y = \{y_1, y_2, \dots, y_n\}$ o segundo conjunto, ambos de mesmo tamanho, pois só é possível calcular covariância se para cada elemento de X corresponde um elemento de Y , então a covariância é calculada em função dos produtos: $(x_i - \bar{x})(y_i - \bar{y})$.

Dessa forma, cada vez que os elementos correspondentes nos dois conjuntos se desviam conjuntamente para cima da média ou para baixo da média, o resultado será positivo. Se um se desvia para cima e outro para baixo, o resultado será negativo. Quanto mais os dois elementos se desviam da média, maior o valor absoluto do produto.

A covariância pode então ser calculada simplesmente como o somatório desses produtos, dividido pelo número de elementos do conjunto de valores menos um, ou seja:

$$s_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

Onde s_{XY} é o símbolo usual para representar a covariância entre os conjuntos X e Y .

Agora, conhecendo a definição matemática da covariância pode-se aplicar esse conceito para identificar entre os dois métodos qual é mais adequado para fazer estimativas de tempo de desenvolvimento de casos de uso. A Tabela 3.3 apresenta o cálculo da covariância para o primeiro método. A média dos tempos conhecidos, \bar{x} , é 20,6 e a média dos valores de classificação, \bar{y} , é 1,8.

Tabela 3.3: Covariância para o primeiro método

Caso de uso	Tempo conhecido (horas): X	Pontos de caso de uso: Y	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$
UC1	1	1	-19,6	-0,8	15,68
UC2	18	2	-2,6	0,2	-0,52
UC3	4	1	-16,6	-0,8	13,28
UC4	67	3	46,4	1,2	55,68
UC5	22	2	1,4	0,2	0,28
UC6	12	2	-8,6	0,2	-1,72
UC7	2	1	-18,6	-0,8	14,88
UC8	7	1	-13,6	-0,8	10,88
UC9	18	2	-2,6	0,2	-0,52
UC10	55	3	34,4	1,2	41,28

Então, o valor da covariância entre os dois conjuntos consiste na somatória da última coluna da Tabela 3.3 dividido por $n-1$, ou seja: $149,2/9 = 16,57777....$ Mas o que significa esse número? A partir dele é possível concluir que os valores de um conjunto são afetados pelos valores do outro conjunto? A covariância não produz um número normalizado, como será explicado em seguida. Então, o valor da covariância serve basicamente para comparar um par de conjuntos com outro par de conjuntos, desde que as unidades de medida dos dois pares sejam idênticas.

Seguindo esse raciocínio, é de se esperar que o método arbitrário, apresentado na Tabela 3.2, designado agora como Y' apresente um valor de covariância bem mais baixo do que o primeiro método, designado na Tabela 3.3 como Y . A Tabela 3.4 apresenta o cálculo da covariância para o método arbitrário, levando em conta que a média dos valores obtidos pelo método arbitrário é 1,9.

Tabela 3.4: Covariância para o método arbitrário

Caso de uso	Tempo conhecido (horas): X	Pontos de caso de uso: Y'	$(x_i - \bar{x})$	$(y'_i - \bar{y}')$	$(x_i - \bar{x})(y'_i - \bar{y}')$
UC1	1	1	-19,6	-0,9	17,64
UC2	18	2	-2,6	0,1	-0,26
UC3	4	3	-16,6	1,1	-18,26
UC4	67	1	46,4	-0,9	-41,76
UC5	22	2	1,4	0,1	0,14
UC6	12	3	-8,6	1,1	-9,46
UC7	2	1	-18,6	-0,9	16,74
UC8	7	2	-13,6	0,1	-1,36
UC9	18	3	-2,6	1,1	-2,86
UC10	55	1	34,4	-0,9	-30,96

Assim, a covariância entre esses dois conjuntos é a somatória da última coluna da Tabela 3.4 dividido por 9, ou seja, $-70,4/9 = -7,822222\dots$. O sinal negativo indica que se existe eventualmente alguma covariância, esta é negativa, mas o valor absoluto obtido para o método arbitrário é menos do que a metade do valor obtido para o primeiro método. Ou seja, a covariância do primeiro método em relação aos valores de tempo é maior do que a covariância do método arbitrário.

3.4.8. Correlação

Como mencionado anteriormente, o valor absoluto da covariância não diz muita coisa sobre um conjunto estar influenciando o outro ou não. Por esse motivo usa-se mais frequentemente a *correlação*. Esta tem basicamente o mesmo significado intuitivo da covariância, mas o resultado numérico varia entre -1 e 1 , em que perto de -1 significa correlação negativa muito forte, perto de 1 significa correlação positiva muito forte e perto de 0 significa ausência de correlação.

Existem vários tipos de cálculo de correlação possíveis. O mais conhecido é o coeficiente de correlação de Pearson, obtido dividindo-se o resultado de $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$, que (dividido por n-1) é usado para obter a covariância, pelo produto do desvio-padrão de cada um dos dois conjuntos de valores, ou seja:

$$r_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_X s_Y}$$

Retornando ao exemplo anterior, pode-se calcular a correlação do primeiro método da seguinte forma:

- a) Já se tem o valor de $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$, que é 149,2.
- b) Calcula-se o desvio-padrão do conjunto dos tempos como a raiz quadrada da variância desse conjunto: $s_X = 67,94409467$.
- c) Calcula-se o desvio-padrão do conjunto dos resultados do primeiro método como a raiz quadrada da variância desse conjunto: $s_Y = 2,366431913$.
- d) Calcula-se o coeficiente de correlação de acordo com a fórmula anterior: $r_{XY} = 149,2 / (67,94409467 * 2,3661913) = 0,928041193$.

Para o método arbitrário tem-se:

- a) O valor de $\sum_{i=1}^n (x_i - \bar{x})(y'_i - \bar{y'})$ é -70,4.
- b) Desvio-padrão do conjunto dos tempos é o mesmo: $s_X = 67,94409467$.
- c) Desvio-padrão dos resultados do método arbitrário: $s_Y = 2,626785107$.
- d) Coeficiente de correlação é, portanto, $r_{XY} = -70,4 / (67,94409467 * 2,626785107) = -0,39445403$.

Pode-se tirar duas conclusões desses valores. Em primeiro lugar, o primeiro método tem um índice de correlação com os tempos

conhecidos bastante próximo a 1 e, portanto, parece ser um método de boa qualidade. Já o método arbitrário tem uma correlação negativa e mais próxima de 0 do que de 1, não podendo ser considerado, portanto, como um bom método para estimativa de tempo.

A intuição, porém, diz que a correlação do método arbitrário deveria ser ainda mais próxima de zero. O método arbitrário atribui valores sequenciais que nenhuma relação tem com os tempos dos casos de uso. Por que não é assim? É que o conjunto de valores examinado é pequeno. Pode-se observar que em alguns casos há até coincidência entre o primeiro método e o método arbitrário. Essas coincidências é que podem levar a crer que o método funcione. Levando a um extremo, se em vez de usar os 10 valores que foram usados para cada método se tivesse trabalhado apenas com os dois primeiros valores, os dois métodos dariam exatamente o mesmo resultado e seriam considerados equivalentes. Mas isso é apenas uma coincidência devido ao pequeno número de situações analisadas.

Um problema a ser considerado, portanto, quando se faz esse tipo de experimento para encontrar correlação entre dois conjuntos de valores é saber se o tamanho da amostra é suficiente para considerar a correlação como efetivamente significativa ou se o valor foi obtido apenas por coincidência. Felizmente existe um método para testar isso. Como a correlação é um índice que deve convergir para um valor à medida que se aumenta o tamanho do conjunto considerado, então não existe certeza absoluta sobre o índice, a não ser que o conjunto seja infinito. Porém, conjuntos infinitos são impossíveis de trabalhar empiricamente. Trabalhando então sempre com conjuntos finitos, é necessário estabelecer um índice de confiabilidade aceitável. Em geral, em estatística, trabalha-se com o índice de 95% de certeza, considerado suficientemente alto para a maioria das aplicações.

O teste para verificar então se um coeficiente de correlação é ou não significativo vai considerar o índice de certeza. Se esse índice for estabelecido em 95%, então, dependendo do tamanho do conjunto examinado, existirá um valor limite para que um coeficiente de cor-

relação seja considerado significativo. O cálculo desses valores foge ao escopo deste livro, mas a Tabela 3.5 apresenta alguns valores que podem ser usados como referência.

Tabela 3.5: Valores mínimos de correlação para ser considerada significativa com 95% de certeza

n	r_{XY} mínimo	n	r_{XY} mínimo	n	r_{XY} mínimo
3	.99692	13	.5529	27	.3809
4	.95000	14	.5324	32	.3494
5	.8783	15	.5139	37	.3246
6	.8114	16	.4973	42	.3044
7	.7545	17	.4821	47	.2875
8	.7067	18	.4683	52	.2732
9	.6664	19	.4555	62	.2500
10	.6319	20	.4438	72	.2319
11	.6021	21	.4329	82	.2172
12	.5760	22	.4227	92	.2050

Considerando agora que os experimentos anteriores foram realizados com 10 casos de uso, o valor mínimo de r_{XY} para que a correlação seja considerada significativa é, pela Tabela 3.5, de 0,6319 no caso de correlação positiva e máximo de -0,6319 no caso de correlação negativa. Então, o primeiro método, com correlação 0,928041193 já pode ser considerado suficientemente analisado para se concluir que efetivamente existe uma forte relação entre o número de pontos de caso de uso que ele produz e o tempo esperado para programar esses casos de uso.

Já o método arbitrário, com correlação de -0,39445403, não atingiu um valor de correlação suficiente para que possa ser considerado que possui correlação negativa com os tempos.

Portanto, o resultado é conclusivo para o primeiro método e inconclusivo para o método arbitrário. Foi demonstrado que o primeiro método atende à expectativa, mas não foi demonstrado nada a respeito do segundo método.

Caso o pesquisador estivesse tentado demonstrar que o método arbitrário funciona, ele teria de aumentar o tamanho do conjunto de valores analisados e calcular um novo índice de correlação, aplicando então outros valores mínimos de acordo com a Tabela 3.5. Porém, o bom-senso diz que, considerando-se a natureza totalmente independente entre os valores dos tempos e os resultados do método arbitrário, por maior que seja o conjunto de valores nunca se chegará a provar qualquer tipo de correlação entre os dois conjuntos. No limite haverá tantos elementos no conjunto que o índice de correlação será praticamente igual a zero.

3.4.9. A Hipótese de Pesquisa

Um aspecto que diferencia o trabalho científico do trabalho técnico é a existência de uma hipótese de pesquisa. A *hipótese* é uma afirmação da qual não se sabe a princípio se é verdadeira ou falsa. O trabalho de pesquisa consiste justamente em tentar provar a veracidade ou falsidade da hipótese.

Um objetivo sem uma boa hipótese pode ser muito arriscado. Anteriormente foi dito que o objetivo consiste em tentar produzir algum conhecimento que ainda não existe. Mas se não houver uma boa hipótese para justificar esse objetivo, corre-se o risco de realizar a pesquisa sem obter resultados. Por exemplo, ter como objetivo de pesquisa provar que $P=NP$ é perfeitamente válido, pois esse problema é relevante à sociedade e o conhecimento necessário para resolver o problema ainda não existe. Porém, com que hipótese o pesquisador vai trabalhar? Se o problema de pesquisa for colocado simplesmente como “provar que $P=NP$ ”, o pesquisador poderá ficar tateando a esmo, e o risco de fracasso será muito grande.

É necessário, portanto, ter uma hipótese.

Segundo Comer (2008), uma tese é uma hipótese ou conjectura. O texto da tese ou monografia é um documento em que o aluno apresenta argumentos a favor de sua tese. Daí a confusão que muitas vezes se faz com o termo “tese”, que pode representar tanto o documento escrito, quanto a hipótese de pesquisa.

O método, como discutido anteriormente, deverá indicar como os testes deverão ser feitos. Ao final dos experimentos haverá evidências a favor ou contra a hipótese inicial. Nesse caso, alguém poderá perguntar, “e se não se conseguir provar que a hipótese era válida?”. A resposta para essa pergunta dependerá de quão relevante era a hipótese original. Uma hipótese qualquer escolhida a esmo, sem nenhum tipo de justificativa, se não for confirmada, não traz nenhuma informação nova para a área de pesquisa. Mas uma hipótese sólida e bem justificada, com evidências de validade, que ao final é invalidada, pode produzir a informação interessante. No pior dos casos provará que aquilo que eventualmente se poderia aceitar intuitivamente como verdadeiro não resistiu à prova. É dessa forma que muitos mitos podem ser derrubados.

Sendo assim, além do objetivo, hipótese e metodologia, é fundamental que o trabalho de pesquisa tenha como base uma boa justificativa para a escolha da hipótese. Uma hipótese bem justificada no início do trabalho aumenta as chances de sucesso. Em primeiro lugar, é mais provável que ela seja verdadeira do que uma hipótese sem justificativa. Em segundo lugar, se ela for falsa, o trabalho terá o mérito de ter derrubado algum mito.

O trabalho científico na área de Computação consiste então em formular uma hipótese e coletar evidências para comprovar a sua validade. Essas evidências podem ser obtidas basicamente de três formas:

- a) Construindo uma teoria, que a partir de fatos aceitos e deduções válidas prove que a hipótese é verdadeira.
- b) Realizando certo número de experimentos controlados, que estatisticamente comprovem a validade da hipótese. Porém, deve-se ter em mente que esse tipo de comprovação é sempre sujeito a erros. Normalmente se aceita que hipóteses sejam comprovadas com 95% de certeza.
- c) Realizar estudos de caso, comparativos, argumentações, colher opiniões através de questionários e outras formas que dificilmente constituem uma prova, mas que podem ser evidências da validade da hipótese.

É possível também realizar um trabalho estruturado sobre combinações de duas ou três formas anteriores. De qualquer maneira, o importante é que a partir da formulação da hipótese, o pesquisador esteja engajado no processo de busca de evidências, e que essas evidências sejam estruturadas em um discurso coerente, em que ele apresenta a sua argumentação para a comunidade científica.

Um problema de pesquisa, então, em geral vai perguntar como duas ou mais variáveis se relacionam, e se existe correlação positiva ou negativa entre os valores das variáveis. A existência dessas correlações, porém, ainda não prova causas. Uma teoria consistente que explique causa e efeito precisa também ser elaborada, além da validação empírica. Isso acontece porque algumas vezes duas variáveis até se correlacionam com alto índice, mas as causas envolvidas podem não ser tão diretas.

Cita-se como exemplo o caso de uma empresa que decidiu verificar se funcionários bem alimentados trabalhavam melhor. A empresa passou a servir um café da manhã saudável para seus funcionários em todas as filiais e, em praticamente todas, a produtividade aumentou. Mas será realmente verdade? Qual a explicação? Será que foi mesmo a alimentação que fez os funcionários trabalharem mais? Na sequência a empresa testou retirar o café da manhã saudável para verificar se os trabalhadores retornavam ao ritmo anterior. Para surpresa de todos, a produtividade cresceu ainda mais.

Como um cientista lidaria com esses dados aparentemente contraditórios? O caso aqui é que se trata de verificar o real motivo do aumento de produtividade. Este não ocorreu por conta de uma alimentação melhor, mas pelo fato de que houve uma significativa modificação no dia a dia da empresa. Funcionários estão acostumados a perceber que modificações na empresa em geral implicam demissões. Por isso, cada vez que se observa uma modificação no ambiente (introdução ou retirada do café da manhã), os funcionários tendem a trabalhar mais para serem notados e garantirem seus empregos.

3.5. Justificativa

Foi comentado anteriormente que uma hipótese de trabalho é muito arriscada se não estiver solidamente apoiada em uma boa justificativa que apresente evidências de que vale a pena investir tempo e recursos na tentativa de comprovar a hipótese. Quem em sã consciência se proporia a trabalhar dois anos para provar que o método arbitrário definido anteriormente faz boas previsões em estimativa de esforço? Uma boa hipótese precisa ser *justificável*.

Em uma monografia, pode-se justificar o *tema* de pesquisa, mas mais importante ainda é justificar a escolha do objetivo e da *hipótese*. Por exemplo, se o tema de pesquisa é “compactação de texto”, o objetivo de pesquisa é obter um algoritmo com maior grau de compactação do que os algoritmos comerciais, e a hipótese de pesquisa pode consistir em utilizar um determinado modelo de rede neural para realizar essa compactação, então a justificativa do *tema* deverá se concentrar em mostrar que é necessário obter algoritmos de compactação melhores. Adicionalmente, a justificativa da *hipótese* deverá se concentrar em apresentar evidências de que o modelo de rede neural escolhido poderá produzir resultados melhores do que os algoritmos comerciais.

Em geral, a justificativa do tema aparece na contextualização do trabalho, em que se tenta justificar ao leitor que o problema escolhido realmente é relevante (no exemplo anterior, compactação de textos). Mas na maior parte das vezes esse convencimento é pacífico. Mais difícil é justificar uma hipótese de trabalho, pois para isso será necessário apresentar alguma evidência de que uma determinada linha de pesquisa pode levar a bons resultados quando ainda não se efetuou essa pesquisa (no exemplo anterior, justificar o uso do modelo específico de redes neurais para compactar textos). Essas evidências podem ser referências a outros trabalhos que eventualmente mostraram algum tipo de resultado que aponte para a viabilidade da hipótese escolhida, ou ainda em dados colhidos preliminarmente pelo próprio autor do trabalho ou em um estudo de caso.

3.6. Resultados Esperados

Em geral, os resultados esperados são situações que o autor de um trabalho espera que ocorram, caso seus objetivos sejam atingidos. Os resultados esperados normalmente fogem ao escopo do trabalho. O autor da pesquisa não tentará obter os resultados esperados ao final da pesquisa. Eles são posteriores.

Isso diferencia os resultados esperados dos objetivos. Os objetivos serão perseguidos pelo autor, e ao final do trabalho ele dirá se foram ou não atingidos. Os resultados esperados possivelmente ocorrerão após a conclusão do trabalho.

Por exemplo, o objetivo do trabalho poderá ser definir um método de cálculo de esforço para desenvolvimento de software mais preciso do que os métodos do estado da arte. O autor da pesquisa deverá ter uma boa hipótese para fundamentar esse objetivo em primeiro lugar. Depois, ele deverá realizar um conjunto de experimentos que, juntamente com uma base teórica, demonstrarão a validade ou não da hipótese.

Esse autor poderá apresentar, inicialmente, como resultados esperados de seu trabalho, a adoção do seu método pela indústria e um melhor desempenho das empresas produtoras de software que venham a utilizar esse método.

Como se vê aqui, é praticamente impossível que o autor obtenha esses resultados esperados durante a realização de sua pesquisa. Mas eles poderão eventualmente ocorrer depois. Também é possível que não ocorram, pois, por outros motivos quaisquer, poderá acontecer que nenhuma empresa venha a adotar o seu método.

Assim, pode-se dizer que os objetivos devem ser verificáveis ao final do trabalho, inclusive os objetivos específicos. Já os resultados esperados são apenas esperanças e não podem necessariamente ser verificados ao final do trabalho.

No início do trabalho de pesquisa, uma forma de se tentar determinar quais são os resultados esperados do trabalho consiste em fazer a pergunta “o que possivelmente mudaria no mundo se eu atingisse os objetivos da minha tese/monografia?”.

3.7. Limitações do Trabalho

Ao contrário do que alunos iniciantes muitas vezes pensam, não é possível resolver todos os problemas da humanidade em dois ou três anos de trabalho (“síndrome de querer mudar o mundo”, ou “síndrome do Prêmio Nobel”).

Um trabalho de pesquisa pode começar muitas vezes com um objetivo demasiadamente amplo e, portanto, inalcançável durante o tempo disponível para a realização do curso. Sendo assim, muitas vezes é necessário realizar cortes nos objetivos, ou limitar a forma de perseguí-los. Em vez de demonstrar que uma hipótese é sempre verdadeira, pode-se optar por demonstrar que ela é verdadeira apenas em determinadas condições, para as quais foi possível realizar testes convincentes. Por exemplo, o método de estimativa de esforço mencionado na seção anterior poderia ser comprovadamente mais preciso apenas para uma determinada classe de sistemas, como, por exemplo, sistemas baseados em Web. O fato de que o método não foi testado com outros tipos de sistemas impõe uma limitação ao trabalho.

As limitações são, portanto, aspectos do trabalho dos quais o autor tem consciência e reconhece a importância, mas não tem condições de abordar no tempo disponível.

É importante, em trabalhos de pesquisa, que as limitações conhecidas sejam claramente identificadas pelo autor desde o início. Isso evitará que o próprio autor muitas vezes se perca em divagações ou buscando aspectos que extrapolam os objetivos iniciais. Isso evita também que o leitor crie expectativas demasiadamente amplas sobre o trabalho, que serão depois frustradas.

