



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-10425-TDI/922

**UMA TEORIA DE COBERTURAS NEBULOSAS PARA
DIAGNÓSTICO, INVESTIGAÇÃO E TRATAMENTO**

Silvia Maria Fonseca Silveira Massruhá

Tese de Doutorado do Curso da Pós-Graduação em Computação Aplicada, orientada
pelos Drs. Sandra Aparecida Sandri e Jacques Wainer, aprovada em
12 de dezembro de 2003.

681.3.019

MASSRUHÁ, S. M. F. S.

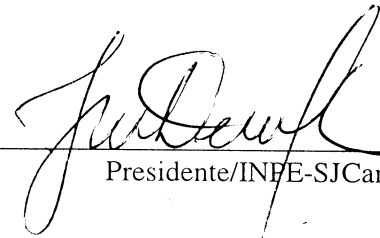
Uma teoria de coberturas nebulosas para diagnóstico, investigação e tratamento / S. M. F. S. Massruhá. – São José dos Campos: INPE, 2003.

251p. – (INPE-10425-TDI/922).

1.Diagnóstico. 2.Investigação. 3.Tratamento. 4.Com-juntos nebulosos. 5.Tomada de decisão. 6.Inferência abdu-tiva. 7.Ferramenta automatizada para diagnóstico. 8.Inteli-gência artificial. I.Título.

Aprovada pela Banca Examinadora em cumprimento a requisito exigido para a obtenção do Título de **Doutora em Computação Aplicada.**

Dr. José Demisio Simões da Silva



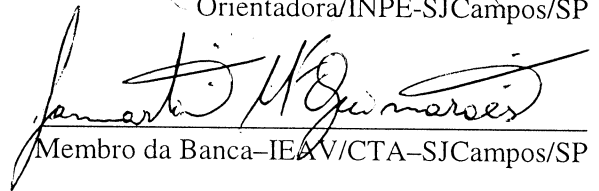
Presidente/INPE-SJCampos/SP

Dr^a Sandra Aparecida Sandri



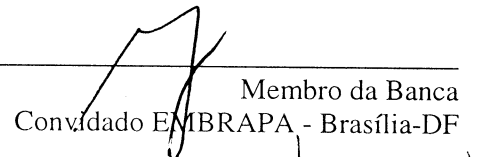
Orientadora/INPE-SJCampos/SP

Dr. Lamartine Nogueira Frutuoso Guimarães




Membro da Banca-IEAV/CTA-SJCampos/SP

Dr. Moacir Pedroso Júnior



Membro da Banca
Convidado EMBRAPA - Brasília-DF

Dr. Flávio Soares Correa da Silva



Membro da Banca
Convidado USP-São Paulo/SP

Candidato (a): Silvia Maria Fonseca Silveira Massruhá

São José dos Campos, 12 de dezembro de 2003.

“ Ter idéias é reunir flores. Pensar é tecer as flores em ramalhetes ”.
Mme.Suetchine

Aos meus pais, **Francisco e Nezita**, ao meu marido **Ério** e minha filha **Karina** que,
cada um a seu modo, contribuíram para o meu sucesso, dedico esta tese.

AGRADECIMENTOS

A Deus, pela força maior, e por Ter permitido que eu convivesse com pessoas tão sábias que contribuíram para o meu crescimento.

À Profa. Doutora SANDRA SANDRI, Orientadora sempre presente, com sugestões valiosas e estímulo constante, pelo exemplo de sabedoria, equilíbrio e competência profissional.

Ao Prof. Doutor JACQUES WAINER, Co-Orientador, que com suas observações possibilitou o meu crescimento como pesquisadora.

Aos Professores do INPE, pelos diálogos mantidos ao longo do doutorado.

À EMBRAPA, pela crença no meu potencial profissional.

Aos colegas da Embrapa Informática Agropecuária, especialmente, ao Dr. Kleber Xavier Sampaio de Souza pelo apoio como conselheiro acadêmico.

Ao fitopatologista Dr. Marcelo Morandi da Embrapa Meio Ambiente e ao meu cunhado médico Dr. Ícaro Zafalon Bozza pelas contribuições nas soluções de dúvidas na área de diagnóstico durante todo o processo de desenvolvimento da tese.

Aos meus pais, Francisco e Nezita que, desde os primeiros anos de minha vida, me ensinaram a entender os mistérios do existir e sempre dedicaram carinho, atenção e incentivo a todos os meus desafios.

Ao meu esposo Ério e minha filha Karina pelo apoio, afeto e compreensão sem os quais seria impossível a realização de mais este passo.

Aos meus irmãos Luiz Fernando, Vânia, Júnior e Carlos Roberto, aos meus cunhados Carlos, Adriane e Eliane e às tias Ana Maria, 'Julinha' e Dinha cuja amizade sempre me foram imprescindíveis.

Aos meus sobrinhos Camila, Carlos Henrique, André, Amanda e Clara e meu afilhado Marcelo esperando que, de alguma forma, este trabalho os ajude a viver melhor.

RESUMO

Este trabalho tem por objetivo apresentar uma abordagem integrada para diagnóstico, investigação e tratamentos, permitindo organizar os vários tipos de informação gerados nestas etapas durante processo de raciocínio clínico. Nesta nova abordagem, denominada Teoria das Coberturas Nebulosas (TCN), o conhecimento é basicamente modelado através de associações causais e a inferência é abdutiva. Conceitos de Parsimonious Covering Theory (PCT), lógica nebulosa e teoria de decisão são também integrados, de maneira a tratar os vários aspectos inerentes aos processos envolvidos em raciocínio clínico, tais como a possibilidade de diversas desordens estarem conjuntamente causando um conjunto de manifestações, a manipulação de informações temporais, a consideração de condições favoráveis ao desenvolvimento de uma desordem, a incapacidade do especialista em oferecer conhecimento generalizado desprovido de incerteza e/ou imprecisão, a manipulação de fatores cruciais na tomada de decisão nas tarefas de investigação e tratamento como o custo e o risco. A validação do modelo teórico foi realizada em fitopatologia, abordando especificamente a diagnose, investigação e tratamento de doenças de milho. Entretanto, devido a sua generalidade, espera-se que os desenvolvimentos teóricos obtidos possam ser aplicados para resolver problemas de diagnóstico em outras áreas de conhecimento.

A FUZZY COVERING THEORY FOR DIAGNOSIS, INVESTIGATION AND TREATMENT

ABSTRACT

This work presents an integrated framework for diagnosis, investigation and treatment, allowing the organization of various pieces of information generated in these tasks during a clinical problem-solving process. In this new framework, called Fuzzy Covering Theory (FCT), knowledge is basically modeled through causal associations and inference is abductive. Concepts of Parsimonious Covering Theory (PCT), fuzzy sets logic and decision theory are also integrated, in order to address the various inherent aspects involved in clinical reasoning, such as the possibility of several disorders to conjunctively cause a series of manifestations, the manipulation of temporal information, the consideration of favorable conditions for the development of a disorder, the incapacity of an expert to yield generalized knowledge devoid of uncertainty/imprecision, the manipulation of crucial factors in decision making in investigation and treatment tasks like cost and risk. The validation of the theoretical model has been performed in phytopathology, specifically, at the diagnosis, investigation and treatment of the corn plants diseases. However, due to its generality, it is expected that the theoretical developments obtained in this work can successfully be employed to solve diagnosis problems in other domains.

SUMÁRIO

	<u>Pág.</u>
LISTA DE FIGURAS	17
LISTA DE TABELAS	23
CAPÍTULO 1 – INTRODUÇÃO	27
1.1 – Motivação do Trabalho	29
1.1.1 – Diagnóstico de Doenças de Milho	30
1.1.2 – Diagnóstico de Falhas em Satélite	31
1.2 – Escopo do Trabalho	32
1.3 – Esboço Geral	33
CAPÍTULO 2 – MODELOS COMPUTACIONAIS PARA SOLUÇÃO DE PROBLEMAS DE DIAGNÓSTICO E TRATA- MENTOS	35
2.1 – Sistemas para Solução de Problemas Baseados em Conhecimento	35
2.1.1 – Representação do Conhecimento	36
2.1.2 – Mecanismos de Inferência	37
2.2 – Modelos Computacionais na Área de Diagnóstico	38
2.2.1 – Sistemas Estatísticos de Classificação: Redes Bayesianas	38
2.2.2 – Sistemas de Produção Baseados em Regras	40
2.2.3 – Sistemas Abdutivos Baseado em Associações	41
2.3 – Modelos Computacionais na Área de Tratamentos	43
2.3.1 – Matrizes de Retorno	45
2.3.2 – Árvores de Decisão	46
2.3.3 – Diagramas de Influência	48
CAPÍTULO 3 – MODELOS ABDUTIVOS	51
3.1 – Ambiente Abduativo	51
3.2 – Teoria das Coberturas Parcimoniosas (TCP)	53
3.2.1 – Formulação do Problema de Diagnóstico	54
3.2.2 – Soluções para o Problema de Diagnóstico	57
3.2.3 – Algoritmos	58
3.2.4 – Limitações e Extensões da TCP	61
3.3 – Extensão Nebulosa à TCP Temporal	63

3.3.1 – Base de Conhecimento e o Caso Particular	64
3.3.2 – Consistência entre o Modelo de Desordens e o Caso	68
3.3.3 – Explicação para o Diagnóstico	72
3.3.4 – Limitações	72
CAPÍTULO 4 – MODELO PROPOSTO	75
4.1 – Escopo do Problema	75
4.1.1 – Processo de Solução de Problemas Clínicos	76
4.1.2 – Pré-diagnóstico	77
4.1.3 – Investigação	78
4.1.4 – Diagnóstico	79
4.1.5 – Prognóstico e Tratamento	79
4.2 – Abordagens Adotadas na Literatura para o Problema	81
4.3 – Abordagem Integrada para Diagnóstico e Tratamento	84
4.4 – Formalização do Problema	89
4.5 – Solução Final para o Problema de Raciocínio Clínico	91
CAPÍTULO 5 – AMBIENTE ABDUTIVO PARA DIAGNÓSTICO	99
5.1 – Representação do Conhecimento em Diagnóstico	99
5.1.1 – Teoria causal	99
5.1.2 – Evolução Temporal	101
5.1.3 – Condições de Controle	102
5.1.4 – Base de Conhecimento	104
5.2 – Especificação do Caso particular	105
5.3 – Solução para o Problema de Diagnóstico	106
5.3.1 – Consistência entre o Modelo de Desordens e o Caso	106
5.3.2 – Solução para o Diagnóstico	111
CAPÍTULO 6 – AMBIENTE ABDUTIVO PARA INVESTIGAÇÃO	113
6.1 – Representação do Conhecimento em Investigação	113
6.2 – Especificação do Caso particular	115
6.3 – Tomada de Decisão para Investigação	116
6.3.1 – Consistência entre o Modelo de Exames e o Caso	116
6.3.2 – Funções de Entropia	117
6.3.3 – Modelo das Distribuições	124
6.3.4 – Estudo de Caso	126
6.3.5 – Solução para Investigação	131

CAPÍTULO 7 – AMBIENTE ABDUTIVO PARA TRATAMENTOS .	135
7.1 – Representação do Conhecimento para Tratamentos	135
7.2 – Especificação do Caso Particular	138
7.3 – Tomada de Decisão em Tratamentos	139
7.3.1 – Consistência entre o Modelo de Tratamentos e o Caso	139
7.3.2 – Solução para Tratamentos	144
CAPÍTULO 8 – VALIDAÇÃO DA ABORDAGEM PROPOSTA	145
8.1 – Diagnóstico e Controle de Desordens em Cultura de Milho	146
8.1.1 – Descrição do Problema	146
8.1.2 – Representação e Implementação da Base de conhecimento KB_d	147
8.1.3 – Representação e Implementação da Base de conhecimento KB_e	149
8.1.4 – Representação e Implementação da Base de conhecimento KB_t	151
8.1.5 – Representação de um Caso Particular	153
8.1.6 – Solução para o Problema de Raciocínio Clínico	154
8.2 – Diagnóstico e Controle em Falhas de Satélite	160
8.2.1 – Descrição do Problema	160
8.2.2 – Representação do Conhecimento	160
8.2.3 – Solução do Problema	162
8.3 – Considerações Finais sobre a Abordagem Proposta	164
CAPÍTULO 9 – CONCLUSÕES	171
9.1 – Contribuições	171
9.2 – Limitações	173
9.3 – Trabalhos Futuros	173
REFERÊNCIAS BIBLIOGRÁFICAS	175
APÊNDICE A – CONCEITOS BÁSICOS DA TEORIA DE CONJUN-	
TOS NEBULOSOS E TEORIA DE POSSIBILIDADES	179
A.1 – Teoria de Conjuntos Nebulosos	179
A.2 – Teoria de Possibilidades	182
APÊNDICE B – IMPLEMENTAÇÃO DO MODELO PROPOSTO . .	185
B.1 – Algoritmos	185
B.1.1 – Algoritmos para Diagnóstico	186
B.1.2 – Algoritmos para Investigação	188
B.1.3 – Algoritmos para Tratamentos	188

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Uma rede de inferência para um sistema baseado em regras.	41
2.2 Exemplo de árvore de decisão.	47
2.3 Exemplo de árvore de decisão.	48
2.4 Diagrama de influências.	49
3.1 Modelo de uma rede causal em TCP.	54
3.2 Modelo de um grafo temporal associado a uma desordem.	63
3.3 Modelo de um grafo de eventos associado a uma desordem no escopo da TCN.	66
3.4 Modelo parcial de um grafo mínimo associado a uma desordem d_i	67
3.5 Exemplo do grau de consistência temporal.	69
4.1 As decisões diagnóstica e terapêutica.	93
4.2 Escopo da TCN.	94
4.3 Um ambiente integrado para diagnose e tratamento.	95
4.4 Integração dos processos de diagnose e tratamento.	96
4.5 Um ambiente abduativo genérico integrado para diagnose e tratamento.	97
4.6 Síntese do processo de raciocínio clínico suportado pela TCN.	98
5.1 Um ambiente abduativo causal para diagnóstico.	99
5.2 Associação causal da ferrugem comum (d_3) e da ferrugem <i>polysora</i> (d_4).	101

5.3	Modelo de um grafo de eventos associado a uma desordem no escopo da TCN.	102
5.4	Modelo parcial de um grafo mínimo associado a uma desordem no escopo da TCN.	103
5.5	Representação nebulosa para os graus de temperatura.	104
5.6	Exemplo do grau de consistência temporal.	108
6.1	Um ambiente abdutivo para investigação.	114
6.2	Distribuição da plausibilidade das hipóteses, supondo-se a confirmação de (a) m_1 e de (b) m_2	118
6.3	Vetores OWA criados usando: (a) $g(i) = i$ (b) $g(i) = i^2$	120
6.4	Exemplos de distribuições δ	122
6.5	Distribuições δ normalizadas.	132
6.6	Distribuições δ não normalizadas.	133
7.1	Um ambiente abdutivo para tratamentos.	136
7.2	Intervalos nebulosos referentes à $Retorno(t_1)$ e $Retorno(t_2)$	143
8.1	Escala fenológica da cultura do milho.	147
8.2	Grafo temporal da cultura do milho.	168
8.3	Associação gradual $R = \langle d_{10}, m_j \rangle$	169
8.4	Grafo mínimo de eventos de d_{10}	169
8.5	Exemplo de uma árvore de decisão no tratamento da Cercosporiose.	170
8.6	Rede causal de um problema de diagnóstico de falhas em satélite $\langle D, M, C \rangle$	170
A.1	Intervalo nebuloso trapezoidal.	180

C.1	Escala Fenológica da Cultura do Milho.	191
C.2	PODRIDÃO DO CARTUCHO POR <i>Erwinia</i> (d_1).	198
C.3	QUEIMA POR <i>Pseudomonas</i> (d_2).	199
C.4	FERRUGEM COMUM (d_3).	200
C.5	FERRUGEM COMUM - Variáveis de Decisão (d_3).	201
C.6	FERRUGEM <i>polysora</i> (d_4).	202
C.7	FERRUGEM <i>polysora</i> - Variáveis de Decisão (d_4).	203
C.8	FERRUGEM BRANCA OU TROPICAL (d_5).	204
C.9	FERRUGEM BRANCA OU TROPICAL - Variáveis de Decisão (d_5).	205
C.10	MANCHA POR <i>Physoderma</i> (d_6).	206
C.11	ANTRACNOSE (d_7).	207
C.12	MANCHA POR <i>Helminthosporium maydis</i> Raça O (d_8).	208
C.13	MANCHA POR <i>Helminthosporium maydis</i> Raça T (d_9).	209
C.14	MANCHA POR <i>Cercospora</i> (d_{10}).	210
C.15	MANCHA POR <i>Cercospora</i> - Variáveis de Decisão (d_{10}).	211
C.16	MANCHA POR <i>Diplodia Macrospora</i> (d_{11}).	212
C.17	MANCHA POR <i>Diplodia Macrospora</i> - Variáveis de Decisão (d_{11}).	213
C.18	MANCHA POR <i>Helminthosporium Turcicum</i> (d_{12}).	214
C.19	MANCHA POR <i>Helminthosporium Turcicum</i> - Variáveis de Decisão(d_{12}).	215
C.20	MANCHA POR <i>Physoderma</i> (d_{13}).	216
C.21	MANCHA POR <i>Curvularia</i> (d_{14}).	217
C.22	MANCHA POR <i>Helminthosporium Carbonum</i> (raça 1)(d_{15}).	218

C.23 MANCHA POR <i>Helminthosporium Carbonum</i> (raça 2)(d_{16}).	219
C.24 MANCHA POR <i>Helminthosporium Carbonum</i> (raça 3)(d_{17}).	220
C.25 MANCHA POR <i>Phaeosphaeria</i> (d_{18}).	221
C.26 MANCHA POR <i>Phaeosphaeria</i> (d_{18}) - Variáveis de Decisão.	222
C.27 Queima por Herbicida (d_{19}).	223
C.28 Míldio do Sorgo (d_{20}).	224
C.29 Virose Rayado Fino(d_{21}).	225
C.30 Virose Rayado Fino(d_{21}).	226
C.31 Mosaico Comum do Milho(d_{22}).	227
C.32 Efeito de Frio(d_{23}).	228
C.33 Deficiência de Fósforo (d_{24}).	229
C.34 Enfezamento Vermelho(d_{25}).	230
C.35 Enfezamento Vermelho(d_{25}) - Variáveis de Decisão.	231
C.36 Enfezamento Pálido(d_{26}).	232
C.37 Enfezamento Pálido(d_{26}) - Variáveis de Decisão.	233
C.38 Carvão Comum da Espiga(d_{27}).	234
C.39 Causas Diversas(d_{28}).	235
C.40 Podridão Rosada da Espiga (<i>Fusarium</i>) (d_{29}).	236
C.41 Podridão Rosada da Espiga (<i>Fusarium</i>) (d_{29}) - Variáveis de Decisão.	237
C.42 Podridão da Espiga por <i>Gibberella</i> (d_{30}).	238
C.43 Podridão Branca da Espiga (<i>Diplodia Maydis</i>) (d_{31}).	239
C.44 Podridão da Espiga (<i>Helminthosporium Maydis</i>) (d_{32}).	240
C.45 Podridão do Colmo por <i>Pythium</i> (d_{33}).	241

C.46 Podridão do Colmo por <i>Erwinia Carotovora</i> (d_{34}).	242
C.47 Antracnose do Colmo (d_{35}).	243
C.48 Podridão do Colmo por <i>Rhizoctonia</i> (d_{36}).	244
C.49 Podridão do Colmo por <i>Fusarium</i> (<i>Gibberella</i>) (d_{37}).	245
C.50 Podridão do Colmo por <i>Fusarium</i> (<i>Gibberella</i>) (d_{37}) - Variáveis de Decisão.	246
C.51 Podridão do Colmo por <i>Diplodia</i> (d_{38}).	247
C.52 Podridão do Colmo por <i>Macrophomina</i> (d_{39}).	248
C.53 Carvão do Pendão (d_{40}).	249
C.54 Míldio do Sorgo no Pendão (d_{41}).	250
C.55 Míldio do Sorgo no Pendão (d_{41}) - Variáveis de Decisão.	251

LISTA DE TABELAS

	<u>Pág.</u>
1.1 Exemplo de falhas em satélites.	32
2.1 Modelos para construção de sistemas automatizados.	43
2.2 Exemplo de matriz de retorno.	45
2.3 Exemplo de matriz de retorno com aplicação de fungicidas.	46
5.1 Associação causal nebulosa $\langle d_i, m_j \rangle$	100
5.2 Graus de temperatura e umidade.	102
6.1 Exemplos de custos associados aos exames.	117
6.2 Comparação das medidas f_{owa} e f_U para distribuições normalizadas.	123
6.3 Comparação das medidas f_{owa} e f_U para distribuições não normalizadas.	124
6.4 Cálculo da entropia f_{owa} para manifestações m_j	129
6.5 Cálculo da entropia f_U para manifestações m_j	130
7.1 Resultados do exemplo.	144
8.1 Etapas de desenvolvimento da cultura de milho.	147
8.2 Etapas de desenvolvimento da cultura de milho.	148
8.3 Níveis de severidade da cultura do milho.	148
8.4 Exemplo de doenças foliares de milho.	148
8.5 Principais características da Cercosporiose.	149

8.6	Ênupla correspondente a representação da desordem Cercospora (d_{10}) em $KB_d >$	149
8.7	Exemplo de testes laboratoriais.	150
8.8	Tupla correspondente a representação dos testes laboratoriais que identificam (d_{10}) em $KB_e >$	150
8.9	Exemplo de fungicidas.	151
8.10	Consequências da árvore de decisão da Figura 8.5.	152
8.11	Tupla correspondente a representação dos tratamentos para a desordem Cercospora (d_{10}) em $KB_t >$	153
8.12	Exemplo de itens críticos que podem causar falhas em satélite.	161
8.13	Exemplo de falhas em um missão de satélite.	161
8.14	Exemplo de métodos de detecção de falhas em satélite.	162
8.15	Exemplo de ações corretivas para controle de falhas em satélite.	162
8.16	Comparação de algoritmos na fase de diagnóstico.	165
8.17	Comparação de algoritmos na fase de diagnóstico.	165
8.18	Análise dos algoritmos na fase de investigação.	166
8.19	Análise dos algoritmos na fase de tratamentos.	167
A.1	Principais \top -normas e \top -conormas duais.	182
C.1	Etapas de desenvolvimento da cultura de milho.	192
C.2	Níveis de severidade da cultura do milho.	192
C.3	Graus de temperatura e umidade.	192
C.4	Tabela das principais desordens da cultura de milho.	193
C.5	Tabela das principais manifestações da cultura de milho.	194

C.6	Principais testes laboratoriais.	195
C.7	Fungicidas para milho fornecidos pelo Ministério da Agricultura.	195
C.8	Fungicidas sistêmicos X desordens.	196
C.9	Consequências da árvore de decisão de cada desordem d_i	197

CAPÍTULO 1

INTRODUÇÃO

O problema de diagnóstico consiste em encontrar as explicações para a presença de um conjunto de manifestações (sintomas, falhas) utilizando o conhecimento sobre o mundo. O raciocínio de diagnóstico pode ser visto como um processo cognitivo que abrange conhecimento genérico sobre falhas e explicações para estas falhas (um modelo de diagnóstico), o conhecimento sobre um domínio particular e heurísticas específicas do domínio.

Os problemas de diagnóstico podem ser encontrados em diversas áreas, como o diagnóstico dos sintomas de um paciente, de uma planta ou de um animal, a determinação de falha de um programa de um computador, a causa de ruídos em um sistema hidráulico, a localização de falhas num dispositivo eletrônico em um satélite, dentre outros. Os sistemas de diagnóstico também são usados em indústrias e tem sido uma área ativa de pesquisa em Inteligência Artificial (IA).

Esta área tem gerado muitas metodologias, teorias e aplicações nas duas últimas décadas. Sistemas de diagnósticos variam desde sistemas baseados em regras e raciocínio baseado em casos, a teorias baseadas em conjuntos e teorias baseadas em lógicas. Entretanto, foi somente em meados dos anos 80 que os aspectos conceituais e formais do processo de diagnóstico começaram a ser investigados por vários autores (Peng e Reggia, 1990).

Nos primeiros sistemas de diagnóstico, o conhecimento dos especialistas eram capturados na forma de regras de classificação empírica. Em seguida, cresceram os sistemas baseados em modelos, isto é, modela-se o comportamento normal do sistema e detectam-se as falhas a partir de comportamentos não esperados. Embora a introdução da abordagem baseada em modelos para construção de sistemas de diagnóstico tenha causado um impacto significativo no campo de diagnóstico, o entendimento fundamental da natureza de diagnóstico começou a ser esclarecido com as pesquisas sobre os aspectos formais do diagnóstico.

Uma das primeiras teorias formais de diagnóstico foi proposta por Peng e Reggia em termos de teoria de conjuntos, denominado *set-covering theory* ou *parsimonious covering theory*, traduzido para Teoria das Coberturas Parcimoniosas (TCP). Resumidamente, no modelo TCP, o conhecimento do domínio é representado por uma relação causal binária que é comparada com os fatos observados para se determinar um diagnóstico. Vários

algoritmos têm sido propostos nesta linha de trabalho embora este problema tenha sido geralmente classificado como NP-difícil (Peng e Reggia, 1990).

Uma modelagem formal de diagnóstico empregando conhecimento causal utiliza a lógica abdutiva como ferramenta básica (McIlraith, 1998). Na teoria lógica de diagnóstico abdutivo, o raciocínio de diagnóstico é formalizado dos efeitos para as causas, com conhecimento causal representado com implicações lógicas do tipo *causas* \rightarrow *efeitos* onde as causas são geralmente anormalidades ou falhas mas também podem incluir situações normais. Na maioria das vezes o raciocínio abdutivo não gera conclusões sólidas, pois somente encontra (e eventualmente classifica) as causas plausíveis para efeitos observados. No entanto, o raciocínio abdutivo pode ser muitas vezes empregado como paliativo para casos onde o raciocínio dedutivo não é possível.

A teoria lógica do diagnóstico abdutivo é mais expressiva que a da TCP, pois permite representar vários tipos de interações que não são possíveis na TCP original. Vários autores têm investigado tanto a teoria quanto a implementação da teoria abdutiva (Peng e Reggia, 1990; McIlraith, 1998; Lucas, 1996).

Informalmente, a abdução é um forma de raciocínio hipotético e ajuda na caracterização de várias tarefas de raciocínio humano. Embora a abdução tenha sido mais empregada na área de diagnóstico por pesquisadores de IA, este mecanismo de inferência também tem sido aplicado em outros domínios, tais como, processamento de imagens, reconhecimento de padrões, entendimento da linguagem natural e atualização de base de dados.

Por outro lado, tomadas de decisão pelos especialistas é outro ponto importante na área de diagnóstico e tratamentos. Por exemplo, um médico tem que decidir que exames laboratoriais deve pedir ao paciente visando confirmar um diagnóstico. Após o diagnóstico, o médico tem que decidir dentre os possíveis tratamentos recomendados. Um agrônomo, por sua vez, tem que decidir dentre os possíveis tratamentos, aqueles que podem controlar melhor uma determinada desordem em uma cultura. Para tal, ele tem que considerar algumas variáveis, como a época do aparecimento dos sintomas, sua severidade e as condições ambientais (temperatura e umidade). No primeiro caso a tomada de decisão é utilizada para auxiliar na geração das hipóteses mais plausíveis. No segundo caso, a tomada de decisão é utilizada para auxiliar na seleção dos tratamentos mais plausíveis para controlar a desordem.

Embora tarefas para diagnóstico e tratamentos estejam fortemente relacionadas, elas não são frequentemente investigadas juntas. Mecanismos de tomada de decisão usados para tratamentos tem sido estudados mais detalhadamente na área de Teoria de Decisão

(matrizes de "payoffs", árvores de decisão, diagramas de influência entre outros) enquanto mecanismos de inferência usados em diagnóstico têm sido mais pesquisados na área de IA.

O objetivo deste trabalho é apresentar uma abordagem integrada para diagnóstico, investigação e tratamentos permitindo organizar os vários tipos de informação gerados nestas etapas. Esta nova abordagem, denominada Teoria das Coberturas Nebulosas (TCN), contempla a descrição de um novo modelo para solução de problemas de diagnóstico, investigação e tratamento que integra conceitos de inferência abductiva, lógica nebulosa e tomada de decisão.

1.1 Motivação do Trabalho

A área de Diagnóstico, explorada no domínio da Medicina, tem permitido grandes avanços para a Inteligência Artificial através da geração de muitas metodologias, teorias e aplicações, nas duas últimas décadas dentre as quais se destaca a de sistemas especialistas (SE). Entretanto, as tecnologias disponíveis na área de IA para formalização e representação do conhecimento dos especialistas e, conseqüentemente, implementação através de um SE não são triviais e, muitas vezes, também não são aderentes à outros domínios, tais como, diagnóstico de plantas, diagnóstico de falhas de computador, diagnóstico de falhas de satélite entre outros.

