UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
COORDENAÇÃO DE PÓS-GRADUAÇÃO EM INFORMÁTICA

DISSERTAÇÃO DE MESTRADO

Sobre a Questão da Formalização do Raciocínio Abdutivo via Sistemas Dedutivos Rotulados

por

Nicia Cristina Rocha Riccio

Campina Grande, 10 de agosto de 1995

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
COORDENAÇÃO DE PÓS-GRADUAÇÃO EM INFORMÁTICA

NICIA CRISTINA ROCHA RICCIO

Sobre a Questão da Formalização do Raciocínio Abdutivo via Sistemas Dedutivos Rotulados

Dissertação apresentada ao Curso de Mestrado em Informática da Universidade Federal da Paraíba, em cumprimento às exigências para obtenção do Grau de Mestre.

Orientador: *Prof. Manoel Agamemnon Lopes*
Co-orientador: *Prof. Ruy J. G. B. de Queiroz*

Campina Grande, 10 de agosto de 1995

R489s Riccio, Nicia Cristina Rocha
Sobre a questao da formalizacao do raciocinio abdutivo via sistemas dedutivos rotulados / Nicia Cristina Rocha Riccio. - Campina Grande, 1995.
184 f. : il.

Dissertacao (Mestrado em Informatica) - Universidade Federal da Paraiba, Centro de Ciencias e Tecnologia.

1. Inteligencia Artificial - 2. Raciocinios Nao-Monotonicos 3. Logica Matematica 4. Dissertacao I. Lopes, Manoel Agamemnon, Dr. II. Queiroz, Ruy J. G. B. de, Dr. III. Universidade Federal da Paraiba - Campina Grande (PB) IV. Título

CDU 004.81:159.953.5(043)

Universidade Federal da Paraíba - UFPB
Centro de Ciências e Tecnologia - DSC
COORDENAÇÃO DE PÓS-GRADUAÇÃO EM INFORMÁTICA - COPIN
Av. Aprígio Veloso, 882 - 58.109-970 Campina Grande, PB
Fone/Fax: (++55) (+83) 333-1698 - e-mail: copin@dsc.ufpb.br

SOBRE A QUESTÃO DA FORMALIZAÇÃO DO RACIOCÍNIO ABDUTIVO VIA SISTEMAS DEDUTIVOS ROTULADOS"

NICIA CRISTINA ROCHA RICCIO

DISSERTAÇÃO APROVADA EM 10.08.95

MANOEL AGAMEMNON LOPES, D.Sc.
Presidente

RUY JOSÉ GUERRA BARRETO DE QUEIROZ, Ph.D
Examinador

Mario R. F. Benevides

MARIO ROBERTO FOLHADELLA BENEVIDES, Ph.D
Examinador

EDILSON FERNEDA, Dr.
Examinador

CAMPINA GRANDE - Pb.

A Adolfo e Victor, meus dois amores.

Longe é um lugar que não existe...
Richard Bach

Agradecimentos

Muitas foram as pessoas ou entidades que colaboraram, de uma maneira ou de outra, para que este trabalho fosse finalizado com sucesso. Assim, tento, aqui, agradecer a todas elas.

Primeiramente agradeço à Universidade Federal da Bahia e ao CNPq pela viabilização financeira deste trabalho e ao Centro de Processamento de Dados da UFBA pela minha liberação para cursar o mestrado.

Agradeço ao Departamento de Sistemas e Computação da UFPB de uma maneira geral, pela minha aceitação no programa de mestrado e pelo acompanhamento que recebi durante todo o período em que lá estive. Em especial, agradeço a Aninha, pelo carinho com que me recebeu e que recebe a todos.

Agradeço aos meus orientadores, Agamemnon e Ruy, pela grande força e ajuda.

Agradeço ao Departamento de Informática da UFPE por ter me permitido usufruir de suas instalações e de seu ambiente de pesquisa extremamente incentivador.

Aos amigos, agradeço a todos aqueles que de alguma maneira torceram por mim; ao pessoal do DSC, pelos nossos bons momento, em especial, " às meninas" pela grande amizade que conseguimos construir; ao pessoal do DI por me considerarem como parte do grupo.

Agradeço a meus pais e a minha irmã pelo apoio e carinho.

Agradeço a Victor, pela paciência que foi obrigado a ter; e a minha mãe, D. Nicinha, Sonca e Tânia pelo carinho e cuidados com ele nas minhas ausências.

E agradeço a Adolfo, pelo amor, pela compreensão e pelo grande incentivo.

Resumo

O objetivo desta dissertação é desenvolver uma teoria de aprendizado de máquina (em particular, raciocínio abdutivo proposicional) usando a estrutura dos Sistemas Dedutivos Rotulados (*LDS*) de D. Gabbay, sob a perspectiva de D. Gillies na dicotomia 'lógica dedutiva versus lógica indutiva' onde a inferência controlada funciona como a noção que aproxima os dois conceitos. Desse ponto de vista, o raciocínio abdutivo, um tópico que atrai bastante interesse na comunidade de IA e na pesquisa sobre automatização do raciocínio, pode ser visto como um tipo de dedução controlada onde o componente de controle será representado pelos rótulos do *LDS*. Este trabalho investiga a possibilidade de tratar lado a lado os aspectos do nível objeto e do meta-nível de um problema abdutivo: o conjunto de explicações possíveis para um fato observado é gerado dedutivamente (o nível objeto) e, depois disso, a explicação mais interessante é escolhida considerando o critério de preferência dado pelo usuário (o meta-nível). Dessa maneira, é possível dar uma solução mais plausível a um problema abdutivo. Além disso, já que o critério de preferência, que depende do significado dos rótulos, pode variar, essa caracterização da abdução pode ser facilmente adaptada para diferentes domínios.

Abstract

The intention here is to develop a theory of machine learning (in particular, propositional abductive reasoning) using the framework of D. Gabbay's Labelled Deductive Systems (*LDS*), in the light of Gillies' perspective on the dichotomy 'deductive versus inductive logic' where controlled inference serves as the bridging notion. From this point of view, abductive reasoning, a topic that attracts much interest in AI and automated reasoning research, can be seen as a kind of controlled deduction where the control component will be represented by the labels of the *LDS* framework. This work investigates the possibility of treating both object and meta-level aspects of an abductive problem side by side: the set of possible explanations to an observed fact is generated deductively (the object level) and, after that, the most interesting explanation is chosen considering the preference criterion stated by the user (the meta-level). This way, it is possible to give an abductive problem a more plausible solution. Moreover, since the preference criterion, which depends on the meaning of the labels, may vary, this characterization of abduction can easily be adapted to different problem domains.

Conteúdo

1 Introdução 1

1.1 Introdução 2

1.2 Organização da dissertação 6

2 Uma Visão Geral do Aprendizado de Máquina 8

2.1 Conceitos gerais sobre o aprendizado de máquina 9

2.2 As origens do problema da indução 16

2.3 As perspectivas indutiva e dedutiva 21

2.3.1 Controle na lógica dedutiva 23

2.3.2 Inferência na lógica indutiva 23

2.4 A indução como caso especial da abdução 25

2.5 O aprendizado indutivo 28

2.5.1 Importância de teorias de aprendizado 33

2.5.2 Algumas abordagens em indução 34

3 Abdução 45

3.1 A filosofia da abdução 46

3.1.1 Peirce e as origens da abdução 46

3.1.2 Sobre a abdução de Sherlock Holmes 51

3.1.3 Conclusões sobre a filosofia da abdução 53

3.2 A abdução formalizada 55

3.3 Trabalhos correlatos 56

3.3.1 Utilizando *predicate completion* para relacionar dedução e abdução 56

3.3.2 Raciocínio abdutivo via tableau e cálculo de seqüentes 67

3.3.3 Uma caracterização lógica da abdução 79

3.4 Outras abordagens 87

3.4.1 Uma regra de inferência para a geração de hipóteses 88

3.4.2 Abdução proposicional para a lógica modal 88

3.4.3 Uma organização de conhecimento para a abdução 89

3.4.4 Heurísticas cognitivamente plausíveis para alcançar a complexidade da abdução 90

3.4.5 As regras de abdução de Zadrozny 91

3.4.6 Raciocínio abdutivo em lógica tri-valorada 93

3.4.7 Abdução rotulada e raciocínio relevante 96

4 Os Sistemas Dedutivos Rotulados **100**

4.1 O por quê dos sistemas rotulados 101

4.2 O *LDS* como um *framework* geral para representar lógicas 103

4.3 Da interpretação funcional de Curry-Howard aos sistemas de dedução natural rotulada 105

4.3.1 A teoria da funcionalidade de Curry 106

4.3.2 A interpretação funcional da implicação 108

4.3.3 As contribuições de Howard e Tait 111

4.4 O sistema de dedução natural rotulado 111

4.4.1 A dedução natural de Gentzen 112

4.4.2 O *LND* 115

4.5 De volta a Frege 119

4.6 Abdução rotulada 119

5 Teoria de Aprendizado Abdutivo via *LDS* **124**

5.1 Definição do problema abdutivo 125

5.1.1 Os rótulos e sua ordenação 125

5.1.2 Os componentes de um problema abdutivo 126

5.2 Abdução proposicional baseada em dedução natural (o nível objeto) 129

5.2.1 Minimalidade e consistência 131

5.2.2 Restrições de integridade 132

5.2.3 Conseqüências dedutivas das hipóteses 132

5.3 O cálculo lógico e o cálculo funcional 133

5.3.1 Regras de introdução e eliminação de conectivos lógicos 133

5.4 Etapas para a solução de um problema abdutivo 135

5.4.1 A construção da prova e a geração dos candidatos 136

5.4.2 A propagação dos rótulos 150

5.4.3 A comparação dos candidatos 153

5.5 Algumas considerações sobre o método proposto 170

5.5.1 Sobre a não monotonicidade na abdução 170

5.5.2 Sobre a extensão para lógica de primeira ordem 172

6 Conclusão **174**

6.1 Considerações finais 175

Lista de Figuras

5.1 Ordenação de rótulos via diagrama 126

5.2 Rótulos com o mesmo grau de preferência 127

5.3 Ordenação de rótulos utilizada no critério da ordem preferencial 156

5.4 Ordenação de rótulos do exemplo 5.14 160

5.5 Ordenação de rótulos do exemplo 5.15 162

5.6 Ordenação de rótulos do exemplo 5.16 164

5.7 Ordenação de rótulos do exemplo das conseqüências dedutivas 166

5.8 Ordenação de rótulos do exemplo 5.18 167

5.9 Ordenação de rótulos do exemplo 5.19 169

Lista de Tabelas

3.1 Propriedades da relação abdutiva 83

4.1 Regras do Sistema de Dedução Natural 113

4.2 Regras da Dedução Natural Rotulada 116

4.3 Correspondência do *LND* com o sistema de Heyting 117

Capítulo 1

Introdução

Neste capítulo iremos introduzir o tema principal desta dissertação, ressaltando seus objetivos e principais contribuições. Particularmente, definiremos os conceitos de abdução e Sistemas Dedutivos Rotulados, dois pontos centrais da dissertação e de cuja compreensão depende o entendimento dos capítulos que seguem. Depois disso, descreveremos a estrutura da dissertação com o conteúdo de cada capítulo.

1.1 Introdução

Na comunidade de Inteligência Artificial a abdução é um termo comumente utilizado para designar a geração de explicações para um conjunto de eventos de uma dada teoria, isto é, o raciocínio abdutivo é um modo de resolver problemas onde um evento observado φ, não questionável, não é explicado pela presente teoria Θ adotada e uma explicação para φ deve ser buscada. Seu esquema lógico característico é a inferência de α a partir de $\alpha \rightarrow \varphi$ e φ. É uma forma de inferência que não preserva a verdade e que reflete algumas formas de raciocínio do senso comum [MP93].

No raciocínio abdutivo, as explicações devem respeitar algumas condições fundamentais para que sejam aceitas como 'interessantes'. Algumas dessas condições podem ser expressas no nível objeto onde a teoria Θ está definida; por exemplo a restrição de que φ deve ser conseqüência lógica de $\Theta \cup \alpha$ e que α deve ser consistente com Θ ($\Theta \not\vdash \neg\alpha$) [MP95b]. Mas a maioria das características que tornam uma explicação interessante são considerações do meta-nível que não podem ser expressas no nível objeto [CDT91]. Sob esse ângulo, abdução pode ser vista como um tipo de inferência controlada, e então um sistema dedutivo pode ser usado para solucionar o nível objeto de problemas abdutivos [MP95a] e as restrições do meta-nível deverão ser tratadas de um modo diferente.

O termo abdução foi primeiramente introduzido por Charles Sanders Peirce, que considerava a abdução como um tipo de inferência lógica, assim como a dedução e a indução. Segundo as definições de Peirce, temos os seguintes esquemas lógicos básicos para os três tipos de inferência:

Dedução - um processo analítico baseado na aplicação de regras gerais a casos particulares, com a inferência de um resultado:

Sócrates é homem	(CASO)
Todo homem é mortal	(REGRA)
Sócrates é mortal	(RESULTADO)

Indução - raciocínio sintético que infere a regra a partir do caso e do resultado:

Sócrates é homem	(CASO)
Sócrates é mortal	(RESULTADO)
Todo homem é mortal	(REGRA)

Abdução - raciocínio sintético que infere o caso a partir da regra e do resultado:

Sócrates é mortal	(RESULTADO)
Todo homem é mortal	(REGRA)
Sócrates é homem	(CASO)

Indução e abdução sempre foram consideradas muito próximas uma da outra, e muitas vezes foram erroneamente tratadas como um argumento único (frequentemente misturadas também com dedução). O próprio Peirce, citado em [SS91], admite ter confundido de certo modo abdução e indução e atribui a confusão entre esses dois tipos de raciocínio ao caráter estreito e formalista da concepção de inferência adotada pelos lógicos (como tendo-se que necessariamente formular julgamentos a partir das premissas). Além do que, abdução e indução levam ambas à aceitação de uma hipótese porque os fatos observados são exatamente aqueles que resultariam necessaria ou provavelmente como conseqüências daquela hipótese. Sob o nosso ponto de vista, a indução pode ser entendida como um caso especial da abdução, já que ambas oferecem (ou buscam) como conclusão a explicação para fatos observados. A diferença básica é que a indução necessariamente acontece com a observação de fatos similares e a explicação encontrada deve ser em forma de regra. Nenhuma dessas duas condições, no entanto, descaracterizam a abdução.

Por outro lado, a lógica indutiva e a lógica dedutiva durante muito tempo foram consideradas como áreas completamente distintas e desvinculadas. A fórmula proposta por D. Gillies '*lógica = inferência + controle*' apresenta uma nova perspectiva sobre a dicotomia 'indução $\times$ dedução', e aponta para um conceito chave presente na maioria dos trabalhos recentes em lógica e computação: 'inferência controlada' [Gil94]. A observação principal é que enquanto o controle vem crescendo na lógica dedutiva, inferência controlada vem se tornando a regra na lógica indutiva. Isso retrata a chamada reaproximação entre as lógicas dedutiva e indutiva. Em termos de lógica e computação, a inferência controlada pode ser vista como um tipo de conceito generalizado de dedução no nível objeto onde os elementos do chamado meta-nível têm um papel decisivo. Uma das áreas da computação onde a 'inferência controlada' é comumente utilizada e que, de certa forma, promove essa chamada reaproximação é o aprendizado de máquina.

Aprendizado pode ser considerado como qualquer mudança em um sistema que o permita executar melhor pela segunda vez uma determinada tarefa ou uma outra tarefa semelhante a uma primeira que já tenha sido executada [Sim83]. A pesquisa na área de aprendizado de máquina visa fazer com que programas de computador alcancem essa característica. Dos diversos tipos de aprendizado já propostos na literatura o que mais diz respeito a esta dissertação é o aprendizado indutivo, justamente devido à proximidade entre abdução e indução que citamos acima. No aprendizado indutivo, o objetivo básico é induzir conceitos gerais a partir de casos particulares; ou seja, é a busca de descrições gerais plausíveis (assertivas indutivas) que explicam os dados de entrada e são úteis para predizer novos dados. Para que um programa de computador possa formular esse tipo de descrição, deve-se usar uma linguagem de descrição apropriada, e para qualquer linguagem de descrição não trivial, um grande número de assertivas indutivas pode ser formulado. A escolha das assertivas indutivas que mais se adequam ao conjunto de dados de entrada é um dos pontos cruciais da indução.

O objetivo básico desta dissertação é esboçar, sob o ponto de vista da perspectiva de Gillies e considerando a proximidade entre abdução e indução, os elementos de uma teoria de aprendizado de máquina enfocando o raciocínio abdutivo, usando a estrutura dos Sistemas Dedutivos Rotulados (*LDS* - *Labelled Deductive Sistems*) proposta por D. Gabbay.

Considerando a estrutura do *LDS*, a unidade declarativa de um sistema lógico é uma fórmula rotulada '$t : A$', e enquanto as fórmulas são parte de um cálculo lógico, os termos são expressões do cálculo funcional. Além disso, existe uma certa harmonia entre o cálculo funcional nos rótulos e o cálculo lógico nas fórmulas [Gab94]. Desse modo é estabelecida uma maneira de representar sistemas lógicos onde as características do meta-nível (os rótulos) possam coexistir lado a lado com as características do nível objeto (as fórmulas). O *LDS* surge dentro dessa perspectiva; a grosso modo, podemos dizer que um sistema de prova baseado em *LDS* é uma tripla (A, L, M), onde L é a linguagem lógica (conectivos e fórmulas bem formadas), A é uma álgebra (com algumas operações) sobre rótulos e M é a disciplina que estabelece o uso dos rótulos nas fórmulas da lógica, juntamente com as regras de dedução em harmonia com uma disciplina de propagação dos rótulos através das regras de dedução [Gab94]. Lógicas tradicionais manipulam fórmulas, o *LDS* visa a manipulação de pares, fórmulas e rótulos. Os rótulos devem ser vistos como informação suplementar sobre as fórmulas, não presente nas mesmas.

Nossa intenção aqui é utilizar a estrutura do *LDS* para tratar ambos os níveis da abdução (nível objeto e meta-nível), restringindo-nos, no entanto à abdução para a lógica proposicional. Particularmente nesse caso (uma teoria de raciocínio abdutivo), os rótulos do *LDS* irão representar as características do meta-nível do referido problema. A idéia é lidar com uma ordem (possivelmente parcial) entre os rótulos que representará o critério de preferência para cada problema abdutivo. Os candidatos a serem uma explicação para o evento φ serão gerados dedutivamente utilizando-se as regras de dedução natural e a explicação mais interessante será escolhida baseado na ordem parcial dos rótulos (as características do meta-nível).

Uma definição de explicações abdutivas precisa é obviamente dependente da lógica que caracteriza a teoria Θ. Então, assumimos que uma lógica L será dada ao sistema como um parâmetro (como o permite o *LDS*). O mesmo ocorre com o critério de preferência (as restrições do meta-nível), dependendo do tipo de problema, o significado dos rótulos pode variar; por exemplo, se estamos tratando com incerteza, os rótulos podem ser entendidos como ordem de preferência, grau de confiabilidade, medidas de probabilidades (como as vistas em [Kyb93], [Kyb90a] e [Kyb90b]) ou distribuições de possibilidade (como as vistas em [DLP93] e [DP90]). A comparação dos rótulos a fim de buscar a explicação mais interessante dependerá do significado dos mesmos, e algum tipo de interação com o usuário deverá ser provido pelo sistema para permitir que o usuário escolha o tipo de comparação que ele considere mais adequada a depender da área de aplicação que esteja tratando.

Nosso trabalho visa a formulação de uma teoria geral de aprendizado abdutivo, não voltado a uma área de aplicação ou problema específicos. Na prática, os métodos aplicados a problemas específicos parecem (e são em verdade) mais eficientes; no entanto, o desenvolvimento apenas desse tipo de método leva a diversas dificuldades de integração entre diferentes pesquisas. Cada projeto utiliza suas próprias formalizações e maneiras de representação, levando inclusive a uma prejudicial redundância de terminologias. Os métodos teóricos mais genéricos, apesar de considerados por muitos como apenas teoricamente elegantes e não úteis na prática, são importantes no sentido de promoverem uma maior integração entre projetos diferentes, o que sem dúvida favorece a evolução da pesquisa na área de aprendizado de máquina.

A utilização de rótulos com a estrutura do LDS, por seu lado, se justifica pela crescente demanda de sistemas lógicos motivada pelo grande número de aplicações na teoria da computação, inteligência artificial e programação em lógica. São as considerações do meta-nível que diferenciam uma lógica de uma outra. Além disso, no caso específico do raciocínio abdutivo, a presença dos rótulos do *LDS* vem justamente permitir que o tratamento diferenciado do meta-nível da abdução leve a uma melhor seleção das hipóteses mais plausíveis de modo a fornecer como resultado final um conjunto de hipóteses abdutivas menor que o conjunto inicial gerado dedutivamente. A presença dos rótulos também permite que se possa fazer a seleção das hipóteses de acordo com o critério de preferência que seja mais relevante para a área de aplicação em questão, já que o significado dos rótulos será dado em função do problema.

1.2 Organização da dissertação

Esta dissertação está organizada em seis capítulos.

No Capítulo 1 introduzimos o tema principal da dissertação, ressaltando seus objetivos e principais contribuições; definimos os conceitos de abdução e Sistemas Dedutivos Rotulados, dois pontos centrais da dissertação e de cuja compreensão depende o entendimento dos capítulos que seguem. Além disso, descrevemos, aqui, a estrutura da dissertação com o conteúdo de cada capítulo e do apêndice.

No Capítulo 2 abordaremos o aprendizado de máquina. Serão apresentados conceitos já definidos da área de aprendizado de máquina, particularmente em aprendizado indutivo, e serão abordadas algumas técnicas em indução ressaltando sua similaridade com o raciocinio abdutivo, sem, no entanto, pretendermos apresentar uma revisão bibliográfica completa do aprendizado de máquina.

No Capítulo 3 abordaremos a abdução sob dois pontos de vista: primeiro a filosofia da abdução, suas origens, sua participação no raciocínio do senso comum, seu relacionamento com outros tipos de raciocínio; depois trataremos a abdução de maneira mais formal, tentaremos dar uma caracterização genérica da abdução que seja condizente com as abordagens estudadas e, em seguida, discutiremos diversas propostas para teorias abdutivas.

No Capítulo 4 serão introduzidos os sistemas dedutivos rotulados (do inglês *Labelled Deductive Systems - LDS*). Descreveremos o *LDS*, como *framework* geral, e uma instância do mesmo, o sistema de dedução natural rotulada (do inglês *Labelled Natural Deduction - LND*). Finalizando, abordaremos a abdução rotulada como é vista por Dov Gabbay.

No Capítulo 5 estudaremos diversos aspectos de uma formalização do raciocínio abdutivo utilizando a estrutura dos Sistemas Dedutivos Rotulados. A abdução é vista como a busca de candidatos para explicar determinado fato observado, considerando o conhecimento de *background* existente. A solução de um problema abdutivo será dividida em dois níveis básicos: o nível objeto, onde serão gerados os candidatos, e o meta-nível, onde serão escolhidos os candidatos mais plausíveis, de acordo com um critério meta-lógico.

No Capítulo 6 apresentaremos as principais conclusões obtidas com o trabalho proposto, as principais vantagens, as atuais deficiências. Em seguida relacionaremos as propostas de trabalhos futuros relacionados com o tema da dissertação, a maioria delas envolvendo extensões da teoria aqui proposta.

Capítulo 2

Uma Visão Geral do Aprendizado de Máquina

Como o tema central desta dissertação é o raciocínio abdutivo, que pode ser encarado como uma forma de aprendizado, nossa intenção, neste capítulo, é apresentar conceitos já definidos da área de aprendizado de máquina (embora não haja um completo consenso sobre o assunto entre os pesquisadores da área), particularmente em aprendizado indutivo, e abordar algumas técnicas em indução ressaltando sua similaridade com o raciocínio abdutivo, sem, no entanto, pretendermos apresentar uma revisão bibliográfica completa do aprendizado de máquina.

2.1 Conceitos gerais sobre o aprendizado de máquina

Com a invenção dos computadores, altamente eficazes em atividades mecanizadas como cálculos, contagens, buscas, seleções, começa a surgir o desejo de utilizar essas máquinas para atividades não tão mecanizadas, mas mais especializadas e 'inteligentes', nascendo daí a pesquisa em Inteligência Artificial (IA). Não nos cabe aqui defender ou não a utilização do termo 'inteligência', apenas o tomamos como um consenso e o adotaremos sem considerar as, talvez relevantes, críticas a ele.

Um dos grandes sucessos reconhecidos da IA são os sistemas especialistas, que se difundiram largamente e hoje já se tornam produtos comercializáveis. O desenvolvimento de sistemas especialistas, no entanto, apresenta um grande 'gargalo', a aquisição de conhecimento. A dificuldade em conseguir obter do especialista o conhecimento especializado necessário, e de formalizar adequadamente esse conhecimento, começa então a incentivar outra área de pesquisa dentro da IA: a aquisição automática de conhecimento, ou aprendizado automático ou ainda aprendizado de máquina. Consideramos aqui que os três termos referem-se ao mesmo conceito (com a pequena diferença de que o primeiro é mais empregado quando se trata de adquirir conhecimentos especificamente para sistemas especialistas). Utilizaremos sempre o terceiro termo, aprendizado de máquina, buscando uma maior correspondência com o termo inglês *machine learning*, que nos parece gozar de um maior consenso.

O aprendizado de máquina passa a ser, então, de grande importância, e começa, de certa forma, a adiquirir força como tema de pesquisa e a desvincular-se um pouco da aquisição de conhecimento para sistemas especialistas. Começa a ser considerado necessário para que se tenha sistemas inteligentes, já que o conceito de inteligência prediz, dentre outras coisas, a capacidade de auto melhoramento, que pode ser entendido como um tipo de aprendizado [For89]. Além disso, o aprendizado de máquina pode abrir a possibilidade de sintetizar conhecimentos totalmente novos quando da automação do processo de descoberta científica, criando padrões e conhecimentos até então desconhecidos.

Existe tão pouco consenso na definição de aprendizado quanto na definição de inteligência. Não buscaremos aqui uma definição que agrade a todos, mas sim exaltaremos aquelas características do aprendizado que consideramos relevantes para o nosso trabalho. Podemos considerar o aprendizado como qualquer mudança em um sistema que o permita executar melhor pela segunda vez uma determinada tarefa ou uma outra tarefa semelhante a uma primeira que já tenha sido executada antes [Sim83]. Com essa conceituação bastante abrangente, podemos ver o aprendizado de máquina como fazendo parte de um conceito maior, o de aprendizado de uma maneira geral; ou podemos dizer que o termo aprendizado de máquina está razoavelmente bem aplicado.

À medida que a pesquisa na área de aprendizado de máquina vai crescendo, vão surgindo diversas tentativas de se criar uma classificação dos tipos de aprendizado, no entanto, tão pouco existe um consenso na literatura com relação a essa classificação.

Antes de qualquer classificação de aprendizado de máquina, temos a própria divisão da IA em simbólica e conexionista; esta última abrange a pesquisa em redes neurais, também considerada como um tipo de aprendizado, o qual, no entanto, não abordaremos aqui. Trataremos, então, apenas da IA simbólica e traçaremos uma classificação de aprendizado relativa apenas a essa parte da IA, classificação esta que foi proposta por Michalski [CMM83, Mic86].

O critério de classificação que abordaremos é a estratégia do aprendizado, dando importância à quantidade de inferência executada pelo sistema aprendiz em cima da informação disponível:

Aprendizado roteado (*rote learning*)

Nenhuma inferência é executada pelo sistema, ele será programado por uma entidade externa no estilo usual de um programa de computador ou apenas 'memorizará' fatos (armazenando-os em bancos de dados); seu conhecimento cresce mas nenhuma inferência é executada de sua parte, todo o esforço cognitivo fica a cargo do programador.

Aprendizado por instrução (*learning by being told*)

O sistema adquire conhecimento de uma fonte organizada (como um livro ou um professor) e deverá reestruturar esse conhecimento em uma representação interna e útil ao sistema e que possa ser integrada com os conhecimentos já existentes. O aprendiz executará alguma inferência, mas a grande fração da tarefa fica com o mestre que deverá organizar e apresentar o conhecimento, de modo a incrementar o conhecimento já existente no aprendiz. A tarefa do aprendizado de máquina, nesse caso, é construir um sistema que pode aceitar instruções e conselhos e pode armazenar e utilizar esse conhecimento aprendido. Esse método corresponde à maioria dos métodos formais de educação.

Aprendizado por analogia (*learning by analogy*)

O sistema utiliza conhecimentos já adquiridos anteriormente para solucionar problemas semelhantes, executando as transformações necessárias no conhecimento preliminar adaptando-o à nova situação. Por exemplo, uma pessoa que nunca dirigiu um caminhão pequeno, mas que dirige automóveis, pode muito bem transformar seu conhecimento prévio para a nova tarefa. Similarmente, um sistema de aprendizado por analogia pode ser utilizado para converter um programa de computador já existente, em outro que execute uma função similar para a qual não foi originalmente projetado.

O aprendizado por analogia requer uma maior quantidade de inferência por parte do aprendiz que os dois anteriores. Um fato análogo deve ser guardado na memória, essa informação deverá ser transformada, aplicada à nova situação e então armazenada para uma futura utilização.

Aprendizado indutivo (*inductive learning*)

No aprendizado indutivo o objetivo básico é induzir conceitos gerais a partir de informações que são fornecidas ao sistema. Esse tipo de aprendizado pode ser dividido em dois grandes grupos: aprendizado por exemplos e aprendizado por observação e descoberta.

- Aprendizado por exemplos (*learning from examples*)

  Dado um conjunto de exemplos e contra-exemplos de um conceito, o aprendiz induz uma descrição conceitual geral que descreva todos os exemplos positivos e nenhum dos contra-exemplos.

  Dentro do aprendizado por exemplos podemos ter várias subdivisões:

  Quanto ao tipo de generalizações [Mic86]:

  - Generalizações de instâncias para classes - onde são dadas instâncias independentes de alguma classe de objetos e o sistema deve induzir uma descrição geral da classe.
  - Generalizações de parte para todo - onde a tarefa é criar uma hipótese de descrição de um objeto (cena, situação, processo) sendo dadas partes selecionadas do objeto.

  Quanto à fonte dos exemplos:

- A fonte pode ser um professor que conhece o conceito e gera sequências de exemplos as mais úteis possíveis. Se o mestre sabe (ou infere) o estado de conhecimento do aprendiz, os exemplos podem ser selecionados para otimizar a convergência para o conceito desejado.
- A fonte pode ser o próprio aprendiz que conhece seu estado de conhecimento mas não sabe exatamente qual o conceito a ser aprendido. Nesse caso o aprendiz gera instâncias baseadas nas informações que ele considera necessárias para discriminar entre as descrições de conceito que ele já tem. Por exemplo, um aprendiz tentando construir o conceito de 'substância ferromagnética', pode gerar como um possível candidato 'todos os metais'; uma vez testando metais como o cobre, ele verá o erro da generalização, e tentará gerar instâncias para definir as discriminações necessárias.
- A fonte pode ser um equipamento externo e nesse caso o processo de gerar os exemplos é randômico e o aprendiz deve saber lidar com observações relativamente não controladas. Por exemplo, um geólogo tentando inferir precursores de terremotos deve lidar com uma apresentação de dados não estruturada (apesar do geólogo saber o conceito de terremoto, ele não pode saber onde e quando vai ocorrer, nem pode fazer com que ocorra).

Quanto ao tipo de exemplos disponíveis ao aprendiz, podemos ter sistemas que lidam apenas com exemplos positivos ou aqueles que lidam com exemplos positivos e negativos ou contra-exemplos; neste último caso os contra-exemplos funcionam como uma prevenção contra a super generalização. Quando o sistema lida apenas com exemplos positivos, algum mecanismo deve ser previsto para evitar a super generalização, como considerar apenas generalizações 'mínimas' necessárias ou com a existência de um conhecimento *a priori* (conhecimento de *background*) que, de alguma maneira, restrinja o conceito a ser inferido.

O aprendizado por exemplos pode ser ainda incremental ou não-incremental. No primeiro caso o sistema deve formar uma ou mais hipóteses de um conceito, consistente com os dados disponíveis, e subsequentemente refinar as hipóteses após considerar exemplos adicionais. No segundo caso, todos os exemplos são apresentados de uma só vez. A técnica incremental assemelha-se mais ao aprendizado humano e permite ao mestre enfocar aspectos que ele considere importantes diante das hipóteses geradas inicialmente.

- Aprendizado por observação e descoberta (*learning from observation and discovery*)

  O aprendizado por observação e descoberta é também conhecido com aprendizado não supervisionado. Nesse caso o aprendiz não tem à sua disposição um conjunto de instâncias de um conceito particular nem lhe é dado acesso a algum tipo de oráculo que lhe informe se determinada instância é um exemplo ou contra-exemplo de qualquer conceito dado. O aprendiz deve contar apenas com um conhecimento de *background* bastante significativo e através de observações deve construir novos conceitos. Além disso, ao invés de concentrar-se em um conceito por vez, as observações podem ser relevantes para diversos conceitos diferentes que devem ser adquiridos. Esse tipo de aprendizado não supervisionado requer do aprendiz uma maior quantidade de inferência que os discutidos anteriormente. É um tipo bastante geral de aprendizado indutivo que engloba sistemas de descoberta, formação de teorias, criação de hierarquias taxonômicas, e outras tarefas semelhantes.

Podemos subdividir essa classe de aprendizado em dois tipos:

- a observação passiva onde o aprendiz classifica e cria taxonomias a partir de observações de vários aspectos do equipamento, sem, no entanto, interagir com o mesmo;

- a experimentação ativa onde o aprendiz pode interagir de alguma maneira com sua fonte de informação, provocando o fornecimento de informações observáveis, ou seja, o aprendiz perturba a fonte para observar os resultados da perturbação. Geralmente esse tipo de aprendizado envolve a geração de exemplos para testar hipóteses de conceitos ou conceitos parcialmente adquiridos.

Resumindo a classificação apresentada, temos:

- Aprendizado
  - Roteado
  - Por instrução
  - Por analogia
  - Indutivo
    - Por exemplos
      - Instância para classe
      - Parte para todo
      - Professor como fonte
      - Aprendiz como fonte
      - Equipamento externo como fonte
      - Somente exemplos
      - Exemplos e contra-exemplos
      - Incremental
      - Não incremental
    - Por observação e descoberta
      - Observação passiva
      - Experimentação ativa

Uma outra interessante classificação, também proposta por Michalski [CMM83] é a classificação de acordo com o tipo de representação do conhecimento adquirido pelo aprendiz:

- **Parâmetros em expressões algébricas** - É o ajuste de parâmetros numéricos ou coeficientes em expressões algébricas de forma funcional fixa de modo a obter o desempenho desejado.

- **Árvores de decisão** - Provendo a discriminação entre classes de objetos. Os nós em uma árvore de decisão correspondem a atributos dos objetos e os arcos correspondem a valores predeterminados desses atributos. As folhas representam conjuntos de objetos com classificação idêntica.

- **Gramáticas formais** - Induzidas a partir de uma seqüência de expressões em uma linguagem (geralmente artificial), a qual se deseja reconhecer. Essas gramáticas são tipicamente representadas como expressões regulares, autômatos finitos, regras de gramáticas livres de contexto, e outros.

- **Regras de produção** - São pares de condições e ações, ações estas que devem ser executadas se as respectivas condições são satisfeitas. Devido à sua simplicidade e fácil interpretação, esse tipo de representação de conhecimento é muito usado em sistemas de aprendizado de máquina.

- **Expressões lógicas formais** - São utilizadas para formular descrições de objetos individuais (entrada para um sistema de aprendizado) e para formular a descrição do conceito resultante (saída do sistema de aprendizado). Essas descrições tomam a forma de expressões lógicas formais cujos componentes são proposições, variáveis de valores finitos, sentenças restringindo o tamanho de variáveis, etc.

- **Grafos e redes** - Em muitos domínios grafos e redes são mais convenientes e eficientes como representação de conhecimento que expressões lógicas.

- **Frames e esquemas** - Nesse caso temos unidades de representação bem maiores que expressões lógicas ou regras de produção que se mostraram bastante úteis em muitas aplicações de IA. Por exemplo, um sistema que adquire conhecimento a partir de mapas genéricos deve ser capaz de lidar com esses mapas como unidades, apesar de sua estrutura interna poder tornar-se arbitrariamente complexa.

- **Programas** - O objetivo de diversos sistemas de aprendizado é adquirir habilidade para executar um específico procedimento eficientemente ao invés de raciocinar sobre a estrutura interna do processo. Esse tipo de sistema pode gerar programas de computador ou codificações procedurais como, por exemplo, sequências de instruções para manipulação de robôs.

- **Taxonomias** - Normalmente utilizadas no aprendizado por observação e descoberta por não serem tão restritivas como conceitos descritos numa linguagem mais formal.

- **Representações múltiplas** - Alguns sistemas de aprendizado de máquina usam vários esquemas de representação para o conhecimento adquirido. Em alguns casos mais notáveis, o sistema deve selecionar a forma ou formas de representação mais adequada(s), a depender do conhecimento adquirido.

Dos tipos de aprendizado que tratamos, o que diz respeito diretamente a esta dissertação é o aprendizado indutivo, o qual está bastante relacionado com o raciocínio abdutivo, tema central da dissertação.

Indução e abdução sempre andaram juntas, e muitas vezes foram erroneamente tratadas como um argumento único (frequentemente misturadas também com dedução). O próprio C. S. Peirce (considerado o pai da abdução), citado em [SS91], admite ter confundido de certo modo abdução e indução e atribui a confusão entre esses dois tipos de raciocínio ao caráter estreito e formalista da concepção de inferência adotada pelos lógicos (como tendo-se que necessariamente formular julgamentos a partir das premissas). Abdução e indução levam ambas à aceitação de uma hipótese porque os fatos observados são exatamente aqueles que resultariam necessaria ou provavelmente como conseqüências daquela hipótese. Porém, segundo Peirce: "A abdução se inicia a partir dos fatos, sem que nesse começo haja qualquer teoria particular em vista, embora seja motivada pelo sentido de que a teoria é necessária para explicar os fatos surpreendentes. A indução se inicia de uma hipótese que parece recomendar a si própria, sem que, nesse começo, hajam quaisquer fatos em particular à vista embora sinta a necessidade de fatos para sustentar a teoria. A abdução persegue uma teoria, a indução persegue fatos. Na abdução a consideração dos fatos segue a hipótese. Na indução o estudo da hipótese segue a experimentação que traz à luz os próprios fatos, para os quais a hipótese havia apontado." (C. S. Peirce, citado em [SS91]).

Sob o nosso ponto de vista, no entanto, a proximidade existente entre abdução e indução é muito maior que a definida por Peirce. Acreditamos que a indução pode ser vista como um caso especial da abdução, e na seção 2.4 mostraremos nossos argumentos a favor disso.

2.2 As origens do problema da indução

Segundo Newton da Costa em seu ensaio *Lógica indutiva e probabilidade* [dC93], que esboça uma nova solução para o problema da indução, " toda a ciência é um vasto sistema conceitual que nos permite, entre outras coisas, sistematizar o real. O próprio senso comum se articula por meio de conceitos. Não há razão sem conceituação". Dado um sistema de categorias[1], o qual pode servir de fundamento para certas sistematizações racionais, a ele em geral acha-se associada uma lógica, a qual determina as inferências válidas correlativas ao sistema de categorias considerado. Qualquer lógica L deve possuir associada uma família de linguagens a fim de que se possa expressar suas regras. Dada a lógica L, as inferências que ela estipula como válidas denominam-se L-deduções. As outras inferências codificadas na linguagem L, mas que não são L-dedutíveis, chamam-se L-paralogismos.

Além da lógica clássica, surgiram, especialmente nesse século, novas lógicas que diferem da postura clássica. Tais lógicas são chamadas heterodoxas ou não clássicas. Ainda segundo

[1]Segundo Newton da Costa, "denominam-se categorias os conceitos muito gerais que pertencem a vários ou todos os ramos da ciência" [dC93].

Newton da Costa, há dois tipos de lógicas heterodoxas: primeiro as que foram propostas como complemento da clássica (por exemplo, lógica modal e lógica temporal); segundo as formuladas com o intuito de substituir a lógica clássica em todos ou determinados contextos racionais (como a lógica intuicionística e a paraconsistente).

Muitas vezes a utilização de paralogismos é da máxima relevância. Na vida cotidiana, na medida em que recorremos a alguma lógica, essa lógica parece ser a clássica. Porém, se uma pessoa quisesse fazer apenas inferências válidas no seu dia-a-dia, dificilmente sobreviveria. Grande parte das inferências realmente importantes da vida comum constituem paralogismos. Se alguém infere, depois de várias constatações, que a bebida alcoólica lhe causa dores de cabeça, não está raciocinando de modo logicamente válido.

Nas ciências empíricas, o emprego de raciocínios que não são logicamente válidos é óbvio, e não haveria ciência empírica se os cientistas procurassem empregar unicamente formas válidas de inferência. Não há dúvidas que a dedução, o raciocínio demonstrativo, tem grandes e importantes aplicações nas ciências empíricas. Por exemplo, quando se está interessado em obter as conseqüências de uma teoria ou as implicações de determinada hipótese, deve-se recorrer à dedução. No entanto, quando se faz realmente avançar a ciência, quando se formulam as leis e teorias, recorre-se à inferência não dedutiva.

Segundo Lakatos [Lak68], a epistemologia clássica pode ser caracterizada por seus dois principais problemas: (1) o problema da fundamentação do conhecimento (lógica da justificação) e (2) o problema do crescimento do conhecimento (lógica da descoberta). O conhecimento somente pode ser gerado através de fatos observáveis, e suas conseqüências indutivas ou dedutivas, cujos significados não eram claramente distintos nos séculos XVII e XVIII. A lógica da justificação do empirismo clássico vem a ser definida, então, como: "proposições factuais e suas conseqüências - dedutivas/indutivas - são conhecimento, o resto é lixo".

De fato, segundo uma linha do empirismo, não só a verdade mas também o significado só podem ser iluminados pela luz da experiência. Mas então, para o crescimento da ciência, a indução não só de proposições mas também de conceitos passa a ser indispensável. O problema da definição indutiva tornou-se crucial para o empirismo lógico e as sucessivas falhas na busca de soluções levou ao chamado liberalismo do critério de verificabilidade do conhecimento [Lak68].

A prática demoliu a idéia clássica da validade das inferências de crescimento de conteúdo, e separou a dedução válida da prova informal e indução inválidas. Somente inferências que não acrescentavam conteúdo lógico poderiam ser validadas. Esse foi o fim da lógica da justificação do empirismo clássico.

A maioria dos empiristas se recusava a concluir que a ciência teorética era apenas uma ilusão. Mas como alguém poderia ser um bom empirista sem desistir da ciência? Alguns deles começaram a pensar na ciência como um conhecimento. Não um conhecimento no sentido clássico, mas num sentido mais fraco, como falível, conhecimento conjectural.

Então dois novos problemas surgiram. O primeiro era a avaliação do conhecimento conjectural que não poderia ser exato como na avaliação clássica. Nem mesmo estava claro se tal avaliação era possível; se ela teria que ser conjectural também; se existia a possibilidade de uma avaliação quantitativa; e tudo mais. O segundo problema era o crescimento do conhecimento conjectural.

Nessa situação, duas escolas de pensamento emergiram. Uma escola - o empirismo neo-clássico - iniciou com o primeiro problema e nunca chegou no segundo. A outra escola - o empirismo crítico - começou resolvendo o segundo problema e seguiu tentando mostrar que essa solução resolvia também os aspectos mais importantes do primeiro problema.

A primeira escola - culminando no empirismo neo-clássico de Carnap - abordou o problema do ponto de vista clássico da lógica da justificação. Já que era claro que as teorias não poderiam ser classificadas como provavelmente verdadeira ou falsa, elas teriam que ser classificadas como confirmadas (pelos fatos) em um certo grau. Foi pensado que esse grau de suporte evidencial ou grau de confirmação poderia de algum modo ser igualado com probabilidade no sentido do cálculo probabilístico.

Assim, Carnap, Reichenbach, e outros começaram a buscar as soluções para os seguintes problemas: (1) justificar que o grau de confirmação satisfazia aos axiomas da probabilidade de Kolmogorov; (2) encontrar e justificar requerimentos para a determinação da função de medida; (3) construir uma linguagem completa e perfeita na qual todas as proposições da ciência pudessem ser expressas e (4) oferecer uma definição de uma função de medida que satisfaria as condições ditas em (1) e (2) [Lak68].

Carnap pensava que enquanto a ciência era conjectural, a teoria de confirmação probabilística seria *a priori* e infalível: "a ciência é falível, mas o grau de sua falibilidade é precisamente e infalivelmente mensurável por uma máquina". No entanto, ele logo percebeu que mensurar o grau de falibilidade da ciência é, aparentemente impossível, e que a construção de uma linguagem completa para a ciência era possivelmente um processo sem fim, e que as dificuldades na construção da função de confirmação cresce passo a passo com a complexidade da linguagem.

Submerso em seu ambicioso programa, Carnap e sua escola ignoraram completamente o problema do método científico. Os problemas clássicos da indução eram a justificação de teorias e a descoberta de teorias a partir de fatos. A solução neo-clássica de Carnap no máximo provê uma solução para o problema da justificação fraca. Mas o problema da descoberta, do crescimento do conhecimento, foi deixado completamente intocado.

Por outro lado, de acordo com o empirismo clássico, a ciência acontece através da inferência indutiva: primeiro alguém coleta alguns fatos e então se faz uma inferência para uma teoria, o que os indutivistas chamam perfeitamente de generalização. De acordo com o empirismo crítico de Popper, começa-se com teorias especulativas que devem ser testadas severamente. E aqui existe apenas inferência dedutiva, e nenhuma generalização.

Nesse ponto, Carnap e Popper concordavam: os fatos não são necessariamente o ponto de partida da descoberta.

Em certo ponto dessa disputa intelectual entre Carnap e Popper, foi sugerido que o trabalho se dividisse: os carnapianos se concentrariam principalmente em uma reconstrução racional e sincrônica da ciência e os popperianos ficariam mais interessados no diacrônico crescimento da ciência. Essa divisão parecia implicar que os dois problemas eram independentes, mas na verdade, não o são. Segundo Lakatos, a falta de reconhecimento dessa interdependência é um importante defeito do empirismo lógico em geral e da teoria de confirmação de Carnap em particular.

Carnap seguiu considerando que tudo em que ele estava interessado era em como julgar teorias prontas, e não como descobrí-las; que o julgamento de teorias poderia ser reduzido ao julgamento de predições particulares, e a descoberta de teorias não poderia ser reduzida à descoberta de teorias particulares:

> "The task of inductive logic is not to find a law for the explanation of given phenomena. This task cannot be solved by any mechanical procedure or by fixed rules; it is rather through the intuition, the inspiration, and the good luck of the scientist. The function of inductive logic begins after a hypothesis is offered for examination. Its task is to measure the support which the given evidence supplies for the tentatively assumed hypothesis."[2] (Carnap, citado em [Lak68])

[2] "A tarefa da lógica indutiva não é encontrar uma lei para a explicação de um dado fenômeno. Essa tarefa não pode ser resolvida por um procedimento mecânico ou por regras fixas; é mais através da intuição, da inspiração, e da boa sorte do cientista. A função da lógica indutiva começa depois que uma hipótese é fornecida para ser examinada. Sua tarefa é medir o suporte que uma dada evidência dá para a hipótese assumida."

O programa de pesquisa de Carnap é detalhadamente descrito em [Lak68], inclusive com as interferências de Popper, sempre tentando refutar alguma teoria proposta por Carnap. Uma das idéias que Popper rejeitava, era a ligação da teoria de Carnap sobre o grau de suporte evidencial com a teoria da probabilidade.

Popper era também um nítido crítico do indutivismo. A seguinte passagem citada por Donald Gillies em [Gil94], deixa essa crítica bastante clara:

> "... my view of the matter, for what it is worth, is that there is no such thing as a logical method of having new ideas, or a logical reconstruction of this process. My view may be expressed by saying that every discovery contains 'an irrational element', or 'a creative intuition', in Bergson's sense. In a similar way, Einstein speaks of the 'search for those highly universal laws ... from which a picture of the world can be obtained by pure deduction. There is no logical path', he says, 'leading to these ... laws. They can only be reached by intuition, based upon something like an intellectual love of the objects of experience." [3](Popper, citado por Donald Gillies em [Gil94])

Nessa passagem Popper exclui a possibilidade de que possam existir regras de inferência indutiva de caráter lógico, e que pudessem ser usadas para obter leis científicas e generalizações. Popper acreditava que tais leis e generalizações só podem ser obtidas usando alguma intuição criativa irracional. Nisso Popper e Carnap concordavam plenamente.

Hoje, com os avanços na área de aprendizado de máquina, as posições de Popper e Carnap não são mais consideradas plausíveis. Nos anos 50, porém, suas posições eram bastante válidas, mas suas conhecidas divergências intelectuais não eram tantas como parecia a princípio [Gil94]. Carnap (o defensor da lógica indutiva) e Popper (o openente) concordavam que era inútil tentar formular regras de inferência indutiva. Além disso, ambos tentavam desenvolver uma teoria da confirmação, isto é, uma teoria que permitiria a um cientista calcular (ou estimar) o grau de confirmação que uma evidência e dava a uma hipótese h (simbolicamente, $C(h, e)$).

A diferença entre os dois filósofos é que para Carnap $C(h, e)$ obedecia os axiomas ordinários da

[3]"... minha visão do problema, o que é pior, é que não existe tal coisa chamada de método lógico para ter novas idéias, ou uma reconstrução lógica desse processo. Minha visão pode ser expressa dizendo que toda descoberta contém 'um elemento irracional', ou 'uma intuição criativa' no sentido de Bergson. De modo similar, Einstein fala sobre a 'busca de leis altamente universais ... a partir das quais uma visão do mundo pode ser obtida por pura dedução. Não há nenhum caminho lógico', ele diz, 'levando a essas ... leis. Elas só podem ser alcançadas por intuição, baseadas em algo do tipo um amor intelectual pelos objetos de experiência."

probabilidade, isto é, era uma função probabilística; enquanto Popper, um crítico do bayesianismo, argumentava que $C(h, e)$ não satisfazia os axiomas da probabilidade, embora ele acreditasse que $C(h, e)$ poderia ser definida em termos de probabilidades.

Então, por esse ponto de vista, a lógica indutiva tornara-se teoria da confirmação, um ramo de estudo que tinha fortes ligações com teoria da probabilidade, mas tinha muito pouco a ver com lógica dedutiva [Gil94].

Como as inferências indutivas são inválidas, elas originam um problema sério, que foi formulado explicitamente pela primeira vez por Hume: "Como se justificam as inferências indutivas?" A resposta de Hume é a de que não há justificação para a indução; ela se baseia apenas no hábito e seria, pois, irracional. Tal resposta pode ser estendida a todo mecanismo indutivo, mesmo o probabilístico.

No entanto, como o disse Russel: " É importante descobrir se existe alguma resposta a Hume dentro da estrutura de uma filosofia que é total ou principalmente empírica. Senão, não há nenhuma diferença intelectual entre a sensatez e a loucura" (Bertrand Russell, citado em [dC93]).

Não há dúvida de que o processo indutivo é falível. No entanto, uma das coisas fundamentais que buscamos na natureza, através da indução, é a regularidade. Somente por meio de regularidades é que conseguimos compreender a realidade e fazer previsões.

2.3 As perspectivas indutiva e dedutiva

A fórmula de D. Gillies 'lógica = inferência + controle', uma perspectiva recente na dicotomia 'lógica dedutiva versus lógica indutiva', aponta para um conceito chave presente na maioria dos trabalhos recentes em lógica e computação: 'inferência controlada'. A observação principal é que enquanto o controle começa a fazer parte da lógica dedutiva, a inferência controlada passa a ser a norma na lógica indutiva. Em termos de lógica e computação, inferência controlada pode ser vista como um tipo de conceito generalizado de dedução no nível objeto onde elementos do meta-nível têm um papel decisivo.

No período de Frege, por volta de 1879, até a década de 70 do nosso século, as lógicas dedutiva e indutiva tendiam a divergir e a tornar-se disciplinas profundamente diferentes em caráter. O efeito de alguns avanços em IA expandiram os conceitos de indução e dedução tornando-os mais similares que antes, permitindo o que Donald Gillies chama de reaproximação entre dedução e indução

[Gil94].