Novamente, espera-se que uma boa interação com o orientador ajude o aluno a colocar as devidas limitações nos seus objetivos, para que o trabalho possa ser concluído com sucesso no tempo disponível.

3.8. Discussão

De acordo com o que foi visto neste capítulo, o trabalho de pesquisa deverá estar enquadrado em um tema que, como área de conhecimento, deverá ser plenamente conhecido pelo pesquisador. Dentro do tema, o pesquisador deverá estabelecer um objetivo a ser buscado. Esse objetivo deverá estar baseado em uma hipótese de trabalho, que deve ter uma boa justificativa para ter sido escolhida. O método vai esclarecer como a hipótese será comprovada pelo autor do trabalho, e as limitações deixarão claros quais aspectos não serão abordados.

É compreensível a dificuldade de muitos alunos que ingressam, especialmente no mestrado, em compreender essa estrutura e realizar um trabalho organizado dessa forma. Surgem assim dissertações que muitas vezes são meramente uma apresentação de um sistema, ou uma proposta testada em apenas uma ou duas situações, ou, ainda, dissertações que se concentram em coletar dados e não elaboram adequadamente os conceitos que os dados representam.

Essa dificuldade deve-se, especialmente, ao fato de que pela primeira vez, talvez, em sua vida o aluno será colocado diante de um trabalho individual extenso, em que a sua iniciativa será fundamental para o sucesso. Trabalhos escolares, mesmo na graduação e na especialização, resumem-se, muitas vezes, apenas à pesquisa bibliográfica. O aluno simplesmente coleta material de várias fontes e organiza essa informação de uma maneira pessoal. A estrutura da pesquisa científica, especialmente no mestrado e no doutorado, vai muito além da pesquisa bibliográfica, como se procurou mostrar neste capítulo.

Capítulo 4

Análise Crítica de Propostas de Monografia

No curso da disciplina de Metodologia da Pesquisa no Programa de Pós-Graduação em Ciência da Computação da UFSC, dezenas de alunos ingressantes foram convidados a apresentar um seminário com uma proposta de pesquisa. As apresentações foram utilizadas para ajudar os alunos a perceberem problemas em suas propostas e melhorá-las. Neste capítulo analisaremos alguns textos contidos nessas propostas antes de sua melhoria, ou seja, os textos apresentados são material não revisado pelos alunos. Dessa forma considerou-se apropriado não mencionar os nomes dos autores deste material, sendo colocado entre aspas o texto que é de produção dos alunos e os comentários logo em seguida.

4.1. Análise da Contextualização e Colocação do Problema

Nesta seção são analisados textos referentes à contextualização do problema. A avaliação verifica se o problema é, de fato, um problema de pesquisa e se está bem claro. Os resultados são apresentados a seguir.

“Uma das grandes dificuldades que o acadêmico encontra ao ingressar em uma universidade é confrontar-se com a disciplina Metodologia da Pesquisa. Desconhece a ABNT e suas normas bem como as propostas de formatação e estruturação de um relatório de pesquisa. Para amenizar o problema, pretende-se elaborar um sistema multimídia, utilizando-se de

técnicas referentes ao desenvolvimento de sistemas deste tipo como: ergonomia de software, um método de apoio para aprimorar a interação usuário-máquina e a distribuição dos recursos no sistema.”

Em relação a esse texto, cabem os seguintes comentários: de onde vem a informação sobre “isto ser uma grande dificuldade”? Observação do autor (experimentos)? Ou pesquisa bibliográfica? Quem demonstrou que essa informação é verdadeira?

Mesmo supondo que se trate de um problema real, caberia ainda perguntar por que o aluno considera que um sistema multimídia será uma solução. Não existem outras possibilidades? Aqui parece que já se escolheu a ferramenta sem analisar o problema em si detalhadamente.

Deve-se considerar em casos como esse que a observação do problema pode estar localizada, como no caso do aluno que observou que o rio que cruzava a cidade onde ele morava não podia ser atravessado. O problema referido pode então ser consequência das técnicas de ensino de um professor em particular, que o aluno tenha observado e não se tratar de problema generalizado que afeta todas as instituições de ensino.

“No Departamento de Tecnologia da Informação da *** não há um método específico para gerenciar projetos de desenvolvimento de software terceirizado. Projetos terceirizados de desenvolvimento de software têm sido entregues com atrasos e com os requisitos não sendo plenamente atendidos.”

O problema refere-se a uma empresa específica. Se essa empresa não usa um método para gerenciar projetos, então a solução é implantar um método existente que já tenha sido testado, ou seja, trata-se de um trabalho técnico.

Esse problema, da forma como está colocado, não justifica uma proposta de pesquisa, como, por exemplo, neste caso, em que será proposto o desenvolvimento de um método para gerenciamento de terceirizados.

“O sistema *** está constituído por 14 instituições (...) de ensino superior. Não houve um aumento no número destas instituições, mas as instituições estão se expandindo. O número de alunos matriculados vem aumentando ano a ano. As informações serão retiradas do questionário socioeconômico e da secretaria acadêmica que formam uma grande base de dados, possibilitando a utilização de *data mining* nesta base de dados.”

Aqui existe uma oportunidade, não um problema. A oportunidade é a existência de dados em abundância, portanto, propõe-se fazer *data mining* nesses dados. Mas falta informar o que se pretende descobrir ao analisar os dados ou pelo menos que pista se está seguindo. Como dizem, “se você não sabe para onde quer ir, qualquer estrada serve”.

“Octrees são usadas para representação espacial de cenários e objetos. São empregadas por possibilitar a rápida determinação de partes não visíveis. Necesitam de precisão infinita para representar curvas (distorcem a realidade).”

Aqui temos um bom exemplo de contextualização. A mensagem é rápida, mas coloca um problema de forma clara. Mesmo que o leitor não saiba o que são *octrees*, o autor informa que elas são usadas para representação espacial de cenários e objetos. Além disso, o autor informa que elas sofrem de um problema sério: necessitam de precisão infinita.

“Serviços distribuídos têm sido usados para obter transparência, desempenho e confiabilidade em sistemas. Dentre os problemas encontrados, está a dificuldade em se obter alta confiabilidade com perda mínima de desempenho.”

Aqui o aluno apresenta uma clara relação de custo-benefício: é difícil melhorar confiabilidade sem perder desempenho. A afirmação inicial, porém, é um tanto genérica e poderia ser mais bem detalhada.

“Aplicação do Raciocínio Baseado em Casos (RBC) na Previsão Meteorológica. Utilização de casos passados (tempo) na busca de uma possível solução (previsão).”

Aqui não foi identificado um problema. Previsão meteorológica é o *tema*. Aparentemente o aluno já escolheu a ferramenta (RBC) e agora quer ver como ela funciona. Mas não especificou o problema nem justificou a escolha da ferramenta.

“Um sistema elétrico é composto por vários componentes que necessitam ser desligados para manutenção durante certo período: geradores, linha de transmissão etc. O funcionamento do sistema elétrico pode ser representado por equações que representam as limitações físicas dos componentes. O problema é aplicar a maximização da carga máxima atendida em vários cenários. O sistema deve considerar o modelo estocástico relacionado aos reservatórios.”

Aqui parece que o aluno confunde o problema com o objetivo. Maximizar a carga atendida parece ser o objetivo desejado, mas qual o problema de fato? Não existem métodos matemáticos ou sistemas para fazer essa maximização? Se existem, eles sofrem de quais limitações? Essas limitações, que possivelmente existem, é que consistiriam no verdadeiro problema de pesquisa.

“A reflexão computacional é um mecanismo amplamente reconhecido para adaptar e reconfigurar software em tempo de execução.”

Observe como o advérbio “amplamente” pode ser dispensado da frase sem lhe prejudicar o sentido. De resto, trata-se de uma informação, meramente, não de um problema.

“Reconstrução 3D é crítica em vários domínios de aplicação e é pesquisada hoje quanto à exatidão (acurácia), precisão (completude e nível de detalhe) e desempenho (velocidade), em que o problema é encontrar a geometria 3D correspondente a partes da cena observada ou a toda ela. No atual estado de desenvolvimento da área, não há uma teoria ge-

ral que unifique o problema. Portanto, estudos são feitos em contextos específicos e aplicando-se restrições. Desse modo, existem abordagens que usam múltiplas visões da cena, algum conhecimento prévio do objeto, indexação por aspectos da imagem etc. A união e melhoramento de algumas dessas abordagens é um campo promissor de pesquisa.”

A contextualização limita bem o problema. A afirmação de que não há uma teoria geral da área deve ser justificada por uma bibliografia que faça essa afirmação ou por um estudo do próprio autor, em que, ao comparar diversos trabalhos, chega à conclusão de que tal teoria é efetivamente inexistente. O aluno indica claramente um campo de pesquisa, mas não necessariamente ainda um problema de pesquisa. Como a área é caracterizada por várias abordagens que se complementam, é possível tentar realizar junções e aprimoramentos entre as diversas teorias. Faltaria, possivelmente, definir quais seriam os ganhos dessas abordagens. Eficiência? Eficácia? Ou apenas uma organização melhor da área de conhecimento?

“GIS (Geographical Information System). A integração dos dados geográficos e alfanuméricos ainda é um grande desafio. Os dados manipulados comumente fazem parte de um sistema maior. A interoperabilidade é fundamental.”

O texto originalmente estava estruturado em tópicos. Aqui é possível perceber claramente um problema; interoperabilidade entre sistemas de base de dados textuais e sistemas de dados geográficos. Resta caracterizar isso como um problema de pesquisa e não como problema técnico. Da forma como está, parece ser apenas um desafio técnico dar solução a esse caso.

“Grande parte das MPEs não possuem um processo definido. Limitações de recursos humanos e financeiros para adoção de processos densos ou serviços de consultoria. Os processos ágeis prometem simplicidade e desempenho para pequenas equipes, porém são escassas a evidências empíricas de aplicações em

MPEs. Os processos ágeis não evidenciam de forma clara todos os ciclos de seu processo.”

A primeira afirmação sobre micro e pequenas empresas (MPE) precisaria de uma fundamentação bibliográfica ou empírica. A afirmação sobre escassez de evidências empíricas de aplicações de processos ágeis em MPE não implica necessariamente um problema de pesquisa. Seria necessário ainda identificar alguma característica especial dessas MPE que as tornassem diferentes em outras empresas em relação à adoção de métodos ágeis.

“XML tem sido utilizada largamente no intercâmbio eletrônico de dados (EDI), o que tem aumentado o interesse na manipulação de dados XML persistentes. Pesquisas desenvolvidas em BD's XML nativos têm procurado atingir o ‘estado da arte’ na gerência de dados semiestruturados.”

O alegado aumento de interesse em uma área (que precisaria ser justificado por evidências bibliográficas recentes) não implica existência de um problema de pesquisa, mas apenas de um tema de pesquisa.

“Um dos grandes problemas na distribuição de vídeos para dispositivos móveis é a sua baixa disponibilidade. Usando um tempo menor na transferência de vídeos se tem um melhor aproveitamento da banda de conexão e uma economia no consumo das baterias, que possuem capacidade limitada. O Grid é uma versão segura de Web Services para o compartilhamento de dados e recursos (Foster, 2001). Entidades como a OGSI e WSRF definem as especificações para os *tools kit* em que temos controle de usuário, segurança e interface aberta para garantir a compatibilidade. Dessa forma podemos propor uma malha de servidores de vídeo em que um aplicativo cliente pode copiar fragmentos de vídeo de diversos pontos fornecendo uma alta disponibilidade dos recursos e diminuir a sobrecarga existente num sistema centralizado.”

Existe um problema claramente caracterizado. Porém, não fica claro se a solução proposta ainda não foi tentada. Se foi, quais os problemas atuais com as soluções existentes para que uma nova tecnologia seja proposta?

“Alguns processos de integração de fontes de dados XML heterogêneas definem um esquema conceitual único (esquema global) representativo do esquema de todas essas fontes. Através desse esquema global é possível realizar consultas de forma transparente, uma vez que ele abstrai as complexidades inerentes às fontes de dados, como diferenças estruturais e semânticas. Trabalhos como [***] tratam esse problema.”

Embora o texto assim mencione, ele não apresenta um problema, mas uma solução. O texto referenciado aparentemente apresenta soluções para a combinação de bases XML heterogêneas. Então, resta ao autor da proposta de pesquisa identificar um problema, possivelmente no trabalho referenciado, que merecesse atenção para ser resolvido. Pela simples leitura do texto anterior não é possível saber qual é o problema que será resolvido na monografia.

“Em supermercados, por exemplo, existe a necessidade de prever a quantidade de produtos que serão vendidos em uma faixa de tempo, para a otimização da quantidade comprada, enxugando custos e necessidades logísticas. Por isso foi estudada a previsão de séries temporais, a qual possui modelos estatísticos que vêm sendo utilizados no mercado. Estes, porém, ora não contemplam perfeitamente os objetivos, ora exigem muitos recursos para tal, portanto há a necessidade de se encontrar um modelo que atenda mais apropriadamente este caso. Alguns estudos em redes neurais sugerem que há a possibilidade de que seja encontrado este modelo neste campo de pesquisa.”

O texto coloca um problema real, menciona que as soluções existentes têm limitações e se propõe a pesquisar uma forma de superá-las usando uma ferramenta que, segundo o autor, a literatura apresenta como uma tendência. Porém, se redes neurais já têm sido usadas para esse fim, será absolutamente necessário um trabalho de comparação entre os resultados de modelos existentes e o modelo definido pelo autor.

“A garantia de QoS é uma preocupação para algumas aplicações distribuídas, as quais necessitam desta para poder funcionar corretamente. Assim, torna-se necessário adotar uma técnica que melhor se encaixe na aplicação de forma a prover os requisitos necessários podendo com isso superar os problemas que poderão surgir, como, por exemplo, atraso (*delay*), extravio de pacotes, entrega desordenada e prioridades, ou seja, requisitos temporais de qualidade e de segurança [***].”

O problema é fundamentado corretamente em uma citação bibliográfica. Esta aponta no sentido da melhoria de vários aspectos relacionados ao tema.

“WebServices têm sido amplamente utilizados com intuito de prover interoperabilidade entre aplicações. Algumas aplicações necessitam que os serviços acessados estejam altamente disponíveis, ou seja, ativos e em funcionamento o máximo de tempo possível.”

Trata-se aqui da descrição de uma necessidade. Para que se torne um problema de pesquisa é necessário indicar ainda como essa necessidade tem sido suprida e quais as limitações das técnicas correntes.

“Há um crescente número de máquinas conectadas à internet. Máquinas de capacidade cada vez maior. Estudos comprovaram que 93% da capacidade de processamento dos computadores permanece ociosa. Essa capacidade pode ser utilizada para proces-

samento de projetos que requerem alto desempenho computacional a um baixo custo operacional. O Ambiente *** é uma alternativa para isto.”

Aqui, a informação apresentada, ociosidade de máquinas na Internet, é quantificada e embasada em um estudo bibliográfico (embora não seja explicitamente citado aqui). Porém, esse texto não apresenta um problema de pesquisa, mas uma *oportunidade*. Máquinas ociosas são uma oportunidade para aproveitar seu potencial de processamento, mas falta estabelecer qual o novo conhecimento que se busca gerar no trabalho.

“Jurisprudência (do Latim: *juris prudentia*) é um termo jurídico com diversos significados. O mais comum refere-se à aplicação de estudo de casos jurídicos na tomada de decisões judiciais. Assim, ‘jurisprudência’ pode referir-se a ‘lei baseada em casos’, ou as decisões legais que se desenvolveram e que acompanham estatutos na aplicação de leis em situações de fato. Grande parte do conhecimento jurídico está formatado em documentos que contêm decisões proferidas em sentenças ou acórdãos. A pesquisa visa contribuir para definir um modelo de arquitetura de pesquisa em documentos com significado semântico incorporado. Criando uma conexão entre a linguagem técnica jurídica e sua representação utilizando ontologias aptas para seu processamento em ferramentas que auxiliem no refinamento do conhecimento.”

Essa contextualização apresenta um tema bem delimitado, mas falha ao não caracterizar o problema de pesquisa. A proposta é definir um modelo de arquitetura de pesquisa, mas não estabelece quais são os problemas com as eventuais arquiteturas de pesquisa existentes. Mesmo a menção de uma aplicação direta no domínio jurídico não permite concluir pelo ineditismo desse trabalho, pois esse tipo de aplicação possivelmente já existia no momento da elaboração dessa proposta. Mesmo que não existisse, deveria ser deixado

claro o que esse domínio tem de diferente em relação a outros, e por que não funcionariam outros sistemas anteriores a este que será proposto.

“A atividade de exploração de poços de petróleo envolve um grau de incerteza muito grande, e essa incerteza gera riscos elevados em termos de custos. Para fazer essa Análise de Risco, a técnica de Simulação vem sendo utilizada, necessitando-se conhecer as distribuições dos valores de tempo das operações. Para a descoberta dessas distribuições, existem vários métodos que avaliam a aderência a algum modelo de probabilidade. Mas para isso é necessário que exista uma série histórica dos tempos de operações semelhantes. Quando não há esse histórico, como realizar a descoberta dos modelos de probabilidade?”

Aqui a contextualização apresenta um problema de fato: para realizar uma simulação é necessário ter uma série histórica. É colocado o problema de como fazer essa simulação quando uma série não está disponível, o que de fato pode acontecer.

4.2. Análise de Objetivo Geral

Nesta seção procura-se avaliar se um objetivo geral está bem escrito, se é um objetivo de pesquisa e se deixa claro que pode ser verificado ao final do trabalho. Atenção especial é dada ao verbo que apresenta o objetivo. Analisa-se também se o objetivo apresentado define claramente uma pesquisa científica ou um objetivo tecnológico, como, por exemplo, a implementação de um sistema. Os resultados estão apresentados a seguir.

“Elaborar um sistema Hipermídia observando critérios ergonômicos e princípios de design como apoio ao ensino da Metodologia Científica.”

O verbo inicial não aponta para um objetivo de pesquisa, mas um objetivo técnico: elaborar um sistema. Critérios ergonômicos e

princípios de design também são requisitos que qualquer projeto técnico deveria observar.

“Obter confiabilidade através de replicação com 5 servidores, obtendo perda máxima de 20% de desempenho comparado a um sistema sem replicação.”

O objetivo é claro: “obter confiabilidade”. Falta especificar, porém, como se mede confiabilidade para que se possa concluir ao final o trabalho se ela foi obtida. A comparação proposta com um sistema sem replicação implicaria que a técnica em questão ainda não foi explorada em outros trabalhos.

“O objetivo geral deste trabalho é desenvolver, implantar e avaliar o impacto de uma abordagem para a gerência de projetos terceirizados de desenvolvimento de software, em uma empresa transmissora de energia elétrica.”

Os dois primeiros verbos, desenvolver e implantar, são objetivos técnicos e não de pesquisa. O terceiro verbo “avaliar” aponta para um objetivo de pesquisa, mas “impacto” é algo muito subjetivo para ser avaliado. Deveria haver uma melhor especificação sobre o que se pretende descobrir, caso contrário, pode-se estar procurando à toa.

“Estudar as técnicas de IA, mais especificamente o RBC. Demonstrar a eficiência da utilização do RBC na previsão meteorológica. Desenvolvimento de um protótipo para demonstração de resultados.”

A primeira frase não pode ser um objetivo de pesquisa. Pelo menos não de pesquisa científica. O aluno vai, sim, estudar técnicas de IA, mas ele faz isso para suprir conhecimentos que ele próprio não possui ainda. Esses conhecimentos são, porém, de domínio público. Então, esse estudo não produz conhecimento novo e não pode ser apresentado como objetivo de pesquisa.

A segunda frase já está melhor. Algo será demonstrado. Porém, não está muito claro onde se pretende chegar com esse objetivo. Como medir a eficiência? Ela deve ser melhor do que alguma outra técnica?

O desenvolvimento do protótipo, apresentado na terceira frase, é parte do método de trabalho, e o protótipo pode ser um subproduto interessante, mas não parece ser um objetivo de pesquisa de fato.

“O objetivo é desenvolver um método de refinamento da reconstrução 3D, através de uma memória adaptativa dos objetos já reconstruídos, de forma que novos objetos analisados são armazenados e aqueles já conhecidos são refinados. Assim, (hipótese) poderá haver uma maior completude e um aumento da acurácia da reconstrução, dada a natureza cumulativa do processo. Trata-se de uma memória com representação geométrica implícita, mediante indexação/mapeamento de objetos a partir de aspectos da imagem, como contornos etc.”