Por exemplo, é notável a crescente demanda por tecnologias de IA na agricultura, mais especificamente de SEs, na área de diagnóstico de doenças de plantas e animais, dada a escassez de informações na literatura especializada sobre a identificação de doenças e métodos de controle (Fernandes e Oliveira, 1997). Devido à falta de informação, muitas vezes produtores e técnicos agrícolas se utilizam de um arsenal de produtos químicos para combater uma doença, com resultados duvidosos, além de colocar em risco a saúde dos próprios aplicadores e causar danos ao ecossistema.

A utilização de SE em geral na agricultura já vem sendo explorada em vários trabalhos. Dentre os tipos de problemas que são resolvidos com SE, encontra-se em primeiro lugar a área de Diagnósticos. Entretanto, a maioria dos sistemas desenvolvidos se baseiam no paradigma de regras de produção do tipo *efeito* → *causa*, onde o mecanismo de inferência utilizado é o modelo dedutivo, que torna tais sistemas difíceis de desenvolver e muito pouco reutilizáveis. Tais problemas puderam ser observados durante o desenvolvimento de um SE para diagnóstico de doenças do milho (disponível em <http://diagnose.cnptia.embrapa.br>) que foi desenvolvido na Embrapa Informática Agropecuária em parceria com a Embrapa Milho e Sorgo (Massruhá et al., 1999).

1.1.1 Diagnóstico de Doenças de Milho

O sistema para diagnóstico de doenças do Milho de Massruhá et al. (1999) é baseado em regras do tipo "se *efeito* então *causa*", possui um mecanismo de inferência dedutivo e sua base teórica é a lógica de primeira ordem. A base de conhecimento deste sistema é constituída de regras de produção geradas a partir de uma árvore de decisão, gerada a partir de entrevistas com especialistas e da literatura na área. Os especialistas descrevem o conhecimento do domínio em direção oposta das regras de produção, isto é, "se uma doença está presente então algumas manifestações ocorrem".

Para fins ilustrativos, abaixo está um parágrafo correspondente à descrição de uma doença de milho denominada Ferrugem Comum (d_3) retirado de Fernandes e Oliveira (1997).

A Ferrugem Comum, sob condições favoráveis, pode causar seca prematura das plantas, comprometendo a produção. Essa ferrugem pode ser encontrada em qualquer fase do desenvolvimento das plantas de milho, inicialmente nas folhas baixas. Ferrugem Comum forma pústulas predominantemente alongadas, de cor marrom-clara em plantas jovens que tornam-se marrom-escuro à medida em que a planta se aproxima da fase de maturação. Essas pústulas podem ser encontradas em ambas as faces das folhas, e devido ao rompimento da epiderme, apresentam caracteristicamente uma fenda. É uma doença favorecida por temperaturas entre 16 e 23° C e umidade relativa alta.

O sistema de diagnóstico de doenças do milho de Massruhá et al. (1999) foi baseado em regras de produção que geralmente têm a forma básica:

IF <manifestações> THEN <causas>

Um exemplo das regras de produção utilizadas na implementação do sistema é:

B1: IF local='folhas' AND aparência = 'Lesão' AND forma='individualizadas' AND tipo='pústula' AND cor='marrom' AND formato = 'alongada' THEN desordem='ferrugem_comum';

É interessante observar que na implementação destas regras de produção, não foram consideradas características temporais bem como condições favoráveis para o desenvolvimento da doença, tais como temperatura e umidade. Incertezas das manifestações causadas por determinadas doenças e dos fatos observados também não são contemplados neste sistema, o que o torna bastante limitado.

A versão atual do sistema para diagnóstico de doenças de milho (disponível em <http://diagnose.cnptia.embrapa.br>) foi desenvolvida com a ferramenta ServCLIPS que possibilita a construção e execução de sistemas especialistas via Internet. Maiores detalhes podem ser encontrados em Cruz et al. (2001). A base de conhecimento de doenças de milho deste SE está sendo utilizada como estudo de caso no presente trabalho.

1.1.2 Diagnóstico de Falhas em Satélite

Uma outra área de aplicação é o diagnóstico de falhas de satélites, que é um ponto crucial na área espacial. Os satélites em órbita se comunicam com seus centros de controle através de telemetria (satélite para terra) e telecomandos (terra para satélite). Parte da telemetria é dedicada ao monitoramento do satélite. Quando uma falha é detectada pelos serviços de monitoramento em tempo-real dos centros de controle, é necessário localizar o mais rápido possível a origem da falha de modo a tomar as ações corretivas para garantir a continuidade da missão (A822007-ANL-61/D1, 2001).

Quanto ao raciocínio de diagnóstico de falhas de satélite, o mecanismo de inferência abduutivo pode ser utilizado para encontrar as explicações possíveis para os fatos observados (manifestações presentes e ausentes). Na maioria das vezes, estas informações são incompletas o que leva os usuários expressarem incertezas sobre os fatos observados.

A preocupação por parte de engenheiros e técnicos é tão grande que existe uma metodologia detalhada de análise de risco e gerenciamento para a plataforma multi-missão (MMP) do INPE (A822007-ANL-61/D1, 2001). Esta análise de risco do sistema é desenvolvida na fase de especificação do projeto.

Para ilustrar um problema de diagnóstico aplicado neste domínio é apresentado um pequeno exemplo. Baseado na análise de risco existe uma lista de itens críticos:

- Tanque de hidrazina (d_1);
- Propulsores do satélite - Thrusters (d_2);
- Válvulas de enchimento e drenagem (d_3);
- Filtros (d_4);
- Rodas de reação (d_5);
- Partes elétricas dos painéis solares - EPSA (d_6);
- Mecanismo de retenção, liberação e abertura (d_7);

- Bateria (d_8);
- Unidade de condicionamento e distribuição de potência - PCDU (d_9).

TABELA 1.1 – Exemplo de falhas em satélites.

Cenários de risco	Equipamentos
Perda da função de propulsão (m_1)	Tanque de Hidrazina e Válvulas (d_1, d_3)
Perda da velocidade mínima (m_2)	Propulsores e filtros (d_2, d_4)
Deterioração da célula solar devido ao ambiente espacial (m_3)	EPSA (d_6)
Redução da capacidade de armazenamento de energia (m_4)	Bateria (d_8)
Perda da distribuição e condicionamento de energia (m_5)	PCDU (d_9)

Baseado na análise de risco, é recomendado manter o gerenciamento de riscos como descrito na tabela 1.1. Por exemplo, uma falha m_1 no equipamento em que há perda da função de propulsão pode ser causada pelo tanque de hidrazina (d_1) ou pelas válvulas de enchimento e drenagem (d_3). Já para a perda de velocidade mínima (m_2), existem duas possíveis causas: d_2 e d_4 .

Estas falhas são detectadas através de telemetria e para cada uma delas existe um formulário onde é preenchido a falha, suas prováveis causas e as ações corretivas que devem ser tomadas para cada falha através de telecomandos. Entretanto, este modelo de falhas também pode contemplar incertezas na associação entre os equipamentos e as falhas. Estas incertezas também devem ser representadas para uma melhor solução de diagnóstico. Um exemplo do problema diagnóstico de falhas em satélite será apresentado no capítulo 8.

1.2 Escopo do Trabalho

A motivação inicial deste trabalho estava na área de diagnóstico. Entretanto, no decorrer da pesquisa observou-se que embora as tarefas para diagnóstico e tratamentos estejam fortemente relacionadas, elas não são frequentemente investigadas juntas. Consequentemente, aspectos relevantes para diagnóstico e tratamentos não são considerados. Adicionalmente, a tarefa de investigação (ex., obtenção de exames complementares) também deveria ser tratada de maneira integrada às tarefas de diagnóstico e tratamentos, pois quando a investigação é bem feita, obtém-se maior eficácia e eficiência no diagnóstico, além de redução de custos e riscos.

No presente trabalho é apresentado uma abordagem integrada para diagnóstico, investigação e tratamentos permitindo organizar os vários tipos de informação gerados nestas etapas durante o processo de raciocínio clínico. Esta nova abordagem, denominada

Teoria das Coberturas Nebulosas (TCN), contempla a descrição de um novo modelo para solução de problemas de diagnóstico, investigação e tratamento que integra conceitos de inferência abdutiva, lógica nebulosa e tomada de decisão.

Inicialmente, antes de apresentar esta nova abordagem, que é o cerne deste trabalho, é apresentado um breve resumo dos modelos computacionais mais utilizados para suportar o processo de raciocínio clínico e tomada de decisão, bem como a modelagem de incertezas, durante o processo de decisão diagnóstica e terapêutica sob o ponto de vista de Inteligência Artificial (IA) e Teoria de Decisão (TD). Aspectos conceituais do mecanismo de inferência abductivo que será utilizado como fundamento teórico para o modelo proposto bem como conceitos básicos da TCP e suas extensões também são apresentados. Na sequência, é apresentada a Teoria das Coberturas Nebulosas propriamente dita bem como a abordagem adotada para diagnóstico, investigação e tratamentos que incorpora conceitos de inferência abdutiva, lógica nebulosa e teoria de decisão. Para tanto, primeiramente são apresentadas as etapas do processo de raciocínio clínico bem como as tarefas de tomada de decisão inerentes a cada etapa.

Neste trabalho também é apresentada a experiência da aplicação da TCN a um problema real de diagnóstico e controle de doenças de milho. Finalmente, são apresentadas as principais contribuições do modelo proposto e suas limitações bem como possíveis trabalhos futuros a partir dos resultados obtidos. Uma síntese dos capítulos da tese estão descritos a seguir.

1.3 Esboço Geral

Este trabalho foi dividido em mais oito capítulos onde:

- *CAPÍTULO 2 - Revisão de Literatura:* Este capítulo contempla a descrição de alguns mecanismos de inferência, modelos de incerteza e tomada de decisão para suportar decisões de natureza diagnóstica e terapêutica.
- *CAPÍTULO 3 - Modelos abductivos:* Conceitos básicos e formais de inferência abductiva são apresentados neste capítulo.
- *CAPÍTULO 4 - Formalização do modelo proposto:* Neste capítulo são apresentados os limites do problema a ser abordado bem como o modelo formal para suportar o ambiente abductivo para raciocínio clínico que engloba diagnóstico, investigação e tratamentos.

- *CAPÍTULO 5 - Ambiente abdutivo para diagnóstico:* Neste capítulo é descrito detalhadamente o modelo formal para suportar o ambiente abdutivo para diagnóstico.
- *CAPÍTULO 6 - Ambiente abdutivo para investigação:* Neste capítulo é descrito detalhadamente o modelo formal para suportar o ambiente abdutivo para investigação.
- *CAPÍTULO 7 - Ambiente abdutivo para tratamentos:* Neste capítulo é descrito detalhadamente o modelo formal para suportar o ambiente abdutivo para tratamentos.
- *CAPÍTULO 8 - Estudo de caso:* Neste capítulo é apresentado um exemplo completo no domínio de diagnóstico de doenças de milho utilizando a abordagem proposta e um pequeno exemplo no diagnóstico de falhas de satélite.
- *CAPÍTULO 9 - Conclusões:* Contribuições, limitações e perspectivas de trabalhos futuros.

Adicionalmente aos 9 capítulos acima, no apêndice [A](#) são apresentados alguns conceitos básicos da teoria de conjuntos nebulosos e teoria de possibilidades que serão utilizados no decorrer do trabalho. No apêndice [B](#) está descrita a implementação do modelo proposto onde são apresentados os algoritmos desenvolvidos para suportar o ambiente genérico para raciocínio clínico. No apêndice [C](#) é apresentada a especificação completa da base de conhecimento para diagnóstico, investigação e tratamentos de doenças de milho.

CAPÍTULO 2

MODELOS COMPUTACIONAIS PARA SOLUÇÃO DE PROBLEMAS DE DIAGNÓSTICO E TRATAMENTOS

Neste capítulo é apresentado um breve resumo dos modelos computacionais mais utilizados para suportar o processo de raciocínio e tomada de decisão, bem como a modelagem de incertezas, durante o processo de decisão diagnóstica e terapêutica sob o ponto de vista de Inteligência Artificial (IA) e Teoria de Decisão (TD). As tecnologias oferecidas por ambas as áreas visam suportar o raciocínio humano e a tomada de decisão pela formalização e automatização do conhecimento do especialista do domínio. Cada vez mais, nota-se a integração das tecnologias desenvolvidas em cada uma destas áreas embora sejam baseadas em princípios diferentes.

Primeiramente, na próxima seção são descritos alguns conceitos básicos de sistemas baseados em conhecimento onde se enquadra o problema de diagnóstico e tratamento. Na seção (2.2) estão descritos sucintamente os principais métodos de IA utilizados em sistemas de diagnósticos: sistemas de classificação estatística, sistemas baseados em regras e sistemas abduativos baseados em associações. Na seção (2.3) estão descritos os principais métodos de TD utilizados para expressar conhecimento qualitativo na área de tomada de decisão terapêutica: árvores de decisão e diagramas de influência.

2.1 Sistemas para Solução de Problemas Baseados em Conhecimento

O projeto e o uso de sistemas (especialistas) baseados em conhecimento tem sido uma área importante de IA. Segundo Peng e Reggia (1990), a característica principal destes sistemas, além de seu alto nível de desempenho, é a separação do conhecimento do mecanismo de inferência. O conhecimento do programa abstraído do especialista sobre o domínio é armazenado na base de conhecimento. O mecanismo de inferência utiliza a informação da base de conhecimento para gerar conclusões sobre os dados de entrada.

Existem vários problemas que devem ser resolvidos no desenvolvimento de um sistema baseado no conhecimento. Estes incluem problemas de representação de conhecimento (que tipo de representação é o melhor para armazenar conhecimento na base de conhecimento), aquisição de conhecimento (como capturar o conhecimento dos especialistas do domínio), modelos de raciocínio (que tipo de mecanismo de inferência usar) e justificativa de respostas (como mostrar as respostas de uma maneira amigável aos usuários). Soluções para estes problemas são interdependentes e têm variado de aplicação para aplicação (Peng e Reggia, 1990).

Geralmente, as aplicações na área de diagnóstico englobam 3 tarefas principais:

- Tarefa de Diagnóstico: Como pode ser inferido as causas mais prováveis de problemas observados (doenças ou falhas em máquinas) a partir de um conjunto de evidências (sintomas, características do paciente, condições operacionais ou resultados de teste)?
- Tarefa de Aquisição de Informação: Que tipo de informações ou testes adicionais devem ser requeridos?
- Tarefa de Tomada de Decisão de Tratamentos: O que pode ser feito para controlar ou tratar o problema?

Mecanismos de inferência e tomada de decisão, bem como mecanismos para expressar a incerteza inerente a estas tarefas, têm sido amplamente investigados na área de Inteligência Artificial e Teoria de Decisão. Nas próximas seções estão descritos os principais modelos computacionais em IA e em TD utilizados para suportar as tarefas acima.

2.1.1 Representação do Conhecimento

Segundo Peng e Reggia (1990), existem pelo menos dois tipos de conhecimento que são usados pelos diagnosticadores humanos e pelos sistemas de diagnóstico mais avançados de IA: conhecimento estrutural e probabilístico.

O conhecimento estrutural (usualmente na forma simbólica) nos sistemas de diagnóstico baseados em conhecimento especifica quais são as entidades do domínio da aplicação, quais entidades são associadas entre si e quais são os tipos de associação entre elas. Dentre os diferentes tipos de associação, relações de causa-efeito são provavelmente o mais importante em sistemas para solução de problemas de diagnóstico. Por exemplo, em diagnóstico médico, a frase "a doença d_i pode ser a causa da manifestação m_j " ou " m_j pode ser uma manifestação de d_i " associa d_i com m_j . Um outro exemplo de associação em medicina é a frase "o tratamento t_k é o mais apropriado para o paciente X com uma desordem d_i " ou "o exame E_x deve ser solicitado para confirmar d_i ". Estas associações são obtidas da compilação do conhecimento em profundidade (p.ex., em medicina, por certas razões fisiológicas, d_i e m_j , t_k e d_i estão associadas) ou da extração de regularidades da experiência passada (p.ex., d_i e m_j são casualmente associadas porque o sintoma m_j tem sido encontrado em certos casos de pacientes com a doença d_i ou t_k e d_i estão associadas porque o tratamento t_k tem mostrado eficiente em pacientes com a doença d_i).

O conhecimento chamado de probabilístico pelos autores, por outro lado, reflete a incerteza em tais associações. No entanto, este conceito pode ser visto de uma maneira mais geral, como a modelagem de incerteza que pode ser feita pela utilização da teoria de probabilidade, da teoria de possibilidades (Dubois e Prade, 1988) ou teoria de Dempster-Shafer e fatores de certeza. O conhecimento de incerteza (usualmente, em forma numérica em sistemas baseados em conhecimento) especifica a força das associações e quantifica as evidências individuais dos casos. A incerteza pode ser representada diretamente através de números de dentro da modelagem de incerteza ou na forma simbólica não-numérica ("muito comum", "moderado", "muito raro", etc.), termos estes que podem ser tratados como tal, no que se chama atualmente de "computing with words" ou traduzidos em números.

A maioria dos sistemas de diagnóstico que utilizam técnicas de IA tem utilizado medidas probabilísticas ou outros esquemas de peso para um bom desempenho. Entretanto, como isto é realizado de uma maneira informal e heurística, muitas vezes não pode ser justificado rigorosamente ou formalmente em situações gerais.

Um modelo consistente e unificado que incorpora ambos os tipos de conhecimento (estrutural e probabilístico) é muito desejável no desenvolvimento de lógica abdutiva formal em geral, e nos sistemas de diagnóstico baseados em conhecimento em particular.

Uma grande variedade de sistemas de diagnóstico automatizados tem sido desenvolvido durante as últimas décadas. Embora muitos destes sistemas tenham alcançado alto nível de desempenho em várias aplicações, eles têm uma série de limitações.

2.1.2 Mecanismos de Inferência

Além da representação do conhecimento, o mecanismo de inferência utilizado pelo especialista para gerar conclusões a partir da base de conhecimento também tem sido muito investigado nas últimas décadas. Os 3 tipos de mecanismos de inferência mais utilizados em sistemas baseados em conhecimento são:

- Dedutivos;
- Indutivos;
- Abdutivos.

Simplificadamente, dada a premissa $a \rightarrow b$, a premissa menor a e a conclusão b tem-se que na:

- Dedução raciocina-se a partir de $a \rightarrow b$ e a produz a conclusão b ;

- Indução raciocina-se a partir de a e b para produzir a regra plausível $a \rightarrow b$;
- Abdução infere-se a explicação a , a partir da premissa $a \rightarrow b$ e b .

A partir desta descrição simplista, abdução aparece como uma regra de inferência *modus ponens reversa*. Peirce (McClelland, 1998) viu isto como a seleção de uma melhor explicação para ocorrência de b que deveria ser confirmada pelo processo científico. Em IA, abdução é definida como inferência para a melhor explicação, sem confirmação subsequente.

Os mecanismos de inferência dedutivos a partir do conhecimento representado através de regras de produção, são os mais utilizados no desenvolvimento de sistemas na área de IA. Recentemente, os mecanismos de inferência abdutiva têm sido amplamente incorporados nos sistemas. Antes de apresentar os modelos abdutivos baseados em associações, serão introduzidos alguns conceitos básicos de modelos estatísticos de classificação, cuja inferência é realizada através do cálculo das probabilidades posteriores e sistemas de produção baseados em regras.

2.2 Modelos Computacionais na Área de Diagnóstico

2.2.1 Sistemas Estatísticos de Classificação: Redes Bayesianas

A classificação Bayesiana é o modelo estatístico mais utilizado na solução de problemas de diagnóstico desde o advento de computadores digitais, às vezes atingindo excelentes resultados (Peng e Reggia, 1990). O conhecimento dos sistemas de diagnósticos que utilizam classificação Bayesiana é representado por probabilidades anteriores de desordens e probabilidades condicionais entre desordens e manifestações. Por exemplo, na medicina há situações em que cada uma das n desordens podem individualmente ser a causa dos sintomas de um paciente.

Suponhamos um conjunto H de n hipóteses possíveis, onde $H = \{h_1, h_2, \dots, h_n\}$ e um conjunto M de m manifestações onde $M = \{m_1, m_2, \dots, m_m\}$. Em uma aplicação médica, as hipóteses são as desordens possíveis e as manifestações incluem história do paciente, sinais físicos, sintomas e resultados de exames laboratoriais. Para simplificação do problema é assumido que todas as variáveis (hipóteses e manifestações) assumem somente dois valores (verdadeiro ou falso). Um diagnóstico ou uma explicação D é um subconjunto de H . Assim, D é um conjunto de hipóteses que devem estar presentes implicando que todas as outras desordens de H estão ausentes. $P(D, s)$ representa a probabilidade anterior de D a partir do conhecimento do especialista s . Seja E o conjunto de evidências contemplando todas as manifestações presentes, ausentes ou não observadas

de um paciente em particular, e seja $P(E|D, s)$ a probabilidade condicional de E na presença de D . O teorema de Bayes pode ser aplicado para calcular a probabilidade posterior de cada explicação D a partir uma evidência E observada:

$$P(D|E, s) = \frac{P(E|D, s)P(D, s)}{\sum_{\forall D \subseteq H} P(E|D, s)P(D, s)} \quad (2.1)$$

Esta fórmula geral é muito complexa para calcular, pois, um paciente pode ter mais de uma desordem das n possíveis. Assim, o número de explicações possíveis seria 2^n e o número de parâmetros independentes necessários para especificar uma distribuição anterior completa seria $2^n - 1$. Para m pedaços de evidência, a distribuição condicional geral tem $2^m - 1$ para cada hipótese. Portanto, este problema requer uma especificação de $2^n(2^m - 1)$ parâmetros independentes para todas diagnoses. O que torna esta abordagem impraticável para problemas com muitas hipóteses e manifestações.

Duas simplificações são assumidas para aplicação desta abordagem bayesiana: que as desordens são mutuamente exclusivas e exaustivas. Por exemplo, considerá-las mutuamente exclusivas garante que um paciente não pode ter mais que uma desordem e considerá-las exaustivas garante a independência condicional de cada manifestação (i.e, $P(m_i|s) = P(m_i|m_j, s)$).

Em uma rede bayesiana onde são assumida estas duas simplificações, a distribuição da probabilidade condicional da evidência E para cada desordem h_i pode ser decomposta no produto das probabilidades condicionais das manifestações como representado a seguir:

$$P(E|h_i, s) = \prod_{\forall m_i \in E} P(m_i|h_i, s) \quad (2.2)$$

Então, considerando-se que as desordens são mutuamente exclusivas e exaustivas, somente $m \times n$ probabilidades condicionais e $n - 1$ probabilidades anteriores são necessárias.

Desta forma, o teorema de Bayes pode ser diretamente aplicado para obter $P(d_i/M_j)$, isto é, a probabilidade posterior de cada d_i para um dado paciente, através da fórmula 2.3

$$P(d_i/M_j) = \frac{P(M_j/d_i)P(d_i)}{P(M_j)} = \frac{P(M_j/d_i)P(d_i)}{\sum_{k=1}^n P(M_j/d_k)P(d_k)} \quad (2.3)$$

A probabilidade posterior $P(d_i/M_j)$ fornece a probabilidade de cada doença d_i ocorrer na presença das manifestações M_j . A doença com maior probabilidade posterior calculada pode ser escolhida como diagnose.

Esta abordagem atingiu alta performance em domínios específicos como na diagnose da causa da dor abdominal aguda, da doença cardíaca congênita e hemorragia cerebral (Henrion e Horvitz, 1992). Embora um sistema de classificação Bayesiana seja capaz de calcular a probabilidade posterior de um dado conjunto de hipóteses múltiplas, este método tornou-se falho devido algumas limitações dos sistemas tais como: a própria restrição de exclusividade e independência condicional e, principalmente, devido a abordagem quantitativa, formal e rigorosa da inferência probabilística que é muito distante da abordagem qualitativa e informal do especialista humano.

2.2.2 Sistemas de Produção Baseados em Regras

O golpe decisivo nos modelos estatísticos de classificação foi o aparecimento dos sistemas de produção baseado em regras (SPBR) como uma abordagem alternativa derivada de IA que utiliza regras de produção e lógica para representar e inferir o conhecimento. O mecanismo de inferência mais comum em SPBR é o dedutivo.

Os sistemas de diagnóstico baseado em regras produziram importantes resultados nas últimas décadas. O MYCIN, para doenças infecciosas é o mais conhecido (Shortliffe, 1976). Neste modelo o conhecimento é representado por um conjunto de regras ou produções. Associado a cada regra pode haver estimativas de certeza que são combinadas para produzir uma estimativa final das conclusões destes sistemas.

Na Figura 2.1, está apresentada uma rede de inferência para um sistema baseado em regras. A direção das setas indicam a direção das regras de diagnóstico, das observações para desordens e para a decisão de tratamento (Henrion e Horvitz, 1992).

O primeiro problema de sistemas baseados em regras é o de representar o conhecimento através de regras, já que os especialistas usualmente descrevem o conhecimento do domínio em direção oposta das regras de produção.

Por exemplo, seja d_1 a causa das manifestações m_1, m_2 e m_3 e d_2 a causa de m_1, m_2, m_3 e m_4 . Para modelar tais situações, um sistema baseado em regras teria usualmente duas regras:

R1: IF m_1, m_2, m_3 THEN d_1

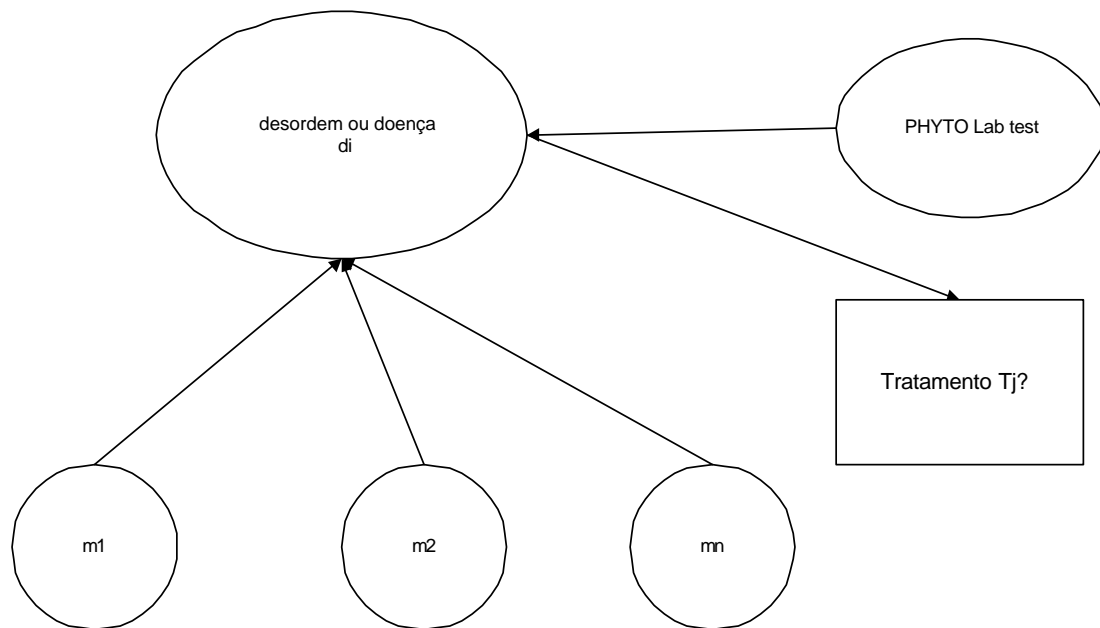


FIGURA 2.1 – Uma rede de inferência para um sistema baseado em regras.

$R2$: IF m_1, m_2, m_3, m_4 THEN d_2

Suponhamos o caso particular em que as manifestações m_1 , m_2 e m_3 estão presentes. O diagnosticador deveria considerar d_1 e d_2 como hipóteses. Entretanto, em um sistema baseado em regras que contemple $R1$ e $R2$, somente d_1 seria considerado como hipótese. Dado que m_4 não está presente a regra $R2$ não seria disparada.

Ou seja, em SPBR usualmente usa-se manifestações nas premissas das regras e causas nas conclusões, ao contrário da maneira com que ocorre na natureza.

Além disso, os especialistas nem sempre conseguem ser precisos e nem têm certeza absoluta sobre a informação que fornecem. Devido a essas limitações, muitas vezes o mecanismo de inferência abdutivo parece mais adequado ao raciocínio de diagnóstico.

2.2.3 Sistemas Abdutivos Baseado em Associações

2.2.3.1 Representação do Conhecimento

As redes associativas, ou semânticas, tem sido estudadas extensamente como método de representação de conhecimento em IA. Uma rede associativa consiste usualmente de nós, representando entidades como objetos, conceitos e eventos, e arcos representando inter-relações associações entre nós. O significativo poder das redes associativas as

torna particularmente adequadas para representação das associações causais em diversas aplicações de raciocínio abdutivo.

2.2.3.2 Inferência Abdutiva

Conforme apresentado em Peng e Reggia (1990), o uso das associações simbólicas causa-efeito entre entidades individuais e um processo repetitivo de formação de hipótese e teste como mecanismo básico de inferência, são as características dos modelos abduativos baseados em associações para a solução do problema de diagnóstico.

Alguns sistemas de diagnóstico auxiliados por um computador podem ser classificados nesta categoria, como o INTERNIST-1 para medicina interna, o NEUROLOGIST para neurologia, o PIP para edema, IDT para diagnose de problemas de hardware de computador, o KMS.HT para sistemas independentes de domínio, e outros (Peng e Reggia, 1990).