A partir da época de Frege, havia uma tendência crescente de ver-se a lógica dedutiva como consistindo de regras de inferência. Lógica indutiva também era vista como consistindo de inferências, com a diferença de que trabalhava a partir de uma massa indefinida de casos particulares para alcançar uma conclusão geral.

A principal crítica à indução era que as premissas da inferência poderiam ser todas verdade, mas a conclusão falsa. Porém, apesar das conclusões das inferências indutivas não serem certas, elas são sem dúvida prováveis dadas as premissas. O desenvolvimento do bayesianismo, porém, fez com que o problema da indução deixasse de ser o de inferir conclusões indutivas de premissas factuais para o de calcular a probabilidade de uma hipótese dadas algumas evidências relevantes (como visto na seção 2.2). Nos anos 70, então, a divergência entre as lógicas dedutiva e indutiva eram representadas da seguinte maneira: a primeira era crescentemente pensada quase que exclusivamente em termos de regras de inferência, e os trabalhos na área de lógica indutiva tinham desistido de encontrar regras de inferência indutiva e estavam tentando, usando teoria da probabilidade, encontrar formas de estimar com que grau uma evidência confirmava uma hipótese.

Por volta de 1979, Kowalski, quando desenvolvendo suas idéias sobre programação em lógica, propôs a seguinte fórmula:

$$\text{ALGORITMO} = \text{LÓGICA} + \text{CONTROLE}$$

Segundo sua fórmula, o componente CONTROLE está fora do programa que expressaria somente o componente LÓGICA.

Gillies, por seu lado, propôs a fórmula:

$$\text{LÓGICA} = \text{INFERÊNCIA} + \text{CONTROLE}$$

Gillies argumenta [Gil94] que em termos de sua nova fórmula, no desenvolvimento de lógica clássica fora da área de IA, a atenção foi largamente voltada para o componente de INFERÊNCIA. O principal problema era o de analisar a lógica de provas matemáticas, que seriam construídas pelos humanos. Isto é, o CONTROLE necessário na construção de uma prova era deixado para os matemáticos humanos.

Por outro lado, filósofos estudando a lógica indutiva concentravam-se em teoria da confirmação, que em termos da nova fórmula de Gillies é parte do componente de CONTROLE da lógica indutiva. A inferência necessária era deixada aos cientistas humanos.

Então, na lógica dedutiva a inferência era estudada enquanto na lógica indutiva, o controle era o principal objeto de estudo. A reaproximação entre os dois ramos da lógica tornou-se possível desde que a IA introduziu no lado dedutivo o componente de CONTROLE, através do estudo de prova automática de teoremas, que levou ao surgimento de sistemas de programação em lógica como PROLOG. No outro lado, o estudo de aprendizado automático, desenvolvendo-se no campo da lógica indutiva, trouxe a introdução de INFERÊNCIA na lógica indutiva.

2.3.1 Controle na lógica dedutiva

O PROLOG foi desenvolvido usando lógica clássica de predicados, mas logo percebeu-se que a negação do PROLOG, a chamada negação por falha, era diferente da negação clássica [Gil94]. Isso deu ao PROLOG uma característica não-monotônica: em lógica clássica, se uma conclusão segue de um conjunto de premissas e acrescentamos uma nova premissa ao conjunto, a conclusão ainda segue. Isso dá à lógica clássica seu caráter monotônico. No PROLOG, a adição de novas premissas ao conjunto pode algumas vezes significar que a conclusão que tinha sido deduzida previamente, não mais segue do novo conjunto de premissas, então o PROLOG não preserva a verdade, perante a adição de novas premissas.

O outro ponto importante é que a lógica clássica pode ser vista como a mecanização do processo de verificação da validade de uma prova, mas ela deixa a construção da prova inteiramente nas mãos dos matemáticos humanos. PROLOG leva o processo de mecanização um passo adiante, mecanizando a construção das provas. Para fazer isso, o PROLOG contém um conjunto de instruções para pesquisa sistemática através de várias possibilidades. Essas instruções são claramente parte de um sistema de controle que foi adicionado aos procedimentos de inferência da lógica clássica.

De fato, PROLOG pode ser corretamente considerado como um *sistema lógico* que contém não somente uma linguagem formal, mas também regras de inferência e um elaborado mecanismo de controle projetado para executar pesquisas e construir provas.

2.3.2 Inferência na lógica indutiva

Como vimos anteriormente, o ponto principal do aprendizado de máquina [CMM83] é o estudo e modelação do processo de aprendizado nas suas múltiplas manifestações. Aprendizado indutivo é uma classe de sistemas de aprendizado de máquina que pode ser vista como a busca de descrições gerais plausíveis (assertivas ou regras indutivas) que explicam dados fornecidos e são úteis para predizer novos dados [Mic83].

Na década de 80, o ID3 de Quinlan, um programa de aprendizado de máquina, representou um grande passo no campo. Ele tinha um algoritmo para construir uma árvore de decisão a partir de dados fornecidos. Seu sucesso foi considerável mas sua limitação mais significativa é que a técnica lida apenas com objetos descritos por um conjunto de predicados unários e induz somente regras de classificação na forma de árvores de decisão [Gil94].

Muggleton e Feng desenvolveram um programa de aprendizado de máquina, chamado GOLEM, que é um procedimento para construir o que é conhecido como a menor generalização geral relativa (*relative least general generalization* ou r.l.g.g. - um conceito introduzido por Plotkin e baseado na definição de uma ordem parcial da generalidade das cláusulas). Esse sistema de aprendizado de sucesso envolve regras de inferência indutiva de caráter lógico que permitem a inferência de leis e generalizações a partir de dados.

Sumarizando, a lógica dedutiva estava concentrada em regras de inferência e ignorava questões de controle; enquanto a lógica indutiva preocupava-se em definir medidas de confirmação (parte do controle) e desistia da tarefa de formular regras de inferência indutiva, considerada sem esperanças. A programação em lógica introduziu controle na lógica dedutiva enquanto o estudo de aprendizado de máquina levou à formulação de regras de inferência indutiva. Essa é a chamada reaproximação entre lógica dedutiva e lógica indutiva.

Gillies [Gil94] conclui a idéia da seguinte forma:

> "This is not to say of course that the two fields have completely merged. Even in the AI context, deduction looks very different from induction. The main difference is perhaps this. While there are a whole variety of quite simple and straightforward rules of inference for deductive inference, the successful rules of inductive inference (...) are relatively few in number, and complicated in character. Of course with the continuing study of machine learning this situation will change. Almost certainly more successful rules of inductive inference will be discovered, and their nature and interrelation will become clarified." [4](Donald Gillies em [Gil94])

[4]"Não quer dizer que os dois campos mesclaram-se completamente. Mesmo no contexto da IA, a dedução é vista de forma bem diferente da indução. A principal diferença talvez seja essa. Enquanto existe toda uma variedade de regras de inferência dedutiva, as regras bem sucedidas para a inferência indutiva (...) são relativamente poucas em número, e complicadas em caráter. Claro que com a continuação no estudo de aprendizado de máquina essa situação

Essa perspectiva de Gillies justifica, de certo modo, a utilização de sistemas dedutivos para tentar solucionar problemas indutivos. O mesmo se dá com problemas abdutivos: como veremos na seção 2.4, podemos ver a indução com um caso particular da abdução, por isso, podemos estender as idéias de Gillies para o caso da abdução e justificar a nossa proposta de um método abdutivo utilizando os sistemas dedutivos rotulados.

2.4 A indução como caso especial da abdução

Como veremos no capítulo 3, a abdução é um raciocínio do senso comum que busca explicações para fatos observados, considerando o conhecimento de *background* existente. Nesta seção, tentaremos mostrar que a indução também pode ser vista como uma busca de explicações, e, por isso, pode ser considerada um caso especial da abdução.

A abdução é muito bem aplicada no caso de de diagnóstico médico. Por exemplo, podemos ter o seguinte problema:

Um paciente apresenta o sintoma S e desejamos explicar a presença desse sintoma diagnosticando uma determinada enfermidade. Para isso temos que ter conhecimento das enfermidades que causam o sintoma S. Essa informação deverá estar presente no nosso conhecimento de *background* da seguinte forma:

$$\Theta = \{ \begin{array}{l} D_1 \rightarrow S \\ D_2 \rightarrow S \\ D_3 \rightarrow S\} \end{array}$$

Assim, podemos abduzir que a explicação para a presença de S é D_1 ou D_2 ou D_3, de acordo com as seguintes regras de inferência abdutiva:

$$\frac{D_1 \rightarrow S \quad S}{D_1} \quad \frac{D_2 \rightarrow S \quad S}{D_2} \quad \frac{D_3 \rightarrow S \quad S}{D_3}$$

Tentando utilizar a dedução para solucionar um problema abdutivo, teríamos que buscar as fórmulas que, acrescentadas à teoria disponível, completariam uma prova para o fato observado. Então o nosso esquema dedutivo para a abdução seria:

irá mudar. Quase que certamente, mais regras bem sucedidas de inferência indutiva serão descobertas e sua natureza e interrelação serão esclarecidas."

$$\frac{D_1 \quad D_1 \rightarrow S}{S}$$

onde, seguindo da conclusão para as premissas, temos S como um fato observado (a nossa conclusão), na nossa teoria temos que $D_1 \rightarrow S$, tudo que precisamos, então, é a fórmula D_1 para completar dedutivamente uma prova de S. Logo, D_1 é a nossa hipótese abdutiva.

O mesmo se daria com as outras hipóteses D_2 e D_3:

$$\frac{D_2 \quad D_2 \rightarrow S}{S} \qquad \frac{D_3 \quad D_3 \rightarrow S}{S}$$

Qual seria o papel da indução num domínio semelhante: seria o de gerar as regras que utilizamos na abdução. Ou seja, quando estudados vários casos da doença D_1, verificou-se que em todos os casos (ou na maioria deles - nesse nomento entrariam as questões probabilísticas da indução), o sintoma S estava presente.

O esquema normalmente apresentado para a indução seria o seguinte:

$$\begin{array}{c} D_1(x_1) \quad S(x_1) \\ D_1(x_2) \quad S(x_2) \\ \vdots \\ D_1(x_n) \quad S(x_n) \\ \hline D_1(x) \rightarrow S(x) \end{array}$$

Ou seja, por indução, posso criar uma regra que nos diz que qualquer que seja o meu paciente x, se x tem a doença D_1, ele apresentará o sintoma S.

O mesmo se daria para as outras doenças D_2, D_3, etc.

Considerando que estamos trabalhando em um domínio específico, podemos dizer que os dados que temos como entrada para a indução já estão, de algum modo, em forma de implicação, ou seja: teríamos algo do tipo: a doença D_1 no paciente x_1 implica em um sintoma S, a doença D_1 no paciente x_2 implica em um sintoma S, e assim por diante. Então, ao invés de considerarmos o esquema da indução como o vimos acima, vamos modificá-lo um pouco, sem, no entanto, alterar o seu significado básico:

$$\frac{\begin{array}{c} D_1(x_1) \rightarrow S(x_1) \\ D_1(x_2) \rightarrow S(x_2) \\ \vdots \\ D_1(x_n) \rightarrow S(x_n) \end{array}}{\forall x.D_1(x) \rightarrow S(x)}$$

Obtemos assim, uma generalização, bastante similar ao esquema anterior, simbolizando a indução. É claro que existem algumas restrições para que essa generalização seja efetivada e considerada correta. Aí entram as questões sobre, por exemplo, a super generalização, que deve ser evitada com a utilização de contra-exemplos, as questões probabilísticas, onde a depender do método utilizado, a regra geral encontrada deve satisfazer a um certo percentual de casos estudados, e assim por diante; Plotkin, por exemplo, tratava essa questão utilizando o conceito de *rlgg - relative least general generalization* - que busca uma generalização geral o suficiente para abranger todos os casos e que ao mesmo tempo seja mínima para que possa ser considerada útil.

Observando o esquema final que propomos para a indução, podemos entender uma generalização, no caso, a regra $\forall x.D_1(x) \rightarrow S(x)$, como a explicação para a veracidade das instâncias da regra (as premissas). E isso nada mais é que uma abdução.

Tentando esquematizar agora a indução em termos de dedução, como o fizemos com a abdução, o que teríamos:

$$\frac{\forall x.D_1(x) \rightarrow S(x)}{(D_1(x_1) \rightarrow S(x_1)) \wedge (D_1(x_2) \rightarrow S(x_2)) \wedge \ldots \wedge (D_1(x_n) \rightarrow S(x_n))}$$

Mais uma vez, o que fizemos foi encontrar uma regra ($\forall x.D_1(x) \rightarrow S(x)$) que explicasse os fatos observados.

Com essa argumentação chegamos à conclusão que a indução nada mais é que uma abdução (de primeira ordem) com algumas restrições específicas. A principal dessas restrições é que os elementos do conjunto de fatos observados devem apresentar uma grande similaridade entre si para que se possa explicar a presença desses fatos (similares) com a introdução da fórmula com o quantificador universal. Com isso chegamos ao ponto principal que queremos abordar nesta seção que é: aonde está a ligação entre a indução e a abdução? Justamente na conclusão que tiramos: a indução é um caso particular da abdução de primeira ordem.

Assim colocado, todos os conceitos ligados à indução e, consequentemente, ao aprendizado indutivo, são relevantes para um estudo em abdução, principalmente porque esta última só recentemente tem despertado maior interesse na comunidade científica e, por isso, suas bases são muito pouco definidas.

Em resumo, nos valemos, então, dessa importante ligação entre indução e abdução para estudarmos os conceitos relevantes para o aprendizado indutivo sempre tendo por trás o ponto de vista da abdução. Sob esse ponto de vista, podemos, sempre que falarmos em indução, visualizar um conceito mais abrangente, o de abdução.

2.5 O aprendizado indutivo

Nesta seção discutiremos várias características do aprendizado indutivo com o objetivo de nos familiarizarmos com diversas abordagens já propostas para esse tipo de aprendizado. No final da seção descreveremos rapidamente alguns sistemas de aprendizado indutivo, e paulatinamente tentaremos fazer uma analogia entre o aqui visto e os métodos abdutivos. No decorrer da dissertação perceberemos a similaridade entre diversos pontos aqui discutidos para o aprendizado indutivo e as características do raciocínio abdutivo.

O processo de aprendizado indutivo pode ser visto como a busca de descrições gerais plausíveis (assertivas indutivas) que explicam os dados de entrada e são úteis para predizer novos dados. Para que um programa de computador possa formular esse tipo de descrição deve-se usar uma linguagem de descrição apropriada. Para qualquer conjunto de dados e qualquer linguagem de descrição não trivial, um grande número de assertivas indutivas pode ser formulado. Essas assertivas formam um conjunto de descrições parcialmente ordenado pela relação de generalidade relativa [DM83]. Os elementos minimais desse conjunto são as descrições mais específicas dos dados de entrada na linguagem dada, e os elementos maximais são as descrições mais gerais dos dados de entrada.

Segundo Richard Forsyth [For89], a indução de máquina abre a possibilidade de sintetizar conhecimentos totalmente novos, de automatizar o processo da descoberta científica e a criação de padrões e conceitos que não se havia pensado antes.

Colocado de maneira simples, aprender é fundamental para o comportamento inteligente.

A indução, isto é, a derivação de leis gerais através do exame de instâncias particulares, é tão importante quanto intratável, pelo menos para uma mente escolada em raciocínio silogístico: ninguém conseguiu até hoje fazer com que a indução preserve a verdade, exceto em um sistema fechado.

Segundo Richard Forsyth [For89], não só a indução forma a base do dia a dia, como também é o fundamento para tudo na descoberta científica. Apesar disso, nunca houve uma prova da inferência indutiva. O problema é que regras indutivamente geradas não podem ser provadas, e esse antigo problema filosófico povoou as mentes de grandes pensadores da antigüidade que tentavam inutilmente uma maneira de validar as inferências indutivas. Infelizmente, sob o ponto de vista de um engenheiro de software, esses filósofos estavam menos preocupados em como fazer induções que em como justificá-las (como vimos na seção 2.2). Seu objetivo era especificar regras governando o argumento indutivo da mesma maneira que os lógicos o fizeram para o argumento dedutivo.

Francis Bacon, por exemplo, era um desses filósofos, e teve uma visão bastante moderna para a época, das deficiências de qualquer ciência puramente dedutiva [For89]. Ele chamou a atenção para a necessidade da indução estar sempre baseada nos fatos observados e também para a taxa de erro (normalmente subestimada) de qualquer inferência indutiva. Também chamou a atenção firmemente para o problema da super generalização.

John Stuart Mill, um outro filósofo preocupado com a indução, citado em [For89], porpôs quatro métodos experimentais primários para induzir regras gerais a partir de casos particulares:

1. **Método da concordânca**: Se dois ou mais exemplos de um fenômeno sob observação têm apenas um fator em comum, esse único fator no qual todos os exemplos concordam é a **causa** ou efeito do dado fenômeno;

2. **O método das diferenças**: Se uma instância positiva de um fenômeno em observação e uma instância negativa do mesmo fenômeno têm todas as circunstâncias em comum, exceto uma, essa única circunstância na qual os exemplos diferem é a causa ou efeito, ou uma parte indispensável da causa do fenômeno em questão;

3. **O método dos resíduos**: Remova de qualquer fenômeno qualquer parte dele conhecida como efeito de certos antecedentes e o resto do fenômeno é o efeito dos antecedentes restantes;

4. **O método da variação concomitante**: Se um fenômeno varia regularmente de alguma maneira sempre que outro fenômeno varia de algum modo particular, o primeiro está conectado com o segundo por alguma cadeia de causas.

Curiosamente, Mill coloca seus quatro métodos cardinais de indução em termos de causa e efeito: a base inquestionável da inferência abdutiva.

Segundo Forsyth, os dois primeiros métodos (que devem ser empregados conjuntamente) trabalham melhor se não existe nenhuma incerteza na cadeia causal que se está tentando explicar. O terceiro método é melhor interpretado como um conselho heurístico para o investigador. O quarto método é o único que faz mais sentido se se admite dados numéricos e informação incerta.

Bertrand Russel foi outro filósofo que, segundo Forsyth, se preocupou com o problema de justificar a indução. Ele acreditava que o princípio indutivo era impossível de ser provado, no entanto, o seu papel no raciocínio humano é tão fundamental que devemos aceitar o princípio indutivo em termos de sua intrínseca evidência ou esquecer qualquer justificativa de nossas espectativas sobre o futuro. Para ele, diferentemente das idéias de Mill, o princípio indutivo era essencialmente probabilístico.

O ponto de vista de Wittgenstein sobre a filosofia da indução, segundo Forsyth, visava principalmente enfatizar a simplicidade: o procedimento de indução consistia em aceitar como verdade a lei mais simples que pode ser "conciliada" com nossa experiência.

Mais uma vez os princípios associados com a indução (nesse caso a simplicidade da lei escolhida), assemelham-se com aspectos do raciocínio abdutivo que, como veremos no capítulo 3, tem como uma de suas características principais a escolha de explicações o mais simples possível.

Colocando os pés um pouco mais na nossa realidade atual, a indução é vista, segundo [DM83], como uma busca dentro de um espaço de descrições generalizadas. Dietterich e Michalski chamam a atenção sobre os seguintes aspectos do aprendizado de máquina como um todo, e , mais especificamente, do aprendizado indutivo:

1. **Representação** - Muitos sistemas representacionais podem ser usados para representar eventos e generalizações de eventos (como vimos na seção 2.1). Por exemplo, cálculo de predicados, regras de produção, descrições hierárquicas, etc. Segundo [DM83] a maioria dos trabalhos em IA em aprendizado indutivo tem utilizado o cálculo de predicados (ou sistemas similares) por ser bem definido sintática e semanticamente.

2. **Tipo de descrição buscada** - Já que a indução é uma busca em um espaço de descrições, devem ser especificados critérios que definam com que objetivo as assertivas indutivas são buscadas. Esses critérios dependem do domínio em questão, mas alguns critérios gerais podem ser citados:

   Pode-se buscar uma **descrição característica** que descreve uma classe de objetos de maneira que os fatos incluídos na descrição são verdadeiros em todos os objetos da classe. Normalmente esse tipo de descrição tem a intenção de discriminar os objetos da classe dada dos objetos de qualquer outra possível classe. Por exemplo, uma descrição característica do conjunto de todas as cadeiras irá discriminar qualquer cadeira de tudo aquilo que não é cadeira. Nesse caso, a descrição caracteriza o conceito de cadeira. Já que é impossível examinar todos os objetos da classe (ou que não pertencem à classe), uma descrição característica normalmente é criada especificando-se todas as características que são verdadeiras em todas as instâncias conhecidas da classe. Em alguns problemas, contra-exemplos estão disponíveis para ajudar na delimitação da classe. Os contra-exemplos mais úteis são os conhecidos como *near misses*, isto é, aqueles que diferem dos exemplos apenas por pequenos detalhes.

Um outro tipo é a **descrição discriminante** que descreve uma classe de objetos no contexto de um conjunto fixo de outras classes. Especifica somente aquelas propriedades dos objetos da classe dada que são necessárias para distinguí-los dos objetos das outras classes.

Podemos ter uma **descrição taxonômica** que é uma descrição que subdivide uma classe em subclasses. Na construção desse tipo de descrição, é assumido que os dados de entrada são membros de diferentes classes. Uma descrição taxonômica é fundamentalmente disjuntiva; a classe superior é descrita pela disjunção das descrições das subclasses.

3. **Regras de generalização** - O espaço parcialmente ordenado de descrições de diferentes níveis de generalidade pode ser descrito pela indicação das transformações que são aplicadas para modificar descrições menos genéricas transformando-as em outras mais genéricas. Consequentemente, a determinação de assertivas indutivas pode ser vista como um processo de aplicação consecutiva de certas 'regras de generalização' sobre descrições iniciais e intermediárias. Em [DM83] é proposta uma formalização para a aplicação de regras de generalização onde tendo-se uma regra de classificação do tipo $S_1 ::> K$ (todos os objetos para os quais S_1 é verdade pertencem à classe K) a aplicação de uma regra de generalização vai gerar uma outra regra de classificação mais geral do tipo $S_2 ::> K$, significando que a implicação $S_1 \Rightarrow S_2$ é verdadeira.

   O conceito de regras de generalização provê a visão da indução como uma busca heurística. As regras de generalização especificam os operadores que a busca utiliza para passar de um modo a outro no espaço de descrições. Várias regras desse tipo são descritas em [Mic83] e algumas serão discutidas na seção 2.5.2 .

4. **Indução construtiva** - A indução construtiva é um tipo de indução que gera novos descritores, não presentes nos dados de entrada. Os métodos que não executam a indução construtiva

são ditos executar a indução seletiva, já que as generalizações são feitas a partir de seleções de informações presentes nos dados de entrada.

A indução construtiva, embora mais complexa, é importante já que muitos problemas de IA não podem ser solucionados sem uma mudança de representação nos dados de entrada.

5. **Estratégia de controle** - A estratégia de controle usada para buscar o espaço de descrições pode ser *bottom-up* (dirigida aos dados) ou *top-down* (dirigida ao modelo), ou uma mistura de ambas. Métodos *bottom-up* processam os dados de entrada um de cada vez, generalizando gradualmente o conjunto de descrições até que uma generalização conjuntiva final é encontrada.

   Os métodos *top-down* buscam um conjunto de possíveis generalizações na tentativa de encontrar umas poucas hipóteses que melhor satisfaçam certos requisitos (critérios de preferência ou restrições). A busca é feita escolhendo-se como hipótese inicial alguns elementos do conjunto parcialmente ordenado das possíveis descrições. Se a hipótese escolhida satisfaz os critérios exigidos, então a busca pára. Caso contrário a hipótese atual é modificada (através de generalizações ou especializações) e a nova hipótese é checada. Esse processo continua até que o critério exigido seja satisfeito. Esse tipo de método é mais imune a ruídos que o anterior, no entanto tem a desvantagem de que a hipótese corrente deve ser repetidamente checada para todos os dados de entrada.

6. **Métodos gerais e orientados ao problema** - Na prática, os métodos aplicados a problemas específicos parecem (e são em verdade) mais eficientes; no entanto, o desenvolvimento apenas desse tipo de método leva a diversas dificuldades de integração entre diferentes pesquisas. Cada projeto utiliza suas próprias formalizações e maneiras de representação, levando inclusive a uma prejudicial redundância de terminologias. Os métodos teóricos mais genéricos, apesar de considerados por muitos como apenas teoricamente elegantes e não úteis na prática, são importantes no sentido de promoverem uma maior integração entre projetos diferentes, o que sem dúvida favorece a evolução da pesquisa na área de aprendizado de máquina. Esse tema (a importância de teorias de aprendizado) é considerado com bastante relevância em [MAL+86].

As seis características discutidas acima podem ser utilizadas como critérios de avaliação de métodos de aprendizado de máquina. Além disso, nos dão uma idéia das várias possibilidades dentre as quais podemos escolher no momento de propor um sistema de aprendizado de máquina.

O paradigma geral para o aprendizado indutivo, formulado em [Mic83], é que, sendo dados um conjunto de declarações observáveis (fatos que chamaremos de F), uma declaração indutiva inicial que poderá ser eventualmente nula, e um conhecimento de *background* consistindo de informações relevantes ao problema e de algum tipo de restrição e critérios de preferência, possa-se chegar a uma hipótese indutiva H, que tautológica ou fracamente implica nos fatos F[5], e satisfaz o conhecimento de *background*.

É inevitável fazermos, mais uma vez, analogia com o raciocínio abdutivo. O paradigma da indução como descrito acima, pode, perfeitamente, ser dito o paradigma da abdução: Dados um conjunto de declarações observáveis F, uma hipótese abdutiva inicial que poderá ser eventualmente nula, e uma teoria de *background* consistindo de, além de informações relevantes ao problema, algum tipo de restrição e critérios de preferência caracterizando as propriedades desejadas para a hipótese abdutiva, o processo deve encontrar uma hipótese abdutiva H, que tautológica ou fracamente implica nos fatos F, e satisfaz o conhecimento de *background*.

Segundo Michalski, permitir a condição de H implicar F só fracamente faz com que esse paradigma inclua métodos probabilísticos de geração de hipóteses, onde as hipóteses propostas atenderiam não todos mas parte dos fatos observados, dentro, é claro, de algumas restrições.

O conhecimento de *background* entra para restringir o conjunto de hipóteses geradas: para qualquer conjunto de fatos, um número potencialmente infinito de hipóteses que implicam esses fatos pode ser gerado; o conhecimento de *background* é então necessário para prover restrições e critérios de preferência para reduzir o conjunto infinito para algumas hipóteses preferenciais.

2.5.1 Importância de teorias de aprendizado

Muito tem-se pesquisado na área de aprendizado automático, pesquisas, no entanto, bastante desintegradas no sentido de que são elaboradas de modo muito independente uma das outras; daí a falta de consenso nas terminologias. Além disso, muito das pesquisas nessa área são orientadas a uma aplicação específica, que, sem querer tirar o mérito inegável de pesquisas com esse objetivo, levam a uma falta de integração e ausência de objetivos teóricos.

A importância da interconexão dos problemas de aprendizado, especialmente na área de descoberta de conceitos é enorme, e a necessidade da integração de trabalhos é reconhecida amplamente

[5] H implica F tautologicamente se F é uma conseqüência lógica de H ($H \rightarrow F$) e H implica F fracamente se F é incerto ou é conseqüência parcial de H

(Amarel e Lenat em [MAL+86]). Isso nos leva a uma outra linha de pensamento que é o de criar-se teorias de aprendizado com métodos multi-aplicativos (Michalski em [MAL+86]) e não voltados a aplicações específicas. Além disso deve-se estar ciente da necessidade da busca de terminologias e linguagens de descrição generalizadas, o que nos levaria exatamente a uma maior integração dos diversos métodos estudados.

O objetivo desta dissertação é exatamente a proposta de uma teoria de aprendizado abdutivo que possa ser utilizada para lógicas e métodos diversos, na busca de uma forma de descrição de problemas abdutivos bastante generalizada. Com isso podemos chegar a uma formalização geral desse tipo de aprendizado.

2.5.2 Algumas abordagens em indução

A utilização de censores para aperfeiçoar as regras

Uma das abordagens que comentaremos nesta seção é a descrita em [Win86] por Patrick H. Winston. Sua proposta é uma extensão de um sistema de aprendizado de máquina por analogia, onde **censores** (*censors*) são adicionados às regras inicialmente geradas a fim de evitar aplicações errôneas das mesmas. O sistema de aprendizado por analogia, sobre o qual é proposta a extensão, apresenta as seguintes características:

- **Raciocínio baseado em analogia usando transferência de restrições** - o raciocínio por analogia (como vimos na seção 2.1), requer a habilidade de determinar como duas situações que são similares com respeito a alguns pontos explícitos, podem assemelhar-se com respeito a outros pontos também. Nesse sistema isso é feito através da transferência de restrições de causa da situação precedente para a situação em exercício.

- **Regras IF-THEN aprendidas** - As regras do tipo IF-THEN emergem automaticamente quando os problemas são resolvidos. O professor suprime o precedente e o exercício deixando o trabalho de formular as regras com o sistema.

- **Raciocínio baseado em regras** - Uma vez aprendidas, as regras podem ser utilizadas como dados. Já que as regras são vistas como situações simples, os programas de transferência de restrições que trabalham com situações precedentes também trabalham com regras. Assim aprendizado e raciocínio residem juntos, harmoniosamente no mesmo sistema.

- **Representação objeto-ator** - As situações são representadas usando relações. Cada relação tem TRUE, FALSE ou UNKNOWN como seu valor verdade e qualquer relação pode ser um objeto envolvido em outra relação.

- **Combinação de importância dominante** - A similaridade entre duas situações é medida pela busca da combinação melhor possível de acordo com o que é importante na situação precedente. Uma relação precedente é considerada importante se ela está conectada a uma outra relação por uma restrição de determinação de importância. A única restrição de determinação de importância considerada nesse trabalho é a conexão de relação de causa.

A proposta de Winston em [Win86] é estender um sistema de aprendizado por analogia, com as características descritas acima, ampliando cada regra, no momento de sua geração, com entradas que correspondem às negações de todas as relações na estrutura causal transferida. Essas entradas são chamadas de "censores" (*censors*) e constituem a parte UNLESS da regra. Funcionarão de acordo com um princípio de bloqueio de regra segundo o qual se uma entrada na parte UNLESS da regra corresponde a alguma coisa que é manifestadamente verdade, então a regra não se aplica.

As regras IF-THEN ampliadas teriam, então, o seguinte formato:

$$\begin{array}{ll} \text{IF} & A_1 \wedge A_2 \wedge ... \wedge A_n \\ \text{THEN} & C \\ \text{UNLESS} & B_1 \vee B_2 \vee ... \vee B_m \end{array}$$

Pode parecer correto descrever essa regra logicamente da seguinte forma:

$$A_1 \wedge A_2 \wedge ... \wedge A_n \wedge \neg(B_1 \vee B_2 \vee ... \vee B_m)$$

No entanto, a notação lógica é enganosa, porque, no caso das regras IF-THEN ampliadas, os A's e B's são tratados diferentemente, em contraste com as convenções da lógica tradicional: um esforço ilimitado deve ser colocado no intuito de mostrar que os A's são verdade, e é feita apenas uma tentativa de mostrar que os B's são verdade, com os B's sendo assumidos como falsos em caso de falha.

O seguinte exemplo, tirado de [Win86], torna clara a idéia de utilizar censores:

Regra-1	IF	There is a man
		and the man is not married
		and the man is an adult
	THEN	the man is a bachelor
	UNLESS	the man is not expected to be married
		or the man is not able to be married
Censor-1	IF	A man is misogamist
	THEN	the man is not expected to be married
Censor-2	IF	A man is a priest
	THEN	the man is not able to be married

Os censores são testados para averiguar a parte UNLESS da regra.

A principal justificativa para a utilização de censores (e para não incorporá-los à regra) é que normalmente eles definem situações de certa maneira incomuns, mas que contribuem para aperfeiçoar o conceito aprendido.

Fazendo uma analogia com o raciocínio abdutivo e suas características, os censores podem ser vistos como restrições de integridade que devem ser consideradas mas que, no entanto, não fazem parte da teoria (veja seção 3.2).

A busca de regularidades

O sistema de aprendizado voltado para descoberta, BACON-6, descrito em [LZSB86], focaliza a descoberta de leis empíricas que sumarizam dados numéricos. Esse programa faz buscas em um espaço de dados e um espaço de leis numéricas.

Dado um conjunto de valores independentes e um conjunto correspondente de valores dependentes, o sistema tenta encontrar uma ou mais leis que predizem os valores observados o mais acuradamente possível. Para isso, BACON-6 requer informações sobre a forma da lei pretendida, por exemplo: $y = ax^2 + bx + c$. A partir dessa informação, o sistema gera um conjunto de estados iniciais para começar a busca. Esses estados são gerados com a atribuição de valor 1 ou -1 a cada coeficiente de cada tipo de fórmula fornecida.

Incrementando ou decrementando os parâmetros, as fórmulas vão sendo aperfeiçoadas até que sejam selecionados os N melhores estados. A estratégia de busca utiliza os dados somente para testar as hipóteses, e não para gerá-las (já que as hipóteses iniciais são fornecidas sob a forma das leis pretendidas) o que, segundo os autores, torna o sistema mais robusto com respeito a ruído numérico.

No entanto, ao nosso ver, a estratégia do BACON-6 é uma busca exaustiva dos melhores coeficientes para abranger a maior parte dos dados observados. Uma proposta dos próprios autores é que uma extensão do BACON-6 poderia integrar algum tipo de interação com o usuário para que este, de certa forma, direcione a busca feita pelo sistema.

O ponto desse trabalho que desejamos chamar a atenção é a questão das hipóteses iniciais fornecidas ao sistema. No sistema abdutivo que propomos, existe um espaço para o fornecimento de hipóteses iniciais (no nosso caso, candidatos fornecidos pelo usuário - veja seção 5.4.1) e em momento oportuno justificaremos a existência dessa opção.

Ainda em [LZSB86], outro dos sistemas descritos é o GLAUBER, concernente à descoberta de leis de estrutura qualitativa, tal como hipóteses de que ácidos reagem com alcalinos para formar sais. O sistema busca em um espaço de leis qualitativas, usando funções de avaliação para dar preferência a leis que cubram o maior número possível de fatos observados.

No exemplo descrito em [LZSB86], são fornecidas uma série de reações químicas observadas, do tipo:

HCl	reage com	$NaOH$	e gera	$NaCl$
HCl	reage com	KOH	e gera	KCl
HNO_3	reage com	$NaOH$	e gera	$NaNO_3$

Além disso são fornecidas determinadas qualidades dos objetos que fazem parte da reação:

HCl	tem sabor	azêdo
HNO_3	tem sabor	azêdo
$NaOH$	tem sabor	amargo
KOH	tem sabor	amargo
$NaCl$	tem sabor	salgado
KCl	tem sabor	salgado
$NaNO_3$	tem sabor	salgado

A partir desses dados o sistema cria classes (ácidos, alcalinos e sais). Seu objetivo é generalizar conceitos qualitativos do tipo:

a classe X reage com a classe Y e gera a classe Z

Para atingir esse objetivo, o GLAUBER inicialmente formula regras quantificadas existencialmente do tipo:

$$\exists \text{ sal tal que (HCL reagindo com NaOH gera sal)}$$
$$\vdots$$

E em seguida tenta generalizar essas regras (ou leis qualitativas) utilizando o quantificador universal:

$$\forall \text{ alcalino } \forall \text{ ácido } \exists \text{ sal tal que (alcalino reage com ácido e gera sal)}$$
$$\vdots$$

A heurística utilizada pelo sistema é, primeiramente, a busca de características que ocorrem na maioria dos fatos. Além disso, aquelas leis que permanecem com o quantificador existencial, ou seja, aquelas para as quais não se conseguiu uma generalização para o quantificador universal, são desconsideradas. Essa última heurística pode ser vista, como os próprios autores colocam, como um ponto negativo do sistema, já que algumas leis importantes, embora não gerais, podem ser perdidas no final do processo.

Esse sistema caracteriza muito bem a busca da generalização da indução, a busca de regras quantificadas universalmente. E vem fortalecer a idéia da indução como uma abdução de primeira ordem (veja seção 2.4).

Um outro ponto importante a observar sobre os sistemas BACON-6 e GLAUBER é que eles também apresentam a característica, de certa forma, probabilística da indução quando buscam generalizações que abrangem a maioria dos casos observados. No GLAUBER, inclusive, são utilizadas funções de avaliação com o intuito de buscar aquelas características que ocorrem na maioria dos fatos.

Utilizando *APC - Annotated Predicate Calculus*

Em [Mic83], Michalski define o aprendizado por exemplos como um processo de aquisição de conhecimentos pela criação de inferências indutivas a partir dos fatos fornecidos. A inovação de Michalski está na forma como são trabalhados esses fatos fornecidos, para se obter uma hipótese indutiva: segundo ele, sobre os fatos, que são nosso conhecimento de *background*, serão aplicadas inferências do tipo: generalização, especialização, transformação, correção, refinamento, que proverão a aquisição de novos conhecimentos. Essas inferências são formalizadas por Michalski de forma bastante inovadora e fugindo, de certa maneira, das formalizações da lógica clássica. A nosso ver, esse distanciamento da lógica clássica é uma grande desvantagem para o método, já que todos os fundamentos existentes e provados para a lógica clássica não poderiam se aproveitados por essa formalização inovadora.

Michalski também chama a atenção para o fato de que os novos conhecimentos gerados a partir da aplicação das inferências citadas não poderão ser completamente validados, já que estamos falando de inferências indutivas que não preservam a verdade e têm um caráter não monotônico. Por essa razão, a pesquisa em aprendizado indutivo normalmente se concentra na geração de hipóteses plausíveis enquanto a adequada validação dessas hipóteses, em última instância, é deixada para os humanos.

No sistema proposto em [Mic83] o conjunto de fatos observados F pode ser visto como uma coleção de implicações do tipo:

$$F : \{e_{ik} ::> K_i\}, i \in I$$

onde e_{ik} denota uma descrição do k-ésimo exemplo da classe K_i (o símbolo ::> significa a implicação ligando uma descrição de conceito a um nome de conceito ou classe), e I é um conjunto de classes indexadoras K_i.

A assertiva indutiva H que deve ser buscada pelo sistema, pode ser caracterizada como um conjunto de regras de reconhecimento de conceitos do tipo:

$$H : \{D_i ::> K_i\}, i \in I$$

onde D_i é uma descrição de conceito da classe K_i, ou seja, uma expressão de condições tal que quando elas são satisfeitas por um objeto, o objeto é considerado uma instância da classe K_i.

A assertiva indutiva, por outro lado, deve satisfazer à seguinte definição:

$$H \mid> F$$

significando que H especializa-se para F, ou seja, através de uma regra de especialização podemos partir de H e chegar em F.

Com essas definições, Michalski chega às seguintes condições:

$$\forall i \in I(E_i \Rightarrow D_i)$$

$$\forall i, j \in I(D_i \Rightarrow \neg E_j), \text{ se } j \neq i$$

onde E_i é uma descrição que é satisfeita por todos os eventos de treinamento da classe K_i, e somente por esses eventos.

A primeira condição caracteriza a completude, ou seja, todo evento de treinamento de alguma classe deve satisfazer à descrição D_i da mesma classe. A segunda condição é dita a condição de consistência e significa que se um evento satisfaz uma descrição de alguma classe, então ele não pode ser um membro de um conjunto de treinamento de outra classe. Essas duas condições são os requerimentos que devem ser satisfeitos para uma assertiva indutiva ser aceita como uma regra de reconhecimento de conceito.

A seguir, Michalski faz algumas observações importantes com relação à linguagem de descrição utilizada. Ele ressalta a importância da compreensibilidade da linguagem e, por outro lado, da capacidade de descrever detalhes, idealmente, de forma ilimitada.

Para o trabalho descrito, o autor propõe a utilização de uma linguagem que ele chama de APC (*Annotated Predicate Calculus*), que seria uma modificação do cálculo de predicados clássico. A principal característica dessa linguagem é que a cada predicado, variável ou função (referidos coletivamente como descritores) é assinalada uma "anotação" que contém informações relevantes orientadas ao problema. A anotação pode conter, por exemplo, uma especificação do domínio e do tipo do descritor; uma especificação dos operadores que podem ser aplicados ao descritor; uma especificação das restrições e das relações entre o descritor em questão e outros descritores; para descritores numéricos, o significado, a variação ou a distribuição de probabilidade dos valores para o referido problema; etc. (veja [Mic83] para descrições mais detalhadas das anotações).

Dessa maneira, cada variável, função ou predicado carrega consigo uma informação adicional que pode ser entendida como parte do conhecimento de *background* necessário ao problema.

Um outro ponto importante do sistema proposto em [Mic83] são as já citadas regras de generalização. Essas regras são descritas como uma transformação de uma descrição em outra mais geral, que tautologicamente implica a descrição inicial.

Uma regra de especialização, que também pode ser usada no processo indutivo, faz a transformação oposta: dada uma descrição, a aplicação de uma regra de especialização gera uma conseqüência lógica da mesma.

As regras de especialização são regras convencionais que preservam a verdade e são utilizadas na lógica dedutiva. Em contraste, as regras de generalização não preservam a verdade, mas sim a falsidade. Isto é, se um evento falsifica uma descrição então ele também falsificará uma descrição mais geral.

A base para a formulação de regras de generalização é a inversão de regras de especialização. O exemplo básico dado por Michalski é o seguinte:

A conhecida lei de simpificação:

$$P\&Q \Rightarrow P$$

pode ser invertida em uma regra de generalização do tipo:

$$P\&Q ::> K \mid< P ::> K$$

onde o símbolo $\mid<$ significa que a expressão à esquerda pode ser generalizada para a expressão à direita.

A seguir apresentaremos algumas das regras de generalização propostas em [Mic83]; em todas elas o termo CTX significa uma expressão arbitrária:

- *dropping condition rule*

$$CTX\&S ::> K \mid< CTX ::> K$$

- *adding alternative rule*

$$CTX_1 \mid< CTX_1 \vee CTX_2 ::> K$$

- *extending reference rule*

$$CTX\&[L = R_1] ::> K \mid< CTX\&[L = R_2] ::> K$$

  onde $R_1 \subseteq R_2 \subseteq Dom(L)$.

- *closing interval rule*

$$\left.\begin{array}{c} CTX\&[L = a] ::> K \\ CTX\&[L = b] ::> K \end{array}\right| < CTX\&[L = a..b] ::> K$$

- *turning constraints into variables rule*

$$\left.\begin{array}{c} F[a] \\ F[b] \\ \vdots \\ F[i] \end{array}\right| < \forall x.F[x]$$

  onde $F[x]$ é uma fórmula dependente da variável x e $a, b, ..., i$ são constantes.

Como pode ser facilmente verificado, em todas as regras acima, todas as expressões à direita do símbolo |< atendem a uma gama maior de objetos que as expressões à esquerda do símbolo |<. No entanto, uma regra de generalização não garante que a descrição obtida seja útil ou plausível.

Não nos deteremos em descrever a metodologia STAR proposta por Michalski que utiliza, dentre outros, os conceitos descritos acima. Nos ateremos à teoria de aprendizado indutivo apresentada, cujos pontos básicos são os que acabamos de ver.

Como conclui Michalski, a teoria apresentada vê o aprendizado indutivo como uma busca heurística através de um espaço de descrições simbólicas, geradas pela aplicação de certas regras de inferência a sentenças observadas. O processo de geração da descrição final - a mais plausível assertiva indutiva - baseia-se nas operações complementares de especialização e generalização da assertiva corrente, a fim de acomodar novos fatos.

O conhecimento de *background* foi mostrado como um componente necessário do aprendizado indutivo, o qual provê restrições e critérios para a seleção da assertiva indutiva preferencial.

Como já foi dito antes, a grande desvantagem da teoria proposta, a nosso ver, é a formulação de regras que não se enquadram na lógica clássica.

Um ponto que devemos chamar a atenção nessa teoria de aprendizado indutivo é a utilização da linguagem APC, onde a anotação, uma informação adicional aos descritores da linguagem, pode ser associada ao conceito de rótulos dos sistemas dedutivos rotulados (descrito no Capítulo 4) que também carregam informações adicionais inerentes às fórmulas que rotulam. Esse conceito é a base de toda a nossa proposta de uma teoria de aprendizado abdutivo.

A indução sob o ponto de vista da probabilidade

Newton da Costa [dC93] considera o termo indução como empregado em conexão com inferências inválidas, mas corretas, ou seja, de alguma forma justificáveis; o problema central da indução, sob essa perspectiva, consiste em se encontrar alguma forma de justificação de todos os tipos de indução correta. Muitas vezes a formulação de critérios de correção acha-se relacionada com a possibilidade de se atribuir pesos às inferências, daí a conexão da lógica indutiva com a probabilidade.

Nos raciocínios indutivos, as premissas não implicam logicamente a conclusão. No entanto, de acordo com diversos especialistas, existe uma relação de probabilidade entre a conjunção das premissas e a conclusão: se as primeiras forem verdadeiras, há uma determinada probabilidade de que a conclusão também o seja. E um argumento indutivo seria correto se tal probabilidade for alta.

Sob certos aspectos, todo o problema da lógica indutiva resumir-se-ia no de construir lógicas probabilísticas. O fundamental não seria mostrar como se deve raciocinar indutivamente, mas tão só associar aos argumentos inválidos probabilidades, que, de certa forma, mediriam o grau de racionalidade que lhes conferimos [dC93].

Alguns autores, por outro lado, pretendem conferir esse grau de racionalidade em função de outras teorias que não a probabilidade. Por exemplo Dubois e Prade em [DLP93] propõem a utilização da chamada lógica possibilística; já W. Spohn em [Spo90] propõe uma teoria não probabilística, baseada no conceito de *plain belief*, para tratar a indução.

Newton da Costa defende a idéia de que, se o termo probabilidade for usado em certo sentido amplo, então toda a lógica indutiva enquadra-se numa lógica probabilística.

O proposto em [dC93] é a chamada teoria pragmática da probabilidade, onde a probabilidade de uma proposição ou raciocínio (implicação) mede o grau de assentimento pragmático que se é levado a dar a essa proposição ou raciocínio. Essa probabilidade é algo subjetiva pois engloba muitos fatores tais como simplicidade, poder explicativo, relação com o conhecimento como um todo, conveniência, plausibilidade intuitiva, etc. Mas apresenta também uma dimensão objetiva.

A probabilidade pragmática não está ligada diretamente à verdade, mas exprime o grau de confiança ou assentimento numa proposição, encarada como candidato para ser aceita provisoriamente e testada. Noutros termos, expressa nosso grau de confiança na conveniência e oportunidade de se admitir uma proposição como hipótese com a finalidade de ser testada e criticada.

São três os tipos de probabilidade pragmática propostos: a não-métrica, a comparativa e a métrica ou probabilística propriamente dita.

A probabilidade pragmática métrica refere-se a enunciados e fica definida por uma função P que associa a certos enunciados de uma linguagem valores pertencentes ao corpo dos números reais. A função P deve satisfazer, dentre outras condições, aos axiomas do cálculo de probabilidades.

A formulação dos princípios formais da probabilidade pragmática tem como ponto de partida uma linguagem L cujo conjunto de sentenças (as quais expressam proposições) é designado por S. Supõe-se S fechado pelos conectivos usuais $(\wedge, \vee, \rightarrow, \neg, \leftrightarrow)$. O símbolo $\vdash$ é definido também da maneira usual ($\vdash \alpha$ significa que α é teorema ou tese de L). Além dos axiomas lógicos comuns, L possui axiomas não lógicos que lhe são característicos. Em certo sentido, a probabilidade pragmática acha-se entre a verdade e a falsidade.

O conceito comparativo de probabilidade pragmática do prisma formal, introduz-se como se segue: entre os elementos de S, ou seja, entre as sentenças de L, suporemos definida uma relação binária, $\leq$, que se lê "menos provável que ou igualmente provável a". Esta relação, pela sua própria natureza, pertence à meta linguagem de L.

Na proposta são definidos postulados para a relação $\leq$. Além disso são colocados os axiomas da teoria (todos eles intuitivos) e são propostos diversos teoremas que retratam definições sobre P e a relação $\leq$.

Newton da Costa conlui que a lógica indutiva está estreitamente ligada à probabilidade pragmática, que é uma probabilidade entre amplos limites objetiva, ou pelo menos, cuja objetividade se iguala à das teorias e leis científicas.

Sob o nosso ponto de vista, a probabilidade é apenas um dos aspectos sob os quais pode ser vista a indução (assim como a abdução, que também tem seu caráter probabilístico). É inegável que o raciocínio indutivo tem um caráter incerto e que a teoria da probabilidade parece enquadrar-se perfeitamente no tratamento de incertezas. No entanto, a área de aplicação do raciocínio indutivo é tão ampla que, certamente, existem domínios onde teorias, não necessariamente probabilísticas, se aplicam muito bem.

Um ponto interessante da proposta de Newton da Costa que gostaríamos de chamar a atenção é a relevância dada a certo caráter meta-lógico (onde entram as considerações de simplicidade, poder explicativo, conveniência, etc.) da indução, relevância esta, completamente condizente com a nossa própria teoria proposta para o raciocínio abdutivo.

Capítulo 3

Abdução

Neste capítulo abordaremos a abdução sob dois pontos de vista: primeiro a filosofia da abdução, suas origens, sua participação no raciocínio do senso comum, seu relacionamento com outros tipos de raciocínio; depois trataremos a abdução de maneira mais formal, tentaremos dar uma caracterização genérica da abdução que seja condizente com as abordagens estudadas e, em seguida, discutiremos diversas propostas para teorias abdutivas.

3.1 A filosofia da abdução

A abdução é tida hoje como um tipo de inferência que segue do efeito para a causa, um raciocínio tipicamente retrospectivo cuja finalidade é fornecer explicações para fatos observados. A origem desse tipo de raciocínio é muito antiga, possivelmente tão antiga quanto a humanidade, mas a abdução como inferência lógica e o próprio termo "abdução", são atribuídos a Charles Sanders Peirce, um filósofo americano cuja obra, datada do final do século passado, somente por volta dos anos 30 é que começou a ser mais amplamente divulgada e reconhecida como importante.

Nesta seção, trataremos da abdução de forma mais ou menos filosófica, tentando relacionar as idéias de Peirce (comentadas por diversos autores) com as definições de abdução que utilizaremos para o nosso trabalho. Discutiremos também algumas analogias interessantes feitas entre esse tipo de raciocínio e atividades 'mundanas' como o raciocíno detetivesco.

3.1.1 Peirce e as origens da abdução

Peirce não conseguiu, em vida, consagração profissional nem reconhecimento dos seus contemporâneos. No entanto, não se deixou desanimar nem descuidou-se de sua obra. De sua incessante atividade de redação, que se estendeu de 1867 até poucos anos antes do seu falecimento em 1914, resultaram os *Collected Papers* dos quais os primeiros seis volumes foram publicados entre 1931 e 1935 e os dois últimos em 1958 [dMH93].

Sem nome ou fama em sua época, embora de valor reconhecido por homens da elite intelectual, Peirce parece ter trabalhado como se estivesse convicto de que, mais cedo ou mais tarde, a posteridade haveria de conceder o que seu tempo lhe negava. Em um manuscrito seu intitulado "*A propósito do autor*" [Pei93b], Peirce, após declarar haver penetrado até a saturação no estudo das ciências físicas e após referir os estudos filosóficos a que se dedicou, descreve sua obra como "tentativa que um físico desenvolve no sentido de fazer conjectura acerca da constituição do universo, utilizando métodos científicos e reconhecendo a ajuda de tudo quanto foi feito por filósofos anteriores". Sua meta era a de despertar no leitor o desejo da investigação [dMH93]. Uma de suas principais ocupações era o estudo dos métodos de investigação científica. Diretamente ligadas com esse estudo, estão as definições dos três tipos de inferência considerados por Peirce: dedução, indução e abdução (este último alternativamente batizado de retrodução ou inferência hipotética). Peirce acreditava que toda investigação, científica ou não, inevitavelmente faria uso desses três tipos de inferência, cada uma no seu devido momento. No entanto, os limites entre cada uma das inferências, principalmente entre abdução e indução, nunca ficaram muito bem definidos.

Segundo Thomas A. Sebeok em [Seb91], Peirce considerava a dedução como dependente de nossa confiança em nossa habilidade de analisar o significado dos signos nos ou pelos quais pensamos; a indução seria dependente de nossa confiança em que o curso de algum tipo de experiência não será interrompido sem qualquer indicação que anteceda a interrupção; e a abdução como dependente de nossa esperança de, cedo ou tarde, supor as condições sob as quais um dado tipo de fenômeno se apresentará.