Desenvolver um método pode ser um bom objetivo de pesquisa, embora algumas vezes os alunos confundam método, processo, procedimento, mecanismo, arquitetura, sistema etc. Nesse caso, de fato trata-se de um método envolvendo técnicas de Computação Gráfica. Há uma hipótese clara, que consiste em uma técnica a ser usada: memória adaptativa. Há um objetivo claro associado a essa hipótese que consiste em obter maior completude e acurácia na reconstrução 3D. Para que o objetivo esteja completo, faltaria apenas indicar quanto os métodos atuais conseguem obter em termo de acurácia e completude para que se possa saber, ao final do projeto, se a técnica tentada melhorou esses valores.

“Identificar os aspectos mais relevantes que influenciam na performance de sistemas reflexivos.”

“Identificar” algo pode ser um objetivo de pesquisa desde que esteja bem claro o que se procura identificar. Não é o caso desse texto, pois ele tem várias palavras com conotação subjetiva. Por exemplo, como se avalia que um aspecto é mais *relevante* do que outro? O que significa *influenciar* a performance? A presença de um vírus no computador, por exemplo, é um aspecto que deveria ser conside-

rado? Ele pode afetar a *performance* de um sistema reflexivo assim como outros sistemas. Dessa forma, a quantidade de aspectos pode ser arbitrariamente grande. Além disso, apenas identificar algo pode ser insuficiente para que se tenha um resultado concreto e útil. O objetivo poderia ser complementado com a proposição de alguma técnica para superar os problemas identificados.

“Classificar os alunos quanto ao seu desempenho acadêmico. Classificar os alunos quanto a sua carência financeira. Comparar o desempenho escolar dos egressos pelo SAEM e pelo vestibular. Predizer o desempenho escolar dos estudantes na universidade. Predizer casos de mudança de instituição. Auxiliar na tomada de decisões.”

As três primeiras frases estão bem conectadas e apontam para um objetivo de pesquisa claro. Porém, esse objetivo não parece ser um objetivo da área de Ciência da Computação. Possivelmente o aluno tencionava utilizar sistemas informatizados para fazer a comparação, mas isso não é suficiente. Deve haver uma produção de conhecimento útil para a área de Computação. Caso contrário, o mestrado pode e deve ser tentado em outra área. A terceira e quarta frases parecem não seguir facilmente as três primeiras. Não fica claro como, a partir do desempenho dos alunos e da sua situação financeira, será possível predizer o seu desempenho escolar e o que aconteceria se mudassem de instituição. A última frase está mais para um resultado esperado do que um objetivo de monografia. Caso o objetivo colocado nas três primeiras frases fosse obtido, isso possivelmente teria consequências no processo de tomada de decisão dos administradores das instituições de ensino, mas como não é algo que o autor da monografia vá fazer não pode ser considerado como objetivo do trabalho.

“Desenvolver um algoritmo baseado no relaxamento de variáveis Lagrangeanas para determinar a escala ótima de manutenção de um sistema elétrico.”

Um algoritmo baseado em relaxamento das variáveis Lagrangeanas parece uma boa solução para otimizar alguma coisa como, por exemplo, a manutenção de um sistema elétrico. Mas, por que não outras técnicas? Existe alguma técnica sendo usada correntemente? Que resultados ela produz? Quanto se pretende melhorar e em que dimensão?

“Gerar aproximação de superfície em nodos externos da Octree.”

Um objetivo sucinto. Talvez por isso seja difícil determinar se está adequado como objetivo de pesquisa. A princípio, “gerar” uma aproximação de superfície não traz novo conhecimento à tona. É uma atividade, não um objetivo de pesquisa. Poderia ser reescrito talvez como, “demonstrar que é possível gerar aproximação de superfície...”. Isso se até o momento da pesquisa nunca tivesse sido feito. Caso contrário, seria reinventar a roda.

“Desenvolver um sistema de gerenciamento hospitalar que, com base na tecnologia de *workflow*, permita a modelagem e execução de processos médicos, interagindo com servidores de dados no padrão ***.”

Desenvolver um sistema, por melhor que seja, nunca será um objetivo de pesquisa. O sistema pode ser usado para demonstrar algo, mas ele em si é um objetivo técnico.

“Oferecer solução para o planejamento de rotas de navegação para um ambiente semiconhecido. Aplicar essa solução a um sistema de planejamento neurocirúrgico. Demonstrar que essa solução está correta através da comparação com outros métodos atualmente utilizados.(manual!)”

“Oferecer solução” está bom como resultado esperado, mas existem muitas formas de fazer isso. Pode-se comprar um sistema, pode-se implementar um algoritmo, pode-se mesmo criar toda uma nova teoria e prática nessa área, com vistas a fornecer uma solução.

Então, isso teria de ser mais bem definido. A aplicação dessa solução numa determinada área pode ser interessante, especialmente se os métodos correntes utilizados são manuais. Deve haver uma referência bibliográfica ou a apresentação de uma pesquisa que demonstre que de fato a técnica corrente é manual. Uma rápida observação no hospital local não seria suficiente como justificativa.

“Propor um modelo para identificação de características para um plano amostral estratificado.”

“Propor” é um verbo que é usado em muitas monografias como objetivo, mas é um verbo perigoso se não vier acompanhado de um objeto direto relevante. No caso anterior, o que está sendo proposto é um modelo para identificação de características. Ok, o modelo pode ser proposto. Mas é necessário indicar claramente o que sucede com outros modelos que possivelmente devem existir. Se não existirem, deve-se avaliar por que não existem e referenciar os modelos mais parecidos com o que se pretende propor.

Em suma, não se pode chegar e simplesmente *propor* alguma coisa. É preciso apresentar uma boa justificativa para tal proposta.

“Portar o Sistema *** para a arquitetura de RSSF da UC Berkeley.”

A não ser que o processo de “portar” apresente desafios que impliquem construção de novos conhecimentos em computação, o trabalho, como proposto, é apenas um objetivo técnico que poderia ser executado por um profissional da área, sem que ele fosse receber ao final disso o título de mestre.

“O objetivo deste trabalho é apresentar um mecanismo sistemático que, através dos contratos definidos em notação OCL, gere diagramas de colaboração, aplicando *design patterns*.”

Seria necessário conhecer bem a área para saber que o mecanismo proposto não existe, portanto, o aluno deveria ter dito isso claramente na contextualização.

4.3. Análise de Objetivos Específicos

Os objetivos específicos devem refletir subprodutos ou um detalhamento do objetivo principal. Não se deve, a princípio, mencionar como objetivo específico passos que são meramente intermediários para atingir o objetivo geral. Vários outros cuidados também devem ser tomados, conforme os comentários nos exemplos a seguir.

- “a) Utilizar critérios de ergonomia de software;
- b) Aplicar princípios de design no desenvolvimento de sistemas multimídia;
- c) Elaborar uma interface amigável em sistema multimídia, ou seja, fácil de usar, aplicar e comunicar;
- d) Adequar o sistema a um método de apoio para sistemas de autoria.”

Todo objetivo deve ter um verbo, mas ele deve indicar a produção de algum conhecimento novo. Isso vale para os objetivos específicos também. O objetivo (a) é apresentado com o verbo “utiliza”, que não se presta bem a um objetivo de pesquisa, pois o fato de o autor estar utilizando algo não quer dizer que esteja produzindo algum tipo de conhecimento novo. O mesmo vale para o objetivo (b). O objetivo (c) realmente implica algo sendo feito, possivelmente algo novo. Porém, é difícil mensurar o que significa ser amigável e fácil de usar. Algum critério de medição dessas características deveria ser mencionado. O objetivo (d) não parece ser de pesquisa, mas tecnológico: existe um sistema que será adaptado.

- “a) Técnicas para identificar padrões em característica;
- b) Técnicas de agrupamento;
- c) Categorização de variáveis contínuas;
- d) Discretização.”

Para que esses itens correspondam a objetivos específicos deveria inicialmente haver um verbo em cada item.

- “a) Desenvolver método de replicação de dados entre servidores;
- b) Mostrar limitações dos métodos existentes para replicação de servidores com balanceamento de carga.”

Considerando que o objetivo geral desse trabalho era “obter confiabilidade”, os objetivos específicos podem ser considerados satisfatórios. Mas ainda assim podem ser questionados. O segundo objetivo específico consiste em mostrar limitações dos métodos existentes. Mas pressupõe-se que tais limitações já sejam conhecidas pelo aluno, pois, caso contrário, por que ele estaria propondo um novo método (primeiro objetivo específico)? Aqui, o que aparece é um estágio ainda pouco maduro da pesquisa. O aluno deseja encontrar limitações nos métodos existentes para que isso justifique a proposição de um novo método.

- “a) Analisar métodos de gerência de projetos de software terceirizados;
- b) Desenvolver um método customizado para gerenciar os aspectos custo, prazo, qualidade e risco de projetos;
- c) Adaptar ou desenvolver uma ferramenta de software (caso não exista nenhuma adequada) para dar suporte à aplicação do método;
- d) Implantar o método;
- e) Coletar e analisar dados empíricos, resultantes da aplicação do método nos projetos;
- f) Propor melhorias no método desenvolvido.”

Aqui aparece claramente a confusão que muitas vezes se estabelece entre objetivos específicos, metodologia e cronograma de atividades. Os objetivos específicos deveriam ser resultados finais mensuráveis, em geral um detalhamento ou subproduto do objetivo geral. Mas o aluno apresenta praticamente uma sequência de passos que levam ao objetivo geral. Não chega a ser um método, pois faltaria complementar com outras informações, mas o que foi apresentado é uma lista de atividades sobre a qual possivelmente o cronograma de trabalho será construído. Pode-se até falar em objetivos intermediários de cada etapa do processo de pesquisa, mas vários deles não são objetivos de pesquisa, mas atividades técnicas.

- “a) Implementar o sistema de inicialização e abstrações de hardware para *** no ***;
- b) Definir, implementar, testar e avaliar uma família de protocolos de comunicação de ***;
- c) Definir, implementar, testar e avaliar um sistema de controle de recursos de energia para ***;
- d) Definir e implementar aplicações que permitam avaliar os sistemas implementados.”

Aqui a divisão do problema em subproblemas é até adequada, mas as atividades de definir, implementar e testar não são objetivos de pesquisa, mas objetivos técnicos.

- “a) Minimização do problema da oclusão de objetos;
- b) Redução do tempo de reconstrução, dada a opção de obter a forma do objeto por classificação de aspectos 2D, se ele já for conhecido.”

Aqui são dois objetivos específicos válidos. Apenas que, no caso do primeiro, não está muito claro o que é exatamente o “problema da oclusão”, nem que tipo de medição será possível para garantir que ele tenha sido minimizado.

- “a) Metodologia de busca de nodos externos de Octrees utilizando operações conhecidas sobre este tipo de árvore;
- b) Determinar pontos de inflexão em Octrees;
- c) Gerar bitmaps a partir de nodos de Octree;
- d) Determinar vetores normais em nodos de Octree para iluminação.”

O primeiro objetivo destoa dos demais por não ter um verbo associado. Os demais parecem mais ser atividades a serem executadas na busca do objetivo geral.

4.4. Análise de Justificativa

A justificativa deve se referir principalmente à hipótese de trabalho, ou seja, deve-se justificar a escolha de uma hipótese em vez de tentar justificar apenas a importância do tema da pesquisa. Usual-

mente a importância do tema da pesquisa já foi abordada na contextualização do problema. Então não há necessidade de repetir essa justificativa. A seguir são analisadas algumas justificativas produzidas em propostas de monografia, com a ideia de que seriam justificativas de hipótese.

“A eficiência de um plano amostral estratificado depende do conhecimento empírico e da experiência do pesquisador. O método proposto auxiliará no desenvolvimento do plano amostral, identificando estratos que fornecerão maior precisão.”

Aqui o aluno apresenta uma vantagem esperada para o método proposto, mas não justifica a pesquisa no sentido que seria esperado aqui. Ele deveria deixar claro a partir da identificação do problema quais foram as tentativas de solução testadas, onde elas falharam e finalmente, na justificativa, explicar por que acredita que a abordagem que ele propõe pode ter sucesso.

“Métodos atuais de replicação geram muito tráfego na rede e reduzem o paralelismo entre eles. Se não houver necessidade de paralelismo, a replicação pode ser feita de maneira muito mais rápida e com menos carga na rede.”

É uma boa justificativa, pois apresenta uma limitação relacionada aos métodos existentes e justifica, com base nisso, uma potencial solução, explicando por que ela seria adequada.

“Atualmente não existe um método para o acompanhamento dos projetos terceirizados de desenvolvimento de software adaptado à realidade das empresas públicas. Com a adoção de um método que permita o gerenciamento dos projetos de desenvolvimento de software terceirizados será possível realizar um melhor controle do andamento dos projetos, permitindo a verificação de desvios e a tomada de decisões para corrigi-los em tempo hábil.”

Não se recomenda fazer comparações com o vazio. No caso, em vez de dizer negativamente que “não existe um método...”, o que pode

ser altamente questionável, o aluno deveria ter verificado o que existe de semelhante e comparado com o que pretende fazer. Ao dizer que não existe alguma coisa ou que não foi encontrada alguma informação, o aluno pode passar a impressão de que não pesquisou o suficiente ou que aquilo que ele busca é tão irrelevante que ninguém nunca se preocupou em procurar. Assim, em vez de afirmar que algo não existe, o aluno deve mostrar o que existe de mais parecido com aquilo que ele quer e então trabalhar nas diferenças entre o existente e o desejado.

“O RBC é utilizado hoje com sucesso em campos como: medicina e mercado financeiro.”

A frase caracteriza mais a contextualização do problema ou importância do tema de pesquisa do que a justificativa do objetivo ou hipótese de trabalho.

“Falta de um software que permita ao especialista modelar o processo hospitalar com base em sua experiência e que seja capaz de executar e gerenciar o processo modelado, controlando inclusive a alocação de recursos. Falta de padronização nos dados médicos dificultando diagnósticos e comprometendo ou impossibilitando cálculos estatísticos.”

Aqui novamente o aluno procura justificar seu trabalho pela ausência de determinadas coisas. Falta software em várias áreas, mas nem sempre a construção de um software implica uma atividade de pesquisa. O mesmo vale para padrões.

4.5. Análise de Método de Pesquisa

A seguir são analisadas propostas de método conforme apresentadas. Todas as análises baseiam-se em versões extremamente resumidas das propostas, ou seja, suficientes para preencher uma transparência. Então todas elas carecem de um maior detalhamento ao serem transformadas em texto da proposta de pesquisa ou do trabalho final. Assim, a análise considerará questões relacionadas ao método no seu nível mais abstrato.

Inicialmente considera-se recomendável que o método de trabalho só seja definido quando já se sabe qual o objetivo da pesquisa. Como o objetivo só é consistentemente definido quando a revisão bibliográfica já está suficientemente madura, não se considera, para efeito de metodologia em computação, que a revisão bibliográfica seja parte do método. A revisão pode ser parte do cronograma do trabalho, mas não do método. O método deve mostrar como o proponente vai fazer para que sua hipótese de trabalho seja testada e que ao final se possa concluir se ela é ou não verdadeira.

- “a) Levantamento bibliográfico sobre os tópicos abordados neste trabalho como: Critérios Ergonômicos, Interatividade, Método de Apoio à autoria e outros;
- b) Pesquisa sobre o Método de Apoio a ser utilizado;
- c) Definição de recursos audiovisuais a serem utilizados no trabalho;
- d) Modelagem do sistema conforme o método de apoio à autoria escolhido;
- e) Elaboração da ferramenta multimídia;
- f) Testes do sistema”.

O método inicia com a revisão bibliográfica (itens a e b), quando fica claro que o aluno ainda não domina suficientemente o tema de pesquisa. Os passos (c) e (d) ficam atrelados ao que o aluno vier a descobrir na revisão bibliográfica e, por isso, ainda não podem ser mais específicos. Já os passos (e) e (f) são tão genéricos que serviriam para uma gama muito grande de propostas de monografia. A questão é: *como* a hipótese de pesquisa será testada. Isso o método não deixa claro.

- “a) Desenvolvimento de um mecanismo que não utiliza *design patterns* (sem conhecimento);
- b) Adição de conhecimento ao mecanismo;
- c) Experimentos + pesquisa bibliográfica = monografia.”

Aqui o aluno parece propor um estilo de projeto no qual vai avaliar uma determinada hipótese (uso de conhecimento ou *design patterns*) contra a opção que não utiliza essa hipótese. Trata-se de um trabalho em que o aluno desenvolve um sistema presumivelmente compatível com o estado da arte e depois insere nele alguma nova característica para avaliar se melhorou. Embora não seja a forma mais madura de pesquisa, pois carece de um *benchmark* mais universal, é uma forma válida, exceto pelo item (c) que não informa nada específico sobre o trabalho. Porém, para que o método fosse mais claro, seria necessário especificar o que significa adição de conhecimento (a hipótese) e mostrar quais testes seriam feitos para testar a validade ou não dessa hipótese, além, claro, de uma métrica bem definida para comparar as duas abordagens.

- “a) Utilizar técnicas existentes de replicação e medir desempenho da rede e do serviço;
- b) Implementar técnica proposta para replicação em serviços sem paralelismo e medir desempenho da rede e serviço.”

A pesquisa proposta aqui é bastante semelhante à imediatamente anterior. Implementa-se o que seria o estado da arte e compara-se com uma versão que tenha alguma característica nova.

- “a) Levantamento Bibliográfico;
- b) Criação do método;
- c) Adaptação ou desenvolvimento da ferramenta de software;
- d) Planejamento da aplicação do método;
- e) Aplicação do método;
- f) Avaliação;
- g) Conclusão.”

Essa apresentação é particularmente interessante porque ela serve para quase qualquer monografia de tão genérica que é. Ela consiste praticamente de um *template*, a partir do qual o cronograma de trabalho poderia ser descrito. Para isso seria necessário instanciar

cada um desses passos em atividades concretas e relacionadas com os objetivos do trabalho.

- “a) Desenvolvimento de um protótipo das quatro ferramentas englobando os conceitos de modelagem e gerenciamento de *workflow* em conformidade com a padronização ***;
- b) Aplicação dessas ferramentas em dois hospitais conveniados ao projeto com a finalidade de coleta de dados estatísticos para validação.”

Essa proposta tem a vantagem, em relação às anteriores, de mencionar explicitamente a aplicação da técnica sendo estudada em casos reais (hospitais). Menciona-se, porém, coleta de dados, sem deixar muito claro que tipo de dados serão coletados e que tipo de análise será feita. Dependendo do tipo de dados a ser levantado, deve-se tomar cuidado, porque estes podem não ter representatividade estatística, como se quer. Um exemplo extremo seria considerar que se trata de dois hospitais especializados em cirurgia cardíaca. Se a maioria dos pacientes que ali se interna tem como objetivo um tratamento cardíaco, essa não é uma informação que possa ser generalizada para outros hospitais. Trata-se de um fato verdadeiro apenas naquela realidade, já que existe uma relação entre causa (o hospital ser especializado em cirurgia cardíaca) e efeito (um maior número de tratamentos cardíacos naquele hospital). Outras relações bem mais sutis podem ocorrer. Então o termo “estatístico” deve ser usado com muito cuidado.

Um último exemplo:

- “a) Design
 - a. Application Oriented System Design
- b) Testes de Corretude
 - a. Técnicas de Depuração para Software Embutido
- c) Testes de Performance
 - a. Executado via Testes Comparativos
 - b. Métricas

- i. Tamanho de Código
- ii. Tempo de Execução de Tarefas
- iii. Energia gasta na Execução
 - 1. Tempo de Operação de Componentes de Hardware
- iv. *Throughput* de Rede
 - 1. Simulação
 - 2. Testes de Campo.”

A descrição do método em si deixa muita coisa subentendida, mas tem o mérito de mencionar explicitamente quais são as métricas buscadas nos testes comparativos.

Capítulo 5

Escrita da Monografia

A escrita de uma monografia depende fundamentalmente da existência de algum conteúdo a ser apresentado. O Capítulo 3 apresenta um caminho para a busca desse conteúdo. Esse caminho se inicia na escolha de um tema, seguido pela revisão bibliográfica, escolha de um objetivo com uma hipótese justificada, definição de um método de trabalho e execução do método assim proposto para a coleta de resultados que serão posteriormente analisados.

5.1. Como os Capítulos de uma Monografia São Ordenados

Para quem lê pela primeira vez uma monografia, pode ficar a impressão de que o trabalho foi escrito sequencialmente da mesma maneira como se apresenta. Porém, normalmente este não é o caso.