Dado o conjunto de uma ou mais características iniciais do problema, o mecanismo de inferência gera um conjunto de hipóteses plausíveis potenciais ou "causas" que possam explicar as características do problema dado. Estas hipóteses são testadas então de duas maneiras.

A primeira utiliza diversos procedimentos que medem suas habilidades em explicar as características presentes, enquanto a segunda gera novas questões cujas respostas auxiliarão na discriminação das hipóteses mais prováveis. Este ciclo de formação de hipótese e teste é então repetido com a informação adicional adquirida. As hipóteses anteriores podem ser atualizadas formando novas hipóteses. Uma vez que este mecanismo de inferência é um modelo direto do processo de raciocínio de diagnóstico humano, o processamento feito por estes sistemas é relativamente de fácil entendimento para os usuários.

Os modelos abduativos baseados em associações, usando conhecimento associativo causa-efeito são bastante promissores. No entanto, esta técnica está ainda em um estágio inicial de desenvolvimento em relação aos modelos estatísticos de classificação e aos sistemas baseados em regras. Diversos dos modelos abduativos desenvolvidos são dependentes do domínio, como INTERNIST-1, PIP e IDT.

A representação do conhecimento nestes sistemas está relacionada a uma aplicação específica, e o mecanismo de inferência conta intensamente com heurísticas extraídas das características destes domínios específicos. Estes modelos de domínios específicos obtiveram sucesso em suas respectivas aplicações, auxiliaram no esclarecimento das idéias

básicas da natureza abduativa da inferência diagnóstica e, conseqüentemente, ocasionaram a diminuição do interesse pelos modelos tradicionais de classificação estatística e baseados em regras.

Por outro lado, a natureza do domínio específico destes modelos limitou suas generalizações e aplicações a outros problemas. Os modelos mais recentes desta categoria incluem aplicações independentes do domínio, como no sistema KMS.HT (Peng e Reggia, 1990). Claramente, separando os aspectos independentes de domínio da representação do conhecimento e o processo de inferência da informação específica da aplicação, estes sistemas usufruem de grande generalidade, mas continuam estacionários na formulação teórica do raciocínio abduativo.

Resumidamente, enquanto os modelos estatísticos de classificação e os sistemas baseados em regras possuem teorias bem definidas, isto é, teoria da probabilidade e lógica dedutiva, respectivamente, os modelos abduativos baseado em associações, no momento, necessitam de uma base teórica para suportá-los.

A Teoria das Coberturas Parcimoniosas (TCP) é uma tentativa de formalizar o raciocínio de diagnóstico para os modelos abduativos baseados em associações. Em outras palavras, a TCP é uma tentativa de remover pelo menos em parte a interrogação da tabela 2.1 conforme descrito em Peng e Reggia (1990).

TABELA 2.1 – Modelos para construção de sistemas automatizados.

Modelo	Representação	Mecanismo de Inferência	Base Teórica
Modelo estatístico de classificação	Probabilidades anteriores e condicionais	Cálculo das probabilidades posteriores	Teoria da probabilidade
Dedução baseada em regras	Regras condicionais (produção)	Deduções (com melhoramentos)	Lógica de primeira ordem
Abdução baseado em associações	Frames e redes semânticas	Formação de hipóteses e testes	?

2.3 Modelos Computacionais na Área de Tratamentos

Decisões devem ser tomadas por indivíduos e organizações constantemente e, apesar de muitas vezes serem tomadas sob condições adversas (pouca informação, incertezas, escassez de tempo, etc .), elas podem ter conseqüências cruciais. Por exemplo, na área de diagnóstico médico, uma decisão pode comprometer a vida de um paciente. Já no domínio agrícola, uma tomada de decisão pode vir a comprometer toda uma produção ou um replanejamento agrícola de uma maneira geral. A utilização de predições durante um processo de diagnóstico é de vital importância pois viabiliza a investigação de manifestações prioritárias mesmo que não se tenha informação completa sobre elas de

modo a selecionar e reduzir hipóteses. Esta funcionalidade pode ser incorporada no processo de diagnóstico através de mecanismos de teoria de decisão.

A teoria de decisão tem um papel importante em várias áreas tais como estatística, economia, teoria de jogos, engenharia, medicina entre outros. A motivação dos trabalhos nesta área é baseada em duas aplicações principais: modelagem de agentes e tomada de decisão. Recentemente, a área de teoria de decisão tem se tornado de grande interesse na comunidade de Inteligência Artificial (IA), principalmente na área de modelagem de agentes inteligentes. Na área de tomada de decisões, a ênfase principal em IA está nos princípios básicos dos procedimentos para tomada de decisão.

A área de planejamento e controle é outra área de aplicação de teoria de decisão. Vários autores têm proposto incorporar teoria de decisão dentro do paradigma de planejamento utilizando modelos clássicos de teoria de decisão, tais como, Processos de Decisão de Markov (MDP).

Uma decisão importante durante um processo de raciocínio clínico é o conjunto de medidas de controle que devem ser adotadas imediatamente após o diagnóstico. Na medicina, estas medidas referem-se ao tratamento do paciente propriamente dito. No domínio agrícola são adotadas medidas como: eliminação das plantas hospedeiras infectadas, aplicação de fungicidas entre outras. Neste domínio, as condições ambientais de temperatura (alta, média, baixa) e umidade relativa do ar (alta, baixa) interferem diretamente na tomada de decisões para controle de doenças. A tomada de decisão sob incertezas é um problema frequentemente encontrado em aplicações reais. Vários critérios de decisão para tomada de decisão considerando incertezas tem sido propostos.

Nesta seção é apresentado um breve resumo dos principais mecanismos de tomada de decisão que são utilizados em processos de diagnóstico e tratamentos durante um problema de raciocínio clínico.

A teoria de decisão é baseada nos axiomas de teoria de probabilidade e utilidade (Savage, 1972). A teoria de probabilidade fornece a infraestrutura para representação de incertezas e a teoria de utilidade adiciona um conjunto de princípios para consistência entre crenças, preferências e decisões. Dentre os métodos mais tradicionais de representação de problemas em teoria de decisão tem-se matrizes de *payoff* (matrizes de retorno), árvores de decisão e diagramas de influência.

2.3.1 Matrizes de Retorno

A representação das matrizes de retorno teve sua origem no trabalho de von Neumann e Morgenstern (1953) em jogos e tornou-se popular por Savage (1972). Esta representação é adequada para problemas onde uma decisão deve ser tomada e existe uma incerteza comum em todas as ações. Um dos critérios de decisão mais utilizados é o *maximin* que define que para escolher uma alternativa de ação deveria ser considerado o maior *payoff* mínimo. Por exemplo, suponhamos uma decisão entre duas ações A_1 e A_2 sobre três estados (s_1, s_2, s_3) com os respectivos *payoffs*, conforme apresentado na tabela 2.2.

TABELA 2.2 – Exemplo de matriz de retorno.

Ações/Estados	s_1	s_2	s_3
A_1	99	30	3
A_2	90	60	0

Neste caso, a alternativa A_1 deveria ser selecionada em relação a A_2 , pois seu retorno mínimo (3) é maior que o de A_2 (0).

Outro critério de decisão é *maximax*, onde a escolha da ação deve se basear no retorno máximo. Suponhamos uma situação onde um produtor agrícola diagnosticou uma doença denominada *Ferrugem Comum* (d_3) em sua plantação de milho e precisa decidir sobre a aplicação de fungicida como uma medida de controle (A_1) ou não (A_2). A aplicação do fungicida pode inibir a evolução da doença causando impacto na produtividade.

Entretanto, vários fatores podem interferir na decisão de aplicar fungicida, tais como: época do aparecimento dos sintomas, intensidade dos sintomas (incidência/severidade), taxa de evolução, condições ambientais e o custo propriamente dito do número de doses. Em média, o custo de uma dose de aplicação de fungicida corresponde a 10 sacas/hectare. É válido lembrar que a produtividade do produtor é medida em número de sacas/hectare. Portanto, considerando que o valor da saca seja R\$20,00, o custo da aplicação será de R\$200,00. Caso o produtor decida pela aplicação (A_1), existem três conseqüências possíveis: um aumento da produtividade (c_1), pequeno aumento da produtividade (c_2) ou nenhuma alteração na produtividade gerando um prejuízo devido o custo da aplicação (c_3). Caso o produtor decida por não aplicar o fungicida (A_2) a produtividade pode não ter nenhuma alteração (c_0). Suponhamos que o produtor tenha os retornos esperados como apresentado na tabela 2.3.

TABELA 2.3 – Exemplo de matriz de retorno com aplicação de fungicidas.

Consequências/Ações	A_1 (R\$)	A_2 (R\$)	Probabilidade de ocorrência
c_1	500,00	200,00	0.4
c_2	300,00	200,00	0.3
c_3	0,00	200,00	0.3
Valores Monetários Esperados(EMV)	290,00	200,00	1.0

O valor monetário esperado (EMV) de A_2 é R\$200,00, entretanto, o valor EMV de A_1 é $(R\$500,00 * 0.4) + (R\$300,00 * 0.3) + (R\$0,00 * 0.3) = R\$290,00$. Sob estes resultados, a melhor opção para o produtor seria A_1 , pois, é a opção que lhe dá um maior retorno.

Uma desvantagem da matriz de retorno é que cada vez que aumenta o número de decisões, isto é, em problemas de decisão multiestágio, uma matriz de retorno requer a enumeração de todas as estratégias possíveis e cálculo probabilístico de todas as variáveis aleatórias. Dependendo do tamanho do problema estas duas tarefas podem tornar-se computacionalmente intratáveis devido a explosão combinatória.

Um método alternativo mais flexível para representar problemas de decisão multiestágio é a árvore de decisão.

2.3.2 Árvores de Decisão

As árvores de decisão também têm sua origem no trabalho de von Neumann e Morgenstern. Uma vantagem da árvore de decisão é que uma estratégia ótima pode ser identificada utilizando métodos de programação dinâmica, portanto, não precisando enumerar todas as estratégias possíveis. Raiffa (1970) chama o método de programação dinâmica para árvores de decisão de *averaging_out_and_folding_back*.

Na árvore de decisão o problema é decomposto em nós e arcos, onde as alternativas de ações são os ramos iniciais principais e os outros eventos que podem interferir no problema de decisão vão sendo apresentados nos outros ramos da árvore com suas respectivas probabilidades de ocorrência.

De modo a clarificar a representação por árvore de decisão será aplicado o exemplo de tomada de decisão do produtor de milho (tabela 2.3), conforme Figura 2.2.

No exemplo acima, o EMV de A_1 é maior que de A_2 então A_1 deve ser selecionada como a melhor opção. A árvore de decisão se torna mais interessante quando se tem que representar mais eventos possíveis interferindo na decisão. Suponhamos um problema mais complexo, onde o produtor tem que considerar as condições de temperatura. Se as

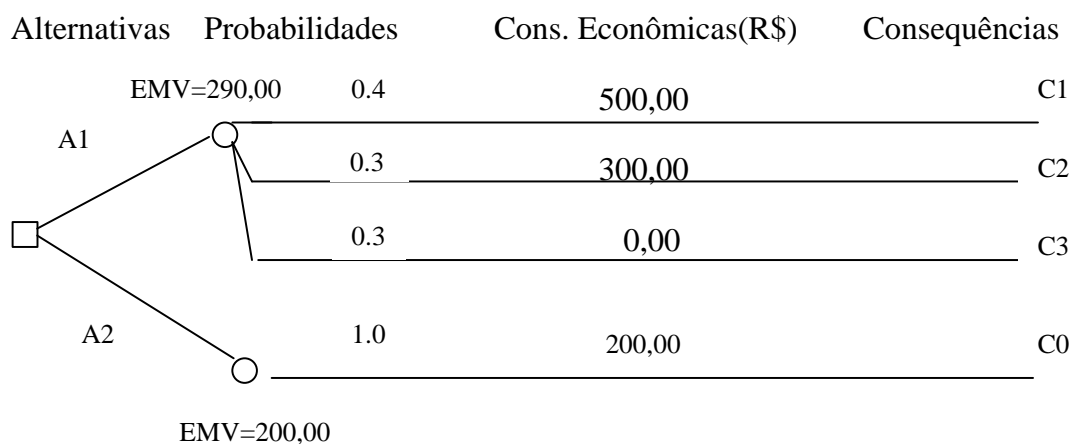


FIGURA 2.2 – Exemplo de árvore de decisão.

condições ambientais forem favoráveis (CF) ou não ($\neg CF$) para evolução da desordem. Caso seja positivo (CF), a aplicação do fungicida se torna mais necessária, caso contrário ($\neg CF$), não necessariamente é preciso aplicar fungicida. Na Figura 2.3 está representado uma nova versão do problema de decisão do produtor de milho considerando as condições ambientais.

Segundo Shenoy (1992), as árvores de decisão têm algumas desvantagens embora sejam mais expressivas e computacionalmente mais eficientes que matrizes de retorno. Na representação das árvores de decisão é necessário representar todas as ações e eventos, portanto, em alguns problemas as árvores de decisão crescem muito rapidamente. Por exemplo, um problema de decisão com n estágios ou estados com m escolhas ou eventos tem no mínimo m^n consequências (nós folha). Outra desvantagem em árvore de decisão é que as probabilidades dos eventos podem não estar disponível na forma necessária. Nestes casos, é necessário calcular estas probabilidades a partir das probabilidades disponíveis através das leis da teoria de probabilidade. Esta é a maior desvantagem da árvore de decisão. Elas deveriam separar claramente a representação do problema de sua solução. Árvores de decisão misturam estas duas tarefas tornando difícil sua automatização. Uma terceira desvantagem da árvore de decisão é que sua representação demanda uma distribuição de probabilidade condicional para cada variável randômica. Esta demanda frequentemente necessita de operações de divisões (no processamento de probabilidades) que podem não ser necessárias o que torna a solução de árvores de decisão ineficientes.

Diagramas de Influência foram inicialmente propostos como alternativa ou complementar às árvores de decisão para representar problemas de decisão (Shenoy, 1992).

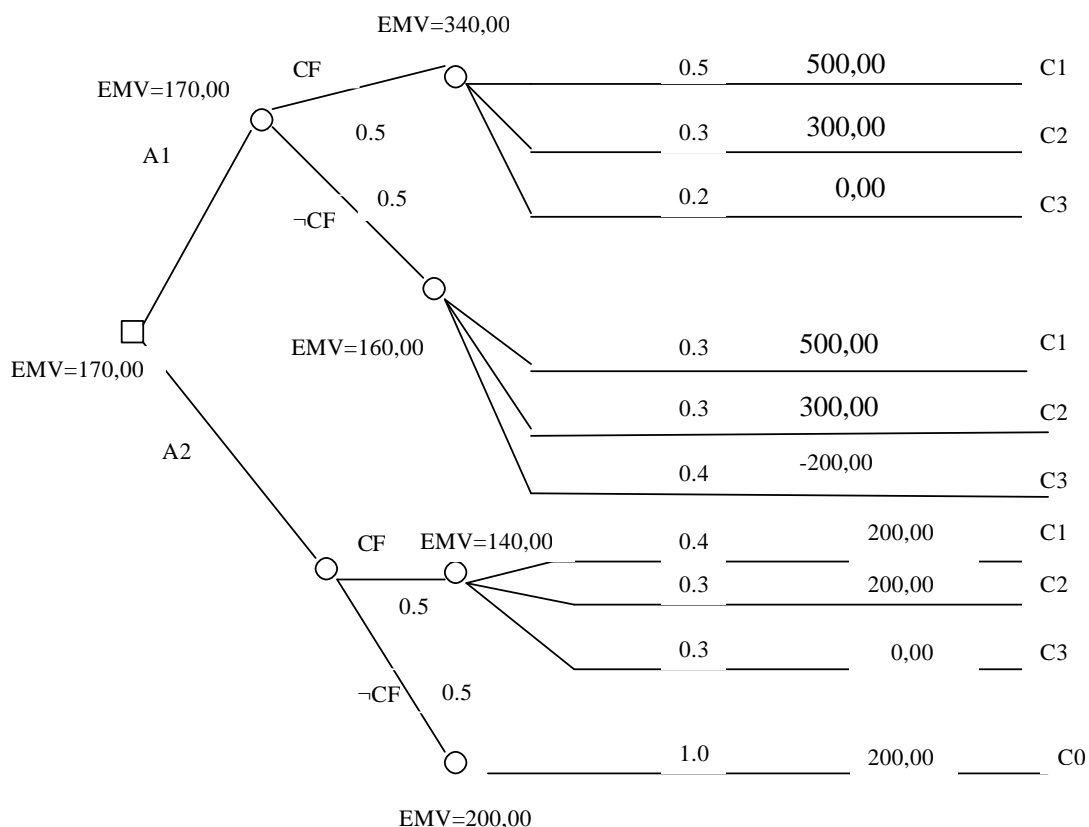


FIGURA 2.3 – Exemplo de árvore de decisão.

2.3.3 Diagramas de Influência

Nas últimas décadas, segundo Henrion e Horvitz (1992), os diagramas de influência têm se tornado mais populares como métodos para representação de problemas de decisão por não apresentar algumas das desvantagens das árvores de decisão. Nos diagramas de influência, as ações e os eventos não são representados graficamente, portanto, não crescem tão rapidamente como as árvores de decisão. As probabilidades condicionais nos diagramas de influência podem ser calculadas diretamente sem ter que calcular as probabilidades posteriores. O cálculo das probabilidades posteriores faz parte do processo de solução dos diagramas de influência. Embora este processo seja mais complexo que o processo de solução das árvores de decisão, este processo pode ser facilmente automatizado (Shenoy, 1992).

Na Figura 2.4, está apresentado um diagrama de influência correspondente a parte apresentada na Figura 2.1. O diagrama de influência pode ser visto como um grafo acíclico direcionado. Os arcos denotam influências representando relações causais incertas. O diagrama de influência inclui um modelo de custo explícito para relações causais incertas. Os losângos representam os nós de valor, os retângulos representam os nós de decisão e

os nós de probabilidade são representados por figuras ovais. Os nós de valor e decisão diferenciam diagramas de influência das redes de crença, que é um diagrama de influência que contém somente nós de probabilidade. As redes de crença permitem executar a inferência probabilística, entretanto, não permitem avaliação de decisões ou modelos de utilidade.

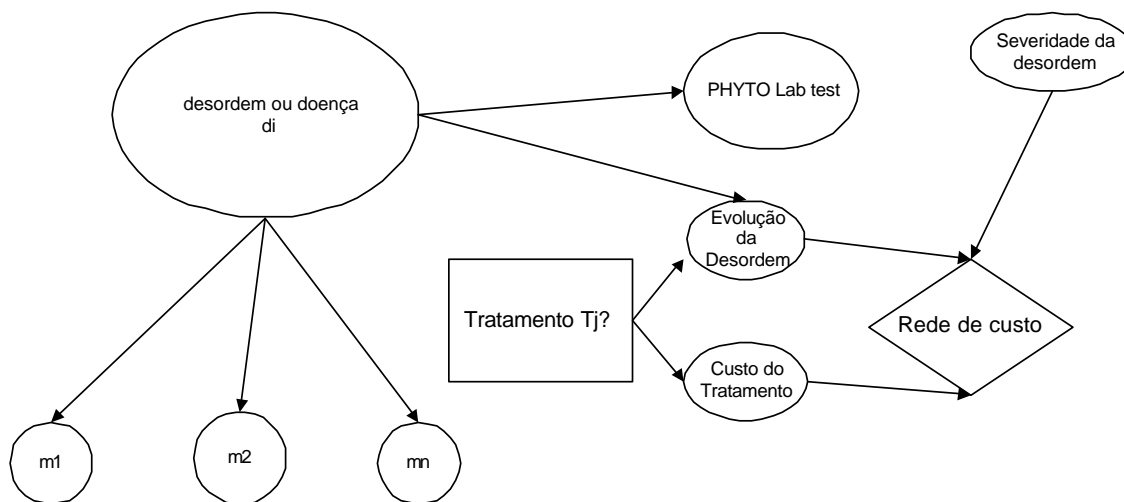


FIGURA 2.4 – Diagrama de Influências.

Neste capítulo foram apresentadas as principais técnicas em IA e TD para tomada de decisão diagnóstica e terapêutica, respectivamente. Embora estas duas áreas sejam baseadas em princípios diferentes, elas tem como objetivo comum suportar o processo de raciocínio e tomada de decisão de especialistas humanos. No próximo capítulo serão detalhados alguns conceitos básicos de abdução que serão utilizados posteriormente no modelo proposto no capítulo 4 para suportar tanto as tarefas de tomada de decisão diagnóstica quanto terapêutica de um processo de raciocínio clínico.

CAPÍTULO 3

MODELOS ABDUTIVOS

Este capítulo descreve sobre aspectos conceituais do mecanismo de inferência abdutivo que será utilizado como fundamento teórico para o modelo proposto no próximo capítulo. Primeiramente, na seção 3.1, são discutidos resumidamente os 3 tipos de definições mais comuns de inferência abdutiva: abordagem baseada em lógica, abordagem de cobertura por conjuntos e abordagem probabilística. Na seção 3.2, são apresentados os conceitos básicos da Teoria das Coberturas Parcimoniosas (TCP) como base teórica para a abordagem de cobertura por conjuntos, bem como suas limitações e extensões propostas por outros autores. Na seção 3.3 são apresentadas algumas definições básicas da TCP temporal/categórica nebulosa que serão utilizadas na Teoria das Coberturas Nebulosas proposta no capítulo 4.

3.1 Ambiente Abdutivo

Em IA, tal como descrito em MClraith (1998), abdução é definida como inferência para a melhor explicação, sem confirmação subsequente.

No escopo de IA, existem 3 abordagens principais de inferência abdutiva que diferem em suas definições no que constitui uma explicação abdutiva e, conseqüentemente, no que constitui uma melhor explicação (MClraith, 1998):

- abordagem baseada em lógica (*logic-based account*);
- abordagem de cobertura por conjuntos (*set-covering account*);
- abordagem probabilística (*probabilistic account*).

A principal definição de inferência abdutiva é a caracterização **baseada em lógica de abdução** como uma teoria. Dada uma teoria e uma observação, uma inferência abdutiva conjectura uma ou mais explicações para a observação a partir da teoria. Cada explicação deve ser consistente com a teoria e a melhor explicação depende do domínio da aplicação. Geralmente, alguma noção de minimalidade e simplicidade deve ser incorporada nos critérios de preferência. Um ponto importante na inferência abdutiva é que ela pode ser uma explicação consistente com a teoria em um determinado momento t e tornar-se inconsistente quando novas observações são adicionadas. As definições 3.1 e 3.2 contemplam a descrição formal de um ambiente abdutivo genérico, conforme descrito em MClraith (1998).

Definição 3.1 *Um ambiente abduativo genérico é uma dupla (Σ, ξ) , onde Σ é uma teoria e ξ é um conjunto do qual explicações são retiradas. Os elementos de ξ são às vezes referenciados como abduáveis.*

Definição 3.2 *Explicação abduativa - dado um ambiente abduativo (Σ, ξ) e uma observação O , $E \in \xi$ é uma explicação abduativa para O sse $\Sigma \cup E \models O$ e $\Sigma \cup E$ é consistente.*

Alguns pesquisadores tem investigado as propriedades da inferência abduativa para uma classe de teorias restrita sintaticamente, referenciada na literatura como teorias causais (ou em alguns casos como teoria de falhas)(McIlraith, 1998). Embora limitado, o ambiente abduativo causal é suficiente para algumas aplicações como na área de diagnóstico. Este tipo de inferência abduativa caracteriza a **abordagem de cobertura por conjuntos**.

Esta abordagem é melhor representada pela Teoria de Coberturas Parcimoniosas (TCP) que utiliza redes causais para representar o relacionamento entre desordens (explicações potenciais) e manifestações (observações potenciais). A partir de uma observação (uma ou mais manifestações), a TCP infere explicações que contemplam aquela observação segundo algum critério de cobertura ou parcimônia.

Formalmente, seja λ uma linguagem proposicional padrão. Tem-se as seguintes definições para um ambiente abduativo causal.

Definição 3.3 *O ambiente abduativo causal é uma tripla (C, E, Σ) onde: C (um conjunto de sentenças atômicas de λ) é o conjunto de causas; E (um conjunto de sentenças atômicas de λ) é o conjunto de efeitos que são observados e cujas causas são procuradas como explicações; Σ , é a teoria causal.*

Definição 3.4 *Dado um ambiente causal (C, E, Σ) , uma explicação abduativa para O , uma conjunção de literais a partir dos E , é K , uma conjunção de literais a partir das C , tal que $\Sigma \cup K \models O$ e $\Sigma \cup K$ é consistente.*

O ambiente abduativo causal difere do genérico baseado em lógica não somente pela restrição sintática em Σ , mas também pelo conjunto de efeitos E a partir do qual todas as observações O são retiradas. O conjunto de causas C corresponde ao ξ do ambiente genérico.

Uma desvantagem potencial da abordagem baseada em lógica, bem como da abordagem de cobertura por conjuntos para abdução, é que elas podem gerar um número muito grande de explicações.

A **abordagem probabilística** para abdução ataca este problema através da integração de redes causais com alguma noção de plausibilidade. A TCP tem sido estendida para incorporar noções de plausibilidade definido como modelo causal probabilístico conforme descrito em Peng e Reggia (1990).

Informalmente, a abdução é um forma de raciocínio hipotético que ajuda na caracterização de várias tarefas de raciocínio humano. Embora a abdução tenha sido mais empregada na área de diagnóstico por pesquisadores de IA, este mecanismo de inferência também tem sido aplicado em outros domínios, tais como, processamento de imagens, reconhecimento de padrões, entendimento da linguagem natural e atualização de base de dados (McIlraith, 1998).

Para ilustrar a aplicação de inferência abdutiva no domínio de diagnóstico, tem-se que a teoria Σ pode ser criada para representar as desordens e as manifestações que elas causam. A partir de um conjunto de observações (manifestações presentes), utilizando o mecanismo de inferência abdutiva, pode-se obter as desordens que explicam aquelas observações.

Neste trabalho, é apresentado um modelo em que a inferência abdutiva além de ser aplicada para tarefas de diagnóstico, também pode ser aplicada para tarefas de tomada de decisão em tratamentos e testes durante o processo de raciocínio clínico. No escopo de tratamentos, a teoria pode ser criada para representar os tratamentos e as desordens que eles controlam. A partir de um conjunto de explicações para um diagnóstico, utilizando o mecanismo de inferência abdutiva, pode-se obter os tratamentos indicados para tais desordens.

Antes de apresentar o modelo propriamente dito, serão apresentados alguns fundamentos da TCP que são importantes para o entendimento do modelo proposto no próximo capítulo.

3.2 Teoria das Coberturas Parcimoniosas (TCP)

A TCP é um modelo formal de inferência de diagnóstico para os modelos abdutivos baseados em associações (Peng e Reggia, 1990). A TCP é uma tentativa de formalizar o raciocínio de diagnóstico, com a vantagem que o conhecimento do domínio, as heurísticas do domínio e a metodologia de resolução do problema de diagnóstico geral são separados claramente um do outro.

A versão básica da TCP define o conhecimento do domínio como um conjunto de desordens (causas), um conjunto de manifestações (efeitos) e uma relação causal entre

desordens e manifestações. A relação causal associa cada desordem com as manifestações que ela pode causar. Se uma ou mais destas manifestações estão presentes em um caso, então a desordem pode ser usada para explicar estas manifestações.

Um problema de diagnóstico particular é definido por um conjunto de manifestações que são observadas em um paciente e a solução para aquele problema consiste em encontrar um conjunto de desordens que explica o conjunto de manifestações presentes.

O conhecimento associativo usado em problemas de diagnóstico gerais podem envolver estados intermediários e cadeias de arcos causais. Por exemplo, usualmente no diagnóstico de falhas em satélite é necessário representar a relação causal entre os equipamentos e as falhas que eles podem causar bem como as falhas que os componentes de cada equipamento podem causar. Entretanto, estes problemas que envolvem hierarquias de arcos causais são muito extensos e complexos (Peng e Reggia, 1990).

Neste trabalho está sendo considerada a versão simplificada da TCP, onde o problema de diagnóstico é restrito apenas às entidades desordens e manifestações, e as relações causais entre estes dois tipos de entidades. Os estados intermediários (patológicos ou síndromes) são desconsiderados.

3.2.1 Formulação do Problema de Diagnóstico

A TCP utiliza-se de dois conjuntos finitos que definem o escopo de problemas de diagnóstico (veja Figura 3.1).

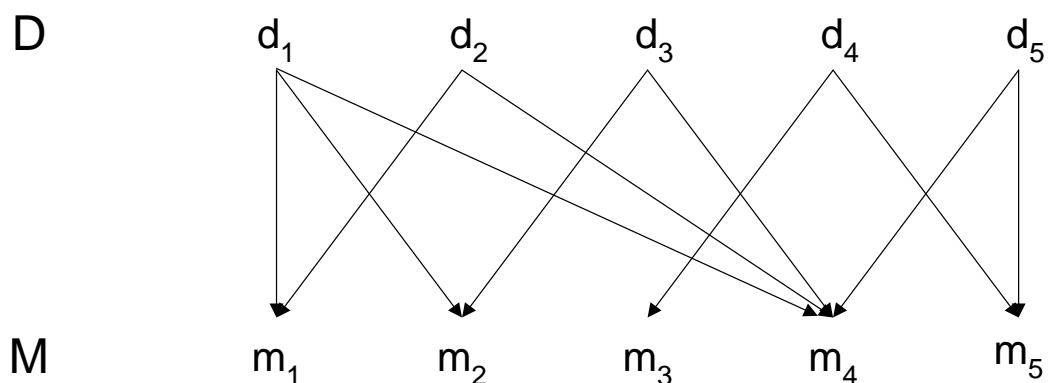


FIGURA 3.1 – Modelo de uma rede causal em TCP.