Para tornar mais concreta a distinção entre os três tipos de raciocínio, tomemos como ilustração o famoso exemplo do saco de feijão, empregado por Peirce em 1878 (citado em [Seb91]):

	Dedução
Regra	Todos os feijões deste saco são brancos
Caso	Estes feijões provêm deste saco
Resultado	Estes feijões são brancos

	Indução
Caso	Estes feijões provêm deste saco
Resultado	Estes feijões são brancos
Regra	Todos os feijões deste saco são brancos

	Abdução
Regra	Todos os feijões deste saco são brancos
Resultado	Estes feijões são brancos
Caso	Estes feijões provêm deste saco

Muitas pesquisas contemporâneas têm identificado abdução com os procedimentos conjecturais de médicos e historiadores (como em [Gin91]). Atualmente um médico busca tanto leis gerais quanto causas específicas, assim como um historiador. Em todos os casos (descobertas científicas, investigações médicas e criminais, reconstituições históricas, etc.) temos casos de pensamento conjectural acerca de uma série de elementos aparentemente desconexos. Por esse motivo, pode-se dizer que a análise dos procedimentos conjecturais na investigação criminal pode lançar nova luz sobre os procedimentos conjecturais na ciência de maneira geral. Por isso o tratamento da relação Peirce-Conan Doyle, que discutiremos mais abaixo, contribui para um esforço epistemológico mais amplo.

Segundo Sebeok, uma abdução nos permite formular um prognóstico geral, mas sem garantia de um resultado bem sucedido; ao mesmo tempo, a abdução enquanto um modo de prognosticação, oferece a única esperança possível de regular racionalmente nossa conduta futura. Em [SS91] é colocado que Peirce encarava a abdução simplesmente como uma conjectura, e que ele (Peirce) argumentava que, para contribuir para o incremento do conhecimento devemos assumir que a mente do homem tem um ajustamento natural para imaginar teorias corretas de certos tipos, certo princípio de abdução o qual coloca um limite para a hipótese admissível, uma espécie de instinto, desenvolvido no curso da evolução. As idéias de Peirce sobre a abdução são bastante vagas e sua sugestão de que uma estrutura biologicamente dada desempenha um papel básico na seleção de hipóteses científicas parece ter exercido muito pouca influência. Segundo Chomsky (citado em [SS91]), Peirce exerceu uma enorme influência, mas não por esse motivo em particular. A idéia de que a escolha de hipóteses corretas (o que segundo Peirce, acontecia com surpreendente freqüência) deve-se a algum instinto, semelhante ao instinto de voar dos pássaros, não foi um pensamento muito aceito, e supomos que existem critérios, não somente intuitivos para a seleção dessas hipóteses.

Acreditamos que assim como nos tempos remotos se acreditava que a inferência indutiva não existia, e que, desse modo jamais poderia ser automatizada, impossibilitando assim o que hoje chamamos de aprendizado de máquina, também o mesmo tipo de pensamento tinha Peirce sobre a abdução, acreditando que só uma ligação direta entre o homem e a natureza tornava possível as inferências abdutivas corretas. Hoje, no entanto, acreditamos que a possibilidade de se fazer inferências abdutivas corretas está diretamente ligada ao grau de conhecimento que o agente possui. E, como a representação de conhecimento há muito deixou de ser uma impossibilidade para os sistemas automatizados, embora ainda aparentemente uma difícil tarefa, a abdução também pode ser pensada em termos de um raciocínio automatizado, desde que o agente, agora artificial, tenha em seu "poder" conhecimento suficiente para inferir abdutivamente correto.

Como acréscimo ao princípio de que a mente humana, como resultado dos processos naturais de evolução, é predisposta a fazer suposições corretas do mundo, Peirce propõe um segundo princípio conjectural para explicar parcialmente o fenômeno da suposição, isto é, que nós freqüentemente retiramos da observação fortes sugestões de verdade, sem sermos capazes de especificar quais foram as circunstâncias por nós observadas que conduziram a essas sugestões, "os diferentes elementos de uma hipótese estão em nossa mente antes que conscientemente cogitemos nela, mas é a idéia de reunir aquilo que nunca antes sonhamos que faz brilhar a sugestão nova diante de nossa contemplação" (Peirce, citado em [SS91].

Peirce parece considerar o processo abdutivo como incontrolável e inconsciente. Talvez ele tenha razão no que se refere ao processo abdutivo executado pela mente humana, já que ela é dotada de procedimentos de analogias e buscas ainda desconhecidos para nós. No entanto, a formalização e automatização da abdução tornam-se possível na medida em que computacionalmente é possível gerar um grande número de hipóteses plausíveis e minimais (como veremos nas seções seguintes), já que, de posse do conhecimento apropriadamente representado, um sistema computacional pode fazer as analogias que lhe sejam 'solicitadas', desde que devidamente especificadas. Assim, a sugestão de hipóteses abdutivas torna-se um procedimento factivelmente mecanizável, o grande porém acontece no momento de se escolher a melhor hipótese abdutiva, o que, segundo nosso ponto de vista, depende de critérios meta-lógicos que deverão ser convenientemente tratados em qualquer que seja o método de formalização.

Segundo Bonfanttini e Proni, em [BP91], a abdução, antes de mais nada, é uma inferência. Isso quer dizer que o último estágio de um argumento abdutivo consiste em extrair uma conclusão de suas premissas. Nesse sentido a abdução é tão formal e mecânica quanto a dedução e a indução: o modo pelo qual se tira a conclusão é rigidamente governado por uma regra. Tão pouco há motivos para se pensar que uma das inferências seja psicologicamente mais fácil ou mais difícil que a outra. Em outras palavras, para usar a terminologia de Peirce, é igualmente mecânico e automático derivar a regra partindo do caso e do resultado (indução), quanto derivar o resultado partindo da regra e do caso (dedução), ou o caso partindo da regra e do resultado (abdução).

Pensando desse modo, isto é, que a abdução tem seu lado mecânico, é perfeitamente plausível a tentativa de formalizar a abdução como um primeiro passo para a criação de um método computacional para executar o raciocínio abdutivo.

Ainda segundo Bonfantini e Proni, "a abdução é sintética e inovativa e, como tal, também contém um elo de risco, uma vez que o valor de verdade da conclusão abdutiva não é naturalmente determinado pela validade das premissas (...) A abdução consiste na atribuição ao sujeito da investigação, identificado na premissa expressando o resultado, as características expressas no antecedente da premissa principal ou regra. Portanto, é completamente compreensível que tanto o elemento de risco - acrescido àquele que pode estar contido nas premissas - quanto o grau de novidade da conclusão abdutiva dependem do relacionamento entre as duas proposições (antecedente e conseqüente) que constituem a premissa principal" [BP91].

Vale ressaltar que o procedimento descrito acima, é exatamente o que propomos para tratar a abdução: atribuir ao fato observado (o resultado no caso de Peirce) as "características" expressas no antecedente da regra que escolhermos (já que a nossa regra terá necessariamente o formato

$A \rightarrow B$ onde A será uma fórmula qualquer e B será o nosso próprio fato observado). É importante chamar a atenção também para o fato de que, como estaremos lidando com um conhecimento de *background* que utilizaremos para explicar o fato observado, a nossa conclusão poderá não ser obtida tão diretamente como mostra o esquema proposto por Peirce. Ela retratará sim, uma relação de conseqüência onde dados da teoria estarão envolvidos no processo. Como veremos nas próximas seções, não estaremos buscando uma fórmula A que sozinha, independente de qualquer contexto, implique na fórmula B (ou seja, $A \rightarrow B$), mas sim uma fórmula α tal que considerando nosso conhecimento de *background*, Θ, $\Theta \cup \{\alpha\} \vdash B$, ou seja, será possível derivar B a partir da união de Θ com α.

Por outro lado, Bonfanttini e Proni sugerem que a busca de hipóteses abdutivas mais confiáveis está diretamente relacionada com a proximidade entre o antecedente e o conseqüente; isso equivaleria, na nossa proposta, à busca de hipóteses com as quais a derivação do fato observado se daria com o menor número de passos de dedução possíveis: o que equivale, especificamente, à intuição do critério de número de elementos utilizado para a comparação dos candidatos (ver seção 5.4.3).

Bonfantini e Proni colocam também que "mais marcante é a novidade da abdução quando a premissa principal vincula o resultado com uma possível causa remota e improvável. E a novidade da abdução é ainda mais nítida e mais forte quando o princípio expresso na premissa principal é uma nova lei teórica mais do que uma lei científica universalmente aceita. Neste caso a conclusão abdutiva é uma idéia nova em termos absolutos: não é apenas a aplicação do princípio geral ao sujeito da investigação que é nova, também o princípio é novo. Por isso é que a conclusão ainda não estava nem mesmo pontecialmente incluída no estoque existente de conhecimento" [BP91].

Com relação ao comentário anterior, questionamos se uma dita abdução desse tipo, onde a regra geral da premissa principal não estava incluída no conhecimento existente, não estaria muito próxima do conceito de indução, que é o tipo de inferência que realmente busca novas regras. O duvidoso é que a indução parte de uma série de fatos similares para buscar a regra, e nesse caso, a regra é, de certa forma, inventada, independente de fatos que levem a ela. Tudo que se tem é uma observação que se deseja explicar, e dentro do conhecimento disponível nenhuma regra parece plausível de ser essa explicação. Então a nova regra é proposta (como um passo intermediário que não caracteriza necessariamente nem a abdução nem a indução).

Em [Pei93a] muito é dito sobre a distinção entre indução e abdução. A indução é dita ocorrer quando generalizamos a partir de certo número de casos em que algo é verdadeiro e inferimos que a mesma coisa será verdadeira para o total da classe. Ou quando verificamos que certa coisa é verdadeira para certa proporção de casos e inferimos que é verdadeira, na mesma proporção, para o total da classe. Hipótese ocorre quando deparamos com uma circunstância curiosa, capaz de ser explicada pela suposição de que se trata de caso particular de certa regra geral, adotando-se, em função disso, a suposição. Ou quando verificamos que sob certos aspectos dois objetos mostram forte semelhança e inferimos que se assemelham fortemente um ao outro sob aspectos diversos.

Um interessante exemplo dado por Peirce [Pei93a] na tentativa de diferenciar indução e abdução, é o exemplo do papel rasgado, que descreveremos a seguir. Num pedaço de papel, está um trecho de um autor anônimo. Suspeita-se que o autor seja determinada pessoa cuja mesa de trabalho, à qual só ele tem acesso, é revistada e ali se encontra um outro pedaço de papel cujo contorno se ajusta exatamente, em todas as suas irregularidades, ao papel anteriormente mencionado. Dizer que o homem apontado como autor foi efetivamente o autor corresponde, segundo Peirce, a uma inferência hipotética. A base dessa inferência, reside, evidentemente no fato de que dois pedaços de papel rasgado muito dificilmente se ajustarão entre si por puro acaso. Se a hipótese fosse tão somente uma indução, tudo que estaríamos justificados a concluir, segundo Peirce, seria que os dois pedaços de papel, ajustados segundo as irregularidades observadas, também se ajustariam com respeito a outras irregularidades. Segundo Peirce, a inferência que relaciona a forma do papel com o seu proprietário é precisamente o que distingue a hipótese da indução, e faz com que esta surja como um passo mais ousado e perigoso.

Analisando cuidadosamente o exemplo do papel rasgado de Peirce, não nos parece que ele caracterize a diferença entre a indução e a abdução. A inferência que relaciona a forma do papel com o seu proprietário na verdade toma como base outros fatos não mencionados aqui, e não somente o encaixe dos dois pedaços de papel. Se tentarmos identificar cada passo do raciocínio que leva a essa inferência chegaremos, certamente, a um outro conjunto de premissas (devidamente habituais para que sejam consideradas conhecidas) e essas sim justificarão a hipótese escolhida tão bem quanto foi justificado o encaixe dos papéis. Uma possível premissa, omitida no exemplo, seria o fato de que todos os papéis encontrados na mesa do suposto autor, foram escritos pelo autor.

3.1.2 Sobre a abdução de Sherlock Holmes

Não existe, tanto quanto saibamos, nenhuma evidência direta de que Peirce tenha lido qualquer das histórias de Sherlock Holmes. No entanto, a sua forma de raciocínio, quando se trata de abdução, é extremamente semelhante ao raciocínio detetivesco relatado por Arthur Conan Doyle nas histórias de Sherlock Holmes.

Um interessante estudo sobre o raciocínio de Sherlock é relatado em uma coletânea organizada por Umberto Eco e Thomas A. Sebeok e intitulada *O signo de três*, onde alguns dos artigos [Seb91, SS91, Tru91, Gin91, BP91, Car91, HH91, Har91, Eco91] dissertam sobre o método desse famoso detetive fictício, na maioria das vezes relacionando-o com o método abdutivo de Peirce.

Um dos pontos mais ressaltados pelos autores dos artigos da coletânea citada acima, é a preocupação de Sherlock com os detalhes e com a comprovação das suposições que faz. Além disso, todo o processo de investigação de Sherlock tem como base um grande conhecimento criminal que ele utiliza para fazer analogias com os casos que tenta resolver.

Muitos aspectos do método de Sherlock estão curiosamente condizentes com a filosofia de Peirce com relação à abdução. Por exemplo, a preocupação em buscar hipóteses simples e naturais e fáceis de serem checadas. Esses critérios de seleção que buscam a hipótese mais simples e mais natural, de certa forma, subjetivos, coincidem de algum modo com o critério de minimalidade proposto em diversas formalizações da abdução (como veremos na seção 3.3). Na nossa própria proposta trabalharemos não só com o critério de minimalidade, como com outros critérios que consideramos do meta-nível, mas que, a grosso modo, refletem essa idéia de preferência pelas hipóteses mais adequadas.

O método de Sherlock também apresenta a característica de utilizar os três tipos de inferência conjuntamente: ele faz induções quando soluciona casos por analogia, ele faz abduções quando propõe hipóteses, e faz deduções quando testa as hipóteses propostas. Essa característica também está condizente com as idéias de Peirce que proclamava o uso conjunto dos três tipos de inferência. Além disso, Peirce afirmava que uma hipótese abdutiva sempre deveria ser testada para que fosse considerada plausível.

Para Sherlock o importante era saber raciocinar retrospectivamente. Raciocinar a partir de um conjunto de eventos em direção a suas conseqüências. É claro que no caso de Sherlock o fator fictício entra como um participante altamente ativo no momento de escolher uma dentre várias hipóteses. Aliás, nada é relatado, além do teste das hipóteses, com relação a critérios para selecionar hipóteses mais plausíveis.

Em [Eco91], Humberto Eco coloca que Sherlock inventa quando, por exemplo advinha o fluxo de pensamento de Watson, "Certamente ele foi obrigado a escolher, dentre os muitos percursos mentais possíveis de Watson (os quais ele provavelmente mentalizou todos ao mesmo tempo), aquele que demonstrava maior coerência estética ou maior elegância. Como o fez Copérnico, adivinhando que o sol deveria estar no centro do universo porque, só dessa maneira, poderia manifestar-se a admirável simetria do mundo criado (...) Os critérios de escolha de hipóteses de Sherlock são apenas critérios subjetivos de elegância e coerência. Ele, no entanto, nunca erra, já que tem o privilégio de viver em um mundo construído por Conan Doyle onde tudo se encaixa perfeitamente, o que faz com que Sherlock não necessite de provas imediatas para verificar suas hipóteses".

Peirce, no entanto, vincula estritamente a fase da abdução com a dedução:

> "A retrodução não proporciona segurança. A hipótese deve ser testada. Esse teste, para ser logicamente válido, deve começar de forma honesta, não como começa a retrodução, com o escrutínio do fenômeno, mas com o exame da hipótese e uma revisão de todos os tipos de conseqüência condicionais experienciais que se seguem de sua veracidade. Isso constitui o segundo estágio da investigação."(C. S. Peirce, citado em [Car91])

3.1.3 Conclusões sobre a filosofia da abdução

A abdução (tenha sido ela chamada de hipótese, retrodução, raciocínio retrospectivo, ou até mesmo de dedução) é, consensualmente, um tipo de raciocínio bastante difundido, e desde os tempos remotos vinha despertando alguns questionamentos.

Alguns pontos chaves sobre a narrativa das seções precedentes devem ser salientados; justamente aqueles que serão importantes no momento da formalização da abdução.

A primeira questão é a necessidade de conhecimento por parte do 'agente' que executa a abdução. Pode-se perceber que essa necessidade é bastante ressaltada, inclusive nos comentários sobre Holmes, que chamava a atenção para o "cuidado com os detalhes". Esse conhecimento é o que chamamos de teoria de *background*. Essa teoria deverá conter o conhecimento relevante ao problema em questão. O problema em si é sempre considerado como a tentativa de averiguar a que caso pertence determinado resultado (fato) observado. Do conhecimento de *background* será obtida a regra que torna essa relação possível.

A teoria de *background*, por seu lado, deve ser a mais completa e correta possível. E qualquer modificação que aconteça pode influenciar na hipótese escolhida.

Uma outra questão que chama bastante a atenção é a utilização conjunta das três inferências lógicas: dedução, indução e abdução. A dedução é inconfundível, e é colocada como necessária para o processo de descoberta tanto no momento de sugerir hipóteses quanto na questão do teste das hipóteses escolhidas: para confirmar se a hipótese escolhida é realmente plausível, deve-se testá-la dedutivamente. Já a relação entre indução e abdução não nos parece tão claramente definida. O processo de descoberta é descrito como uma mistura alternada de abdução e indução, sem que fiquem muito claros os limites entre ambas. Sob o nosso ponto de vista, como já foi discutido na seção 2.4, a indução é um caso particular da abdução, e o que a caracteriza é uma busca de uma explicação generalizadora a partir da observação de fatos semelhantes.

O último ponto a salientar, talvez o mais importante no que diz respeito à nossa proposta em particular, é quanto à escolha da hipótese mais plausível. Não é dito muito sobre esse processo de escolha: Peirce o justifica através de uma 'ligação com o natural'; Holmes, graças à sua perspicácia, faz suposições extremamente bem. Ora, parece haver algo de nebuloso nessas colocações. Em verdade, tudo que é dito é sobre a importância de uma boa escolha de hipóteses. Mas como fazê-lo, é uma incógnita.

Tomando o ponto de vista de Holmes, o interessante das narrativas sobre ele é a forma como o mesmo expõe (e executa) seu processo de solucionar crimes: o raciocínio retrospectivo do efeito para a causa, a forma como ele tira as conclusões e em seguida as testa. No entanto, as conclusões de Holmes são um tanto irrealistas (é óbvio). O fictício atua fortemente para a escolha das hipóteses corretas, e, embora algumas vezes ele proponha mais de uma hipótese e opte por uma delas, os fatores que influenciam nessa opção parecem muito determinísticos, como se não houvesse chance de errar. E ele, não tão incrivelmente, acerta sempre. É claro que nossa realidade é bem diferente desse determinismo.

Já o ponto de vista de Peirce nos sugere muito mais uma justificativa sobrenatural para o desconhecido, que qualquer teoria bem fundamentada.

A nosso ver, a escolha de hipóteses plausíveis está baseada no conhecimento que se tem sobre o problema em questão. E todo o processo de descoberta é em função desse conhecimento e da modificação desse conhecimento. É claro que o processo não é determinístico como o sugere as hitórias de Sherlock Holmes. A abdução é, sem dúvida, um processo falível; não existem garantias de que a hipótese escolhida seja a hipótese correta. No entanto, pode-se buscar a hipótese mais plausível em função do fato observado e de todo o conhecimento que se tem em mãos; e quanto mais correto e completo é o nosso conhecimento, maior a possibilidade de supormos todas as hipóteses possíveis e, dentre elas, aquela que é a correta.

Concluindo, e, de certa forma, resumindo o que temos em mente: o processo de sugerir hipóteses abdutivamente é razoavelmente determinístico, ainda que complexo e dependente do quão completo e correto é o nosso conhecimento. Já a escolha de hipóteses mais plausíveis é uma questão dependente não só do conhecimento que temos, como também dos aspectos que desejamos valorizar nessa escolha. Esses aspectos compõem justamente o que chamamos de considerações do meta-nível da abdução, e caracterizam o principal aspecto do método proposto nesta dissertação.

3.2 A abdução formalizada

Nesta seção a abdução será mostrada sob um aspecto mais formal. Definiremos aqui os conceitos básicos sobre abdução que são comuns a praticamente todos os trabalhos conhecidos na área. Essas definições servirão como conhecimento preliminar para o entendimento das seções seguintes onde descreveremos os trabalhos correlatos.

É consenso que a abdução é um raciocínio do senso comum que busca explicações para um fato observado e inquestionável. Essas explicações são sugeridas em função de um conhecimento preliminar sobre o problema em questão (teoria de *background*).

O resultado de um processo abdutivo é a sugestão de hipóteses para explicar o fato observado. Formalmente podemos definir o conceito de hipótese abdutiva como segue:

Definição 3.1 (Hipótese abdutiva) *Dado um fato φ observado e inquestionável, e dada uma teoria de background, Θ, α será o resultado de um processo abdutivo (ou será uma hipótese abdutiva) se:*

$$\Theta \cup \{\alpha\} \vdash \varphi$$

Em outras palavras, o processo abdutivo busca encontrar uma explicação α para um fato φ, dado o conhecimento de *background* Θ.

É consenso também que a abdução é um raciocínio não monotônico. Isto é, sempre que acrescentarmos novos fatos a Θ, o que antes parecia uma explicação para φ, pode deixar de ser. Isso ocorre em função de que o novo conhecimento pode ser incoerente (inconsistente) com a explicação α ou porque o novo conhecimento pode acrescentar uma explicação α' mais plausível que α.

Quanto à plausibilidade das hipótese abdutivas, a maioria das abordagens em abdução trata apenas dois critérios para classificar as hipóteses como mais plausíveis ou menos plausíveis: (1) o critério de consistência com a teoria, ou seja, $\Theta \cup \{\alpha\}$ não pode gerar uma inconsistência; e (2) o critério da minimalidade a nível objeto, que está relacionado com as questões da implicação lógica: uma hipótese α é mais plausível que outra α' se $\alpha' \rightarrow \alpha$. O critério de minimalidade reflete, de certo modo, a busca de hipóteses mais simples.

Além desses dois critérios básicos, algumas abordagens levam em consideração as conseqüências dedutivas das hipóteses. Nesse caso deve ser verificado quais conseqüências dedutivas $\Theta \cup \{\alpha\}$ gera além de φ; essas conseqüências funcionam como condições adicionais para que a hipótese α seja considerada plausível.

Algumas abordagens colocam a questão da plausibilidade das hipóteses como uma consideração meta-lógica e não tratável no nível da solução proposta. Uma das abordagens que discutiremos na seção 3.4.1 trata essas considerações meta-lógicas como heurísticas que direcionam a escolha das melhores hipóteses [MD94].

Nossa proposta, que será discutida no Capítulo 5, trata essas considerações meta-lógicas, ou do meta-nível com critérios de comparação de hipóteses abdutivas que se baseiam numa ordem preferencial fornecida pelo usuário.

Outro ponto que é considerado na maioria das abordagens em abdução é a adição de restrições de integridade com a finalidade de reduzir o conjunto de hipóteses possíveis. Quase sempre essas restrições são tratadas como uma informação diferenciada do conhecimento de *background* e sua consideração só acontece após a geração do conjunto de hipóteses. Assim, as restrições de integridade atuam para eliminar hipóteses do conjunto inicialmente gerado.

3.3 Trabalhos correlatos

Nesta seção, descreveremos com algum detalhe três trabalhos em abdução que tiveram bastante influência na definição de nossa própria proposta para caracterizar o raciocínio abdutivo. Muitas das definições que veremos aqui, serão utilizadas, de forma adaptada, no Capítulo 5, onde descreveremos nosso método proposto.

3.3.1 Utilizando *predicate completion* para relacionar dedução e abdução

Muitos autores abordam a abdução sob o enfoque da programação em lógica. É o caso do artigo [CDT91] onde é proposta a utilização de *predicate completion* (um conceito da área de programação em lógica) para a caracterização de explicações abdutivas.

No referido artigo é considerada uma definição de meta-nível da abdução em termos de dedução, similar a várias definições propostas na literatura, e uma definição do nível objeto na qual conclusões abdutivas são expressas como conseqüência lógica de observações e de uma simples transformação da teoria de domínio baseada em *predicate completion*. A equivalência entre as duas definições é provada para teorias de domínios consideravelmente expressivos.

Considerando-se a abdução como a geração de explicações para um conjunto de eventos a partir de uma dada teoria, a seguinte caracterização é dita pelos autores uma caracterização do meta-nível:

Dada uma teoria de domínio Θ e uma fórmula φ, uma explicação para φ em Θ é um conjunto de fórmulas E tal que:

(1) $\Theta \cup E$ é consistente;

(2) φ é uma conseqüência de $\Theta \cup E$;

(3) E tem algumas propriedades a mais que a tornam interessante.

O objetivo do artigo de Console et. al. é analisar sobre vários pontos de vista as relações entre abdução e dedução. Nele é proposta uma caracterização alternativa mostrando como as conclusões de um raciocínio abdutivo podem ser obtidas dedutivamente no nível objeto, isto é, no mesmo nível da teoria Θ. Tal caracterização permite explicitar algumas suposições subjacentes a outras definições de abdução. Além disso, já que é baseada em *predicate completion*, ela mostra o relacionamento entre a abdução e os fundamentos da programação em lógica.

O seguinte problema é considerado: determinar uma explicação para um átomo q, dada uma teoria T que contém as seguintes implicações tendo q como seu conseqüente:

$$\begin{gathered} p_1 \rightarrow q \\ p_2 \rightarrow q \\ \vdots \\ p_3 \rightarrow q \end{gathered}$$

onde cada p_i é uma fórmula.

Tais implicações expressam relações de causa e efeito ou, em geral, o fato de que todo p_i pode ser aceito como uma explicação direta para q.

Considerando-se as definições do meta-nível da abdução, o processo para determinar as explicações para q é baseado nas seguintes suposições implícitas:

- se q está presente, então pelo menos uma de suas explicações diretas deve estar presente;
- em particular, a fim de que T possa explicar q, pelo menos um entre os $p_1, p_2, ..., p_n$ deve ser assumido já que são as únicas explicações diretas para q em T.

Desde que as suposições acima são subjacentes à definição de meta-nível da abdução, um problema interessante é buscar uma transformação sintática de T que torne estas suposições explícitas. Com tal transformação as explicações abdutivas podem ser determinadas no nível objeto, isto é, no nível da teoria (transformada).

O seguinte exemplo concreto de um problema simples de interpretação representado pela teoria T_1, é considerado em [CDT91]:

$$\begin{aligned} T_1 = \{ \quad & \textit{rained-last-night} \rightarrow \textit{grass-is-wet}, \\ & \textit{sprinkler-was-on} \rightarrow \textit{grass-is-wet} \\ & \textit{grass-is-wet} \rightarrow \textit{grass-is-cold-and-shiny}, \\ & \textit{grass-is-wet} \rightarrow \textit{shoes-are-wet}\} \end{aligned}$$

Os átomos *rained-last-night* e *sprinkler-was-on* que não aparecem no conseqüente de nenhuma implicação, são considerados como possíveis hipóteses (átomos abduzíveis) que podem ser aceitas como explicações de fatos observados. Por exemplo temos duas explicações minimais para $\varphi = \textit{grass-is-cold-and-shiny}$:

$$E_1 = \{\textit{rained-last-night}\}$$

$$E_2 = \{\textit{sprinkler-was-on}\}$$

Esses são conjuntos minimais de hipóteses que, em conjunção com T_1 implicam a observação φ. Note-se que o conceito de ser interessante aqui significa primeiramente 'conter somente átomos abduzíveis' e depois ser 'minimal'.

A técnica de meta-nível que produz E_1 e E_2 como explicações é baseada na suposição implícita descrita acima: *grass-is-wet* é a única explicação direta de *grass-is-cold-and-shiny* então ela deve ser assumida a fim de explicar *grass-is-wet* (e consequentemente φ).

A fim de tornar essa suposição explícita, considera-se a completação T_{1c} dos predicados não abduzíveis em T_1, aumentada com a fórmula φ, descrevendo as observações:

$$\begin{aligned} T_{1c} = \{ \ & \textit{grass-is-wet} \leftrightarrow \textit{rained-last-night} \vee \textit{sprinkler-was-on}, \\ & \textit{grass-is-cold-and-shiny} \leftrightarrow \textit{grass-is-wet}, \\ & \textit{shoes-are-wet} \leftrightarrow \textit{grass-is-wet}\} \end{aligned}$$

$$\varphi = \textit{grass-is-cold-and-shiny}$$

As duas explicações E_1 e E_2 para φ podem ser representadas no nível objeto através da fórmula:

$$F_1 = \{\textit{rained-last-night} \vee \textit{sprinkler-was-on}\}$$

e é verificado que $T_{1c}, \varphi \models F_1$, isto é, que a fórmula caracterizando a explicação segue logicamente da teoria completada e das observações.

Segundo Console et. al., completar predicados não abduzíveis corresponde a tornar explícito o fato de que pelo menos uma das explicações diretas de um átomo não abduzível q deve estar envolvido na explicação de q. Em outras palavras, na teoria completada as explicações que podem ser obtidas de T_1 são representadas no nível objeto, então a completação permite a redução de abdução a dedução. Os autores também chamam a atenção para o fato de que, pela abdução ser baseada numa completude semântica, não significa que a teoria de domínio esteja completa (no sentido de que toda explicação direta de cada átomo não abduzível deve ser dada pela teoria). Isso significa que a explicação obtida através da abdução pode ser errada se nem todo conhecimento relevante está presente na teoria.

É importante que conhecimento geral (a teoria) e conhecimento factual (os dados observados em um caso específico) sejam mantidos separados. Já que *predicate completion* é usado para significar que um fato é verdade, somente conhecimento genérico deve ser completado. Adicionando *grass-is-cold-and-shiny* (isto é, conhecimento factual) para a teoria a ser completada significará que *grass-is-cold-and-shiny* é incondicionalmente verdade. Essa não é a intenção que desejamos. Num caso desses, de fato, a completação de *grass-is-cold-and-shiny* retornaria:

$$\begin{aligned} & \textit{grass-is-cold-and-shiny} \Leftrightarrow \textit{grass-is-wet} \vee \textit{true} \\ & \equiv \textit{grass-is-cold-and-shiny} \end{aligned}$$

de onde nada pode ser concluído sobre *grass-is-wet*.

Considerando a lógica proposicional, a teoria T deverá conter cláusulas do tipo:

$$L_1 \wedge \ldots \wedge L_n \rightarrow p$$

onde cada L_i é um literal (isto é, um átomo ou a negação do átomo) e p é um átomo. $L_1 \wedge ... \wedge L_n$ é chamado o corpo da cláusula e p é chamado a cabeça. Além disso, é requerido que a teoria seja hierárquica. De acordo com [CDT91], uma teoria é dita hierárquica se é possível assinalar um nível (um número natural) aos predicados da teoria (aos átomos no caso proposicional) tal que para toda cláusula o nível dos predicados ocorrendo no corpo é estritamente menor que o nível do predicado ocorrendo na cabeça; alternativamente, uma teoria é hierárquica se o grafo de dependência (isto é, o diagrama contendo os predicados da teoria como nós e um arco de p para q se e somente se p ocorrer no corpo de uma cláusula em cuja cabeça q ocorre) é acíclico. Tal restrição é considerada razoável para uma teoria de domínio na qual *corpo* $\rightarrow$ *cabeça* é entendida como *corpo diretamente explica cabeça*; de fato, a relação *A está envolvida em uma explicação para B* pode resultar naturalmente em uma ordem parcial, já que é inútil explicar um fato em termos dele mesmo. Para a abordagem em questão, também deve ser assumido que:

(a) o conjunto de símbolos de predicados em T é particionado em dois subconjuntos disjuntos, os símbolos abduzíveis (aqueles que podem ser aceitos como explicações de fatos observados) e os símbolos não abduzíveis;

(b) os símbolos abduzíveis são exatamente aqueles que não ocorrem na cabeça de nenhuma cláusula da teoria.

Em seguida, a seguinte definição para problemas abdutivos é proposta em [CDT91]:

Definição 3.2 *Um problema abdutivo é um par* $\langle T, \varphi \rangle$ *onde:*

- T *(a teoria de domínio) é um programa lógico proposicional e hierárquico cujos átomos abduzíveis são aqueles que não ocorrem na cabeça de nenhuma cláusula;*

- φ *(as observações a serem explicadas) é uma conjunção consistente de literais com nenhuma ocorrência de átomos abduzíveis.*

Uma importante observação é que φ pode conter literais positivos ou negativos e o requerimento de φ ser consistente corresponde ao requerimento de que φ não contenha um literal e sua negação.

A transformação necessária para executar a abdução em T no nível objeto é a completação T_c dos predicados não abduzíveis em T. A completação T_c é um conjunto de equivalências $\{p_i \equiv$

$D_i | i = 1, ..., n\}$, onde $p_1, ..., p_n$ são todos átomos não abduzíveis em T e $D_i \equiv Q_{i1} \vee ... \vee Q_{im}$ caso $\{Q_{ij} \rightarrow p_i, j = 1, ..., m\}$ seja o conjunto de cláusulas em T tendo p_i como cabeça.

As conclusões do raciocínio abdutivo (as explicações) são representadas como a noção de fórmula de explicação definida como segue:

Definição 3.3 *Seja $P = \langle T, \varphi \rangle$ um problema abdutivo e T_c a completação dos predicados não abduzíveis de T. A fórmula de explicação para P é aquela mais específica F na linguagem de átomos abduzíveis tal que:*

$$T_c, \varphi \models F$$

onde F é mais específico que F' se e somente se $\models F \rightarrow F'$

A fórmula de explicação é dita caracterizar todas as soluções para um problema abdutivo. O seguinte procedimento mostra, segundo os autores, que tal fórmula existe, e a busca da mesma pára, sempre que T é hierárquica:

Procedimento ABDUCE

- *Reescreva* φ usando as equivalências em T_c (substituindo cada p_i não abduzível pela fórmula D_i obtida na completação; *até que* a fórmula F contendo somente átomos abduzíveis é obtida.

O procedimento ABDUCE apresenta a seguinte propriedade de corretude:

Teorema 3.1 *ABDUCE determina a fórmula de explicação F para um problema abdutivo $P = \langle T, \varphi \rangle$.*

O teorema acima é provado a partir do seguinte lema, que significa que T_c não impõe por si só qualquer relação entre átomos abduzíveis:

Lema 3.1 *Dada uma teoria de domínio T e uma fórmula F na linguagem dos átomos abduzíveis de T, e dada a completação T_c dos predicados não abduzíveis de T, então: $T_c \models F$ se e somente se $T \models F$.*

A definição 3.3 provê uma nova caracterização de abdução mostrando que as explicações para um problema $P = \langle T, \varphi \rangle$ podem ser expressas como conseqüências de φ, e uma apropriada teoria transformada T_c na qual os passos do raciocínio abdutivo são representados explicitamente. Além disso, a definição mostra também que $T_c \models \varphi \leftrightarrow F$ e, consequentemente, $T_c, \varphi \models F$ são facilmente verificadas, já que F é uma reescrita de φ usando as regras de equivalência em T_c.

Uma questão que é considerada importante no artigo é mostrar que a abordagem de nível objeto descrita de fato caracteriza as explicações abdutivas no sentido entendido pela literatura (e que os autores chamam de definição do meta-nível). Para provar essa correspondência, primeiro é introduzida uma definição do meta-nível da abdução:

Definição 3.4 *Dado um problema abdutivo $P = \langle T, \varphi \rangle$ uma m-explicação para P é um conjunto E de átomos abduzíveis tal que:*

$$T \cup E \vdash_{NF} \varphi$$

onde $\vdash_{NF}$ é a derivação de SLDNF.

Essa definição é dita coincidir com aquelas encontradas na literatura quando T contém somente cláusulas definidas e φ contém somente literais positivos. O teorema seguinte mostra a correspondência entre m-explicações para P e modelos da fórmula de explicação para P:

Teorema 3.2 *Seja $P = \langle T, \varphi \rangle$ um problema abdutivo tendo F como a fórmula de explicação. Seja E um conjunto de átomos abduzíveis e M uma interpretação tal que para todo átomo abduzível α*

$$M \models \alpha \text{ se e somente se } \alpha \in E$$

Então E é uma m-explicação se e somente se $M \models F$.

A prova do teorema 3.2 é construída considerando-se que a completação padrão de $T \cup E$ é $T_c \cup Comp(E)$, onde $Comp(E) = E \cup \{\neg\alpha \mid \alpha \text{ é um átomo abduzível e } \alpha \notin E\}$ e considerando-se também a corretude do procedimento ABDUCE e o lema 3.1.

Segundo os autores, o teorema acima é o resultado fundamental que esclarece a ponte entre abdução e dedução através da *completion semantics*.

Em geral, um problema abdutivo apresenta várias soluções, por isso é necessário o estabelecimento de críterios para escolher uma ou outra explicação como sendo preferencial. Os autores consideram o critério de minimalidade como o mais natural para ser introduzido numa estrutura de nível objeto. Em outras palavras, aquelas explicações que envolvem um conjunto mínimo de literais abduzíveis é preferível a outras. Por exemplo, a solução p é preferível às soluções $p \wedge q$ e $p \wedge \neg q$. Sua definição formal é a seguinte:

Definição 3.5 *Dado um problema abdutivo $P = \langle T, \varphi \rangle$ e duas explicações E_1 e E_2, E_1 é preferível a E_2 se e somente se $E_2 \models E_1$.*

Uma explicação E é uma explicação preferencial se e somente se não existe uma explicação que seja preferível a E.

Esse critério foi utilizado por Kleer (citado em [CDT91]) em sua caracterização de Kernel diagnóstico. As explicações preferenciais podem ser entendidas como geradores do conjunto de todas as explicações. Na nossa própria proposta, também utilizaremos as explicações minimais geradas pelo método para generalizar o conjunto de candidatos.

No artigo é considerada a possibilidade dos átomos abduzíveis não serem independentes uns dos outros e esse conhecimento pode ser tratado pelo método.

Dois diferentes tipos de conhecimento adicional são considerados:

- taxonomias ou relações abstratas entre os átomos abduzíveis, do tipo $\alpha \rightarrow \beta$;
- restrições entre átomos abduzíveis na forma de *denials* (contradições) do tipo: $\neg(\alpha_1 \wedge ... \wedge \alpha_n)$.

O primeiro tipo de conhecimento adicional pode evitar redundâncias do tipo: se $\alpha \wedge \beta$ é uma explicação para um problema abdutivo $\langle T, \varphi \rangle$ e sabe-se que $\alpha \rightarrow \beta$ (através de um conhecimento adicional), pode-se reduzir a solução a α. Deve ficar claro, no entanto, que um átomo abduzível β não deve ser explicado em termos de outro átomo abduzível α, a não ser que α deva ser assumido por outras razões. O segundo tipo de conhecimento adicional, as retrições, podem ser utilizadas para reduzir (podar) o conjunto de explicações candidatas produzidas pela abdução.

É importante notar que a teoria deve ser completada para gerar as explicações, mas o conhecimento adicional é utilizado separadamente para explicitar as explicações preferenciais.

Formalmente a definição de abdução é então estendida para lidar com o conhecimento adicional:

Definição 3.6 *Um problema abdutivo é um par $\langle \langle T, A, I \rangle, \varphi \rangle$ onde:*

- *T, (a teoria) é definida como em 3.2;*
- *A é um conjunto hierárquico de cláusulas $\alpha \rightarrow \beta$ (onde α e β são átomos abduzíveis);*
- *I é um conjunto de denials $\neg(\alpha_1 \wedge ... \wedge \alpha_n)$ onde cada α_i é um átomo abduzível;*
- *φ é como na definição 3.2.*

T será usada como antes, para gerar a fórmula de explicação F, enquanto A e I são usados para simplificar F. Dessa maneira a fórmula de explicação resultante será uma conseqüência lógica de $T_c \cup A \cup I \cup \varphi$.

O processo de gerar a fórmula de explicação simplificada (ou transformada) é definido em três passos:

1. Gerar a fórmula de explicação F para P usando ABDUCE;
2. Simplificar a fórmula F usando as restrições T;
3. Simplificar a fórmula F usando a taxonomia A.

Veja [CDT91] para definição detalhada dos passos 2 e 3.

A seguinte propriedade estende a propriedade de que $T_c \models \varphi \leftrightarrow F$ para o caso da abdução com taxonomias e restrições:

Propriedade 3.1 *Seja F^* o resultado do processo acima (passos 1, 2 e 3) para o problema abdutivo $\langle\langle T, A, I\rangle, \varphi\rangle$; temos:*

$$T_c \cup A \cup I \models \varphi \equiv F^*$$

A prova é conseqüência da já mencionada propriedade $T_c \models \varphi \leftrightarrow F$ e da equivalência de F e F^* dado $A \cup I$. Para estender a correspondência entre o nível objeto e o meta-nível da abdução, a seguinte caracterização de meta-nível é introduzida, em substituição à definição 3.4:

Definição 3.7 *Dado um problema abdutivo $P = \langle\langle T, A, I\rangle, \varphi\rangle$, uma m-explicação para P é um conjunto de átomos abduzíveis tal que:*

1. $T \cup A \cup E \vdash_{NF} \varphi$

2. $A \cup I \cup E$ é consistente

Para estender o resultado expresso no teorema 3.2, é necessário caracterizar como uma m-explicação e uma interpretação satisfazendo a fórmula F^* podem ser vistas como correspondentes.

Teorema 3.3 *Seja $P = \langle\langle T, A, I\rangle, \varphi\rangle$ um problema abdutivo; seja F^* o resultado das transformações acima na fórmula de explicação para P; seja E um conjunto de átomos abduzíveis de T e M uma interpretação tal que para todo átomo abduzível α, $M \models \alpha$ se e somente se $A \cup E \models \alpha$. Então, E é uma m-explicação para P se e somente se ($M \models F^*$ e $M \models I$).*

A prova do teorema 3.3 é baseada no fato de que $Comp(T \cup A \cup E) \equiv T_c \cup Comp(A \cup E)$. Inicialmente é provado que (1) dado o modelo mínimo $M_{A\cup E}$ de $A \cup E$, então $M_{A\cup E} \models I$ se e somente se $A \cup E \cup I$ é consistente, e (2) $Comp(A \cup E) \models I$ se e somente se $M_{A\cup E} \models I$. Em seguida a prova é finalizada utilizando-se a propriedade 3.1 e o lema 3.1.

No artigo em questão, as estruturas definidas para a abdução proposicional são estendidas para a teoria de primeira ordem, considerando o seguinte caso em particular:

- a teoria T é um programa em lógica hierárquico de primeira ordem;
- A é um conjunto de implicações do tipo: $\alpha(x) \rightarrow \beta(x)$; onde α e β são símbolos abduzíveis;
- I é um conjunto de restrições do tipo: $\neg(\alpha_1(t_1) \wedge ... \wedge \alpha_n(t_n))$; onde α_i são símbolos abduzíveis;
- a fórmula φ a ser explicada é uma conjunção de literais (possivelmente com variáveis).

A idéia geral continua a mesma, mas alguns detalhes devem ser modificados. O preço a ser pago é que a teoria de igualdade incluída na completação deve ser levada em conta. A introdução da teoria de igualdade tem duas conseqüências principais: primeiro, os aspectos lógicos tornam-se mais complexos. Segundo, transformar a caracterização declarativa em uma estrutura efetiva para computar explicações abdutivas torna-se mais trabalhoso, por exemplo, a simplificação da fórmula de explicação deve levar em conta a teoria de igualdade.

As modificações propostas utilizam a teoria equacional de Clark (CET) referenciada em [CDT91], como a teoria de igualdade incluída na completação:

- Na definição 3.3 a fórmula F é considerada mais específica que F' se e somente se $CET \models F \rightarrow F'$. Como conseqüência, a fórmula de explicação é única sob equivalência lógica considerando a equivalência em CET.

- No procedimento ABDUCE, a reescrita baseada nas definições completadas torna-se mais complexa se se deseja executar as simplificações associadas com a teoria de igualdade.
- Similarmente ao lema 3.1, as definições completadas em T_c não impõem qualquer restrição nos predicados abduzíveis:

Lema 3.2 *Dada a teoria T e a fórmula F na linguagem de predidados abduzíveis e igualdade, e dada a teoria completada T_c de predicados não abduzíveis em T, então $T_c \models F$ se e somente se $CET \models F$.*

- O teorema 3.1 não muda nada, exceto pela introdução de CET na prova.

Para definir e computar m-explicações, um dos problemas abordados é se variáveis devem ser permitidas nas suposições: a solução comum em tal caso é trocar variáveis por constantes de Skolem; por tal razão uma m-explicação é definida como sendo um conjunto de átomos abduzíveis sem variáveis livres (*graound*).

Definição 3.8 *Dado um problema abdutivo $P = \langle T, \varphi \rangle$, uma m-explicação para P é um conjunto E de átomos abduzíveis sem variáveis livres tal que: $T \cup E \vdash_{NF} \varphi$.*

Segundo os autores, um outro problema em computar m-explicações é o fato de que *negation as failure* não pode ser usada para solucionar problemas com objetivos negativos possuindo variáveis livres. De fato, uma avaliação SLDNF pára sem resposta (*flounders*) se é selecionado um objetivo negativo com variáveis livres. Essa limitação não tem correspondência no nível objeto (de fato, a computação baseada em *predicate completion* é proposta a fim de evitar o problema de *floundering*).

Com respeito à correspondência entre o nível objeto e o meta-nível da abdução, um resultado similar aos discutidos anteriormente pode ser provado. Em particular, a idéia é mostrar que existe uma relação entre as m-explicações e os modelos de Herbrand da fórmula de explicação. Os modelos de Herbrand são considerados na linguagem de T, A, I e φ, estendidos para incluir as constantes de Skolem que podem ser introduzidas nas m-explicações.

Teorema 3.4 *Seja $P = \langle\langle T, A, I\rangle, \varphi\rangle$ um problema abdutivo; seja F^* o resultado das transformações satisfazendo a propriedade 3.1 na fórmula de explicação para P. Seja E um conjunto de átomos abduzíveis de T e M uma interpretação de Herbrand tal que para todo átomo abduzível α sem variáveis livres: $M \models \alpha$ se e somente se $A \cup E \models \alpha$. Se φ não pára sem resposta (flounders) em $T \cup A \cup E$, então: E é uma m-explicação para P se e somente se ($M \models F^*$ e $M \models I$).*

Considerações sobre o método

Como Luca Console et. al. concluem em [CDT91], a caracterização de nível objeto proposta trata as explicações abdutivas simplesmente como uma conseqüência lógica da fórmula observada e da teoria transformada, fazendo com que a abdução seja definida usando formas simples de raciocínio. Uma das contribuições do método, é a possibilidade de tratar uniformemente objetos positivos e negativos além de poder lidar com objetos negativos com variáveis.

Uma desvantagem do método, a nosso ver, é que a teoria deve estar representada em uma forma sintática específica para ser tratada como programa em lógica. Além disso, existe a necessidade de separar o conhecimento factual do conhecimento geral, o que, caso não seja levado em consideração, pode levar a resultados completamente incoerentes.

Outra questão que é tratada de forma limitada são os critérios de preferência para a escolha das explicações abdutivas; primeiro a consistência com a teoria é mencionada na definição geral, mas não é explicitada no método (a consistência é alcançada a partir do procedimento ABDUCE); os únicos critérios considerados para a seleção entre hipóteses são o de restrição de integridade e o de minimalidade no nível objeto, este último, como mencionado pelos próprios autores, é um critério muito fraco cuja vantagem é apenas a de ser suportado logicamente e de ser homogêneo com a definição de nível objeto da abdução.

Um ponto importante que deve ser ressaltado é que o que aqui é considerado como definição de meta-nível é apenas a definição de abdução como ela se encontra na literatura (segundo os autores) e não deve ser confundida com o que chamamos, na nossa própria proposta, de critérios de meta-nível, que são critérios preferenciais extra-lógicos de seleção de explicações, que não são tratados de forma alguma em [CDT91].

3.3.2 Raciocínio abdutivo via tableau e cálculo de seqüentes

Algumas pesquisas em abdução buscam trabalhar fora do contexto da programação em lógica. Esse é o caso do artigo [MP93] onde a análise dos conceitos envolvidos no raciocínio abdutivo é baseada em sistemas de prova tipo tableau e tipo Gentzen. Um método abdutivo teórico para lógica clássica de primeira ordem é definido, baseado no cálculo de seqüentes e um outro dual baseado em tableau semântico. Os métodos são ditos corretos e completos e não requerem nenhuma redução preliminar de fórmulas a formas normais. No caso proposicional, duas diferentes caracterizações são dadas para explicações abdutivas. A primeira corresponde à geração de todo o conjunto de explicações minimais e consistentes, onde a minimalidade é checada através da comparação com os elementos do conjunto. A segunda caracterização corresponde a um algoritmo (não determinístico) para a geração de uma única explicação minimal, consistente com a teoria.

As versões de primeira ordem do sistema abdutivo fazem uso de unificação e skolemização dinâmica de fórmulas. A construção da fórmula abduzida utiliza o conceito de de-skolemização (ou skolemização reversa). No entanto a questão da minimalidade, na abdução de primeira ordem, torna-se um grande problema. Como usualmente definida, minimalidade é indecidível por duas diferentes razões: (i) determinar se uma explicação é melhor que outra é, em geral, indecidível; (ii) o conjunto de explicações pode ser infinito. Além disso, por causa de (ii), um elemento minimal pode não existir. Tal problema sugere que o requerimento de minimalidade deva ser relaxado, possivelmente definindo-o com respeito a uma relação mais forte que a conseqüência lógica.

O método proposto por Mayer e Pirri em [MP93], que descreveremos nesta seção, é definido de acordo com as considerações acima.

Em [MP93] um problema abdutivo é dado por uma teoria de *background* Θ (ou um conjunto de fórmulas) e uma fórmula φ a ser explicada, tal que:

- $\Theta \not\models \varphi$
- $\Theta \not\models \neg\varphi$

Uma solução para o problema dado pelo par $\langle \Theta, \varphi \rangle$ deve ser buscada entre as fórmulas α tal que $\Theta \cup \{\alpha\} \models \varphi$.

Aqui também é considerado o fato de que as explicações para um problema abdutivo devem respeitar algumas condições fundamentais para que sejam aceitas como explicações interessantes. E, segundo Mayer e Pirri, embora não haja um consenso quanto aos limites entre interessante e não interessante, as três restrições seguintes são usualmente impostas às explicações:

(*i*) α é consistente com Θ (ou Θ-consistente);

(*ii*) α é uma explicação minimal para o problema abdutivo $\langle \Theta, \varphi \rangle$, isto é, para qualquer fórmula β, se $\Theta \cup \{\beta\} \models \varphi$ e $\alpha \models \beta$, então $\models \beta \equiv \alpha$.

(*iii*) α tem alguma forma sintática restrita; por exemplo, ela deve ser uma fórmula prenex cuja matriz é uma conjunção de literais.

Em [MP93] a restrição sintática imposta é dada pelas seguintes definições:

Definição 3.9 (C-fórmula) *Uma fórmula α é uma C-fórmula (de primeira ordem) se é construída com literais usando somente quantificadores e conjunção.*

Definição 3.10 (Explicação para problemas abdutivos) *α é uma explicação para um problema abdutivo $\langle\Theta, \varphi\rangle$ se α é uma C-fórmula na linguagem de $\Theta \cup \{\varphi\}$ e $\Theta \cup \{\alpha\} \models \varphi$.*

Na primeira caracterização proposta para o caso proposicional, a geração de todo o conjunto de explicações minimais e Θ-consistentes é construído por um método incremental que usa os ramos do tableau (ou as folhas da árvore de derivação) uma a uma e as descarta após utilizadas. A segunda carcterização requer que um dado conjunto de ramos de tableau (ou folhas de uma árvore de derivação) seja armazenado e usado até que o algoritmo termine.

No caso de primeira ordem, o método proposto constrói o conjunto de explicações de uma maneira incremental e a checagem da minimalidade é feita de maneira muito imprecisa.

Para o caso proposicional, Mayer e Pirri consideram uma linguagem proposicional L_0, contendo dois valores proposicionais distintos, *true* e *false*. Fórmulas, cláusulas e literais são definidas como usualmente. É feita a convenção de que a disjunção vazia equivale ao átomo *false* e a conjunção vazia equivale ao átomo *true*. Duas disjunções (conjunções) são consideradas iguais - e o sinal de igualdade '=' será utilizado - se elas disjuntam (conjuntam) conjuntos de elementos iguais.