A ordem física usual dos capítulos de uma monografia é a seguinte:

- a) Resumo.
- b) Introdução.
- c) Revisão bibliográfica.
- d) Desenvolvimento.
- e) Conclusões.
- f) Referências.

Dependendo do trabalho, a revisão bibliográfica e o desenvolvimento poderão ocupar mais de um capítulo cada. Mas normalmente a ordem em que as seções se apresentam é essa.

Esta é a ordem em que o trabalho será lido por alguém que esteja iniciando a pesquisa nessa área de conhecimento. Primeiramente, é necessário conhecer o resumo para saber de que se trata o trabalho. A introdução apresenta objetivos, limitações e metodologia do trabalho, além de situar o trabalho resumidamente no estado da arte. Será necessário ler a revisão bibliográfica para que se conheça algo sobre outros trabalhos similares, bem como os conceitos fundamentais para a compreensão do trabalho propriamente dito que é apresentado no capítulo do desenvolvimento. Finalmente, o leitor observará as conclusões, e, se quiser aprender mais sobre o assunto, consultará as referências listadas.

Porém, essa não é necessariamente a ordem em que o trabalho será lido pela banca examinadora.

5.2. Como uma Monografia poderá ser Lida pela Banca Examinadora

Deve-se presumir que a banca seja composta por especialistas no assunto do trabalho, os quais já conhecem suficientemente os principais conceitos e trabalhos correlatos. Esses especialistas possivelmente farão uma leitura em ordem distinta, procurando obter inicialmente as informações mais relevantes sobre a contribuição do aluno, para em seguida analisar os aspectos mais triviais do trabalho. Os especialistas vão direto ao cerne da questão, deixando os complementos para avaliar em um segundo momento.

A forma como um avaliador lê uma monografia depende da personalidade e experiência do avaliador, porém, uma sequência de leitura muito interessante pode ser definida da seguinte forma:

- a) Resumo.
- b) Referências.
- c) Introdução.
- d) Conclusão.
- e) Desenvolvimento.
- f) Revisão bibliográfica.

O especialista inicialmente lê o resumo para saber do que trata o trabalho. Ele verifica se o resumo é compatível com o título do trabalho. Ele espera que o resumo apresente o problema sendo tratado, uma justificativa para a hipótese escolhida e, principalmente, uma descrição rápida dos resultados obtidos ou contribuições.

Um trabalho cujo resumo não indique claramente a existência de contribuições relevantes possivelmente não será bem avaliado. Deve-se evitar, portanto, que o resumo apresente apenas informações do tipo “Este trabalho apresenta um estudo sobre...”, ou “Este trabalho propõe um método para...”.

É necessário que o resumo apresente alguma informação do tipo “O principal resultado obtido a partir deste trabalho é...”. Esse resultado, supõe-se, deve ser relevante, no sentido que foi comentado anteriormente, ou seja, deve corresponder a algum conhecimento que não estava disponível antes da execução do trabalho e que foi descoberto durante a sua elaboração. Não será suficiente, por exemplo, escrever “O principal resultado obtido a partir deste trabalho foi o estudo de...”, pois, como foi dito, o estudo é um objetivo pessoal do aluno e não o objetivo propriamente dito do trabalho de pesquisa.

Na sequência, tendo o especialista compreendido o resumo e conhecido a real contribuição do trabalho, vai verificar se a alegada contribuição realmente foi obtida. Inicialmente ele verificará as referências citadas pelo aluno para ver se os principais trabalhos da área estão ali. Verificará também a existência de artigos recentes em eventos e periódicos, bem como a relevância desses mesmos eventos e periódicos.

É possível reprovar um aluno em função das referências citadas no trabalho dele? No nível de mestrado e doutorado, sim. Por exemplo, considere-se uma monografia sobre XML que apresente como referências 12 trabalhos, sendo três livros técnicos de XML e nove páginas na Internet que consistem em manuais técnicos de XML. Nesse caso, não há evidência de trabalho científico, mas apenas de trabalho técnico. O aluno pode até ter feito uma implementação muito bonita de alguma coisa com XML, mas possivelmente

não terá valor científico se não estiver fundamentado em trabalhos relatados em eventos ou periódicos.

Estando as referências adequadas, o avaliador observará o capítulo de introdução com atenção. Neste capítulo ele entenderá com mais detalhes o problema sendo resolvido, a técnica usada para solução e a forma como os resultados foram validados. O avaliador estará especialmente atento aos objetivos do trabalho, incluindo os objetivos específicos, que ele procurará identificar claramente nas conclusões.

Ao ler as conclusões, um bom avaliador esquadrinhará o texto buscando um comentário conclusivo sobre cada um dos objetivos do trabalho. A falta dessa ligação direta entre objetivos e conclusões certamente dará margem a críticas.

Após verificar as conclusões, o avaliador desviará sua atenção ao capítulo de desenvolvimento, a fim de verificar como o aluno chegou àquelas conclusões. As conclusões terão de ser todas consequências de avaliações feitas no capítulo de desenvolvimento. Não se admite, por exemplo, que um aluno conclua que seu sistema é fácil de usar, se no desenvolvimento testou apenas a eficiência do sistema, sem checar a usabilidade. Concluir algo que não foi testado ao longo do trabalho é denominado “conclusão forte”, sendo inaceitável em trabalhos científicos.

Tendo compreendido a real contribuição do aluno, o avaliador finalmente dará uma vista de olhos no capítulo da revisão bibliográfica para ver se os principais conceitos foram bem apresentados e se os trabalhos correlatos estão adequadamente descritos.

A forma de leitura descrita parece ir das pontas para o centro da monografia. O especialista concentra-se primeiro nos pontos críticos do trabalho, para depois ler as partes menos sujeitas a problemas.

5.3. Como uma Monografia poderia ser Escrita

Já a ordem sugerida para escrever os diferentes capítulos de uma monografia difere das duas anteriores. Pode-se escrever a monografia exatamente na ordem em que os capítulos se apresentam.

Mas essa abordagem tem alguns inconvenientes, como, por exemplo, produzir uma revisão bibliográfica desnecessariamente longa (por não saber quais conceitos realmente serão usados no texto, coloca-se tudo sobre o que se leu), e, por vezes, deixa o aluno cansado na hora de escrever o mais importante: as conclusões. Quantas pessoas já não entregaram seu trabalho ao orientador dizendo “Está pronto! Só faltam as conclusões”. Porém, o mais importante num trabalho científico são as conclusões. O resto é apenas um meio convincente para se chegar a elas.

Comer (2008) diz *“the easiest way to build a dissertation is inside-out. Begin by writing the chapters that describe your research (...). Collect terms as they arise and keep a definition for each. Define each technical term, even if you use it in a conventional manner.”*¹

Recomenda-se que só se inicie a escrita do texto final depois de ter concluído pelo menos a maior parte dos experimentos. Nada impede a escrita de rascunhos, mas esses textos não precisam ainda ter a preocupação de formatação de um texto finalizado.

Então, quando o aluno terminar os experimentos e já tiver uma ideia muito clara do que fez e quais resultados obteve, ele vai transformar seus rascunhos em um texto acabado. Mas em que ordem ele deveria escrever esse texto? Segue uma sugestão:

- a) Introdução.
- b) Desenvolvimento.
- c) Conclusões.
- d) Revisão Bibliográfica.
- e) Referências.
- f) Resumo.

Recomenda-se iniciar pela introdução porque possivelmente é o rascunho que estará mais próximo da forma final neste momento. O capítulo de introdução frequentemente é uma reescrita da pro-

¹ Tradução: A maneira mais fácil de construir uma dissertação é de dentro para fora. Comece escrevendo os capítulos que descrevem sua pesquisa (...). Crie termos à medida que eles surgem e mantenha uma definição para cada um deles. Defina cada termo técnico, mesmo se você o usar da maneira convencional.

posta de monografia. Troca-se o tempo futuro pelo tempo presente e a monografia passa a ter objetivo, justificativa, hipóteses etc., que eram os mesmos, muitas vezes, da proposta de pesquisa.

Em seguida, o desenvolvimento deve ser escrito porque os experimentos ainda estarão fresquinhos para serem relatados. Esse é um capítulo importante, por isso deve ser escrito com muita atenção.

As conclusões deveriam ser escritas logo depois. Após terminar o relato do desenvolvimento e já tendo o capítulo de introdução sido revisado com seus objetivos claramente colocados, a conclusão deverá ser uma consequência daquilo que foi relatado no desenvolvimento, bem como ter ligação com cada um dos objetivos, conforme já observado.

Na sequência escreve-se então a revisão bibliográfica e listam-se as referências. Deve-se evitar colocar na revisão bibliográfica a totalidade das fichas de leitura, pois isso não serve a um objetivo claro. Deve-se colocar, isso sim, as comparações com trabalhos correlacionados e os conceitos principais usados no capítulo de desenvolvimento. Um conceito que não é usado no capítulo de desenvolvimento, por mais interessante que seja, não precisa estar na revisão bibliográfica. As referências vão mencionar apenas os trabalhos efetivamente citados. Por isso é importante que essas duas sessões sejam verificadas em conjunto.

Finalmente, escreve-se o resumo, que, como o nome diz, resume o trabalho. Mais adiante será discutido o que deveria ser dito nesta seção do trabalho.

5.4. O Título

O título do trabalho é o primeiro meio para chamar a atenção de um potencial leitor. Um título deve descrever a principal contribuição do trabalho de forma sintética. Alguns títulos, por serem muito genéricos, não motivam a leitura. Por exemplo, “um estudo sobre redes semânticas”. Para que fosse um bom título, seria melhor dizer que tipo de resultado esse estudo produziu. Outro exemplo de um título não motivador é “XYZ: uma nova técnica de modelagem

de dados”. Novamente, seria mais interessante se o título pudesse esclarecer que tipo de vantagem essa nova técnica teria em relação a outras. Apenas o fato de ser nova não garante que seja interessante.

Seguem alguns exemplos de bons títulos obtidos na Biblioteca Digital da Sociedade Brasileira de Computação, SBC (<http://bibliotecadigital.sbc.org.br/?subject=144>):

- a) Formal Approaches to Ensuring the Safety of Space Software.²
- b) Automação de Métodos e Técnicas para Teste Funcional de Componentes.
- c) Electric Wheelchair Simulator for Rehabilitation of Persons with Motor Disability.³
- d) Análise Comparativa dos Dicionários LBG e SOA sob o Ponto de Vista da Complexidade Computacional Envolvida na Fase de Codificação da Quantização Vetorial.

O título deve apresentar, então, claramente, a ideia ou contribuição central do trabalho. Porém, o título não pode ser muito longo. A apresentação mais detalhada da ideia ou contribuição central será feita então no resumo do trabalho, que normalmente aparece logo após o título.

5.5. O Resumo

O resumo de uma monografia não é, como alguns parecem pensar, um *trailer* de um filme, em que se começa a contar uma história, mas não se conta o final. O resumo de um trabalho científico deve contar o *final* da história, ou seja, o leitor vai querer saber, em primeiro lugar, qual foi o resultado científico a que esse trabalho chegou. Se ele achar o resultado interessante no resumo, vai querer ler o resto para ver como o aluno chegou a tal resultado.

Centenas de monografias em Computação são defendidas a cada ano apenas no Brasil. Se contarmos outros países, ainda teremos uma infinidade de material científico disponível cujo índice de

² Tradução: Abordagens Formais para Garantir a Segurança de Software Espacial.

³ Tradução: Simulador de Cadeira de Rodas Elétrica para Reabilitação de Pessoas com Deficiência Motora.

produção aumenta cada vez mais. Esperar que alguém leia uma monografia cujo resumo diz “Este trabalho apresenta um estudo sobre bancos de dados” é esperar demais. Afinal, essa frase diz muito pouco. O que efetivamente esse “estudo” poderia ter gerado em termos de informação nova que poderia interessar a alguém que trabalhe com bancos de dados?

Seria muito mais informativo um resumo que dissesse algo do tipo “Este trabalho demonstra que as sete formas normais de bancos de dados não são suficientes para evitar um problema de inconsistência dos dados identificado aqui como...”. Se o leitor está acostumado a trabalhar com as sete formas normais e acha que elas explicam como deve ser um bom banco de dados relacional, então ele ficará curioso para ver que caso estranho é esse que não é atendido pelas formas conhecidas. Ao longo do texto, mais detalhes serão dados, mas a atenção do leitor já foi conquistada.

Portanto, o resumo da tese ou monografia é efetivamente o lugar para vender o peixe. Se o autor não conseguir deixar um leitor interessado no resumo, não conseguirá fazer com que ele leia sua monografia quando há tanto outro material de boa qualidade disponível. Além disso, sistemas de indexação em bases de dados de *abstracts* também não vão identificar o trabalho adequadamente.

Alguém poderia argumentar que o resumo, usualmente com menos de uma página, é um espaço muito pequeno para apresentar uma grande contribuição obtida em mais de dois anos de trabalho. Mas o problema é o seguinte: se o autor não consegue explicar a contribuição de seu trabalho em uma página (resumo), então deve haver algo muito errado no seu trabalho, ou na sua capacidade de ser sucinto.

Uma coisa que não se faz no resumo é revisão bibliográfica. A não ser que seja vital para a compreensão do trabalho, não se deve fazer citações bibliográficas no resumo. Não é razoável perder valiosas linhas citando trabalhos de outras pessoas. Esse espaço é reservado para o autor da monografia dizer a que veio e o que trouxe.

O resumo deve conter uma explicação bastante clara sobre o real problema abordado no trabalho, pois pessoas com problemas semelhantes poderão se interessar. Além disso, um esboço da solução usada deve ser também apresentado, pois pessoas que usem tecnologias parecidas poderão também se interessar em ver uma possível nova classe de problemas sendo resolvidos por essa tecnologia.

Segundo Rugaber (1995), o propósito de uma monografia é a apresentação de uma tese. Então, faz sentido apresentar essa tese o mais cedo possível, ou seja, no resumo. A tese é definida como sendo uma afirmação, que se procura comprovar verdadeira. Se um trabalho em nível de mestrado ou doutorado não puder ser definido a partir de uma tese, que possa ser expressa em uma frase, então possivelmente algo está errado na concepção do trabalho.

5.6. A Introdução

O capítulo de introdução apresentará de forma mais detalhada o tema e o problema de pesquisa. Em relação ao tema, espera-se uma descrição geral da área e da abrangência do estudo. Deve-se evitar, porém, introduções muito longas, por exemplo, iniciando na pré-história, para chegar a explicar que o tema do trabalho é relativo a redes de computadores.

A introdução deve conter os elementos que já foram mencionados no projeto de pesquisa, ou seja, os objetivos geral e específicos, resultados esperados, limitações do trabalho, metodologia utilizada e justificativa. Em geral, o capítulo de introdução é fechado por uma descrição sucinta dos demais capítulos do trabalho.

5.7. O Capítulo de Revisão Bibliográfica

O capítulo de revisão bibliográfica contém trabalhos de outros autores que de alguma forma se relacionam com o texto da monografia. Neste capítulo é importante ser objetivo na apresentação, pois a quantidade de textos a serem consultados na maioria das áreas é muito grande. Não se recomenda fazer grandes digressões

sobre trabalhos que não sejam diretamente relacionados ao tema da monografia. Por exemplo, se o trabalho trata da comparação entre técnicas de mutação em algoritmos genéticos, então não seria necessário neste capítulo abordar uma revisão bibliográfica sobre robótica ou redes semânticas, que, embora sejam temas diretamente ligados à Inteligência Artificial, são colaterais ao tema do trabalho.

O capítulo de revisão bibliográfica deve, isso sim, abordar os principais conceitos da área de pesquisa para que possam servir de referência a eventuais leitores que não sejam exatamente especialistas no assunto. Recomenda-se que esses conceitos sejam, se possível, citados a partir de mais de uma fonte, e que a forma de organização não seja copiada de um único trabalho. Por exemplo, se os autores A, B e C apresentam os conceitos 1, 2 e 3, em vez de citar os conceitos 1, 2 e 3 de acordo com A, seguido dos conceitos 1, 2 e 3 de acordo com B e finalmente 1, 2 e 3 de acordo com C, sugere-se que sejam citados o conceito 1, de acordo com A, B e C, se possível com alguma comparação feita pelo próprio autor; posteriormente citam-se as definições para o conceito 2, de acordo com A, B e C, e finalmente o conceito 3, de acordo com A, B e C.

Assim, o capítulo de revisão bibliográfica organiza-se por conceitos e não por autores. A exceção a essa regra se dá apenas no caso da apresentação de trabalhos correlatos. A seção de trabalhos correlatos deve descrever pesquisas semelhantes àquela do autor.

A comparação do trabalho atual com trabalhos correlatos pode ser feita em dois momentos: na revisão bibliográfica e/ou nas considerações finais. No final do capítulo de revisão bibliográfica pode ser comparada a hipótese do trabalho com as hipóteses de trabalhos correlatos. Ainda não se têm dados efetivos do trabalho porque esses dados vão aparecer apenas no capítulo de desenvolvimento. Mas já é possível discutir sobre vantagens e desvantagens das diferentes hipóteses de pesquisa de cada um dos trabalhos.

O segundo momento em que a comparação pode ser feita é no final do documento, em geral, no final do capítulo de desenvolvimento ou no capítulo de conclusão ou considerações finais. Nesse ponto

já se têm os resultados do trabalho em questão, os quais podem ser comparados com os resultados de outros trabalhos correlatos. Trata-se, portanto, nesse caso, de uma comparação bem mais objetiva.

Um ponto a ser sempre mencionado é a questão do plágio. Em hipótese alguma pode-se utilizar textos escritos por outra pessoa, mesmo em um trabalho de cunho eminentemente escolar, sem colocar o texto citado entre aspas e, quando possível, mencionar a fonte de consulta. Utilizar o trabalho de outro como se fosse seu é considerado crime de plágio pela legislação brasileira.

Mesmo traduções devem ser evitadas. Se for o caso de mencionar algum texto em outra língua, recomenda-se que seja mantido no original entre aspas e com a fonte citada. Opcionalmente pode-se colocar em nota de rodapé uma tradução. Se a fonte de consulta for a obra traduzida, então pode-se citar conforme a língua em que foi traduzida (no caso Português), mas se a tradução for feita pelo próprio autor do trabalho, recomenda-se manter o original e utilizar a nota de rodapé para a tradução, visto que não se trata de uma tradução oficial e, portanto, o sentido da tradução não pode ser atribuído ao autor do texto original.

Rugaber (1995) afirma que a seção de revisão bibliográfica em geral é maçante e mal usada. Isso acontece porque o aluno, ao escrever, perde a oportunidade de usar o trabalho de outros autores para motivar o seu próprio estudo. Em vez disso, muitas vezes perde-se tempo e espaço fazendo um estafante e desnecessário inventário de tudo o que foi lido ou ainda uma sequência de citações de outros autores que não levam a lugar nenhum.

Ainda segundo Rugaber, uma boa seção de revisão bibliográfica inclui uma ontologia e uma ontogenia. A *ontologia* vai apresentar os principais conceitos utilizados no trabalho. Já a *ontogenia* mostra como as diferentes ideias evoluíram com o passar do tempo até chegar ao estado da arte. O final da ontogenia naturalmente vai incluir o trabalho sendo desenvolvido, mostrando como ele se encaixa na evolução da área.

5.8. O Capítulo de Desenvolvimento

O capítulo de desenvolvimento marca o início da contribuição pessoal do autor do trabalho. Portanto, não se deve fazer do capítulo de desenvolvimento uma nova revisão bibliográfica. De preferência, todos os conceitos que serão necessários nesse capítulo já devem ter sido citados no capítulo de revisão bibliográfica. Se alguma comparação for feita com trabalhos correlatos nesse capítulo, então apenas a comparação objetiva deve ser feita aqui, sendo que a apresentação pura e simples dos trabalhos correlatos já terá ocorrido no capítulo anterior.

O capítulo de desenvolvimento deve apresentar a construção da teoria, modelo ou proposta, seja de que natureza for. Conceitos criados pelo autor da monografia devem ser descritos aqui e não na revisão bibliográfica. Na sequência, o autor deve trabalhar as evidências de que sua hipótese é verdadeira. Serão então apresentados dados, gráficos, testes, provas formais, estudos de casos, transcrição de entrevistas ou quaisquer outros meios julgados adequados para provar o seu ponto, ou seja, para mostrar que a hipótese é verdadeira.