O conhecimento do domínio é representado por um conjunto de desordens D , um conjunto de manifestações M e uma relação causal $C \subseteq D \times M$. Uma associação $\langle d_i, m_j \rangle$ em C significa que d_i pode causar diretamente m_j . Portanto, os conjuntos D , M e C juntos constituem a base de conhecimento KB de um problema de diagnóstico. Formalmente, uma base de conhecimento é definida como uma tripla $KB = \langle D, M, C \rangle$.

Para completar a formulação do problema necessita-se de um caso particular. Utiliza-se M^+ como um subconjunto de M para denotar o conjunto de manifestações presentes, isto é, as manifestações conhecidas para um determinado caso.

A partir das idéias acima, define-se formalmente um problema de diagnóstico.

Definição 3.5 *Um problema de diagnóstico P é definido como um par $\langle KB, CA \rangle$ em que:*

- $KB = \langle D, M, C \rangle$ é a base de conhecimento composta por:
 - $D = \{d_1, d_2, \dots, d_n\}$ que é um conjunto finito, não vazio, de elementos denominados desordens;
 - $M = \{m_1, m_2, \dots, m_k\}$ que é um conjunto finito, não vazio, de elementos denominados manifestações;
 - $C \subseteq D \times M$ que é uma relação causal;
- $CA = \langle M^+ \rangle$ é o caso, e $M^+ \subseteq M$ é o conjunto de manifestações presentes no caso.

Pela definição, os conjuntos D , M , C juntos especificam o ambiente do problema em que o processo de diagnóstico é conduzido. Por exemplo, para o problema simples de diagnóstico da Figura 3.1, D é o conjunto formado por cinco desordens d_1, d_2, \dots, d_5 , enquanto M por cinco manifestações m_1, m_2, \dots, m_5 . O produto cartesiano $D \times M$, na definição 3.5, representa um conjunto de pares $\langle d_i, m_j \rangle$ para cada desordem distinta d_i em D e para cada manifestação distinta m_j em M . Para o problema da Figura 3.1, há no total $5 \times 5 = 25$ pares em $D \times M$. A relação C é tipicamente um pequeno subconjunto de $D \times M$, conforme ilustrado na Figura 3.1. Por exemplo, d_1 e m_1 estão casualmente associados, o que não ocorre com d_1 e m_3 . Assim, $\langle d_1, m_1 \rangle$ pertence a C , o que não é o caso de $\langle d_1, m_3 \rangle$. Na Figura 3.1, há dez pares do tipo desordem-manifestação na relação C .

A informação sobre um caso particular é descrita por um conjunto de manifestações M^+ que um indivíduo (p.ex., uma planta ou um paciente) apresenta em um dado período

de tempo (ao contrário de D , M e C , M^+ é o conhecimento de um caso específico ou instância). Este conjunto corresponde a entrada de um sistema baseado em conhecimento. Um problema de diagnóstico particular é definido por uma base de conhecimento descrita pelos conjuntos D , M e C , e pelas informações descritas em M^+ .

Na definição do problema acima, é importante ressaltar que não há nenhuma restrição na relação causal C , e na entrada do problema M^+ : qualquer desordem pode estar associada potencialmente a qualquer manifestação, e qualquer manifestação pode pertencer à M^+ .

Para um problema de diagnóstico P , é conveniente e útil definir os conjuntos ou funções baseadas na relação C :

Definição 3.6 Para todo $d_i \in D$ e $m_j \in M$ num problema de diagnóstico P ,

- $efeitos(d_i) = \{m_j \mid \langle d_i, m_j \rangle \in C\}$, o conjunto dos elementos diretamente causados por d_i ;
- $causas(m_j) = \{d_i \mid \langle d_i, m_j \rangle \in C\}$, o conjunto dos elementos que podem diretamente causar m_j .

Os conjuntos definidos acima podem facilmente ser generalizados de desordens e manifestações individualmente para conjuntos de desordens e manifestações conforme apresentado a seguir.

Definição 3.7 Para todo $D_L \subseteq D$ e $M_J \subseteq M$ num problema de diagnóstico P ,

- $efeitos(D_L) = \bigcup_{d_i \in D_L} efeitos(d_i)$,
- $causas(M_J) = \bigcup_{m_j \in M_J} causas(m_j)$.

Assim, os efeitos de um conjunto de desordens são apenas a união dos efeitos individuais de cada desordem do conjunto. Por exemplo, no problema de diagnóstico apresentado na Figura 3.1 tem-se:

$$efeitos(d_1) = \{m_1, m_2, m_4\}, \quad efeitos(d_4) = \{m_3, m_5\},$$

$$causas(m_4) = \{d_1, d_2, d_3, d_5\} \quad e \quad causas(m_5) = \{d_4, d_5\}.$$

Assim,

$$efeitos(d_1, d_4) = \{m_1, m_2, m_3, m_4, m_5\} \quad e \quad causas(m_4, m_5) = \{d_1, d_2, d_3, d_4, d_5\}.$$

3.2.2 Soluções para o Problema de Diagnóstico

Uma vez caracterizado os problemas de diagnóstico, caracteriza-se agora suas soluções. Para isto, define-se a noção de "cobertura", conforme descrito a seguir.

Definição 3.8 O conjunto $D_L \subseteq D$ é dito ser uma **cobertura** de $M_J \subseteq M$ se $M_J \subseteq \text{efeitos}(D_L)$.

Baseado na noção de "cobertura", que formaliza o termo impreciso "causalmente explica", define-se o conceito de explicação diagnóstica.

Definição 3.9 Um conjunto $E \subseteq D$ é dito ser uma **explicação** para M^+ para um problema de diagnóstico, se somente se, E cobre M^+ e satisfaz um certo critério de parcimônia.

Os critérios de parcimônia mais comuns estão descritos a seguir:

- Uma cobertura D_L de M_J é dita ser **mínima** se sua cardinalidade for a menor de todas as coberturas de M_J .
- Uma cobertura D_L de M_J é dita ser **irredundante** se nenhum de seus subconjuntos próprios for também uma cobertura de M_J ; caso contrário a cobertura é dita ser **redundante**.
- Uma cobertura D_L de M_J é dita ser **relevante** se for um subconjunto de $\text{causas}(M_J)$; caso contrário, a cobertura é dita ser **irrelevante**.

Há outros critérios de cobertura possíveis. Entretanto, por razões conceituais, a irredundância é a escolha preferível. O principal motivo é que através do conjunto de todas as coberturas irredundantes é possível gerar todas as coberturas. Assim, a não ser que se afirme ao contrário, no escopo deste trabalho quando se falar de cobertura ou explicação parcimoniosa, estar-se-á se referindo a uma cobertura irredundante.

Em diversos problemas de diagnóstico, tal como ocorrem na medicina e na fitopatologia, o diagnosticador está geralmente interessado em conhecer todas as explicações plausíveis para um caso. Estas explicações, como alternativas, podem interferir no andamento das ações a serem tomadas pelo diagnosticador. Este fato conduz á seguinte definição de solução do problema:

Definição 3.10 A **solução** de um problema de diagnóstico $P = \langle KB, CA \rangle$, denotado por $Sol(P)$, é o conjunto de todas as explicações para M^+ .

Considerando o exemplo da Figura 3.1, tem-se 2^d o número de subconjuntos que pode ser gerado a partir de $D = \{d_1, d_2, d_3, d_4, d_5\}$ onde $|D| = 5$. Para $M^+ = \{m_1, m_2, m_3\}$ tem-se que $\{d_1, d_4\}$ é um cobertura mínima. Se for considerado $\{d_2, d_3, d_4\}$ além de $\{d_1, d_4\}$ tem-se as únicas coberturas irredundantes de M^+ . Portanto, as duas coberturas irredundantes $\{d_1, d_4\}$ e $\{d_2, d_3, d_4\}$ são as únicas explicações plausíveis, isto é, constituem a solução do problema.

Além dessas duas, existem mais 3 coberturas relevantes $\{d_1, d_4, d_2\}$, $\{d_1, d_4, d_3\}$ e $\{d_1, d_2, d_3, d_4\}$, geradas pela adição das desordens extras nas coberturas irredundantes. Desde que a desordem d_5 é casualmente irrelevante para M^+ , adicionando d_5 para qualquer cobertura relevante resulta em uma cobertura irrelevante. Assim existe um total de 10 coberturas, em que 5 são irrelevantes. Os outros 22 subconjuntos restantes de D não são coberturas (como por exemplo $\{\}$, $\{d_1\}$, $\{d_2, d_4\}$, *etc...*).

Esta ferramenta separa claramente o conhecimento do domínio (conjuntos M , D e relação C), o papel do raciocínio de diagnóstico geral (critério de parcimônia e a definição da cobertura), e heurísticas do domínio diferentemente dos sistemas baseados em regras.

As incertezas também devem ser representadas para o sucesso de um problema de diagnóstico. Em Peng e Reggia (1990), a teoria de probabilidade é integrada a TCP para modelar estas incertezas gerando o que é denominado pelos autores de modelo causal probabilístico.

Neste modelo, uma probabilidade anterior p_i é associada com cada desordem d_i . Cada relação causal $\langle d_i, m_j \rangle$ tem associada uma força causal c_{ij} que representa a frequência que d_i causa m_j . Assim, o grau de verossimilhança $L(D_I, M^+)$ de qualquer explicação abduativa potencial dada a presença de M^+ pode ser calculada através dos p'_i s e c'_{ij} s relevantes. Embora o modelo causal probabilístico ter sido desenvolvido para a versão básica do TCP, ele tem sido estendido por outros pesquisadores para ser aplicado em redes causais com estados intermediários.

3.2.3 Algoritmos

Basicamente, existem duas abordagens para o desenvolvimento de algoritmos em TCP dependendo da forma em que o conjunto M^+ é apresentado. O conjunto pode ser apresentado não interativamente ou interativamente.

A primeira abordagem parece mais apropriada em situações que se pode monitorar todas as manifestações possíveis. Na segunda abordagem, as observações em M^+ são apresentadas uma de cada vez como respostas as questões do sistema de diagnóstico.

Esta abordagem parece mais adequada em situações onde o custo para se obter todas as observações é muito alto.

Os algoritmos também podem diferir pelo critério de parcimônia utilizado para definir uma explicação: irredundância ou cardinalidade mínima. Peng e Reggia (1990) apresentam dois algoritmos que utilizam cardinalidade mínima como critério de parcimônia: HT, um algoritmo iterativo e SOLVE um algoritmo não iterativo. Em Peng e Reggia (1990) também é apresentado o algoritmo BIPARTITE que usa irredundância como critério de parcimônia e que será utilizado como núcleo dos algoritmos para suportar o novo modelo proposto no trabalho corrente.

O BIPARTITE utiliza uma álgebra de geradores, uma representação compacta de explicações alternativas para o caso. Algumas definições básicas das operações que compõem a álgebra de geradores serão apresentadas, entretanto, maiores detalhes podem ser encontrados em Peng e Reggia (1990).

Se $g_1, g_2, g_3, \dots, g_m$ são subconjuntos disjuntos dois a dois e não vazios de D , então $G_I = \{g_1, g_2, g_3, \dots, g_m\}$ é um gerador, e a classe gerada por G_I é $[G_I] = \{\{d_1, d_2, d_3, \dots, d_m\} | d_i \in g_i\}$. $G = \{G_1, G_2, G_3, \dots, G_n\}$ é um conjunto gerador se cada G_I é um gerador, e $[G_I] \cap [G_J] = \emptyset$.

As operações *res*, *div*, *augres* e *revise* são definidas onde G e Q são conjunto geradores, $G_I \in G$ e $Q_J \in Q$ são geradores, $S_D \subseteq D$ é um conjunto de desordens, e $q_j \in Q_J$ é também um conjunto de desordens. De maneira a explicar melhor as operações acima, estas serão apresentadas em termos de um conjunto de explicações para um problema de diagnóstico.

Dado um conjunto de explicações para um conjunto de manifestações (M^+), representada como um conjunto gerador, e as desordens evocadas por uma nova manifestação m , representada como um conjunto de desordens, o operador *div* retorna as explicações da original M^+ que também podem explicar a nova manifestação m .

$$div(G, S_D) = \bigcup_{G_I \in G} div(G_I, S_D)$$

$$div(G_I, S_D) = \{Q_k | Q_k = \{q_{k1}, q_{k2}, \dots, q_{kn}\}\}$$

onde,

$$q_{kj} = \begin{cases} g_j - S_D & \text{se } j < k; \\ g_j \cap S_D & \text{se } j = k; \\ g_j & \text{se } j > k \end{cases}$$

A operação *res* é o dual de *div*, dado um conjunto de explicações M^+ e as desordens evocadas pela nova manifestação m , o operador *res* retorna as explicações de M^+ que não explicam a nova manifestação m .

$$res(G, S_D) = \bigcup_{G_I \in G} res(G_I, S_D)$$

$$res(G_I, S_D) = \begin{cases} \{g_1 - S_D, \dots, g_n - S_D\} & \text{se } g_i - S_D \neq \emptyset, \quad 1 \leq i \leq n \\ \emptyset & \text{cc} \end{cases}$$

As operações *div* e *res* são estendidas para trabalhar com conjuntos de conjuntos de desordens (geradores e conjuntos geradores) como seu segundo argumento.

$$div(G, Q_J) = \bigcup_{G_I \in G} div(G_I, Q_J)$$

$$div(G_I, Q_J) = \begin{cases} \{G_I\} & \text{se } Q_J = \emptyset \\ div(div(G_I, q_j), Q_J - \{q_j\}) & \end{cases}$$

$$res(G, Q) = \begin{cases} G & \text{se } Q = \emptyset \\ res(res(G, Q_J), Q - \{Q_J\}) & \end{cases}$$

$$res(G, Q_J) = \bigcup_{G_I \in G} res(G_I, Q_J)$$

$$res(G_I, Q_J) = \begin{cases} \emptyset & \text{se } Q = \emptyset \\ res(G_I, q_j) \cup res(div(G_I, q_j), Q_J - \{q_j\}) & \text{cc} \end{cases}$$

O operador *augres* é uma modificação do operador *res* que ao invés de retornar o conjunto de explicações de M^+ , que não explicam uma nova manifestação m , adiciona novas desordens para o conjunto de explicações que explica $M^+ \cup \{m\}$.

$$augres(G, S_D) = \bigcup_{G_I \in G} augres(G_I, S_D)$$

$$augres(G_I, S_D) = \begin{cases} \{\{g_1 - S_D, \dots, g_n - S_D, A\}\} & \text{se } g_i - S_D \neq \emptyset, \quad A \neq \emptyset \\ \emptyset & \text{cc} \end{cases}$$

onde,

$$A = S_D - \bigcup_{i=1}^n g_i$$

Portanto, dado um conjunto de explicações de M^+ e um conjunto de desordens evocadas por uma nova manifestação m , o conjunto de explicações $M^+ \cup \{m\}$ pode ser obtido da combinação dos operadores *div* e *augres* em uma operação denominada *revise*:

$$revise(G, S_D) = Q \cap res(Q', Q), \quad \text{onde } Q = div(G, S_D) \text{ e } Q' = augres(G, S_D)$$

O algoritmo BIPARTITE, aqui denominado BIPARTIDO, trabalha de forma sequencial e construtiva. A rede causal do problema de diagnóstico definida na seção 3.2.1 pode ser vista como um grafo bipartido e daí a origem do nome do algoritmo. O algoritmo BIPARTIDO representa suas tentativas de solução (explicações) e solução final na forma de conjunto gerador e se baseia nas operações *div*, *res* e *augres* apresentadas acima.

A função *revise* que é o núcleo do algoritmo é definida para construir as novas hipóteses através das hipóteses existentes e das desordens evocadas pela nova manifestação, conforme descrito abaixo.

```
function BIPARTIDO(D,M,C)
```

```
variables
```

```
    mnova: manifestation;
```

```
    hipoteses: generator-set;
```

```
begin
```

```
    hipoteses={ };
```

```
    while Maismanifs do
```

```
        mnova=Proxman;
```

```
        hipoteses=revise(hipoteses,causas(mnova));
```

```
    endwhile;
```

```
    return hipoteses;
```

```
end.
```

3.2.4 Limitações e Extensões da TCP

Embora a TCP seja uma ferramenta simples e poderosa para raciocínio de diagnóstico, o modelo original tem algumas limitações (Peng e Reggia, 1990). Uma de suas limitações é para representar formas mais complexas de relacionamento causal entre desordens e manifestações. Por exemplo, a TCP assume que as manifestações de duas desordens não interferem uma na outra. Não é possível representar a interferência da presença

de uma desordem nas manifestações de outra desordem, ou que se duas desordens ocorrem simultaneamente pode acontecer uma manifestação não prevista em uma dessas desordens isoladamente.

Outro problema da TCP é que a solução de um problema tende a ter muitas explicações alternativas. Irredundância como critério de parcimônia é muito fraco para reduzir significativamente o número de explicações alternativas. Uma abordagem para reduzir o tamanho da solução do problema de diagnóstico é adicionar informação probabilística na relação causal como apresentado em Peng e Reggia (1990).

No entanto, este conceito pode ser visto de uma maneira mais geral, como na modelagem de incerteza que pode ser feita pela utilização da teoria de probabilidade, da teoria de possibilidades ou da teoria de Dempster-Shafer e fatores de certeza. O conhecimento de incerteza (usualmente, em forma numérica em sistemas de diagnóstico baseados em conhecimento) especifica a força das associações e quantifica as evidências individuais dos casos. A incerteza pode ser representada diretamente através de números de dentro da modelagem de incerteza ou na forma simbólica não-numérica ("muito comum", "moderado", "muito raro", etc.), termos estes que podem ser tratados como tal, no que se chama atualmente de "*computing with word*" ou traduzidos em números.

Outra solução para reduzir o tamanho da solução do problema de diagnóstico é usar heurísticas específicas do domínio que ajudam a selecionar dentro de um subconjunto de explicações irredundantes as mais plausíveis e, conseqüentemente, gerando uma solução mais eficiente.

Vários autores têm investigado tanto a teoria quanto a implementação da teoria da diagnose e proposto outros modelos. Lucas (1996) permite representar vários tipos de interações que não são possíveis na TCP original.

Dubois e Prade (1995) propõem uma extensão possibilista à TCP para modelar as incertezas. Wainer e Rezende (1997) propõem uma extensão temporal à TCP de modo que informações categóricas e temporais possam ser incorporadas na base de conhecimento. A informação temporal permite representar a evolução esperada das manifestações causadas por determinada desordem. A informação categórica permite distinguir entre manifestações que são somente possíveis daquelas que são necessárias durante a evolução de uma desordem.

O trabalho de Wainer e Sandri (1998) mostra que melhor do que representar informação temporal como intervalos bem definidos (crisp) é representá-los como intervalos nebu-

losos. Na seção 3.3 serão apresentados os conceitos de extensão nebulosa à TCP propostos em Wainer e Sandri (1998) que posteriormente serão utilizados no capítulo 4.

No apêndice A são introduzidas algumas definições básicas de conjuntos nebulosos e teoria de possibilidades que serão utilizadas na próxima seção e no decorrer dos próximos capítulos.

3.3 Extensão Nebulosa à TCP Temporal

A representação da informação temporal proposta em Wainer e Rezende (1997) é feita através de um grafo $G_l = \langle V_l, A_l \rangle$ para cada desordem d_l no qual os nós V_l representam as manifestações e os arcos A_l a precedência temporal. A informação quantitativa sobre a duração de uma dada manifestação está associada com seu nó correspondente e a informação quantitativa relacionada ao tempo entre o início de duas manifestações está associada com o arco correspondente. A informação quantitativa é modelada por intervalos $I = [I^-, I^+]$ definidos em uma escala de tempo onde I^- e I^+ são pontos extremos do intervalo I . A Figura 3.2 ilustra o grafo associado com uma desordem d_i . Neste exemplo, a manifestação m_2 ocorre de "24 a 48 horas" após a ocorrência da manifestação m_1 , a manifestação m_3 ocorre de "36 a 48 horas" após a ocorrência da manifestação m_2 e assim sucessivamente.

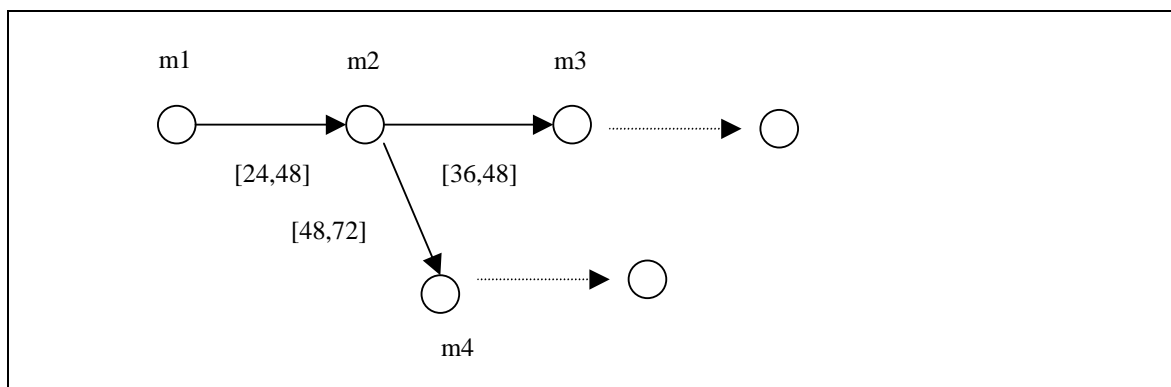


FIGURA 3.2 – Modelo de um grafo temporal associado a uma desordem.

O trabalho de Wainer e Sandri (1998) mostra que melhor do que representar informação temporal como intervalos bem definidos (*crisp*) é representá-los como intervalos nebulosos. Na maioria dos domínios, as restrições temporais no modelo usualmente representam o conhecimento acumulado destes intervalos para um grande número de casos.

Cada um desses pedaços de informação temporal pode ser modelado como um intervalo bem definido, entretanto, estes intervalos podem ser muito pequenos para conter casos falsos ou muito grandes para conter informação útil. Intervalos nebulosos podem auxiliar a definir melhor os casos típicos e possíveis de um determinado domínio.

Por exemplo, seja d_1 uma doença na qual o sintoma m_1 ocorre 'por volta de 24 a 48 horas' antes do sintoma m_2 como representado na Figura 3.2. Seja um caso X onde o sintoma m_1 precedeu o sintoma m_2 'por volta de 21 horas'. Em um sistema automatizado em que o intervalo fosse modelado por $I = [24, 48]$, d_1 seria descartado para este caso. Entretanto, isto não deveria acontecer em certos domínios. Seja agora um intervalo I_1 mais amplo tal como $[21, 52]$ para modelar 'por volta de 24 a 48 horas' para d_1 e um intervalo $I_2 = [18, 26]$ usado para modelar 'por volta de 20 a 24 horas' entre m_1 e m_2 em d_2 . O sistema não perceberia que d_2 é uma melhor explicação para X do que d_1 . Entretanto, isto não deveria acontecer se intervalos nebulosos fossem usados para modelar estas informações temporais. Através da extensão nebulosa, além das explicações para um determinado caso serem melhor definidas em relação a um modelo, também se pode classificar melhor as explicações mais plausíveis.

3.3.1 Base de Conhecimento e o Caso Particular

A base de conhecimento proposta em Wainer e Sandri (1998) para um problema de diagnóstico temporal nebuloso é a informação sobre uma desordem particular e como ela evolui.

Definição 3.11 *A base de conhecimento é dada por $\langle \theta, D, M, N, P, V, T \rangle$ onde:*

- θ é a escala de tempo;
- D é o conjunto de desordens;
- M é o conjunto de manifestações;
- N é a função de efeitos necessários que associa a cada desordem d_i um conjunto $M_i \subseteq M$ de manifestações que d_i necessariamente causa. Por exemplo, se $N(d_1) = \{m_4, m_5, m_7\}$ então não é possível ter a desordem d_1 sem ter eventualmente os sintomas m_4, m_5 e m_7 ;
- P é a função de efeitos possíveis que associa a cada desordem d_i um conjunto $M_i \subseteq M$ de manifestações que d_i pode causar;
- V associa a cada desordem d um conjunto de eventos. Estes eventos serão usados para descrever a evolução da desordem. Em $V(d)$, devem ser incluídos os eventos que correspondem ao início de todas as manifestações em $E(d)$

onde $E(d) = N(d) \cup P(d)$ é a função derivada que produz todos os efeitos de uma desordem. Adicionalmente, $V(d)$ pode incluir eventos que correspondem ao final de algumas manifestações em $E(d)$ e pode incluir outros eventos que também não são observáveis. Por exemplo, um evento não observável em medicina é a própria ingestão de um alimento contaminado.

- T é uma função que associa a alguns pares de eventos $e_i, e_j \in V(d)$ um intervalo temporal nebuloso $T(d)(e_i, e_j) = p$ que representa que o tempo entre a ocorrência de e_i e e_j devem estar dentro de um intervalo temporal nebuloso p (se e_i ocorre antes de e_j então $T(d)(e_i, e_j)$ é um conjunto nebuloso positivo).

As funções $V(d)$ e $T(d)$ podem ser melhor representadas em termos de um grafo de eventos como descrito na definição 3.12.

Definição 3.12 O grafo temporal de eventos de uma desordem $d_i \in D$, $G_i = (V, T)$, é um grafo direcionado acíclico e transitivo (veja Figura 3.3), onde:

- $V(d)$ são os nós do grafo que representam os eventos de início ou término de uma manifestação $m_j \in M$ rotulados por (m_1^b, m_1^e) ;
- $T(d)$ é definido sobre o par de eventos (e_i, e_j) constituindo os arcos do grafo rotulados de $\pi(e_i, e_j)$ que representam os intervalos nebulosos mínimos entre quaisquer dois eventos que possam ocorrer na progressão de uma doença d_i .

Para ilustrar o grafo temporal de eventos, consideremos a seguinte base de conhecimento para a desordem d_1 apresentada na Figura 3.1:

- θ é a escala de tempo em horas;
- $D = \{d_1\}$;
- $M = \{m_1, m_2, m_4\}$;
- $N = \{m_1\}$;
- $P = \{m_2, m_4\}$;

Na Figura 3.3 está representado um modelo simplificado do grafo temporal de eventos $G(V, T)$ da desordem d_1 . Os intervalos 'menos de 24 horas', 'por volta de 6 a 24 horas' e 'por volta de 24 a 48 horas' são dados fictícios que correspondem aos intervalos $\langle 0, 24 \rangle$, $\langle 6, 24 \rangle$ e $\langle 24, 48 \rangle$, respectivamente.

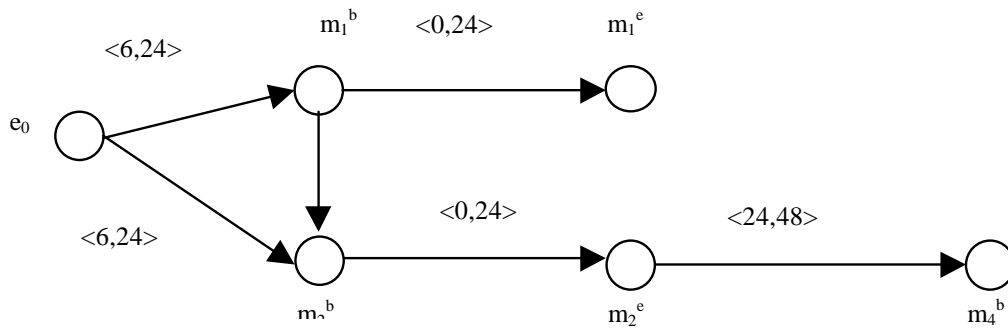


FIGURA 3.3 – Modelo de um grafo de eventos associado a uma desordem no escopo da TCN.

Na Figura 3.3 estão representados os intervalos nebulosos entre alguns eventos. Entretanto, em Wainer e Sandri (1998) é proposto como representar o intervalo mínimo entre todos os pares de eventos de uma desordem d_i . Dado um conjunto de intervalos entre alguns eventos como apresentado na Figura 3.3, a rede mínima é uma maneira de calcular os intervalos mínimos entre quaisquer dois eventos.

O grafo mínimo para cada desordem d_i pode ser calculado pelo algoritmo de Floyd-Warshall, que calcula o caminho mínimo entre todos os pares como descrito em Wainer e Sandri (1998).

É assumido que para todo d os eventos em $V(d)$ são numerados, e que $|V(d)| = n$. O algoritmo calcula cada valor t_{ij} que é o intervalo entre os eventos e_i e e_j na rede mínima de uma desordem d específica.