A seguir serão colocadas as definições básicas sobre tableau e cálculo de seqüentes (como definidas em [MP93]), necessárias para a compreensão do método.

Tableaus semânticos são usados como sistemas de refutação. São construídos a partir do seguinte conjunto de regras de expansão:

$$(\neg - regras) \quad \frac{\neg\neg\varphi}{\varphi} \qquad \frac{\neg false}{true}$$

$$(\alpha - regras) \quad \frac{\alpha_1 \wedge \alpha_2}{\alpha_1 ... \alpha_2} \qquad \frac{\neg(\alpha_1 \rightarrow \alpha_2)}{\alpha_1 ... \neg\alpha_2} \qquad \frac{\neg(\alpha_1 \vee \alpha_2)}{\neg\alpha_1 ... \neg\alpha_2}$$

$$(\beta - regras) \quad \frac{\beta_1 \vee \beta_2}{\beta_1 \mid \beta_2} \qquad \frac{\beta_1 \rightarrow \beta_2}{\neg\beta_1 \mid \beta_2} \qquad \frac{\neg(\beta_1 \wedge \beta_2)}{\neg\beta_1 \vee \neg\beta_2}$$

Se φ é uma fórmula, um tableau para φ é uma árvore cuja raiz é rotulada por φ e todo nó diferente da raiz é obtido pela aplicação de uma regra de expansão em um nó precedente do mesmo tableau.

Definição 3.11 (Tableau) *Seja* $\{\varphi_1, ..., \varphi_n\}$ *um conjunto finito de fórmulas de* L_0*:*

1. *A seguinte árvore de um único ramo é um tableau para* $\{\varphi_1, ..., \varphi_n\}$*:*

$$\begin{array}{c} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{array}$$

2. *Se* T *é um tableau para* $\{\varphi_1, ..., \varphi_n\}$ *e* T^* *resulta de* T *pela aplicação de uma regra de expansão de tableau, então* T^* *é também um tableau para* $\{\varphi_1, ..., \varphi_n\}$*.*

Cada ramo de um tableau pode ser visto como a conjunção das fórmulas que aparecem nele e o tableau como um todo é a disjunção de seus ramos. Um ramo de tableau é satisfatível se a conjunção de todas as fórmulas rotulando o ramo é satisfatível. Um tableau é satisfatível se um de seus ramos é satisfatível. O sistema de tableau preserva a satisfatibilidade.

Um ramo B de um tableau é chamado fechado se tanto φ quanto $\neg\varphi$ ocorrem em B, para alguma fórmula φ, ou se o átomo *false* ocorre em B; caso contrário ele é dito aberto. Um tableau é fechado se e somente se todos os seus ramos são fechados.

Definição 3.12 (Refutação de tableau) *Uma refutação de tableau para* φ *é um tableau fechado para* φ*.*

Definição 3.13 (Prova de tableau) *Uma prova de tableau pra* φ *é uma refutação de tableau para* $\neg\varphi$*.* φ *é um teorema do sistema de tableau se existe uma prova de tableau para* φ*.*

O sistema de tableau é correto e completo, isto é, φ é uma tautologia se e somente se tem uma prova de tableau.

Um cálculo tipo Gentzen é um sistema de provas dado por um conjunto de regras que preservam a validade. Provas são árvores rotuladas por seqüentes, isto é, construções do tipo $\Gamma \Rightarrow \Delta$, onde Γ e Δ são conjuntos finitos de fórmulas e $\Rightarrow$ é um novo símbolo. Um seqüente $\Gamma \Rightarrow \Delta$ é verdade em uma interpretação M se ou alguma fórmula em Γ é falsa em M ou alguma fórmula em Δ é verdade em M. Em outras palavras, o seqüente $\{\gamma_1, ..., \gamma_n\} \Rightarrow \{\delta_1, ..., \delta_m\}$ é interpretado como a implicação $(\gamma_1 \wedge ... \wedge \gamma_n) \rightarrow (\delta_1 \vee ... \vee \delta_m)$.

As seguintes regras são as regras de inferência proposicionais do cálculo de seqüentes clássico:

$$(\neg \Rightarrow) \quad \frac{\Gamma \Rightarrow \alpha, \Delta}{\Gamma, \neg\alpha \Rightarrow \Delta} \qquad (\Rightarrow \neg) \quad \frac{\Gamma, \alpha \Rightarrow \Delta}{\Gamma \Rightarrow \neg\alpha, \Delta}$$

$$(\wedge \Rightarrow) \quad \frac{\Gamma, \alpha, \beta \Rightarrow \Delta}{\Gamma, \alpha \wedge \beta \Rightarrow \Delta} \qquad (\Rightarrow \wedge) \quad \frac{\Gamma \Rightarrow \alpha, \Delta; \Gamma \Rightarrow \beta, \Delta}{\Gamma \Rightarrow \alpha \wedge \beta, \Delta}$$

$$(\vee \Rightarrow) \quad \frac{\Gamma, \alpha \Rightarrow \Delta; \Gamma, \beta \Rightarrow \Delta}{\Gamma, \alpha \vee \beta \Rightarrow \Delta} \qquad (\Rightarrow \vee) \quad \frac{\Gamma \Rightarrow \alpha, \beta, \Delta}{\Gamma \Rightarrow \alpha \vee \beta, \Delta}$$

$$(\rightarrow \Rightarrow) \quad \frac{\Gamma \Rightarrow \alpha, \Delta; \Gamma, \beta \Rightarrow \Delta}{\Gamma, \alpha \rightarrow \beta \Rightarrow \Delta} \qquad (\Rightarrow \rightarrow) \quad \frac{\Gamma, \alpha \Rightarrow \beta, \Delta}{\Gamma \Rightarrow \alpha \rightarrow \beta, \Delta}$$

Um axioma do cálculo de seqüentes é um seqüente $\Gamma \Rightarrow \Delta$ tal que $\Gamma \cap \Delta \neq \emptyset$.

De acordo com o exposto por Mayer e Pirri, as regras acima podem ser lidas de cima para baixo (permitindo a elaboração de deduções e a construção de provas) ou de baixo para cima, onde a questão da validade de um seqüente é **reduzida** à questão da validade de um ou dois seqüentes mais simples. Como o método proposto em [MP93] considera principalmente a leitura de baixo para cima, o referido cálculo é chamado de cálculo de redução (*reduction calculus*).

Definição 3.14 (Árvore de redução - R-árvore) *Seja Σ um seqüente. Uma árvore de redução T para Σ é uma árvore finita cuja raiz (seqüente final) é Σ e todo nó Σ_i em T ou é uma folha ou é a conclusão de uma regra de inferência do cálculo de redução, onde a(s) premissa(s) é (são) o(s) nó(s) em T imediatamente sobre Σ_i.*

Definição 3.15 (Prova de seqüente) *Uma R-árvore T sobre um seqüente Σ é uma prova de seqüente para Σ se todas as suas folhas são axiomas.*

O cálculo de seqüentes é correto e completo: um seqüente proposicional é válido se e somente se existe uma árvore de prova para ele.

Aplicando as definições acima à questão da abdução, os autores definem o seguinte: seja $\langle\Theta, \varphi\rangle$ um problema abdutivo. Como $\Theta \not\models \varphi$, não há uma refutação de tableau para $\Theta \cup \{\neg\varphi\}$, e não há uma prova de seqüente para $\Theta \Rightarrow \varphi$. A solução do problema abdutivo $\langle\Theta,\varphi\rangle$ pode ser encontrada entre as fórmulas que forçam o fechamento de um tableau para $\Theta \cup \{\neg\varphi\}$; equivalentemente, entre as fórmulas que, quando adicionadas ao lado esquerdo dos seqüentes do topo de uma R-árvore para $\Theta \Rightarrow \varphi$, os torna válidos.

Claro que não tem sentido fechar um tableau que não pode ser expandido simplesmente adicionando o átomo falso ou a própria fórmula φ a ser explicada. As árvores devem ser expandidas o máximo possível e uma cuidadosa escolha dos literais que fecham cada ramo deve ser feita.

Definição 3.16 (Árvore aceitável) *Seja T um tableau. Um ramo B de T é* **fundamental** *se cada fórmula não literal de B foi expandida. T é dito* **aceitável** *se todas as suas folhas são fundamentais.*

Seja T uma R-árvore. Um seqüente Σ de T é fundamental se e somente se ele contém somente átomos. T é aceitável se todas as suas folhas são fundamentais.

Definição 3.17 (*Threads* e conjuntos de fechamento) *Seja B um ramo fundamental de T. O* **thread** *associado a B, $p(B)$ é o conjunto de literais rotulando os nodos de B. O* **conjunto de fechamento** *para B é $\tau(B) = \{\neg\lambda \mid \lambda \in p(B)\}$ onde $\neg\lambda$ é o complemento de λ.*

$S(T)$ é igual ao conjunto dos conjuntos de fechamento mínimos de $T = \{\tau(B) \mid B$ é um ramo aberto de T e não existe um ramo B' em T tal que $\tau(B') \subset \tau(B)\}$.

Seja $\Sigma = \Gamma \Rightarrow \Delta$ um seqüente fundamental. O **thread** *associado a Σ, $p(\Sigma)$ é $\Gamma \cup \{\neg\lambda \mid \lambda \in \Delta\}$. O* **conjunto de fechamento** *para Σ é $\tau(\Sigma) = \{\neg\lambda \mid \lambda \in p(\Sigma)\} = \{\neg\lambda \mid \lambda \in \Gamma\} \cup \Delta$.*

Se T é uma R-árvore aceitável, o conjunto $S(T)$ dos conjuntos de fechamento minimais de T é $\{\tau(\Sigma) \mid \Sigma$ é uma folha não axioma em T e não existe nehuma folha Σ' em T tal que $\tau(\Sigma') \subset \tau(\Sigma)\}$.

$S(T)$ coleciona os subconjuntos minimais dos conjuntos de fechamento dos ramos abertos (ou folhas não axiomas) de T.

Definição 3.18 (Fechamentos (*closures*)) *Seja T um tableau aceitável (uma R-árvore aceitável) para o problema abdutivo $\langle\Theta,\varphi\rangle$ e $S(T) = \{\tau_1, ..., \tau_n\}$ os conjuntos de fechamento de T. Seja g qualquer função escolhida para os elementos $S(T)$, isto é, $g(\tau_i) \in \tau_i$.*

(i) Se $S(T) = \emptyset$, T tem um único fechamento, o átomo true *(é o caso em que $\Theta \models \varphi$);*

(ii) *Se* $S(T) \neq \emptyset$ *e para qualquer função* g *para os elementos de* $S(T)$ *o conjunto* $\{g(\tau_1), ..., g(\tau_n)\}$ *contém um par de literais complementares, então o único fechamento para* T *é o átomo* false *(é o caso em que* $\Theta \models \neg\varphi$*);*

(iii) *Caso contrário* $\alpha = g(\tau_1) \wedge ... \wedge g(\tau_n)$ *é um fechamento para* T *se e somente se* α *não contém um par de literais complementares.*

$\mathcal{E}(T)$ é o conjunto de todos os α's que são um fechamento para T. Como pode-se ter diferentes árvores aceitáveis para um mesmo problema abdutivo com os mesmos fechamentos, então $\mathcal{E}(\Theta, \varphi) = \mathcal{E}(T)$ para qualquer árvore T para $\langle\Theta, \varphi\rangle$.

Mayer e Pirri propõem o seguinte teorema que estabelece completude e um tipo de corretude do procedimento que resolve problemas abdutivos pela geração do conjunto $\mathcal{E}(\Theta, \varphi)$:

Teorema 3.5 *Seja* $\langle\Theta, \varphi\rangle$ *um problema abdutivo. Então, todo elemento de* $\mathcal{E}(\Theta, \varphi)$ *é uma explicação não contraditória para* $\langle\Theta, \varphi\rangle$ *(corretude) e qualquer explicação minimal e não contraditória para* $\langle\Theta, \varphi\rangle$ *é um elemento de* $\mathcal{E}(\Theta, \varphi)$ *(completude).*

Com respeito à minimalidade, em [MP93] é argumentado que do teorema 3.5 tem-se que o conjunto das explicações minimais para $\langle\Theta, \varphi\rangle$ é igual ao conjunto $min(\mathcal{E}(\Theta, \varphi))$, ou seja, os fechamentos minimais para qualquer árvore T para $\langle\Theta, \varphi\rangle$: $min(\mathcal{E}(\Theta, \varphi)) = min(\{\alpha \mid$ para todo $\tau \in S(T), (\alpha \cup \tau) \neq \emptyset\})$ onde, se Γ é um conjunto de fórmulas, $min(\Gamma)$ denota seu subconjunto que contém somente elementos minimais com respeito a $\models$.

Definição 3.19 *Seja* $\langle\tau_1, ..., \tau_m\rangle$ *um conjunto ordenado não vazio qualquer de conjuntos de literais e* α *uma conjunção de literais. Então* α *é minimamente determinado por* $\langle\tau_1, ..., \tau_m\rangle$ *se e somente se uma das seguintes condições ocorre:*

(a) $m = 1$ *e* $\alpha \in \tau_1$*;*

(b) $m > 1$, α *é minimamente determinado por* $\langle\tau_1, ..., \tau_{m-1}\rangle$ *e* $\alpha \cap \tau_m \neq \emptyset$*;*

(c) $m > 1$, *a condição (b) não é satisfeita e existe um* $\lambda \in \alpha \cap (\tau_m - (\tau_1 \cup ... \cup \tau_{m-1}))$ *tal que se* $\alpha \equiv \gamma \wedge \lambda$, *então* $\neg\lambda \notin \gamma$ *e* γ *é minimamente determinado por* $\langle\tau_1, ..., \tau_{m-1}\rangle$.

Definição 3.20 *Seja $\tau_1, ..., \tau_m$ conjuntos de literais e α uma conjunção de literais. Então α é minimamente determinada por $\{\tau_1, ..., \tau_m\}$ se e somente se existe uma permutação $\langle \tau_{p1}, ..., \tau_{pm} \rangle$ de $\tau_1, ..., \tau_m$ tal que α é minimamente determinado por $\langle \tau_{p1}, ..., \tau_{pm} \rangle$.*

Lema 3.3 *Seja T uma árvore com um conjunto $S(T)$ não vazio e tal que $false$ não é um fechamento minimal para T. Então uma C-fórmula α é um fechamento minimal para T se e somente se α é minimamente determinado por $S(T)$.*

Com respeito à consistência, tem-se em [MP93] que se Θ é um conjunto consistente de fórmulas e T é algum tableau para Θ, então T será genericamente chamado de árvore analítica para Θ e $\mathcal{E}(\Theta)$ denotará $\mathcal{E}(T)$, ou seja, $\mathcal{E}(\Theta, false)$.

Para obter de $\mathcal{E}(\Theta, \varphi)$ as explicações para $\langle \Theta, \varphi \rangle$ que são consistentes com Θ, deve-se notar que o conjunto de todas as explicações minimais e Θ-consistentes para $\langle \Theta, \varphi \rangle$ é $min(\mathcal{E}(\Theta, \varphi)) - min(\mathcal{E}(\Theta))$. Então pode-se primeiro gerar todas as explicações minimais para a negação da teoria (isto é, o conjunto $min(\mathcal{E}(\Theta)))$ e então removê-las de $min(\mathcal{E}(\Theta, \varphi))$. Mayer e Pirri argumentam que isso é razoável se assumirmos que a teoria Θ não é frequentemente modificada, então o conjunto $min(\mathcal{E}(\Theta))$ pode ser armazenado uma única vez e usado cada vez que um problema abdutivo deva ser resolvido.

Baseado nas observações acima, os autores seguem com uma caracterização de explicação minimal e Θ-consistente, que não se refere ao conjunto de todas as explicações para o problema abdutivo, mas faz uso da coleção de conjuntos de fechamento de alguma árvore analítica para o problema abdutivo.

Se T é alguma árvore para $\langle \Theta, \varphi \rangle$, seja T' uma sub-árvore de T onde nenhuma regra é aplicada a $\neg\varphi(\varphi)$. A caracterização de explicação Θ-consistente para $\langle \Theta, \varphi \rangle$ se refere aos dois conjuntos de conjuntos de fechamento $S(T)$ e $S(T')$ e é justificada pelo seguinte lema:

Lema 3.4 *Seja T uma árvore analítica para Θ. Então uma C-fórmula α é consistente com Θ se e somente se existe $\tau \in S(T)$ tal que $\alpha \cap \tau = \emptyset$.*

De fato, se α é consistente com Θ, então existe algum ramo em alguma árvore para Θ que não é fechada adicionando (e expandindo) α. Então existe também um ramo minimal que não é fechado por α. Então para algum $\tau \in S(T)$ deve acontecer que $\alpha \cap \tau = \emptyset$.

Os lemas acima se juntam no seguinte:

Teorema 3.6 *Seja T uma árvore para $\langle\Theta,\varphi\rangle$ e T' uma sub-árvore de T onde nenhuma regra é aplicada para $\neg\varphi(\varphi)$. Então uma C-fórmula α é uma explicação minimal e Θ-consistente para $\langle\Theta,\varphi\rangle$ se e somente se α é minimamente determinada por $S(T)$ e existe $\tau \in S(T')$ tal que $\alpha\cap\tau = \emptyset$.*

O teorema acima, segundo os autores, oferece um caminho para se definir um algoritmo correto e completo para a construção de uma explicação única para um problema abdutivo.

Abdução de primeira ordem

A seguir as principais definições utilizadas em [MP93] para os casos de tableau e cálculo de seqüentes para a lógica de primeira ordem, são introduzidas.

Seja L uma linguagem de primeira ordem estendendo L_0. Termos, fórmulas, literais, ocorrência de variáveis livres e ligadas em uma fórmula são definidos como de costume. L^{SKO} é a extensão de L, obtida pela adição de um conjunto (enumerável) de novos símbolos, as meta-variáveis $d_0, d_1, d_2, \ldots$ e, para qualquer n, um conjunto (enumerável) de símbolos de função com aridade n, chamadas funções de Skolem ou h-funções $h_0^n, h_1^n, \ldots$ (o expoente será sempre omitido).

As regras de tableau de primeira ordem incluem as regras de expansão proposicionais e as seguintes regras onde d é uma meta-variável que não ocorre em outro lugar senão no tableau, h é uma nova função de Skolem e $d_1, \ldots, d_n$ são todas as meta-variáveis ocorrendo no ramo:

$$(\gamma - regras) \qquad \frac{\forall x\gamma(x)}{\gamma(d)} \qquad \frac{\neg\exists x\gamma(x)}{\neg\gamma(d)}$$

$$(\delta - regras) \qquad \frac{\exists x\delta(x)}{\delta(h(d_1,\ldots,d_n))} \qquad \frac{\neg\forall x\delta(x)}{\neg\delta(h(d_1,\ldots,d_n))}$$

O cálculo de redução que é adotado no artigo contém as seguintes regras de quantificadores, onde d é uma nova meta-variável que não ocorre em outro lugar senão na R-árvore, h é uma nova função de Skolem e $d_1, \ldots, d_n$ são todas as meta-variáveis que ocorrem em $\Gamma, \forall x\alpha(x), \Delta$.

$$(\forall \Rightarrow) \quad \frac{\Gamma, \alpha(d), \forall x\alpha(x) \Rightarrow \Delta}{\Gamma, \forall x\alpha(x) \Rightarrow \Delta} \qquad (\Rightarrow \forall) \quad \frac{\Gamma \Rightarrow \alpha(h(d_1,\ldots,d_n), \Delta}{\Gamma \Rightarrow \forall x\alpha(x), \Delta}$$

$$(\exists \Rightarrow) \quad \frac{\Gamma, \alpha(h(d_1,\ldots,d_n)) \Rightarrow \Delta}{\Gamma, \exists x\alpha(x) \Rightarrow \Delta} \qquad (\Rightarrow \exists) \quad \frac{\Gamma \Rightarrow \alpha(d), \exists x\alpha(x), \Delta}{\Gamma \Rightarrow \exists x\alpha(x), \Delta}$$

As regras irrestritas, isto é, ($\forall \Rightarrow$) e ($\Rightarrow \exists$) correspondem às γ-regras para tableau, enquanto ($\exists \Rightarrow$) e ($\Rightarrow \forall$) correspondem às δ-regras.

Claramente, como as γ-regras podem ser aplicadas muitas vezes infinitamente, árvores de primeira ordem podem ser infinitas. Porém, as definições de R-ávores e tableaus são estendidas para a primeira ordem de forma que somente árvores finitas sejam consideradas.

Definição 3.21 (Tableau e árvore de redução não instanciadas - U-árvores) *Seja Σ um conjunto de fórmulas. Um tableau não instanciado T para Σ é uma árvore finita que é, ou um tableau de um ramo para Σ, ou é obtida de um tableau para Σ pela aplicação de regras de expansão de primeira ordem.*

Seja Σ um seqüente. Uma árvore de redução não instanciada T sobre Σ é uma árvore finita cuja raiz (seqüente final) é Σ e todo nó Σ_i em T ou é uma folha ou é a conclusão de uma regra de inferência do cálculo de redução onde as premissas são os nós em T imediatamente sobre Σ_i.

Substituições são definidas como usualmente. Elas devem afetar apenas meta-variáveis, que são todas distintas das variáveis ligadas, de forma que as substituições são sempre livres.

Definição 3.22 (Árvore instanciada - I-árvore) *Π é uma árvore instanciada (ou um tableau ou uma R-árvore) se é obtida pela aplicação de uma substituição θ de termos para meta-variáveis para toda fórmula de uma U-árvore. Isto é, para toda U-árvore T e substituição θ, $T\theta$ é uma I-árvore.*

Definição 3.23 (Refutações e provas) *Um tableau instanciado Π para Σ é uma refutação de Σ se ele é fechado. Π é uma prova para φ se ele é fechado para $\neg\varphi$.*

Uma árvore de redução instanciada Π para Σ é uma prova de Σ se todas as suas folhas são axiomas.

Seja Σ um conjunto de fórmulas (um seqüente). Como observado em [MP93], no caso geral de primeira ordem, uma coleção possivelmente infinita de U-árvores pode ser construída para Σ: $T_1, T_2, \ldots$ Duas delas podem diferir ou na ordem da aplicação das regras do cálculo, ou porque uma é a extensão da outra, ou porque ambas são extensões de duas outras U-árvores que são permutações uma da outra. De fato, apesar da reusabilidade de γ-regras, as regras δ e γ não são permutáveis.

Além disso, toda U-árvore T_i para Σ pode corresponder a diferentes I-árvores, uma para cada substituição aplicada a T_i.

As seguintes definições determinam quando uma árvore é aceitável para abdução, isto é, quando sua construção pode ser interrompida para que seja gerado um conjunto de explicações:

Definição 3.24 (U-árvore aceitável - AU-árvore) *Um ramo em um tableau é fundamental se para toda ocorrência não literal no ramo foi aplicada a apropriada regra de expansão pelo menos uma vez.*

Um seqüente é fundamental se e somente se ele contém somente átomos ou fórmulas da forma $\exists x\alpha$ no conseqüente ou $\forall x\alpha$ no antecedente.

Uma U-árvore é aceitável se todos os seus ramos são fundamentais.

Definição 3.25 (I-árvore aceitável - AI-árvore) *Uma I-árvore $\Pi = T\theta$ é aceitável se e somente se T é uma AU-árvore.*

Em seguida, Mayer e Pirri mostram como construir um conjunto finito de explicações $\mathcal{E}^{FOL}(\Pi)$ para $\langle\Theta,\varphi\rangle$, para cada uma das (finitas) AI-árvores Π para um problema abdutivo $\langle\Theta,\varphi\rangle$.

Segundo os autores, corretude vai significar dizer que se uma dada fórmula α é um elemento de $\mathcal{E}^{FOL}(\Pi)$, então α é não contraditória em $\Theta \cup \{\alpha\} \models \varphi$. Porém, α não é necessariamente minimal. Em particular, pode acontecer o caso de $\Theta \models \varphi$ mesmo se explicações não válidas são geradas. Completude vai garantir que, para qualquer fórmula contraditória α tal que $\Theta \cup \{\alpha\} \models \varphi$, existe uma AI-árvore Π para $\langle\Theta,\varphi\rangle$ tal que $\mathcal{E}^{FOL}(\Pi)$ contém uma conseqüência lógica γ de α.

As noções de *thread*, conjunto de fechamento e fechamento para AI-árvores Π são extensões imediatas das correspondentes proposicionais. Porém, no caso de seqüentes, só literais são colecionados nos *threads* e conjuntos de fechamento. Se Π é uma AI-árvore, então $\mathcal{E}(\Pi)$ significa, como antes, o conjunto de todos os fechamentos para Π.

Agora, se um elemento de $\mathcal{E}(\Pi)$ está na linguagem $L^{SKO} \supseteq L$, então ele pode não ser uma explicação para $\langle\Theta,\varphi\rangle$. A fim de obter explicações, todas as meta-variáveis e os termos de Skolem introduzidos na árvore para $\langle\Theta,\varphi\rangle$ serão trocados por variáveis convenientemente quantificadas.

Definição 3.26 (Skolemização reversa) *Seja α uma fórmula $\in L^{SKO}$ e seja $st(\alpha) = \{t_1, ..., t_k\}$ o conjunto de termos ocorrendo em α que não estão em L. Seja $\langle t_{p1}, ..., t_{pk}\rangle$ uma ordem total* *qualquer dos elementos de* $st(\alpha)$ *tal que para todo* i *e* j *se* t_{pi} *ocorre propriamente em* t_{pj}*, ou seja, se* t_{pi} *é um subconjunto próprio de* t_{pj}*, então* $i < j$*. Então* $Q_{x1}, ...Q_{xk}\alpha'$ *é obtido a partir de* α *por skolemização reversa na base de* $\langle t_{p1}, ..., t_{pk}\rangle$ *se e somente se:*

- α' *é obtido de* α *pela troca de cada termo* t_{pi} *com a nova variável* x_i*;*
- *para todo* i*, se* t_{pi} *é uma meta-variável, então* x_i *é existencialmente quantificado, caso contrário, se* t_{pi} *é um termo de Skolem, então* x_i *é universalmente quantificado.*

O conjunto de todas as fórmulas que podem ser obtidas por skolemização reversa de α será denotado por $desk(\alpha)$.

Definição 3.27 (Fechamento de primeira ordem) *Seja* Π *uma I-árvore qualquer para o problema abdutivo* $\langle\Theta, \varphi\rangle$ *e* $\mathcal{E}(\Pi)$ *o conjunto de todos os fechamentos para* Π*. Então o conjunto dos fechamentos de primeira ordem de* Π *é:*

$$\mathcal{E}^{FOL}(\Pi) = \{\alpha \mid \alpha \in min(desk(\gamma)) \textit{ para algum } \gamma \in min(\mathcal{E}(\Pi))\}$$

Se $\langle\Theta, \varphi\rangle$ *é um problema abdutivo e* $R = \{\Pi \mid \Pi$ *é uma AI-árvore para* $\langle\Theta, \varphi\rangle\}$*, então*

$$\mathcal{E}^{FOL}(\Theta, \varphi) = \bigcup_{\Pi \in R} \mathcal{E}^{FOL}(\Pi)$$

O teorema seguinte é proposto em [MP93] e estabelece que o cálculo abdutivo que gera elementos de $\mathcal{E}^{FOL}(\Theta, \varphi)$ é correto e completo.

Teorema 3.7 *Seja* $\langle\Theta, \varphi\rangle$ *um problema abdutivo. Então para qualquer elemento* α *de* $\mathcal{E}^{FOL}(\Theta, \varphi)$*,* $\Theta \cup \{\alpha\} \models \varphi$ *(corretude). Se* γ *é uma C-fórmula consistente tal que* $\Theta \cup \{\gamma\} \models \varphi$*, então existe um elemento* α *de* $\mathcal{E}^{FOL}(\Theta, \varphi)$ *tal que* $\gamma \models \alpha$ *(completude).*

Considerações sobre o método

Em [MP93] é apresentada uma interessante caracterização de abdução com uma conceituação simples e bem definida, que, inclusive, utilizaremos (de forma adaptada) em nossa própria proposta.

No caso proposicional da caracterização de Mayer e Pirri, é alcançada a geração do conjunto de todas as explicações minimais e Θ-consistentes ou a geração de uma única explicação. Nesse último caso, porém, a explicação gerada é a primeira a ser encontrada que satisfaça certos requisitos. Nada garante que melhores explicações não poderiam ser encontradas posteriormente.

Ainda no caso proposicional, apesar do próprio artigo mencionar a importante questão da existência de critérios meta-lógicos para a seleção de explicações interessantes, esses critérios não são considerados. A seleção das soluções se dá apenas com respeito à minimalidade e à consistência com a teoria.

Para o caso de primeira ordem, bem mais complexo, uma idéia interessante é proposta para a geração de explicações: o processo de de-skolemização (ou skolemização reversa). No entanto, como os próprios autores comentam, a característica de ser infinito não pode ser eliminada na lógica de primeira ordem. Porém, os problemas a que se chega com a indecidibilidade do cálculo de predicados refletem principalmente sobre os critérios de preferência de consistência com a teoria (obviamente) e minimalidade. De fato, não somente a implicação entre C-fórmulas de primeira ordem é indecidível como também um problema abdutivo pode ter um infinito número de explicações; então, em geral, não se pode determinar se uma dada explicação é minimal somente pela comparação com outras. Existem até casos em que nenhuma explicação é minimal. Essas observações sugerem que ambos os critérios de consistência com a teoria e de minimalidade devem ser relaxados de alguma maneira. Essa questão não é desenvolvida no artigo, apenas é sugerido que a minimalidade poderia ser tratada em termos de uma relação de ordem $\sqsubseteq$ entre C-fórmulas que seja decidível, tal que para qualquer α e β, se $\alpha \sqsubseteq \beta$, então $\beta \models \alpha$.

Como comentário final, os autores notam que na lógica de primeira ordem é indecidível até mesmo determinar se um par $\langle \Theta, \varphi \rangle$ é um problema abdutivo genuíno. Consequentemente, qualquer método para executar abdução em lógica de primeira ordem não deve basear-se nas suposições de que $\Theta \not\models \varphi$ e $\Theta \not\models \neg\varphi$.

3.3.3 Uma caracterização lógica da abdução

Muitas são as técnicas que objetivam formalizar o raciocínio explicativo, no entanto a natureza lógica da abdução ainda está longe de tornar-se clara e diferentes especificações dos conceitos chaves subjacentes têm sido dadas, o que torna difícil falar de abdução como uma forma de raciocínio única e bem definida.

Em [MP95b] é investigada a natureza lógica da abdução buscando unir os aspectos comuns entre as técnicas baseadas em lógica e ainda adicionando diferentes caracterizações dos conceitos básicos. É dada uma definição parametrizada da relação abdutiva que permite diferentes instanciações e cujas propriedades lógicas podem ser estudadas dependentemente das propriedades dos parâmetros. Essas definições abrem caminho para a especificação de uma lógica de abdução, tal que identificaria as propriedades estruturais da relação abdutiva e as condições que acarretam a possibilidade de determinar completamente suas propriedades lógicas.

Segundo [MP95b], um dos aspectos chaves que distingue implicação material de implicação explicativa é que explicações são suposições minimais que implicam o *explanandum* (o que se deseja explicar). Apesar da minimalidade por si só estar longe de capturar o conceito de causalidade, ela é uma peculiaridade da abdução e a principal responsável por seus problemas computacionais.

A minimalidade em [MP95b] é definida com respeito a uma ordem parcial das fórmulas: um critério de preferência (meta-lógico), que permite a comparação de duas explicações em potencial a fim de determinar qual delas é melhor que a outra. O critério de preferência, segundo os autores, tem sido definido de diferentes maneiras, desde absolutamente gerais a outras muito empíricas. Normalmente a preferência é dada a hipóteses mais simples, onde a simplicidade é entendida como uma propriedade qualitativa que não pode ser reduzida a uma propriedade lógica. Em contextos diferentes o critério de preferência é entendido de maneiras diferentes.

O papel chave da minimalidade e do critério de preferência é salientado pelo reconhecimento de que outros requisitos nas explicações podem ser embutidos em uma definição apropriada da relação de preferência. Até mesmo consistência com a teoria, por exemplo, pode ser atendida pelo requisito de que qualquer solução consistente com a teoria é preferível a outras inconsistentes. Na abordagem de programação em lógica onde as explicações devem conter somente átomos abduzíveis (como visto na seção 3.3.1), qualquer fórmula satisfazendo tal propriedade é preferível a fórmulas contendo predicados não abduzíveis. Uma fórmula satisfazendo um dado conjunto de restrições é preferível a fórmulas que não satisfazem.

Nenhuma relação de preferência pode ser dita como apropriada a qualquer lógica ou aplicação. Por isso, uma abordagem lógica genérica para a abdução não deveria se comprometer com nenhuma definição dada para a relação de preferência.

Segundo Mayer e Pirri, uma definição precisa de explicações abdutivas é obviamente dependente da lógica subjacente à teoria de *background*: alguns fatos α podem explicar φ em lógica clássica, mas não em lógica intuicionística ou linear. Então é considerado que uma lógica L é dada como parâmetro, com uma relação de conseqüência lógica correspondente. E assim, uma definição de abdução será dependente não só das propriedades da relação de preferência que é usada como também da lógica subjacente à teoria de *background*.

A lógica e a relação de preferência não são os únicos parâmetros de uma caracterização genérica do raciocínio abdutivo. Um terceiro parâmetro é o conjunto de hipóteses alternativas com o qual uma boa explicação deve ser comparada.

A seguinte definição de explicação abdutiva é considerada. Trata-se de uma simplificação (omitindo-se o critério de consistência) da definição central em muitas abordagens para abdução. Ela é chamada de 'o caso geral' no artigo, exercendo o papel de um limite inferior (*lower bound*) de relações abdutivas aceitáveis, isto é, a mais forte noção de abdução que é considerada no referido trabalho:

($\diamond$) α é uma boa explicação para β no contexto da teoria Θ se $\Theta \vdash_L \alpha \rightarrow \beta$ e, para qualquer γ tal que $\Theta \vdash_L \gamma \rightarrow \beta$, se $\vdash_L \alpha \rightarrow \gamma$, então $\vdash_L \gamma \rightarrow \alpha$.

Aqui α é aceita como uma boa explicação se não há outra estritamente melhor (mais fraca com respeito à implicação lógica) no conjunto dado: o conjunto de todos os outros possíveis candidatos à explicação, isto é, o conjunto daqueles γ's tal que $\Theta \vdash_L \gamma \rightarrow \beta$. No raciocínio prático, na verdade, uma hipótese é aceita como boa razão para a verdade de um dado fato até que uma melhor seja encontrada. Isto é, ser uma 'boa' explicação é um julgamento com respeito a um conjunto de hipóteses alternativas, que não contém necessariamente todos os γ's tal que $\Theta \vdash_L \gamma \rightarrow \beta$: na prática, a explicação absolutamente melhor não existe.

Se a relação ternária $\Theta \vdash_L \alpha \rightsquigarrow \beta$ é usada para expressar a relação abdutiva 'na teoria Θ, α é uma boa razão para β', a relação de preferência é denotada pelo símbolo $\sqsubseteq$ e a função *Cand* aplicada à teoria Θ e a um *explanandum* β retorna o conjunto de candidatos alternativos que constituem o teste pretendido para qualquer explicação aceitável de β em Θ, então define-se a relação de abdução da seguinte forma generalizada:

($\star$) $\Theta \vdash_L \alpha \rightsquigarrow \beta$ se $\Theta \vdash_L \alpha \rightarrow \beta$ e para todo $\gamma \in Cand(\Theta, \beta)$, se $\gamma \sqsubseteq \alpha$, então $\alpha \sqsubseteq \gamma$.

O caso geral ($\diamond$) é a instância de ($\star$), onde:

$$\alpha \sqsubseteq \gamma \equiv_{def} \vdash_L \gamma \rightarrow \alpha$$

$$Cand(\Theta, \beta) \equiv_{def} \{\gamma : \Theta \vdash_L \gamma \rightarrow \beta\}$$

Note-se que ($\star$) não requer que uma explicação abdutiva α pertença a $Cand(\Theta, \beta)$. Isto é, α não é um elemento minimal de $Cand(\Theta, \beta)$, mas um limite inferior.

Qualquer instância de ($\star$) possui a seguinte propriedade trivial: se $\Theta \vdash_L \alpha \rightsquigarrow \beta$ então $\Theta \vdash_L \alpha \rightarrow \beta$. Isto é, a relação $\rightsquigarrow$ é um subconjunto da implicação material.

Obviamente, não seria qualquer definição de $\sqsubseteq$ e $Cand$ que seria aceitável como uma significativa caracterização de abdução. Qualquer caracterização deve satisfazer algumas propriedades essenciais, em particular: se $\alpha \in Cand(\Theta, \beta)$ então $\Theta \vdash_L \alpha \rightarrow \beta$.

A seguir em [MP95b], assumindo que os requisitos básicos são atendidos, outras propriedades da relação abdutiva são estudadas, relacionando cada uma delas com propriedades correspondentes dos parâmetros e estabelecendo quais devem se dar e quais não correspondem a qualquer noção intuitiva. A tabela 3.1, proposta em [MP95b], mapeia propriedades razoáveis da relação abdutiva (coluna esquerda) em condições suficientes na relação de ordem e na função $Cand$ (coluna direita), garantindo sua validade. Nem todas as condições do lado direito são necessárias para que a propriedade correspondente se verifique.

Segundo Mayer e Pirri, as duas primeiras propriedades são facilmente obtidas e dizem respeito à relação entre conseqüência lógica e relação abdutiva. Posteriormente são definidas algumas propriedades do comportamento da relação abdutiva com respeito a conectivos lógicos. Finalmente algumas propriedades estruturais são consideradas.

As propriedades são listadas seguindo uma ordem crescente no sentido de que as primeiras são requeridas por qualquer noção razoável de abdução, enquanto podem existir noções interessantes de abdução que não respeitem as últimas. Todas as propriedades na tabela são satisfeitas pelo caso geral ($\diamond$).

Muitas propriedades que geralmente são válidas para a conseqüência lógica são definitivamente inaceitáveis para a relação abdutiva. Em [MP95b] são mostrados alguns exmplos:

1. Reflexividade na forma:

$$\Theta \vdash_L \alpha \rightsquigarrow \alpha$$

Essa propriedade só é válida se $Cand(\Theta, \alpha)$ não contém nenhuma fórmula γ que seja estritamente preferível a α, isto é, tal que $\gamma \sqsubseteq \alpha$ e $\alpha \not\sqsubseteq \gamma$. Como contra-exemplo, considere a

Tabela 3.1: Propriedades da relação abdutiva

$\Theta \vdash_L \beta$ sse $\Theta \vdash_L \top \rightsquigarrow \beta$	Para qualquer α, $\top \sqsubseteq \alpha$
$\dfrac{\alpha \sqsubseteq \gamma; \Theta \vdash_L \alpha \rightarrow \beta; \Theta \vdash_L \gamma \rightsquigarrow \beta}{\Theta \vdash_L \alpha \rightsquigarrow \beta}$	A relação $\alpha \sqsubseteq \beta$ é transitiva
Conjunção-direita: $\dfrac{\Theta \vdash_L \alpha \rightsquigarrow \beta; \Theta \vdash_L \alpha \rightsquigarrow \gamma}{\Theta \vdash_L \alpha \rightsquigarrow \beta \wedge \gamma}$	$Cand(\Theta, (\alpha \wedge \beta)) \subseteq Cand(\Theta, \alpha) \cup Cand(\Theta, \beta)$
Disjunção-direita: $\dfrac{\Theta \vdash_L \alpha \rightsquigarrow \beta; \Theta \vdash_L \alpha \rightsquigarrow \gamma}{\Theta \vdash_L \alpha \rightsquigarrow \beta \vee \gamma}$	$Cand(\Theta, (\alpha \vee \beta)) \subseteq Cand(\Theta, \alpha) \cup Cand(\Theta, \beta)$
Implicação-direita: $\dfrac{\Theta \vdash_L \alpha \wedge \beta \rightsquigarrow \gamma}{\Theta \vdash_L \alpha \rightsquigarrow \beta \rightarrow \gamma}$	a. Se $\varphi \in Cand(\Theta, \beta \rightarrow \gamma)$, então $\varphi \wedge \beta \in Cand(\Theta, \gamma)$ b. $\varphi \sqsubseteq \alpha$ sse $(\varphi \wedge \beta) \sqsubseteq (\alpha \wedge \beta)$
Modus Ponens: $\dfrac{\Theta \vdash_L \alpha \rightsquigarrow \beta \rightsquigarrow \gamma; \Theta \vdash_L \alpha \rightarrow \beta}{\Theta \vdash_L \beta \rightsquigarrow \gamma}$	$Cand(\Theta, \gamma) \subseteq Cand(\Theta, \beta \rightarrow \gamma)$
Corte: $\dfrac{\Theta \vdash_L \alpha; \Theta, \alpha \vdash_L \beta \rightsquigarrow \gamma}{\Theta \vdash_L \beta \rightsquigarrow \gamma}$	Se $\Theta \vdash_L \alpha$, então $Cand(\Theta, \gamma) \subseteq Cand(\Theta \cup \{\alpha\}, \gamma)$
Semi-compactação: Se $\Theta \vdash_L \alpha \rightsquigarrow \beta$, então para algum conjunto finito $\Delta \subseteq \Theta, \Delta \vdash_L \alpha \rightsquigarrow \beta$	Se $\Delta \subseteq \Theta$, então $Cand(\Delta, \beta) \subseteq Cand(\Theta, \beta)$
E-monotonicidade: $\dfrac{\Theta \vdash_L \alpha; \Theta \vdash_L \beta \rightsquigarrow \gamma}{\Theta, \alpha \vdash_L \beta \rightsquigarrow \gamma}$	Se $\Theta \vdash_L \alpha$, então $Cand(\Theta \cup \{\alpha\}, \gamma) \subseteq Cand(\Theta, \gamma)$
Equivalência lógica à esquerda: $\dfrac{\vdash_L \alpha \equiv \beta; \Theta \vdash_L \alpha \rightsquigarrow \gamma}{\Theta \vdash_L \beta \rightsquigarrow \gamma}$	Se $\vdash_L \alpha \equiv \beta$, então $\varphi \sqsubseteq \beta$ sse $\varphi \sqsubseteq \alpha$ e $\beta \sqsubseteq \varphi$ sse $\alpha \sqsubseteq \varphi$

lógica proposicional clássica, $\Theta = \{p \rightarrow q\}$ e $\alpha = p \wedge q$. Sempre que $Cand(\Theta, \alpha)$ contém p e $p \sqsubseteq p \wedge q$ - que não são condições fortes - a reflexividade não é válida. De fato, $p \wedge q \not\sqsubseteq q$, porque $\not\vdash_L q \rightarrow p \wedge q$.

2. Transitividade na forma:

$$\frac{\Theta \vdash_L \alpha \rightsquigarrow \beta; \Theta \vdash_L \beta \rightsquigarrow \gamma}{\Theta \vdash_L \alpha \rightsquigarrow \gamma}$$

   Um contra-exemplo é dado por $\Theta = \{p \rightarrow q, r \rightarrow s, p \rightarrow t, q \wedge s \rightarrow t\}, \alpha = p \wedge r, \beta = q \wedge s, \gamma = t$. A transitividade não é válida pois $p \in Cand(\Theta, t)$ e $p \sqsubseteq (p \wedge r)$.

3. Enfraquecimento à direita, na forma:

$$\frac{\Theta \vdash_L \alpha \rightsquigarrow \beta; \vdash_L \beta \rightarrow \gamma}{\Theta \vdash_L \alpha \rightsquigarrow \gamma}$$

   Para essa propriedade valer seria o caso em que se $\vdash_L \beta \rightarrow \gamma$, então $Cand(\Theta, \gamma) \subseteq Cand(\Theta, \beta)$. Seja agora $\Theta = \emptyset, \alpha = \beta = \perp, \gamma = p$. Se $p \in Cand(\Theta, p)$ e $p \sqsubseteq \perp$, enfraquecimento à direita não é válido; de fato $p \notin Cand(\Theta, \perp)$ porque $\Theta \not\vdash_L p \rightarrow \perp$ e $\perp \sqsubseteq q$ porque $\not\vdash_L q \rightarrow \perp$.

4. Monotonicidade, na forma

$$\frac{\Theta \vdash_L \alpha \rightsquigarrow \beta}{\Theta, \gamma \vdash_L \alpha \rightsquigarrow \beta}$$

   Para essa propriedade valer teria que ser o caso em que $Cand((\Theta \cup \{\gamma\}), \beta) \subseteq Cand(\Theta, \beta)$. Como contra-exemplo, seja $\Theta = \emptyset, \alpha = \beta = \gamma = p$, se $p \in Cand(\Theta, p)$ e $\top \in Cand(\{p\}, p)$ e $\top \sqsubseteq p$, a monotonicidade não vale.

5. Corte, na forma:

$$\frac{\Theta \vdash_L \alpha \rightsquigarrow \beta; \Theta, \beta \vdash_L \gamma \rightsquigarrow \sigma}{\Theta, \alpha \vdash_L \gamma \rightsquigarrow \sigma}$$

   Essa propriedade é válida quando $Cand((\Theta \cup \{\alpha\}), \sigma) \subseteq Cand(\Theta \cup \{\beta\}, \sigma)$; isso é razoável somente se $\alpha = \top$ e $\Theta \vdash_L \beta$, e esse caso corresponde à forma aceitável do corte definido na tabela 3.1. Como contra-exemplo da forma geral, considere $\Theta = \{p \wedge q \rightarrow r, p \wedge s \rightarrow t, r \wedge s \wedge u \rightarrow t\}, \alpha = p \wedge q, \beta = r, \gamma = s \wedge u, \sigma = t$. Então $\Theta \vdash_L p \wedge q \rightsquigarrow r$ e $\Theta, rvdash_L s \wedge u \rightsquigarrow t$. Mas se $s \in Cand(\Theta, t)$ e $s \sqsubseteq s \wedge u$, então$\Theta, p \wedge q \not\vdash_L s \wedge u \rightsquigarrow t$.

Como dito em [MP95b], o conjunto de propriedades considerado aqui não tem a intenção de ser exaustivo (completo). De acordo com Mayer e Pirri, nem sempre é possível encontrar uma solução para o caso geral da abdução. Essa não é uma situação desejável. A propriedade seguinte relaciona a existência de soluções para os aspectos da relação de ordem.

Propriedade 3.2 *Se $\sqsubseteq$ é uma relação de ordem bem fundada, então para qualquer Θ e β existe um α tal que $\Theta \vdash_L \alpha \rightsquigarrow \beta$.*

Na verdade, para qualquer Θ e β existe α tal que $\Theta \vdash_L \alpha \rightarrow \beta$; se não é o caso que $\Theta \vdash_L \alpha \rightsquigarrow \beta$, então $Cand(\Theta, \beta)$ é não vazio; se $\sqsubseteq$ é bem fundada, então $Cand(\Theta, \beta)$ tem um elemento minimal φ, logo $\Theta \vdash_L \varphi \rightsquigarrow \beta$.

Na sentença de condições gerais que garantem a decidibilidade da relação de abdução, a seguinte definição é utilizada:

Definição 3.28 *Um conjunto é finitamente gerado se existe um procedimento efetivo que gera todos os seus elementos.*

Propriedade 3.3 *Se, para todo Θ, α e β:*

1. *$\sqsubseteq$ é uma relação de ordem decidível;*
2. *o conjunto $Cand(\Theta, \beta)$ é decidível;*
3. *ou o conjunto $\{\gamma : \gamma \sqsubseteq \alpha\}$ é finitamente gerado ou $Cand(\Theta, \beta)$ é finitamente gerado, e*
4. *a lógica L é decidível,*

então a relação abdutiva $\Theta \vdash_L \alpha \rightsquigarrow \beta$ é decidível.

Na verdade, determinar se $\Theta \vdash_L \alpha \rightsquigarrow \beta$ resulta em:

- checar se $\Theta \vdash_L \alpha \rightarrow \beta$, que é decidível pela condição 4.
- se $\{\gamma : \gamma \sqsubseteq \alpha\}$ é finitamente gerado, gerar todos os seus elementos γ e checar se algum deles está em $Cand(\Theta, \beta)$, então $\alpha \sqsubseteq \gamma$; esse procedimento termina porque $\{\gamma : \gamma \sqsubseteq \alpha\}$ é finito, $Cand(\Theta, \beta)$ é decidível e $\sqsubseteq$ é decidível;

- caso contrário, se $Cand(\Theta, \beta)$ é finitamente gerado, gerar todos os seus elementos γ e checar que nunca acontece o caso em que $\gamma \sqsubseteq \alpha$ e $\alpha \not\sqsubseteq \gamma$. Analogamente ao caso anterior, esse procedimento termina.

Apesar das condições na propriedade 3.3 parecerem muito fortes, elas são naturalmente satisfeitas em muitos casos não triviais de L decidíveis.

O caso de lógicas semi-decidíveis é mais difícil de ser tratado. Uma propriedade análoga à propriedade 3.3 é definida [MP95b]:

Propriedade 3.4 *Se, para todo Θ, α e β:*

1. *$\sqsubseteq$ é uma relação de ordem decidível;*
2. *o conjunto $Cand(\Theta, \beta)$ é decidível;*
3. *ou o conjunto $\{\gamma : \gamma \sqsubseteq \alpha\}$ é finitamente gerado ou $Cand(\Theta, \beta)$ é finitamente gerado, e*
4. *a lógica L é semi-decidível,*

então a relação abdutiva $\Theta \vdash_L \alpha \rightsquigarrow \beta$ é semi-decidível.

As mesmas três primeiras condições garantindo a decidibilidade da abdução em lógicas decidíveis, são necessárias para garantir a semi-decidibilidade da abdução em lógicas semi-decidíveis.

Assumindo que $\sqsubseteq$ é não decidível, então, comparar duas soluções em potencial α e γ para um problema abdutivo pode claramente resultar em uma resposta indefinida para $\Theta \vdash_L \alpha \rightsquigarrow \beta$, mesmo quando $\Theta \vdash_L \alpha \rightsquigarrow \beta$ é verdade. Se o segundo requesito é somente semi-decidível, então checar se $\gamma \in Cand(\Theta, \beta)$ para algum $\gamma \notin Cand(\Theta, \beta)$ pode impedir que a minimalidde de α seja positivamente testada. Finalmente, se a terceira condição é válida, não minimalidade é semi-decidível (em qualquer linguagem enumerável) mas minimalidade não é: se o conjunto $\{\gamma : \gamma \sqsubseteq \alpha\} \cap Cand(\Theta, \beta)$ não é finitamente gerado - nenhum processo que gera seus elementos termina - então testar um α minimal para minimalidade pode claramente iniciar um processo não-terminal.

Considerações sobre o método

A abordagem mostra-se bastante interessante já que pretende uma caracterização lógica parametrizada da abdução. Desse modo, tal caracterização pode ser utilizada de maneira generalizada, sem estar, necessariamente, dependente de métodos ou lógicas específicas. Nossa própria abordagem da abdução via Sistemas Dedutivos Rotulados fará uso dessa definição parametrizada.

Além dessa interessante contribuição, o estudo das propriedades aplicáveis ou não ao raciocínio abdutivo (obviamente dependente de características dos parâmetros utilizados), também nos parece um importante aspecto a ser investigado, ainda que estejam muito fracamente definidas no artigo em questão.

Como colocado por Mayer e Pirri, o objetivo final desse trabalho é uma especificação completa de um sistema de prova para o raciocínio abdutivo. As considerações vistas mostram que há casos não triviais onde essa tarefa pode ser aperfeiçoada e um cálculo pode ser definido, tratando a expressão $\Theta \vdash_L \alpha \rightsquigarrow \beta$ como um objeto sintático, no mesmo estilo de um seqüente. A formulação de um sistema de inferência razoável e o estudo de suas propriedades é o objetivo principal de futuras pesquisas, assim como o desenvolvimento de um exemplo concreto e a análise de métodos computacionais particulares a fim de verificar que propriedades lógicas eles disfrutam.

Além disso, as variações na lógica a ser utilizada precisam ser investigadas em maior detalhe.