Deve-se evitar sempre transformar o capítulo de desenvolvimento em uma apresentação de um sistema computacional. Se um sistema foi desenvolvido, foi para servir a algum propósito de descobrir novo conhecimento. A monografia deve ser sobre o conhecimento gerado, não sobre o sistema em si. Apresentações detalhadas sobre telas de software, incluindo telas de *login*, menu principal etc., são enfadonhas e desnecessárias em um trabalho científico. O seguinte texto, do Prof. John W. Chinneck (1988), resume tudo: “*The purpose of your thesis is to clearly document an original contribution to knowledge. You may develop computer programs, prototypes, or other tools as a means of proving your points, but remember, the thesis is not about the tool, it is about the contribution to knowledge. Tools such as computer programs are fine and useful products, but you can't get*

an advanced degree just for the tool. You must use the tool to demonstrate that you have made an original contribution to knowledge; e.g., through its use, or ideas it embodies.”⁴

5.8.1. Definições Constitutivas e Operacionais

No capítulo de desenvolvimento frequentemente o pesquisador precisará definir termos que esteja usando. Existem duas formas clássicas de criar definições: as definições constitutivas e as operacionais. Dependendo do tipo de variável, será necessário usar uma ou outra.

Trabalhos eminentemente formais tendem a usar mais as definições constitutivas. Segundo Kerlinger (1980), “*definições constitutivas são definições de dicionário*”. As definições constitutivas procuram definir um termo em função de seus constituintes. Uma gramática formal, por exemplo, pode ser definida como um conjunto de regras de produção; uma regra de produção pode ser definida como duas sequências de símbolos e assim por diante.

Porém, pesquisas que utilizem termos não formais como “facilidade”, “adequação”, “flexibilidade” etc. dificilmente poderão utilizar apenas definições constitutivas para esses termos, até porque em geral tais definições sequer existem.

É necessário, nesses casos, utilizar uma definição operacional que, segundo Kerlinger (1980): “*atribui significado a um constructo ou variável especificando as atividades ou ‘operações’ necessárias para*

⁴ Tradução: “O propósito da sua monografia é documentar claramente uma contribuição original ao conhecimento. Você pode desenvolver programas de computador, protótipos e outras ferramentas como forma de provar suas ideias, mas lembre-se, a monografia não é sobre a ferramenta, é sobre a contribuição ao conhecimento. Ferramentas tais como programas de computador são produtos bons e úteis, mas você não pode obter um título de pós-graduação somente pela ferramenta. A ferramenta deve ser usada para demonstrar que você fez uma contribuição original ao conhecimento; por exemplo, através de seu uso, ou pelas ideias que através dela são materializadas.” Tradução de Américo E. de Oliveira Costa e Vitória Pureza, disponível em <http://www.sce.carleton.ca/faculty/chinneck/thesis/ThesisPortuguese.html>.

medi-lo ou manipulá-lo”. A definição operacional é, então, uma definição pragmática. Ela não define a natureza de um fenômeno, mas os meios para obter uma medição e caracteriza o resultado dessa medição como sendo o próprio fenômeno.

Por exemplo, o termo “facilidade” pode ser definido como o número de toques no teclado ou mouse para realizar uma determinada tarefa. O termo “adequação” pode ser definido como a nota obtida em um teste padrão aplicado por especialistas. O termo “flexibilidade” pode ser definido como o tempo médio que um programador leva para introduzir um conjunto predefinido de características.

Mas esses são apenas exemplos do que poderia ser feito. Não se pretende aqui fornecer definições operacionais para esses termos. Cada trabalho usará as definições que forem mais adequadas ao objetivo da pesquisa.

O importante aqui é enfatizar que no caso de variáveis que representem características não formais é necessário utilizar definições operacionais para que o fenômeno associado à variável possa efetivamente ser medido. Sem isso o trabalho seria apenas especulativo.

5.9. O Capítulo de Conclusões

O capítulo das conclusões é, em geral, a pedra no sapato do estudante. Aparentemente tudo já foi dito sobre o trabalho no capítulo de desenvolvimento; então o que escrever nesse capítulo final?

A primeira dica é observar os objetivos geral e específicos do trabalho no capítulo de introdução e colocar no capítulo das conclusões um comentário sobre como o desenvolvimento apresentado ajudou a chegar a cada um desses objetivos, ou seja, como o trabalho de pesquisa permite concluir que cada um dos objetivos foi atingido.

Outro ponto importante é apresentar não apenas os pontos positivos do trabalho, mas também os negativos. Não se espera de nenhum trabalho científico que ele resolva todos os problemas do mundo. Pelo contrário, espera-se que o pesquisador seja suficiente-

mente honesto para descrever de forma clara as fraquezas e limitações de seu próprio trabalho.

A seguinte máxima segue implacavelmente das leis da lógica: “Se você não for o maior crítico de seu próprio trabalho, outra pessoa será”.

Outro tópico a ser abordado no capítulo de conclusões são as lições aprendidas. O aluno passou dois anos ou mais estudando um tema e realizando experimentos com ele. Além dos objetivos do trabalho, claramente colocados e atingidos, ele deve ter aprendido muita coisa no processo. Talvez essa informação possa ser útil a outras pessoas. Então se deve descrever no capítulo de conclusões quais foram essas lições aprendidas ao longo do trabalho.

Pode-se descrever também outras situações nas quais se imagina que essas lições possam ser aplicadas. Por exemplo, ao comparar o resultado de questionários aplicados em uma empresa com a situação real observada *in locu*, um aluno percebeu que, por conta do medo de retaliações por parte da chefia, a maioria dos funcionários procurava apresentar nos questionários uma situação mais bonita do que realmente era. Dessa forma, o aluno aprendeu que questionários não são fontes confiáveis de informação se não houver uma validação das respostas no ambiente de estudo. Essa lição aprendida teria de ser colocada no capítulo de conclusões.

Segundo Chinneck (1988), o capítulo final de uma monografia deve ter pelo menos três partes: a conclusão, as contribuições e os trabalhos futuros.

Na conclusão o aluno fará de forma concisa uma referência ao problema examinado e resolvido. A conclusão propriamente dita teria o seguinte formato: “o problema descrito na seção x foi resolvido como demonstrado nas sessões y e z, em que foi desenvolvido um algoritmo/método/abordagem etc. para tratar as situações mencionadas” (Chinneck, 1988).

Ainda segundo Chinneck, o resumo das contribuições, que viria em seguida, deve ser organizado em ordem decrescente de sua importância, por exemplo:

- a) Desenvolveu-se um algoritmo muito mais rápido para problemas de *Zylon* de grande porte.
- b) Demonstrou-se pela primeira vez o uso do mecanismo de *Grooty* para os cálculos de *Zylon*.
- c) Etc.

As contribuições mais importantes do trabalho serão aquelas que geraram conhecimento novo. Ferramentas, protótipos e outros artefatos tecnológicos usualmente são contribuições secundárias.

Artigos publicados e relatórios de pesquisa não são contribuições, nesse sentido da palavra, mas relatos da pesquisa. Portanto, tais referências não deveriam ser mencionadas aqui.

Finalmente, os trabalhos futuros são a contribuição que o aluno deixa para que outros possam continuar sua pesquisa. Trabalhos futuros também devem tratar de futuras contribuições ao conhecimento com mais ênfase do que futuras contribuições às ferramentas, protótipos etc., que eventualmente possam ser desenvolvidas.

5.9.1. Trabalhos Futuros

Espera-se sempre que uma monografia não seja apenas o final de uma pesquisa, mas também o início de uma caminhada. Assim, a seção final das conclusões normalmente é dedicada a deixar para os leitores ideias de oportunidades de pesquisa com as quais o autor se deparou ao longo do seu trabalho, mas que não teve tempo ou possibilidade de perseguir.

Nessa seção esperam-se dicas sobre trabalhos de pesquisa futura, e não trabalhos técnicos futuros. Por exemplo, o leitor terá pouco interesse em saber que o autor pretende futuramente implementar o sistema em Java, já que a versão atual está em C. Isso é apenas uma questão técnica. O leitor vai querer saber, sim, quais situações não foram testadas com a ferramenta atual e que poderiam ser relevantes para compreender seu comportamento. Isso seria um trabalho futuro de pesquisa relevante.

5.10. Seção de Bibliografia ou Referências Bibliográficas

Não é objetivo deste livro apresentar normas de citação bibliográfica, já que estas encontram-se disponíveis em qualquer biblioteca que se preze. Como existem diferentes padrões de citação, sugere-se que se siga as normas estabelecidas pelo curso, periódico ou evento no qual o trabalho será publicado. Como referência pode-se consultar na Internet o trabalho de Alves e Arruda (2007), baseado na NBR 6023/2002.

Apenas uma dica: após colocar todas as referências bibliográficas, deve-se verificar se cada uma delas é citada no texto e se todas as que são citadas no texto aparecem na lista de referências.

5.11. A Forma do Texto Científico

Comer (2008) apresenta uma série de recomendações sobre a forma do texto científico. Primeiramente, ele deve comunicar uma ideia de pesquisa e seus resultados. Isso, claro, é uma questão semântica. Mas há uma série de pequenos vícios que muitos alunos têm, e que, por serem tão comuns, estão sendo citados aqui para que possam ser evitados.

Esta seção não tem a intenção de ensinar a escrever português correto, nem a formatar trabalhos monográficos ou artigos. O objetivo da seção é apresentar e discutir erros frequentes, cometidos por alunos de computação em seus trabalhos, que podem ser facilmente minimizados.

Primeiramente, uma monografia deve evitar sempre que possível o uso de advérbios. Estranho, não? O advérbio é uma palavra que modifica um verbo e é bastante comum na língua portuguesa. Porém, o uso do advérbio no texto científico frequentemente estraga uma frase que sem ele poderia ficar bem melhor. Embora o advérbio possa ser muito útil no texto dissertativo, como, por exemplo, neste livro, para enfatizar ideias, no texto científico, o uso do advérbio deve ser minimizado, pois pode reforçar desnecessariamente certas afirmações. Por exemplo, dizer que “a experiência demonstra que as

abordagens são equivalentes” é uma coisa, mas dizer “a experiência demonstra *definitivamente* que as abordagens são equivalentes” dá um ar de prepotência ao texto, o que não é necessário para a frase em questão. Pode-se notar que as duas frases têm o mesmo sentido, mas a primeira soa bem melhor do que a segunda.

Algo a ser evitado em um texto científico também são as brincadeiras, piadas ou ironia. Eventualmente, periódicos científicos até publicam textos com esse tipo de expediente, mas nesses casos, usualmente o autor é algum “Papa” da área. Quem não é o “Papa” deve se abster de usar esses recursos no texto científico.

Em um texto científico não se espera que o autor utilize julgamento de valor sobre temas que não podem ser avaliados como ruim ou bom de forma maniqueísta. Dizer, por exemplo, que orientação a objetos é *bom* enquanto projeto estruturado é *ruim* para o desenvolvimento de software é uma questão de opinião. Em vez de dizer que algo é bom, deve-se procurar salientar uma das qualidades que se julga boa. Em vez de dizer que algo é ruim, deve-se apresentar uma explicação baseada em fatos verificáveis sobre os defeitos que se julga serem um problema.

Nunca se deve dizer que alguma coisa é “perfeita”, porque na natureza nada é.

Expressões como “hoje em dia” e “atualmente” também devem ser evitadas no texto científico, porque a monografia é um trabalho atemporal. Então não se deveria evitar dizer que “*atualmente* a Internet é bastante usada”. Pode-se dizer, em vez disso, que em 2009, tantos milhões de pessoas usam a Internet. Dessa forma, a frase ficará mais precisa e dará melhor embasamento científico a qualquer outra informação que se for apresentar em seguida.

Não se deve usar expressões do tipo “ficamos surpresos ao perceber que...”. Não é da conta de ninguém se o aluno ficou surpreso. Deve-se apenas dizer o que foi percebido e, de preferência, mostrar *como* isso foi percebido de forma que também possa ser percebido por outras pessoas.

Afirmações como “uma nova abordagem”, “uma técnica diferente” etc. devem ser evitadas a todo o custo. Qualquer abordagem que estiver sendo proposta será nova e diferente das outras, caso contrário, não seria tema de uma monografia. Não faria sentido propor algo que seja velho ou igual ao que já existe, não? Então, deve-se evitar qualificar o trabalho como *novo* e *diferente*, pois isso já é o que todos esperam. É melhor apenas dizer, afinal de contas, qual é o trabalho e em que ele é diferente ou melhor que os anteriores.

Usar palavras como “obviamente” ou “claramente” pode insultar o leitor, pois o autor diz estar falando algo que é óbvio. Se for óbvio, não precisa ser dito, se não é óbvio então não se deve dizer que é. Então, *obviamente*, deve-se evitar usar esse termo.⁵

Sempre que o texto usar a expressão “na verdade” pode dar a impressão de que aquilo que foi escrito antes era mentira. Então, *na verdade*, essa não é uma boa escolha.⁶

Deve-se evitar sempre o uso da primeira pessoa, mesmo o plural majestático, reservado apenas ao Papa e aos reis, bem como a segunda pessoa no texto científico. A monografia é impessoal. Ela não é narrativa, para que se pudesse usar primeira ou segunda pessoa. Então o impessoal sempre deve prevalecer.

Quando for usado um dos pronomes “todos”, “muitos”, “alguns” ou “nenhum”, deve-se ter certeza de que se tem uma evidência ou prova de que a afirmação efetivamente possa ser assim qualificada. Não se pode dizer, por exemplo, que “muitos estudantes têm problemas com a monografia”, em um texto científico, sem ter havido um estudo, observação ou medição sobre isso. Outra possibilidade é usar uma citação. Dizer “segundo Fulano (200x), muitos alunos têm problemas com a monografia” está correto, porque passa o problema de provar a afirmação para o autor do texto citado. No entanto, nesse caso, é preciso verificar se a referência é um trabalho confiável publicado em um bom veículo.

⁵ Sugere-se que o leitor releia a frase sem a palavra em itálico.

⁶ Sugere-se que o leitor releia a frase sem a expressão em itálico.

Outras recomendações são: usar a voz ativa ao invés da passiva, escrever sempre no tempo presente e colocar negações no início da frase. Essas recomendações de estilo ajudam a facilitar a compreensão do texto. Seria difícil ler uma frase longa na qual se vai afirmando uma série de coisas e no final se usa uma expressão para indicar que tudo o que estava sendo dito é falso. Se o objetivo é negar algo, então é preferível iniciar negando. Assim ficará mais fácil compreender o texto. Em todo o caso, sempre é preferível usar frases afirmativas quando possível. Por exemplo, em vez de dizer “nenhum programa rodou em menos de dez segundos” é preferível dizer “todos os programas rodaram em dez segundos ou mais.

Chinneck (1988) apresenta algumas dicas interessantes para o texto:

- a) Ter sempre em mente o *background* do leitor. Deve-se saber qual a capacidade de compreensão que o leitor-alvo do texto terá. O texto deve ser então suficientemente informativo a esse leitor-alvo. Não detalhar demais conceitos que seriam triviais e não deixar de explicar conceitos que não são provavelmente de conhecimento do leitor-alvo.
- b) Não fazer com que o leitor tenha de dar duro. Sabendo quais são as obrigações do aluno no texto (deixar claro o problema de pesquisa, mostrar que ele ainda não tinha sido resolvido, mostrar que valia a pena resolver o problema e mostrar que o aluno efetivamente resolveu o problema), deve-se deixar o texto o mais acessível possível. Quanto mais dificuldade os avaliadores tiverem para encontrar as respostas às questões fundamentais sobre o trabalho, pior será sua impressão sobre o texto e maior a probabilidade que exijam grandes mudanças no texto final.
- c) Escrever de forma que seja impossível ser mais claro. Deve-se escrever cada frase com muito cuidado, verificar se ela faz sentido e se apresenta alguma informação útil de forma clara. Deve-se verificar se cada termo usado em cada sentença já foi devidamente explicado no nível de compreensão do leitor-alvo e se todas as possíveis ambiguidades foram eliminadas.

- d) Lembrar-se de que a monografia não é uma história. Ela não é uma cronologia das coisas que o aluno tentou fazer, mas um documento formal que apresenta resultados de uma pesquisa.
- e) Evitar declarações fortes como “o software é a parte mais importante de um sistema computacional”, as quais são apenas uma opinião e não uma informação substanciada na literatura corrente. Os examinadores provavelmente pegarão frases como essas e perguntarão “você pode demonstrar que o software é a parte mais importante de um sistema computacional?”.

Moro (2009) apresenta os sete pecados capitais do texto científico:

- a) “Frases longas (repletas de vírgulas ou não!)”. Sempre que forem detectadas frases muito longas com várias orações coordenadas, deve-se procurar dividi-las em frases menores, usando pontos para isso. Mas deve-se cuidar para que cada frase lida individualmente faça sentido, tendo sujeito, verbo e objeto, quando for o caso.
- b) “Erros ortográficos”. Nada desqualifica mais um autor do que erros ortográficos, ou seja, palavras mal escritas. Um bom conteúdo pode até passar despercebido se o autor cometer erros desse tipo.
- c) “Tradução literal e *imbromation*”. Um autor que não domine a língua inglesa deve procurar a ajuda de um tradutor e, na maioria dos casos também, de um revisor profissional dessa língua. Textos em inglês que são traduções literais do português como, por exemplo, “*the cow went to the swamp*”,⁷ “*between hundred beat*”⁸ e “*I am with you and don ’t open*”⁹ são motivo de piada.
- d) “Imagens/tabelas ilegíveis”. Letras muito pequenas ou borradasservem para comunicar alguma coisa? Legendas que coloridas são perfeitas, mas quando impressas em preto e branco ficam indistinguíveis também devem ser revisadas e evitadas.

⁷ Tradução literal: “A vaca foi pro brejo”.

⁸ Tradução literal: “Entre sem bater”.

⁹ Tradução literal: “Estou contigo e não abro”.

- e) “Erros gramaticais (paralelismo, concordância, conjugação, crase)”. Erros de concordância acabam sendo muito comuns devido a revisões malfeitas do texto. Algumas vezes muda-se o sujeito de uma frase do masculino para o feminino ou do singular para o plural e deixa-se o verbo ou complementos como estavam. Por exemplo, a frase original estava assim: “o método foi devidamente demonstrado”. Uma revisão trocou o termo “método” por “passos” e o texto final ficou assim: “os passos foram devidamente *demonstrado*”. O erro aparece apenas na última palavra.
- f) “Cópia literal”. Se houver cópia literal de outros textos sem uso de aspas e citação da fonte incorre-se em plágio.
- g) “Blábláblá (encher linguiça)”. Quantas páginas deve ter uma monografia? Algumas vezes alunos que acham que seu trabalho está curto e resolvem preencher páginas com textos que não informam nada apenas para dar mais sensação de volume.

Além disso, Moro (2008) apresenta algumas dicas para a produção de textos de boa qualidade. Entre elas destacam-se:

- a) Usar revisores automáticos de texto, embora nada substitua uma leitura atenta por parte do autor, seu orientador e eventualmente por terceiros também.
- b) Dividir os parágrafos cuidadosamente. Cada parágrafo deve apresentar uma ideia central que pode ser introduzida e comentada no mesmo parágrafo. Mas quando se introduz uma nova ideia, usualmente inicia-se um novo parágrafo. Parágrafos longos demais devem ser evitados.
- c) Uma seção ou capítulo devem ser formados por mais de um parágrafo. Sessões numeradas são formadas por texto. É um erro de estilo criar uma seção (por exemplo, “3.2.1 Exemplos de algoritmos”) e preenchê-la apenas com uma lista de itens. A seção sempre inicia com um texto. Listas de itens podem ser parte da seção, mas nunca sua totalidade.

- d) Cada frase deve ter um sujeito e um verbo. Apenas os títulos de sessões, figuras e tabelas podem ser compostos por sentenças sem verbo (por exemplo, “Testes Finais”). Mas frases incluídas no texto devem sempre ter pelo menos um verbo. Por outro lado, o texto de uma seção não pode ser continuação do título da seção. Por exemplo, seria errado dar a uma seção o título “3.2.1. Testes Finais” e iniciar o parágrafo imediatamente seguinte por “Foram realizados a contento”. O correto seria iniciar o parágrafo com uma frase com sujeito e verbo que possa ser lida independentemente do título da seção, no caso, “Os testes finais foram realizados a contento”.
- e) Siglas esclarecidas. Sempre que for usada uma sigla pela primeira vez, ela deve ser definida por extenso. Mesmo que ela apareça na lista de abreviações no início da monografia, deve ainda assim ser apresentada por extenso no texto na primeira vez em que for usada. Isso vale inclusive para siglas bastante famosas em determinadas áreas (como, por exemplo, XP ou RUP no caso de engenharia de software).

Hexsel (2004) acrescenta mais algumas sugestões:

- a) Destacar termos usando *italico* e não **negrito**, pois o primeiro tem efeito mais agradável do que o segundo.
- b) Usar gráficos planos, que são mais claros do que os pseudotridimensionais.
- c) Evitar anglicismos sempre que possível, por exemplo, usando “enlace” ou “ligação” em vez de *link*.
- d) Inserir as referências bibliográficas de forma que não atrapalhem o fluxo do texto, por exemplo, no final de frases (antes do ponto).