- 1) for $i = 1$ to n do
 - 2) for $j = 1$ to n do
 - 3) if $i = j$ then $t_{ii} = I_{=0}$
 - 4) else if $T(d_l)(e_i, e_j)$ then $t_{ij} = T(d_l)(e_i, e_j)$
 - 5) else if $T(d_l)(e_j, e_i)$ then $t_{ij} = -T(d_l)(e_i, e_j)$
 - 6) else $t_{ij} = I_{anytime}$
 - 7) for $k = 1$ to n do
 - 8) for $i = 1$ to n do
 - 9) for $j = 1$ to n do
 - 10) $t_{ij} = t_{ij} \cap (t_{ik} \oplus t_{kj})$

Em Wainer e Sandri (1998) é proposta uma função $\pi(e_i, e_j)$ que retorna o valor de t_{ij} na rede mínima para a desordem d_l . Em termos da analogia do grafo de V e T , a rede mínima calcula o fecho transitivo do grafo. A principal restrição do grafo mínimo de

eventos $G_i = (V, T)$ é a impossibilidade de definir ciclos, assim não é possível representar eventos recorrentes. Por outro lado, uma grande vantagem é que a transitividade do grafo é garantida através da função $T(d)$. Por exemplo, seja $\pi_1 = \langle a_1, b_1, c_1, d_1 \rangle$ e $\pi_2 = \langle a_2, b_2, c_2, d_2 \rangle$, $T(d)(e_i, e_j) = \pi_1$ e $T(d)(e_j, e_k) = \pi_2$ então $T(d)(e_i, e_k) = \pi_1 \oplus \pi_2 = \langle a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2 \rangle$.

A Figura 3.4 representa parte do grafo mínimo de eventos para a desordem (d_1) . Nesta Figura estão representados apenas os intervalos nebulosos que serão utilizados nos exemplos desta seção.

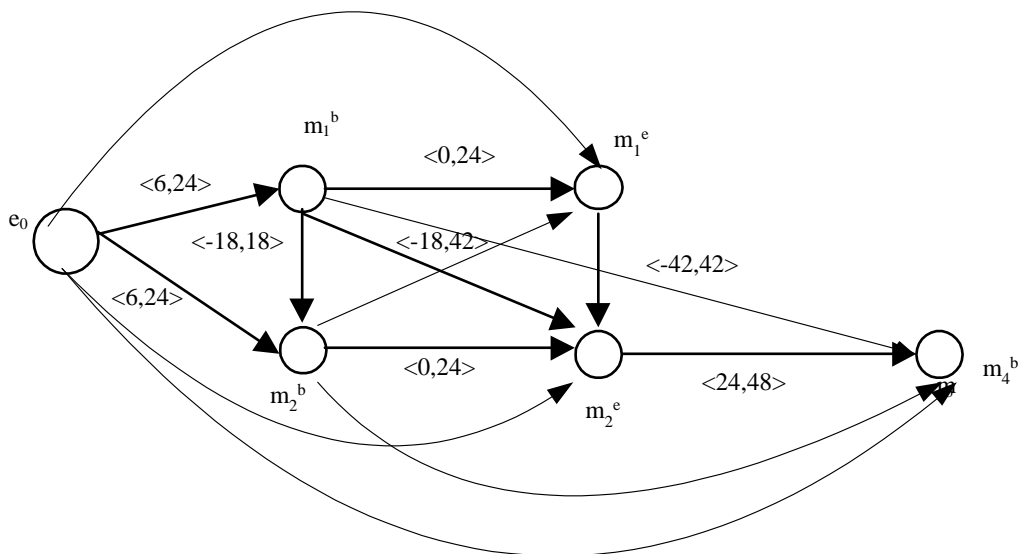


FIGURA 3.4 – Modelo parcial de um grafo mínimo associado a uma desordem d_i .

Em um problema de diagnóstico, além da base de conhecimento, precisa-se definir a informação sobre um caso particular. A informação sobre um caso particular no escopo de Wainer e Sandri (1998) precisa informar sobre as manifestações que ocorreram, quando elas começaram e finalizaram, e sobre manifestações que não ocorreram. Formalmente, a informação sobre um dado caso pode ser definida como a seguir.

Definição 3.13 A informação sobre um dado caso é modelada por uma quintupla $C_a = \langle M^+, M^-, EV^+, TIME^+, \theta_0 \rangle$, onde:

- M^+ é o conjunto de manifestações presentes conhecidas no caso;
- M^- é o conjunto de manifestações ausentes conhecidas no caso;
- EV^+ é o conjunto de eventos de início e fim de cada manifestação em M^+ ;

- $TIME^+$ é uma função que associa para cada evento $e_i \in EV^+$ um intervalo nebuloso que representa os momentos possíveis no qual o evento ocorreu;
- θ_0 é o momento do diagnóstico.

Um exemplo de um caso particular para exemplificar como diagnosticar a desordem d_1 no escopo do trabalho de Wainer e Sandri (1998) poderia ser modelado como $M^+ = \{m_2\}$, $M^- = \{m_1\}$, e $EV^+ = \{m_2^b, m_2^e\}$. Se for considerado o momento do diagnóstico $\theta_0 = 120$ então, $TIME^+(m_2^b) = \langle 96, 96 \rangle$ e $TIME^+(m_2^e) = \langle 101, 102 \rangle$.

3.3.2 Consistência entre o Modelo de Desordens e o Caso

Segundo Wainer e Sandri (1998), a consistência temporal entre o caso e o modelo é calculada pela comparação do tempo transcorrido entre os eventos do caso (os eventos em EV^+) e os intervalos nebulosos correspondentes no modelo.

Primeiramente, deve ser calculada a distância temporal entre pares de todos os eventos de EV^+ como descrito na definição a seguir.

Definição 3.14 Dado que e_i e $e_j \in EV^+$ de um caso particular C_a , calcula-se a distância temporal nebulosa entre as ocorrências reais dos eventos e_i e e_j através da fórmula $DIST^+(e_i, e_j) = TIME^+(e_j) \ominus TIME^+(e_i)$.

Na sequência, para comparar como estes dois eventos se encaixam no modelo de uma desordem particular d_i deve-se comparar a $DIST^+(e_i, e_j)$ com $\pi(e_i, e_j)$. Um caso é consistente temporalmente com o modelo de uma desordem se seu grau de consistência for maior que zero. O grau de consistência temporal de dois eventos e_i e e_j é a altura da intersecção de $DIST^+(e_i, e_j)$ com $\pi(e_i, e_j)$. Formalmente, a consistência temporal entre um caso C_a e uma desordem d_i é definido como a seguir.

Definição 3.15 Um caso particular C_a é consistente temporalmente com o modelo de uma desordem d_i se seu grau de consistência temporal for maior que zero, isto é, se $\alpha(d_i) = \min_{e_i, e_j \in EV^+ \cap V(d_i)} [h(DIST^+(e_i, e_j) \cap \pi_i(e_i, e_j))] > 0$.

Para exemplificar a consistência temporal consideremos novamente o conjunto de manifestações $M = \{m_1, m_2, m_4\}$ que caracterizam a desordem (d_1).

Suponhamos o caso hipotético 1 em que $M^+ = \{m_1, m_2\}$; $EV^+ = \{m_1^e, m_2^e\}$; $TIME^+(m_1^e) = \langle 101, 102 \rangle$; $TIME^+(m_2^e) = \langle 57, 58 \rangle$ e $DIST^+(m_1^e, m_2^e) = \langle 43, 45 \rangle$.

A partir do grafo mínimo (Figura 3.4) tem-se $\pi(m_1^e, m_2^e) = \langle -42, 42 \rangle$ e consequentemente $\alpha(d_i) = h(DIST^+(m_1^e, m_2^e) \cap \pi(m_1^e, m_2^e)) = h(\langle 43, 45 \rangle \cap \langle -42, 42 \rangle) = 0$.

Portanto, o modelo e o caso são completamente incompatíveis em termos temporais. Realmente, se considerarmos o $TIME^+(m_1^e)$ de m_1 , o evento e_0 deveria ter ocorrido no intervalo $\langle 53, 96 \rangle$ e considerando o $TIME^+(m_2^e)$, e_0 deveria ter ocorrido no intervalo $\langle 9, 52 \rangle$, o que os torna totalmente incompatíveis.

Consideremos um segundo exemplo em que $M^+ = \{m_1, m_2\}$; $EV^+ = \{m_1^e, m_2^e\}$; $TIME^+(m_1^e) = \langle 100, 101, 102, 103 \rangle$; $TIME^+(m_2^e) = \langle 56, 57, 58, 59 \rangle$ e $DIST^+(m_1^e, m_2^e) = \langle 41, 43, 45, 47 \rangle$.

A partir do grafo mínimo (Figura 3.4) tem-se $\pi(m_1^e, m_2^e) = \langle -42, 42 \rangle$ e consequentemente $\alpha(d_i) = h(DIST^+(m_1^e, m_2^e) \cap \pi(m_1^e, m_2^e)) = h(\langle 41, 43, 45, 47 \rangle \cap \langle -42, 42 \rangle) = 0.5$ (veja Figura 3.5).

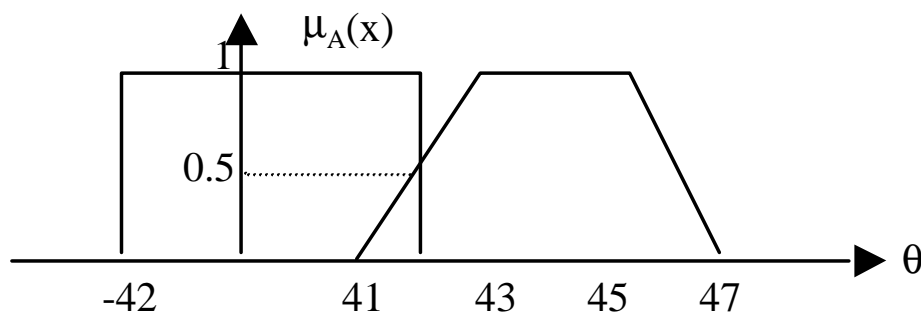


FIGURA 3.5 – Exemplo do grau de consistência temporal.

Assim, a desordem (d_1) poderia ser considerada uma explicação plausível para M^+ .

Adicionalmente, conforme apresentado em Wainer e Sandri (1998), a consistência categórica está intimamente ligada à consistência temporal, pois, um caso é categoricamente inconsistente com um modelo caso uma manifestação necessária não tenha ocorrido e já não há mais tempo suficiente para que ela ocorra. Uma manifestação m_j é dita ter tempo suficiente para ter ocorrido em d_i se:

- existe um evento e_j que deveria ocorrer após m_j e ele já ocorreu;

- ou existe um evento e_j que era previsto ocorrer antes de m_j , ele ocorreu e o tempo previsto para ocorrer m_j após e_j já estourou.

Assim, o índice de consistência temporal $\alpha(d_i)$ refletirá ambos, consistência temporal e categórica, e será denominado $\alpha_{ct}(d_i)$ (definição 3.15).

Wainer e Sandri (1998) propuseram o uso de quatro intervalos nebulosos para descrever o momento de ocorrência de um evento que serão utilizados neste trabalho.

Definição 3.16 *Seja θ_0 o momento do diagnóstico Θ . Os intervalos nebulosos descrevem a possibilidade de um evento ocorrer a qualquer momento, exatamente no momento θ_0 , após o momento do diagnóstico, ou antes do momento do diagnóstico, que são definidos respectivamente por:*

- $I_{anytime} = A$, tal que $\forall x \in \Theta, \mu_A(x) = 1$;
- $I_{I=\theta_0} = A$, tal que $\mu_A(x) = 1$ se $x = \theta_0$, e $\mu_A(x) = 0$ caso contrário;
- $I_{afternow} = I_{I \geq \theta_0} = A$, tal que $\forall x \in \theta$, se $x \geq \theta_0$, $\mu_A(x) = 1$ caso contrário $\mu_A(x) = 0$;
- $I_{beforenow} = I_{I \leq \theta_0} = A$, tal que $\forall x \in \theta$, se $x \leq \theta_0$, $\mu_A(x) = 1$ caso contrário $\mu_A(x) = 0$.

Agora, já se pode definir quando um caso particular é consistente temporal e categoricamente com o modelo de uma desordem como definido a seguir.

Definição 3.17 *Um caso particular C_a é consistente temporal e categoricamente com o modelo de uma desordem d_i se:*

- *tiver a seguinte inicialização, $\forall m_j \in M^- \cap N(d)$, $TIME^+(m_j^b) = I_{afternow} = I_{I \geq 0}$ (veja apêndice A)*

Para exemplificar a consistência categórica, consideremos o caso hipotético em que o momento de diagnóstico é igual a 120 (semanas): $M^+ = m_2, m_4$; $M^- = m_1$; $EV^+ = m_2^e, m_4^b$; $TIME^+(m_2^e) = \langle 91, 92 \rangle$ e $TIME^+(m_4^b) = \langle 114, 116 \rangle$ $DIST^+(m_2^e, m_4^b) = \langle 22, 25 \rangle$ e $\pi_1(m_2^e, m_4^b) = \langle 24, 48 \rangle$.

a) Se forem considerados m_2 e m_4 sem considerar m_1 , tem-se que o caso é compatível com a desordem d_1 conforme apresentado abaixo:

$$\alpha(d_1) = h(DIST^+(m_2^e, m_4^b) \cap \pi_1(m_2^e, m_4^b)) = h(\langle 22, 25 \rangle \cap \langle 24, 28 \rangle) = 1.0.$$

b) Se for considerada a ausência da manifestação m_1 , com $TIME^+(m_1^b) = I_{afternow} = \langle 120, \infty \rangle$, o caso já não é compatível com o modelo de d_1 , conforme apresentado abaixo:

$$\alpha(d_1) = \min \left\{ \begin{array}{l} (h(DIST^+(m_2^e, m_4^b) \cap \pi_l(m_2^e, m_4^b))), \\ (h(DIST^+(m_1^b, m_2^e) \cap \pi_l(m_1^b, m_2^e))), \\ (h(DIST^+(m_1^b, m_4^b) \cap \pi_l(m_1^b, m_4^b))) \end{array} \right.$$

$$\alpha(d_1) = \min \left\{ \begin{array}{l} h(\langle 22, 25 \rangle \cap \langle 24, 48 \rangle), \\ h(\langle -\infty, -28 \rangle \cap \langle -18, -42 \rangle), \\ h(\langle -\infty, -4 \rangle \cap \langle 6, 90 \rangle), \end{array} \right.$$

$$\alpha(d_1) = \min(1, 0, 0) = 0.$$

Adicionalmente à consistência temporal e categórica, em algumas desordens é importante quantificar a intensidade com que a manifestação ocorre. Por exemplo, para uma determinada doença a temperatura tem que ser alta para que ela se desenvolva. Portanto, em um caso em que a temperatura é baixa talvez diminua a possibilidade de ser aquela desordem a explicação para os sintomas apresentados pelo caso.

Em Wainer e Sandri (1998), para obter consistência de intensidade, utiliza-se uma função INT na base de conhecimento que atribui para cada nó m de $E(d)$ um conjunto nebuloso $INT(m)$ descrevendo a intensidade com que aquela manifestação deve ocorrer. Cada conjunto nebuloso $INT(m)$ é definido sobre um domínio particular $\Omega_{INT(m)}$.

Quando a intensidade não é relevante para um determinado problema, $INT(m)$ é :

$$\forall x \in \Omega, \mu_{INT(m)}(x) = 1 \quad (3.1)$$

Quando a intensidade pode ser quantificada por uma constante precisa x^* em X, então $INT(m)$ é:

$$\mu_{INT(m)}(x) = 1 \text{ se } x = x^* \text{ e } \mu_{INT(m)}(x) = 0 \text{ cc} \quad (3.2)$$

Do mesmo modo, para a informação do caso existe uma função INT^+ que atribui para cada $m_j \in CA = \{M^+\}$ um conjunto nebuloso $INT^+(m)$ descrevendo a intensidade do qual a manifestação ocorreu.

Portanto, a consistência de intensidade de uma manifestação m_j em relação a uma doença d_i é medida por:

$$\gamma(m_j) = h(INT(m_j) \cap INT^+(m_j)) \quad (3.3)$$

Finalmente, para uma determinada doença d_i a consistência de intensidade é:

$$\gamma(d_i) = \text{inf}_{m_j \in KB_d} \gamma(m_j) \quad (3.4)$$

Para exemplificar o índice de intensidade, seja o caso em que a manifestação m_1 corresponde á alta temperatura do paciente. Seja $\Omega_{INT(m_1)} = [0, 10]$ um domínio para intensidade de temperatura e seja m_1 modelado por $INT(m_1) = \langle 7.5, 8.5, 9.5, 10 \rangle$.

Suponhamos o caso modelado pelo intervalo nebuloso $INT^+(m_1) = \langle 4, 5, 7, 8 \rangle$ sobre o mesmo domínio $\Omega_{INT(m_1)}$. Assim, o grau de consistência para m_1 é

$$\gamma(m_1) = h \langle 7.5, 8.5, 9.5, 10 \rangle \cap \langle 4, 5, 7, 8 \rangle = 1.0$$

Aplicando a fórmula 3.4 tem-se que as manifestações presentes em C_a são consistentes com a evolução da desordem d_1 .

3.3.3 Explicação para o Diagnóstico

A melhor explicação para o diagnóstico assumida em Wainer e Sandri (1998) é uma doença que tem consistência temporal, categórica e de intensidade com todos os sintomas e explica todos os sintomas presentes no caso. Assim, d_i é um diagnóstico para o caso $C_a = \langle M^+, M^-, EV^+, TIME^+, \theta_0 \rangle$, se:

- $\alpha_{ct}(d_i) > 0$,
- $\gamma_{ct}(d_i) > 0$ e
- $\forall m_i \in M^+ \in E(d_i)$.

3.3.4 Limitações

No modelo de Wainer e Sandri (1998) descrito acima é proposta a incorporação de conjuntos nebulosos em alguns conceitos de TCP temporal. Entretanto, este trabalho está simplificado em relação ao modelo temporal original (Wainer e Rezende, 1997), pois está somente interessado em verificar como uma desordem única cobre um conjunto de manifestações ao invés de construir uma teoria de diagnóstico completa.

Um outro ponto em aberto no modelo de Wainer e Sandri (1998) é como os vários índices de consistência podem ser usados em uma situação prática de diagnóstico. Por exemplo, se os índices de consistência para uma desordem não são todos iguais a 1 existe mais de uma maneira de usá-los: combiná-los em um único índice, definir qual índice é

mais importante (este grau de importância pode depender de heurísticas do domínio). Além disso, estes índices podem ser usados para classificar hipóteses, utilizando alguma heurística.

Um outra limitação do modelo apresentado se refere á informação categórica nebulosa. É necessário investigar melhor a informação categórica nebulosa, como por exemplo 'na doença d_i a manifestação m_i raramente ocorre'. Isto faria com que N e P fossem substituídos por distribuições de possibilidades. Outro ponto é modelar a incerteza de um determinada manifestação ter ocorrido ou não, que é diferente de modelar a intensidade da manifestação como foi apresentado neste trabalho (Dubois e Prade, 1995).

Um outro ponto interessante de ser investigado em diagnóstico, e não contemplado pelo modelo de diagnóstico de Wainer e Sandri (1998), são as condições favoráveis para evolução de uma desordem. Em fitopatologia, por exemplo, temperatura, umidade bem como a época do aparecimento dos sintomas podem favorecer o desenvolvimento de uma determinada desordem d_i . No diagnóstico de satélites, temperatura e níveis de radiação podem favorecer a ocorrência de determinadas falhas.

Um ponto também crucial em diagnóstico, que não está contemplado no modelo de Wainer e Sandri (1998), é identificar quais as manifestações prioritárias a serem investigadas para obter um diagnóstico mais rápido. No domínio agrícola, isto pode ser vital, inclusive podendo vir a comprometer toda uma produção ou um replanejamento agrícola. Isto pode ser feito através de uma função de utilidade em cima das manifestações de modo a reduzir hipóteses.

Além das etapas de diagnóstico e investigação, um outro ponto crucial no processo de raciocínio clínico que não é suportado pelo modelo de Wainer e Sandri (1998) é o mapeamento correto do diagnóstico em tratamentos. No capítulo 4 é apresentado o modelo proposto no presente trabalho para suportar uma teoria de diagnóstico completa englobando investigação e tratamentos além de incorporar algumas das funcionalidades não suportadas nos modelos de TD e IA apresentados nos capítulos 2 e 3.

CAPÍTULO 4

MODELO PROPOSTO

Neste capítulo é apresentada a Teoria das Coberturas Nebulosas (TCN) proposta para suportar a abordagem integrada para diagnóstico, investigação e tratamentos que incorpora conceitos de inferência abdutiva, lógica nebulosa e teoria de decisão. Primeiramente, na seção 4.1 é definido o escopo do problema de raciocínio clínico que engloba diagnóstico, investigação e tratamentos. Na seção 4.2 é apresentada uma breve revisão da abordagem adotada na literatura para este problema. Em seguida, na seção 4.3 é apresentada a infraestrutura geral proposta no presente trabalho para suportar o problema de raciocínio clínico. Na seção 4.4 é apresentada a definição formal para o problema. Finalmente, na seção 4.5 está caracterizada formalmente a solução de um problema de raciocínio clínico que será detalhada posteriormente nos capítulos 5, 6 e 7.

4.1 Escopo do Problema

Nas últimas décadas, tem havido um grande crescimento na capacidade das pessoas de compreender o raciocínio humano e, em particular, o raciocínio clínico. As pesquisas realizadas nas disciplinas da ciência cognitiva, teoria de decisão e ciência da computação têm fornecido uma ampla visão do processo cognitivo que forma a base das decisões diagnósticas e terapêuticas que explicam um processo de raciocínio clínico, onde

- a decisão diagnóstica é realizada quando uma hipótese atinge uma certo grau de verossimilhança (grau de verdade ou falsidade de uma hipótese);
- a decisão terapêutica (medidas de controle) depende dos objetivos pretendidos e da efetividade esperada.

Segundo Réa-Neto (1998), o raciocínio clínico é uma função essencial da atividade médica e a eficiência de um atendimento médico é altamente dependente da análise e síntese adequada dos dados clínicos e da qualidade das decisões envolvendo riscos e benefícios dos testes e tratamento.

Conforme apresentado em Réa-Neto (1998), o sucesso de um problema clínico depende de 2 fatores: diagnóstico correto e tratamento efetivo. A solução para estes dois objetivos envolve duas grandes fases de tomada de decisão:

- a designação de um diagnóstico num nível de especificidade adequado para as considerações terapêuticas;
- seleção de um tratamento que afete o problema de modo a resolvê-lo ou aliviá-lo.

Além dessas 2 fases macros, existem outros estágios de tomada de decisão de menor ordem:

- selecionar perguntas a fazer;
- decidir que respostas são confiáveis ;
- interpretar um sinal físico;
- selecionar um ou mais testes de laboratório;
- escolher uma ou mais das alternativas de tratamento.

Resumidamente, o processo de solução dos problemas clínicos é constituído por dois grandes componentes que usualmente são considerados separadamente, embora eles provavelmente não possam ser separados na prática, no caso de especialistas humanos:

- uma base de conhecimento rica e extensa;
- um mecanismo de inferência, que é o método de aplicação do conhecimento utilizado pelo diagnosticador na busca da solução do problema.

O método de aplicação do conhecimento utilizado pelo diagnosticador pode ser descrito segundo vários métodos, dentre os mais utilizados tem-se: o método hipotético-dedutivo e o método abduutivo. No escopo deste trabalho o processo de utilização do conhecimento é descrito segundo o método abduutivo, conforme será apresentado na seção 4.3 em contrapartida ao método hipotético-dedutivo.

Na próxima seção está descrito detalhadamente o processo de solução para problemas clínicos visando facilitar o entendimento dos próximos capítulos.

4.1.1 Processo de Solução de Problemas Clínicos

Conforme apresentado em Réa-Neto (1998), o processo de raciocínio para tomada de decisão diagnóstica e terapêutica é constituído das seguintes tarefas:

- Formulação do conceito inicial (percepção), a partir de informações do paciente (entrevista, anamnese);
- Geração de múltiplas hipóteses diagnósticas: 1 a 5 hipóteses iniciais a partir da base de conhecimento (centrada no dado ou centrada no modelo de desordens);
- Avaliação e regeneração de hipóteses (refinamento das hipóteses iniciais) ;

- Formulação de uma estratégia de avaliação (investigação ou rastreamento);
- Desenvolvimento da síntese do problema;
- Decisão diagnóstica: princípio lógico do diagnóstico e validação diagnóstica (coerência, adequação e parcimônia);
- Decisão terapêutica: o tratamento deve levar em conta o diagnóstico em si (conjunto de desordens que explicam as manifestações), mas também características do paciente (por exemplo, no caso médico, idade, sexo, etc.)

É válido lembrar que embora as tarefas acima tenham sido descritas no domínio de aplicação médica, no contexto deste trabalho estendeu-se o conceito de raciocínio clínico tanto para o raciocínio de um médico, fitopatologista ou veterinário quanto para o raciocínio de um engenheiro ou técnico no diagnóstico e controle de falhas em computadores, satélites, entre outros.

Na Figura 4.1, está apresentado o processo de raciocínio para tomada de decisão diagnóstica e terapêutica, no escopo do presente trabalho. Na Figura 4.1 está representado, através de um fluxograma, o mecanismo de inferência que é utilizado pelo diagnosticador durante o processo de raciocínio clínico onde as tarefas acima estão descritas em um contexto mais geral. Na Figura, os retângulos representam tarefas e os losângos as tarefas de tomada de decisão. No lado esquerdo da Figura 4.1 estão descritas as etapas que agrupam estas tarefas em subconjuntos que aqui serão denominadas: pré-diagnóstico, investigação, diagnóstico e tratamentos.

4.1.2 Pré-diagnóstico

A etapa de pré-diagnóstico abrange as tarefas iniciais do processo de raciocínio clínico. O primeiro elemento na tentativa de solucionar um problema clínico é obter informações relacionadas ao problema (do paciente, da plantação ou do computador) através de entrevistas ou anamnese (termo utilizado em medicina e em fitopatologia) e exames físicos ou inspeção no local (termo usado na área agrícola). Neste momento o diagnosticador coleta, além dos sintomas presentes, outras informações que lhe pareçam importantes na busca de uma solução. Essa é a síntese inicial do problema.

Logo que o diagnosticador completa sua síntese inicial, várias hipóteses iniciais surgem na sua mente. As hipóteses geradas nesta etapa dependem da natureza da síntese inicial e da capacidade do diagnosticador em conceber explicações plausíveis. A base de conhecimento utilizada pelo diagnosticador nesta etapa pode ser dividida em

conhecimento centrado no dado e conhecimento centrado na doença segundo Réa-Neto (1998). O conhecimento centrado no dado capacita o diagnosticador a avaliar um sintoma, ou um sinal, ou um resultado laboratorial em particular. Com esse conhecimento, quando um determinado dado ou sintoma é observado, suas possíveis causas são lembradas e avaliadas. O conhecimento centrado na doença permite o diagnosticador conhecer as manifestações clínicas que, tipicamente, caracterizam uma desordem (como por exemplo, descrito nos livros textos).

Devido à grande incerteza que caracteriza esta fase inicial do processo, as hipóteses têm uma função primordial: elas estruturam o problema clínico, restringem o número de explicações possíveis, limitam as ações na busca da solução para o problema e fornecem uma base para as expectativas. Essas expectativas são predições de achados clínicos baseados no modelo mental das desordens consideradas plausíveis. Por isso, a representação mental que o diagnosticador tem das desordens ou doenças é um fator crítico na eficiência do processo de solução dos problemas clínicos.

Cada hipótese diagnóstica evoca um modelo com o qual as manifestações clínicas observadas podem ser comparadas. Dentro deste contexto, também novos dados podem ser coletados e avaliados, preservando e refinando a hipótese ou rejeitando-a.

No início do processo de avaliação das hipóteses, quando apenas um pequeno número de dados clínicos significativos estão disponíveis, as hipóteses tendem a ser mais numerosas e abertas. Nesse estágio a entropia diagnóstica (incerteza) é alta e a diferenciação entre as hipóteses é pequena.

Assim, após ter construído a síntese inicial e várias hipóteses terem surgido na cabeça do diagnosticador, é necessário iniciar um processo de avaliação ou investigação das hipóteses (advindas de entrevistas, de inspeção, do laboratório ou de procedimentos) para estabelecer uma hipótese apropriada. Neste ponto, inicia-se a etapa de investigação.

4.1.3 Investigação

Durante o processo de investigação, novas hipóteses podem ser geradas e novas estratégias elaboradas em direção à decisão diagnóstica e terapêutica. Quando uma nova informação relevante é obtida, positiva ou negativa, ela deve ser adicionada à síntese do problema. Essa adição de um novo dado ao conceito prévio aumenta e modifica o conteúdo significativo do problema clínico. Com um processo contínuo e cíclico de raciocínio, a adição de novas manifestações ou sintomas transformam o conjunto de hipóteses plausíveis.

4.1.4 Diagnóstico

A síntese do problema é a primeira tarefa da terceira etapa denominada diagnóstico. Mesmo que a síntese do problema seja muito sugestiva de um diagnóstico, a hipótese desse diagnóstico é somente um rótulo conveniente. A síntese do problema, neste momento, é a verdadeira representação do problema (do paciente, da plantação, etc.).