É importante observar que o critério de preferência como é tratado em [MP95b] está bastante condizente com a nossa proposta, já que no artigo é colocado que somente a implicação material não é suficiente para capturar a idéia de causalidade, e uma definição mais exata da abdução deve impor condições externas adicionais (meta-lógicas) às explicações, com o objetivo de selecionar boas razões para um *explanandum* β entre as fórmulas que logicamente implicam β na teoria de *background*. No entanto, não é explicitado no artigo em questão como se deve lidar com esse critério meta-lógico. No Capítulo 5, onde detalharemos nossa proposta, tentaremos ir um pouco mais além, especificando com maiores detalhes a utilização das chamadas condições do meta-nível.

3.4 Outras abordagens

Nesta seção descreveremos rapidamente outros trabalhos em abdução que não tiveram tanta influência na nossa proposta, mas que de algum modo contribuiram para o aprofundamento de conceitos importantes, e que, cada um com sua abordagem específica, caracterizam os diversos ramos que a pesquisa em abdução vem tomando.

3.4.1 Uma regra de inferência para a geração de hipóteses

Em [DdC92] Robert Demolombe e Luis Fariñas del Cerro propõem uma regra de inferência e uma estratégia associada a essa regra projetadas para computar explicações minimais eficientemente.

O conceito de explicação é definido em função de 'perguntas' (*queries*): para um dado banco de dados DB, e uma dada pergunta Q que não é derivável a partir de DB, uma explicação para Q é o conjunto de hipóteses X tal que $X \rightarrow Q$ é derivável a partir de DB e X é tão geral quanto possível.

No referido artigo a linguagem de primeira ordem é cónsiderada e as fórmulas devem estar sempre em forma de cláusulas. O problema abdutivo tratado pelos autores deve gerar como solução as cláusulas contendo instâncias de um dado literal L (a pergunta). A negação dos outros literais contidos em tais cláusulas são hipóteses que permitem a derivação de L.

A regra de inferência apresentada é chamada de L-inferência e foi projetada de modo a derivar cláusulas do tipo discutido acima. A L-estratégia associada à L-inferência considera um algoritmo de saturação com o objetivo de excluir do conjunto de respostas obtido as tautologias e as cláusulas não minimais.

A completude e a corretude da L-inferência são o principal resultado do artigo, segundo os autores. Além disso, a completude da L-estratégia também é mostrada.

3.4.2 Abdução proposicional para a lógica modal

Em [MP95a] Mayer e Pirri consideram o problema de executar a abdução em lógica modal. Esse trabalho é uma extensão de [MP93], dos mesmos autores (veja seção 3.3.2), e aplica a mesma metodologia para resolver a abdução modal.

A proposta faz uso do conceito de tableau modal e regras de expansão modal são definidas. Em seguida é especificado um méodo de construção de explicações abdutivas para os sistemas K, D, T e $S4$ que são corretos e completos. Além disso são feitas considerações com respeito à minimalidade das hipóteses.

Segundo Mayer e Pirri, a construção de hipóteses abdutivas está estritamente relacionada com as regras de expansão para as lógicas modais de uma maneira modular que torna possível modificações locais. Desse modo o método proposto é considerado geral no sentido de que pode ser adaptado para qualquer lógica proposicional modal para a qual exista um tableau analítico.

Além disso os autores argumentam que a expansão para a lógica modal de primeira ordem pode ser feita de maneira direta.

3.4.3 Uma organização de conhecimento para a abdução

Em [Add88] T. R. Addis examina os três tipos de inferência (dedução, indução e abdução) sob o ponto de vista da organização de conhecimentos dentro de programas de IA. Addis argumenta que que a única inferência suportada pelos computadores é a dedução, mas que a abdução, quando uma organização particular é adotada, pode ser explicada em termos de dedução.

Addis defende que a 'inferência' (de modo geral) é o que nos permite interpretar fatos e chegar a hipóteses, e que sem ela não existiriam novos conhecimentos. Com essa idéia em mente, ele ressalta a importância da utilização dos três tipos de inferência em conjunto.

No referido artigo, a dedução é definida como normalmente, salientando que os sistemas dedutivos são fechados em relação a novos conhecimentos e que a dedução é a única forma de inferência que pode ser realizada por computador e a única que pode ser formalmente justificada.

A abdução é definida por Addis como um processo de criação de novas hipóteses quando se tem fenômenos 'embaraçosos', e, segundo ele, abdução e indução funcionam conjuntamente e são normalmente chamadas simplesmente de indução. O propósito da 'indução' (no sentido mais amplo) seria prover suportes para as hipóteses, esse suporte seria usualmente alcançado pela comparação das conseqüências dedutivas com os fatos observados.

O processo de indução/abdução é definido, então, como a busca de 'premissas suprimidas' que expliquem em termos dedutivos os fatos observados. Essas 'premissas suprimidas' são também chamadas no artigo de 'generalizações'. Addis defende que existem mecanismos que usam instâncias para derivar generalizações que estão abertas a serem testadas. Esses mecanismos seguem um ciclo de abdução, dedução e indução que garante a formação de hipóteses viáveis.

Finalizando, Addis coloca que o processo de generalização a partir de observações deve gerar uma hipótese que seria uma proposição que abrange fatos semelhantes pela aplicação da dedução. Uma melhor hipótese seria uma proposição mais simples que abrange mais fatos com maior precisão.

3.4.4 Heurísticas cognitivamente plausíveis para alcançar a complexidade da abdução

Em [Fis92], Olivier Fischer aborda a abdução considerando três linhas de trabalho: técnicas baseadas em conhecimento; a análise formal da complexidade computacional do raciocínio abdutivo; e o estudo de protocolos utilizados por especialistas humanos para executar uma tarefa abdutiva.

A abdução é considerada como a inferência de uma hipótese que melhor explique um conjunto de dados D. O autor defende que geralmente uma boa hipótese abdutiva não está diretamente disponível e deve ser construída a partir da composição de hipóteses elementares, onde cada hipótese elementar seria uma hipótese abdutiva plausível para explicar parte do conjunto D.

Fischer chama a atenção para a falta de consenso na caracterização da melhor hipótese, mas considera razoável que uma hipótese H_1 seja melhor que outra H_2 se H_1 explica mais dados do conjunto D que H_2, e H_1 é mais plausível que H_2, e H_1 é menos redundante que H_2. No entanto, no resumo não fica claro o que seria plausível e redundante.

Quanto à análise computacional da abdução, o autor coloca que a complexidade de alcançar uma explicação abdutiva é dependente das propriedades apresentadas pelas hipóteses elementares. As propriedades analisadas por Fischer são a dependência entre as hipóteses, a monotonicidade do poder de explicação das hipóteses, o efeito de cancelamento entre os processos causais associados a diferentes hipóteses, e a existência de uma ordem de plausibilidade das hipóteses serem verdade.

Fischer argumenta ainda que, embora dependente das propriedades acima, a abdução é em geral um problema NP-completo, considerando-se a complexidade de computar os dados explicados pelas hipóteses.

Com respeito à análise dos protocolos dos especialistas para executar a abdução, Fischer tem como principal motivação o fato de que os humanos conseguem executar atividades abdutivas com um sucesso relativamente grande. Daí o autor acreditar na existência de heurísticas que tornem a abdução uma tarefa tratável. Essas heurísticas seriam dependentes da área de aplicação do problema e deveriam ser identificadas para cada caso.

Seguindo essas idéias, Fischer propõe dois níveis para solucionar um problema abdutivo: no primeiro nível devem existir processos para geração e teste de hipóteses abdutivas seguindo um modelo abstrato, independente do domínio do problema; o segundo nível deve conter processos de geração, seleção e testes de hipóteses abdutivas concernentes ao problema em questão. Nesse último nível estariam presentes as heurísticas identificadas para cada domínio.

O autor implementou duas heurísticas cognitivamente plausíveis aplicadas à área de identificação de anticorpos para transfusão de sangue que foram experimentadas com sucesso. Além disso, o trabalho foi generalizado para a solução de problemas abdutivos independentemente do domínio de aplicação.

3.4.5 As regras de abdução de Zadrozny

Wlodek Zadrozny em [Zad94] propõe uma teoria de abdução baseada na idéia de que a abdução é um tipo de raciocínio separado da indução e dedução, e que é possível formular regras abdutivas que não estão necessariamente restritas ao reverso do *modus ponens*.

A idéia é que os três tipos de inferência podem ser entendidos em termos de simetria, isto é, o tipo da relação que ela preserva. Então uma regra de inferência é dedutiva se ela preserva a verdade das consequências de uma teoria; é indutiva se preserva a falsidade de uma teoria; é abdutiva se preserva conjuntos de explicações.

Assim, uma regra de abdução é definida informalmente como alguma coisa que produz uma nova explicação a partir de outras explicações; essa regra é considerada válida se a explicação produzida estiver de acordo com a semântica em questão.

Zadrozny define o que seria um sistema de explicação:

Definição 3.29 *Um sistema de explicação é uma quíntupla* $(P, e-, F, T_E, Mods(T_E))$ *onde:*

- P *é um conjunto de explicações com uma ordem parcial associada a ele;*
- F *é uma coleção de fórmulas bem formadas, em uma linguagem formal, que o sistema tentará explicar;*
- $e-$ *é uma relação de explicação sobre* $P \times F$*;*
- T_E *é a teoria de explicação que é formulada em alguma meta-linguagem formal* $L_M(e-)$ *tal como teoria dos conjuntos ou lógica de alta ordem, com* $e-$ *como um predicado de dois argumentos;*
- $Mods(T_E)$ *contém modelos da teoria* T_E *que são completamente determinados pelos quatro elementos acima (possivelmente acrescidos de alguma restrição).*

As regras de abdução do sistema são definidas da seguinte forma:

Definição 3.30 *Dada uma linguagem $L_M(e-)$, uma regra de abdução é uma regra do tipo:*

$$\frac{\ldots \quad \ldots}{\ldots \quad e- \quad \ldots}$$

permitindo que se tire conclusões sobre explicações.

Segundo Zadrozny, os exemplos mais simples de sistemas de explicação são aqueles em que o conceito de explicação está relacionado com a idéia de prova; a explicação de uma fórmula consiste de um conjunto minimal de literais a partir do qual a fórmula é provada. Nesse caso, T_E seria algo do tipo:

$$l\ e - t \text{ se e somente se } l \vdash t \text{ e } l \text{ é minimal}$$

O autor argumenta que para diferentes relações de provabilidade, como por exemplo na lógica relevante, diferentes teorias podem ser obtidas. Partindo desse princípio, Zadrozny sugere uma série de regras abdutivas que podem ser utilizadas em diferentes domínios. Vejamos, a seguir, alguns exemplos de regras abdutivas propostas em [Zad94]:

- A seguinte regra permite a explicação da conjunção de literais, onde os literais já foram explicados

$$\frac{u\ e - X, w\ e - Y}{u \cup w\ e - X \cup Y}$$

- Uma regra de corte para a abdução

$$\frac{l\ e - X \leftarrow Y, l'\ e - Y \leftarrow Z}{l \cup l'\ e - X \leftarrow Z}$$

- Outra versão do corte para a abdução

$$\frac{l\ e - X \leftarrow Y, l'\ e - Y}{l \cup l'\ e - X}$$

- Regra que lida com conjunções no corpo das cláusulas

$$\frac{l\ e - X, l'\ e - Y}{l \cup l'\ e - X\&Y}$$

Além de propor regras como as mostradas acima, o autor caracteriza a validade sintática de regras abdutivas, relacionando-a com a equivalência entre as explicações envolvidas. Esse critério de validade - a invariância das explicações - fornece uma base formal para avaliar as regras de inferência abdutiva.

Ainda em [Zad94] é formulada uma teoria dos modelos para abdução e é feita a ligação com uma teoria da prova. As regras abdutivas propostas são utilizadas para descrever tarefas como resolução de ambigüidade e resolução de anáforas no processamento de linguagem natural, assim como o diagnóstico abdutivo.

Em [Zad93] as mesmas idéias são desenvolvidas com o acréscimo do conceito de coerência ou coesão. Zadrozny argumenta que algumas tarefas bastante diferentes podem ser modeladas por uma regra abdutiva denominada de $coh(2)$. Essa regra pode ser informalmente descrita como sendo possível partir da conclusão para as premissas se as premissas são coerentes, isto é, se existe alguma ligação semântica entre elas.

A justificativa para a utilização desse conceito de coerência é baseada no fato de que geralmente as explicações abdutivas são baseadas na proximidade dos argumentos. Assim, Zadrozny propõe o esquema abaixo argumentando que essa regra pode ser usada em diversos tipos de raciocínio baseados em proximidade:

$$\frac{l_1 e - t_1, l_2\ e - t_2, c(l_1, l_2)}{l_1 + l_2\ e - t_1.t_2}$$

onde o operador + define uma combinação entre as explicações e o operador . define uma combinação entre os termos.

Ainda em [Zad93] é discutida uma teoria dos modelos para explicações e a validade das regras de coerência. Tudo isso leva a alguns resultados matemáticos com relação à corretude do efeito de algumas estratégias de uso de regras de coerência e à incorretude do efeito de outras estratégias para uso das mesmas regras.

3.4.6 Raciocínio abdutivo em lógica tri-valorada

Em [MD94], Philippe Mathieu e Jean-Paul Delahaye propõem um método abdutivo para computar perguntas inteligentes, no contexto de sistemas especialistas, a fim de atingir determinado objetivo.

O que é tratado nesse artigo como perguntas a serem feitas ao usuário corresponde a nossas hipóteses abdutivas; e o que é chamado de objetivo corresponde a nosso fato observado a ser explicado.

No artigo é proposta a utilização de uma lógica tri-valorada (*true, false, unknown*), já que se trabalhará com fatos desconhecidos, e o próprio usuário (o especialista, no caso) pode responder 'não sei' a uma pergunta feita pelo sistema.

O conceito de 'memória de trabalho' é estabelecido (WM -*work memory*) com o intuito de guardar as novas informações recebidas ou deduzidas. O conteúdo da WM deve estar condizente com o princípio da não contradição: dois literais opostos não podem estar ambos na WM; se for o caso, uma contradição é detectada e aquele caminho para a busca da solução é descartado.

O sistema utiliza o conceito de '*mixed chaning*': utiliza um algoritmo do tipo *forward chaining* para deduzir novos fatos e um algoritmo de *backward chaining* para computar perguntas a serem feitas ao usuário. Segundo os autores, é a relevância das perguntas escolhidas (as hipóteses abdutivas) que levam o usuário a qualificar o sistema como inteligente. Portanto, a idéia central do sistema proposto por Mathieu e Delahaye é computar corretamente questões a serem feitas ao usuário.

O sistema apresenta, então, um algoritmo para dedução e outro para computar questões. As questões são convencionadas como sendo sempre literais. A fim de definir o que seria uma questão útil, são introduzidas duas noções:

Definição 3.31 *Seja Kb uma base de conhecimento, L uma conjunção de literais básicos desconhecidos, e x um literal. Dizemos que L é uma* **lista de perguntas inteligentes** *com respeito a x se e somente se:*

- *$Kb \cup L$ é consistente;*
- *$Kb \cup L \models_T x \vee \neg x$, ($L \rightarrow x$ ou $L \rightarrow \neg x$ é uma conseqüência tri-valorada de Kb);*
- *Não existe nenhuma lista de perguntas inteligentes L' com respeito a x tal que $L' \subsetneq L$.*

Definição 3.32 *Seja Kb uma base de conhecimento, L uma conjunção de literais básicos desconhecidos, e x um literal. Dizemos que L é uma* **lista de perguntas relevantes** *com respeito a x se e somente se existe um subconjunto R de Kb para o qual:*

- *$R \cup L$ é consistente;*
- *$R \cup L \models_T x \vee \neg x$, ($L \rightarrow x$ ou $L \rightarrow \neg x$ é uma conseqüência tri-valorada de R);*

- *Não existe nenhuma lista de perguntas inteligentes L' com respeito a x tal que $L' \subseteq L$.*

Pode-se dizer que a computação de perguntas relevantes é uma estratégia abdutiva fraca e a computação de perguntas inteligentes é uma estratégia dedutiva forte.

No artigo são propostos dois algoritmos para a computação, respectivamente, de listas de perguntas relevantes e de listas de perguntas inteligentes. É mostrado que, mesmo para o caso de perguntas relevantes (que é um problema mais simples por considerar apenas parte da base de conhecimentos) o problema é NP-completo. Na computação das listas são considerados os conceitos de consistência com a WM, consistência com futuras deduções e minimalidade da lista.

Além da computação das perguntas, que é considerado pelos autores como o nível objeto do problema, é proposto um nível heurístico do problema: freqüentemente existem muitas listas de perguntas relevantes ou inteligentes, e deve-se escolher qual delas usar para fazer as perguntas ao usuário.

É interessante notar a similaridade do nível heurístico proposto nesse artigo com o passo de comparação dos candidatos proposto por nós na seção 5.4.3.

No nível heurístico, são propostas algumas heurísticas para a escolha da lista a ser usada para fazer as perguntas ao usuário. Duas abordagens são possíveis: (1) escolher uma lista durante a geração das mesmas ou (2) computar todas as listas e, então, escolher uma. Para a primeira abordagem, as seguintes heurísticas são propostas:

- Prioridade nas regras - se muitas regras têm a mesma conclusão, o usuário (especialista) deve atribuir uma prioridade a cada regra. O sistema irá examinar primeiro as regras com menor (maior) prioridade.
- Regras que têm menos premissas desconhecidas - é uma heurística dinâmica que supõe que quanto menos premissas desconhecidas uma regra tiver, maior a chance de se chegar a uma solução mais rapidamente.

Para a segunda abordagem, onde todas as listas possíveis são computadas, os autores propõem as seguintes heurísticas:

- Lista com menor número de perguntas - escolhe-se a menor lista e a primeira pergunta a ser feita é a correspondente ao primeiro literal da lista.

- Escolher o átomo que aparece com maior freqüência em todas as listas.
- Coeficientes de satisfatibilidade - nesse caso deverá estar associado a cada átomo um valor correspondendo à chance existente de se obter uma resposta *true* ou *false* quando o usuário é questionado sobre aquele fato.
- Método acumulativo - nesse caso a proposta é utilizar as heurísticas anteriores e generalizá-las conjuntamente a fim de obter um método acumulativo.

Os algoritmos propostos em [MD94] foram implementados no sistema BIVOUAC.

Dentre as abordagens estudadas, a proposta de Mathieu e Delahaye é a que mais se aproxima da nossa própria proposta, no que se refere ao meta-nível da abdução. As heurísticas propostas pelos autores têm o mesmo objetivo que os critérios de preferência que descreveremos no Capítulo 5: o de reduzir o número de hipóteses abdutivas geradas no nível objeto, selecionando aquelas mais plausíveis em função dos critérios (heurísticas) especificados.

3.4.7 Abdução rotulada e raciocínio relevante

Um tema bastante atual na área de IA e formalização do raciocínio é a tentativa de lidar com interpretação de linguagem natural sob o ponto de vista da abdução. Uma abordagem desse tema é feita por Dov Gabbay, Ruth Kempson e Jeremy Pitt em [GKP94]. Nesse artigo é explorada a modelagem de uma interação menos padrão entre um homem e um banco de dados. A proposta é equipar o banco de dados com a capacidade de entender perguntas em linguagem natural e respondê-las de um modo mais inteligente que o simples *yes/no*. Para isso, segundo os autores, são necessários dois componentes: primeiro o banco de dados deve possuir um módulo interativo capaz de analisar (*parsing*) a entrada em linguagem natural e traduzí-la para sua própria linguagem lógica. Em segundo lugar, o computador deve saber lidar com os padrões de pergunta/resposta humanos que estão muito mais para o raciocínio do senso comum não monotônico, que para linguagens comuns do tipo *query*.

A abdução é utilizada nessa proposta sob dois aspectos: primeiro, na interpretação de linguagem natural a abdução é utilizada para a formulação de escolhas quando a informação disponível é insuficiente (principalmente no que diz respeito à interpretação de pronomes); e depois, no segundo componente proposto pelos autores, onde o processo abdutivo entra mais ativamente na modelagem do raciocínio subjacente aos diálogos do tipo pergunta/resposta, que envolve respostas indiretas, mudanças de crenças, objetivos, escolhas, e outros.

Para exemplificar informalmente o tipo de problema que os autores desejam resolver, vejamos o seguinte exemplo tirado de [GKP94]:

Seja o seguinte banco de dados Δ_1:

$$\begin{aligned} \Delta_1 = \{ \quad & Holiday(x,y) if Available(y) \wedge Afford(x,y) \\ & Afford(x,y) \text{ if } Cost(y,z) \wedge Credit(x,z) \\ & Credit(x,y) \text{ if } Card\text{-}limit(x,y) \wedge y \leq z \\ & Sunnyplace(a_i), i = 1...3 \\ & Cost(a_1, 200) \\ & Cost(a_2, 250) \\ & Cost(a_3, 700) \\ & Sunnyplace(b) \\ & Cost(b, 500)\} \end{aligned}$$

Um cliente acessa o banco de dados inserindo seu cartão de crédito e digitando uma pergunta $(holiday, a_1, a_2, a_3)$. O banco de dados pode automaticamente obter do cartão o nome do cliente, x_0, e outros detalhes de crédito, e então faz a si mesmo as perguntas:

$$holiday(x_0, a_i), i = 1, 2, 3$$

Um banco de dados não inteligente vai simplesmente checar se cada uma das perguntas tem sucesso. Se sim, o banco de dados responde com os detalhes (preço, etc.), se não, ele responde com uma 'falha'. Um banco de dados inteligente deve ter como saída uma explicação no caso de falha, tal como:

You can't have a_1 *because it is not available*

ou

You can't have a_3 *because you cannot get credit*

Um banco de dados mais inteligente ainda deve reconhecer, por exemplo, que nenhum dos a_i's estão disponíveis, mas todos eles são lugares de sol e b é também um lugar de sol dentro da mesma faixa de preço. Então, o banco de dados deve perguntar ao usuário (seja qual for o modo em que ele se comunica):

Would you like (holiday,b)?

e informar os detalhes.

Os autores sugerem ainda que o usuário possa dialogar em linguagem natural, que pode transceder o simples sim/não. Por exemplo, uma forma natural de responder à pergunta acima seria:

b is expensive

O computador deveria entender o que o usuário quer dizer e deduzir (como um humano o faria) que a resposta à pergunta é 'não' e que qualquer férias custando mais que $500 não é um candidato. E assim por diante.

Para modelar diálogos como o descrito acima, deve-se reconstruir o raciocínio por trás desse rico processo de uma maneira o mais geral possível, aplicando-o tanto a decisões envolvidas no estabelecimento do significado de uma sentença como também às informações adicionais que são obtidas a partir dela.

Para atender esse objetivo, os autores adotam a estrutura do *LDS* (descrito no Capítulo 4 desta dissertação) para modelar o processo dinâmico de interpretação de linguagem natural, e descrevem como o banco de dados resultante é usado em processos mais gerais de raciocínio *on-line* [GKP94].

A forma que as sentenças são analisadas é assumindo que a cada palavra é assinalada uma fórmula bem formada de alguma lógica, com um rótulo que corresponde ao próprio ítem léxico, e escolhas abdutivas adicionais são feitas quando necessário. Com esse procedimento (que nos absteremos de descrever em detalhes), constrói-se um banco de dados rotulado expressando o significado da sentença de entrada.

A outra parte do problema de interpretação será considerar os possíveis efeitos inferenciais indiretos a que a interpretação obtida pode levar. Nessa fase é feita uma livre manipulação das escolhas abdutivas. Em muitos casos, os participantes de um diálogo respondem às questões de forma deliberadamente indireta. O efeito é sempre o mesmo: o ouvinte obtém mais informação do que obteria com uma resposta direta (supondo que estamos falando de participantes cooperativos).

Diante desses fatos, os autores utilizam a regra de suposições relevantes (*RRA - rule of relevant assumptions*) para tentar explicar as colocações indiretas e assim fazer deduções a partir delas. Essa regra, RRA, significa o princípio do menor enriquecimento relativo consistente para um procedimento de prova, onde:

1. Dado um procedimento de prova para encontrar respostas a perguntas do tipo $\Delta?Q$ (isto é, 'Q é derivável a partir de Δ?'), então se Δ é um banco de dados parcial e fazemos a pergunta $\Delta?Q$ e obtemos a entrada B, então tenta-se provar Q a partir de $\Delta \cup \{B\}$ e assume-se que qualquer informação $B_1, B_2, ...$ requerida para o sucesso da prova, deve ser adicionada a Δ juntamente com B, e $\Delta \cup \{B, B_1, B_2, ...\}$ deve ser consistente. Quando se fala de prova ou dedução, não nos restringimos à dedução monotônica. As regras de dedução utilizadas na tentativa de provar $\Delta \cup \{B\} \vdash Q$ podem ser anuláveis ou não monotônicas. Essas regras correspondem melhor ao raciocínio do senso comum. A natureza da lógica envolvida não foi especificada no artigo.

2. Se existem várias maneiras de adicionar $\{B_i\}$ chegando ao sucesso da prova, deve-se escolher aquela que envolve o menor esforço dedutivo para algum efeito inferencial.

Os dois ítens acima relacionam estreitamente a segunda fase proposta pelos autores com a abdução. Inclusive o algoritmo abdutivo usado como RRA é um algoritmo abdutivo proposicional que foi implementado em PROLOG pelos autores.

O segunto item, em particular, chama a atenção para o fato das várias possibilidades que acontecem na abdução. No entanto, nenhum mecanismo de escolha é especificado. Os autores chamam a atenção para a sua necessidade, e o colocam como trabalho futuro.

Essa proposta retrata muito bem a relação existente entre o raciocínio abdutivo e a interpretação de linguagem natural. Além disso, os autores argumentam que o *LDS* provê a fundamentação necessária dentro da qual as atividades de raciocínio envolvidas na interpretação de linguagem natural podem ser especificadas, relacionadas e absorvidas em um único sistema de raciocínio.

Capítulo 4

Os Sistemas Dedutivos Rotulados

Os sistemas dedutivos rotulados (do inglês *Labelled Deductive Systems - LDS*) foram propostos por Dov Gabbay com o intuito de fornecer um *framework* geral para a representação de lógicas diversas, onde os aspectos do meta-nível possam ser tratados lado a lado com os aspectos do nível objeto. Neste capítulo, descreveremos o *LDS*, como *framework* geral, e uma instância do mesmo, o sistema de dedução natural rotulada (do inglês *Labelled Natural Deduction - LND*). Finalizando, abordaremos a abdução rotulada como é vista por Dov Gabbay.

4.1 O por quê dos sistemas rotulados

Existe uma demanda crescente na ciência da computação, lingüística e filosofia para uma variedade de sistemas lógicos. Nos campos da IA, programação em lógica e ciência da computação teórica existe uma crescente necessidade de sistemas lógicos com significado semântico e representações algorítmicas que possam servir para aplicações diversas. Dessa maneira, muitas atividades de pesquisa têm sido devotadas à análise e alterações de velhas e novas lógicas.

Esse tipo de atividade levou a uma maior complexidade na definição de sistemas lógicos que passam então a lidar não só com conjuntos de fórmulas mas estruturas de fórmulas cada vez mais ricas. Com isso, surgem novos problemas que clamam por uma estrutura geral aprimorada na qual as novas lógicas emergentes das aplicações da ciência da computação possam ser representadas e investigadas.

Tradicionalmente, qualquer relação binária do tipo $\Delta \vdash T$, satisfazendo certas condições (reflexividade, monotonicidade e transitividade) é um sistema lógico. Tal relação deve ser representada matematicamente, e isso pode ser feito semanticamente ou algoritmicamente.

Ainda tradicionalmente, para representar uma lógica L, nós precisamos definir primeiro um conjunto de fórmulas bem formadas da lógica. Essa é a linguagem da lógica. Definimos então o conjunto de fórmulas atômicas, conectivos, quantificadores e o conjunto de fórmulas arbitrárias. Depois nós definimos matematicamente a noção de conseqüência, ou seja, para um dado conjunto de fórmulas Δ, e uma dada fórmula Q, nós definimos a relação de conseqüência $\Delta \vdash_L Q$, lendo-se "Q seque de Δ na lógica L".

A relação de conseqüência deve satisfazer as seguintes propriedades intuitivas:

- **Reflexividade**

$$\Delta \vdash Q \text{ se } Q \in \Delta$$

- **Monotonicidade**

$$\Delta \vdash Q \text{ implica } \Delta, \Delta' \vdash Q$$

- **Transitividade**

$$\Delta \vdash A; \Delta, A \vdash Q \text{ implica } \Delta \vdash Q$$

Então, uma lógica é obtida pela especificação de L e $\vdash$. Dois sistemas algorítmicos S_1 e S_2 que levam à mesma relação de conseqüência ($\vdash$) são considerados como a mesma lógica.

A noção acima foi essencialmente colocada por Tarski [Tar56] e é referida como a relação de conseqüência de Tarski. Essas propriedades pareciam ser minimais e as mais naturais para um sistema lógico possuir, dado que as principais aplicações da lógica eram em matemática e filosofia.

Essa noção foi generalizada por Gabbay em 1969 e em seguida por Scott em 1974 no sentido de que Q poderia ser um conjunto de fórmulas T. A relação básica, então, passa a ser do tipo $\Delta \vdash T$, satisfazendo as mesmas propriedades [Gab94].

As noções acima são monotônicas, porém, o crescente uso da lógica em IA levou ao surgimento de sistemas lógicos não monotônicos. Existem muitos sistemas desse tipo satisfazendo uma variedade de condições representadas de diversas formas diferentes. Todas essas diferentes representações levam a alguma noção de conseqüência, mas todas elas parecem estar de acordo somente com relação a uma forma restrita de reflexividade ($A \vdash A$). A diferença essencial entre essas lógicas (comumente chamadas de lógicas não monotônicas) e as lógicas tradicionais é o fato de que $\Delta \vdash A$ acontece no caso monotônico por causa de algum $\Delta_A \subseteq \Delta$, enquanto que no caso não monotônico todo o conjunto Δ é utilizado para derivar A. Então, se Δ é incrementado para um Δ', não há nenhuma alteração no caso monotônico, enquanto que pode acontecer alguma mudança no caso não monotônico.

No início dos anos 80, era essa a situação. Um grande número de sistemas eram geralmente aceitos como lógicas, sem uma teoria subjacente unificada, e muitas tinham semântica sem teoria da prova e outras tinham teoria da prova sem semântica [Gab94]. Apesar disso, quase todas eram baseadas em intuições corretas, de uma forma ou de outra. Claramente existia a necessidade de uma estrutura geral unificada que fosse sensível à variedade de características dos diversos sistemas.

Atualmente, por outro lado, a relativa importância de dedução automática está em crescimento, dada a sua larga aplicabilidade. Novos métodos de dedução automática têm sido desenvolvidos para lógicas não clássicas, e o princípio de resolução tem sido generalizado e modificado para ser aplicável a essas lógicas [Gab94]. Em geral, por causa do valor dessas lógicas para a ciência da computação teórica e para a IA, uma maior consciência dos aspectos computacionais de sistemas lógicos está se desenvolvendo e mais atenção tem sido devotada a representações de teoria da prova. Isso torna aparente que um aspecto chave no estudo de teoria da prova dessas lógicas é que uma pequena variação em um sistema de prova de uma lógica L_1 pode levar a outra lógica L_2 [Gab94].

Diante dessa proliferação de lógicas, Dov Gabbay começou a perguntar-se o que tornava uma lógica diferente da outra. A essa pergunta ele responde dizendo que a maioria das lógicas são baseadas em *modus ponens* e que as regras de quantificadores são normalmente as mesmas;

as diferenças entre uma lógica e outra são considerações do meta-nível na teoria da prova ou na semântica. Se pudermos encontrar um modo de representar sistemas lógicos onde aspectos do meta-nível possam residir lado a lado com aspectos do nível objeto, então podemos dizer que temos um *framework* geral para a representaçào de lógicas.

4.2 O *LDS* como um *framework* geral para representar lógicas

No contexto descrito acima, surgem os Sistemas Dedutivos Rotulados (do inglês *Labelled Deductive Systems - LDS*), propostos por Dov Gabbay [Gab94], com o objetivo de suprir tal necessidade de um *framework* geral para a representação de lógicas.

Formalmente, o *LDS* é definido da seguinte maneira:

Definição 4.1 (LDS) *O LDS é um sistema de prova caracterizado por uma tripla (A,L,M), onde:*

- *A é uma álgebra dos rótulos;*
- *L é uma linguagem lógica (conectivos, fórmulas bem formadas, etc.);*
- *M é uma disciplina de rotular fórmulas da lógica (a partir da álgebra A dos rótulos).*

Além disso, devem ser definidas as regras de dedução e uma maneira de propagar os rótulos através da aplicação das regras de dedução.

No *LDS*, a tradicional relação de conseqüência "$A_1, A_2, ..., A_n \vdash A$" entre fórmulas, é substituída por uma noção mais geral "$t_1 : A_1, t_2 : A_2, ..., t_n : A_n$", onde os t_i's são rótulos das fórmulas A_i's. Os rótulos carregam alguma informação do meta-nível em relação aos A_i's, e têm o papel de trazer essa informação (do meta-nível) para o nível objeto. Essas duas dimensões (nível objeto e meta-nível/fórmulas e rótulos) são trabalhadas de forma separada, porém harmoniosa e disciplinadamente. Desse modo, o sistema lógico é enriquecido sem adicionar complicações na linguagem lógica propriamente dita, que é mantida sem o acréscimo de outros conectivos.

Definido dessa maneira, o *LDS* busca atender aos seguintes objetivos [Gab94]:

- dispor de uma perspectiva unificadora que permita a 'fatoração' (em inglês, *factoring out*) das características de meta-nível daquelas essencialmente de nível objeto;

- manter as regras de inferência tão simples quanto possível, ao mesmo tempo em que dispomos de um cálculo de manuseie as características de meta-nível em separado e de forma tal que mantenha uma certa harmonia com o cálculo lógico;
- dispor de meios para a estruturação e a combinação de lógicas;
- fazer com que as suposições relevantes em uma dedução sejam apropriadamente trazidas à superfície (veja seção 4.4.2).

A idéia central, então, é que a unidade declarativa de um sistema lógico passa a ser uma fórmula rotulada, '$t : A$' (leia t rotula A). Sob tal perspectiva, um sistema lógico é visto não apenas como um cálculo de deduções lógicas sobre fórmulas, mas como uma combinação de um cálculo funcional sobre os rótulos e um cálculo lógico sobre as fórmulas.

A informação adicionada ao longo das fórmulas tem o papel, como já vimos, de trazer algumas características de meta-nível de volta para o nível objeto de forma disciplinada e tal que não complique demasiadamente a linguagem das fórmulas. O rótulo, com sua origem num domínio essencialmente procedural, constitui informação de natureza bem diferente daquela contida nas fórmulas (estas últimas consistindo essencialmente de declarações). De modo geral, podemos fazer uso de várias leituras para a unidade declarativa '$t : A$' dos sistemas dedutivos rotulados [Gab94]:

1. Um valor de confiabilidade difuso (um número real x, $0 \leq x \leq 1$). Essa leitura pode ser útil em sistemas especialistas que usam lógicas difusas.
2. A origem da fórmula (de onde vem a informação). Pode ser útil em sistemas de gerenciamento de bancos de dados complexos.
3. A prioridade da fórmula. Útil na modelagem de sistemas dedutivos com prioridade. Particularmente, na nossa proposta de formalizar o raciocínio abdutivo os rótulos terão um papel que pode ser visto como prioridade, a fim de direcionar a escolha de hipóteses abdutivas, como veremos no Capítulo 5.
4. A construção de prova da fórmula. Aqui a interpretação funcional de Curry-Howard se encarrega do resto, como veremos a seguir.

Em muitas ocasiões a apresentação de um elemento de informação da forma '$t : A$', onde t é o rótulo e A é a fórmula, é bastante intuitiva. Em geral, o uso de t como rótulo significa que t

representa uma certa informação necessária à modificação e/ou suplementação da informação que A representa, mas que não é do mesmo tipo da informação representada por A propriamente.

O seguinte exemplo, mostrado em [Gab94], ilustra a natureza da informação contida nos rótulos:

Exemplo 4.1 *Considere uma linguagem com o predicado* $VS900(x,t)$. *Esse predicado denota o vôo "Londres-Tóquio" da companhia "Virgin", onde t é a data do vôo e x é o nome de um indivíduo. Por exemplo,* $VS9000(Dov, 15.11.91)$ *pode ser colocado em um banco de dados, significando que Dov está com uma reserva para esse vôo, cujo embarque está previsto para 15.11.91.*

Se a companhia utiliza procedimentos de cancelamento ou outros tipos de procedimento como "overbooking", seria útil guardar informações adicionais do tipo:

- *tempo de reserva;*
- *se a reserva é de um grupo de pessoas ou individual;*
- *tipo da passagem;*
- *se a pessoa é ou não VIP; etc.*

Esse tipo de informação é de natureza diferente daquela contida no predicado principal e, por esse motivo, é mais conveniente mantê-la como uma anotação ou rótulo. Além disso, a manipulação dessa informação extra é de natureza diferente daquela do predicado.

4.3 Da interpretação funcional de Curry-Howard aos sistemas de dedução natural rotulada

O sistema de dedução natural rotulada (do inglês *Labelled Natural Deduction - LND*) [dQG92] é uma das instâncias dos sistemas rotulados propostos por Gabbay [Gab94], no qual se trabalha com a interpretação funcional de Curry-Howard, conhecida como fórmulas-como-tipos (do inglês *formulae-as-types*) [How80]. Ao contrário da concepção de Curry-Howard, que se restringe à lógica intuicionística, o *LND* se apresenta como um *framework* geral no qual o paradigma fórmula-como-tipos é substituído pelo paradigma rótulos-e-fórmulas, solucionando problemas anteriormente não resolvidos. Além disso, a interpretação funcional é estentida para outras lógicas (clássica, linear, relevante, etc.) [GdQ92].

Esta seção apresenta o estudo originalmente desenvolvido no artigo de Gabbay e Queiroz [GdQ92] onde é feita a extensão da interpretação funcional de Curry-Howard. Tal interpretação tem sido associada à interpretação intuicionística da lógica e da matemática. De fato, a interpretação de Curry-Howard, da maneira como foi originalmente formulada, é especialmente adequada para o cálculo de predicados intuicionístico de Heyting. Ela está fundamentada em uma certa harmonia entre, por um lado, um cálculo funcional sobre as expressões construídas a partir do registro dos passos de dedução (isto é, a informação procedural: os rótulos), e, por outro lado, um cálculo lógico sobre as fórmulas (isto é, a informação declarativa). A interpretação está intimamente associada com a descoberta, no contexto da teoria da funcionalidade de Curry, da correspondência entre os axiomas da implicação intuicionística e os esquemas de tipo dos chamados combinadores da lógica combinatória. Em vista de sua própria origem, e de sua associação com a interpretação intuicionística de conectivos lógicos *a la* Heyting [1930], onde a noção de prova era fundamental na definição do significado de proposições, ficou associado à interpretação de Curry o paradigma denominado fórmulas-como-tipos (*formulae-as-types*). O trabalho de Howard, que já em 1969 demonstrava a extensão do paradigma a outros conectivos lógicos (conjunção, disjunção, quantificadores) resultou no que hoje se chama de "interpretação funcional de Curry-Howard".

4.3.1 A teoria da funcionalidade de Curry

Conforme o exposto em [GdQ92], Schönfinkel, em 1924, definiu os "blocos básicos da lógica matemática", conhecidos como combinadores, através de regras de conversão. Tomando *aplicação* como a noção mais primitiva, cada combinador foi então caracterizado pela maneira em que sua aplicação a uma lista ordenada de objetos combinatórios operaria naquela lista, que objeto resultante seria obtido. Assim, escrevendo-se '$\rightarrow$' para abreviar 'converte para', variáveis $(x, y, z, ...)$ para denotar objetos combinatórios arbitrários, e '$app(x, y)$' para significar 'a aplicação de x a y', os combinadores seriam definidos como:

$$
\begin{aligned}
\mathsf{I}x &\rightarrow x \\
\mathsf{B}xyz &\rightarrow app(x, app(y, z)) \\
\mathsf{B}'xyz &\rightarrow app(y, app(x, z)) \\
\mathsf{C}xyz &\rightarrow app(app(x, z), y) \\
\mathsf{W}xy &\rightarrow app(app(x, y), y) \\
\mathsf{S}xyz &\rightarrow app(app(x, z), app(y, z)) \\
\mathsf{K}xy &\rightarrow x
\end{aligned}
$$

Esta foi essencialmente a caracterização original dos combinadores dada (independentemente) por Schönfinkel e Curry no final dos anos 20. Mais tarde, em seu artigo de 1934, entretanto, Curry quis caracterizar os combinadores pelo aspecto de suas 'funcionalidades' (isto é, dados objetos de determinados tipos, o combinador retornaria um objeto de certo tipo). Para fazer isso, Curry começou a atribuir 'tipos' aos respectivos objetos combinatórios. Obviamente, o princípio de aplicação permaneceria válido, apenas estendido com símbolos para os tipos dos objetos, da seguinte forma:

$$
(\text{APP}) \qquad \frac{y : \mathbf{A} \qquad x : \mathbf{A} \rightarrow \mathbf{B}}{app(x, y) : \mathbf{B}}
$$

significando: se y é do tipo $\mathbf{A}$ e x é do tipo $\mathbf{A} \rightarrow \mathbf{B}$ então a aplicação de x a y resulta em um objeto do tipo $\mathbf{B}$.

Assim, ao chegar numa conclusão que um objeto '$app(x, y)$' é do tipo '$\mathbf{B}$' poderíamos seguir a regra (APP) em direção oposta (da conclusão às premissas) e atribuir a x um tipo da forma '$\mathbf{A} \rightarrow \mathbf{B}$' para um certo '$\mathbf{A}$', e a y o tipo '$\mathbf{A}$' propriamente dito. Seguindo tal procedimento, poderia-se chegar na atribuição correta do tipo dos combinadores.

De modo a ver o procedimento funcionando, tomemos um dos combinadores, digamos 'B', e procuremos sua atribuição de tipo. Inicialmente temos que:

$$
\mathsf{B}xyz \rightarrow app(x, app(y, z))
$$

A partir desse ponto começamos a atribuir ao objeto resultante, ou seja '$app(x, app(y, z))$', um tipo, digamos '$\mathbf{B}$'. Daí,

$$
app(x, app(y, z)) : \mathbf{B}
$$

Dito isso, sigamos (APP) no sentido inverso: enquanto o tipo x deve ser da forma '$\mathbf{A} \rightarrow \mathbf{B}$' para um certo '$\mathbf{A}$', $app(y, z)$ deve ter este mesmo '$\mathbf{A}$' como seu tipo. Por isso, nesse estágio teremos:

$$
\begin{aligned}
x :&\ \mathbf{A} \rightarrow \mathbf{B} \\
app(y, z) :&\ \mathbf{A}
\end{aligned}
$$

Da última linha deduzimos que enquanto y deve ser do tipo '$\mathbf{C} \rightarrow \mathbf{A}$' para um certo '$\mathbf{C}$', z deve ter o mesmo '$\mathbf{C}$' como seu tipo. Nesse ponto identificamos o tipo de nossos objetos combinatórios arbitrários:

$$\begin{aligned} x &: \ \mathbf{A} \rightarrow \mathbf{B} \\ y &: \ \mathbf{C} \rightarrow \mathbf{A} \\ z &: \ \mathbf{C} \end{aligned}$$

Agora precisamos encontrar o tipo do combinador B propriamente dito, que toma x, y, z (nessa ordem) e produz $app(x, app(y, z))$ do tipo $\mathbf{B}$. Para isso, faremos o movimento inverso do tipo do objeto resultante em direção ao início da lista ordenada dos objetos de entrada, introduzindo um '$\rightarrow$' sempre que movermos uma posição na lista. Daí construimos os seguintes passos:

$$\begin{aligned} \mathsf{B}xyz &: \ \mathbf{B} \\ \mathsf{B}xy &: \ (\mathbf{C} \rightarrow \mathbf{B}) \\ \mathsf{B}x &: \ ((\mathbf{C} \rightarrow \mathbf{A}) \rightarrow (\mathbf{C} \rightarrow \mathbf{B})) \\ \mathsf{B} &: \ (\mathbf{A} \rightarrow \mathbf{B}) \rightarrow ((\mathbf{C} \rightarrow \mathbf{A}) \rightarrow (\mathbf{C} \rightarrow \mathbf{B})) \end{aligned}$$

e nós obtemos a funcionalidade (tipo) de B como:

$$(\mathbf{A} \rightarrow \mathbf{B}) \rightarrow ((\mathbf{C} \rightarrow \mathbf{A}) \rightarrow (\mathbf{C} \rightarrow \mathbf{B}))$$

Se repetirmos o mesmo procedimento para os outros combinadores, encontramos o seguinte esquema de tipos:

$$\begin{aligned} \mathsf{I} &: \ \mathbf{A} \rightarrow \mathbf{A} \\ \mathsf{C} &: \ (\mathbf{A} \rightarrow (\mathbf{B} \rightarrow \mathbf{C})) \rightarrow (\mathbf{B} \rightarrow (\mathbf{A} \rightarrow \mathbf{C})) \\ \mathsf{W} &: \ (\mathbf{A} \rightarrow (\mathbf{A} \rightarrow \mathbf{B})) \rightarrow (\mathbf{A} \rightarrow \mathbf{B}) \\ \mathsf{S} &: \ (\mathbf{A} \rightarrow (\mathbf{B} \rightarrow \mathbf{C})) \rightarrow ((\mathbf{A} \rightarrow \mathbf{B}) \rightarrow (\mathbf{A} \rightarrow \mathbf{C})) \\ \mathsf{K} &: \ \mathbf{A} \rightarrow (\mathbf{B} \rightarrow \mathbf{A}) \end{aligned}$$

e que podem ser lidos como esquemas axiomáticos da lógica implicacional se '$\rightarrow$' é lido como 'implica'.

4.3.2 A interpretação funcional da implicação

O paradigma fórmula-como-tipos é uma conseqüência direta da teoria da funcionanlidade de Curry, descrita anteriormente. De um lado tem-se fórmulas, e do outro termos funcionais. Os termos λ são as provas (ou construções) dos axiomas implicacionais, ou seja, os termos possuem um tipo '$\rightarrow$'.

De modo a observar mais atentamente a regra de abstração, vamos ver a definição padrão para se obter o lambda cálculo:

Os termos λ

O λ-cálculo é uma coleção de vários sistemas formais baseados em uma notação funcional inventada por Alonso Church em 1930. Eles são projetados para capturar os aspectos mais básicos do modo como operadores ou funções podem ser combinados para formar outros operadores [Bar81].

O uso da notação λ tem como motivação simplificar o manuseio de funções de alta ordem (funções que atuam sobre outras funções). Consideremos o seguinte exemplo tirado de [Bar81]:

Exemplo 4.2 *Seja a expressão matemática '$x - y$'. Ela pode ser vista como definindo uma função f de x ou uma função g de y:*

$$f : x \rightarrow x - y$$

$$g : y \rightarrow x - y$$

Existe a necessidade de uma notação que dê a f e g nomes diferentes de uma maneira sistemática. Na prática, os matemáticos utilizam notações de maneira *ad-hoc*, mas isso pode trazer complicações ao se trabalhar com funções de alta ordem.

A notação de Church é uma maneira sistemática de construir, para cada expressão envolvendo x, uma notação para a correspondente função de x (e similarmente de y, etc.). Church escolheu o λ como um símbolo auxiliar usado para ligar as variáveis livres das quais a função depende. Para o exemplo acima, as funções seriam definidas da seguinte maneira:

$$f : \lambda x.x - y$$

$$g : \lambda y.x - y$$

Formalmente, os termos λ são definidos em [Bar81] como segue:

Definição 4.2 (Termos λ [Bar81]) *Suponha-se que é dada uma seqüência infinita de símbolos distintos chamados* variáveis, *uma seqüência infinita, finita ou vazia de símbolos distintos chamados* constantes. *O conjunto de expressões chamadas* termos λ *é definido indutivamente como segue:*

1. *Todas as variáveis e constantes são termos λ (chamados átomos);*

2. *Se M e N são termos λ quaisquer, então (MN) é um termo λ (chamado uma aplicação);*

3. *Se M é um termo λ qualquer e x uma variável qualquer, então $(\lambda x.M)$ é um termo λ (chamado uma abstração).*

Para atribuir tipos aos termos λ é necessária uma nova regra (além da APP), que corresponde à seguinte construção[1]:

$$(\lambda x^{\alpha}.M^{\beta})^{\alpha\rightarrow\beta}$$

A regra utilizada é a regra de introdução do $\rightarrow$, no estilo do sistema de dedução natural de Gentzen:

$$\dfrac{\begin{array}{c}[x \in \alpha]\\ M \in \beta\end{array}}{\lambda x.M \in \alpha \rightarrow \beta}$$

Desse modo, juntamente com a regra APP, a prova dos axiomas implicacionais referentes a cada combinador é construída. O seguinte exemplo mostra como obter a correspondência entre os termos λ e o combinador B, usando a terminologia do *LND*, mostrada na tabela 4.2:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[z:\mathbf{R}] \quad [y:\mathbf{R}\rightarrow\mathbf{A}]}{\mathtt{APP}(y,z):\mathbf{A}} \qquad [x:\mathbf{A}\rightarrow\mathbf{P}]}{\mathtt{APP}(x,\mathtt{APP}(y,z)):\mathbf{P}}}{\lambda z.\mathtt{APP}(x,\mathtt{APP}(y,z)):\mathbf{R}\rightarrow\mathbf{P}}}{\lambda y.\lambda z.\mathtt{APP}(x,\mathtt{APP}(y,z)):(\mathbf{R}\rightarrow\mathbf{A})\rightarrow(\mathbf{R}\rightarrow\mathbf{P})}}{\lambda x.\lambda y.\lambda z.\mathtt{APP}(x,\mathtt{APP}(y,z)):(\mathbf{A}\rightarrow\mathbf{P})\rightarrow((\mathbf{R}\rightarrow\mathbf{A})\rightarrow(\mathbf{R}\rightarrow\mathbf{P}))}$$

[1] Essa construção faz parte da definição de termos λ com tipos ([Bar81], capítulo 13):

Definição 4.3 (Termos λ com tipo) *Para cada tipo α, assumindo que se tenha um número infinito de variáveis v^{α} de tipo α, e algumas constantes c^{α}, os termos λ com tipo são definidos como segue:*

- *Cada v^{α} e c^{α} é um termo λ de tipo α;*
- *Se $M^{\alpha\rightarrow\beta}$ e N^{α} são termos λ de tipo $\alpha\rightarrow\beta$ e α, respectivamente, então $(M^{\alpha\rightarrow\beta}N^{\alpha})^{\beta}$ é um termo λ do tipo β;*
- *Se x^{α} é uma variável de tipo α e M^{β} é um termo λ de tipo β, então $(\lambda x^{\alpha}.M^{\beta})^{\alpha\rightarrow\beta}$ é um termo λ de tipo $\alpha\rightarrow\beta$.*

4.3.3 As contribuições de Howard e Tait

Howard estende o paradigma fórmulas-como-tipos para a lógica intuicionística de primeira ordem, permitindo, assim, que a interpretação funcional seja aplicada a outros conectivos lógicos (conjunção, disjunção, quantificadores) [How80]. Ele apresentou uma investigação mais precisa do isomorfismo entre as provas de dedução natural dos teoremas de Heyting e os termos do λ-cálculo. Por esse motivo, a interpretação funcional é conhecida também como o isomorfismo de Curry-Howard. No entanto, Gabbay e de Queiroz, em [dQG92, GdQ92], reinvidicam que é bastante razoável referir-se à interpretação funcional como sendo de Curry-Howard-Tait. Em [GdQ92], eles ressaltam a importância do trabalho de Tait citando a seguinte observação feita por Howard em [How80]:

> "H. Curry (1958) has observed that there is a close correspondence between *axioms* of positive implicational propositional logic, on the one hand, and *basic combinators* on the other hand. (...) The following notion of construction, for positive implicational propositional logic, was motivated by Curry's observation. More precisely, Curry's observation provided *half* the motivation. The other half was provided by W. Tait's discovery of the close correspondence between cut elimination and reduction of λ-terms (W. W. Tait 1965)[2]." (citado em [GdQ92]).