Devido ao costume de trabalhar com linguagens de programação, por vezes, alunos de computação também esquecem como usar apropriadamente os sinais de pontuação. Em um texto nunca se coloca espaço antes de um sinal de pontuação, mas sempre se coloca

espaço após o sinal se houver uma palavra em seguida. Veja a seguir exemplos e contraexemplos.

- a) Esta é uma frase. O ponto final da frase anterior deveria ter um espaço depois dele.
- b) Esta é uma frase. O ponto final da frase anterior tem um espaço antes que não deveria existir.
- c) Esta é uma frase. O ponto final está bem colocado.

No caso de parênteses, colchetes e chaves, nunca deve haver espaço na parte interna desses símbolos. Se do lado de fora houver uma palavra, então se usa espaço. Porém, se após um fecha parêntese ou similar houver um sinal de pontuação, então não se usa espaço. Exemplos a seguir:

- a) Este texto (corretamente formatado) está entre parênteses.
- b) Este outro(mal formatado) ficou com o abre parêntese colado na palavra anterior.
- c) Este aqui (mal formatado) tem um espaço a mais depois do abre parêntese.
- d) Finalmente (bem formatado), este aqui suprime o espaço após o fecha parênteses devido à vírgula.

Na dúvida, o aluno sempre deve procurar ler bastante e ver como os textos são formatados e não inventar formatações novas.

Um último aspecto que consiste em erro bastante comum em monografias em computação é o uso inadequado de letras maiúsculas no início de palavras. No geral, há duas regras que devem ser observadas. Em primeiro lugar, títulos de capítulos ou seções do texto usualmente não levam sinal de pontuação no final, mas os substantivos, adjetivos e a maioria dos verbos devem iniciar em maiúscula. Seguem exemplos:

Título do Trabalho com Letras Maiúsculas Corretamente Usadas

Título do Trabalho com Ponto Final que não Deveria Existir.

Título do trabalho sem as letras que deveriam estar maiúsculas

Exceções são os verbos de ligação ser, estar, ter etc. que não devem iniciar em maiúscula quando aparecerem no meio de um título.

Um título somente terá pontuação final se for uma interrogação ou exclamação. Por exemplo:

Como fazer um título correto?

Agora, em relação ao uso da letra maiúscula no texto em geral, usa-se apenas no início de frases (novo parágrafo ou após um ponto), ou no caso de nomes próprios, ou ainda no caso de substantivos que representem conceitos com um único exemplar, como, por exemplo, Humanidade e Universo. Outro caso em que se usa maiúscula é para nomear as ciências, como Ciência da Computação, Física, Matemática etc.

É errado então usar letra maiúscula aleatoriamente, como muitos alunos fazem. Veja estes exemplos:

- a) O Sistema que foi desenvolvido... (“sistema” deveria estar em minúscula).
- b) A Engenharia de Software trata de conceitos... (está correto, pois Engenharia de Software pode ser considerada uma ciência).
- c) Esta Dissertação foi escrita... (“dissertação” deveria estar em minúsculas).

Essas são regras de estilo que em geral não afetam o conteúdo de um trabalho, mas ajudam a torná-lo mais agradável de ler.

Capítulo 6

Escrta de Artigo Científico

O artigo científico é a forma academicamente reconhecida de divulgação de um trabalho de pesquisa. No nível de mestrado, e especialmente no nível de doutorado, espera-se, e em alguns casos, exige-se, a publicação de um artigo em um evento ou periódico de boa qualidade. Este capítulo apresenta algumas dicas sobre escrita de artigos e finaliza com comentários sobre o sistema brasileiro de avaliação da qualidade de veículos de publicação em Ciência da Computação, o Qualis-CC, definido por uma comissão de especialistas vinculada ao Ministério da Educação.

Todas as recomendações relativas ao texto da monografia valem também para artigos científicos. Adiciona-se ainda a recomendação de que o artigo deve ser muito mais sucinto do que a monografia. Então a clareza e objetividade são muito mais críticos em um artigo do que em uma monografia.

6.1. Autores

Ao contrário da monografia, que é um trabalho individual, o artigo científico muitas vezes será um trabalho colaborativo. Em países como o Brasil, considera-se estranho que um artigo publicado a respeito de uma monografia não possua o respectivo orientador como co-autor.

Segundo Moro (2008), não existe um consenso sobre qual deve ser a ordem em que os autores devem aparecer no artigo. Em alguns casos recomenda-se, por simplicidade, o uso da ordem alfabética.

Porém, o primeiro autor de um artigo costuma ser considerado o mais importante. E o uso da ordem alfabética pode criar certa confusão. Pode ser difícil, no caso de trabalhos cooperativos entre pesquisadores ou instituições, chegar a uma conclusão sobre quem é o autor principal. Mas no caso de artigos inspirados em monografias essa decisão é bem mais simples:

- a) Em primeiro lugar, deve aparecer o nome do aluno que foi o autor da monografia em questão, visto que esse é considerado como o autor principal do trabalho monográfico que deu origem ao artigo.
- b) Em segundo lugar, deve aparecer o nome do orientador, já que este tem também uma grande responsabilidade sobre o trabalho original, bem como sua revisão final.
- c) Em terceiro lugar, poderão aparecer, se houver, co-orientadores, colegas ou outros pesquisadores que tenham de alguma forma contribuído com o texto.

O usual é que sejam considerados co-autores apenas pessoas que participaram da confecção do *texto*, embora algumas vezes acabem sendo citadas também pessoas que tenham ajudado na coleta de dados ou na implementação dos protótipos que deram origem ao trabalho. Não se recomenda, entretanto, uma lista muito grande de co-autores, porque pode dar a impressão de que se tenta artificialmente melhorar o currículo de alguém.

6.2. Motivação para Escrever

Não se deve escrever um artigo se você não souber o que vai dizer. Parece senso comum, mas às vezes estudantes começam assim um trabalho.

Primeiramente o candidato a autor deve pensar em uma frase que resuma a contribuição do artigo e então desenvolver essa frase apresentando antecedentes, detalhamento e consequências dessa ideia. Caso não consiga pensar em uma frase que resuma o artigo, então o autor estará em maus lençóis. Talvez seja interessante parar

e pensar mais um pouco, ou ainda, desenvolver melhor a pesquisa, organizar as ideias, procurar o orientador e então tentar novamente.

Um artigo consiste na comunicação de uma ideia. Não se deve falar por falar. Não se deve escrever à toa, nem desperdiçar preciosas linhas com informação irrelevante ou desconexa.

Um artigo científico em geral é um texto curto, com 8 a 12 páginas. Raramente um artigo terá mais do que 16 páginas. Então, o artigo não pode e não deve ser um tratado sobre uma área do conhecimento, mas a transcrição objetiva e precisa de uma ideia de pesquisa, do desenvolvimento que a validou e das suas consequências no mundo.

Assim, o artigo deve enfatizar o resultado concreto obtido na pesquisa. Importante também é mostrar ao leitor como o autor chegou nesse resultado e, afinal de contas, qual é o problema real que o resultado resolve.

Para a melhor compreensão do assunto, o leitor deve receber, no início do artigo, um resumo dos principais conceitos, apenas aqueles imprescindíveis para compreender os resultados. Nota-se que estabelecer quais são os conceitos necessários para a compreensão do artigo depende muitas vezes do tipo de veículo de publicação. Por exemplo, para escrever um artigo sobre redes neurais aplicadas a sistemas de previsão de cotações da bolsa e publicar em um evento de Computação, não é necessário definir o que são redes neurais, basta mencionar qual modelo foi usado, por que foi escolhido e colocar uma citação bibliográfica. Por outro lado, os conceitos de Economia usados possivelmente terão de ser mais detalhadamente explicados para que um leitor da área de Computação possa entender a pesquisa. Por outro lado, se a publicação for ocorrer em um evento de Economia, acontece o contrário. Deve-se explicar claramente o que é uma rede neural e como ela funciona e pode-se ser mais econômico em relação aos conceitos de Economia.

Snyder (1991) propõe que o autor de um artigo faça a si mesmo algumas perguntas antes de submeter o artigo a um evento. A primeira dessas perguntas é: “por que estou escrevendo este arti-

go?”. Se a resposta for “para documentar o que tenho feito nos últimos dois anos”, o autor corre um sério risco de ter seu trabalho rejeitado. Poucas pessoas estarão interessadas em saber o que alguém fez nos últimos dois anos. Se o objetivo for documentar essas atividades, o autor deveria escrever um relatório de pesquisa, não um artigo.

Outra resposta errada seria “para melhorar meu currículo”. Essa até poderia ser a motivação inicial para alguém escrever um artigo, mas dificilmente motivará outros a aceitarem o texto para publicação. Então outra resposta tem de ser buscada.

A resposta correta para a questão estaria na linha de “comunicar uma ideia a alguém”. Então, ainda segundo Snyder, as questões seguintes seriam: “O que o meu artigo está tentando comunicar?” e “Qual é o público-alvo de meu artigo?”. Se o autor não conseguir responder categoricamente a essas duas questões, então há uma grande chance de ser um artigo fraco.

Outra questão é que um artigo focado tem mais chance de ser bom do que um artigo disperso. É melhor tomar uma ideia e trabalhá-la claramente no artigo do que passar superficialmente sobre um conjunto de grandes ideias que o autor teve ao longo da vida. No caso de várias ideias, é melhor escrever vários artigos.

Depois de saber claramente o que o artigo está comunicando, o autor ainda deveria se perguntar se vale a pena apresentar essa comunicação, ou seja, ele está realmente comunicando uma nova ideia ou é apenas uma nova maneira de apresentar uma velha ideia, que já é bastante conhecida? É uma ideia relevante ou é trivial? É apenas uma conjectura ou uma informação baseada em sólidas evidências?

6.3. Trabalhos Correlacionados

Não só um artigo, como também a monografia final deverão mencionar trabalhos correlacionados. Não se aceitam desculpas do tipo “ninguém nunca fez algo parecido” (pois “parecido” é um termo difuso e bastante flexível), ou “não encontrei nada na bibliografia sobre isso” (pois o leitor vai *sempre* pensar que não se procurou o suficiente). Um aluno inteligente vai evitar cair nessa armadilha!

Mas, o que fazer se efetivamente não for encontrado nada? Por exemplo, se alguém está pesquisando as aplicações de redes neurais na bolsa e por mais que procure não encontra outro trabalho sobre esse assunto. Como escapar dessa armadilha?

A primeira dica é delimitar claramente o escopo da pesquisa bibliográfica. Em vez de dizer “procurei e não achei nada”, o que é um suicídio acadêmico, deve-se usar uma abordagem sistemática, como a apresentada anteriormente: escolher os melhores periódicos e eventos na área de redes neurais e/ou de bolsa de valores, no caso, fixar um período de tempo razoável para realizar as pesquisas nesses veículos, por exemplo, os últimos cinco ou dez anos, e finalmente fazer a pesquisa sistematicamente nesses veículos no período de tempo delimitado verificando título e resumo dos artigos. Os artigos podem então ser classificados em “não relacionado”, “moderadamente relacionado” e “fortemente relacionado”. Se o aluno não encontrar nenhum artigo fortemente relacionado, então ele pode mencionar no seu trabalho os moderadamente relacionados. Ele deve dizer sempre onde pesquisou. Pode dizer, por exemplo, “Ao proceder à revisão bibliográfica nos periódicos x, y e z de 2003 a 2009 não foram encontrados trabalhos que abordassem o uso de redes neurais para previsões da bolsa de valores. Porém os seguintes trabalhos moderadamente relacionados foram encontrados: (a) Fulano de tal (200x) apresenta uma abordagem que utiliza redes neurais para prever a demanda de energia em transformadores elétricos de rua, (b) etc.”.

Agindo dessa forma o leitor fica mais tranquilo. Ninguém pode ser culpado por não ter encontrado um artigo sobre o assunto que eventualmente tenha sido publicado em um obscuro evento na Groenlândia ou no Marrocos, pois ninguém é capaz de saber tudo, não? Mas o aluno estará bem respaldado se esse artigo existir e se algum avaliador o mencionar, pois ele delimitou sua revisão bibliográfica aos melhores periódicos e eventos da área. Se aquele trabalho da Groenlândia ou Marrocos é realmente bom, por que não foi submetido aos bons veículos de publicação reconhecidos pela comunidade?

6.4. A Contribuição do Artigo

Sobre a contribuição do artigo pode-se recomendar o seguinte:

- a) Não ser modesto!
- b) Não exagerar!

Para Platão (2006), os vícios estão nos extremos, e a virtude, no equilíbrio. Então o autor deve ser realista em relação aos resultados e à contribuição de seu artigo.

O autor deve convencer o comitê avaliador de que seus resultados estão corretos. Não se deve esperar que eles simplesmente acreditem ou simpatizem com as ideias. Eles precisam ser convencidos. O comitê avaliador fará uma leitura crítica do artigo, procurando por quaisquer lapsos que possam invalidar o trabalho. Para ser publicado, o trabalho precisa passar por esse crivo. Então, um convencimento crítico é necessário. Deve-se apresentar provas, evidências e exemplos que possam ajudar.

Finalmente, a contribuição do artigo deve estar clara desde o *abstract* ou resumo. Ela não pode ser deixada para o final. Os resultados devem ser apresentados logo no início do artigo para interessar o leitor. Depois o autor pode ir explicando como chegou a eles.

6.5. Tipos de Artigos

Há vários tipos ou estilos de artigos, cada um dos quais com suas próprias características e seus próprios veículos de publicação. Esta seção destaca alguns.

6.5.1. Artigo Teórico

Um artigo teórico basicamente apresenta um conjunto de definições, conhecido como “teoria”, e posteriormente passa a provar propriedades lógicas desse conjunto. Exemplos de técnicas de prova usadas neste tipo de artigo são *indução* (matemática ou estrutural) e *redução ao absurdo*. Existem bons livros de lógica que podem ajudar o leitor interessado a estruturar seu trabalho de acordo com esse paradigma.

Em um artigo teórico, cada afirmação precisa ser colocada cuidadosamente, e todas precisam ser fundamentadas. Pode-se fundamentar uma afirmação através de referência bibliográfica, prova lógica, relato de observação direta, ou ainda como hipótese ou definição.

Porém, não adianta fazer apenas a abordagem teórica de uma questão. É preciso mostrar qual o problema real que essa teoria resolve. Então não adianta criar uma teoria bonita que não sirva para nada. Mesmo um artigo teórico deve ter algum tipo de consequência no mundo real.

6.5.2. Relato de Experiência

Um relato de experiência conta uma história informativa sobre um experimento e suas observações. Este relato deve mostrar como a situação observada se reflete em situações mais gerais. Ou seja, o relato de experiência deve, sempre que possível, não se fixar à instância específica sendo observada, mas apresentar a possível generalização das observações para outras situações.

Deve-se evitar entrar em detalhes irrelevantes sobre o experimento. Apenas as informações necessárias para compreender ou validar o relato devem ser incorporadas.

O relato deve se concentrar nas ideias, e não no experimento em si. Ou seja, o relato de experiência não é uma narrativa sobre todos os passos que o autor deu no caminho da observação, mas uma estruturação das ideias aprendidas durante a observação. Essa estruturação, então, raramente será apresentada de forma temporal narrativa, mas sim de forma dissertativa, organizada por conceitos e suas implicações.

6.5.3. Artigos sobre Métodos

Artigos sobre métodos são especialmente comuns na Ciência da Computação. Um bom artigo sobre método não pode ser simplesmente uma apresentação do método. Ele deve se concentrar nas vantagens

que o método sendo apresentado tem sobre outros anteriormente propostos para o mesmo problema ou problemas semelhantes.

Um artigo sobre métodos deve ter um objetivo informativo bastante claro, ou seja, o novo método deve ter um ponto focal. O artigo deve relacionar as vantagens do novo método sobre abordagens anteriores. Então aqui, a comparação bibliográfica é fundamental para a aceitação do artigo.

Recomenda-se que um artigo sobre métodos permita a aplicação do método em um projeto real. Quando isso não for possível, uma referência a um texto mais completo pode ajudar. Além disso, como o artigo deve conter uma comparação com métodos anteriores, ele deve deixar bem claro qual a métrica usada para a comparação. Uma comparação subjetiva terá pouco valor, especialmente se o autor não deixar claro de onde vieram os resultados. Por exemplo, uma tabela de características, na qual se compararam diversos métodos avaliando cada um com o valor “atende”, “não atende” ou “atende parcialmente”, não terá muito valor a não ser que o autor deixe bastante claro como essas avaliações foram obtidas, de forma que possam ser repetidas por observadores independentes.

O artigo deve ser mais do que uma apresentação do método com explanação. O importante não é listar os procedimentos operacionais do novo método, mas apresentar as ideias que o novo método incorpora.

Além do mais, o artigo deve ser equilibrado. Novas abordagens normalmente não são a panaceia universal. Se o novo método tem vantagens, possivelmente também terá limitações, que devem ser descritas e analisadas no artigo.

6.6. Veículos de Publicação

Os resultados de uma pesquisa podem ser publicados em uma série de veículos reconhecidos pela comunidade científica. Quanto maior o impacto do veículo, ou seja, o número de pessoas que ele efetivamente atinge, maior a dificuldade relativa de se conseguir publicar um artigo ali.

Então, dependendo da real contribuição e inovação do trabalho, diferentes veículos de publicação deverão ser escolhidos. O estudante não deve ficar desestimulado se sua primeira tentativa de publicação for frustrada. Usualmente avaliações críticas consistentes sobre o trabalho são entregues explicando os motivos da recusa. Essas avaliações podem e devem ser usadas como um estímulo e um guia para produzir uma versão melhorada do artigo visando à publicação em outro veículo.

Existem veículos que são bastante específicos de cada área, como, por exemplo, o Simpósio Brasileiro de Qualidade de Software ou o IEEE Transactions on Software Engineering. Mas estes não são os únicos veículos em que se pode publicar. Há periódicos e eventos mais genéricos que em geral aceitam artigos de todas as subáreas da Computação, como, por exemplo, o SEMISH, que ocorre junto ao Congresso da Sociedade Brasileira de Computação e o periódico *Communications of ACM*. Ambos são considerados bons veículos.

Em relação ao estilo de veículo, pode-se fazer as seguintes distinções:

- a) Periódico: É considerada a publicação mais importante por todas as áreas da ciência. Os melhores artigos usualmente são destinados aos periódicos mais reconhecidos dentro de cada área. A Ciência da Computação no mundo todo, entretanto, conta com poucos periódicos quando comparada a outras áreas do conhecimento. No Brasil, periódicos de qualidade em Ciência da Computação são praticamente inexistentes.
- b) Eventos ou conferências: A Ciência da Computação privilegia a publicação em conferências, o que muitas vezes cria problemas em relação à avaliação relativa com outras áreas como Física e Química, cujos pesquisadores publicam quase que exclusivamente em periódicos. Embora não tão valorizados quanto periódicos por outras áreas, as conferências em Computação podem ter peso relativo bastante relevante na produção científica de um pesquisador.

- c) *Workshops* e seminários: Em geral, são eventos satélites de conferências maiores. Como são normalmente muitíssimo restritos em termos de abrangência temática e número de participantes, são considerados publicações de menor impacto, embora alguns *workshops* tenham se firmado como boas conferências ao longo de uma história consistente de boas edições.
- d) Livros e Capítulos de Livros: Embora bastante valorizados, livros e capítulos de livros são publicações que usualmente não resultam de teses e monografia (com poucas exceções), visto que, em geral, o objetivo desse tipo de publicação é apresentar um conteúdo didático para compreensão por parte de um público bem mais amplo do que o conjunto de pesquisadores de uma determinada área da ciência.

Existe uma diferença fundamental no estilo de processo de revisão de eventos e periódicos. Como os eventos ocorrem em data predeterminada, as publicações em geral são submetidas até um determinado prazo (*deadline*) e então avaliadas por um comitê de programa. Trata-se normalmente de um processo competitivo em que os melhores artigos são aceitos para publicação, com algumas poucas sugestões de modificação no texto. A maioria dos eventos, especialmente os mais bem conceituados, aceitará uma porcentagem relativamente pequena dos artigos submetidos. Então, no caso de envio de artigo a um evento, o autor terá apenas uma chance de publicar. O artigo deve estar pronto e em condições de concorrer com outros artigos. Se estiver entre os melhores, será publicado, caso contrário, será rejeitado.