Após estabelecer a síntese do problema, o diagnosticador deve decidir qual das hipóteses ativas tem maior poder explicativo para solucionar o problema clínico. Para se chegar a essa decisão, o diagnosticador avalia se a síntese do problema se encaixa em uma das hipóteses ativas. Esse “encaixe” ocorre quando o caso apresenta um número suficiente de achados positivos e negativos esperados em uma determinada hipótese diagnóstica, para dar ao diagnosticador a segurança na hipótese que explica o problema do paciente.

Antes que um diagnóstico seja aceito como base para uma ação (prognóstica ou terapêutica), ele deve ser submetido a uma avaliação de sua validade. Esse processo de verificação da validade diagnóstica compõe-se de uma comparação entre os achados clínicos (presentes e ausentes) e a(s) desordem(s) suspeitas. Quando o problema clínico é idêntico a uma entidade clínica conhecida, pouca ou nenhuma investigação diagnóstica futura, geralmente, é necessária e uma ação pode ser tomada. Quando alguma manifestação difere do padrão clínico conhecido, uma decisão sobre se a manifestação é meramente uma variação clínica ou se ela invalida o diagnóstico deve ser feita pelo diagnosticador. Para diminuir as possibilidades de erro, o diagnosticador deve buscar, então, uma validação do seu diagnóstico.

Este teste de validade envolve avaliar a coerência, adequação e parcimônia de cada hipótese. A coerência busca uma consistência entre as manifestações e o modelo da desordem hipotética. A adequação requer uma hipótese que explique todos os achados clínicos normais e anormais do problema. A parcimônia é a procura da hipótese mais simples para explicar todos os achados clínicos. Uma vez feita a decisão diagnóstica, o diagnosticador deve selecionar o tratamento apropriado dando, início à última etapa da Figura 4.1.

4.1.5 Prognóstico e Tratamento

Conforme está descrito em Réa-Neto (1998), apesar dos esforços da medicina moderna em procurar estabelecer o melhor tratamento para cada doença em particular, a escolha terapêutica é influenciada por diversos fatores, tais como:

- condições clínicas do paciente;

- presença de doenças intercorrentes;
- possíveis complicações;
- riscos terapêuticos;
- disponibilidade de recursos;
- custos;
- experiência do médico.

Portanto, antes de finalizar sua decisão quanto ao tratamento, o diagnosticador deve tentar estreitar suas possíveis decisões terapêuticas, que são lançadas e testadas de modo semelhante a avaliação das hipóteses diagnósticas. Durante o processo de decisão terapêutica duas perguntas são úteis.

- Qual o objetivo do tratamento: cura, correção do estado patológico, alívio dos sintomas, prevenção de complicações, prolongamento da vida?
- Qual é o grau de efetividade esperada: prolongar a vida (e neste caso, por quanto tempo), aliviar totalmente ou parcialmente os sintomas ? Neste caso, qual a base de conhecimento que permite esperar tal efetividade, os estudos que mostraram efetividade foram feitos em pacientes semelhantes ao seu?

Além das questões anteriores, relacionadas com o benefício potencial da escolha terapêutica, o diagnosticador também precisa levar em consideração seus custos e riscos.

Teoricamente, a decisão terapêutica com menores custos e riscos e com maiores benefícios deve ser escolhida. Muitas vezes, isso não é tão simples e o diagnosticador tem de avaliar se benefícios adicionais compensam maiores custos e riscos. Ainda, respeitadas outras considerações, quando a eficácia do tratamento disponível para uma dada condição clínica é baixa ou o risco do tratamento é alto, esse tratamento só deve ser dado se a probabilidade da desordem for alta. Se o risco do tratamento é insignificante e sua eficácia é muito grande, a decisão pode ser iniciar o tratamento mesmo quando a probabilidade da desordem não é muito alta. Nestes casos, a experiência prévia do diagnosticador é essencial para a escolha.

Segundo Réa-Neto (1998), em complementação à decisão terapêutica deve-se ter um plano educacional do paciente e um procedimento de monitoramento (efeitos do tratamento na progressão da desordem). Os resultados do monitoramento devem modificar constantemente a síntese do problema como um processo cíclico e dinâmico.

4.2 Abordagens Adotadas na Literatura para o Problema

Conforme apresentado na seção 4.1, o sucesso de um problema clínico depende de 2 fatores: diagnóstico correto e tratamento efetivo. Além disso, quando a tarefa de investigação (ex., obtenção de exames complementares) é bem feita, obtém-se maior eficácia e eficiência no diagnóstico além de redução de custos e riscos.

No capítulo 2 foi apresentado um breve resumo dos principais modelos computacionais mais utilizados para suportar o processo de raciocínio e tomada de decisão, bem como a modelagem de incertezas, durante o processo de decisão diagnóstica e terapêutica sob o ponto de vista de Inteligência Artificial (IA) e Teoria da Decisão (TD). Os mecanismos de inferência utilizados na etapa de diagnóstico tem sido mais pesquisados na área de IA enquanto os mecanismos de tomada de decisão têm sido estudados mais detalhadamente na área de Teoria de Decisão.

Na década de 80 foram desenvolvidos vários sistemas em IA para automatizar a etapa de diagnóstico. Uma das características em comum destes sistemas é a de se comparar um modelo estabelecido para cada desordem com os dados do caso particular. Usualmente, estes sistemas são baseados em regras, onde as manifestações aparecem nas premissas e as causas nas conclusões (i.e. *efeitos* \rightarrow *causas*). Em consequência, o raciocínio adotado para a inferência nestes sistemas é do tipo dedutivo; a partir das regras (*efeitos* \rightarrow *causas*) e fatos (*efeitos*), deduzem-se novas informações (*causas*). Como as causas dificilmente podem ser deduzidas com certeza a partir de um conjunto de manifestações, estes sistemas geralmente adotam algum modelo de representação de incerteza para permitir ao especialista exprimir sua confiança regra.

O MYCIN (Shortliffe, 1976), o sistema mais representativo desta abordagem, foi criado para o diagnóstico e tratamento de doenças infecciosas no sangue. As regras são modeladas a partir das manifestações para as desordens (i.e, *efeitos* \rightarrow *causas*) e das desordens para os tratamentos. Além disso, um coeficiente de certeza é associado a cada regra.

Os sistemas baseados nesta abordagem tiveram um grande sucesso na época de seu surgimento, e geraram uma grande expectativa sobre a aplicação deste tipo de abordagem em qualquer área de conhecimento. No entanto, muitos desenvolvedores de aplicações se frustraram na tentativa de modelar o conhecimento de seus especialistas no formato adotado por estes sistemas, e o interesse neste tipo de abordagem diminuiu, tendo como consequência levado a uma diminuição de interesse na área de IA em si, tanto no mundo acadêmico quanto nos outros setores.

Pode-se argumentar que o problema principal com este tipo de abordagem é que o conhecimento do tipo *efeitos* \rightarrow *causas* é contrário à maneira com que os fenômenos ocorrem na natureza (*causas* \rightarrow *efeitos*). Embora seja possível a criação de regras que concluem uma causa plausível para um conjunto de efeitos, os especialistas nem sempre conseguem ser precisos e raramente tem certeza absoluta sobre a informação que fornecem.

O conhecimento modelado através de regras do tipo *causas* \rightarrow *efeitos* é pois mais fácil de ser obtido, mesmo que algum tipo de incerteza/imprecisão permeie este conhecimento. Na década de 90, o conhecimento na forma *causas* \rightarrow *efeitos* passou a ser modelado, com o conseqüente uso de raciocínio do tipo abdutivo; a partir das associações (*causas* \rightarrow *efeitos*) e fatos (*efeitos*), criam-se novas hipóteses (*causas*).

A Teoria das Coberturas Parcimoniosas (TCP) (Peng e Reggia, 1990), conforme descrito no capítulo 2, representou uma primeira tentativa para formalizar o raciocínio causal abdutivo de diagnóstico. Embora a TCP original apresentasse como vantagem em relação à abordagem anterior, a capacidade de modelar regras do tipo *causas* \rightarrow *efeitos* e a ausência de um modelamento de incerteza faziam com que a solução de um problema tendesse a ter muitas explicações alternativas. Uma versão posterior do sistema, permitiu o uso de informação probabilística na relação causal, possibilitando reduzir o tamanho das soluções do problema de diagnóstico, além de ordená-las, como apresentado em Peng e Reggia (1990).

No entanto, a modelagem da imperfeição da informação fornecida por especialistas não deveria necessariamente se limitar ao modelo de representação probabilista. Por exemplo, especialistas muitas vezes quantificam o quanto uma desordem está associada a uma desordem de uma forma simbólica não-numérica ("muito comum", "moderado", "muito raro", etc.), termos estes que podem ser tratados como tal, no que se chama atualmente de "*computing with words*", ou traduzidos em números, usando seja probabilidade, seja outros modelos alternativos, como teoria de possibilidades, teoria de Dempster-Shafer ou fatores de certeza.

Desde o advento da TCP, vários autores têm investigado teorias da diagnose alternativas, ou extensões/modificações da TCP. Lucas (1996) permite representar vários tipos de interações que não são possíveis na TCP original. Dubois e Prade (1995) propõem uma extensão possibilista à TCP para modelar as incertezas. Wainer e Rezende (1997) propõem uma extensão à TCP de modo que informações do tipo categórica e temporal possam ser incorporadas na base de conhecimento.

Conforme descrito no capítulo 2, o trabalho de Wainer e Sandri (1998) mostra que é melhor representar informação temporal como intervalos nebulosos. Entretanto, este trabalho se limita em verificar como uma desordem única cobre um conjunto de manifestações, ao invés de construir uma teoria de diagnóstico completa. O modelo proposto em Wainer e Rezende (1997) tem ainda algumas outras limitações, descritas a seguir.

- A representação da informação categórica se limita a classificar uma dada manifestação como efeito necessário ou possível de uma dada desordem. Por exemplo, não é possível distinguir-se uma informação categórica do tipo “na doença d_l a manifestação m_i raramente ocorre” de outra do tipo “na doença d_l a manifestação m_i frequentemente ocorre”. Em ambos os casos, a manifestação é tratada como simplesmente possível, não permitindo assim um melhor aproveitamento da informação disponível.
- As condições favoráveis para evolução de uma desordem não são contempladas pelo modelo. Em fitopatologia, por exemplo, temperatura, umidade bem como a época do aparecimento dos sintomas podem favorecer o desenvolvimento de uma determinada desordem d_i . No diagnóstico de satélites, temperatura e níveis de radiação podem favorecer que determinadas falhas aconteçam.
- O modelo não identifica quais as manifestações prioritárias a serem investigadas para completar um diagnóstico. No domínio agrícola, isto pode ser vital, inclusive podendo vir a comprometer toda uma produção ou um replanejamento agrícola.
- O modelo trata da intensidade com que uma manifestação ocorre mas não trata da incerteza em si de uma determinada manifestação ter ocorrido ou não.
- Um ponto crucial no processo de raciocínio clínico que não é suportado pelo modelo é o mapeamento de diagnósticos em tratamentos.

Os trabalhos citados acima têm apresentado soluções parciais na área de diagnóstico e não uma solução completa para o problema de raciocínio clínico. As tecnologias oferecidas em TD (matrizes de ‘*payoffs*’, árvores de decisão, diagramas de influência entre outros), por sua vez, propõem soluções para as tarefas de tomada de decisão nas áreas de investigação e tratamentos. Entretanto, cada vez mais nota-se a necessidade de integração das tecnologias desenvolvidas nas áreas de IA e TD, embora sejam baseadas em princípios diferentes, para suportar as etapas de diagnóstico, investigação e tratamentos de um processo de raciocínio clínico.

Resumidamente, as tecnologias oferecidas pelas áreas de IA e TD visam suportar o raciocínio humano e a tomada de decisão pela formalização e automatização do conhecimento do especialista do domínio. Apesar de objetivos comuns, estes dois paradigmas têm divergências e diferenças fundamentais nos princípios e na prática. Conseqüentemente, a integração das tecnologias desenvolvidas em cada uma destas áreas não é trivial.

No presente trabalho é apresentada uma nova abordagem para tomada de decisão no processo de raciocínio clínico que engloba diagnóstico, investigação e tratamentos que será descrita a seguir.

4.3 Abordagem Integrada para Diagnóstico e Tratamento

Neste trabalho, propõe-se uma teoria integrada para diagnóstico, investigação e tratamento, denominada Teoria das Coberturas Nebulosas (TCN), que utiliza conceitos de inferência causal abdutiva, lógica nebulosa e teoria de decisão, conforme apresentado na Figura 4.2.

A TCN contempla a possibilidade de se ter um conjunto de desordens para explicar as manifestações de um caso particular, e não somente uma desordem, como proposto em Wainer e Sandri (1998). Além da classificação de hipóteses mais plausíveis para explicar um conjunto de manifestações, a TCN também classifica os exames prioritários para completar um diagnóstico e, dado um diagnóstico, classifica os tratamentos mais efetivos para ele.

Na Figura 4.3 está apresentada a estrutura geral da TCN. A arquitetura proposta para suportar o modelo apresentado acima é composta por 3 módulos principais conforme apresentado na Figura 4.3. Em todos os 3 módulos da TCN, o conhecimento envolvido é do tipo causal e todo o raciocínio é abdutivo. Na tarefa de diagnóstico, as associações que modelam o conhecimento são do tipo *desordem* \rightarrow *manifestações*, que, como visto anteriormente, são mais naturais que a associação reversa, o que facilita a interação com o especialista e que garante resultados mais confiáveis. Como o conhecimento sobre o caso é dado pelas manifestações presentes (ou ausentes) o raciocínio é naturalmente abdutivo.

Na tarefa de investigação, as associações são do tipo *exames* \rightarrow *manifestações*, pois cada exame é criado para verificar um conjunto de manifestações. Como o conhecimento sobre o caso é dado pelas manifestações cuja presença se pretende verificar, o raciocínio é também abdutivo.

Finalmente, na tarefa de tratamento, as associações são do tipo *tratamentos* \rightarrow *desordens* pois cada tratamento é criado visando ser útil para um conjunto de desordens (ou manifestações causadas por desordens). Como o conhecimento sobre o caso é dado pelas desordens presentes no caso, o raciocínio é mais uma vez abduutivo.

A etapa de pré-diagnóstico apresentada na seção 4.1, geralmente utilizada por diagnosticadores humanos, envolve dois aspectos principais: o levantamento de informações pertinentes ao problema em questão e a criação de um subconjunto de hipóteses iniciais que, aparentemente, são obtidas através da utilização de regras intuitivas do tipo *manifestações* \rightarrow *desordens*. A criação de um subconjunto de hipóteses inicial pode ser extremamente útil para um diagnosticador humano, que assim não terá a necessidade de comparar os dados de um caso com os modelos de todas as desordens possíveis, ganhando desta forma tempo e evitando confusões. Esta heurística humana de focalização, no entanto, não é necessária em um sistema computacional no qual todo o conhecimento está modelado na forma de associações causais. Isto porque não existe restrição efetiva para um sistema varrer toda a base de conhecimento, comparando todos os modelos com os dados do caso. Na TCN, por exemplo, ao contrário do diagnosticador humano, o tempo dispendido nesta tarefa é pouco relevante e o desempenho do sistema não se degrada com o aumento de hipóteses. Por estas razões, a fase de pré-diagnóstico na TCN se resume ao levantamento de informações relevantes.

Conforme apresentado na Figura 4.3, o ciclo inicia com a obtenção das manifestações presentes ou ausentes coletada através de entrevistas (ou anamnese), exames físicos ou inspeção no local na fase de pré-diagnóstico. A partir das manifestações presentes e ausentes com suas possíveis informações temporais, o módulo de Diagnóstico evoca as causas potenciais através da inferência abduativa, isto é, seleciona as desordens das manifestações presentes e ausentes a partir da base de conhecimento.

É importante lembrar que este módulo deve suportar tanto a associação gradual entre desordens e manifestações na base de conhecimento como a incerteza dos fatos observados em um caso específico. Em seguida, são calculados os índices de consistência temporal, categórica e intensidade para as desordens evocadas no módulo de Diagnóstico. As informações nebulosas da base de conhecimento, das manifestações presentes (M^+) com seus intervalos temporais nebulosos e das manifestações ausentes (M^-) são utilizadas para calcular estes índices. O conjunto de desordens inconsistentes é atualizado por este módulo. As hipóteses evocadas que não se tornaram inconsistentes são incorporadas às hipóteses resultantes. As hipóteses geradas por este módulo são consistentes com a informação nebulosa do caso particular apresentado.

A partir de um conjunto de observações e das desordens mais plausíveis, utilizando o mecanismo de inferência abdutiva, um conjunto de exames laboratoriais são selecionados de modo a investigar melhor os sintomas presentes para confirmar as hipóteses. Nesta etapa é fundamental a seleção de um ou mais testes de laboratório que contribuam para diminuir a entropia inicial e aumentar a diferenciação entre as hipóteses. Para tal, no escopo deste trabalho é proposta uma função para seleção das manifestações prioritárias a serem investigadas que é implementada no módulo de Investigação.

Como se trata de um processo contínuo e cíclico de raciocínio, o processo se repete enquanto existir manifestações presentes ou ausentes para serem investigadas. Ao final de todas as manifestações, o sistema fornecerá todas as explicações consistentes com o caso.

Antes que um diagnóstico seja aceito como base para uma ação (prognóstica ou terapêutica), ele deve ser submetido a uma avaliação de sua validade. O resultado final do processo de validação diagnóstica deve englobar dois objetivos principais como descrito em Peng e Reggia (1990):

- objetivo de cobertura: o objetivo de explicar todas as manifestações que estão presentes com coerência e adequação.
- objetivo de parcimônia: o objetivo de minimizar a complexidade da explicação.

Estes objetivos são suportados pela abordagem abdutiva, através do formalismo da TCP e de suas extensões apresentadas no capítulo 3.

Neste trabalho são propostas algumas extensões ao modelo de diagnóstico de Wainer e Sandri (1998) que propõem a incorporação de conjuntos nebulosos em alguns conceitos de TCP temporal. Uma das funcionalidades incorporadas à TCN, refere-se a informação categórica nebulosa onde a relação causal entre as desordens e manifestações é representada por uma distribuição de causalidades.

Uma outra funcionalidade da TCN é a incorporação de um índice de consistência relativo às condições favoráveis à existência de uma desordem. Por exemplo, em fitopatologia existem algumas condições favoráveis de temperatura, umidade e época do ano que podem influenciar o desenvolvimento de uma desordem. No diagnóstico médico, sexo, idade e hereditariedade são condições que podem influenciar no desenvolvimento de uma doença. No diagnóstico de falhas em satélite existem algumas condições favoráveis

para determinadas falhas como temperatura, níveis de radiação e época do ano. Adicionalmente, no escopo da TCN é proposta uma combinação dos vários índices de consistência (temporal, de intensidade) de modo que possam ser usados na classificação e redução de hipóteses.

Uma vez feita a decisão diagnóstica, conforme apresentado na Figura 4.3, o diagnosticador deve executar a tarefa de selecionar o tratamento mais apropriado.

Quanto ao módulo de Tratamento é utilizado para extrair informação da base de conhecimento sobre as condições favoráveis para evolução de uma desordem (temperatura, umidade) de modo auxiliar na tomada de decisão das medidas de controle. Mais uma vez, o conhecimento do diagnosticador pode ser representado pelo modelo abdutivo, isto é, a partir dos tratamentos para as desordens (i.e, *tratamentos* \rightarrow *desordens*).

No contexto deste trabalho, é proposto um índice de consistência relativo às condições favoráveis à eficácia de um tratamento que devem ser levadas em conta durante uma decisão terapêutica além de outras variáveis como CUSTO e RISCO. Por exemplo, em fitopatologia condições de temperatura, umidade, severidade e época do aparecimento dos sintomas podem influenciar na eficácia do tratamento. Além disso, em uma decisão terapêutica para plantas devem ser considerados o CUSTO da aplicação do tratamento e o RISCO na produtividade da colheita. Em um outro exemplo, no diagnóstico de falhas em satélite a temperatura e os níveis de radiação podem interferir na ação corretiva a ser tomada para uma determinada falha no satélite. Adicionalmente, deve-se levar em conta o CUSTO e o RISCO da ação corretiva para a missão do satélite.

Conforme apresentado na seção 4.1, em complementação a uma decisão terapêutica deve-se ter um procedimento de monitoramento dos efeitos do tratamento na progressão da desordem. Os resultados do monitoramento devem modificar constantemente a síntese do problema como um processo cíclico e dinâmico. Entretanto, a automatização do procedimento de monitoramento não está incluído no escopo deste trabalho.

Uma síntese do processo de tomada de decisão diagnóstica e terapêutica em fitopatologia suportado pela TCN está representada na Figura 4.4.

Os nós $m_1, m_2, ..m_n$ representam as variáveis de observação que correspondem às manifestações causadas por uma desordem d_i . Os arcos representam as influências das relações causais incertas. A relação causal $R(d_i, m_j)$ entre desordens e manifestações pode ser observada no diagrama assim como a relação causal $R(A_c, d_i)$ onde o nó (A_c) é o agente que causa a desordem d_i . A relação entre tratamentos e desordens está expressa

na relação $Z(t_k, d_i)$. Os retângulos representam as variáveis de decisão de *tratamentos* e *testes laboratoriais*. A relação entre um teste e uma desordem é expressa através da confirmação de um resultado de teste que confirma se o agente causal da desordem está presente. Caso seja confirmada a hipótese da desordem, um tratamento é selecionado considerando algumas outras variáveis de observação, tais como, severidade, época ou idade, condições favoráveis e custo. O nó losângulo representa a *medida de valor* calculada para esta tomada de decisão.

É válido lembrar que as incertezas nas relações causais acima podem ser representadas por modelos probabilísticos. Muitas vezes, no entanto, é difícil obter informações quantitativas ou estatísticas dos diagnosticadores. No escopo deste trabalho, a teoria de possibilidades é usada como alternativa para a representação de incertezas.

Para dar suporte à teoria acima e obter uma solução para o problema raciocínio clínico, o modelo TCN contempla algumas características principais:

- um mecanismo de inferência abdutivo para representar o conhecimento do diagnosticador;
- um mecanismo de representação de incertezas inerentes ao processo de raciocínio clínico;
- um modelo para classificação de hipóteses em uma situação prática de diagnóstico que pode ser incorporado através de uma generalização do modelo TCP temporal-categórico nebuloso proposto por Wainer e Sandri (1998) de modo a combinar os vários índices de consistência visando usá-los na classificação de hipóteses em uma situação prática de diagnóstico;
- uma ou mais funções de utilidade associadas a outros conceitos de teoria da decisão de modo a auxiliar na seleção e classificação de manifestações prioritárias para investigação;
- um modelo para identificar casos onde a associação entre desordens e manifestações são incompletas;
- um mecanismo de identificação de condições favoráveis à evolução de uma desordem visando previsões futuras e tomada de decisões para tratamento.

Na próxima seção é apresentado um ambiente abdutivo nebuloso proposto para dar suporte às funcionalidades acima.

4.4 Formalização do Problema

O ambiente abdutivo genérico proposto no escopo da TCN para representar o conhecimento do diagnosticador está apresentado na Figura 4.5.

O conhecimento utilizado na TCN é representado como uma rede associativa de relações binárias. Uma rede associativa consiste de nós, usualmente representando entidades como objetos, conceitos e eventos, e arcos representando associações entre nós. O significativo poder das redes associativas as torna particularmente adequadas para representação das associações causais em diversas aplicações de raciocínio abdutivo.

Os nós da Figura 4.5 representam os diversos tratamentos, desordens, manifestações e exames. Os tratamentos (conjunto T) estão relacionados com as desordens (conjunto D) que eles podem tratar ou controlar. As desordens (conjunto D) estão relacionadas com os efeitos ou as manifestações que eles podem causar. Por outro lado, cada manifestação tem associada os exames ou testes laboratoriais (conjunto E) que podem confirmar sua existência ou confirmar o agente causal que as ocasionou.

A associação causal $R(A_c, d_i)$ está implícita em d_i , isto é, esta relação não está explícita na Figura 4.5. Por exemplo, no domínio de doenças de milho existe uma desordem denominada 'Mancha por Cercospora ou Cercosporiose' cujo agente causal é o fungo 'Cercospora zea-maydis'. Portanto, quando é inferido que a plantação está com 'Cercosporiose' implicitamente infere-se que a plantação está contaminada pelo fungo 'Cercospora zea-maydis'. Similarmente, quando em medicina um paciente está com uma desordem d_i infere-se que ele está contaminado por algum agente causal (fungo, bactéria, etc.).

Da mesma maneira, o conhecimento sobre as desordens e os efeitos que elas causam pode ser representado por uma teoria criada para suportar a relação causal $R(d_i, m_j)$. A partir de um conjunto de observações em um caso particular (um paciente, uma plantação ou um satélite), utilizando-se o mecanismo de inferência abdutiva, pode-se obter as desordens mais plausíveis que explicam aquelas observações. A partir deste conjunto de observações e das desordens mais plausíveis, também utilizando o mecanismo de inferência abdutiva, um conjunto de exames laboratoriais são selecionados de modo a investigar melhor os sintomas presentes para confirmar as hipóteses utilizando a relação binária $I(e_k, m_j)$.

Similarmente, os tratamentos (fungicidas, remédios, determinado comando por telemetria) disponíveis para determinadas desordens ou falhas em um domínio específico (doenças de milho, doenças cardíacas, falhas em satélite, etc ...) podem ser representados

por uma teoria que suporte esta relação $Z(t_k, d_i)$. A partir de um conjunto de hipóteses (um conjunto de desordens), utilizando-se o mecanismo de inferência abductiva, pode-se obter os tratamentos mais plausíveis que explicam aquelas hipóteses. Portanto, o mecanismo de inferência abductiva pode ser utilizado tanto para tomada de decisão diagnóstica quanto terapêutica.

A partir das idéias acima, define-se formalmente o problema de raciocínio clínico. As definições apresentadas neste capítulo constituem a base teórica para esta nova abordagem.

Definição 4.1 *Um problema de raciocínio clínico P para diagnóstico, investigação e tratamentos é uma quadrúpla $P = \langle KB_d, KB_e, KB_t, CA \rangle$ em que:*

- *KB_d é a base de conhecimento para representar informação correspondente às desordens e os efeitos que elas causam;*
- *KB_e é a base de conhecimento para representar informação correspondente aos testes laboratoriais e às desordens ou manifestações que eles identificam;*
- *KB_t é a base de conhecimento para representar informação correspondente aos tratamentos e às desordens que eles controlam;*
- *CA é a especificação de um caso particular (um paciente, uma cultura, um satélite, etc.).*

O conhecimento associativo usado em problemas de diagnóstico, investigação e tratamentos é usualmente muito extenso e complexo. Portanto, neste trabalho o problema de raciocínio clínico é dividido em três subproblemas. Cada subproblema é representado por um ambiente abductivo que contém uma rede associativa com apenas duas entidades e uma relação binária entre elas.

A base de conhecimento para diagnóstico KB_d é representada pela sextupla $\langle D, M, R, G, CC, \theta \rangle$ onde: D é um conjunto de desordens ou causas; M , é o conjunto de efeitos que são observados e cujas causas ou desordens são procuradas como explicações; R é a teoria causal; G corresponde ao grafo que representa a informação temporal das desordens, CC é um conjunto de variáveis que representa as condições favoráveis para existência de uma desordem e θ corresponde a escala de tempo utilizada.

A base de conhecimento para investigação KB_e é representada pela quintupla $\langle E, M, D, I, R \rangle$ onde: E é um conjunto de exames ou testes laboratoriais; M é o conjunto de efeitos observados, I é a teoria que associa o conjunto de efeitos que são observados com os exames que os confirmam, D e R correspondem ao conjunto de desordens e a teoria causal, respectivamente.

A base de conhecimento para tratamento KB_t é representada pela quintupla $\langle T, D, Z, CC, V \rangle$ onde: T é um conjunto de tratamentos; D , é o conjunto de desordens que são gerados como hipóteses; e Z , é a teoria que associa o conjunto de desordens geradas como hipóteses com os tratamentos que as controlam, CC corresponde a certas condições que visam garantir a eficácia de um tratamento e V corresponde as variáveis CUSTO e RISCO.

4.5 Solução Final para o Problema de Raciocínio Clínico

Dado um problema de raciocínio P (4.1), pode-se definir a solução final para o problema de raciocínio clínico como descrito a seguir.

Definição 4.2 (*Solução abdutiva*): Dada uma base de conhecimento $\Theta = \{KB_d, KB_e, KB_t\}$ para o problema de raciocínio clínico P e um caso particular CA , $SOL(P) = \{SOL_d, SOL_e, SOL_t\}$ é uma solução abdutiva para o caso particular CA sse $\Theta \cup SOL(P)$ cobre CA com adequação, coerência e parcimônia e $\Theta \cup SOL(P)$ é consistente.

Nos capítulos 5, 6 e 7, as definições de cada base de conhecimento (KB_d, KB_e, KB_t) bem como de cada solução $\{SOL_d, SOL_e, SOL_t\}$ serão detalhadas de modo a representar os problemas do mundo real nas áreas de diagnóstico, investigação e tratamentos, respectivamente. Para facilitar o entendimento dos próximos capítulos, na Figura 4.6 está apresentada uma síntese do mecanismo de inferência do processo de raciocínio clínico da TCN onde é representada cada etapa do processo e o capítulo correspondente.