4.4 O sistema de dedução natural rotulado

Dentro do propósito de Gabbay de estudar as características do meta-nível na linguagem objeto, a interpretação funcional fornece a fundamentação adequada, por ser um paradigma onde se tem um cálculo funcional nos termos, harmonizado com um cálculo lógico nas fórmulas. As características do meta-nível, as informações procedurais, são trabalhadas no cálculo funcional nos termos (rótulos), enquanto que as informações de natureza declarativa são tratadas no cálculo lógico das fórmulas. O sistema de dedução natural rotulada atende a esse propósito, substituindo o paradigma fórmulas-como-tipos pelo paradigma rótulo-e-fórmula. Como resultado, tem-se um *framework* geral, no qual

[2] "H. Curry (1958) observou que existe uma correspondência entre os axiomas da lógica proposicional implicacional positiva, de um lado, e os combinadores básicoa do outro lado. (...) A noção de construção, surgida a partir daí, para a lógica proposicional implicacional positiva, foi motivada pela observação de Curry. Mais precisamente, a observação de Curry forneceu metade da motivação. A outra metade foi fornecida pela descoberta de W. Tait da correspondência entre a eliminação do corte e a redução de λ-termos (W. W. Tait 1965)"

é possível trabalhar com outras lógicas além da intuicionística.

4.4.1 A dedução natural de Gentzen

Em [Gen35], Gentzen propôs uma teoria da dedução baseada nas regras de manuseio de suposições, em lugar da usual formulação axiomática, na qual praticamente não há lugar para o raciocínio hipotético. O seu propósito era construir um sistema formal que se aproximasse ao máximo do modo de raciocínio humano. Como resultado desse propósito, Gentzen apresenta um cálculo de dedução para a lógica intuicionística, o cálculo **NJ**, e um cálculo para a lógica clássica, o cálculo **NK**. Esse último é obtido a partir no cálculo **NJ**, através da inclusão da lei do meio excluído como um axioma. Esses sistemas ficaram conhecidos como sistema de Dedução Natural.

Uma característica importante da Dedução Natural é que para cada conectivo lógico existem dois tipos de regras: as regras de introdução e as regras de eliminação. Desse modo, há uma análise da dedução do ponto de vista da capacidade de cada conectivo lógico. As regras de inferência, projetadas num padrão de regras de introdução e regras de eliminação, seguem o *princípio da inversão*: as regras de elimição são o inverso das regras de introdução.

Na tabela 4.1 são mostradas as regras do sistema de Dedução Natural. As suposições são representadas entre colchetes ('[' e ']') e cada regra de inferência possui a indicação do conectivo e do passo dedutivo dado (*Intr - introdução e Elim - eliminação*):

As seguintes restrições devem ser observadas nas regras de dedução natural:

(i) Na '$\forall - Intr$', 'a' não deve ocorrer em nenhuma suposição da qual '$P(a)$' dependa;

(ii) Na '$\exists - Elim$', 'a' não deve ocorrer em '$\exists x.P(x)$', em 'C' ou em qualquer outra suposição da qual a premissa 'C' dependa, exceto '$P(a)$';

(iii) Nas regras para a negação, o $\mathcal{F}$ é a contradição, o falso, assim, 'A' deve ser atômico e diferente de '$\mathcal{F}$'. É importante notar que a negação é encarada como um caso particular da implicação: $\neg A$ equivale a $A \rightarrow \mathcal{F}$

Os cálculos **NJ** e **NK**, mostraram-se inadequados para a prova do *Hauptsatz* (teorema principal ou teorema da eliminação do corte[3]), devido à falta de simetria do cálculo **NJ** (com relação

[3]O enunciado do teorema da eliminação do corte é o seguinte [Gen35]:

Teorema 4.1 (Eliminação do corte) *Toda derivação* **LJ** *ou* **LK** *(cálculo de seqüentes intuicionístico ou clássico)*

Introdução	Eliminação
$\dfrac{A \quad B}{A \wedge B}(\wedge\text{-}Intr)$	$\dfrac{A \wedge B}{A}(\wedge\text{-}Elim1)$ $\quad$ $\dfrac{A \wedge B}{B}(\wedge\text{-}Elim2)$
$\dfrac{A}{A \vee B}(\vee\text{-}Intr1)$ $\quad$ $\dfrac{B}{A \vee B}(\vee\text{-}Intr2)$	$\dfrac{A \vee B \quad \begin{matrix}[A]\\ C\end{matrix} \quad \begin{matrix}[B]\\ C\end{matrix}}{C}(\vee\text{-}Elim)$
$\dfrac{\begin{matrix}[A]\\ B\end{matrix}}{B}(\rightarrow\text{-}Intr)$	$\dfrac{A \quad A \rightarrow B}{B}(\rightarrow\text{-}Elim)$
$\dfrac{P(a)}{\forall x.P(x)}(\forall\text{-}Intr)$	$\dfrac{\forall x.P(x)}{P(t)}(\forall\text{-}Elim)$
$\dfrac{P(t)}{\exists x.P(x)}(\exists\text{-}Intr)$	$\dfrac{\exists x.P(x) \quad \begin{matrix}[P(a)]\\ C\end{matrix}}{C}(\exists\text{-}Elim)$
$\dfrac{\begin{matrix}[A]\\ \mathcal{F}\end{matrix}}{\neg A}(\neg\text{-}Intr)$	$\dfrac{A \quad \neg A}{\mathcal{F}}(\neg\text{-}Elim)$

Tabela 4.1: Regras do Sistema de Dedução Natural às regras para a negação) e a falta de elegância do cálculo **NK**, embora nele estivessem presentes as propriedades essenciais para a validade do *Hauptsatz*. Assim, Gentzen propõe un novo cálculo que ficou conhecido como 'cálculo de seqüentes', com o objetivo de provar o *Hauptsatz*.

No cálculo de seqüentes as derivações são elaboradas através de regras que promovem a substituição de seqüentes por outros seqüentes. Um seqüente é uma expressão do tipo $\Gamma \vdash \Delta$, onde Γ e Δ são seqüências finitas de fórmulas $A_1, A_2, ..., A_n$ e $B_1, B_2, ..., B_m$. A seqüência de fórmulas à esquerda do sinal $\vdash$ corresponde a uma conjução de fórmulas e a seqüência à direita, a uma disjunção. As regras do cálculo de seqüentes são de dois tipos: estruturais e operacionais. As primeiras são o enfraquecimento, a permutação e o corte; as regras operacionais atuam do lado esquerdo e direito do sinal $\vdash$ e correspondem, respectivamente, à eliminação e introdução de cada conectivo lógico.

Dentre as características do cálculo de seqüentes, destaca-se o fato de que esse cálculo pode ser visto como um meta-cálculo, pois ele guarda o histórico da dedução. A todo momento se tem acesso às suposições utilizadas e às conclusões obtidas. A seqüência de fórmulas do lado esquerdo do sinal $\vdash$ são suposições e as fórmulas à direita do sinal são conclusões, provadas a partir das suposições do lado esquerdo. Por esse motivo o cálculo de seqüentes pode ser visto como uma ferramenta para se estudar o "raciocínio sobre o raciocínio".

Posteriormente aos trabalhos de Gentzen, o trabalho de Prawitz [Pra65, Pra71] se destaca por oferecer contribuições importantes dentro da teoria da prova. Prawitz estende os resultados que Gentzen obteve no cálculo de seqüentes para a dedução natural e inclui a aplicação da dedução natural na lógica modal.

Uma das principais contribuições de Prawitz foi a definição de normalização para a dedução natural, onde uma prova é transformada para a forma normal através das regras de redução. Uma prova está na forma normal quando não contém redundâncias, como por exemplo, a introdução de um conectivo seguida de sua eliminação. Esse tipo de redundância equivale à existência da regra do corte no cálculo de seqüentes. A normalização é o equivalente do *Hauptsatz* para a dedução

---

pode ser transformada em uma outra derivação **LJ** *ou* **LK** *com o mesmo seqüente final e no qual a regra de inferência do corte não ocorra.*

A regra do corte é formulada como a seguir:

$$\frac{\Gamma \vdash D, \Theta \quad \Delta, D \vdash \Lambda}{\Gamma, \Delta \vdash \Theta, \Lambda}$$

onde $\Gamma, \Delta, \Theta, \Lambda$ são seqüências finitas de fórmulas e D é uma fórmula.

natural.

4.4.2 O *LND*

As regras de inferência para o *LND* são mostradas na tabela 4.2, e seguem o mesmo padrão das regras para a dedução natural (maiores detalhes sobre as regras para o *LND* podem ser encontrados em [dQG92]).

As regras de introdução dos conectivos lógicos na dedução natural rotulada possuem uma correspondência com o sistema de Heyting. A concepção de Heyting não pergunta se uma sentença A é verdadeira, mas sim qual a prova de A. Em [dQG92], as provas canônicas à la Heyting são explicadas de acordo com a tabela 4.3.

O *LND* apresenta as seguintes potencialidades:

- O dispositivo de abstrair variáveis é mais geral, não é necessariamente ligado ao conectivo implicacional, mas sim ao descarte das suposições no cálculo lógico (ver as regras $\exists - Elim$ e $\vee - Elim$, por exemplo);
- Há um melhor tratamento das regras classificadas como *impróprias* na terminologia de Prawitz [Pra65] (as regras que descartam suposições). A falta de controle local é um problema para a dedução natural pura. Quando se introduz rótulos ao longo das fórmulas, o descarte das hipóteses é refletido no rótulo da conclusão pretendida, através do dispositivo de abstrair variáveis. O nome arbitrário, introduzido como rótulo da suposição, perde sua identidade na conclusão, pois é abstraído. Desse modo é fácil verificar se uma prova ainda depende de alguma suposição: basta checar se os nomes arbitrários das hipóteses estão ligados por qualquer um dos abstratores disponíveis.
- É possível classificar diferentes sistemas de implicação proposicional (linear, relevante, intuicionística, clássica, etc.) através da extensão da interpretação funcional, obtida com o controle das abstrações λ do sistema. A regra de introdução do $\rightarrow$ é sujeita a condições laterais, ou seja, se a abstração liga todas as variáveis correspondentes ou não, se é permitido uma abstração lambda vácua[4], etc. Desse modo, os sistemas de lógica são classificados através da disciplina de abstração λ adotada. Por exemplo, no caso da lógica relevante, não se aceita abstrações vácuas.

[4] Em uma abstração vácua a variável abstraída não faz parte do escopo do termo λ; por exemplo, no termo $\lambda x.g(y)$, λx é uma abstração vácua.

Introdução	Eliminação
$\dfrac{x:\mathbf{A} \quad y:\mathbf{B}}{\langle x,y\rangle:\mathbf{A}\wedge\mathbf{B}}(\wedge\text{-}Intr)$	$\dfrac{x:\mathbf{A}\wedge\mathbf{B}}{\mathtt{FST}(x):\mathbf{A}}(\wedge\text{-}Elim1) \qquad \dfrac{x:\mathbf{A}\wedge\mathbf{B}}{\mathtt{SND}(x):\mathbf{B}}(\wedge\text{-}Elim2)$
$\dfrac{x:\mathbf{A}}{\mathtt{inl}(x):\mathbf{A}\vee\mathbf{B}}(\vee\text{-}Intr1) \qquad \dfrac{y:\mathbf{B}}{\mathtt{inr}(y):\mathbf{A}\vee\mathbf{B}}(\vee\text{-}Intr2)$	$\dfrac{x:\mathbf{A}\vee\mathbf{B} \quad \begin{array}{c}[y:\mathbf{A}]\\ \vdots\\ f(y):\mathbf{C}\end{array} \quad \begin{array}{c}[z:\mathbf{B}]\\ \vdots\\ g(z):\mathbf{C}\end{array}}{\mathtt{CASE}(x,\upsilon y.f(y),\upsilon z.g(z)):\mathbf{C}}(\vee\text{-}Elim)$
$\dfrac{\begin{array}{c}[x:\mathbf{A}]\\ \vdots\\ f(x):\mathbf{B}\end{array}}{\lambda x.f(x):\mathbf{A}\to\mathbf{B}}(\to\text{-}Intr)$	$\dfrac{x:\mathbf{A} \quad f:\mathbf{A}\to\mathbf{B}}{\mathtt{APP}(f,x):\mathbf{B}}(\to\text{-}Elim)$
$\dfrac{\begin{array}{c}[x:\mathbf{D}]\\ \vdots\\ f(x):\mathbf{P}(x)\end{array}}{\Lambda x.f(x):\forall x^{\mathbf{D}}.\mathbf{P}(x)}(\forall\text{-}Intr)$	$\dfrac{x:\mathbf{D} \quad c:\forall x^{\mathbf{D}}.\mathbf{P}(x)}{\mathtt{EXTR}(c,a):\mathbf{P}(a)}(\forall\text{-}Elim)$
$\dfrac{a:\mathbf{D} \quad f(a):\mathbf{P}(a)}{\epsilon x.(f(x),a):\exists x^{\mathrm{D}}.\mathbf{P}(x)}(\exists\text{-}Intr)$	$\dfrac{e:\exists x\mathbf{P}(x) \qquad \begin{array}{c}[t:\mathbf{D},g(t):\mathbf{P}(t)]\\ \vdots\\ d(g,t):\mathbf{C}\end{array}}{\mathtt{INST}(e,\sigma g.\sigma t.d(g,t)):\mathbf{C}}(\exists\text{-}Elim)$

Tabela 4.2: Regras da Dedução Natural Rotulada

Proposições	Provas canônicas
$\mathbf{A}_1 \wedge \mathbf{A}_2$	$\langle a_1, a_2 \rangle$ onde a_1 é a prova de $\mathbf{A}_1$ and a_2 é a prova de $\mathbf{A}_2$
$\mathbf{A}_1 \vee \mathbf{A}_2$	$\mathtt{inl}(a_1)$ onde a_1 é a prova de $\mathbf{A}_1$ ou $\mathtt{inr}(a_2)$ onde a_2 é a prova de $\mathbf{A}_2$ ('inl' e 'inr' abreviam '*into left disjunct*' e '*into right disjunct*', respectivamente)
$\mathbf{A} \rightarrow \mathbf{B}$	$\lambda x.b(x)$ onde $b(a)$ é a prova de $\mathbf{B}$ dado que a é a prova de $\mathbf{A}$
$\forall x^{\mathrm{D}}.\mathbf{P}(x)$	$\Lambda x.f(x)$ onde $f(a)$ é a prova de $\mathbf{P}(a)$, assegurado que a é um indivíduo arbitrário escolhido do domínio $\mathbf{D}$
$\exists x^{\mathrm{D}}.\mathbf{P}(x)$	$\varepsilon x.(f(x), a)$ onde a é um indivíduo (testemunha) do domínio $\mathbf{D}$, e $f(a)$ é a prova $\mathbf{P}(a)$

Tabela 4.3: Correspondência do *LND* com o sistema de Heyting

- Há um tratamento mais explícito dos contextos e dos escopos; desta forma, os elementos são introduzidos explicitamente como rótulos. O domínio de quantificação é introduzido explicitamente como uma fórmula, por exemplo: "a:D", que se lê: "a é um elemento do domínio D".

Devido a essa última característica, as dificuldades do sistema de dedução natural sem rótulos em lidar com lógicas inclusivas[5] são superadas. No *LND* a dificuldade de formular um sistema para lógica inclusiva não existe simplesmente porque os indivíduos são tomados como parte do cálculo funcional. Não é o caso para o sistema de dedução natural sem rótulos que não tem uma forma direta de manipular termos ou símbolos funcionais nas deduções.

Por exemplo, no seguinte caso:

$$\dfrac{\dfrac{\dfrac{[\forall x.\mathbf{F}(x)]}{\mathbf{F}(t)}}{\exists x.\mathbf{F}(x)}}{\forall x.\mathbf{F}(x) \to \exists x.\mathbf{F}(x)}$$

o termo t não é introduzido explicitamente como uma suposição extra, como seria no caso de uma leitura informal da dedução acima ('seja t um elemento arbitrário do domínio').

No caso do *LND*, essa mesma dedução ficaria como segue:

$$\dfrac{\dfrac{\dfrac{[t:\mathbf{D}] \qquad [z:\forall x^{\mathbf{D}}.\mathbf{F}(x)]}{\mathtt{EXTR}(z,t):\mathbf{F}(t)}}{\varepsilon x.(\mathtt{EXTR}(z,x),t):\exists x^{\mathbf{D}}.\mathbf{F}(x)}}{\lambda z.\varepsilon x.(\mathtt{EXTR}(z,x),\boxed{t}):\forall x^{\mathbf{D}}.\mathbf{F}(x) \to \exists x^{\mathbf{D}}.\mathbf{F}(x)}$$

A presença da variável livre t indica que a suposição $[t : D]$ ainda não foi descartada, ou seja, a prova da proposição ainda depende da suposição 'seja t um elemento arbitrário do domínio'. Para ser categórica, a prova acima deverá prosseguir um passo a mais:

$$\dfrac{\dfrac{\dfrac{\dfrac{[t:\mathbf{D}] \qquad [z:\forall x^{\mathbf{D}}.\mathbf{F}(x)]}{\mathtt{EXTR}(z,t):\mathbf{F}(t)}}{\varepsilon x.(\mathtt{EXTR}(z,x),t):\exists x^{\mathbf{D}}.\mathbf{F}(x)}}{\lambda z.\varepsilon x.(\mathtt{EXTR}(z,x),\boxed{t}):\forall x^{\mathbf{D}}.\mathbf{F}(x) \to \exists x^{\mathbf{D}}.\mathbf{F}(x)}}{\underbrace{\lambda t.\lambda z.\varepsilon x.(\mathtt{EXTR}(z,x),t)}_{\text{sem variável livre}}:\mathbf{D} \to (\forall x^{\mathbf{D}}.\mathbf{F}(x) \to \exists x^{\mathbf{D}}.\mathbf{F}(x))}$$

[5] A lógica inclusiva deve ser correta para os domínios vazios e não vazios.

Agora, olhando para a construção da prova, isto é, o rótulo '$\lambda t.\lambda z.\varepsilon x.(\texttt{EXTR}(z,x),t)$' verifica-se que não há variáveis livres, portanto a prova não está baseada em nenhuma suposição, e é, por isso, categórica.

Além das características acima, o *LND* possui a seguinte propriedade fundamental:

Definição 4.4 (Propriedade da subdedução) *Em toda dedução do LND, cada passo da prova é registrado nos rótulos. Dessa maneira, os rótulos contêm um histórico da dedução. Conseqüentemente, na dedução de uma fórmula rotulada t : $\mathbf{A}$, todas as subdeduções utilizadas para obter $\mathbf{A}$ estão registradas em t.*

4.5 De volta a Frege

O *framework* dos *Labelled Deductive Systems* de Gabbay, é o resultado de uma tentativa de reinterpretação do cálculo lógico de Frege, onde abstratores e operadores funcionais trabalham simultaneamente e harmoniosamente com conectivos lógicos e quantificadores. Em outras palavras, a interpretação funcional pode ser vista dentro de uma perspectiva mais ampla, proveniente da metodologia de sistemas dedutivos rotulados, tal que se permite estudar a matemática da dedução em uma gama de lógicas, incluindo algumas que podem até não se limitar aos princípios da interpretação intuicionística. Como contra-partida concreta, é apresentada uma teoria da prova baseada em um sistema de dedução natural rotulada onde a noção de subdedução substitui a tradicional subfórmula. Sua fundamentação conceitual pode ser vista como uma reinterpretação da perspectiva funcional de Frege à lógica: é como se a teoria de funções definida no *Grundgesetze* de Frege fosse combinada com a teoria de predicados e quantificadores apresentada em *Begriffsschrift*, de tal maneira que uma fórmula é verdadeira (válida) se e somente se uma dedução da fórmula pode ser construída tal que o rótulo da conclusão não contém nenhuma variável livre (isto é, sua prova é um objeto completo, o que significa que a verdade da fórmula não depende de qualquer suposição) [dQG92].

4.6 Abdução rotulada

Em [Gab94] o capítulo 16 é dedicado à abdução em sistemas dedutivos rotulados. Apenas algumas noções intuitivas sobre o tema são colocadas, mas vale a pena ressaltar alguns pontos.

Duas idéias são colocadas como chave:

- a abdução depende dos procedimentos de prova utilizados;
- os princípios abdutivos (que nós chamamos de hipóteses abdutivas) podem ser parte dos dados. Em outras palavras, um ítem declarativo de dado pode ser uma fórmula ou princípio abdutivo.

A situação básica é apresentada da seguinte forma:

Δ	$\vdash ?!Q$
dado	? pergunta ou ! entrada

Aqui temos um relacionamento entre um banco de dados e uma fórmula. O relacionamento pode ser declarativo ($?Q$, Q uma pergunta) ou imperativo ($!Q$, Q como uma entrada ou uma demanda para executar abdução, ou uma demanda para uma explicação, etc.). No caso imperativo existe uma interação entre Δ e Q e um novo banco de dados Δ' emerge.

As regras de abdução são vistas como movimentos extras que ajudam a responder a pergunta ou ajudam a alterar o banco de dados como resultado da pergunta ou da entrada.

Segundo Gabbay, para executar a abdução são necessários procedimentos de prova mais precisos ou procedimentos de atualização estruturados e então pode-se definir sobre eles as regras abdutivas.

O seguinte exemplo ilustra a intuição da abdução em [Gab94]:

Exemplo 4.3 *Seja o seguinte banco de dados de cláusulas tipo Horn. Ele é rotulado no sentido de que cada cláusula é nomeada. A pergunta em questão é D, que não é derivável do banco de dados como ele está:*

	Dados	*Pergunta*
a_1 :	$I \wedge T \rightarrow D$	$?D$
a_2 :	$L \rightarrow I$	
a_3 :	$L \wedge S \rightarrow T$	
a_4 :	$O \wedge P \rightarrow T$	
a_5 :	L	

Dando uma possível interpretação ao banco de dados acima, tem-se:

	Dados	*Pergunta*
a_1 :	$Interview \wedge Thesis \rightarrow Degree$	$?Degree$
a_2 :	$Lecture \rightarrow Interview$	
a_3 :	$Lecture \wedge ScholarlySurvey \rightarrow Thesis$	
a_4 :	$(job)Offer \wedge Publication \rightarrow Thesis$	
a_5 :	$Lecture$	

onde pode-se entender o seguinte: na universidade de Stanford no departamento de inglês, existem duas maneiras de alcançar o título de PhD. Pode-se propor uma tese, ficar no departamento por 4 a 5 anos adquirindo e produzindo conhecimentos e passar em uma entrevista, ou pode-se escrever uma publicação muito boa e conseguir um emprego oferecido por outra universidade que esteja entre as 10 melhores do paı́s.

Para computar a pergunta $?D$ ($?Degree$), do exemplo acima, Gabbay coloca as seguintes opções:

1. A primeira opção é abduzir a primeira coisa que for encontrada. Nesse caso, adicionaríamos o próprio D ao banco de dados, ou seja, o princípio abdutivo seria D. Mas essa abdução não tem sentido: "Para conseguir o título você consegue o título"!?

2. A segunda opção é abduzir somente literais que não aparecem na cabeça das cláusulas (seguindo o princípio de literais abduzíveis da programação em lógica, visto na seção 3.3.1). Nesse caso, adicionamos S, porque S é o primeiro literal abduzível encontrado na execução *top down*. Note-se que aqui não é utilizado um conjunto de literais abduzíveis, a estrutura do banco de dados é que determina o que adicionar.

3. Se a lógica subjacente não é a lógica clássica, mas sim alguma lógica de recurso, adicionar S não será bem sucedido para alcançar D porque será necessário o uso de L duas vezes, uma para conseguir I na cláusula a_2 e outra para conseguir T na cláusula a_3. Assim, para uma lógica de recurso, o resultado da abdução será $O \wedge P$, a não ser que se possa adicionar outra cópia de L.

4. Se queremos a suposição lógica mais fraca (em lógica clássica) que nos faz alcançar o objetivo, então deve ser adicionaddo $S \vee (O \wedge P)$. Esse princípio abdutivo é independente da computacão.

5. Em respostas cooperativas o princípio abdutivo pega a cláusula no nível mais alto. Nesse caso a resposta é T. À pergunta $?D$ responde-se 'sim, se T'.

6. O poder do mecanismo de rotulação do *LDS* pode ser facilmente ilustrado por um uso mais refinado dos rótulos. Se os átomos são rotulados com valores de custo, por exemplo, o princípio abdutivo pode buscar um custo minimal. Pode-se também colocar custo na própria computação e buscar abduzir fórmulas que dão uma maior provabilidade, com um número menor de instâncias de *modus ponens*.

Além do exemplo acima, outros exemplos são colocados em [Gab94], inclusive faz-se referência ao uso da abdução na interpretação de linguagem natural, onde é citado o artigo [GKP94], abordado na seção 3.4.2.

Além dos exemplos, Gabbay coloca alguns poucos pontos sobre os mecanismos da abdução rotulada: O *LDS* é um sistema preciso de regras, permitindo mostrar (ou falhar ao mostrar) que $\Delta \vdash \Gamma$, sendo Δ uma estrutura de dados e Γ uma estrutura-objetivo. Mais útil ainda é a estrutura da forma $t : G$. Então, para o propósito de explicar o que é a abdução no *framework* do *LDS*, assume-se que a noção:

$$\Delta \vdash t : G$$

é precisa e algoritmicamente definida, dado um *LDS L* particular.

Considerando, então, um banco de dados Δ contendo $\alpha : X$, que escreve-se como $\Delta[\alpha : X]$, em algum lugar na estrutura Δ, α é uma variável tipo rótulo e X é uma variável proposicional. Para qualquer escolha de X e de α, por exemplo, $\alpha = \alpha_0$ e $X = A$, $\Delta[\alpha_0 : A]$ é um banco de dados próprio.

Suponha-se que se deseje provar $t : G$. Então, para alguma (ou possivelmente nenhuma) fórmula bem formada A_i e rótulo α_i, devemos ter:

$$\Delta[\alpha_i : A_i] \vdash t : G$$

Um princípio abdutivo *Abduce*$(\alpha : X)$ é uma computação que pode escolher um ou mais dos α_i's e A_i's acima.

Segundo Gabbay, a importância do ponto de vista acima é que: (1) os bancos de dados podem tomar princípios abdutivos como parte dos seus dados, guardados em lugares corretos e (2) o princípio abdutivo é relativo ao procedimento de computação e ao resto do banco de dados.

Então, quando novos dados são colocados no banco de dados, o princípio abdutivo muda. Ganha-se um forte correspondente de aprendizado no banco de dados.

Para fechar, Gabbay coloca que o princípio indutivo é um caso especial do princípio abdutivo que aprende uma regra do tipo $A \rightarrow B$: matematicamente não há diferença alguma. Uma argumentação sobre essa colocação pode ser vista na seção 2.4 desta dissertação.

Capítulo 5

Teoria de Aprendizado Abdutivo via LDS

Neste capítulo, estudaremos diversos aspectos de uma formalização do raciocínio abdutivo utilizando a estrutura dos Sistemas Dedutivos Rotulados. A abdução é vista como a busca de candidatos para explicar determinado fato observado, considerando o conhecimento de *background* existente. A solução de um problema abdutivo será dividida em dois níveis básicos: o nível objeto, onde serão gerados os candidatos, e o meta-nível, onde serão escolhidos os candidatos mais plausíveis, de acordo com um critério meta-lógico.

5.1 Definição do problema abdutivo

Como vimos no capítulo 3, o objetivo da abdução é buscar uma explicação para um fato φ não questionável, explicação esta não presente no nosso conhecimento de *background*, ou nossa teoria Θ. Nossa proposta é que essa explicação será buscada tentando-se construir uma prova de φ a partir da teoria Θ. Em determinado momento da construção da prova não poderemos prosseguir por falta de dados, nesse ponto começamos a gerar nossos candidatos à explicação de φ ($Cand(\Theta, \varphi)$). Para escolher dentre os candidatos disponíveis aquele mais 'plausível' vamos então tratar o meta-nível da abdução lado a lado com o nível objeto utilizando os rótulos do *LDS* (a informação do meta-nível associada às fórmulas).

A principal característica da nossa proposta é o tratamento do meta-nível da abdução. O que aqui chamamos de meta-nível é a fase de escolha entre os diversos candidatos que podem ser gerados dedutivamente como possíveis hipóteses abdutivas. Obviamente, essa é uma tarefa dependente do domínio de aplicação do problema, e, na maioria das abordagens em abdução, essa fase é deixada para o especialista humano. Nosso objetivo é ir um passo à frente e fornecer ao especialista apenas aqueles candidatos que se mostrarem mais plausíveis, de acordo com determinado critério de preferência fornecido pelo usuário; ou ainda, fornecer uma lista de candidatos, ordenados de acordo com sua plausibilidade.

Para atingir esse objetivo, lançamos mão de uma das principais características do *LDS*, que é o tratamento do meta-nível lado a lado com o nível objeto. Assim, os rótulos do *LDS* retratarão uma ordem preferencial entre as fórmulas da teoria de *background*. Essa ordem, como já foi dito, será fornecida pelo usuário (especialista) e o significado dos rótulos poderá variar, a depender da aplicação. O sistema, então, utilizará essa informação (do meta-nível) dos rótulos para avaliar, metalogicamente, a plausibilidade dos candidatos gerados, possivelmente reduzindo a lista inicial gerada dedutivamente.

5.1.1 Os rótulos e sua ordenação

Um problema abdutivo será definido de forma similar à que foi usada em [CDT91]. Porém, para manipular as características do meta-nível do problema abdutivo a teoria de *background* deverá ser constituída de fórmulas rotuladas segundo um critério meta-lógico definido pelo usuário em função do problema. Além disso, deverá ser definida uma ordem, possivelmente parcial, entre os rótulos que corresponderá ao critério de preferência (meta-lógico) do referido problema. Esse critério deverá ser definido de maneira dependente do domínio e poderá ser interpretado como, por exemplo, medidas de probabilidade, distribuições de possibilidade ([DP90, DLP93]), ordem de preferência, grau de confiabilidade, custo, etc. A ordenação entre os rótulos será utilizada no momento da escolha da solução (ou soluções) mais plausível ao problema abdutivo.

Formalmente, a ordem poderá ser definida em termos de pares ordenados retratando o relacionamento entre os rótulos. No decorrer desta dissertação os exemplos trabalhados terão sua ordenação definida por diagramas, por uma questão apenas de ser visualmente mais compreensível.

Por exemplo, com a teoria Θ definida da seguinte forma:

$$\begin{aligned} \Theta = \{ \quad l_1 : & \quad A \rightarrow D, \\ l_2 : & \quad E \rightarrow C \wedge D, \\ l_3 : & \quad B \vee F \rightarrow D\} \end{aligned}$$

a ordem entre os elementos do conjunto de rótulos $\{l_1, l_2, l_3\}$ pode ser definida por um diagrama como na figura 5.1, ou através de um conjunto de pares ordenados: $\{\langle l_2, l_1\rangle, \langle l_2, l_3\rangle\}$. Significando que $l_2 \sqsubseteq l_1$ (l_2 é preferível a l_1) e $l_2 \sqsubseteq l_3$ (l_2 é preferível a l_3).

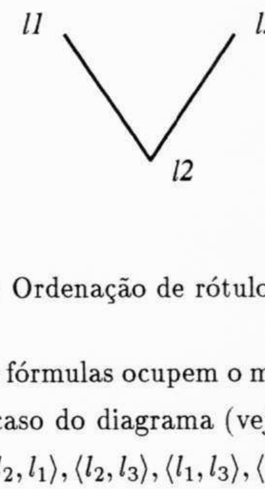

Figura 5.1: Ordenação de rótulos via diagrama

No caso particular em que duas fórmulas ocupem o mesmo lugar em uma ordenação, teríamos seus rótulos num mesmo ponto, no caso do diagrama (veja figura 5.2) ou teríamos os dois pares ordenados presentes no conjunto, $\{\langle l_2, l_1\rangle, \langle l_2, l_3\rangle, \langle l_1, l_3\rangle, \langle l_3, l_1\rangle\}$, significando que $l_2 \sqsubseteq l_1$, $l_2 \sqsubseteq l_3$, $l_1 \sqsubseteq l_3$ e $l_3 \sqsubseteq l_1$, ou seja, l_1 e l_3 têm o mesmo grau de preferência.

5.1.2 Os componentes de um problema abdutivo

Como já foi dito, definiremos formalmente um problema abdutivo de maneira semelhante à definição dada em [CDT91]; a diferença básica é a inclusão da ordem preferencial dos rótulos na definição.

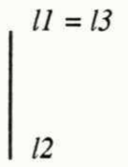

Figura 5.2: Rótulos com o mesmo grau de preferência

Definição 5.1 (Problema abdutivo) *Um problema abdutivo é dado por uma tripla* $\langle \Theta, \varphi, \sqsubseteq \rangle$, *onde:*

- Θ - *é uma teoria de background composta por um conjunto consistente de fórmulas, rotuladas segundo um critério de preferência qualquer.*

$$\Theta = \{ \begin{array}{l} l_1 : \alpha_1, \\ l_2 : \alpha_2, \\ \vdots \\ l_n : \alpha_n \} \end{array}$$

- φ - *é a fórmula a ser explicada, tal que:*

  (i) $\Theta \nvdash_L \varphi$ *(φ não é demonstrável a partir de Θ usando a lógica L);*

  (ii) $\Theta \nvdash_L \neg\varphi$ *(a negação de φ não é demonstrável a partir de Θ usando a lógica L).*

- $\sqsubseteq$ - *é uma relação de ordem, possivelmente parcial, definida entre as assertivas da teoria de background, denotando o critério de preferência entre as fórmulas.*

Uma solução para o problema dado pela tripla $\langle \Theta, \varphi, \sqsubseteq \rangle$ deve ser buscada entre as fórmulas α tal que:

$$\Theta \cup \{\alpha\} \vdash_L \varphi$$

Podemos definir formalmente a solução α, como feito em [MP95b]:

Definição 5.2 (Solução α para um problema abdutivo, a nível objeto) *Dado um problema abdutivo* $\langle \Theta, \varphi, \sqsubseteq \rangle$, *$\alpha$ é formalmente definida como uma solução para o referido problema como a seguir:*

$\Theta \vdash_L \alpha \rightsquigarrow \varphi$[1] *se:*

(i) $\Theta \vdash_L \alpha \rightarrow \varphi$ *e,*

(ii) *para todo* $\gamma \in Cand(\Theta, \varphi)$, *se* $\gamma \sqsubseteq_c \alpha$[2], *então* $\alpha \sqsubseteq_c \gamma$.

onde:

$$\alpha \sqsubseteq \gamma \equiv_{def} \vdash_L \gamma \rightarrow \alpha$$

$$Cand(\Theta, \varphi) \equiv_{def} \{\gamma : \Theta \vdash_L \gamma \rightarrow \varphi\}$$

$Cand(\Theta, \varphi)$ foi originalmente definido em [MP95b] como um conjunto de candidatos alternativos que constituem um conjunto de teste para qualquer explicação aceitável para φ em Θ. Na nossa proposta consideraremos $Cand(\Theta, \varphi)$ como o conjunto de candidatos gerados pelo sistema, ou eventualmente fornecidos pelo usuário (veja seção 5.4.1).

Outra distinção com relação ao especificado em [MP95b] é que ali o critério de preferência, $\sqsubseteq$, era considerado apenas como uma restrição lógica. Aqui nossa intenção é ampliar esse conceito, e o critério de preferência levará em conta também a ordenação dos rótulos.

A lógica (ou linguagem) L será dada como parâmetro e, a princípio, pode tratar-se de qualquer lógica com uma relação de conseqüência e regras bem definidas, como o permite o *LDS*. Nesta dissertação nos restringiremos a tratar problemas abdutivos proposicionais dentro da lógica clássica implicacional ou da lógica proposicional intuicionística. Essa restrição se dá devido ao fato de que a dedução natural (sistema que utilizaremos para a solução do problema a nível objeto) não se presta para tratar a disjunção clássica. Assim, na nossa proposta, incluimos a regra de redução ao absurdo mas não a lei do terceiro excluído. Uma investigação para extensão da proposta para a utilização de outras lógicas é colocada como trabalho futuro.

Na verdade, as condições fundamentais para que uma explicação seja aceita como uma 'interessante' solução para o problema dado, podem ser divididas em dois grupos. No primeiro grupo temos as condições de consistência e minimalidade, e possivelmente alguma restrição sintática, todas elas definidas no nível objeto da teoria Θ; no segundo grupo temos as condições ou critérios de preferência do meta-nível da abdução (dependentes do domínio) que serão tratados através dos rótulos.

[1] Leia-se 'dentro da teoria Θ com a lógica L, α é uma boa explicação para φ'.

[2] Leia-se 'γ é um candidato preferível a α'

Com relação às restrições do nível objeto, apesar de não haver um consenso geral no que diz respeito à distinção entre explicações interessantes e não interessantes, as três restrições abaixo são geralmente impostas às explicações α para um problema abdutivo $\langle\Theta, \varphi, \sqsubseteq\rangle$:

(i) α é consistente com Θ (Θ-consistente), ou seja, $\Theta \nvdash_L \neg\alpha$;

(ii) α é uma explicação minimal para o problema abdutivo $\langle\Theta, \varphi, \sqsubseteq\rangle$, ou seja, para qualquer fórmula γ, se $\Theta \cup \{\gamma\} \vdash_L \varphi$ e $\alpha \vdash_L \gamma$ então $\gamma \vdash_L \alpha$;

(iii) α tem alguma forma sintática restrita; por exemplo, α deve ser uma fórmula construída por literais utilizando apenas quantificadores e a conjunção (geralmente usada nas abordagens de programação em lógica).

As restrições do nível objeto consideradas nesse trabalho são as relativas aos ítens (i) e (ii) (de acordo com a definição 5.2 vista anteriormente); além disso trataremos também as restrições de integridade e as conseqüências dedutivas das hipóteses. A restrição sintática poderá ser adicionada à formalização aqui proposta sem que se perca a generalidade da caracterização.

Como já foi dito no Capítulo 3, pode-se usar um sistema dedutivo para executar a abdução. Neste trabalho propomos um método teórico para a abdução baseado em um sistema de dedução natural e utilizando a estrutura do *LDS*[3]. Consideraremos apenas a lógica proposicional, onde a geração do conjunto de explicações é fadada a terminar. O método gera todo o conjunto de explicações minimais e Θ-consistentes, considerando o nível objeto da teoria; em seguida as explicações mais interessantes, segundo um critério meta-lógico, são selecionadas.

5.2 Abdução proposicional baseada em dedução natural (o nível objeto)

Seja $\langle\Theta, \varphi, \sqsubseteq\rangle$ um problema abdutivo. Por definição, $\Theta \nvdash_L \varphi$, logo, não é possível construir uma prova de φ a partir da teoria Θ.

A solução para o prolema abdutivo $\langle\Theta, \varphi, \sqsubseteq\rangle$ pode ser encontrada entre as fórmulas que completam a prova de φ na teoria Θ. Nesse sentido será gerado um espaço de busca composto pelas possíveis árvores de dedução geradas na tentativa de construir uma prova para φ.

[3]A dedução natural e o *LDS* foram descritos no Capítulo 4.

Claro que não tem sentido completar a prova de φ simplesmente adicionando a própria fórmula φ à teoria; se assim o fosse, estaríamos explicando o fato φ com o próprio fato φ. As árvores de dedução devem ser expandidas o máximo possível e uma cuidadosa escolha dos átomos que encerram cada ramo deve ser feita.

As seguintes definições serão utilizadas na caracterização do processo de geração de explicações abdutivas, onde consideraremos as regras de dedução natural para a geração das árvores de prova. Essas definições foram adaptadas de [MP93] onde um método abdutivo via tableau e cálculo de seqüentes é especificado (ver seção 3.3.2).

Definição 5.3 (Árvore de dedução aceitável) *Seja R uma árvore de dedução. R é dita* **aceitável** *se para cada fórmula das folhas de R não pode ser aplicada mais nenhuma regra de introdução ou eliminação de conectivos lógicos, considerando a teoria adotada, ou se a fórmula é um átomo da teoria.*

Definição 5.4 (Conjunto de explicações para uma árvore aceitável) *Seja R uma árvore de dedução aceitável associada a um problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$. $F(R)$ é o conjunto das fórmulas presentes nas folhas da árvore R. $Exp(R)$ é definido da seguinte maneira:*

(i) $Exp(R) = \emptyset$ *, quando todas as folhas de R pertencem à teoria Θ, ou seja, esse é o caso em que $\Theta \vdash \varphi$;*

(ii) $Exp(R) = \perp$ *, ou indefinido, quando φ é inconsistente com Θ, ou seja, $\Theta \vdash \neg\varphi$;*

(iii) $Exp(R) = \{\alpha_1 \wedge \alpha_2 \wedge ... \wedge \alpha_n$*, tal que* $\alpha_i \in F(R)$ *e* $\alpha_i \notin \Theta\}$ *, quando não ocorre (i) nem (ii); ou seja, $Exp(R)$ é a conjunção das fórmulas presentes nas folhas de uma árvore de dedução associada a um problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$ sendo que cada uma dessas fórmulas não pertence à teoria Θ, e sendo que $\Theta \nvdash \neg\varphi$*

Definição 5.5 (Conjunto de candidatos) *O conjunto de todos os candidatos a serem uma solução para o problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$ será representado pelo conjunto $Cand(\Theta, \varphi)$ que será formado por todos os $Exp(R_i)$, onde cada R_i é uma árvore de dedução aceitável associada ao problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$.*

A partir dessas definições podemos chegar a uma noção de completude e corretude na geração dos candidatos a solução para um problema abdutivo. O teorema abaixo retrata essas noções.

Teorema 5.1 (Corretude e completude da geração de candidatos) *Seja $\langle \Theta, \varphi, \sqsubseteq \rangle$ um problema abdutivo. Então, todo elemento de $Cand(\Theta, \varphi)$ é uma explicação não contraditória para $\langle \Theta, \varphi, \sqsubseteq \rangle$ (corretude) e qualquer explicação minimal e não contraditória para $\langle \Theta, \varphi, \sqsubseteq \rangle$ é um elemento de $Cand(\Theta, \varphi)$ (completude).*

Argumentação

Considerando-se que as árvores de dedução foram exaustivamente geradas e expandidas e que por isso cada uma delas é aceitável (segundo definição 5.3) no sentido de que as fórmulas das folhas ou pertencem à teoria Θ ou a elas não pode ser aplicada nenhuma regra de eliminação ou introdução de conectivos lógicos (condição de falha para uma prova), temos que qualquer explicação minimal dentro da teoria Θ foi gerada (completude). Considerando também que os elementos de $Cand(\Theta, \varphi)$ foram formados pela conjunção das fórmulas das folhas de árvores aceitáveis associadas ao problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$, e que esse conjunto, $Cand(\Theta, \varphi)$ é definido quando $\Theta \not\vdash \neg\alpha$ (o caso (ii) da definição 5.4 não ocorre), temos que cada um dos elementos de $Cand(\Theta, \varphi)$ é uma explicação plausível para o problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$, pois cracteriza uma fórmula que completa uma prova de φ na teoria Θ.

5.2.1 Minimalidade e consistência

A condição de minimalidade correspondente às características do nível objeto de um problema abdutivo dizem respeito, como visto anteriormente, à relação de conseqüência lógica ($\vdash$). Então, do conjunto de candidatos à explicação de φ, $Cand(\Theta, \varphi)$ podemos selecionar os elementos minimais. Ou seja:

$$min(Cand(\Theta, \varphi)) = \{\alpha_i \in Cand(\Theta, \varphi), \text{ tal que } \forall \alpha_j \in Cand(\Theta, \varphi), \text{ se } i \neq j \text{ então } \Theta \not\vdash \alpha_i \rightarrow \alpha_j\}$$

A condição de consistência com a teoria pode ser garantida através da exclusão de todas as explicações que negam a teoria, como proposto por Mayer e Pirri em [MP93] (veja seção 3.3.2). Então podemos primeiro gerar todas as explicações minimais para a negação da teoria (o que seria equivalente à solução do problema $\langle \Theta, \perp, \sqsubseteq \rangle$) e então removê-las do conjunto $Cand(\Theta, \varphi)$ (ou $min(Cand(\Theta, \varphi))$, se o critério de minimalidade já houver sido aplicado). Isso é razoável se assumirmos que a teoria Θ não é modificada com freqüência, então o conjunto $Cand(\Theta, \perp)$ pode ser gerado uma única vez e usado toda vez que um problema abdutivo, dentro da mesma teoria Θ deva ser solucionado.

Por questão de simplificação na formalização, consideraremos o conjunto $Cand(\Theta, \varphi)$ como o conjunto de candidatos final, após a solução a nível objeto (inclusive após a aplicação do critério de minimalidade), ao invés de $min(Cand(\Theta, \varphi))$.

5.2.2 Restrições de integridade

O processo abdutivo como definido aqui também pode ser restringido com o uso de restrições de integridade como visto na seção 3.2. Nesse caso, as restrições de integridade devem ser expressas em termos de um conjunto de fórmulas I que devem ser adicionadas à teoria Θ, de modo que $\Theta \cup \{\alpha\}$ deverá satisfazer I. As restrições de integridade serão verificadas juntamente com a verificação de consistência com a teoria, bastando para isso adicionar as fórmulas $\gamma \in I$ à teoria no momento da geração das explicações que negam a teoria. Assim, ao invés de gerar os elementos do conjunto $Cand(\Theta, \bot)$, seriam gerados os elementos pertencentes ao conjunto $Cand(\Theta \cup I, \bot)$.

5.2.3 Conseqüências dedutivas das hipóteses

Uma importante consideração a fazer é que, no momento que optamos por determinada solução abdutiva, ou antes de optarmos, no momento que geramos candidatos à solução, esses candidatos podem implicar na existência de outros elementos relevantes ao problema. O que queremos dizer é que alguns candidatos podem ter como conseqüência lógica, considerando a teoria adotada, outras fórmulas que podem ou não estar presentes na teoria. Se essas conseqüências dedutivas estão presentes na teoria, isso pode ser considerado como um ponto a favor da plausibilidade do candidato gerado. No entanto, se o candidato tem conseqüências dedutivas que não estão presentes na teoria, essa informação deve ser considerada no momento de selecionar os candidatos mais plausíveis. Assim, para viabilizar a consideração dessa informação na comparação dos candidatos, após a geração do conjunto $Cand(\Theta, \varphi)$ deverão ser deduzidas as conseqüências lógicas da união de Θ com cada um dos candidatos gerados.

Uma referência a esse tipo de procedimento é encontrada em [MD94] (veja seção 3.4.6) onde é proposto um algoritmo de *forward chaining* para saturar a teoria com as conclusões de regras cujas premissas já estão presentes na teoria. Um estudo mais profundo sobre esse procedimento (geração das conseqüências dedutivas dos candidatos) e suas complexidades será proposto como trabalho futuro.

Podemos definir formalmente o conjunto das conseqüências dedutivas da seguinte maneira:

Definição 5.6 (Conseqüências dedutivas) *Seja $Cand(\Theta, \varphi)$ o conjunto dos candidatos gerados para o problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$. O conjunto das conseqüências dedutivas de cada elemento $\alpha_i \in Cand(\Theta, \varphi)$ é chamado de $Conseq(\Theta, \alpha_i)$ e é definido como segue:*

$$Conseq(\Theta, \alpha_i) = \{\beta \text{ tal que } \Theta \nvdash \beta, \Theta \cup \{\alpha_i\} \vdash \beta \text{ e } \beta \neq \varphi\}$$

Até aqui definimos a solução a nível objeto do problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$ e chegamos a um conjunto de candidatos minimais à explicação de φ. Logicamente falando, todos os $\alpha \in Cand(\Theta, \varphi)$ são passíveis de ser a explicação da fórmula φ, mas provavelmente somente um deles, ou até um outro diferente de todos eles, foi realmente a causa do fato φ. Essa é a questão crucial da abdução: como dito no Capítulo 3, não podemos garantir que a solução encontrada para um problema abdutivo seja a solução correta. O que pretendemos garantir é que, dados os critérios lógicos (do nível objeto, definidos acima) e os critérios meta-lógicos fornecidos pelo usuário (os rótulos e sua ordenação), a solução fornecida é uma das mais plausíveis.

5.3 O cálculo lógico e o cálculo funcional

Para tratar nosso problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$, necessitaremos de ferramentas para a solução a nível objeto e também para a solução (ou o melhoramento da solução) a nível meta-lógico. Para isso teremos um cálculo lógico para as fórmulas e um cálculo funcional para os rótulos.

O cálculo lógico que utilizaremos será a dedução natural, como vista no Capítulo 4. Para o cálculo funcional, consideramos as regras de dedução natural rotulada [dQG92] vistas também no Capítulo 4, e formulamos regras similares de introdução e eliminação dos conectivos lógicos, que serão aplicadas no momento da propagação dos rótulos.

5.3.1 Regras de introdução e eliminação de conectivos lógicos

As regras para cada conectivo lógico são as seguintes:

Conjunção

$$\frac{x : \mathbf{A} \quad y : \mathbf{B}}{I_{\wedge}(x, y) : \mathbf{A} \wedge \mathbf{B}} (\wedge - Intr)$$

$$\frac{x : \mathbf{A} \wedge \mathbf{B}}{E_{1\wedge}(x) : \mathbf{A}} (\wedge - Elim_1) \qquad \frac{x : \mathbf{A} \wedge \mathbf{B}}{E_{2\wedge}(x) : \mathbf{B}} (\wedge - Elim_2)$$

Disjunção

$$\frac{x : \mathbf{A}}{I_{1\vee}(x) : \mathbf{A} \vee \mathbf{B}} (\vee - Intr_1) \qquad \frac{y : \mathbf{B}}{I_{2\vee}(y) : \mathbf{A} \vee \mathbf{B}} (\vee - Intr_2)$$

$$\frac{x : \mathbf{A} \vee \mathbf{B} \quad \begin{matrix}[y : \mathbf{A}] \\ \vdots \\ f(y) : \mathbf{C}\end{matrix} \quad \begin{matrix}[z : \mathbf{B}] \\ \vdots \\ g(z) : \mathbf{C}\end{matrix}}{E_{\vee}(x, \lambda y.f(y), \lambda z.g(z)) : \mathbf{C}} (\vee - Elim)$$

Implicação

$$\frac{\begin{matrix}[x : \mathbf{A}] \\ \vdots \\ f(x) : \mathbf{B}\end{matrix}}{I_{\rightarrow}(\lambda x.f(x)) : \mathbf{A} \rightarrow \mathbf{B}} (\rightarrow -Intr) \qquad \frac{x : \mathbf{A} \quad y : \mathbf{A} \rightarrow \mathbf{B}}{E_{\rightarrow}(x, y) : \mathbf{B}} (\rightarrow -Elim)$$

Negação

$$\frac{\begin{matrix}[x : \mathbf{A}] \\ \vdots \\ f(x) : \perp\end{matrix}}{I_{\neg}(\lambda x.f(x)) : \neg\mathbf{A}} (\neg - Intr) \qquad \frac{x : \mathbf{A} \quad y : \neg\mathbf{A}}{E_{\neg}(x, y) : \perp} (\neg - Elim)$$

Reductio ad Absurdum[4]

$$\frac{\begin{matrix}[x : \neg\mathbf{A}] \\ \vdots \\ f(x) : \perp\end{matrix}}{R(\lambda x.f(x)) : \mathbf{A}}$$

[4] A justificativa para a inclusao da regra de Redução ad Absurdum pode ser encontrada em [GdQ92].

Quantificador Universal

$$\frac{\begin{array}{c}[x:\mathbf{D}]\\ \vdots \\ f(x):\mathbf{P}(x)\end{array}}{I_\forall(\lambda x.f(x)):\forall x^{\mathbf{D}}.\mathbf{P}(x)}(\forall\text{-}Intr) \qquad \frac{x:\mathbf{D} \quad c:\forall x^{\mathbf{D}}.\mathbf{P}(x)}{E_\forall(c,a):\mathbf{P}(a)}(\forall\text{-}Elim)$$

Na introdução do $\forall$, x é um elemento arbitrário qualquer.

Quantificador Existencial

$$\frac{a:\mathbf{D} \quad f(a):\mathbf{P}(a)}{I_\exists\lambda x.(f(x),a):\exists x^{\mathbf{D}}.\mathbf{P}(x)}(\exists\text{-}Intr) \qquad \frac{e:\exists x\mathbf{P}(x) \quad \begin{array}{c}[t:\mathbf{D},g(t):\mathbf{P}(t)]\\ \vdots \\ d(g,t):\mathbf{C}\end{array}}{E_\exists(e,\lambda g.\lambda t.d(g,t)):\mathbf{C}}(\exists\text{-}Elim)$$

Na eliminação do $\exists$ a restrição é que $\mathbf{C}$ é uma fórmula que não contém t e não depende de nenhuma suposição que contenha t.