Já o processo de revisão em periódicos acontece de forma diferenciada. Exceto no caso de números especiais temáticos, os periódicos funcionam com regime de envio contínuo, ou seja, não há *deadline*. Assim, um artigo eventualmente aceito no periódico entra em uma fila e será publicado quando chegar a sua vez. O processo de revisão, então, pode ser bem mais interativo do que no caso de eventos.

Um artigo submetido será revisado pelo comitê editorial, e, possivelmente, várias sugestões serão feitas ao texto antes que este possa ser aceito para publicação. Poderão, inclusive, acontecer várias rodadas de avaliação do texto, em que os revisores solicitam modificações e os autores as incorporam ao texto, se possível. Esse processo interativo de revisão de um texto pode até levar anos, em alguns casos, mas procura garantir que o material final estará adequado ao público leitor do periódico, de acordo com os critérios do comitê editorial.

Deve-se ter em mente também que o processo de revisão em periódicos é muito mais detalhado do que em conferências. No caso de conferências, os revisores trabalham com *deadlines*, e às vezes recebem um grande fardo de avaliações para fazer, estando, dessa forma, mais sujeitos a cometer erros de avaliação. No caso de periódicos, a revisão é feita com mais tempo e, portanto, é bem mais detalhada.

Além do tipo e publicação, a abrangência também deve ser considerada. Todos os diferentes veículos têm uma abrangência estimada, que pode ser:

- a) Internacional: veículos publicados em língua inglesa que são distribuídos ou que contam com a participação de autores de vários países, sem predominância de nenhuma nação, como, por exemplo, a maioria dos periódicos e conferências da IFIP, ACM e IEEE.
- b) Nacional: não são apenas os veículos publicados no Brasil, como se poderia pensar. Existem periódicos e eventos de abrangência nacional publicados em outros países. Caracteriza-se como um veículo nacional aquele que é publicado em uma língua diferente do inglês ou que, embora publicado em inglês, tenha participantes predominantemente de um único país ou região, por exemplo, a Conferência Latino-americana de Informática (CLEI) e a maioria dos simpósios da SBC.

- c) Regional: são veículos que abrangem apenas uma fração de um país, por exemplo, um estado brasileiro e ou região geográfica. Exemplos desse tipo de veículo são os anais da Escola Regional de Bancos de Dados e o Seminário Catarinense de Imagens Médicas.
- d) Local: são veículos publicados por uma única universidade ou faculdade. Em geral, contam com um comitê avaliador local ou de pouca abrangência, e a maioria dos autores de artigos pertence à própria instituição. Ex.: Semana de Informática da Universidade de Água Suja do Norte.

Existem então as combinações entre os diversos tipos de publicação e sua abrangência, como, por exemplo, periódico internacional, ou evento regional etc. O documento de área da Ciência da Computação na CAPES distingue os seguintes tipos: Periódico Internacional (PI), Periódico Nacional (PN), Conferência Internacional (CI), Conferência Nacional (CN), Livro Científico Nacional (LCN), Livro Didático Nacional (LDN), Livro Científico Internacional (LCI), Livro Didático Internacional (LDI), Capítulo de Livro Nacional (CLN) e Capítulo de Livro Internacional (CLI).

Deve-se prestar atenção ao seguinte: algumas conferências publicam seus anais como livros, com ISBN, inclusive. Não existe consenso na comunidade sobre se isso caracteriza uma publicação em livro ou se ainda seria publicação em evento. Porém, pela lógica, deveria ser considerada como um livro apenas a obra literária que foi proposta e construída com esse propósito. Um livro tem uma estrutura sequencial e lógica, o que não ocorre normalmente com conferências, porque estas publicam os melhores artigos. Mas estes são submetidos de forma independente pelos seus autores, baseado no trabalho de pesquisa que tenham efetuado, sem uma sequência lógica predeterminada.

A era da informática, entretanto, tem facilitado tanto a criação de publicações que possivelmente as fronteiras claras que havia antigamente entre os diferentes tipos de publicação talvez não mais existam no futuro.

6.7. Ética no Envio de Artigos

Ao enviar um ou mais artigos, o autor deve estar atento aos aspectos éticos considerados pela comunidade científica.

Em primeiro lugar, a publicação de um artigo em um evento ou periódico implica um compromisso de não enviar ou publicar novamente esse mesmo artigo ou partes dele em outro local.

Outro aspecto não tão evidente, mas também importante, é que um artigo, mesmo que ainda não tenha sido publicado, deve ser submetido a apenas um veículo de cada vez. Ou seja, não se deve cair na tentação de enviar o mesmo artigo a vários eventos ao mesmo tempo na esperança de que um deles vai aceitar. Snyder (1993) aponta que “*simultaneous submission without notice is considered highly unethical.*”¹ Alguns eventos ou periódicos até aceitam envio simultâneo a outro veículo, desde que seja explicitamente informado.

A forma correta de se tentar uma publicação é, portanto, enviar uma primeira versão do artigo a um veículo e aguardar o resultado. Sendo aprovado para publicação, ótimo! Caso contrário, os avaliadores enviarão juntamente com a informação da recusa de publicação uma lista de motivos. Essa lista vai conter sugestões para melhoria do trabalho que poderão ser usadas pelo autor de modo a aprimorar o artigo antes de enviá-lo a outro veículo.

Ainda outro aspecto a ser considerado é o caso da publicação de dois ou mais artigos semelhantes sobre o mesmo assunto. Não é considerado ético fazer várias versões de um mesmo artigo e enviá-las a diferentes veículos. Cada artigo apresenta uma ou mais ideias de pesquisa que foram avaliadas. Assim, artigos, mesmo que diferentes na forma, mas que apresentam as mesmas ideias, são considerados uma forma de autoplágio, e, portanto, antiéticos. É possível, por outro lado, publicar vários artigos a partir do mesmo trabalho de pesquisa, desde que cada um trate de aspectos diferentes do trabalho, ou seja, cada artigo deve explorar uma ideia de pesquisa diferente ou, no caso de publicações mais antigas, pode-se gerar um novo artigo aprofundando ou estendendo os resultados já apresentados.

¹ Tradução: “Envios simultâneos sem aviso são considerados altamente antiéticos.”

Não se pode copiar diretamente o texto de um artigo já publicado em outro, mesmo em se tratando do próprio autor. Uma vez que o texto tenha sido publicado, ele pode ser citado em um novo artigo através de referência bibliográfica, entre aspas, como se fosse outro texto qualquer. Cada artigo deve, portanto, ser escrito com texto 100% original.

6.8. Qualis²

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), órgão do Ministério da Educação do Brasil, publica uma “lista de veículos utilizados para a divulgação da produção intelectual dos programas de pós-graduação *stricto sensu* (mestrado e doutorado), classificados quanto ao âmbito de circulação (Local, Regional, Nacional, Internacional) e à qualidade (A, B, C), por área de avaliação”. Essa lista, denominada Qualis, é usada pela CAPES para fundamentar o processo de avaliação do Sistema Nacional de Pós-Graduação.

O Qualis é organizado por área do conhecimento. Na área de Ciência da Computação, a lista denomina-se Qualis-CC e é organizada por um comitê de especialistas. Um mesmo veículo de publicação pode estar classificado em mais de uma área de conhecimento, às vezes, inclusive, com uma avaliação diferente em termos da qualidade. Isso acontece porque cada área define os seus próprios critérios de qualidade.

Na sequência serão apresentados os critérios de qualidade usados pela comissão de Ciência da Computação.

6.8.1. Periódicos

Na área de Ciência da Computação, são considerados os índices de impacto registrados no ISI/JCR (*Journal of Citation Records*) e no CiteSeer Computer Science Citation Index.

² No momento em que este livro estava sendo escrito, a CAPES estava realizando uma revisão do sistema Qualis. Possivelmente algumas das regras aqui mostradas já tenham mudado. Recomenda-se consultar www.capes.gov.br para obter uma versão atualizada dessas regras.

Os periódicos que constam no JCR são divididos em três grupos de acordo com o índice de impacto:

- a) Nível A: 60% superior.
- b) Nível B: 30% médio.
- c) Nível C: 10% inferior.

Os periódicos que constam no CiteSeer também são divididos em três grupos:

- a) Nível A: 40% superior.
- b) Nível B: 40% médio.
- c) Nível C: 20% inferior.

No caso de periódicos que constam nos dois índices, é considerada a nota mais alta entre os dois.

Além dessa regra, baseada em índices de impacto, considera-se que os periódicos das sociedades científicas internacionais ACM, IEEE, INFORMS e SIAM, mesmo que não apareçam em nenhum índice, são classificadas como B.

Periódicos nacionais só são considerados se forem indexados. Por uma deferência especial da comissão, o JCBS, *Journal of the Brazilian Computer Society*, é avaliado como Internacional C.

Periódicos de áreas afins com interface com a Ciência da Computação recebem a maior avaliação dentre as áreas afins. São consideradas áreas afins: Engenharia Eletrônica, Matemática, Matemática Aplicada, Pesquisa Operacional e Estatística.

6.8.2. Eventos Internacionais

Os principais fatores de avaliação de um evento internacional são, segundo o comitê, o índice de impacto, a tradição do evento, a sociedade promotora, a avaliação por comitê de programa internacional, sem predominância de nenhum país, e por artigo completo (eventos que avaliam artigos a partir de resumos não são bem classificados).

O JCR não avalia o impacto de eventos. Apenas o CiteSeer tem informação sobre eventos. Assim, o índice de um evento será dado

basicamente por dois critérios, o *CiteSeer* e uma combinação das qualidades mencionadas. No caso de atribuição de notas divergentes por esses dois critérios, prevalece a mais alta.

Para a avaliação pelo *CiteSeer* são usados os mesmos pontos de corte indicados anteriormente. O critério de combinação de qualidades do evento é aplicado assim:³

- a) Um evento é de nível A se for “patrocinado (i.e., *sponsored*) por Sociedades Científicas Internacionais como: IEEE, IFIP, ACM, SIAM, INFORMS, W3C etc., desde que a publicação seja artigo completo (*full paper*) e tenha avaliação por revisores (i.e., *referees*) e conferência de tradição. Conferências de tradição devem ter tido pelo menos quatro edições. O Comitê tem constatado uma grande variabilidade na qualidade dos eventos que publicam seus anais (i.e., *proceedings*) como Lectures Notes. Sendo assim, o Comitê irá julgar caso a caso a qualidade desses eventos para efeitos de classificação”.
- b) Um evento é de nível B “se os critérios anteriores são satisfeitos, mas a conferência é considerada ainda recente, i.e., menos de quatro edições”.
- c) Um evento é de nível C se avalia artigo completo (*full paper*) utilizando revisores e cujas publicações são *impressas* pelas Sociedades Científicas como IEEE, ACM IFIP, SIAM etc., embora não sejam patrocinados por essas sociedades. Também são de nível C “conferências patrocinadas por Sociedades Científicas, mas com caráter claramente regional... (por exemplo, Asian International Conference, Pacific-rim conference, Latin-american Conference)”. Finalmente, “workshops e similares associados a Conferências Internacionais QUALIS A patrocinadas (i.e., *sponsored*) por Sociedades Científicas Internacionais como: IEEE, IFIP, ACM, SIAM, INFORMS, W3C etc., com trabalho completo, recebem classificação internacional C”, sendo, porém, analisados caso a caso.

³ Fonte www.capes.gov.br.

Veículos de publicação que não aparecem no Qualis não são necessariamente ruins. O Qualis-CC não é uma lista de todos os veículos de Ciência da Computação. Apenas os eventos e periódicos reportados pelos programas de pós-graduação aparecem no índice. Novos veículos são adicionados anualmente quando algum programa informa que um de seus pesquisadores ali publicou. Assim, eventos ou periódicos que não aparecem no Qualis podem receber uma boa classificação quando da revisão do índice. Porém, quando veículos aparecem no Qualis-CC, mas não têm nota (nem A, nem B e nem C), é porque já foram avaliados, mas não foram encontrados indícios de que possam receber nota A, B ou C.

6.8.3. Eventos Nacionais

A avaliação dos eventos nacionais é feita de forma mais subjetiva. Em primeiro lugar, porque estes não aparecem, normalmente, nos indexadores internacionais. Em segundo lugar, porque a maioria dos eventos nacionais considerados são aqueles que ocorrem no Brasil, sendo, portanto, conhecidos, e seu número não é tão grande.

Ao contrário dos eventos internacionais, espera-se que todos os eventos brasileiros de bom nível já estejam no Qualis-CC, justamente pela proximidade e pelo pequeno número destes.

Os eventos nacionais são classificados de acordo com os seguintes critérios (do documento de área):

- a) Características do evento.
- b) Comitê de programa.
- c) Número de envios.
- d) Percentual de aceitação.
- e) Apoio de entidades científicas.
- f) Tradição do evento.

Em relação às características do evento, espera-se que este seja representativo de uma subárea da Computação. Eventos com abrangência temática genérica e abrangência geográfica limitada dificilmente terão uma boa avaliação. Espera-se que o evento seja realizado no âmbito de uma sociedade científica, como, por exem-

pto, a Sociedade Brasileira de Computação (SBC). Eventos ligados a empresas nem sempre têm caráter científico. De preferência, eventos de nível A devem ter abrangência internacional, com comitê de programa sem predominância de brasileiros, página na Internet em inglês e anais publicados por uma editora internacional.

Em relação ao comitê de programa, espera-se que seja constituído de pesquisadores reconhecidos na área do evento e que se distribuam em vários países.

Em relação ao número de envios, espera-se que seja compatível com o tamanho da área de pesquisa. Um evento com poucos envios indica pouca procura e pouco interesse por parte da academia, o que não é um bom indicativo de qualidade.

O percentual de aceitação é a razão entre o número de artigos aceitos para publicação e o número de artigos enviados ao evento. Espera-se que eventos de nível A tenham menos de 35% de aceitação, e que eventos de nível B tenham entre 35% e 55% de aceitação.

O apoio de sociedades científicas internacionais como ACM e IEEE deve ser uma constante para eventos de nível A.

Finalmente, um evento novo, mesmo que bem organizado, dificilmente terá conceito máximo. Para que um evento obtenha nível A, em geral, deve ter tido pelo menos três edições anteriores organizadas regularmente e com boa qualidade. Por esse motivo também, eventos pontuais, ou seja, que não acontecem regularmente, dificilmente serão classificados.

São examinados caso a caso os *workshops* ligados a eventos nacionais de nível A, podendo receber nível C ou até superior, dependendo do caso.

Capítulo 7

Plágio

Plágio é a apropriação indevida de ideias ou textos de outras pessoas. A prática da cópia do trabalho alheio era comum e aceita entre os escribas antigos e os músicos da renascença e do barroco, mas, com o passar do tempo e com a consolidação do direito à propriedade e sua exploração, o plágio adquiriu *status* de procedimento antiético. Porém, sempre continuou acontecendo. Na era da Internet, nunca foi tão fácil copiar o trabalho alheio, porém, também nunca foi tão fácil detectar essas cópias.

Independentemente da questão da exploração comercial de direitos autorais, o plágio, no meio acadêmico, é extremamente nocivo se não for detectado, pois o plagiador apresenta um resultado que não é de sua autoria e recebe um título que não merece. Nessa condição, ele próprio pode ser prejudicado, ao não dominar conhecimentos que seriam necessários para exercer sua profissão, ou, pior ainda, se exercer sua profissão, prejudicará a terceiros, por apresentar soluções inadequadas, de acordo com sua própria incompetência.

Há pelo menos duas formas de plágio: a cópia literal de textos de outras pessoas constituindo integral ou parcialmente um trabalho que deveria ser do autor, e a cópia de ideias, em que o autor, apesar de não repetir as palavras como foram escritas, apresenta as mesmas ideias, na mesma sequência lógica, como se fossem suas.

Não é considerado plágio o uso de ideias de terceiros, desde que a fonte apareça claramente identificada. No caso de cópias literais, devem aparecer entre aspas.

Conta-se que certa vez um estudante de doutorado plagiou uma tese, copiando o texto integral de outro autor e trocando apenas o nome do autor original pelo seu próprio nome. Durante a defesa, um dos membros da banca, o convidado externo, elogiou copiosamente o trabalho durante vários minutos. No final acrescentou: “mas você não pode obter o doutorado com esta tese, porque este trabalho é meu”. O plagiador foi tão displicente que sequer olhou o nome de quem estava plagiando e acabou convidando essa pessoa para a banca.

Neste capítulo serão discutidos casos de plágio e suas consequências, e serão mostradas também algumas situações nas quais fica caracterizado o desvio de conduta dos plagiadores e seus apologistas.

7.1. Antecedentes

Até a invenção da imprensa, as obras escritas eram reproduzidas por escribas e copiadores profissionais. Naquela época, apenas esses profissionais da cópia é que eram remunerados pelo seu trabalho. Ao autor cabia apenas o mérito pela obra (mas às vezes nem isso).

Com a invenção da imprensa, no século XV, a cópia de textos se tornou uma atividade de massa e, por esse motivo, suscitou a questão da proteção jurídica ao trabalho do autor. Não apenas a questão de proteger o direito ao patrimônio da obra, mas também da proteção à sua integridade.

No início, a censura caminhou junto com a proteção ao direito autoral. Em 1662, por exemplo, na Inglaterra, o Licensing Act proibia a impressão de qualquer livro que não tivesse sido previamente autorizado.

Já o Copyright Act, de 1709, protegia por até 21 anos a propriedade intelectual e patrimonial de obras impressas.

Na França, após a Revolução de 1789, os valores iluministas passaram a imperar e, com isso, a primazia do autor sobre a obra intelectual. A proteção do direito autoral passou a ser por toda a vida do autor e inclusive transmitida aos seus herdeiros legais.

No Brasil, a primeira menção sobre proteção dos direitos autorais surge em 1827, na lei que criava os cursos jurídicos. Em 1830 a matéria é regulamentada através da promulgação do código de direito criminal.

7.2. Proteção aos Direitos Autorais

O Brasil tem uma das legislações mais fortes em relação a direitos autorais no mundo. Nos Estados Unidos, por exemplo, a proteção ao direito autoral depende de registro. No Brasil, o registro não é precondição para a proteção do direito autoral. Basta a prova da autoria. Pode-se, por exemplo, lacrar a obra literária ou técnica em um envelope do correio e remetê-la para si mesmo ou para uma pessoa de sua confiança. Não se deve abrir esse envelope, exceto em juízo. O carimbo do correio é uma prova da data da produção, ou pelo menos de que na data do envio a produção estava com a pessoa em questão. Assim, no caso de um processo por plágio, o autor tem como provar que o texto estava com ele numa determinada data. Se essa data for anterior à produção tida como plágio, então boa parte do processo terá sido resolvido em favor do autor.

Nos cursos de graduação e programas de pós-graduação, é grande a preocupação com o plágio. Em especial, porque nunca foi tão fácil copiar textos usando a Internet. Por outro lado, também, nunca foi tão fácil detectar cópias. Sequer são necessárias ferramentas sofisticadas para isso. Basta um site de busca e uma cópia da monografia. Três ou quatro palavras são escolhidas aleatoriamente em qualquer ponto da monografia e pesquisa-se no site de busca. Dificilmente a ocorrência dessas palavras juntas em um texto será mera coincidência. Com esse recurso é possível descobrir a grande maioria dos casos de plágio.

Textos traduzidos de outras línguas para o português são mais difíceis de detectar, mas nem tanto. Normalmente, o próprio estilo de escrita permite perceber que um texto é uma tradução e não um texto originalmente escrito em português. Por exemplo, brasileiros dificilmente usam a palavra “eventualmente”, ou “fornece” em tex-

tos científicos quando escrevem em português, mas traduzem essas palavras literalmente a partir de textos em inglês, em que são mais comuns (no caso, respectivamente, “*eventually*” e “*provides*”).

Acredita-se que boa parte do plágio acadêmico ocorria porque os alunos não eram corretamente orientados em relação ao que podia e o que não podia ser copiado. Então, para que não haja dúvidas, aqui vai a resposta: *nada* pode ser copiado, a não ser que seja colocado entre aspas e com a citação da fonte bibliográfica. Mesmo assim, deve-se agir com parcimônia, pois as citações não podem predominar em um trabalho científico. É necessário haver a contribuição do autor.

Certa vez um aluno de especialização entregou uma monografia com 20 páginas. Na primeira página ele escreveu algo como “Outro dia li um artigo interessante na Internet:”. Em seguida ele abria aspas e incluía o dito artigo literalmente na sua monografia. Vinte páginas depois, após fechar aspas, ele concluía dizendo: “Por isso achei o artigo tão interessante”. O aluno foi reprovado. Ele questionou dizendo que, como tinha colocado o texto entre aspas e citado a fonte, isso não era plágio. Mas foi reprovado mesmo assim, pois tirando o que estava entre aspas não sobrava praticamente nada do trabalho dele. Assim, o trabalho foi considerado insuficiente para a obtenção do título.