Os retângulos da Figura 4.6 correspondem aos módulos de Diagnóstico, Investigação e Tratamentos. A partir de um caso particular CA , a solução SOL_d corresponde a solução abdutiva para o problema de diagnóstico gerada no módulo de Diagnóstico. No escopo da TCN, quando uma solução SOL_d não atinge um certo grau de verossimilhança (i.e, $SOL_d < \Delta$) pode ser assumido que os efeitos causados por cada d_i tal que $d_i \in SOL_d$ devem ser melhor investigados (i.e, através de exames complementares ou testes laboratoriais) visando um refinamento das hipóteses. A partir de SOL_e um conjunto de manifestações são confirmadas ou refutadas, conseqüentemente, atualizando os conjuntos de manifestações de M^+ e M^- que pertencem a CA gerando um novo caso C_{atual} . Dado a base de conhecimento KB_d e o caso particular C_{atual} , infere-se um nova explicação abdutiva SOL_d . Para nova solução SOL_d pode-se inferir os melhores tratamentos, através de um ambiente abdutivo para tratamentos, gerando uma solução abdutiva para Tratamentos denominada SOL_t .

Os exemplos utilizados no decorrer dos próximos capítulos referem-se à diagnose, investigação e tratamento de doenças de plantas, mais especificamente, doenças de milho. Entretanto, espera-se que os desenvolvimentos teóricos obtidos possam ser aplicados para outros problemas de diagnóstico dinâmico, conforme apresentado no capítulo 8.

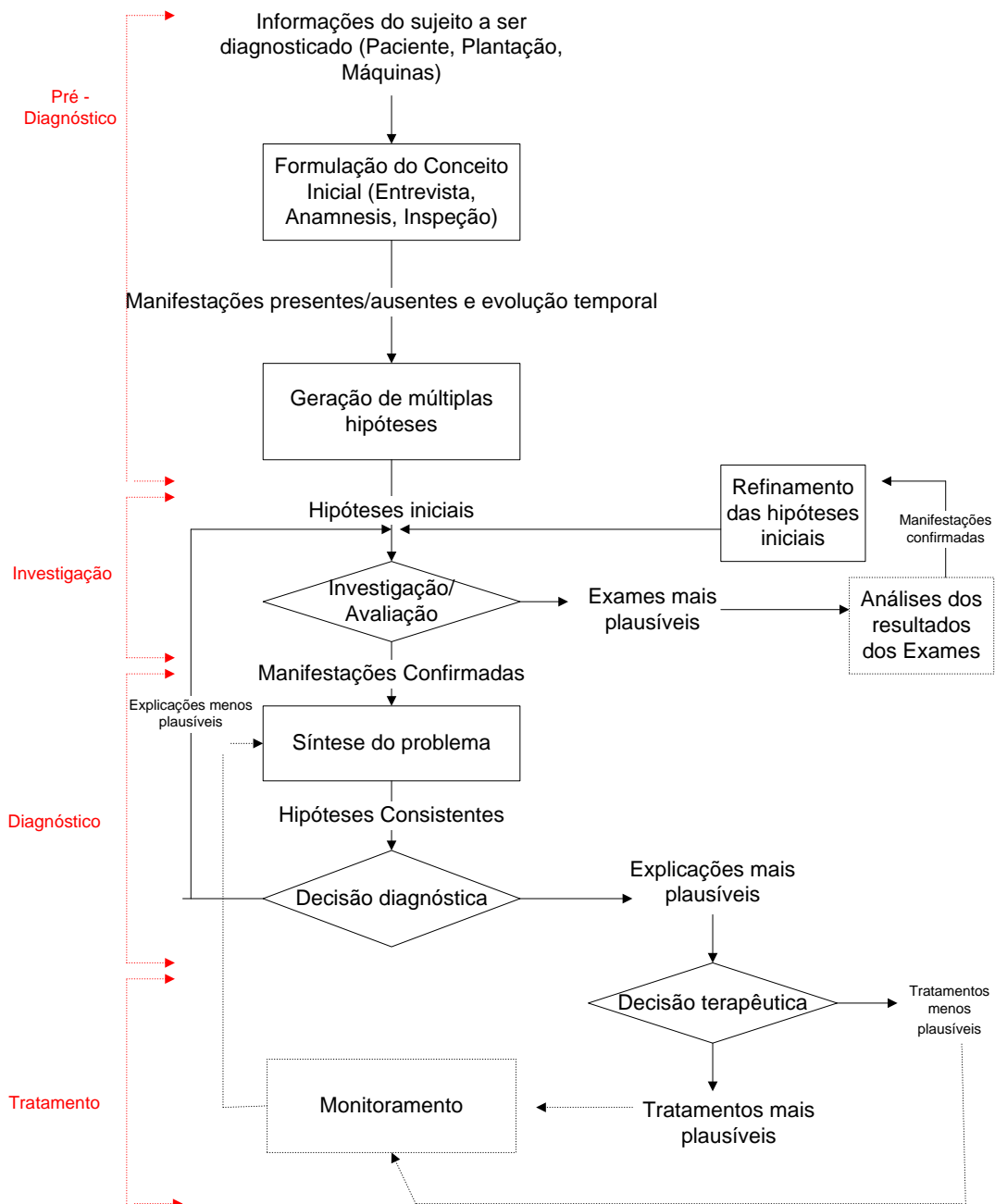


FIGURA 4.1 – As decisões diagnóstica e terapêutica.

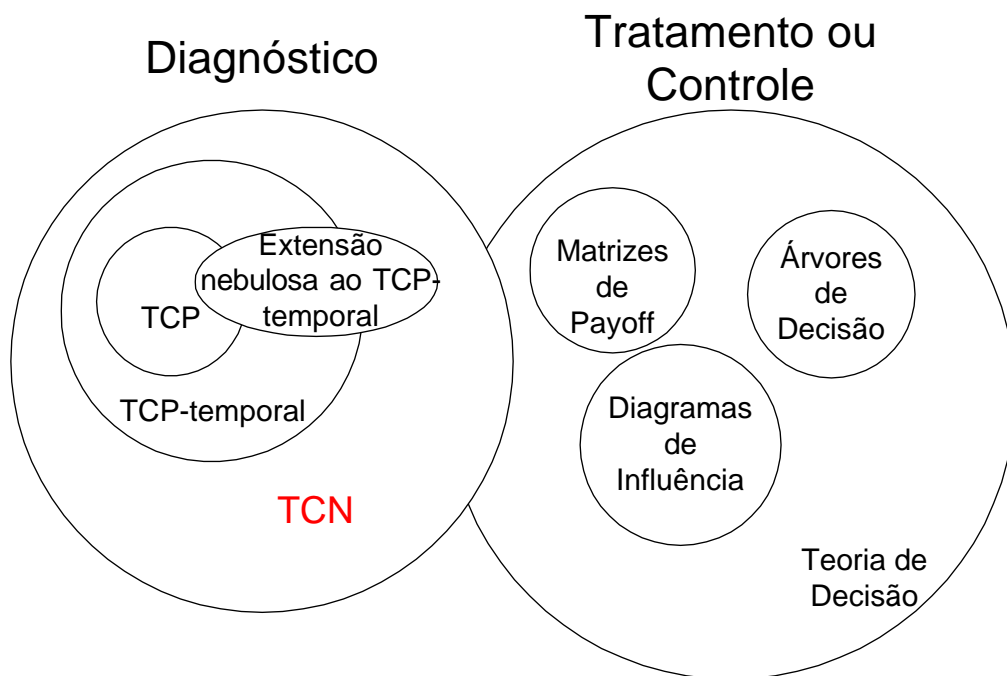


FIGURA 4.2 – Escopo da TCN.

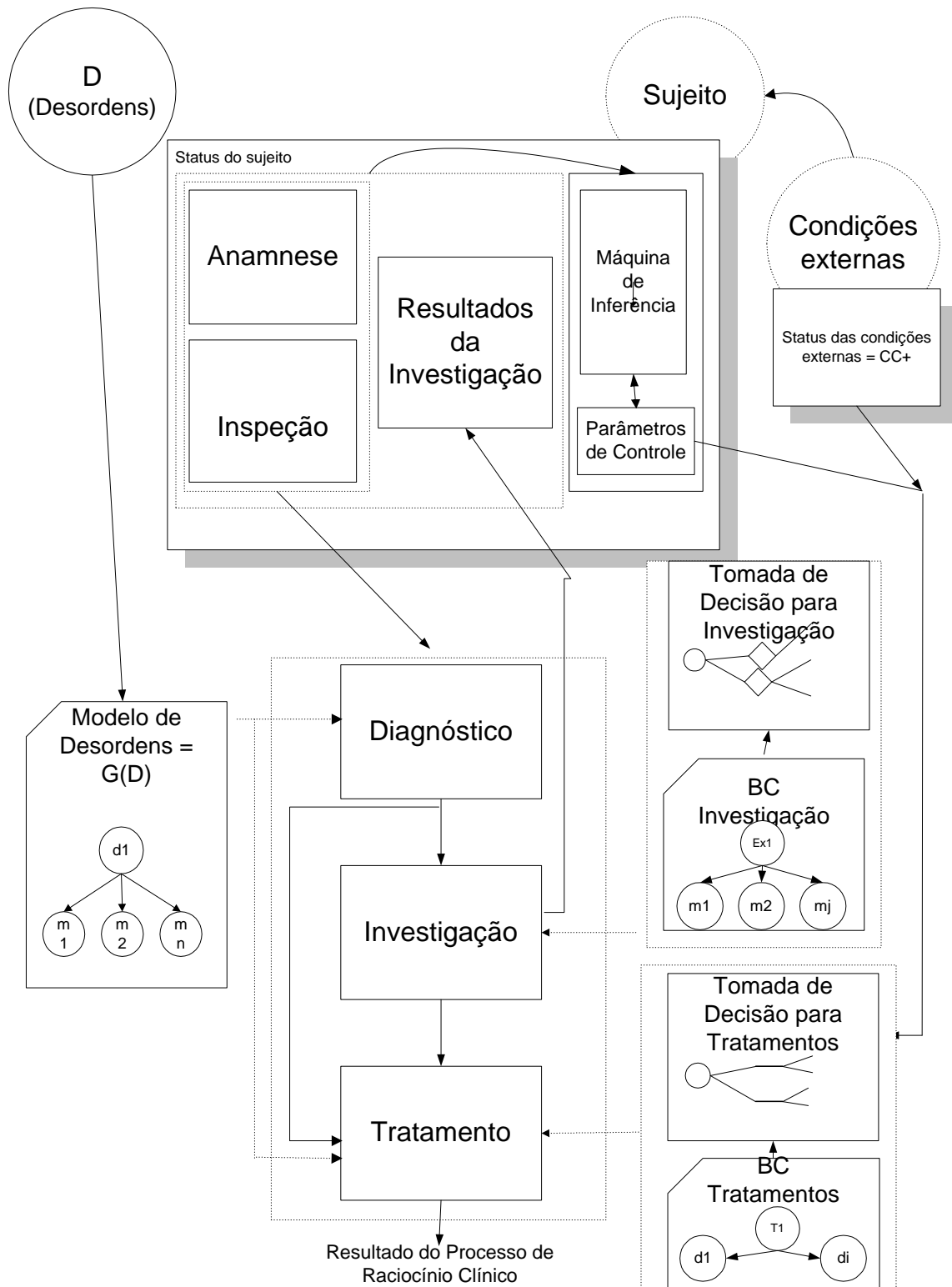


FIGURA 4.3 – Um ambiente integrado para diagnose e tratamento.

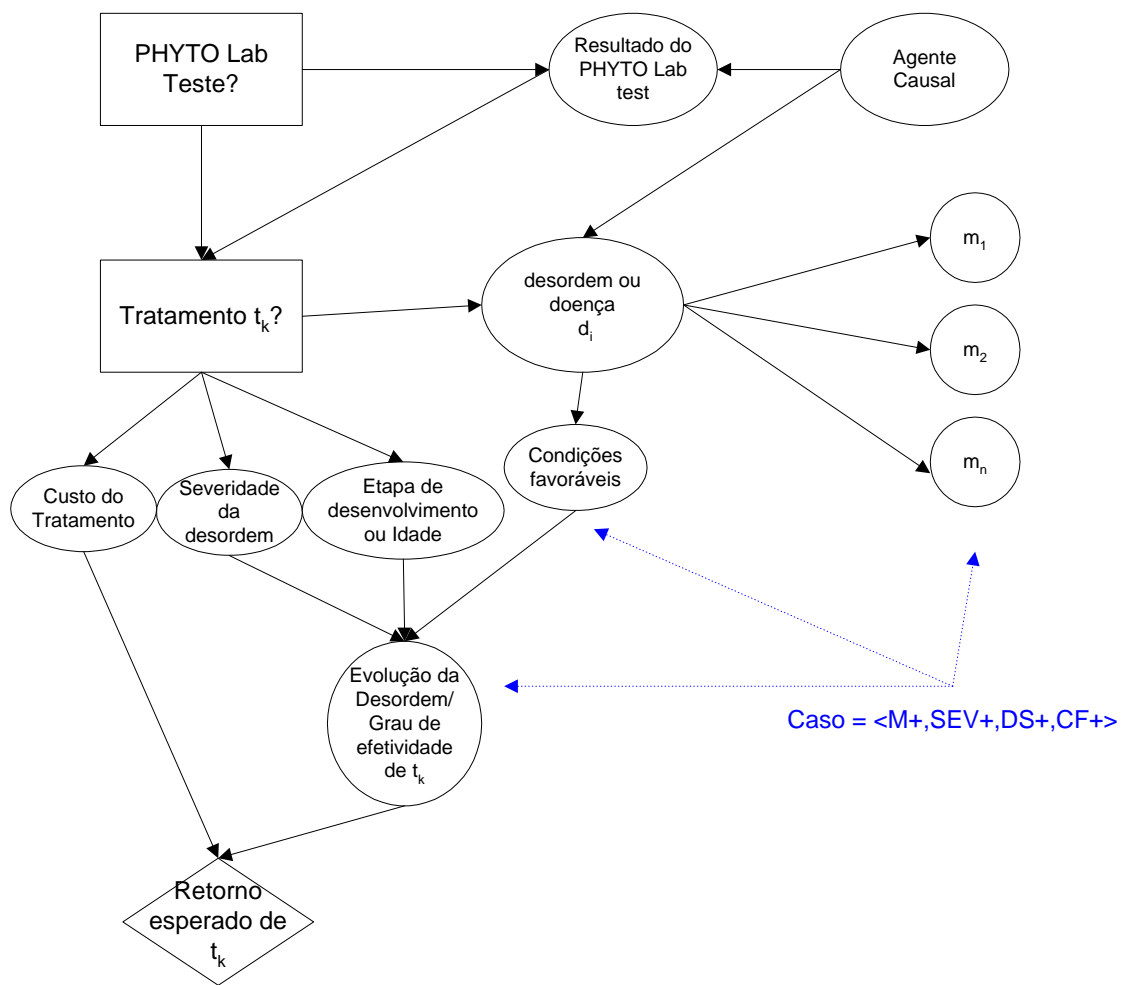
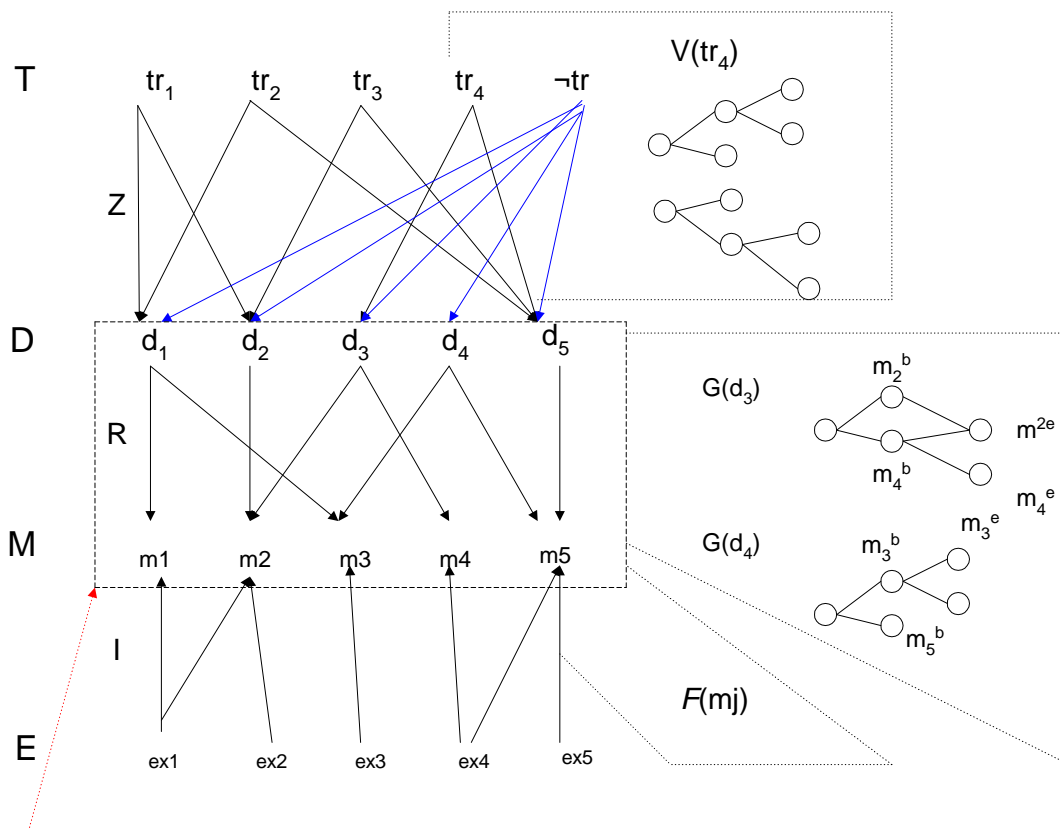


FIGURA 4.4 – Integração dos processos de diagnose e tratamento.



CA=<M+>

FIGURA 4.5 – Um ambiente abdutivo genérico integrado para diagnose e tratamento.

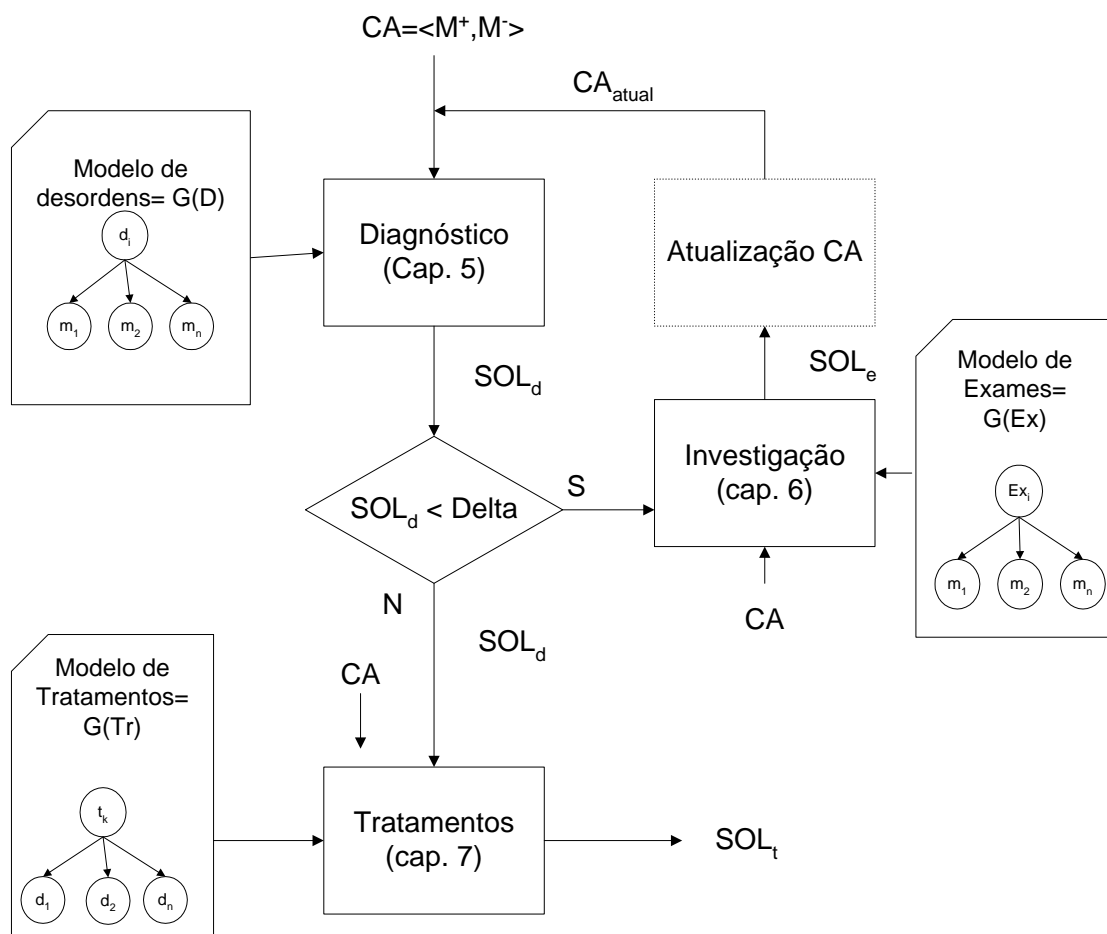


FIGURA 4.6 – Síntese do processo de raciocínio clínico suportado pela TCN.

CAPÍTULO 5

AMBIENTE ABDUTIVO PARA DIAGNÓSTICO

Neste capítulo é apresentado detalhadamente o ambiente abduativo para diagnóstico que incorpora conceitos de inferência abduativa, lógica nebulosa e teoria de decisão no escopo da TCN. Primeiramente, nas seções 5.1 e 5.2 são apresentados detalhadamente os mecanismos para representação do conhecimento do problema de diagnóstico e de um caso particular, bem como das incertezas inerentes a estes problemas, no escopo desta nova abordagem. Na seção 5.3 são apresentados os mecanismos de inferência e tomada de decisão utilizados para diagnóstico.

5.1 Representação do Conhecimento em Diagnóstico

No escopo da TCN, são utilizados dois conjuntos finitos e discretos (veja Figura 5.1) para modelar o subproblema de diagnóstico. O conjunto D representa o conjunto de todas as desordens possíveis d_i que podem ocorrer em um domínio e o conjunto M representa todas as manifestações m_j que podem surgir quando uma ou mais desordens estiverem presentes. Assume-se que D e M não possuem elementos em comum. Para modelar a noção intuitiva de causa utiliza-se uma relação causal R , de D para M , associando individualmente desordens e manifestações.

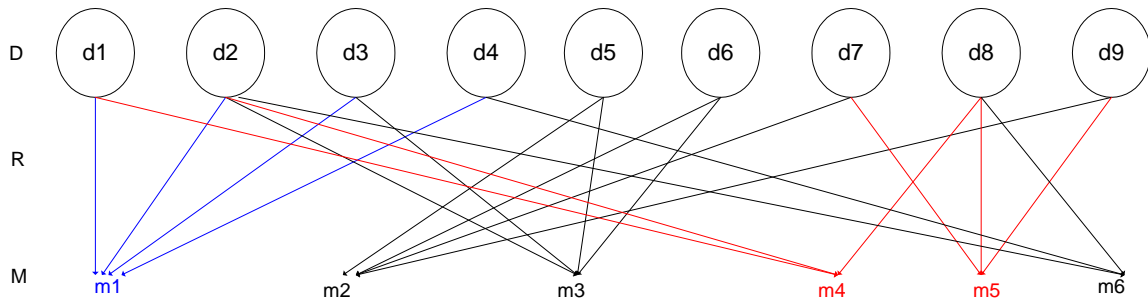


FIGURA 5.1 – Um ambiente abduativo causal para diagnóstico.

Conforme descrito na seção 4.4, o problema de diagnóstico pode ser modelado por um ambiente abduativo causal (D, M, R) onde a relação causal R entre manifestações e desordens corresponde a teoria Σ restrita sintaticamente às teorias causais.

5.1.1 Teoria causal

A representação do ambiente abduativo causal adotada neste trabalho é a TCP apresentada na seção 3.2. No contexto deste trabalho, será utilizado o modelo TCP mais simples

que não utiliza estados intermediários na rede causal conforme apresentado na Figura 5.1.

Uma associação $\langle d_i, m_j \rangle$ em R significa que ' d_i pode diretamente causar m_j '. Em outras palavras, a probabilidade associada a uma associação causal é suposta não ser zero (nenhuma associação causal neste caso) mas algum valor intermediário entre]0, 1]. Este fato reflete a situação do mundo real para vários casos. Em Peng e Reggia (1990) é apresentado um modelo causal probabilístico para representar as noções numéricas de plausibilidade da relação causal R. No escopo da TCN, será utilizado um modelo baseado em teoria de possibilidades e conjuntos nebulosos para manipular as incertezas desta associação causal.

Conforme apresentado na tabela 5.1, os tipos de causalidade existentes entre $\langle d_i, m_j \rangle$ são aqui divididas em cinco classes: ' d_i necessariamente causa m_j ', ' d_i predominantemente causa m_j ', ' d_i possivelmente causa m_j ', ' d_i raramente causa m_j ' e ' d_i não causa m_j ' e estão aqui respectivamente associados aos pesos 1, .5, 0, .5, -1. Estes valores foram derivados a partir da interpretação dos tipos em termos de possibilidades e necessidades através da fórmula:

$$R(d_i, m_j) = \frac{\Pi(A) - \Pi(\sim A) - l_{inferior}}{l_{superior} - l_{inferior}}$$

TABELA 5.1 – Associação causal nebulosa $\langle d_i, m_j \rangle$.

$C = \langle d_i, m_j \rangle$	$\Pi(A)$	$\Pi(\sim A)$	$N(A)$	$N(\sim A)$	p_{ij}	$R(d_i, m_j) = Norm(p_{ij})$
" d_i necessariamente causa m_j "	1.0	0	1.0	0	1.0	$1.0 + -1.0 /2 = 1.0$
" d_i predominantemente causa m_j "	1.0	0.5	0.5	0	0.5	$0.5 + -1.0 /2 = 0.75$
" d_i possivelmente causa m_j "	1.0	1.0	0	0	0	$0 + -1.0 /2 = 0.5$
" d_i raramente causa m_j "	0.5	1.0	0	1.0	-0.5	$-0.5 + -1.0 = 0.25$
" d_i não causa m_j "	0	1.0	0	1.0	-1.0	$-1.0 =+ -1.0 = 0$

Nas colunas 2 e 3 estão calculadas as medidas de possibilidades da associação $\langle d_i, m_j \rangle$ ocorrer ou não, respectivamente. As colunas 4 e 5 representam as medidas de necessidade para cada classe de associação e no caso do evento contrário, respectivamente. Na coluna 6 é calculada a função $p_{ij} = \Pi(A) - \Pi(\sim A) = N(A) - N(\sim A)$ que se refere às diferenças entre as medidas de possibilidades (colunas 2 e 3) e as medidas de necessidade (colunas 4 e 5). Finalmente, na coluna 7 é apresentado o grau nebuloso associado a cada classe de $\langle d_i, m_j \rangle$ que corresponde à normalização de p_{ij} .

Um exemplo de como representar a associação causal nebulosa no escopo da TCN pode ser vista na Figura 5.2 que representa a desordem '*ferrugem comum*(d_3)' (descrita na

seção 1.1.1) e da ‘*ferrugem polysora* (d_4)’ que tem características similares à desordem d_3 .

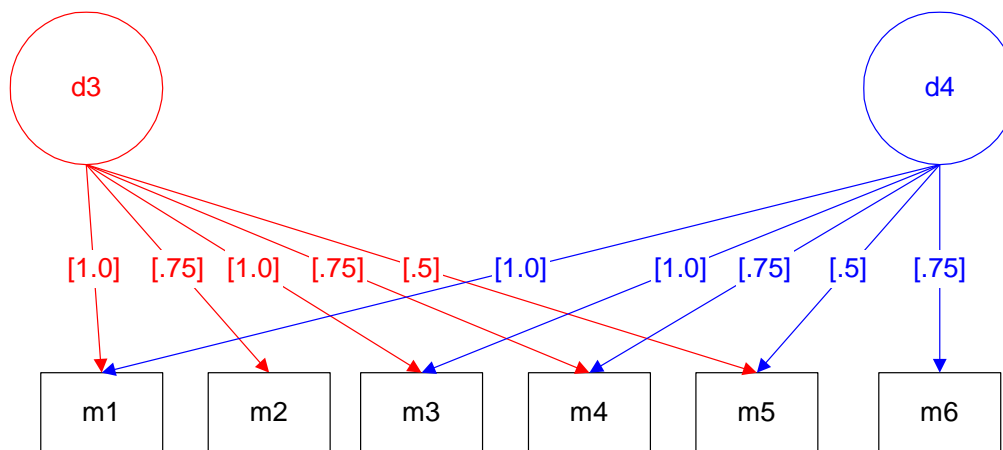


FIGURA 5.2 – Associação causal da ferrugem comum (d_3) e da ferrugem *polysora* (d_4).

As manifestações $m_1, m_2, m_3, m_4, m_5, m_6$ correspondem à formação de pústulas individualizadas marrom-claro, formação de fendas, formação de pústulas marrom-escuro, seca prematura das plantas, baixa produção e *adquirem aspecto pulverulento*, respectivamente (retiradas da descrição na seção 1.1.1).