Como podemos ver, as regras seguem o mesmo padrão das regras do *LND* descritas no Capítulo 4. A diferença básica é o nome das funções aqui utilizadas. O porque de não mantermos as mesmas funções da dedução natural rotulada é que, a depender do significado do rótulo, essas funções poderão ser interpretadas de diversas maneiras. A idéia é que se o rótulo representar, por exemplo, uma medida de probabilidade, a álgebra dos rótulos poderá ser especificada como um cálculo probabilístico. Assim, as funções aqui especificadas não têm, *a priori*, nenhuma interpretação, no estudo de cada caso particular é que seu significado deverá ser precisamente definido.

5.4 Etapas para a solução de um problema abdutivo

Nesta seção abordaremos o processo abdutivo via *LDS* em todas as suas etapas, etapas estas que podemos resumir da seguinte maneira:

- A tentativa de construção da prova de φ a partir da teoria Θ (o nível objeto da solução);
- A geração dos candidatos minimais para completar a prova de φ;
- A propagação dos rótulos a partir dos candidatos gerados na etapa anterior;
- A comparação dos candidatos selecionando aquele ou aqueles mais plausíveis.

5.4.1 A construção da prova e a geração dos candidatos

Os dois primeiros passos do processo abdutivo via *LDS*, a construção da prova e a geração dos candidatos, serão executados dedutivamente no sentido da conclusão para as premissas, usando-se as regras de dedução natural e considerando-se a teoria Θ. O objetivo final dessa etapa é a geração do conjunto $Cand(\Theta, \varphi)$.

Em verdade, no raciocínio abdutivo a prova de φ, o fato a ser explicado, não se completará apenas com a teoria de *background*, o que é feito é apenas a tentativa de construção da prova. Em determinado ponto da tentativa de construção da prova de φ não teremos mais elementos em Θ que nos permitam prosseguir, a não ser que a fórmula φ seja conseqüência lógica da própria teoria; nesse caso específico o problema torna-se trivial, significando que a explicação para φ é dada a partir da própria teoria e não há, assim, uma hipótese abdutiva propriamente dita. Os exemplos 5.1 e 5.2 retratam esse caso.

Exemplo 5.1 (Caso em que o fato φ é conseqüência lógica da teoria)

$$\Theta = \{l_1 : B \to C\}$$

$$\varphi = (A \wedge B) \to (A \wedge C)$$

A árvore de dedução com a tentativa da prova, seria:

$$\dfrac{\dfrac{\dfrac{[A \wedge B]}{A} \qquad \dfrac{\dfrac{[A \wedge B]}{B} \quad B \to C}{C}}{A \wedge C}}{(A \wedge B) \to (A \wedge C)}$$

A solução então seria:

$$\Theta \vdash_L B \to C \leadsto (A \wedge B) \to (A \wedge C)$$

ou

$$\Theta \vdash_L (A \wedge B) \to (A \wedge C)$$

já que a fórmula φ é dedutível a partir de Θ.

Quando não temos a explicação de φ na própria teoria Θ, a tentativa de construção da prova pára em determinado ponto onde temos uma fórmula não presente em Θ e sobre a qual não podemos aplicar nenhuma regra de eliminação ou introdução de conectivos lógicos. É essa fórmula que passamos a chamar de candidato à explicação de φ. Em alguns casos, a árvore de dedução pode ramificar-se e várias folhas podem ficar 'pendentes' (no sentido de que sua fórmula não pertence à teoria Θ e tão pouco pode-se aplicar alguma regra de introdução ou eliminação de conectivos lógicos). É nesse caso que temos que aplicar o ítem (iii) da definição 5.4 e o candidato associado àquela árvore de dedução aceitável será a conjunção das fórmulas das folhas pendentes. O exemplo 5.2 retrata o caso em que φ é explicado pela conjunção de fórmulas da teoria; o exemplo 5.3 retrata o caso em que φ é explicado pela conjunção de fórmulas que não pertencem à teoria.

Exemplo 5.2 (O fato φ é explicado pela conjunção de fórmulas da própria teoria)

$$\begin{aligned} \Theta = \{ \;\; & l_1 : \; A \to B, \\ & l_2 : \; C \to D\} \end{aligned}$$

$$\varphi = (A \wedge C) \to (B \wedge D)$$

A árvore de dedução com a tentativa da prova, seria:

$$\dfrac{\dfrac{\dfrac{\dfrac{[A \wedge C]}{A} \quad A \to B}{B} \qquad \dfrac{\dfrac{[A \wedge C]}{C} \quad C \to D}{D}}{B \wedge D}}{(A \wedge C) \to (B \wedge D)}$$

A solução então seria:

$$\Theta \vdash_L (A \to B) \wedge (C \to D) \rightsquigarrow (A \wedge C) \to (B \wedge D)$$

ou

$$\Theta \vdash_L (A \wedge C) \to (B \wedge D)$$

Exemplo 5.3 (O fato φ é explicado pela conjunção de fórmulas)

$$\begin{aligned} \Theta = \{ \;\; & l_1 : \; A, \\ & l_2 : \; A \wedge B \to E, \\ & l_3 : \; C \to D \wedge J, \\ & l_4 : \; D \wedge E \to F, \\ & l_5 : \; B \vee G \to F, \\ & l_6 : \; H \to F\} \end{aligned}$$

No entanto, a geração de todos os candidatos (minimais e não minimais) passíveis de serem a explicação de φ é perfeitamente possível através da generalização do conjunto $Cand(\Theta,\varphi)$. Em [CDT91] Console et al. faz menção ao processo de generalização de candidatos considerando que as explicações preferenciais (com respeito à minimalidade) podem ser entendidas como geradoras do conjunto de todas as explicações. O procedimento de generalizar candidatos pode ser útil em determinados casos específicos (como veremos a seguir) e pode ser definido da seguinte maneira:

Definição 5.7 (Generalização do conjunto de candidatos) *Seja $Cand(\Theta,\varphi)$ o conjunto de candidatos minimais para o problema abdutivo $\langle\Theta,\varphi,\sqsubseteq\rangle$. O conjunto generalizado dos candidatos para $\langle\Theta,\varphi,\sqsubseteq\rangle$, $Gen(Cand(\Theta,\varphi))$ é definido da seguinte maneira:*

(i) $\alpha \in Gen(Cand(\Theta,\varphi))$, *para todo* $\alpha \in Cand(\Theta,\varphi)$;

(ii) $\gamma \rightarrow \varphi \in Gen(Cand(\Theta,\varphi))$ *se* $\gamma \in \Theta$ *ou se* $\Theta \vdash_L \gamma$;

(iii) $\gamma \wedge \alpha \in Gen(Cand(\Theta,\varphi))$, *para todo* $\alpha \in Cand(\Theta,\varphi)$, *para todo* γ *Θ-consistente;*

(iv) $\gamma \rightarrow \alpha \in Gen(Cand(\Theta,\varphi))$ *se* $\gamma \in \Theta$ *ou se* $\Theta \vdash_L \gamma$, *para todo* $\alpha \in Cand(\Theta,\varphi)$;

(v) $\gamma \in Gen(Cand(\Theta,\varphi))$, *se* γ *for gerada a partir da aplicação dos ítens (iii) ou (iv) acima, sendo* $\alpha \in Gen(Cand(\Theta,\varphi))$.

Teorema 5.2 (Corretude do procedimento de generalização de candidatos) *Todos os elementos de $Gen(Cand(\Theta,\varphi))$ são candidatos à explicação de φ.*

Prova

Todos os membros de $Gen(Cand(\Theta,\varphi))$ são formados de uma das 5 maneiras vistas na definição 5.7. Então, para cada um dos casos, temos:

(*i*) $\alpha \in Gen(Cand(\Theta,\varphi))$, para todo $\alpha \in Cand(\Theta,\varphi)$;

Se $\alpha \in Cand(\Theta,\varphi)$, logo α é um candidato.

(*ii*) $\gamma \rightarrow \varphi \in Gen(Cand(\Theta,\varphi))$ se $\gamma \in \Theta$ ou se $\Theta \vdash_L \gamma$;

Como $\Theta \vdash_L \gamma$, então $\Theta \cup \{\gamma \rightarrow \varphi\} \vdash_L \varphi$, logo, $\gamma \rightarrow \varphi$ é um candidato.

(iii) $\gamma \wedge \alpha \in Gen(Cand(\Theta, \varphi))$, para todo $\alpha \in Cand(\Theta, \varphi), \forall\gamma\Theta$-consistente;

Se $\alpha \in Cand(\Theta, \varphi)$, então $\Theta \cup \{\alpha\} \vdash_L \varphi$; mas como $\gamma \wedge \alpha \rightarrow \alpha$, então $\Theta \cup \{\gamma \wedge \alpha\} \vdash_L \varphi$, logo, $\gamma \wedge \alpha$ é um candidato.

(iv) $\gamma \rightarrow \alpha \in Gen(Cand(\Theta, \varphi))$ se $\gamma \in \Theta$ ou se $\Theta \vdash_L \gamma$, para todo $\alpha \in Cand(\Theta, \varphi)$;

Se $\alpha \in Cand(\Theta, \varphi)$, então $\Theta \cup \{\alpha\} \vdash_L \varphi$; mas, por hipótese, $\Theta \vdash_L \gamma$, logo $\Theta \cup \{\gamma \rightarrow \alpha\} \vdash_L \varphi$ portanto, $\gamma \rightarrow \alpha$ é um candidato.

(v) Todos os elementos de $Gen(Cand(\Theta, \varphi))$ são candidatos, logo, para todo $\gamma \in Gen(Cand(\Theta, \varphi))$, $\Theta \cup \{\gamma\} \vdash_L \varphi$. Mas essa é a mesma característica apresentada pelos elementos de $Cand(\Theta, \varphi)$, logo, como foi demonstrado acima, se γ é gerada pela aplicação das regras (iii) ou (iv) a elementos de $Gen(Cand(\Theta, \varphi))$, então γ é candidato. □

Eventualmente pode-se ter problemas abdutivos onde, dentro da teoria adotada, apenas um único candidato pode ser gerado. Num caso como esse, os passos seguintes do processo abdutivo (a propagação dos rótulos e a comparação dos candidatos) tornam-se desnecessários e a solução dada ao problema é (como não poderia deixar de ser) o candidato único gerado no processo. O exemplo 5.4 enquadra-se nesse caso.

Exemplo 5.4 (Caso em que temos um único candidato gerado pela teoria)

$$\begin{aligned} \Theta = \{ \quad & l_1: \; A \rightarrow B, \\ & l_2: \; A, \\ & l_3: \; E \rightarrow C \wedge D\} \end{aligned}$$

E a fórmula a ser explicada, $\varphi = D$.

A tentativa de provar D *a partir da teoria* Θ *nos dá a seguinte árvore de dedução:*

$$\dfrac{\dfrac{E \quad E \rightarrow C \wedge D}{C \wedge D}}{D}$$

Nesse caso, como temos uma única árvore de dedução aceitável, o nosso conjunto de candidatos possui um único elemento, E*, que é, portanto, a única solução possível para o referido problema. Então:*

$$Cand(\Theta, D) = \{E\}$$

$$\Theta \vdash_L E \leadsto D$$

Tratando-se de um método abdutivo genérico, devemos considerar também a possibilidade de os candidatos serem fornecidos pelo usuário. Gerando candidatos a partir da teoria Θ ficamos limitados ao que nos diz a teoria, e em alguns casos isso pode levar-nos a poucas opções de solução como no caso do exemplo 5.4 onde temos um único candidato gerado a partir da teoria. Pode haver casos em que seja viável o fornecimento de candidatos à explicação da fórmula φ pelo próprio usuário e então a abdução pode tornar-se mais passível de estar correta, já que não estaremos limitados à teoria.

Num problema caracterizado dessa maneira, a geração de candidatos a partir da tentativa de construção da prova deve acontecer afim de que se possa confirmar que o conjunto de candidatos fornecido pelo usuário é realmente composto de elementos plausíveis de serem uma explicação para o fato φ. Para isso devem ser executados os passos de construção da prova e geração dos candidatos e, em seguida, deve ser executada a generalização dos candidatos como descrita na definição 5.7 acima. Se os candidatos fornecidos, que chamaremos de $CandF(\Theta, \varphi)$, forem também elementos do conjunto generalizado de candidatos, $Gen(Cand(\Theta, \varphi))$, então cada $\alpha \in CandF(\Theta, \varphi)$ é realmente um candidato logicamente plausível.

Feito isso, passaríamos então, ao passo seguinte de propagação dos rótulos. Nesse caso, vale ressaltar que essas fórmulas (representando os candidatos) também deverão estar rotuladas, seguindo o mesmo critério meta-lógico da teoria Θ. E a ordenação seria definida para todos os rótulos do problema (os da teoria e os dos candidatos). Isso deve acontecer porque, quando da comparação dos candidatos, estaremos lidando com hipóteses abdutivas fornecidas pelo usuário ($CandF(\Theta, \varphi)$); nos parece óbvio que o fornecimento dessas hipóteses tenha se dado devido a alguma crença nas mesmas (se fossem fornecidos candidatos aleatoriamente, o próprio sistema poderia fazê-lo). Logo, a informação adicional (do meta-nível) contida nos rótulos deve ser fornecida também para os elementos de $CandF(\Theta, \varphi)$. Além disso, como a comparação dos candidatos se dará de acordo com os passos de dedução utilizados na construção da prova (como veremos na seção 5.4.3) e como os rótulos dos candidatos fornecidos serão tratados juntamente com os rótulos da teoria, a mesma interpretação deve ser dada aos dois grupos de rótulos, assim como a ordem entre os rótulos, também fornecida pelo usuário, deve englobar ambos os grupos.

O exemplo 5.5 trata o mesmo problema abdutivo do exemplo 5.4 (onde apenas um candidato foi gerado a partir da teoria Θ), agora considerando candidatos fornecidos pelo usuário. A comparação desses dois exemplos nos dá claramente um caso onde o fornecimento de candidatos pelo usuário (quando é viável) leva-nos a uma proposta de solução muito mais plausível.

Exemplo 5.5 (Caso em que temos um conjunto de candidatos fornecido pelo usuário)

$$\Theta = \{ \; l_1 : \; A \rightarrow B, \quad l_2 : \; A, \quad l_3 : \; E \rightarrow C \wedge D\} \qquad CandF(\Theta, D) = \{ \; l_4 : B \rightarrow E, \; l_5 : A \rightarrow E, \; l_6 : B \rightarrow D, \; l_7 : E\}$$

E a fórmula a ser explicada: $\varphi = D$.

A geração dos candidatos é idêntica à do exemplo 5.4. Então temos o seguinte conjunto de candidatos gerados:

$$Cand(\Theta, D) = \{E\}$$

Agora passamos à generalização dos candidatos:

(*i*) $\gamma \in Gen(Cand(\Theta, D))$ *se* $\gamma \in Cand(\Theta, D)$*;*

(*ii*) $\gamma \rightarrow D \in Gen(Cand(\Theta, D))$ *se* $\gamma \in Cand(\Theta, D)$ *ou se* $\Theta \vdash_L \gamma$*;*

(*iii*) $\gamma \wedge \alpha \in Gen(Cand(\Theta, D))$ *se* γ *é* Θ*-consistente e* $\alpha \in Cand(\Theta, D)$*, logo* $\gamma \wedge E \in Gen(Cand(\Theta, D))$ *se* γ *é* Θ*-consistente;*

(*iv*) $\gamma \rightarrow \alpha \in Gen(Cand(\Theta, D))$ *se* $\gamma \in \Theta$ *ou* $\Theta \vdash_L \gamma$ *e* $\alpha \in Cand(\Theta, D)$*, logo* $\gamma \rightarrow E \in Gen(Cand(\Theta, D))$ *se* $\gamma \in \Theta$ *ou se* $\Theta \vdash_L \gamma$*;*

Verificando se os elementos de $CandF(\Theta, D)$ *são realmente candidatos, temos:*

1. *Como* $\Theta \vdash_L B$*, então* $B \rightarrow E$ *é candidato pois se enquadra no caso (iv) acima;*
2. *Como* $A \in \Theta$*, então* $A \rightarrow E$ *é candidato pois se enquadra no caso (iv) acima;*
3. *Como* $\Theta \vdash_L B$*, então* $B \rightarrow D$ *é candidato pois se enquadra no caso (ii) acima;*
4. *Como* $E \in Cand(\Theta, D)$*, então* E *é candidato pois se enquadra no caso (i) acima.*

Verificado que os candidatos fornecidos pertencem a $Gen(Cand(\Theta, D))$*, seguimos para a construção das provas:*

Para o candidato $l_4 : B \rightarrow E$ temos então a seguinte árvore de dedução:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \quad A \rightarrow B}{B} \quad B \rightarrow E}{E} \quad E \rightarrow C \land D}{C \land D}}{D}$$

Para o candidato $l_5 : A \rightarrow E$ temos:

$$\dfrac{\dfrac{\dfrac{A \quad A \rightarrow E}{E} \quad E \rightarrow C \land D}{C \land D}}{D}$$

Para o candidato $l_6 : B \rightarrow D$ temos:

$$\dfrac{\dfrac{A \quad A \rightarrow B}{B} \quad B \rightarrow D}{D}$$

Para o candidato $l_7 : E$, temos a mesma árvore mostrada anteriormente quando da geração do candidato:

$$\dfrac{\dfrac{E \quad E \rightarrow C \land D}{C \land D}}{D}$$

Então, nosso conjunto de candidatos será $Cand(\Theta, D) = \{B \rightarrow E, A \rightarrow E, B \rightarrow D, E\}$.

Na maioria dos problemas abdutivos, como já foi dito, mais de um candidato é gerado como possível solução. No exemplo 5.6 temos a construção de cinco possíveis árvores de prova com a geração de, respectivamente, cinco candidatos:

Exemplo 5.6 (Exemplo com geração de mais de um candidato) *Suponhamos a teoria Θ composta das seguintes fórmulas rotuladas:*

$$\begin{aligned} \Theta = \{ \ & l_1 : \ A \rightarrow D, \\ & l_2 : \ E \rightarrow C \land D, \\ & l_3 : \ B \lor F \rightarrow D\} \end{aligned}$$

E o fato observado $\varphi = D$.

O primeiro passo, a tentativa de construção da prova de φ a partir de Θ (no sentido da conclusão para as premissas) nos dá as cinco árvores de dedução seguintes:

$$\frac{A \quad A \rightarrow D}{D}$$

$$\frac{\dfrac{E \quad E \rightarrow C \wedge D}{C \wedge D}}{D}$$

$$\frac{B \vee F \quad B \vee F \rightarrow D}{D}$$

$$\frac{\dfrac{B}{B \vee F} \quad B \vee F \rightarrow D}{D}$$

$$\frac{\dfrac{F}{B \vee F} \qquad B \vee F \rightarrow D}{D}$$

Considerando as 5 (cinco) possibilidades, nosso conjunto de candidatos é:

$$\begin{aligned} Cand(\Theta, D) = \{ \ & A, \\ & E, \\ & B \vee F, \\ & B, \\ & F\} \end{aligned}$$

Aplicando-se o critério de minimalidade a nível objeto, temos que:

$$B \rightarrow B \vee F$$

$$F \rightarrow B \vee F$$

logo, podemos excluir os candidatos B e F, por não serem minimais com respeito à conseqüência lógica, e reduzir o conjunto $Cand(\Theta, D)$ a apenas elementos minimais:

$$\begin{aligned} min(Cand(\Theta, D)) = \{ \ & A, \\ & E, \\ & B \vee F\} \end{aligned}$$

Na próxima seção discutiremos a propagação dos rótulos e seguiremos na busca da solução para o exemplo acima.

Com respeito à minimalidade dos candidatos, como discutido na seção 5.2.1, podemos ter casos em que, ainda a nível objeto, podemos excluir determinados candidatos por não serem minimais com respeito à relação de conseqüência. Ou seja, se dentro do conjunto $Cand(\Theta, \varphi) = \{\alpha_1, ..., \alpha_n\}$ para algum i e j menores que n, se $\alpha_i \rightarrow \alpha_j$, então podemos excluir o elemento α_i por não ser minimal. Os exemplos 5.6 e 5.7 mostram casos onde um elemento não mininal é gerado e deve ser excluído do conjunto de candidatos.

Exemplo 5.7 (O critério de minimalidade a nível objeto é aplicado)

$$\begin{aligned} \Theta = \{ \quad l_1 : &\quad A \wedge C \rightarrow D, \\ l_2 : &\quad A \rightarrow B \wedge E, \\ l_3 : &\quad E \rightarrow D\} \end{aligned}$$

$$\varphi = D$$

As árvores de dedução com as tentativas da prova, seriam:

$$\frac{A \wedge C \qquad A \wedge C \rightarrow D}{D}$$

$$\frac{E \qquad E \rightarrow D}{D}$$

$$\cfrac{\cfrac{\cfrac{A \qquad A \rightarrow B \wedge E}{B \wedge E}}{E} \qquad E \rightarrow D}{D}$$

O nosso conjunto de candidatos é então:

$$\begin{aligned} Cand(\Theta, \varphi) = \{ \quad & A \wedge C, \\ & E, \\ & A\} \end{aligned}$$

Pelo critério de minimalidade, como $A \wedge C \rightarrow A$, então $A \wedge C$ não é um candidato minimal e deve ser excluído. Ficamos então com:

$$\begin{aligned} min(Cand(\Theta, \varphi)) = \{ \quad & E, \\ & A\} \end{aligned}$$

Com respeito à consistência com a teoria, podemos ter casos em que os candidatos à solução de um problema do tipo $\langle \Theta, \varphi, \sqsubseteq \rangle$ sejam elementos que negam a própria teoria. Nesse caso a solução do problema é indefinida e isso pode ser verificado através da geração dos candidatos que negam a teoria Θ ($Cand(\Theta, \bot)$, como discutido na seção 5.2.1); o exemplo 5.8 ilustra esse caso.

Exemplo 5.8 (Caso em que o candidato é também uma fórmula que nega a teoria)

$$\begin{aligned} \Theta = \{ \quad l_1 : &\quad \neg E \rightarrow D, \\ l_2 : &\quad E \rightarrow \neg D, \\ l_3 : &\quad E\} \end{aligned}$$

$$\varphi = D$$

A árvore de dedução com a tentativa da prova, seria:

$$\frac{\neg E \qquad \neg E \rightarrow D}{D}$$

O nosso conjunto de candidatos é então:

$$Cand(\Theta, \varphi) = \{ \quad \neg E \ \}$$

Mas se buscarmos os candidatos que negam a teoria Θ, temos:

$$\frac{\neg E \qquad E}{\bot}$$

$$\frac{\dfrac{E \quad E \rightarrow \neg D}{\neg D} \qquad D}{\bot}$$

E temos então:

$$\begin{aligned} Cand(\Theta, \bot) = \{ \quad & \neg E, \\ & D \ \} \end{aligned}$$

Logo, $\neg E$ é um candidato que nega a teoria e por isso não pode ser considerado como solução para um problema abdutivo utilizando a teoria Θ. Então o problema $\langle \Theta, D, \sqsubseteq \rangle$ não tem solução definida.

Outra possibilidade de inconsistência é quando $\Theta \vdash_L \neg\varphi$, e isso também pode ser verificado dedutivamente. O exemplo 5.9 retrata esse caso de inconsistência.

Exemplo 5.9 (φ não é Θ-consistente)

$$\begin{aligned} \Theta = \{ \quad l_1 : & \quad B \rightarrow D, \\ l_2 : & \quad E \rightarrow \neg D, \\ l_3 : & \quad E\} \end{aligned}$$

$$\varphi = D$$

A árvore de dedução com a tentativa da prova, seria:

$$\frac{B \qquad B \rightarrow D}{D}$$

O nosso conjunto de candidatos é então:

$$Cand(\Theta, \varphi) = \{ \quad B \ \}$$

Mas temos que:

$$\frac{E \qquad E \rightarrow \neg D}{\neg D}$$

logo

$$\Theta \vdash_L \neg D$$

O que implica que o problema abdutivo $\langle \Theta, D, \sqsubseteq \rangle$ não tem solução definida porque D é inconsistente com a teoria Θ.

Outra maneira de solucionar este problema abdutivo seria construindo o conjunto $Cand(\Theta, \perp)$ e verificando que $D \in Cand(\Theta, \perp)$, como no exemplo 5.8.

Outra questão importante é com relação às restrições de integridade. Em alguns casos, a adição de restrições de integridade pode restringir o conjunto de candidatos, fazendo com que a solução proposta seja ainda mais plausível. No caso do exemplo 5.10 abaixo, os 'sapatos molhados' podem ter como causa tanto a chuva quanto o fato de que o molhador de grama estava ligado:

Exemplo 5.10 (O exemplo dos sapatos molhados [CDT91])

$$\begin{aligned} \Theta = \{ \quad & l_1 : \quad \textit{grass-is-wet} \rightarrow \textit{shoes-are-wet}, \\ & l_2 : \quad \textit{rained-last-night} \rightarrow \textit{grass-is-wet}, \\ & l_3 : \quad \textit{sprinkler-was-on} \rightarrow \textit{grass-is-wet}\} \end{aligned}$$

$$\varphi = \textit{shoes-are-wet}$$

ou

$$\begin{aligned} \Theta = \{ \quad & l_1 : \quad G \rightarrow S, \\ & l_2 : \quad R \rightarrow G, \\ & l_3 : \quad SP \rightarrow G\} \end{aligned}$$

$$\varphi = S$$

As árvores de dedução com as tentativas da prova, seriam:

$$\dfrac{\dfrac{R \qquad R \rightarrow G}{G} \qquad G \rightarrow S}{D}$$

$$\dfrac{\dfrac{SP \qquad SP \rightarrow G}{G} \qquad G \rightarrow S}{D}$$

O nosso conjunto de candidatos é então:

$$\begin{aligned} Cand(\Theta, \varphi) = \{ \quad & R, \\ & SP \ \} \end{aligned}$$

Se estendermos um pouco a nossa teoria acrescentando o fato *couldless-sky* e apresentarmos uma restrição de integridade do tipo $\neg(\textit{rained} \wedge \textit{couldless-sky})$ chegaremos à conclusão que a única explicação plausível para os 'sapatos molhados' é que o molhador de grama estava ligado. O exemplo 5.11 ilustra esse caso.

Exemplo 5.11 (O exemplo dos sapatos molhados com restrição de integridade [Ton95])

$$\begin{aligned} \Theta = \{ \quad & l_1 : \quad \textit{grass-is-wet} \rightarrow \textit{shoes-are-wet}, \\ & l_2 : \quad \textit{rained-last-night} \rightarrow \textit{grass-is-wet}, \\ & l_3 : \quad \textit{sprinkler-was-on} \rightarrow \textit{grass-is-wet}, \\ & l_4 : \quad \textit{coudless-sky}\} \end{aligned}$$

$$\varphi = \textit{shoes-are-wet}$$

E a restrição de integridade:

$$I = \{l_5 : \neg(\textit{rained-last-night} \wedge \textit{couldless-sky})\}$$

ou

$$\begin{aligned}\Theta = \{ \quad & l_1 : \quad G \rightarrow S, \\ & l_2 : \quad R \rightarrow G, \\ & l_3 : \quad SP \rightarrow G, \\ & l_4 : \quad C\}\end{aligned}$$

$$\varphi = S$$

$$I = \{l_5 : \neg(R \wedge C)\}$$

As árvores de dedução com as tentativas da prova, seriam as mesmas do exemplo anterior:

$$\cfrac{\cfrac{R \qquad R \rightarrow G}{G} \qquad G \rightarrow S}{D}$$

$$\cfrac{\cfrac{SP \qquad SP \rightarrow G}{G} \qquad G \rightarrow S}{D}$$

E o conjunto de candidatos também seria o mesmo:

$$\begin{aligned}Cand(\Theta, \varphi) = \{ \quad & R, \\ & SP \ \}\end{aligned}$$

Mas se gerarmos os candidatos que negam a teoria Θ*, considerando a restrição de integridade proposta, temos:*

$$\cfrac{\cfrac{R \quad C}{R \wedge C} \qquad \neg(R \wedge C)}{\perp} \qquad \cfrac{\neg C \qquad C}{\perp}$$

logo:

$$Cand(\Theta, \perp) = \{R, \neg C\}$$

O exemplo 5.12 mostra a propagação dos rótulos para o problema do exemplo 5.6.

Exemplo 5.12 (Propagação dos rótulos - continuação do exemplo 5.6) *Para* $\alpha = \mathbf{A}$, *e supondo* $\mathbf{A}$ *rotulado por um rótulo hipotético* x, *temos:*

$$\frac{x : \mathbf{A} \quad l_1 : \mathbf{A} \rightarrow \mathbf{D}}{E_{\rightarrow}(x, l_1) : \mathbf{D}}$$

Para $\alpha = \mathbf{E}$, *supondo* $\mathbf{E}$ *rotulado por* y:

$$\frac{\dfrac{y : \mathbf{E} \quad l_2 : \mathbf{E} \rightarrow \mathbf{C} \wedge \mathbf{D}}{E_{\rightarrow}(y, l_2) : \mathbf{C} \wedge \mathbf{D}}}{E_{2\wedge}(E_{\rightarrow}(y, l_2)) : \mathbf{D}}$$

Para $\alpha = \mathbf{B} \vee \mathbf{F}$, *supondo* $\mathbf{B} \vee \mathbf{F}$ *rotulado por* z:

$$\frac{z : \mathbf{B} \vee \mathbf{F} \quad l_3 : \mathbf{B} \vee \mathbf{F} \rightarrow \mathbf{D}}{E_{\rightarrow}(z, l_3) : \mathbf{D}}$$

Para $\alpha = \mathbf{B}$ *(se não houvéssemos eliminado esse candidato pelo critério da minimalidade), supondo* $\mathbf{B}$ *rotulado por* t:

$$\frac{\dfrac{t : \mathbf{B}}{I_{1\vee}(t) : \mathbf{B} \vee \mathbf{F}} \quad l_3 : \mathbf{B} \vee \mathbf{F} \rightarrow \mathbf{D}}{E_{\rightarrow}(I_{1\vee}(t), l_3) : \mathbf{D}}$$

Para $\alpha = \mathbf{F}$ *(se não houvéssemos eliminado esse candidato pelo critério da minimalidade), supondo* $\mathbf{F}$ *rotulado por* w:

$$\frac{\dfrac{w : \mathbf{F}}{I_{2\vee}(w) : \mathbf{B} \vee \mathbf{F}} \quad l_3 : \mathbf{B} \vee \mathbf{F} \rightarrow \mathbf{D}}{E_{\rightarrow}(I_{2\vee}(w), l_3) : \mathbf{D}}$$

Nesse ponto temos 5 (cinco) possíveis suportes para a fórmula $\mathbf{D}$. *Como já tínhamos eliminado os candidatos* $\mathbf{F}$ *e* $\mathbf{B}$, *passaremos ao passo seguinte apenas considerando os três candidatos restantes e os respectivos suportes encontrados para a fórmula* $\mathbf{D}$. *São eles:*

$$Spt(\mathbf{D}, \mathbf{A}) = E_{\rightarrow}(x, l_1)$$

$$Spt(\mathbf{D}, \mathbf{E}) = E_{2\wedge}(E_{\rightarrow}(y, l_2))$$

$$Spt(\mathbf{D}, \mathbf{B} \vee \mathbf{F}) = E_{\rightarrow}(z, l_3)$$

No caso em que temos candidatos fornecidos pelo usuário, a propagação dos rótulos se dá de maneira similar. A diferença reside em que teríamos candidatos rotulados **não** hipoteticamente, já que o fornecimento de candidatos prevê também a especificação de rótulos reais (segundo o mesmo critério de preferência meta-lógico) para cada candidato fornecido. Nesse caso, os suportes encontrados para a fórmula φ em cada árvore de dedução só conteriam rótulos reais, o que, inclusive, facilitaria no momento da comparação dos mesmos. A propagação de rótulos com candidatos fornecidos pelo usuário pode ser vista no exemplo 5.13 que mostra a propagação dos rótulos para o problema abdutivo do exemplo 5.5.

Exemplo 5.13 (Propagação dos rótulos para o exemplo 5.5) *Para o candidato* $l_4 : \mathbf{B} \rightarrow \mathbf{E}$ *temos a seguinte propagação dos rótulos:*

$$\cfrac{\cfrac{\cfrac{\cfrac{l_2 : \mathbf{A} \quad l_1 : \mathbf{A} \rightarrow \mathbf{B}}{E_{\rightarrow}(l_2, l_1) : \mathbf{B}} \quad l_4 : \mathbf{B} \rightarrow \mathbf{E}}{E_{\rightarrow}(E_{\rightarrow}(l_2, l_1), l_4) : \mathbf{E}} \quad l_3 : \mathbf{E} \rightarrow \mathbf{C} \wedge \mathbf{D}}{E_{\rightarrow}(E_{\rightarrow}(E_{\rightarrow}(l_2, l_1), l_4), l_3) : \mathbf{C} \wedge \mathbf{D}}}{E_{2\wedge}(E_{\rightarrow}(E_{\rightarrow}(E_{\rightarrow}(l_2, l_1), l_4), l_3)) : \mathbf{D}}$$

Nesse primeiro caso obtemos o $Spt(\mathbf{D}, \mathbf{B} \rightarrow \mathbf{E}) = E_{2\wedge}(E_{\rightarrow}(E_{\rightarrow}(E_{\rightarrow}(l_2, l_1), l_4), l_3))$.

Para o candidato $l_5 : \mathbf{A} \rightarrow \mathbf{E}$ *temos:*

$$\cfrac{\cfrac{\cfrac{l_2 : \mathbf{A} \quad l_5 : \mathbf{A} \rightarrow \mathbf{E}}{E_{\rightarrow}(l_2, l_5) : \mathbf{E}} \quad l_3 : \mathbf{E} \rightarrow \mathbf{C} \wedge \mathbf{D}}{E_{\rightarrow}(E_{\rightarrow}(l_2, l_5), l_3) : \mathbf{C} \wedge \mathbf{D}}}{E_{2\wedge}(E_{\rightarrow}(E_{\rightarrow}(l_2, l_5), l_3)) : \mathbf{D}}$$

No segundo caso obtemos $Spt(\mathbf{D}, \mathbf{A} \rightarrow \mathbf{E}) = E_{2\wedge}(E_{\rightarrow}(E_{\rightarrow}(l_2, l_5), l_3))$.

Para o candidato $l_6 : \mathbf{B} \rightarrow \mathbf{D}$ *temos:*

$$\cfrac{\cfrac{l_2 : \mathbf{A} \quad l_1 : \mathbf{A} \rightarrow \mathbf{B}}{E_{\rightarrow}(l_2, l_1) : \mathbf{B}} \quad l_6 : \mathbf{B} \rightarrow \mathbf{D}}{E_{\rightarrow}(E_{\rightarrow}(l_2, l_1), l_6) : \mathbf{D}}$$

Nesse caso temos $Spt(\mathbf{D}, \mathbf{B} \rightarrow \mathbf{D}) = E_{\rightarrow}(E_{\rightarrow}(l_2, l_1), l_6)$.

Para o candidato $l_7 : \mathbf{E}$ *temos:*

$$\cfrac{\cfrac{l_7 : \mathbf{E} \quad l_3 : \mathbf{E} \rightarrow \mathbf{C} \wedge \mathbf{D}}{E_{\rightarrow}(l_7, l_3) : \mathbf{C} \wedge \mathbf{D}}}{E_{2\wedge}(E_{\rightarrow}(l_7, l_3)) : \mathbf{D}}$$

E temos então $Spt(\mathbf{D}, \mathbf{E}) = E_{2\wedge}(E_{\rightarrow}(l_7, l_3))$.

5.4.3 A comparação dos candidatos

Escolher entre os diversos candidatos é o ponto crucial da abdução. Antes de mais nada, não poderemos garantir que o candidato escolhido seja realmente o melhor, já que o melhor pode não ser alcançável através do conhecimento de *background* que temos. Feita essa ressalva, vamos então escolher o melhor dentre aqueles que podem ser a explicação de φ, segundo as informações que temos. Mas a característica de ser o melhor é muito subjetiva, e normalmente depende da área de aplicação do problema. Esse é o principal motivo para que a ordenação das fórmulas da teoria seja fornecida pelo usuário. Mas somente isso não é suficiente. Quando propagamos os rótulos dos diversos candidatos que temos disponíveis, essa propagação está relacionada com uma série de passos de dedução. É natural que se pense que a depender da regra de dedução que estejamos aplicando e consequentemente dos rótulos que viemos propagando desde o topo da árvore de dedução, o suporte da fórmula φ retratará uma maior ou menor plausibilidade de ser esse o candidato (a partir do qual a dedução foi feita) que melhor explica φ. A grande questão é como interpretar os diversos suportes que encontramos para φ, a depender do candidato que usamos na dedução.

A grande vantagem desse passo da solução é que quando da geração dos candidatos, obteremos um conjunto com uma variedade de elementos, e, na maioria dos casos, não é possível reduzir satisfatoriamente esse conjunto apenas com os critérios lógicos. Fazendo a comparação dos candidatos utilizando a ordem preferencial dos rótulos fornecida pelo usuário estamos dando um passo à frente no sentido de que trazemos o meta-nível da abdução para ser tratado ao lado do nível objeto, e conseguimos, assim, reduzir ainda mais o nosso conjunto de candidatos, o que claramente nos leva a uma melhor solução para o problema: ainda que nem sempre consigamos reduzir o nosso conjunto $Cand(\Theta, \varphi)$ a um único elemento, certamente forneceremos uma solução com um conjunto de candidatos reduzido e contendo os candidatos mais plausíveis, meta-logicamente falando.

Acreditamos que a questão de comparar os candidatos é um problema dependente da área de aplicação e por isso a nossa proposta é que essa comparação não seja especificada de forma totalmente rígida. Deveremos ter uma variedade de opções de comparação genéricas e o usuário irá escolher aquelas que considere mais adequadas ao tipo de problema com o qual está lidando.

Uma abordagem para a abdução que apresenta um aspecto semelhante ao aqui proposto como comparação dos candidatos, é encontrada em [MD94] (veja seção 3.4.6). No referido artigo os autores propõem um nível heurístico onde as soluções geradas são comparadas; o que aqui chamamos de critérios de comparação, Mathieu e Delahaye chamam de heurísticas.

Vale lembrar que nesse ponto do processo abdutivo todos os critérios a nível objeto (consistência, minimalidade, restrições de integridade) já foram aplicados. Estamos lidando agora apenas com o critério meta-lógico, objetivando dar preferência a determinado candidato com base na ordem que definiremos entre os suportes.

A seguir descreveremos cinco critérios de comparação de rótulos que podem ser utilizados independentemente do domínio do problema. São eles: critério de unificação, critério da ordem preferencial dos rótulos, critério do número de elementos, critério do tipo de rótulo e critério das conseqüências dedutivas.

Critério de unificação

O primeiro critério de comparação que deve ser aplicado é o **critério de unificação.** Por esse critério devemos tentar eliminar funções e rótulos atômicos[5] iguais para reduzirmos o máximo possível o número de elementos dos suportes.

Formalmente, o critério de unificação pode ser definido da seguinte maneira:

Definição 5.9 (Critério de unificação) *Sejam* $s_1 = Spt(\varphi, \alpha_1) = f_1(t_{11}, ..., t_{1n})$ *e* $s_2 = Spt(\varphi, \alpha_2) = f_2(t_{21}, ..., t_{2m})$ *dois suportes associados a um problema abdutivo* $\langle \Theta, \varphi, \sqsubseteq \rangle$. *Se* $f_1 = f_2$, *ou seja, se os símbolos de função* f_1 *e* f_2 *são os mesmos, então o critério de unificação pode ser aplicado da seguinte maneira:*

1. *Reduza o suporte* s_1 *a* $red(s_1) = (t_{11}, ..., t_{1n})$ *e* s_2 *a* $red(s_2) = (t_{21}, ..., t_{2m})$.

2. *Para todo* i *:*

   - *se* $t_{1i} = t_{2i}$ *então faça* $red(s_1) = (t_{1i+1}, ..., t_{1n})$ *e* $red(s_2) = (t_{2i+1}, ..., t_{2m})$ *e continue no passo 2;*
   - *se* t_{1i} *e* t_{2i} *são rótulos hipotéticos então faça* $red(s_1) = (t_{1i+1}, ..., t_{1n})$ *e* $red(s_2) = (t_{2i+1}, ..., t_{2m})$ *e continue no passo 2;*
   - $t_{1i} \neq t_{2i}$ *então interrompa a aplicação do critério de unificação e* $red(s_1)$ *e* $red(s_2)$ *devem agora ser comparados por outro método.*

[5]Estamos chamando de rótulos atômicos aqueles que rotulam as fórmulas da teoria, ou os rótulos hipotéticos.

Por exemplo, quando temos os seguintes suportes para comparar:

$$s_1 = Spt(\varphi, \alpha_1) = E_{\rightarrow}(I_{\wedge}(l_1, l_2)$$

$$s_2 = Spt(\varphi, \alpha_2) = E_{\rightarrow}(I_{\wedge}(l_1, l_3)$$

podemos reduzir a nossa comparação aos termos l_2 e l_3 já que podemos eliminar por unificação as funções $E_{\rightarrow}$ e $I_{\wedge}$ e o rótulo atômico l_1.

A utilidade e coerência do critério de unificação é facilmente percebida: já que o objetivo da comparação dos rótulos é definir uma ordem de plausibilidade entre os candidatos gerados onde o que consideramos relevante para essa ordem são as regras de dedução aplicadas na tentativa de construção da prova e a ordem preferencial dos rótulos fornecida pelo usuário, se temos suportes cujos elementos são iguais, é porque alguns passos das respectivas deduções foram os mesmos; nada mais natural que desconsiderar esses passos iguais e direcionar a comparação apenas àqueles elementos que demonstram as diferenças nas deduções.

Critério da ordem preferencial dos rótulos atômicos

Quando o critério de unificação pode ser utilizado e, após a sua utilização, se os suportes ficam reduzidos a apenas conjuntos de rótulos atômicos do mesmo tamanho, podemos aplicar diretamente a **ordem preferencial dos rótulos atômicos**, que pode ser definido formalmente da seguinte maneira:

Definição 5.10 (Critério da ordem preferencial dos rótulos) *Sejam $s_1 = Spt(\varphi, \alpha_1) = (t_{11}, ..., t_{1n})$ e $s_2 = Spt(\varphi, \alpha_2) = (t_{21}, ..., t_{2n})$ dois suportes (ou suportes reduzidos) associados a um problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$ onde as seguintes condições são verdadeiras:*

- *$t_{11}, ..., t_{1n}$ são rótulos atômicos;*
- *$t_{21}, ..., t_{2n}$ são rótulos atômicos.*

Dizemos que $s_1 \sqsubseteq s_2$ se para todo j uma das condições abaixo é verdadeira, e a condição (i) é satisfeita pelo menos uma vez:

(i) $t_{1j} \sqsubseteq t_{2j}$, na ordem fornecida pelo usuário;

(ii) $t_{1j} \sqsubseteq t_{2j} \sqsubseteq t_{1j}$, *na ordem fornecida pelo usuário;*

(iii) t_{1j} *e* t_{2j} *são rótulos hipotéticos.*

Se para todo j uma das condições (ii) ou (iii) são verdadeiras e a condição (i) não é satisfeita nem uma vez, dizemos que $s_1 \sqsubseteq s_2 \sqsubseteq s_1$, *ou seja,* s_1 *e* s_2 *são equivalentes. Se nem todos os rótulos atômicos são comparáveis, então o critério não pode ser aplicado.*

Intuitivamente, esse critério nos leva à idéia de que se dois suportes têm o mesmo número de elementos e esses elementos são rótulos atômicos (rotulam fórmulas da teoria ou são rótulos hipotéticos), então é porque eles foram formados pelas mesmas regras de dedução, e por isso seus elementos podem ser comparados um a um. Geralmente, só após a aplicaçação do critério de unificação e da redução dos suportes, de tal forma que os suportes reduzidos fiquem formados por um pequeno número de elementos (atômicos), é que o critério da ordem preferencial dos rótulos atômicos pode ser aplicado.

No caso do exemplo acima basta verificar a ordem de preferência dos rótulos l_2 e l_3. Se temos a ordem definida como no diagrama da figura 5.3, por exemplo, então $l_2 \sqsubseteq l_3$, o que implica que o candidato α_1 a partir do qual chegou-se ao suporte $E_{\rightarrow}(I_{\wedge}(l_1, l_2)$ que foi reduzido a l_2 é preferível ao candidato α_2 a partir do qual chegou-se ao suporte $E_{\rightarrow}(I_{\wedge}(l_1, l_3)$ que foi reduzido a l_3.

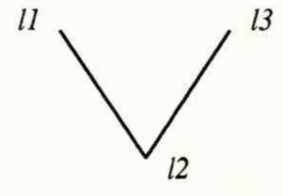

Figura 5.3: Ordenação de rótulos utilizada no critério da ordem preferencial

Critério de número de elementos

Outro critério importante é o critério de **número de elementos** (rótulos atômicos e funções) presentes no suporte. Em alguns casos, após reduzirmos os rótulos a serem comparados o máximo possível, nos deparamos com a comparação de um suporte reduzido com apenas um ou dois rótulos atômicos e funções e outro suporte reduzido no qual ainda restam diversas funções e rótulos atômicos. Nesse caso verificamos qual o rótulo atômico de menor preferência (aquele que se encontra em uma posição mais alta no diagrama de ordenação) que está presente no menor suporte a ser comparado e se ele (o rótulo atômico de menor preferência) não possuir uma preferência menor que qualquer dos rótulos atômicos do suporte reduzido maior a ser comparado, então podemos considerar o candidato a partir do qual chegou-se ao suporte reduzido menor, como sendo mais plausível.

Formalmente, podemos definir o critério de número de elementos segundo a definição abaixo:

Definição 5.11 (Critério de número de elementos) *Sejam* $s_1 = Spt(\varphi, \alpha_1) = (t_{11}, ..., t_{1n})$ *e* $s_2 = Spt(\varphi, \alpha_2) = (t_{21}, ..., t_{2m})$ *dois suportes (ou suportes reduzidos) associados a um mesmo problema abdutivo* $\langle \Theta, \varphi, \sqsubseteq \rangle$ *onde* s_1 *e* s_2 *não podem mais ser reduzidos por unificação.*

Seja t_{1j} *o rótulo atômico não hipotético presente em* s_1 *que possui a menor preferência (encontra-se mais acima no diagrama preferencial dos rótulos), isto é, para todo* i, $t_{1i} \sqsubseteq t_{1j}$, *se* t_{1i} *é um rótulo atômico não hipotético.*

Dizemos que $s_1 \sqsubseteq s_2$ *se as condições abaixo são verdadeiras:*

(i) $m > n$;

(ii) para todo i, *se* t_{2i} *é um rótulo não hipotético, então* $t_{1j} \sqsubseteq t_{2i}$ *na ordem fornecida pelo usuário.*

Com o critério de número de elementos, pretendemos dar preferência àqueles candidatos que são gerados com o menor número de passos na dedução. Esse critério faz sentido se consideramos, por exemplo, que o rótulo retrata um grau de incerteza e assim quanto mais elementos da teoria são necessários para a prova de φ a partir de determinado candidato, menor é a plausibilidade desse candidato ser a causa de φ, já que a cada passo da dedução podemos estar inserindo uma maior incerteza.

Critério do tipo de rótulo

O critério de **tipo de rótulo** pode ser utilizado sempre que estamos comparando dois a dois os elementos de suportes, inclusive quando da aplicação de outros dos critérios de comparação vistos acima, para dar preferência a um ou outro elemento, segundo as regras definidas a seguir.

Definição 5.12 (Critério do tipo de rótulo) *Quando dois elementos de dois suportes associados a um mesmo problema abdutivo $\langle \Theta, \varphi, \sqsubseteq \rangle$ estão sendo comparados, pode-se considerar as seguintes regras:*

1. *$a \sqsubseteq b$, se ambos rotulam fórmulas da teoria e $a \sqsubseteq b$ na ordem fornecida pelo usuário;*

2. *$a \sqsubseteq b$, se $a \sqsubseteq b \sqsubseteq a$, ou $a \not\sqsubseteq b$ e $b \not\sqsubseteq a$ na ordem fornecida pelo usuário e a rotula um fato da teoria e b rotula uma regra da teoria;*

3. *$l \sqsubseteq x$, se l rotula uma fórmula da teoria e x é um rótulo hipotético;*

4. *$f(l) \sqsubseteq x$, se l rotula uma fórmula da teoria , f é uma função qualquer e x é um rótulo hipotético;*

5. *$f(a) \sqsubseteq f(b)$, se a e b rotulam fórmulas da teoria e $a \sqsubseteq b$ na ordem fornecida pelo usuário.*

6. *$f(a,b) \sqsubseteq f(c,d)$, se a, b, c e d rotulam fórmulas da teoria e se $a \sqsubseteq c$, $a \sqsubseteq d$, $b \sqsubseteq c$ e $b \sqsubseteq d$.*

Critério das conseqüências dedutivas

Como foi dito na seção 5.2.3, as conseqüências dedutivas dos candidatos gerados devem ser consideradas, se existirem, no momento da comparação dos candidatos. Assim, o critério das conseqüências dedutivas pode ser formulado em função do número de conseqüências existentes para cada candidato. Podemos definir formalmente este critério como segue:

Definição 5.13 (Critério das conseqüências dedutivas) *Seja α_1 e α_2 dois elementos de $Cand(\Theta, \varphi)$ e $Conseq(\varphi, \alpha_1)$ e $Conseq(\varphi, \alpha_1)$ seus respectivos conjuntos de conseqüências dedutivas. Dizemos que $\alpha_1 \sqsubseteq \alpha_2$ se o número de elementos de $Conseq(\varphi, \alpha_1)$ é menor ou igual ao número de elementos de $Conseq(\varphi, \alpha_2)$.*

No exemplo 5.14 a comparação de candidatos utilizando os critérios de unificação e número de elementos é ilustrada.

Exemplo 5.14 (Utilização dos critérios de unificação e número de elementos)

$$\Theta = \{ \begin{array}{ll} l_1 : & \mathbf{A} \vee \mathbf{B} \to \mathbf{C}, \\ l_2 : & \mathbf{D} \to \mathbf{A}, \\ l_3 : & \mathbf{F} \wedge \mathbf{E} \to \mathbf{D}, \\ l_4 : & \mathbf{G} \to \mathbf{E}, \\ l_5 : & \mathbf{G}, \\ l_6 : & \mathbf{H} \to \mathbf{B}, \\ l_7 : & \mathbf{J} \wedge \mathbf{G} \to \mathbf{B}\} \end{array}$$

$$\varphi = \mathbf{C}$$

As árvores de dedução para as tentativas de construção da prova seriam:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{x : \mathbf{F} \quad \dfrac{l_5 : \mathbf{G} \quad l_4 : \mathbf{G} \to \mathbf{E}}{E_{\to}(l_5, l_4) : \mathbf{E}}}{I_{\wedge}(x, E_{\to}(l_5, l_4)) : \mathbf{F} \wedge \mathbf{E}} \quad l_3 : \mathbf{F} \wedge \mathbf{E} \to \mathbf{D}}{E_{\to}(I_{\wedge}(x, E_{\to}(l_5, l_4)), l_3) : \mathbf{D}} \quad l_2 : \mathbf{D} \to \mathbf{A}}{E_{\to}(E_{\to}(I_{\wedge}(x, E_{\to}(l_5, l_4)), l_3), l_2) : \mathbf{A}}}{I_{1\vee}(E_{\to}(E_{\to}(I_{\wedge}(x, E_{\to}(l_5, l_4)), l_3), l_2)) : \mathbf{A} \vee \mathbf{B}} \quad l_1 : \mathbf{A} \vee \mathbf{B} \to \mathbf{C}}{E_{\to}(I_{1\vee}(E_{\to}(E_{\to}(I_{\wedge}(x, E_{\to}(l_5, l_4)), l_3), l_2)), l_1) : \mathbf{C}}$$

$$\dfrac{\dfrac{\dfrac{y : \mathbf{H} \quad l_6 : \mathbf{H} \to \mathbf{B}}{E_{\to}(y, l_6) : \mathbf{B}}}{I_{2\vee}(E_{\to}(y, l_6)) : \mathbf{A} \vee \mathbf{B}} \quad l_1 : \mathbf{A} \vee \mathbf{B} \to \mathbf{C}}{E_{\to}(I_{2\vee}(E_{\to}(y, l_6)), l_1) : \mathbf{C}}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{z : \mathbf{J} \quad l_5 : \mathbf{G}}{I_{\wedge}(z, l_5) : \mathbf{J} \wedge \mathbf{G}} \quad l_7 : \mathbf{J} \wedge \mathbf{G} \to \mathbf{B}}{E_{\to}(I_{\wedge}(z, l_5), l_7) : \mathbf{B}}}{I_{2\vee}(E_{\to}(I_{\wedge}(z, l_5), l_7)) : \mathbf{A} \vee \mathbf{B}} \quad l_1 : \mathbf{A} \vee \mathbf{B} \to \mathbf{C}}{E_{\to}(I_{2\vee}(E_{\to}(I_{\wedge}(z, l_5), l_7)), l_1) : \mathbf{C}}$$

O conjunto de candidatos seria:

$$Cand(\Theta, \varphi) = \{ \begin{array}{l} \mathbf{F}, \\ \mathbf{H}, \\ \mathbf{J} \; \} \end{array}$$

E os suportes encontrados a partir de cada um dos candidatos gerados, são:

$$s_1 = Spt(\mathbf{C},\mathbf{F}) = E_{\rightarrow}(I_{1\vee}(E_{\rightarrow}(E_{\rightarrow}(I_{\wedge}(x, E_{\rightarrow}(l_5, l_4)), l_3), l_2)), l_1)$$

$$s_2 = Spt(\mathbf{C},\mathbf{H}) = E_{\rightarrow}(I_{2\vee}(E_{\rightarrow}(y, l_6)), l_1)$$

$$s_3 = Spt(\mathbf{C},\mathbf{J}) = E_{\rightarrow}(I_{2\vee}(E_{\rightarrow}(I_{\wedge}(z, l_5), l_7)), l_1)$$

Suponha a ordem preferencial dada pela figura 5.4.