7.3. A Lei Brasileira

É interessante saber o que diz a lei brasileira. A principal referência, no caso de plágio, é a Lei número 9.610, de 19 de fevereiro de 1998. Essa lei altera, atualiza e consolida a legislação sobre direitos autorais.

Uma pergunta frequente na área de Computação é se o governo, ao subvencionar um projeto, torna-se detentor do direito autoral sobre este. Por exemplo, se o aluno recebe uma bolsa da CAPES ou CNPq para fazer sua monografia, esses órgãos têm algum direito sobre a monografia e seus produtos? A lei é clara em seu artigo 6º:

“Não serão de domínio da União, dos estados, do Distrito Federal ou dos municípios as obras por eles simplesmente subvencionadas.”

Já o artigo 7º estipula quais são as obras protegidas pela lei (literalmente):

- a) Os textos de obras literárias, artísticas ou científicas.
- b) As conferências, alocuções, sermões e outras obras da mesma natureza.
- c) As obras dramáticas e dramático-musicais.
- d) As obras coreográficas e pantomímicas, cuja execução cênica se fixa por escrito ou por outra qualquer forma.
- e) As composições musicais, tenham ou não letra.
- f) As obras audiovisuais, sonorizadas ou não, inclusive as cinematográficas.
- g) As obras fotográficas e as produzidas por qualquer processo análogo ao da fotografia.
- h) As obras de desenho, pintura, gravura, escultura, litografia e arte cinética.
- i) As ilustrações, cartas geográficas e outras obras da mesma natureza.
- j) Os projetos, esboços e obras plásticas concernentes à Geografia, Engenharia, Topografia, Arquitetura, Paisagismo, Cenografia e Ciência.
- k) As adaptações, traduções e outras transformações de obras originais, apresentadas como criação intelectual nova.
- l) Os programas de computador.
- m) As coletâneas ou compilações, antologias, enciclopédias, dicionários, bases de dados e outras obras, que, por sua seleção, organização ou disposição de seu conteúdo, constituam uma criação intelectual.

No caso de programas de computador, especificamente, o § 1º desse artigo estabelece que eles ainda são objeto de uma lei específica.

Por outro lado, o artigo 8º estabelece quais as obras que *não* são protegidas por essa lei:

- a) As ideias, procedimentos normativos, sistemas, métodos, projetos ou conceitos matemáticos como tais.
- b) Os esquemas, planos ou regras para realizar atos mentais, jogos ou negócios.
- c) Os formulários em branco para serem preenchidos por qualquer tipo de informação, científica ou não, e suas instruções.
- d) Os textos de tratados ou convenções, leis, decretos, regulamentos, decisões judiciais e demais atos oficiais.
- e) As informações de uso comum tais como calendários, agendas, cadastros ou legendas.
- f) Os nomes e títulos isolados.
- g) O aproveitamento industrial ou comercial das ideias contidas nas obras.

Segundo a lei, o autor de uma obra será sempre a pessoa física que a gerou, não a pessoa jurídica. A proteção à pessoa jurídica poderá, porém, também ser concedida em casos previstos em lei.

A duração da proteção ao direito autoral perdura por 70 anos, a contar do dia primeiro de janeiro do ano subsequente à morte do autor. Os direitos patrimoniais sobre a obra são herdados pelos legítimos herdeiros, obedecida a ordem estabelecida pela Lei Civil.

Segundo o artigo 46 da mesma lei, não constitui ofensa aos direitos autorais a reprodução:

- a) Na imprensa diária ou periódica, de notícia ou de artigo informativo, publicado em diários ou periódicos, com a menção do nome do autor, se assinados, e da publicação da qual foram transcritos.
- b) Em diários ou periódicos, de discursos pronunciados em reuniões públicas de qualquer natureza.
- c) De retratos, ou de outra forma de representação da imagem, feitos sob encomenda, quando realizada pelo proprietário do objeto encomendado, não havendo a oposição da pessoa nele representada ou de seus herdeiros.

- d) De obras literárias, artísticas ou científicas, para uso exclusivo de deficientes visuais, sempre que a reprodução, sem fins comerciais, seja feita mediante o sistema Braile ou outro procedimento em qualquer suporte para esses destinatários.
- e) De um só exemplar de pequenos trechos, para uso privado do copista, desde que feita por este, sem intuito de lucro.
- f) Da citação em livros, jornais, revistas ou qualquer outro meio de comunicação, de passagens de qualquer obra, para fins de estudo, crítica ou polêmica, na medida justificada para o fim a atingir, indicando-se o nome do autor e a origem da obra.
- g) Do apanhado de lições em estabelecimentos de ensino por aquelas a quem se dirigem, vedada sua publicação, integral ou parcial, sem autorização prévia e expressa de quem as ministrou.
- h) De obras literárias, artísticas ou científicas, fonogramas e transmissão de rádio e televisão em estabelecimentos comerciais, exclusivamente para demonstração à clientela, desde que esses estabelecimentos comercializem os suportes ou equipamentos que permitam a sua utilização.
- i) Da representação teatral e execução musical, quando realizadas no recesso familiar ou, para fins exclusivamente didáticos, nos estabelecimentos de ensino, não havendo em qualquer caso intuito de lucro.
- j) De obras literárias, artísticas ou científicas para reproduzir prova judiciária ou administrativa.
- k) De pequenos trechos, em quaisquer obras, de obras preexistentes, de qualquer natureza, ou de obra integral, quando de artes plásticas, sempre que a reprodução em si não seja o objetivo principal da obra nova e que não prejudique a exploração normal da obra reproduzida nem cause um prejuízo injustificado aos legítimos interesses dos autores.

Portanto, a regra é copiar apenas o essencial de outros trabalhos, desde que seja realmente necessário para colocar uma informação sobre esses trabalhos ou traçar um comparativo, lembrando

sempre que o trecho copiado deve constar entre aspas e com citação da fonte para que nunca haja dúvida sobre se tratar ou não de plágio.

O plágio no Brasil é considerado crime, e a Lei prevê pena de multa e prisão. Portanto, não vale a pena transgredir essa lei. Também não existe plágio mais sério ou menos sério. Plágio é crime perante a Lei e academicamente é uma falta ética gravíssima. Houve um caso em que um estudante plagiou apenas o capítulo de método (metodologia) de sua monografia, copiando-o de outra monografia. Ele realmente fez o trabalho, realizou a pesquisa, obteve os dados e gerou as conclusões. Mas pela falta ética de ter copiado parte do trabalho, esse estudante teve seu diploma cassado, independentemente de outros fatores.

Areal (1997) apresenta em seu site uma série de comentários de plagiadores, que são pérolas que evidenciam o tipo de raciocínio ou ignorância que muitas vezes está por trás desse tipo de atitude. Algumas dessas pérolas são transcritas a seguir para exemplificar:

- a) “(...) você deveria estar orgulhoso de ver seu trabalho em um grande sucesso que não é ocasso da sua página” (*sic*). O plagiador sequer consegue escrever corretamente...
- b) “(...) o que de maneira nenhuma constitui-se na ação citada em no assunto (*subject*) de seu e-mail, o qual refere-se a crime de coação sob utilização de arma branca ou de fogo para obtenção de propriedade alheia. Acusação, aliás, que pode perfeitamente ser objeto de processo judicial por calúnia e difamação (considerando que várias pessoas possuem cópia testemunhal do seu delito criminal)” (*sic*). Aqui o plagiador acusa o autor de difamação...
- c) “Pensei que você fosse ficar orgulhoso”.
- d) “se você não quiser que ninguém copie, não ponha na Internet. É bobagem achar que vai ter exclusividade sobre o conteúdo na Internet. Besteira pura (...) isso é a Internet, cara. O jeito correto de lidar com a situação não é falar de (...) de lei de direito autoral, e sim relaxar e aproveitar enquanto a internet ainda é nossa” (*sic*).

- e) “Achei super interessante sua abortagem sobre plagio na rede, porem, constatei que na sua pagina existe alguns gifs (imagens) que acho que nao sao de sua autoria. Seria melhor que os mesmo fossem retirados, pois, e contraditorio falar de plagio e fazer o mesmo, mesmo sendo gifs de dominio publico” (sic). Bem, se é de domínio público pode usar, não?

Em relação a usar materiais de domínio público ou com autorização do autor, apenas deve-se tomar certo cuidado em verificar se a pessoa que autoriza o uso é realmente o autor. Há casos de sites que se apropriam de materiais de terceiros e autorizam seu uso por outras pessoas, mas não teriam autoridade para isso.

Capítulo 8

Níveis de Exigência do Trabalho de Conclusão

Dependendo do nível do curso, deve variar o nível de exigência em relação ao trabalho de conclusão. Embora Eco (1989) defina a monografia como um texto com 100 a 400 páginas, tamanho não é documento. O que se avalia é o grau e o tipo de contribuição que o estudante apresentou no trabalho.

A estrutura do ensino superior brasileiro identifica diferentes tipos de cursos. Inicialmente os cursos de graduação, que podem ser cursados pelos egressos do Ensino Médio. Há vários tipos e modalidades: bacharelado, engenharia, licenciatura são considerados cursos de graduação plena. Cursos de graduação mais rápidos são os cursos de tecnologia, que formam o tecnólogo, e os cursos sequenciais, ainda mais rápidos.

Já a pós-graduação caracteriza-se por dois tipos de curso, o *lato sensu* e o *stricto sensu*. Os cursos *lato sensu*, usualmente denominados “especialização”, ou seguindo modismos do norte, MBA (*Master in Business Administration*), são de cunho mais técnico e de complementação da formação técnica profissional.

Já os cursos *stricto sensu*, em seus dois níveis, mestrado e doutorado, procuram formar pesquisadores e docentes de ensino superior. São cursos, portanto, com características bastante diferentes dos cursos de graduação e *lato sensu*.

Há ainda o mestrado profissionalizante, que busca formar um pesquisador em uma área bastante próxima da aplicação profissional. Não se trata de um meio-termo entre a especialização e o mes-

trado científico, assim como a informática médica não é um meio termo entre a informática e medicina. É simplesmente uma forma diferente de conceber um curso de mestrado, com aplicação direta do conhecimento gerado na indústria.

Em geral, basta ter um diploma de graduação plena para poder ingressar em um programa de mestrado, embora algumas universidades tenham já aceitado tecnólogos.

Para ingresso no mestrado não é necessário ter especialização, embora algumas vezes isso possa valorizar o currículo do candidato no processo seletivo. Algumas universidades também validam algumas disciplinas cursadas em nível de especialização ao aluno que ingressa no mestrado. Deve-se verificar caso a caso como a universidade procede.

Para ingresso no curso de doutorado não é necessário ter o mestrado. Mas é praxe na maioria dos programas de doutorado não admitir alunos sem mestrado. A prática é possível, mas os programas dificilmente se arriscam.

O doutorado é considerado como a titulação plena e definitiva em termos acadêmicos. O título de PhD nada mais é do que um doutorado obtido em país de língua inglesa (Eco, 1989). Existem outros títulos equivalentes também no exterior, e às vezes deve-se tomar cuidado em não confundi-los com os títulos outorgados no Brasil.

Por exemplo, existem países na Europa com cursos de mestrado de um ano, em que o aluno apenas cursa disciplinas e entrega um trabalho escrito, constando basicamente de revisão bibliográfica. Esses cursos não são considerados equivalentes ao mestrado brasileiro, mas a cursos de especialização.

Para que um diploma no exterior seja reconhecido no território brasileiro é necessário que seja revalidado por uma universidade brasileira com delegação de poderes do Ministério da Educação para tal fim (em geral, as universidades federais). Por mais conceituado que seja o diploma obtido no exterior, ele só terá validade no Brasil mediante esse processo de revalidação. Tal processo é na maior

parte das vezes demorado, pois a universidade brasileira que tenha um programa *stricto sensu* na área da monografia apresentada irá constituir uma banca avaliadora que verificará se o trabalho teria qualidade para ser aprovado no próprio programa. Tendo qualidade, o título é revalidado e recebe um carimbo no verso do diploma, atestando sua validade no Brasil. Caso contrário, o pedido é recusado. Mas o interessado pode ainda procurar outra universidade e tentar o processo novamente.

No caso de cursos à distância, deve-se verificar, antes de mais nada, se a instituição tem autorização expressa do Ministério da Educação para oferecer esse tipo de curso (www.mec.gov.br). No caso de instituições estrangeiras que oferecem cursos à distância no Brasil também deve se tomar especial cuidado, pois o diploma é emitido no exterior e não tem validade automática no Brasil.

8.1. Graduação

O que se espera de um trabalho de conclusão em um curso de graduação? Neste nível podem ser feitos dois tipos de trabalho: o tecnológico e o científico.

O trabalho científico deve seguir as linhas metodológicas descritas neste livro.

Já o trabalho tecnológico consiste usualmente em o aluno ser capaz de mostrar que sabe aplicar as técnicas que aprendeu ao longo do curso. O desenvolvimento de um sistema interessante pode ser um bom exemplo de trabalho de final de curso, desde que o aluno o desenvolva usando técnicas aprendidas durante o curso e apresente um relatório mostrando isso.

8.2. Especialização

Os cursos de especialização já foram encarados como um degrau para chegar ao mestrado. Hoje em dia, porém, são vistos muito mais como uma complementação ou atualização profissional. Pela

lei brasileira, todo curso de especialização requer a elaboração e defesa pública de uma monografia.

Essa monografia pode até ser um trabalho de pesquisa, feito nos moldes metodológicos apresentados neste livro. Mas é aceitável também, em muitos cursos, que o aluno desenvolva apenas um estudo bibliográfico e que apresente as ideias aprendidas com alguma pequena contribuição pessoal, consistindo normalmente de comentários à bibliografia ou ao resultado de experimentos simples. Normalmente não se exige nesses casos provas de hipóteses ou uma contribuição científica mais relevante.

Isso, porém, varia muito de curso para curso. O aluno deve estar atento às exigências colocadas pelo curso em que ele está matriculado.

8.3. Mestrado e Doutorado

A maioria das recomendações deste livro se aplica ao mestrado e ao doutorado. Em ambos os casos, espera-se que o aluno apresente uma contribuição à ciência que seja relevante, ou seja, que não seja trivial, que seja útil e que esteja correta.

A diferença entre o que se espera no mestrado e no doutorado reside mais no nível de exigência da contribuição do que na forma. Em ambos os casos, exige-se a aplicação de metodologia científica, comparação com trabalhos correlatos, elaboração de uma hipótese de pesquisa e sua comprovação ou refutação. A diferença está então no impacto esperado dessa contribuição.

Para o mestrado, em geral, basta que o aluno apresente uma informação nova sobre algum tema, que seja relevante para a área. No caso do doutorado, essa informação nova tem de ter importância suficiente para mudar o modo como as pessoas em todo o mundo encaram aquela área de pesquisa. Ou seja, espera-se que um doutorado produza uma contribuição que de fato modifique o estado da arte.

Comer (2008) ainda acrescenta que uma tese deve ser “original” e “substancial”. É difícil muitas vezes avaliar isso. Cabe aos orientadores, com sua experiência, direcionar os alunos de forma a escolher um objetivo que seja compatível com o nível do curso que está obtendo.

Chinneck (1998) afirma que a diferença entre o mestrado e o doutorado não está na forma do documento nem na apresentação, mas apenas na profundidade e na dificuldade do problema sendo tratado. Segundo ele, uma tese de doutorado exige a resolução de um problema mais difícil e, consequentemente, mais contribuições significativas.

No Brasil costuma-se usar os termos “tese de doutorado” e “dissertação de mestrado”. Mas em inglês, ambos os termos são usados para designar os trabalhos de mestrado e doutorado. Rugaber (1995) diferencia os termos “tese” e “dissertação”, afirmando que a *tese* é uma afirmação que se procura provar e a *dissertação* é um texto descriptivo sobre a tese.

Referências

- ALVES, M. B. M.; ARRUDA, S. M. *Como fazer Referências: Bibliográficas, eletrônicas e demais formas de documentos*. Atualizada em fevereiro 2007. Disponível em: <<http://www.bu.ufsc.br/design/framerefer.php>>. Acesso em: 6 fevereiro 2009.
- AREAL, A. C. B. *Plágio e Direito Autoral na Internet Brasileira*. 1997. Disponível em: <<http://www.persocom.com.br/brasilia/plagio1.htm>>. Acesso em: 2003.
- BARBETTA, P. A.; REIS, M. M.; BORNIA, A. C. *Estatística para Cursos de Engenharia e Informática*. 2. ed. São Paulo: Atlas, 2008.
- CHINNECK, J. W. *How to Organize your Thesis*. 1988. Atualizada em 29 setembro 1999. Disponível em: <<http://www.sce.carleton.ca/faculty/chinneck/thesis.html>>. Acesso em: 2 janeiro 2009. Versão em português disponível em: <<http://www.sce.carleton.ca/faculty/chinneck/thesis/ThesisPortuguese.html>>.
- COMER, D. *How to Write a Dissertation or Bedtime Reading for People who do not have Time to Sleep*. Disponível em: <<http://www.cs.purdue.edu/homes/dec/essay.dissertation.html>>. Acesso em: 5 novembro 2008.
- DESCARTES, R. *Discurso do Método*. Tradução de Paulo Neves. Porto Alegre: L&PM, 2004. 123p. Título original: *Discours de la Méthode*.
- ECO, H. *Como se faz uma Tese*. Tradução de Gilson Cesar Cardoso de Souza. São Paulo: Perspectiva, 1989. 170p. Título original: *Como se fa una Tesi di Laurea*.

GRIDDITHS, R. *How to write a dissertation*. Disponível em: <<http://www.it.bton.ac.uk/staff/rng/papers/writediss.html>>. Acesso em: 5 novembro 2008.

HEXSEL, R. A. *Pequeno Manual de Escrita Técnica*. Curitiba, 2004. (Relatório Técnico 004/2004 DINF/UFPR). Disponível em: <http://www.inf.ufpr.br/info/techrep/RT_DINF004_2004.pdf>. Acesso em: 5 janeiro 2009.

KERLINGER, F. N. *Metodologia da Pesquisa em Ciências Sociais – Um tratamento conceitual*. Tradução de Helena Mendes Rotundo. São Paulo: EPU, 1980. Título original: *Behavioral Research – A conceptual approach*.

LAKATOS, E. M.; MARCONI, M. A. *Fundamentos de Metodologia Científica*. 6. ed. São Paulo: Atlas, 2006. 315p.

MORO, M. M. *A Arte de Escrever Artigos Científicos*. Disponível em: <<http://www.inf.ufrgs.br/~mirella/Dicas.html>>. Acesso em: 5 janeiro 2009.

NAGESWARAN, S. Test Effort Estimation Using Use Case Points. *Quality Week 2001*, San Francisco, Califórnia, junho de 2001. Disponível em: <<http://www.bfpug.com.br/artigos.htm>>. Acesso em: 18 fevereiro 2009.

PLATÃO. *A República*. Organização de Jacó Guinsburg. São Paulo: Perspectiva, 2006. 418p. Original em grego.

RUGABER, S. *Thoughts on the Structure of CS Dissertations*. 1995. Disponível em: <<http://www.cc.gatech.edu/fac/Spencer.Rugaber/txt/thesis.html>>. Acesso em: 3 janeiro 2009.

RUSSEL, S.; NORVIG, P. *Inteligência Artificial*. São Paulo: Campus: 2004. 1056p. Título original: *Artificial Intelligence – A Modern Approach*.

SNYDER, A. *How to Get your Paper Accepted at OOPSLA*. 1993. Disponível em: <<http://www.sigplan.org/oopsla/oopsla96/how91.html>>. Acesso em: 4 janeiro 2009.

- TRAINA, A. J. M.; TRAINA Jr., C. *Como Escrever Artigos*. 2002. Disponível em: <<http://gbdi.icmc.sc.usp.br/disciplinas/sce-5845/Como-Escrever/>>. Acesso em: 5 janeiro 2009.
- WAZLAWICK, R. S. *Análise e Projeto de Sistemas de Informação Orientados a Objetos*. Rio de Janeiro: Campus: 2004. 302 p.
- WIKIPÉDIA Hipótese. 2008. Disponível em: <<http://pt.wikipedia.org/wiki/Hipótese>>. Acesso em: 19 setembro 2008.
- WIKIPÉDIA Metodologia. 2009a. Disponível em: <<http://pt.wikipedia.org/wiki/Metodologia>>. Acesso em: 1º fevereiro 2009.
- WIKIPÉDIA Método Científico. 2009b. Disponível em: <http://pt.wikipedia.org/wiki/Método_científico>. Acesso em: 1º fevereiro 2009.