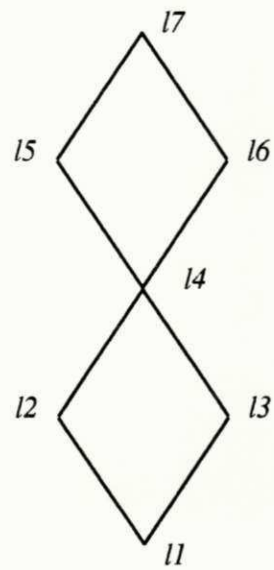

Figura 5.4: Ordenação de rótulos do exemplo 5.14

Pelo critério de unificação, para os candidatos H e J, podemos eliminar as 3 (três) primeiras funções. Nos resta então a seguinte comparação:

$$((y, l_6), l_1) \ e \ ((I_{\wedge}(z, l_5), l_7), l_1)$$

onde comparamos y *com* $I_{\wedge}(z, l_5)$, l_6 *com* l_7 *e* l_1 *com* l_1.

Os primeiros parâmetros a serem comparados possuem rótulos hipotéticos, para os quais não podemos ter preferência alguma. Passamos então ao segundo parâmetro (l_6 *e* l_7*) e verificamos que* $l_6 \sqsubseteq l_7$ *(pela ordenação fornecida pela figura 5.4); os últimos parâmetros apresentam o mesmo elemento* l_1. *Então podemos descartar o candidato J, dando preferência ao candidato H.*

Nos resta agora a comparação dos suportes de F e H. Podemos eliminar, também por unificação, a primeira função. Ficamos então com os dois suportes reduzidos a seguir:

$$s_1 = Spt(\mathbf{C}, \mathbf{F}) = E_{\rightarrow}(I_{1\vee}(E_{\rightarrow}(E_{\rightarrow}(I_{\wedge}(x, E_{\rightarrow}(l_5, l_4)), l_3), l_2)), l_1)$$

$$s_2 = Spt(\mathbf{C}, \mathbf{H}) = E_{\rightarrow}(I_{2\vee}(E_{\rightarrow}(y, l_6)), l_1)$$

Podemos aplicar o critério do número de elementos: temos s_2 como o menor suporte e seu rótulo atômico de menor preferência é o l_6. Como l_6 não tem uma preferência menor que qualquer um dos rótulos atômicos de s_1, que é o maior suporte, podemos dar preferência ao elemento H, a partir do qual foi encontrado o suporte s_2.

Concluímos então que:

$$\Theta \vdash_L \mathbf{H} \leadsto \mathbf{C}$$

No exemplo 5.15 é mostrada a comparação utilizando o critério do tipo de rótulo.

Exemplo 5.15 (Utilização do critério do tipo de rótulo)

$$\begin{aligned} \Theta = \{ \quad & l_1 : \mathbf{A} \\ & l_2 : \mathbf{B} \rightarrow \mathbf{C}, \\ & l_3 : \mathbf{E} \wedge \mathbf{A} \rightarrow \mathbf{X}, \\ & l_4 : \mathbf{D} \wedge (\mathbf{B} \rightarrow \mathbf{C}) \rightarrow \mathbf{X}, \\ & l_5 : \mathbf{E} \rightarrow \mathbf{G}\} \end{aligned}$$

$$\varphi = \mathbf{X}$$

As árvores de dedução para as tentativas de construção da prova seriam:

$$\dfrac{\dfrac{x : \mathbf{D} \quad l_2 : \mathbf{B} \rightarrow \mathbf{C}}{I_{\wedge}(x, l_2) : \mathbf{D} \wedge (\mathbf{B} \rightarrow \mathbf{C})} \quad l_4 : \mathbf{D} \wedge (\mathbf{B} \rightarrow \mathbf{C}) \rightarrow \mathbf{X}}{E_{\rightarrow}(I_{\wedge}(x, l_2), l_4) : \mathbf{X}}$$

$$\dfrac{\dfrac{y : \mathbf{E} \quad l_1 : \mathbf{A}}{I_{\wedge}(y, l_1) : \mathbf{E} \wedge \mathbf{A}} \quad l_3 : \mathbf{E} \wedge \mathbf{A} \rightarrow \mathbf{X}}{E_{\rightarrow}(I_{\wedge}(y, l_1), l_3) : \mathbf{X}}$$

Assim chegamos aos seguintes suportes:

$$s_1 = Spt(\mathbf{X}, \mathbf{D}) = E_{\rightarrow}(I_{\wedge}(x, l_2), l_4)$$

$$s_2 = Spt(\mathbf{X}, \mathbf{E}) = E_{\rightarrow}(I_{\wedge}(y, l_1), l_3)$$

Por unificação reduzimos os dois suportes a:

$$red(s_1) = (l_2, l_4)$$

$$red(s_2) = (l_1, l_3)$$

já que podemos eliminar as funções $E_{\rightarrow}, I_{\wedge}$ e os rótulos hipotéticos x e y.

Se considerarmos o critério da ordem preferencial dos rótulos (supondo a figura 5.5 como representando a ordenação das fórmulas de Θ), como l_1 e l_2 não são comparáveis e o mesmo se dá para l_3 e l_4, então os candidatos deveriam ser considerados equivalentes.

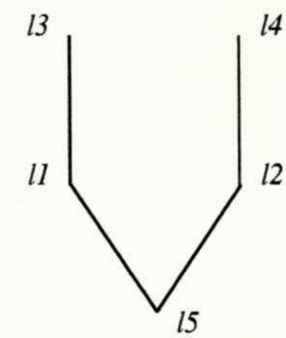

Figura 5.5: Ordenação de rótulos do exemplo 5.15

Mas, pelo critério do tipo de rótulo, verificamos que l_2 rotula uma regra e l_1 rotula um fato. Logo, podemos assumir que $\mathbf{E} \sqsubseteq_c \mathbf{D}$. Então,

$$\Theta \vdash_L \mathbf{E} \rightsquigarrow \mathbf{X}$$

Considerando o caso em que temos candidatos fornecidos pelo usuário, conclusões aparentemente estranhas podem ser tiradas. Por exemplo, podemos ter um conjunto de candidatos (fornecido pelo usuário) ordenado por uma determinada relação preferencial, e podemos chegar a uma diferente ordem de preferência desses mesmos candidatos, após a propagação dos rótulos e posterior comparação dos suportes. Isso se dá devido ao fato de que a construção da prova de φ

se dá através também da utilização de outras sentenças da teoria que terão sua própria ordem de preferência (seguindo o mesmo critério dos candidatos fornecidos), ordem essa que influenciará também no suporte de φ para cada candidato. Esse fato apenas confirma que a maior ou menor plausibilidade de um candidato ser a causa do fato observado dependende dos passos de dedução que ocorrem na construção da prova. O exemplo 5.16 ilustra um caso em que a solução encontrada apresenta uma ordem diversa da dos candidatos fornecidos pelo usuário.

Exemplo 5.16 (Candidatos fornecidos pelo usuário e solução com ordem diversa)

$$\begin{aligned} \Theta = \{ \quad l_1 :& \quad \mathbf{A} \rightarrow \mathbf{B}, \\ l_2 :& \quad \mathbf{C} \rightarrow \mathbf{D}, \\ l_3 :& \quad \mathbf{A} \wedge \mathbf{C} \rightarrow \mathbf{E}, \\ l_4 :& \quad \neg\mathbf{G} \rightarrow \mathbf{D}, \\ l_5 :& \quad \mathbf{A}, \\ l_6 :& \quad \mathbf{A} \rightarrow \mathbf{H}\} \end{aligned}$$

$$\varphi = \mathbf{D}$$

E temos os seguintes candidatos fornecidos pelo usuário:

$$\begin{aligned} CandF(\Theta, \mathbf{D}) = \{ \quad & l_7 : \mathbf{B} \rightarrow \mathbf{D}, \\ & l_8 : \mathbf{H} \rightarrow \mathbf{D} \ \} \end{aligned}$$

Gerando os candidatos, teremos as seguintes árvores de dedução:

$$\frac{x : \mathbf{C} \quad l_2 : \mathbf{C} \rightarrow \mathbf{D}}{E_{\rightarrow}(x, l_2) : \mathbf{D}}$$

$$\frac{y : \neg\mathbf{G} \quad l_5 : \neg\mathbf{G} \rightarrow \mathbf{D}}{E_{\rightarrow}(y, l_5) : \mathbf{D}}$$

Assim, temos o seguinte conjunto de candidatos:

$$\begin{aligned} Cand(\Theta, \mathbf{D}) = \{ \quad & \mathbf{C}, \\ & \neg\mathbf{G} \ \} \end{aligned}$$

Generalizando os candidatos para verificar se os elementos de $CandF(\Theta, D)$ são realmente candidatos, temos:

- $\Theta \vdash_L B$, *logo, pelo item (ii) da definição 5.7,* $\mathbf{B} \to \mathbf{D}$ *é candidato;*

- $\Theta \vdash_L H$, *logo, pelo item (ii) da definição 5.7,* $\mathbf{H} \to \mathbf{D}$ *é candidato;*

Então os elementos de $CandF(\Theta, \mathbf{D})$ *são realmente candidatos; a construção da prova a partir de ambos seria:*

$$\dfrac{\dfrac{l_5 : \mathbf{A} \quad l_1 : \mathbf{A} \to \mathbf{B}}{E_{\to}(l_5, l_1) : \mathbf{B}} \qquad l_7 : \mathbf{B} \to \mathbf{D}}{E_{\to}(E_{\to}(l_5, l_1), l_7) : \mathbf{D}}$$

$$\dfrac{\dfrac{l_5 : \mathbf{A} \quad l_6 : \mathbf{A} \to \mathbf{H}}{E_{\to}(l_5, l_6) : \mathbf{H}} \qquad l_8 : \mathbf{H} \to \mathbf{D}}{E_{\to}(E_{\to}(l_5, l_6), l_8) : \mathbf{D}}$$

Assim chegamos aos seguintes suportes:

$$s_1 = Spt(\mathbf{D}, \mathbf{C}) = E_{\to}(x, l_2)$$

$$s_2 = Spt(\mathbf{D}, \neg\mathbf{G}) = E_{\to}(y, l_5)$$

$$s_3 = Spt(\mathbf{D}, \mathbf{B} \to \mathbf{D}) = E_{\to}(E_{\to}(l_5, l_1), l_7)$$

$$s_4 = Spt(\mathbf{D}, \mathbf{H} \to \mathbf{D}) = E_{\to}(E_{\to}(l_5, l_6), l_8)$$

Vamos supor a ordenação dos rótulos dada pela figura 5.6.

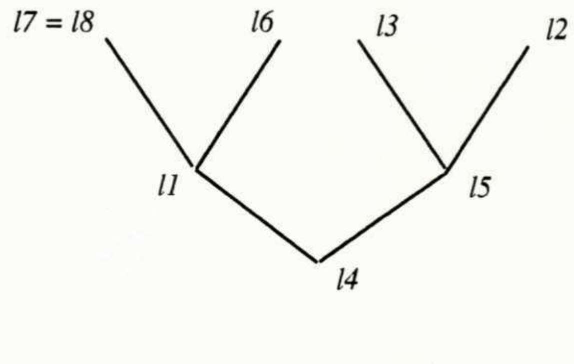

Figura 5.6: Ordenação de rótulos do exemplo 5.16

Por unificação, o primeiro e segundo suportes podem ser reduzidos para l_2 *e* l_5*; logo, como* $l_5 \sqsubseteq l_2$ *na ordem preferencial dos rótulos, podemos eliminar o candidato* $\mathbf{C}$ *pelo critério da ordem preferencial dos rótulos.*

Os dois últimos suportes também podem ser comparados por unificação e reduzidos para (l_1, l_7) *e* (l_6, l_8)*. Como* $l_1 \sqsubseteq l_6$ *e* l_7 *e* l_8 *têm a mesma preferência, pelo critério da ordem preferencial dos rótulos, podemos considerar que* $\mathbf{B} \rightarrow \mathbf{D} \sqsubseteq_c \mathbf{H} \rightarrow \mathbf{D}$ *(contrariando a ordem fornecida pelo usuário que os considera equivalentes).*

Nos resta comparar os suportes $Spt(\mathbf{D}, \neg\mathbf{G})$ *e* $Spt(\mathbf{D}, \mathbf{B} \rightarrow \mathbf{D})$*. Podemos tentar utilizar a critério de número de elementos; no entanto, alguns rótulos atômicos do suporte maior não são comparáveis com o rótulo atômico do suporte menor. Logo, não podemos eliminar mais nenhum candidato, e ficamos com as seguintes soluções como mais plausíveis:*

$$\Theta \vdash_L \neg\mathbf{G} \rightsquigarrow \mathbf{D}$$

$$\Theta \vdash_L \mathbf{B} \rightarrow \mathbf{D} \rightsquigarrow \mathbf{D}$$

O exemplo 5.17 mostra a aplicação do critério das conseqüências dedutivas.

Exemplo 5.17 (Utilização do critério das conseqüências dedutivas)

$$\begin{aligned} \Theta = \{ \quad l_1 : &\quad \mathit{Doença1} \rightarrow \mathit{Sintoma1}, \\ l_2 : &\quad \mathit{Doença2} \rightarrow \mathit{Sintoma1} \wedge \mathit{Sintoma2}\} \end{aligned}$$

$$\varphi = \mathit{Sintoma1}$$

ou

$$\begin{aligned} \Theta = \{ \quad l_1 : &\quad \mathbf{D1} \rightarrow \mathbf{S1}, \\ l_2 : &\quad \mathbf{D2} \rightarrow \mathbf{S1} \wedge \mathbf{S2}\} \end{aligned}$$

$$\varphi = \mathbf{S1}$$

Gerando os candidatos, teremos as seguintes árvores de dedução:

$$\frac{x : \mathbf{D1} \quad l_1 : \mathbf{D1} \rightarrow \mathbf{S1}}{E_{\rightarrow}(x, l_1) : \mathbf{S1}}$$

$$\cfrac{\cfrac{y : \mathbf{D2} \quad l_2 : \mathbf{D2} \rightarrow \mathbf{S1} \wedge \mathbf{S2}}{E_{\rightarrow}(y, l_2) : \mathbf{S1} \wedge \mathbf{S2}}}{E_{1\wedge}(E_{\rightarrow}(y, l_2)) : \mathbf{S1}}$$

Assim, chegamos ao seguinte conjunto de candidatos:

$$Cand(\Theta, \mathbf{S1}) = \{ \quad \mathbf{D1}, \\ \mathbf{D2} \}$$

Com, respectivamente, os seguintes suportes:

$$s_1 = Spt(\mathbf{S1}, \mathbf{D1}) = E_{\rightarrow}(x, l_1)$$

$$s_2 = Spt(\mathbf{S1}, \mathbf{D2}) = E_{1\wedge}(E_{\rightarrow}(y, l_2))$$

Gerando as conseqüências dedutivas dos dois candidatos, temos:

$$Conseq(\Theta, \mathbf{D1}) = \emptyset$$

$$Conseq(\Theta, \mathbf{D2}) = \{\mathbf{S2}\}$$

Vamos supor a ordenação dos rótulos dada pela figura 5.7.

Figura 5.7: Ordenação de rótulos do exemplo das conseqüências dedutivas

Como os dois rótulos atômicos dos dois candidatos (l_1 e l_2) não são comparáveis, não teríamos como eliminar nenhum dos candidatos considerando os quatro primeiros critérios de comparação.

Considerando, porém, o critério das conseqüências dedutivas, como o número de elementos de $Conseq(\Theta, \mathbf{D1})$ é menor que o número de elementos de $Conseq(\Theta, \mathbf{D2})$, podemos concluir que $\mathbf{D1} \sqsubseteq_c \mathbf{D2}$. *logo:*

$$\Theta \vdash_L \mathbf{D1} \rightsquigarrow \mathbf{S1}$$

O exemplo 5.18 abaixo dá continuidade na busca da solução do exemplo 5.12.

Exemplo 5.18 (Continuação da solução para o problema dos exemplos 5.6 e 5.12) *Após a aplicação do critério lógico da minimalidade nos restaram os seguintes suportes:*

$$Spt(\mathbf{D}, \mathbf{A}) = E_{\rightarrow}(x, l_1)$$

$$Spt(\mathbf{D}, \mathbf{E}) = E_{2\wedge}(E_{\rightarrow}(y, l_2))$$

$$Spt(\mathbf{D}, \mathbf{B} \vee \mathbf{F}) = E_{\rightarrow}(z, l_3)$$

que foram derivados respectivamente dos candidatos **A**, **E** *e* **B** $\vee$**F**. *Supondo que a ordem preferencial dos rótulos da teoria seja definida pelo diagrama da figura 5.8, a comparação dos candidatos se processará como descrito abaixo.*

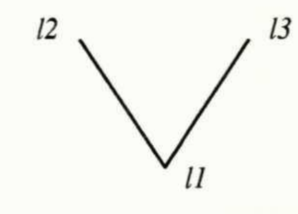

Figura 5.8: Ordenação de rótulos do exemplo 5.18

Pelo critério de unificação, o primeiro e o terceiro suportes podem ser reduzidos a l_1 *e* l_3. *Pelo critério da ordem preferencial dos rótulos temos que* $l_1 \sqsubseteq l_3$, *então:*

$$\mathbf{A} \sqsubseteq_c \mathbf{B} \vee \mathbf{F}$$

Comparando agora o primeiro e o segundo suportes, temos: pelo critério de número de elementos nos suportes, como o segundo suporte ($Spt(\mathbf{D}, \mathbf{E}) = E_{2\wedge}(E_{\rightarrow}(y, l_2))$) *possui mais elementos que o primeiro* ($Spt(\mathbf{D}, \mathbf{A}) = E_{\rightarrow}(x, l_1)$), *e, além disso, o seu único rótulo atômico comparável* (l_2) *tem uma ordem preferencial menor que o rótulo atômico comparável do primeiro suporte* (l_1), *ou seja,* $l_1 \sqsubseteq l_2$, *temos:*

$$\mathbf{A} \sqsubseteq_c \mathbf{E}$$

logo, a solução do nosso exemplo é:

$$\Theta \vdash_L \mathbf{A} \leadsto \mathbf{D}$$

O exemplo 5.19 mostra a solução completa do exemplo 5.3:

Exemplo 5.19 (Continuação do exemplo 5.3)

$$\begin{aligned}\Theta = \{\ & l_1: \ \mathbf{A},\\ & l_2: \ \mathbf{A}\wedge\mathbf{B}\rightarrow\mathbf{E},\\ & l_3: \ \mathbf{C}\rightarrow\mathbf{D}\wedge\mathbf{J},\\ & l_4: \ \mathbf{D}\wedge\mathbf{E}\rightarrow\mathbf{F},\\ & l_5: \ \mathbf{B}\vee\mathbf{G}\rightarrow\mathbf{F},\\ & l_6: \ \mathbf{H}\rightarrow\mathbf{F}\}\end{aligned}$$

$$\varphi = \mathbf{F}$$

As árvores de dedução com as tentativas de prova e com os rótulos devidamente propagados, seriam:

$$\dfrac{\dfrac{\dfrac{\dfrac{x:\mathbf{C}\quad l_3:\mathbf{C}\rightarrow\mathbf{D}\wedge\mathbf{J}}{E_{\rightarrow}(x,l_3):\mathbf{D}\wedge\mathbf{J}}}{E_{1\wedge}(E_{\rightarrow}(x,l_3)):\mathbf{D}}\qquad \dfrac{\dfrac{l_1:\mathbf{A}\quad y:\mathbf{B}}{I_{\wedge}(l_1,y):\mathbf{A}\wedge\mathbf{B}}\quad l_2:\mathbf{A}\wedge\mathbf{B}\rightarrow\mathbf{E}}{E_{\rightarrow}(I_{\wedge}(l_1,y),l_2):\mathbf{E}}}{I_{\wedge}(E_{1\wedge}(E_{\rightarrow}(x,l_3)),E_{\rightarrow}(I_{\wedge}(l_1,y),l_2)):\mathbf{D}\wedge\mathbf{E}}\qquad l_4:\mathbf{D}\wedge\mathbf{E}\rightarrow\mathbf{F}}{E_{\rightarrow}(I_{\wedge}(E_{1\wedge}(E_{\rightarrow}(x,l_3)),E_{\rightarrow}(I_{\wedge}(l_1,y),l_2)),l_4):\mathbf{F}}$$

$$\dfrac{z:\mathbf{B}\vee\mathbf{G}\quad l_5:\mathbf{B}\vee\mathbf{G}\rightarrow\mathbf{F}}{E_{\rightarrow}(z,l_5):\mathbf{F}}$$

$$\dfrac{\dfrac{w:\mathbf{B}}{I_{1\vee}(w):\mathbf{B}\vee\mathbf{G}}\quad l_5:\mathbf{B}\vee\mathbf{G}\rightarrow\mathbf{F}}{E_{\rightarrow}(I_{1\vee}(w),l_5):\mathbf{F}}$$

$$\dfrac{\dfrac{t:\mathbf{G}}{I_{2\vee}(t):\mathbf{B}\vee\mathbf{G}}\quad l_5:\mathbf{B}\vee\mathbf{G}\rightarrow\mathbf{F}}{E_{\rightarrow}(I_{2\vee}(t),l_5):\mathbf{F}}$$

$$\dfrac{s:\mathbf{H}\quad l_6:\mathbf{H}\rightarrow\mathbf{F}}{E_{\rightarrow}(s,l_6):\mathbf{F}}$$

O conjunto de candidatos obtido possui os seguintes elementos:

$$Cand(\Theta,\mathbf{F}) = \{\mathbf{C}\wedge\mathbf{B},\mathbf{B}\vee\mathbf{G},\mathbf{B},\mathbf{G},\mathbf{H}\}$$

cujos respectivos suportes são:

$$Spt(\mathbf{F}, \mathbf{C} \wedge \mathbf{B}) = E_{\rightarrow}(I_{\wedge}(E_{1\wedge}(E_{\rightarrow}(x, l_3)), E_{\rightarrow}(I_{\wedge}(l_1, y), l_2)), l_4)$$

$$Spt(\mathbf{F}, \mathbf{B} \vee \mathbf{G}) = E_{\rightarrow}(z, l_5)$$

$$Spt(\mathbf{F}, \mathbf{B}) = E_{\rightarrow}(I_{1\vee}(w), l_5)$$

$$Spt(\mathbf{F}, \mathbf{G}) = E_{\rightarrow}(I_{2\vee}(t), l_5)$$

$$Spt(\mathbf{F}, \mathbf{H}) = E_{\rightarrow}(s, l_6)$$

Pelo critério da minimalidade a nível objeto, podemos eliminar os candidatos $\mathbf{B}$ *e* $\mathbf{G}$*, já que* $\mathbf{B} \rightarrow \mathbf{B} \vee \mathbf{G}$ *e* $\mathbf{G} \rightarrow \mathbf{B} \vee \mathbf{G}$*; logo os candidatos* $\mathbf{B}$ *e* $\mathbf{G}$ *não são minimais.*

Considerando que a ordenação dos rótulos a teoria é dada pelo diagrama da figura 5.9, por unificação podemos comparar os candidatos $\mathbf{H}$ *e* $\mathbf{B} \vee \mathbf{G}$ *e reduzir os respectivos suportes a* l_6 *e* l_5*, respectivamente. Como* $l_6 \sqsubseteq l_5$ *na ordem de preferência, pelo critério de ordem preferencial dos rótulos posso eliminar o candidato* $\mathbf{B} \vee \mathbf{G}$*.*

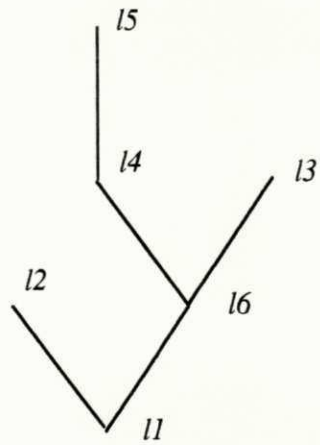

Figura 5.9: Ordenação de rótulos do exemplo 5.19

Os candidatos que nos restam, $\mathbf{H}$ *e* $\mathbf{C} \wedge \mathbf{B}$*, poderiam ser comparados pelo critério de número de elementos. No entanto, fazendo essa tentativa observamos que o rótulo atômico de menor preferência do suporte menor* (l_6) *não tem uma preferência maior que os rótulos atômicos do suporte maior. Logo a nossa solução final apresentaria dois candidatos:*

$$\Theta \vdash_L \mathbf{H} \rightsquigarrow \mathbf{F}$$

$$\Theta \vdash_L \mathbf{C} \wedge \mathbf{B} \rightsquigarrow \mathbf{F}$$

Mas podemos ainda aplicar o critério das conseqüências dedutivas, e assim temos que:

$$Conseq(\Theta, \mathbf{H}) = \emptyset$$

$$Conseq(\Theta, \mathbf{C} \wedge \mathbf{B}) = \{\mathbf{E}, \mathbf{D}, \mathbf{J}\}$$

Elimina-se, então o candidato $\mathbf{C} \wedge \mathbf{B}$ *e temos como solução final:*

$$\Theta \vdash_L \mathbf{H} \rightsquigarrow \mathbf{F}$$

Os critérios de comparação de candidatos discutidos acima são critérios genéricos que podem ser utilizados com problemas abdutivos referentes a qualquer domínio. No entanto, supomos que quando se tratando de um domínio específico, seria interessante propor a aplicação de critérios de comparação mais específicos e apropriados que os aqui apresentados. Um estudo de caso para uma aplicação específica pode esclarecer essa possibilidade. Como veremos no Capítulo 6, esse é um dos ítens sugeridos como trabalho futuro.

Outro ponto importante a salientar é que as definições propostas aqui para os critérios de comparação de candidatos precisam ser algoritmicamente especificadas a fim de que todos os seus aspectos sejam explicitados, inclusive deverá ser definida uma ordem para a aplicação dos critérios.

5.5 Algumas considerações sobre o método proposto

5.5.1 Sobre a não monotonicidade na abdução

No que diz respeito à busca da explicação mais plausível, a abdução é, sem dúvida, não monotônica: quando novos fatos são acrescentados à teoria, melhores explicações podem ser encontradas. Isso é indiscutível.

Mas a abdução também é não monotônica no sentido de que α que era uma explicação plausível, pode deixar de ser uma explicação com o acréscimo de novos fatos. Se modificarmos a teoria, é como se estivéssemos acrescentando restrições de integridade, e dessa maneira (como visto na seção 5.4.1) provavelmente serão excluídos alguns elementos de $Cand(\Theta, \varphi)$, ou seja, o que antes era uma explicação pode deixar de ser.

Em [MP95b] é argumentado que a monotonicidade não ocorre porque, utilizando o exemplo citado em tal referência, verificando-se que '*uma xícara de café está doce*', pode-se tentar explicar esse fato com a seguinte afirmação: '*o café pode estar doce porque foi colocada uma colher de açúcar*'. Mas, se acrescentamos informação à explicação dizendo que '*foi colocada uma colher de açúcar e outra de sal*', então essa não é mais uma explicação para o fato (pois o café não estaria doce dessa maneira).

No entanto, no nosso processo de generalização de candidatos, estamos propondo que se $\alpha \in Cand(\Theta, \varphi)$ então $\gamma \wedge \alpha \in Gen(Cand(\Theta, \varphi))$ (veja seção 5.4.1, definição 5.7). Isso parece incoerente com os argumentos em [MP95b]. Como justificar nosso procedimento?

Bem, segundo os critérios puramente lógicos, é correto dizer o que afirmamos no item (*iii*) da definição 5.7. Mas vale deixar a ressalva que a generalização de candidatos pode levar-nos a candidatos não plausíveis. Esse é mais um motivo para lidarmos apenas com o conjunto de candidatos minimais. Já que o objetivo é propor um procedimento para gerar candidatos, como não podemos incorporar a semântica das fórmulas, devemos nos ater ao que nos diz a teoria.

A generalização que propomos, no entanto, se torna útil para verificar, no caso de candidatos fornecidos pelo usuário, a veracidade desses candidatos: os candidatos fornecidos devem pertencer ao conjunto $Gen(Cand(\Theta, \varphi))$.

No caso do '*café doce*', simplesmente acrescentar à explicação '*foi colocada uma colher de açúcar*' a nova suposição '*foi colocada também uma colher de sal*' (ou seja, acrescentamos a α um γ e queremos que $\alpha \wedge \gamma$ também seja um candidato), não faz com que a frase resultante deixe de ser uma explicação para o fato (sabemos que a conjunção resultante não pode ser uma explicação porque entendemos a **semântica** da frase). Na verdade, para que a conjunção '*foi colocada uma colher de açúcar e outra de sal*' não seja vista como uma explicação para o fato '*o café está doce*', outras informações também devem ser acrescentadas, por exemplo: '*colocar colher de sal implica que o café não estará doce*'. Ora, se acrescentamos novas informações, estamos **modificando a teoria**, ou seja, estamos colocando restrições de integridade. Sendo assim, podemos tratar facilmente essa situação no nível objeto da solução como visto na seção 5.2.2.

Concluindo, o raciocínio abdutivo é não monotônico sim. Mas, desde que a teoria não seja alterada, o processo de generalização dos candidatos é perfeitamente coerente. Se alteramos a teoria, estaremos lidando com restrições de integridade, e isso é uma outra questão.

5.5.2 Sobre a extensão para lógica de primeira ordem

Um dos ítens propostos como trabalhos futuros, no Capítulo 6 desta dissertação, é a extensão da teoria de abdução para a lógica proposicional aqui definida, para a lógica de primeira ordem. Nesta seção serão colocados alguns pontos já identificados inerentes a essa extensão.

Como vimos na abdução proposicional, os principais problemas que surgem são aqueles que se referem ao meta-nível da solução: a comparação dos candidatos é o mais crucial. Na abdução de primeira ordem, porém, os problemas começam a surgir já no nível objeto. Como colocado em [MP95b], a indecidibilidade da lógica de primeira ordem se reflete no fato de que pode ser impossível terminar o processo de geração de explicações para um problema abdutivo. Conseqüentemente, o conjunto de explicações pode ser infinito, e por isso é obvio que determinar se uma explicação é minimal (com respeito a $\vdash$) é em geral indecidível e também que um elemento minimal pode não existir.

Um estudo que, sem dúvida, será necessário para a execução dessa extensão é o que diz respeito a teorias de unificação. Em [CDT91], quando da extensão do método abdutivo proposto, para a lógica de primeira ordem, a teoria equacional de Clark (*CET - Clark equational theory*) é utilizada (veja seção 3.3.1). Já em [MP95b] é utilizado o processo de skolemização reversa (veja seção 3.3.2). Certamente, o estudo desses e outros métodos com a mesma finalidade será necessário a fim de definirmos a abdução de primeira ordem.

Um ponto crucial com que certamente nos depararemos, é a questão da minimalidade. Para gerarmos candidatos não minimais, basta termos em mente a quantificação universal dos mesmos. No entanto, explicações desse tipo não nos parecem muito úteis, e uma das grandes questões da abdução de primeira ordem vem a ser, justamente, a de quando é possível quantificar apenas existencialmente e quando a quantificação universal é indispensável.

A questão acima está diretamente relacionada com o número de candidatos gerados. Isto é, os candidatos que primeiro aparecem na tentativa de construção da prova são candidatos instanciados. Como ampliar esse número de candidatos e como quantificá-los?

Um ponto que certamente nos chama a atenção é quanto ao relacionamento entre abdução e indução. A utilização de quantificadores universais, em situações específicas, está diretamente relacionada com o processo de indução (como vimos na seção 2.4). Assim, o processo indutivo poderá também ser especificado formalmente utilizando-se a formalização da abdução para a lógica de primeira ordem (consideradas as devidas restrições).

Uma outra questão é que o que aqui chamamos de generalização de candidatos (veja seção 5.4.1), para a lógica de primeira ordem parece estar relacionado com a quantificação dos candidatos universalmente.

Por último, vale ressaltar que a comparação dos candidatos (a nível meta-lógico), a princípio, dar-se-á de maneira similar ao caso da lógica proposicional.

Capítulo 6

Conclusão

Neste capítulo apresentaremos as principais conclusões obtidas com o trabalho proposto, as principais vantagens, as atuais deficiências. Em seguida relacionaremos as propostas de trabalhos futuros relacionados com o tema da dissertação, a maioria delas envolvendo extensões da teoria aqui proposta.

6.1 Considerações finais

A importância da modelagem do raciocínio abdutivo é indiscutível, seja do ponto de vista puramente teórico (a busca de uma caracterização de 'Lógica' com 'L' maiúsculo, tal que não se restrinja à dedução e/ou indução, mas incorpore também a abdução), seja do ponto de vista do seu enorme potencial de aplicação. A sua aplicabilidade em diversas áreas comprova essa importância e atrai o investimento em pesquisas abordando o tema. É também indiscutível que a ausência de definições e terminologias consensuais que contribuam para a interação das pesquisas na área é uma das principais dificuldades a serem suplantadas. Diante dessa perspectiva fica clara a importância do tema abordado nesta dissertação.

A maneira como abordamos a formalização do raciocínio abdutivo, tem como objetivo principal ressaltar a conveniência da definição da abdução como composta de dois níveis básicos: o nível objeto e o meta-nível. Para o nível objeto é proposto um método dedutivo de geração de candidatos abdutivos que nos parece bastante claro e simples, e que apresenta grandes semelhanças com outros trabalhos propostos na área. O que destaca o nosso método em relação a outras abordagens é justamente o tratamento do meta-nível.

A maioria das abordagens em abdução trata o problema até a fase da geração dos candidatos à explicação. Algumas delas mencionam a importância da escolha de hipóteses interessantes dentre as muitas geradas; outras vão mais além e colocam alguns critérios para selecionar as melhores hipóteses (como é o caso de [MD94]). Nossa proposta encaixa-se neste último grupo, e essa é sua característica principal e mais vantajosa: propomos critérios meta-lógicos para selecionar as hipóteses mais plausíveis, conseguindo, assim, fornecer ao usuário um conjunto de possíveis hipóteses com um menor número de elementos que o conjunto inicialmente gerado.

Além dessa principal característica, a utilização do *LDS* nos favorece em outros pontos: podemos pensar em lidar com diferentes lógicas, sem que seja necessário alterações no formalismo; e podemos fazer a seleção das hipóteses de acordo com o critério de preferência que seja mais relevante à área de aplicação do problema (já que a relação de preferência e seu significado é dado pelo usuário).

As principais contribuições do trabalho que culmina nesta dissertação, poderiam ser enumeradas assim:

- A geração de candidatos é feita dedutivamente utilizando como base a dedução natural e sendo, por isso, um método bastante intuitivo;

- No nível objeto são tratadas as condições de minimalidade, consistência com a teoria, restrições de integridade e conseqüências dedutivas das hipóteses. Todas essas condições são resolvidas dedutivamente através de métodos semelhantes aos propostos em diversas abordagens em abdução;

- A utilização dos rótulos, que incorporam ao sistema uma informação adicional (proveniente do meta-nível) relevante ao problema em questão, permite que sejam levados em consideração aspectos extra-lógicos da abdução; o que resulta no procedimento de comparação dos candidatos cujo objetivo principal é reduzir o conjunto de hipóteses abdutivas inicialmente gerado;

- A ordenação dos rótulos, que deve refletir um critério de preferência entre as fórmulas da teoria (seja custo, incerteza, probabilidade, etc.) fornecido pelo usuário, favorece a aplicação do método a problemas de diferentes domínios, já que esse critério pode variar de caso para caso;

- Com a utilização da estrutura do *LDS*, abre-se a possibilidade de utilização de lógicas diferentes dentro do mesmo formalismo abdutivo.

De modo geral, todo o trabalho aqui elaborado busca uma caracterização da abdução que seja genérica o suficiente para abranger a maioria dos aspectos do raciocínio abdutivo que são abordados nos trabalhos desenvolvidos (e estudados) na área. Assim, a caracterização cujos primeiros elementos são apresentados nesta dissertação pretende contribuir com a construção de uma conceituação consensual da abdução.

Sem dúvidas, a formalização ainda apresenta muitas limitações; algumas delas são as seguintes:

- O método limita-se à abdução para a lógica proposicional;

- A linguagem utilizada restringe-se à lógica clássica implicacional ou à lógica intuicionística, devido às limitações da dedução natural para tratar a lei do terceiro excluído;

- Um algoritmo para a construção das deduções em dedução natural precisa ser sugerido, salientando-se a necessidade da construção apenas de deduções na forma normal;

- As condições tratadas no nível objeto (em particular, a geração do conjunto de conseqüências dedutivas das hipóteses) devem ser melhor especificadas a fim de que os detalhes computacionais sejam explicitados e solucionados;

- Os teoremas propostos (teorema 5.1 - completude e corretude da geração de candidatos - e teorema 5.2 - corretude da generalização de candidatos) devem ser revisados e devidamente provados e estendidos;

- Os critérios propostos para a comparação de candidatos devem ser especificados algoritmicamente e novos critérios devem ser investigados, como, por exemplo, critérios semelhantes às heurísticas propostas em [MD94].

Os principais pontos para trabalhos futuros são:

- Extensão para a lógica clássica proposicional e para a lógica de primeira ordem;

- Especificação de algoritmos para a efetivação das diversas etapas do método e o estudo de suas complexidades;

- Uma análise das propriedades que o método apresenta, considerando, inicialmente, as propriedades propostas em [MP95b];

- Aplicação do método a casos específicos, provendo estudos de casos onde seriam definidos critérios de comparação de candidatos específicos para o caso em questão;

- Uma investigação do comportamento do método proposto com a utilização de lógicas não clássicas (como linear, relevante, etc.);

- Uma investigação da aplicação do método aqui definido, para tratar a interpretação de linguagem natural. Tal investigação possivelmente teria como ponto de partida o estudo do trabalho proposto em [Gor95] onde é abordada a análise de elementos da linguagem natural utilizando o *LDS*;

- Um estudo da ligação entre o raciocínio abdutivo e a semântica via jogos; tal relação é abordada em [HH91, Hin91], e esse estudo poderia ter início na abordagem de sistemas de jogos rotulados proposta em [Dur95].

Referências Bibliográficas

[Add88] T. R. Addis. A knowledge organisation for abduction. In *Proceedings of the interdisciplinary information technology conference*. Bradford University, April 1988. KSG/KTM/TR/8.

[Bar81] Henk P. Barendregt. *The Lambda Calculus, its Syntax and Semantics.*, volume 103 of *Studies in Logic and the Foundation of Mathematics*. North-Holland. Amsterdam. New York. Oxford., revised edition, 1981. Edited by Barwise, J. , Kaplan, D., Keisler, H.J., Suppes, P., Troelstra, A. S.

[BP91] Massimo A. Bonfantini and Giampaolo Proni. Suposição: sim ou não? Eis a questão. In *O signo de três*. Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês: *The sign of three*. Indiana University Press, 1983.

[Car91] Gian Paolo Carettini. Peirce, Holmes e Popper. In *O signo de três*. Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês : *The sign of three*. Indiana University Press, 1983.

[CDT91] Luca Console, Daniele Theseider Dupre, and Pietro Torasso. On the relationship between abduction and deduction. *Journal of Logic and Computation*, 1(5):661–690, 1991.

[CMM83] Jaime G. Carbonell, Ryszard S. Michalski, and Tom M. Mitchell. An overview of machine learning. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning, an artificial intelligence approach*, volume 1, chapter 1, pages 3–23. Morgan Kaufmann Publishers, Inc, 1983.

[dC93] Newton C. A. da Costa. *Lógica indutiva e probabilidade*. Hucitec: Edusp, 1993.

[DdC92] Robert Demolombe and Luis Fariñas del Cerro. An inference rule for hypothesis generation. *Proceedings of IJCAI*, 1992.

[DLP93] D. Dubois, J. Lang, and H. Prade. Possibilistic logic. *Handbook of Logic in Artificial Inteligence and Logic Programming*, pages 439–513, 1993. edited by D. Gabbay, C. Hogger & J. Robinson, Oxford University Press.

[DM83] Thomas G. Dietterich and Ryszard S. Michalski. A comparative review of selected methods for learning from examples. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning, an artificial intelligence approach*, volume 1, chapter 3, pages 41–81. Morgan Kaufmann Publishers, Inc, 1983.

[dMH93] Octanny Silveira da Mota and Leonidas Hegenberg. Introdução. In Octanny Silveira da Mota e Leonidas Hegenberg, editor, *Semiótica e filosofia, textos escolhidos*. Editora Cultrix, SP, 9 edition, 1993. 3. edição 1984.

[DP90] D. Dubois and H. Prade. Modeling uncertain and vague knowledge in possibility and evidence theories. In R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, editors, *Uncertaint in Artificial Intelligence 4*, volume 9, pages 303–318. North Holland, 1990. Machine Intelligence and Patern Recognition Series.

[dQG92] Ruy J. G. B. de Queiroz and Dov M. Gabbay. An introduction to labelled natural deduction. *Proc 3dr Adv. Summer Sch. in AI, Sept 21-25'92*, 1992. Paper available via anonymous ftp from theory.doc.ic.ac.uk, file intro-lnd.{dvi,ps}.gz in /guests/deQueiroz.

[Dur95] Adolfo Almeida Duran. *Semântica via Jogos para a Dedução Rotulada*. Dissertação de Mestrado, Departamento de Informática da Universidade Federal de Pernambuco, Julho, 1995. (to appear).

[Eco91] Umberto Eco. Chifres, cascos, canelas: algumas hipóteses acerca de três tipos de abdução. In *O signo de três*. Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês: *The sign of three*. Indiana University Press, 1983.

[Fis92] Olivier Fischer. Cognitively plausible heuristics to tackle complexity of abductive reasoning. *AI Magazine*, 13(2):46–49, 1992. PhD dissertation abstract.

[For89] Richard Forsyth. The logic of induction. In Richard Forsyth, editor, *Machine learning - principles and techniques*, chapter 1. Chapman and Hall Computing, 1989.

[Gab94] Dov M. Gabbay. *LDS - Labelled Deductive Systems, Volume I - Foundations.* Oxford University Press, 1994. First Draft 1989. Current Draft, 465pp.,May 1994. Published as MPI-I-94-223, Max-Planck-Institut für Informatik, Saarbrücken, Germany.

[GdQ92] Dov M. Gabbay and Ruy J. G. B. de Queiroz. Extending the Curry-Howard interpretation to linear, relevant and other resource logics. *The Journal of Symbolic Logic*, 1992.

[Gen35] Gerhard Gentzen. Untersuchungen über das logische schliessen. *Mathematische Zeitschrift*, pages 176:210 and 405–431, 1935. English translation:"Investigations into Logical Deduction" in "The Collected Works of Gerhard Gentzen", ed. M.E.Szabo, North-Holland Pub Co.,1969.

[Gil94] Donald Gillies. A rapprochement between deductive and inductive logic. *Bulletin of the Interest Group in Pure and Apllied Logics*, 2(2):149–166, 1994.

[Gin91] Carlo Ginsburg. Chaves do mistério: Morelli, Freud e Sherlock Holmes. In *O signo de três.* Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês : *The sign of three.* Indiana University Press, 1983.

[GKP94] Dov M. Gabbay, Ruth Kempson, and Jeremy Pitt. Labelled abduction and relevance reasoning. In R. Demolombe and T. Imielinski, editors, *Nonstandard queries and nonstandard answers*, Studies in Logic and Computation - 3, chapter 7, pages 155–185. Oxford University Press, 1994. Series editor: D. M. Gabbay.

[Gor95] Sérgio Gorender. *Sobre a Representação de Linguagem Natural usando Dedução Natural Rotulada.* Dissertação de Mestrado, Departamento de Informática da Universidade Federal de Pernambuco, Maio, 1995.

[Har91] Nancy Harrowitz. O arcabouço do modelo de detetive: Charles S. Peirce e Edgar Alan Poe. In *O signo de três.* Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês: *The sign of three.* Indiana University Press, 1983.

[HH91] Jaakko Hintikka and Merril B. Hintikka. Sherlock Holmes em confronto com a lógica moderna: para uma teoria de obtenção de informação através do questionamento. In *O signo de três.* Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês: *The sign of three.* Indiana University Press, 1983.

[Hin91] Jaakko Hintikka. Sherlock Holmes formalizado . In *O signo de três*. Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês: *The sign of three*. Indiana University Press, 1983.

[How80] W. A. Howard. *"The Formulae-as-types Notion of Construction". In "To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism"*. Academic Press, 1980. xxv+606pp.

[Kyb90a] H. E. Kyburg Jr. Epistemological relevance and statistical knowledge. In R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, editors, *Uncertaint in Artificial Intelligence 4*, volume 9, pages 159–168. North Holland, 1990. Machine Intelligence and Patern Recognition Series.

[Kyb90b] H. E. Kyburg Jr. Probabilistic inference and non-monotonic inference. In R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, editors, *Uncertaint in Artificial Intelligence 4*, volume 9, pages 319–326. North Holland, 1990. Machine Intelligence and Patern Recognition Series.

[Kyb93] Henry E. Kyburg Jr. Uncertainty logics. *Handbook of Logic in Artificial Inteligence and Logic Programming*, pages 397–434, 1993. edited by D. Gabbay, C. Hogger & J. Robinson, Oxford University Press.

[Lak68] Imre Lakatos. Changes in the problem of inductive logic. In Imre Lakatos, editor, *Problems in the philosophy of mathematics*, Amsterdam, 1968. North-Holland.

[LZSB86] Pat Langley, Jan M. Zytkow, Herbert A. Simon, and Gary L. Bradshaw. The search of regularity: four aspects of scientific discovery. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning, an artificial intelligence approach*, volume 2, chapter 16, pages 425–469. Morgan Kaufmann Publishers, Inc, 1986.

[MAL+86] Ryszard S. Michalski, Saul Amarel, Douglas B. Lenat, Donald Michie, and Patrick H. Winston. Understanding the nature of learning: Issues and research directions. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning: challenges of the eighties*, volume 2, chapter 2, pages 27–41. Morgan Kaufmann Publishers, Inc, 1986.

[MD94] Philippe Mathieu and Jean-Paul Delahaye. Abductive reasoning in three-valued logic for knowledge bases. In R. Demolombe and T. Imielinski, editors, *Nonstandard queries and* *nonstandard answers*, Studies in Logic and Computation - 3, chapter 6, pages 131–153. Oxford University Press, 1994. Series editor: D. M. Gabbay.

[Mic83] Ryszard S. Michalski. A theory and methodology of inductive learning. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning, an artificial intelligence approach*, volume 1, chapter 4, pages 83–134. Morgan Kaufmann Publishers, Inc, 1983.

[Mic86] Ryszard S. Michalski. Understanding the nature of learning: Issues and research directions. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning, an artificial intelligence approach*, volume 2, chapter 1, pages 3–25. Morgan Kaufmann Publishers, Inc, 1986.

[MP93] Marta Cialdea Mayer and Fiora Pirri. First order abduction via tableau and sequent calculi. *Bulletin of the Interest Group in Pure and Applied Logics*, 1(1):99–117, 1993.

[MP95a] Marta Cialdea Mayer and Fiora Pirri. Propositional abduction in modal logic. *Bulletin of the Interest Group in Pure and Applied Logics*, 1995. Special Issue on *Proc. Tableaux Workshop 1994*, (to appear).

[MP95b] Marta Cialdea Mayer and Fiora Pirri. Towards a logical characterization of abductive reasoning. *Bulletin of the Interest Group in Pure and Applied Logics*, 1995. Special Issue on *Mechanised Deduction in the Logics of Practical Reasoning*, (to appear).

[Pei80] Charles Sanders Peirce. Escritos coligidos. In *Os pensadores*. Abril Cultural, 9 edition, 1980. Seleção: Armando Mora D'Oliveira.

[Pei93a] Charles Sanders Peirce. Dedução, indução e hipótese. In Octanny Silveira da Mota e Leonidas Hegenberg, editor, *Semiótica e filosofia, textos escolhidos*. Editora Cultrix, SP, 9 edition, 1993. 3. edição 1984.

[Pei93b] Charles Sanders Peirce. A propósito do autor. In Octanny Silveira da Mota e Leonidas Hegenberg, editor, *Semiótica e filosofia, textos escolhidos*. Editora Cultrix, SP, 9 edition, 1993. 3. edição 1984.

[Pra65] D. Prawitz. Natural Deduction. A Proof-Theoretical Study. In *Acta Universitatis Stockholmiensis. Stockholm Studies in Philosophy*, volume 3, page 113pp. Almqvist and Wiksell, Stockholm, 1965.

[Pra71] D. Prawitz. Ideas and results in proof theory. *Proceedings of the Second Scandinavian Logic Symposium, Ed. J. E. Festadnn, North-Holland.*, pages 235–307, 1971.

[RdQ95a] Nicia Cristina Rocha Riccio and Ruy J. G. B. de Queiroz. Abductive reasoning via labelled deductive systems. In *Abstracts of the 10th International Congress of Logic Methodology and Philosophy of Science*, August 1995. (to appear).

[RdQ95b] Nicia Cristina Rocha Riccio and Ruy J. G. B. de Queiroz. Labelled abduction. In *Proceedings of XII Brazilian Conference on Artificial Intelligence (SBIA'95)*, 1995. (to appear).

[Seb91] Thomas A. Sebeok. Um, dois, três, uberdade desta vez. In *O signo de três.* Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês : *The sign of three.* Indiana University Press, 1983.

[Sim83] Herbert A. Simon. Why should machine learn? In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning, an artificial intelligence approach*, volume 1, chapter 2, pages 25–37. Morgan Kaufmann Publishers, Inc, 1983.

[Spo90] W. Spohn. A general non-probabilistic theory of inductive reasoning. In R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, editors, *Uncertaint in Artificial Intelligence 4*, volume 9, pages 149–158. North Holland, 1990. Machine Intelligence and Patern Recognition Series.

[SS91] Thomas A. Sebeok and Jean Umiker Sebeok. "Você conhece o meu método": uma justaposição de Charles S. Peirce e Sherlock Holmes. In *O signo de três.* Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês: *The sign of three.* Indiana University Press, 1983.

[Tar56] A. Tarski. On the concept of logical consequence (Polish 1936). *English translation in: Logic, Semantics, Metamathematics*, 1956. Oxford University Press.

[Ton95] Francesca Toni. *Abductive Logic Programming.* PhD thesis, University of London, February 1995. Department of Computing, Imperial College of Science, Technology and Medicine.

[Tru91] Marcelo Truzzi. Sherlock Holmes: psicólogo social aplicado. In *O signo de três.* Editora Perspectiva, 1991. Organizadores: Umberto Eco e Thomas A. Sebeok. Do original em inglês: *The sign of three.* Indiana University Press, 1983.

[Win86] Patrick H. Winston. Learning by augmenting rules and accumulating censors. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning, an artificial intelligence approach*, volume 2, chapter 3, pages 43–61. Morgan Kaufmann Publishers, Inc, 1986.

[Zad93] Wlodek Zadrozny. On rules of abduction. *Annals of Mathematics and Artificial Intelligence*, 9:387–419, 1993.

[Zad94] Wlodek Zadrozny. Is there a prototypical rule of abduction? (yes, e.g. in proximity based explanation). *J. Expt. Theor. Artif. Intell.*, 6:147–162, 1994.