5.1.2 Evolução Temporal

Além da associação causal entre desordens e manifestações, é importante representar a evolução temporal das manifestações para uma determinada desordem. No escopo da TCN, a informação temporal também é definida como intervalos nebulosos e pode ser representada através de um grafo temporal de eventos $G = (V, T)$ como descrito na definição 3.12 em 3.3.1. Na Figura 5.3 está representado um modelo simplificado do grafo temporal de eventos para a ferrugem comum d_3 apresentada na Figura 5.2. Os intervalos ‘menos de 24 semanas’, ‘por volta de 8 a 32 semanas’, ‘em menos de 4 semanas’ e ‘por volta de 16 a 20 semanas’ são dados fictícios que correspondem aos intervalos $\langle 0,24 \rangle$, $\langle 8,32 \rangle$, $\langle 0,4 \rangle$ e $\langle 16,20 \rangle$, respectivamente.

Na Figura 5.3 estão representados os intervalos nebulosos entre alguns eventos. Entretanto, é necessário representar o intervalo mínimo entre todos os eventos de uma

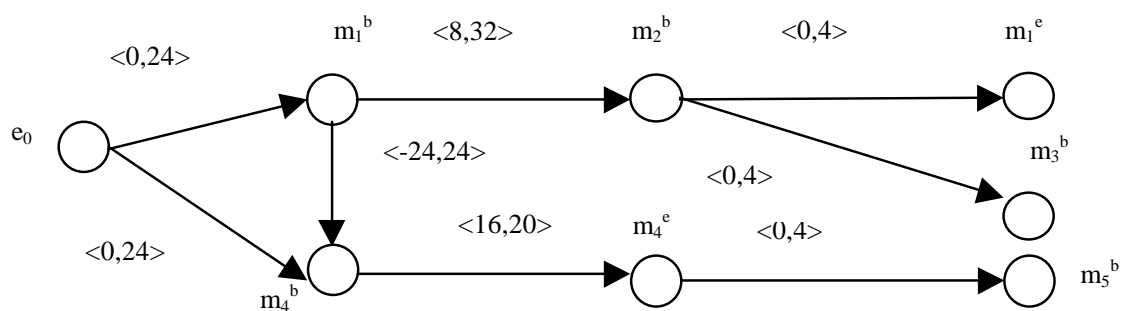


FIGURA 5.3 – Modelo de um grafo de eventos associado a uma desordem no escopo da TCN.

desordem d_i no escopo da TCN. Dado um conjunto de intervalos entre alguns eventos como apresentado na Figura 5.3, a rede mínima é uma maneira de calcular os intervalos mínimos entre quaisquer dois eventos. O grafo mínimo para cada desordem d_i pode ser calculada pelo algoritmo de Floyd-Warshall, que calcula o caminho mínimo entre todos os pares como descrito em Wainer e Sandri (1998). A Figura 5.4 representa parte do grafo mínimo de eventos para a desordem ferrugem comum (d_3). Os intervalos nebulosos que aparecem nos arcos do grafo mínimo representam os intervalos de tempo no melhor caso.

5.1.3 Condições de Controle

Outra propriedade característica de um ambiente de diagnóstico são as condições de controle que interferem na evolução de uma desordem. Por exemplo, as condições favoráveis de temperatura e umidade, a intensidade ou severidade de uma desordem e a época ou idade que favorece a evolução da desordem. No escopo da TCN, este tipo de informação também é representada através de intervalos nebulosos como apresentado no apêndice A. A tabela 5.2 exemplifica os intervalos assumidos para temperatura e umidade.

TABELA 5.2 – Graus de temperatura e umidade.

Graus	Temperatura(T)	Umidade(U)
alta	$T > 25^{\circ}C$	$U > 85\%$
média	$15^{\circ}C < T < 25^{\circ}C$	$70\% < U < 80\%$
baixa	$T < 15^{\circ}C$	$U < 70\%$

A Figura 5.5 mostra graficamente os intervalos nebulosos trapezoidais para os graus de temperatura (alta, média e baixa) expressos na tabela 5.2.

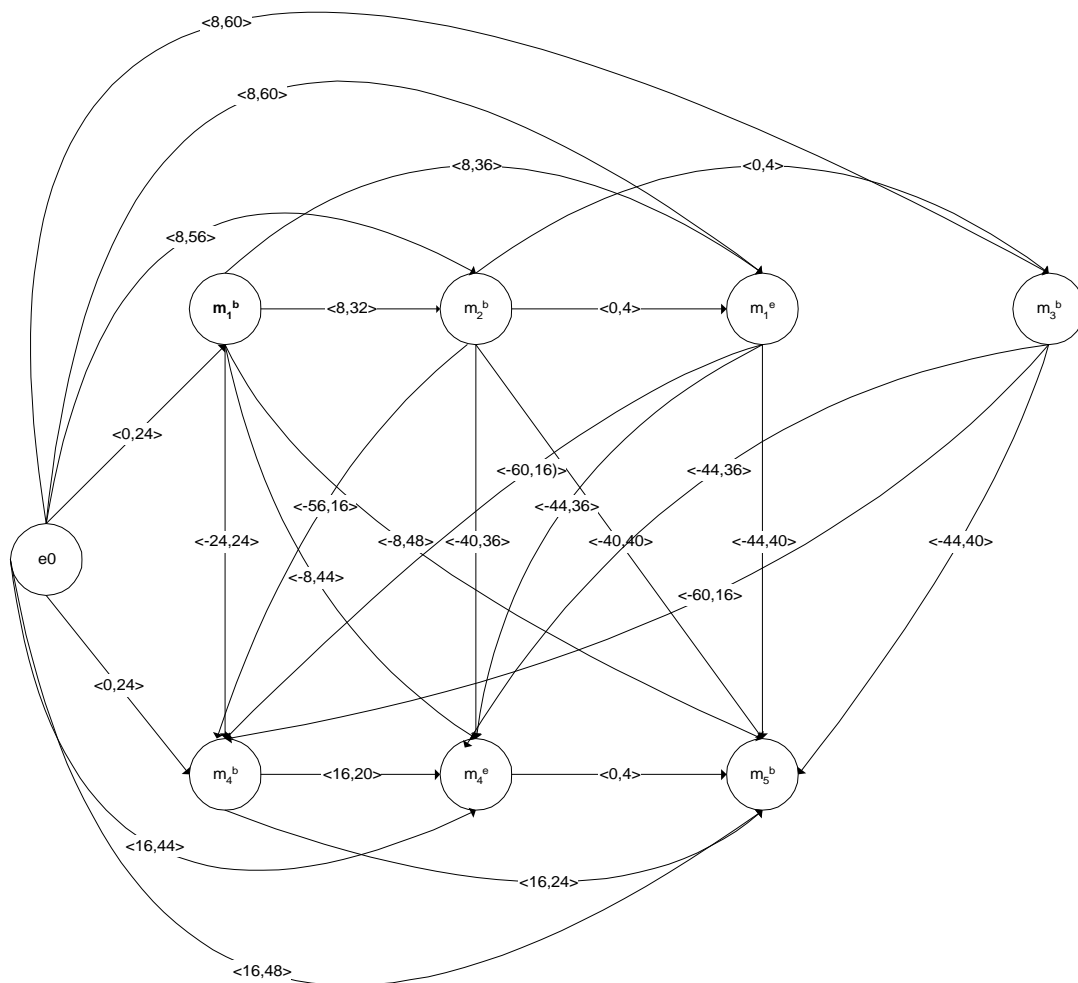


FIGURA 5.4 – Modelo parcial de um grafo mínimo associado a uma desordem no escopo da TCN.

Similarmente, as variáveis de severidade e época são representadas como intervalos nebulosos trapezoidais. A partir das idéias acima, pode-se definir formalmente as condições de controle de uma desordem d_i .

Definição 5.1 As condições de controle CC de uma desordem d_i é representada pela tripla (CF, SEV, EP) onde:

- o par CF corresponde aos intervalos nebulosos trapezoidais que representam as condições favoráveis à evolução de uma desordem d_i (ex. Temperatura e Umidade no domínio agrícola $CF = \langle TEMP, UMID \rangle$);
- SEV corresponde a um grau nebuloso trapezoidal de intensidade ou severidade de uma desordem d_i
- EP corresponde a um intervalo nebuloso que representa a fase de desenvolvimento ou idade de um determinado paciente ou planta que

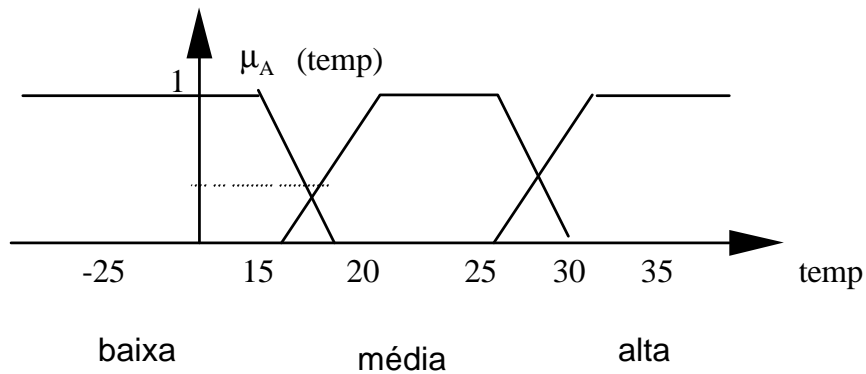


FIGURA 5.5 – Representação nebulosa para os graus de temperatura.
favorece a evolução de uma determinada desordem.

5.1.4 Base de Conhecimento

Para suportar as idéias acima, a base de conhecimento KB_d deve englobar informação possibilística. A definição formal da base de conhecimento possibilística KB_d está descrita a seguir.

Definição 5.2 A base de conhecimento KB_d é representada por uma sextupla $\langle D, M, R, G, CC, \theta \rangle$ onde:

- D é o conjunto de desordens;
- M é o conjunto de manifestações;
- $R : D \times M \rightarrow [0, 1]$ é uma relação nebulosa que modela a associação gradual entre desordens e manifestações como apresentado na tabela 5.1;
- G é o grafo mínimo de eventos associado a cada desordem para representar a evolução de uma desordem (seção 5.1.2).
- CC é representada pela tripla $CC = (CF, SEV, EP)$ (definição 5.1);
- θ é a escala de tempo.

Para um problema de diagnóstico representado por KB_d , é conveniente e útil definir os conjuntos ou funções baseadas na relação C:

Definição 5.3 Para todo $d_i \in D$ e $m_j \in M$ em uma base de conhecimento KB_d ,

- $efeitos(d_i) = \{m_j \mid \langle d_i, m_j \rangle \in C\}$, o conjunto dos elementos diretamente causados por d_i ;
- $causas(m_j) = \{d_i \mid \langle d_i, m_j \rangle \in C\}$, o conjunto dos elementos que podem diretamente causar m_j .

Os conjuntos definidos acima podem facilmente ser generalizados de desordens e manifestações individualmente para conjuntos de desordens e manifestações conforme apresentado a seguir.

Definição 5.4 Para todo $D_L \subseteq D$ e $M_J \subseteq M$ em uma base de conhecimento KB_d ,

- $efeitos(D_L) = \bigcup_{d_i \in D_L} efeitos(d_i)$,
- $causas(M_J) = \bigcup_{m_j \in M_J} causas(m_j)$.

Assim, os efeitos de um conjunto de desordens são apenas a união dos efeitos individuais de cada desordem do conjunto. Por exemplo, no problema de diagnóstico apresentado na Figura 5.1 tem-se:

$$\begin{aligned} efeitos(d_1) &= \{m_1, m_4\}, \\ efeitos(d_2) &= \{m_1, m_3, m_4, m_6\}, \\ causas(m_1) &= \{d_1, d_2, d_3, d_4\}, \\ causas(m_2) &= \{d_5, d_6, d_7, d_9\}. \end{aligned}$$

Utilizando estes dados tem-se que:

$$\begin{aligned} efeitos(d_1, d_2) &= \{m_1, m_3, m_4, m_6\}. \\ causas(m_1, m_2) &= \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_9\}. \end{aligned}$$

5.2 Especificação do Caso particular

Para complementar a formulação do problema de diagnóstico necessita-se de um caso particular. Utilizando a terminologia da TCP, M^+ é utilizado como subconjunto de M para referenciar as manifestações presentes que correspondem as observações do caso e M^- refere-se também a um subconjunto de M que corresponde as manifestações ausentes no caso.

A informação sobre um caso particular deve contemplar observações adicionais, além das manifestações presentes e ausentes, tais como, em que época estas manifestações

ou sintomas ocorreram, quando elas começaram e como evoluíram, a severidade e intensidade de cada uma bem como condições ambientais em que elas ocorreram.

Normalmente, a informação sobre o caso particular é coletada através de um processo denominado anamnese em que o especialista do domínio (médico, fitopatologista ou engenheiro) coleta as informações do próprio paciente, produtor ou computador. Formalmente, a informação sobre um dado caso pode ser definida como a seguir.

Definição 5.5 *A informação sobre um caso particular é modelada por uma quintupla $CA = \langle M^+, M^-, EV^+, TIME^+, CC^+ \rangle$ onde:*

- M^+ é o conjunto de manifestações presentes conhecidas no caso;
- M^- é o conjunto de manifestações ausentes conhecidas no caso;
- EV^+ é o conjunto de eventos de início e fim de cada manifestação em M^+ ;
- $TIME^+$ é uma função que associa para cada evento $e_i \in EV^+$ um intervalo nebuloso que representa os momentos possíveis no qual o evento ocorreu;
- $CC^+ = (CF^+, SEV^+, EP^+)$ são as condições externas em que ocorreram os sintomas, onde no domínio agrícola CF^+ é uma função que associa as condições de temperatura e umidade ($TEMP^+, UMID^+$) no momento do diagnóstico, SEV^+ e EP^+ são as funções que associam a severidade de cada $m_j \in M^+$ e a fase em que elas ocorreram, respectivamente.

Uma vez caracterizado o problema de diagnóstico ($P = \langle KB_d, CA \rangle$), caracteriza-se agora sua solução, denominado $Sol_d(P)$. Para isto, além da noção de "cobertura" (seção 3.2.2), utiliza-se outros índices de consistência e coberturas nebulosas que serão descritos a seguir.

5.3 Solução para o Problema de Diagnóstico

Nesta seção descreve-se as soluções para os problemas de diagnóstico definidos na seção 5.1. Inicialmente, define-se a noção de consistência entre um modelo desordens (seção 5.1) e um caso particular (seção 5.2). Na sequência, uma explicação plausível para um problema de diagnóstico é definida formalmente.

5.3.1 Consistência entre o Modelo de Desordens e o Caso

As explicações ou hipóteses mais plausíveis D_l (desordens ou falhas) devem ser inferidas abduativamente a partir da consistência entre a informação sobre o caso particular

CA (sintomas ou resultados de testes) coletada através da *anamnese* e a base de conhecimento KB_d . Para tal, é utilizada a noção de coberturas irredundantes da TCP apresentada na seção 3.2.2. O critério de parcimônia irredundante garante a noção de minimalidade e simplicidade explicitada na seção 3.1.

No escopo da TCN, antes de aplicar o critérios de parcimônia de irredundância, é necessário verificar a consistência temporal, categórica e das condições de controle do caso particular em relação ao modelo de desordens. Para isto, são utilizados os índices de consistência temporal, categórica e de intensidade propostos em Wainer e Sandri (1998).

5.3.1.1 Consistência Temporal

Segundo Wainer e Sandri (1998), a consistência temporal entre o caso e o modelo é calculada pela comparação do tempo transcorrido entre os eventos do caso (os eventos em EV^+) e os intervalos nebulosos correspondentes no modelo. Primeiramente, deve ser calculada a distância temporal entre pares de todos os eventos de EV^+ como descrito na definição 3.14 (seção 3.3.1).

Na sequência, para comparar como estes dois eventos se encaixam no modelo de uma desordem particular d_i deve-se comparar a $DIST^+(e_i, e_j)$ com $\pi(e_i, e_j)$. Um caso é consistente temporalmente com o modelo de uma desordem se seu grau de consistência for maior que zero. O grau de consistência temporal de dois eventos e_i e e_j é a altura da intersecção de $DIST^+(e_i, e_j)$ com $\pi(e_i, e_j)$. Formalmente, a consistência temporal entre uma caso CA e uma desordem d_i é definido como descrito na definição 3.15.

Para exemplificar a consistência temporal serão apresentados dois casos hipotéticos de produtores de milho. Consideremos novamente o conjunto de manifestações $M = \{m_1, m_2, m_3, m_4, m_5\}$ que caracterizam a desordem ferrugem comum (d_3) apresentada na seção 5.1.1.

Seja o caso hipotético 1 em que $M^+ = \{m_1, m_4\}$; $EV^+ = \{m_1^b, m_4^b\}$; $TIME^+(m_1^b) = \langle 49, 50 \rangle$; $TIME^+(m_4^b) = \langle 23, 24 \rangle$ e $DIST^+(m_1^b, m_4^b) = \langle 25, 27 \rangle$. A partir do grafo mínimo (Figura 5.4) tem-se $\pi(m_1^b, m_4^b) = \langle -24, 24 \rangle$ e conseqüentemente $\alpha(d_i) = h(DIST^+(m_1^b, m_4^b) \cap \pi(m_1^b, m_4^b)) = h(\langle 25, 27 \rangle \cap \langle -24, 24 \rangle) = 0.5$

Portanto, o modelo e o caso são completamente incompatíveis em termos temporais. Realmente, se considerarmos o $TIME^+(m_1^b)$ de m_1 , o evento e_0 deveria ter ocorrido no intervalo $\langle 25, 50 \rangle$ e considerando o $TIME^+(m_4^b)$, e_0 deveria ter ocorrido no intervalo $\langle -1, 24 \rangle$, o que os torna totalmente incompatíveis.

Seja o caso hipotético 2 em que $M^+ = \{m_1, m_4\}$; $EV^+ = \{m_1^b, m_4^b\}$; $TIME^+(m_1^b) = \langle 48, 49, 50, 51 \rangle$; $TIME^+(m_4^b) = \langle 22, 23, 24, 25 \rangle$ e $DIST^+(m_1^b, m_4^b) = \langle 23, 25, 27, 29 \rangle$. A partir do grafo mínimo (Figura 5.4) tem-se $\pi(m_1^b, m_4^b) = \langle -24, 24 \rangle$ e consequentemente $\alpha(d_i) = h(DIST^+(m_1^b, m_4^b) \cap \pi(m_1^b, m_4^b)) = h(\langle 23, 25, 27, 29 \rangle \cap \langle -24, 24 \rangle) = 0.5$ (veja Figura 5.6).

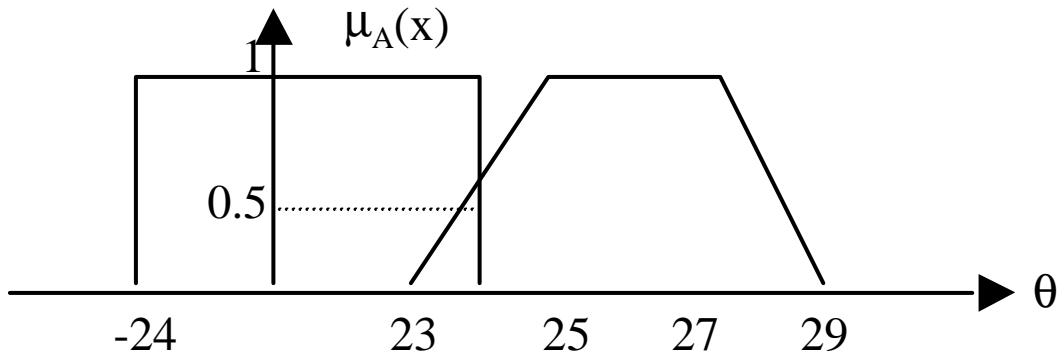


FIGURA 5.6 – Exemplo do grau de consistência temporal.

Assim, a desordem ferrugem comum (d_3) poderia ser considerada como uma explicação plausível para M^+ . Adicionalmente, conforme apresentado em Wainer e Sandri (1998), a consistência categórica está intimamente ligada à consistência temporal, pois, um caso é categoricamente inconsistente com um modelo caso uma manifestação necessária não tenha ocorrido e já não há mais tempo suficiente para que ela ocorra. Uma manifestação m_j é dita ter tempo suficiente para ter ocorrido em d_i se:

- existe um evento e_j que deveria ocorrer após m_j e ele já ocorreu;
- ou existe um evento e_j que era previsto ocorrer antes de m_j , ele ocorreu e o tempo previsto para ocorrer m_j após e_j já estourou.

Assim, o índice de consistência temporal e categórica $\alpha(d_i)$ no escopo da TCN pode ser definido como descrito na definição 5.6 que é uma extensão das definições 3.15 e 3.17 apresentadas na seção 3.3.1.

Definição 5.6 *Um caso particular CA é consistente temporal e categoricamente com o modelo de uma desordem d_i se:*

- *tiver a seguinte inicialização, $\forall m_j \in M^- \cap R(d_i, m_j) = 1$, $TIME^+(m_j^b) = I_{afternow} = I_{I \geq 0}$;*

- e seu grau de consistência temporal e categórica for maior que zero, isto é, se $\alpha(d_i) = \min_{e_i, e_j \in EV^+ \cap V(d_i)} [h(DIST^+(e_i, e_j) \cap \pi_i(e_i, e_j))] > 0$.

Para exemplificar a consistência categórica no escopo da TCN, consideremos o caso hipotético em que o momento de diagnóstico é igual a 80 (semanas): $M^+ = m_4, m_5$; $M^- = m_1$; $R(d_1, m_1) = 1.0$; $EV^+ = m_4^e, m_5^b$; $TIME^+(m_4^e) = \langle 49, 50 \rangle$ e $TIME^+(m_5^b) = \langle 48, 52 \rangle$ $DIST^+(m_4^e, m_5^b) = \langle -2, 3 \rangle$ e $\pi_1(m_4^e, m_5^b) = \langle 0, 4 \rangle$. Se forem considerados m_4 e m_5 sem considerar m_1 , tem-se que o caso é compatível com a doença d_3 conforme apresentado abaixo: $\alpha(d_1) = h(DIST^+(m_4^e, m_5^b) \cap \pi_1(m_4^e, m_5^b)) = h(\langle -2, 3 \rangle \cap \langle 0, 4 \rangle) = 1.0$.

Se for considerada a ausência da manifestação m_1 onde $R(d_3, m_1) = 1.0$, com $TIME^+(m_1^b) = I_{afternow} = \langle 80, \infty \rangle$, o caso já não é compatível com o modelo de d_3 , conforme apresentado abaixo:

$$\alpha(d_3) = \min \left\{ \begin{array}{l} (h(DIST^+(m_4^e, m_5^b) \cap \pi_1(m_4^e, m_5^b))), \\ (h(DIST^+(m_1^b, m_4^e) \cap \pi_1(m_1^b, m_4^e))), \\ (h(DIST^+(m_1^b, m_5^b) \cap \pi_1(m_1^b, m_5^b))) \end{array} \right.$$

$$\alpha(d_3) = \min \left\{ \begin{array}{l} h(\langle -2, 3 \rangle \cap \langle 0, 4 \rangle), \\ h(\langle -\infty, -30 \rangle \cap \langle -8, 44 \rangle), \\ h(\langle -\infty, -28 \rangle \cap \langle -8, 48 \rangle), \end{array} \right.$$

$$\alpha(d_3) = \min(1, 0, 0) = 0.$$

5.3.1.2 Consistência das Condições de Controle

Adicionalmente, ao grau de consistência temporal, é necessário calcular o grau de consistência das variáveis que compõem as condições de controle. No escopo deste trabalho, são utilizados os índices de consistência de intensidade propostos em Wainer e Sandri (1998).

Em Wainer e Sandri (1998), para obter consistência de intensidade, utiliza-se uma função INT na base conhecimento que atribui para cada nó m de $E(d)$ um conjunto nebuloso INT(m) descrevendo a intensidade com que aquela manifestação deve ocorrer. Cada conjunto nebuloso INT(m) é definido sobre um domínio particular $\Omega_{INT(m)}$. Do mesmo modo, para a informação do caso existe uma função INT^+ que atribui para cada $m_j \in CA = \{M^+\}$ um conjunto nebuloso $INT^+(m)$ descrevendo a intensidade do qual a manifestação ocorreu. As equações 3.3 e 3.4 definidas na seção 3.3.1 formalizam a consistência de intensidade de uma manifestação m_j em relação a uma desordem d_i .

Para exemplificar o índice de intensidade no domínio de doenças de milho, suponhamos o caso em que as manifestações manchas marrom claro (m_1) e manchas marrom escuro (m_3) da ferrugem comum são caracterizadas por se encontrar em densamente distribuídas por ambos os lados das folhas. Seja $\Omega_{INT(m_1, m_3)} = [0, 10]$ um domínio para intensidade de pústulas nas folhas de plantação de milho e que pústulas densamente distribuídas seja modelado por $INT(m_1, m_3) = \langle 7.5, 8.5, 9.5, 10 \rangle$.

Suponhamos o caso em que o produtor diz que apareceram muitas pústulas marrom-claro distribuídas pelas folhas da plantação e após quatro semanas a quantidade de pústulas aumentou e escureceram. Este caso pode ser modelado pelo intervalo nebuloso $INT^+(m_1) = \langle 4, 5, 7, 8 \rangle$ e $INT^+(m_3) = \langle 6, 7, 8, 9 \rangle$ sobre o mesmo domínio $\Omega_{INT(m_1, m_3)}$. Assim, os graus de consistência para m_1 e m_3 , ao aplicar a equação 3.3, são $\gamma(m_1) = h \langle 7.5, 8.5, 9.5, 10 \rangle \cap \langle 4, 5, 7, 8 \rangle = 1.0$ e $\gamma(m_3) = h \langle 7.5, 8.5, 9.5, 10 \rangle \cap \langle 6, 7, 8, 9 \rangle = 1.0$, respectivamente.

Aplicando a fórmula 3.4 tem-se que as manifestações presentes em CA são consistentes com a evolução da desordem d_3 : $\gamma(d_3) = \inf(\gamma(m_1), \gamma(m_3)) = 1.0$. Similarmente, é definido o índice de intensidade das condições de controle no diagnóstico de doenças de milho.

Definição 5.7 O grau de consistência das condições de controle de uma variável $c_k \in CC$ e um caso particular CA é calculado através da fórmula $\gamma(d_i) = \inf_{c_k \in CC} \gamma(c_k)$ de modo que $\gamma(c_k) = h(INT(c_k) \cap INT^+(c_k))$, $INT(c_k)$ corresponde a intensidade de $c_k \in CC$ em KB_d e $INT^+(c_k)$ corresponde a intensidade de cada $c_k \in CC^+$.

Primeiramente, define-se a fórmula para consistência das condições favoráveis a partir das equações 3.3 e 3.4 :

$$\omega_{CF}(d_i) = \sup(\gamma(TEMP), \gamma(UMID)) \quad (5.1)$$

Para exemplificar o índice de intensidade das condições de controle no diagnóstico de doenças de milho, suponhamos novamente o caso hipotético 1 $CA_1 = \langle M^+, M^-, EV^+, TIME^+, CC^+ \rangle$ onde $CC^+ = CF^+, SEV^+, EP^+$ onde $EP^+ = \langle 7, 8, 9, 10 \rangle$, $SEV^+ = \langle 3, 4, 5, 6 \rangle$ e $CF^+ = TEMP^+ = \langle 10, 15, 15, 20 \rangle$, $UMID^+ = \langle 80, 85, 90, 95 \rangle$.

Aplicando a fórmula 3.3 tem-se o grau de consistência de intensidade de temperatura e umidade como descrito abaixo. $\gamma(TEMP) = h \langle 15, 20, 25, 30 \rangle \cap \langle 10, 15, 15, 20 \rangle = 0.5$ e $\gamma(UMID) = h \langle 85, 90, 95, 100 \rangle \cap \langle 80, 85, 90, 95 \rangle = 1.0$.

Ao aplicar a fórmula 7.2 tem-se o grau de consistência de das condições favoráveis: $\omega_{CF}(d_3) = 1.0$.

Observemos que o domínio de intensidade da variável $CF \in V$ é $\Omega_{INT(CF)} = [0, 1]$ tal que $\Omega_{INT(CF)} = [0, 0.5] = \neg CF$ (quando CF é falsa) e $\Omega_{INT(CF)} = [0.5, 1.0] = CF$ (quando CF é verdadeira).

No exemplo acima, as condições de temperatura e umidade do caso CA são compatíveis com as condições favoráveis da desordem d_3 . Em um segundo passo, conforme descrito na definição 5.7, calcula-se o grau de consistência entre uma variável de controle $c_k \in CC$ e um caso particular $c_k \in CC$ em CA é através da fórmula $\gamma(d_i) = \inf_{c_k \in V} \gamma(c_k)$. No exemplo dado acima, calcula-se o grau de consistência para a desordem d_3 ((d_3)):

$$\begin{aligned}\gamma(EP) &= h \langle 2, 4, 10, 12 \rangle \cap \langle 7, 8, 9, 10 \rangle = 1.0 \\ \gamma(SEV) &= h \langle 1, 3, 5, 7 \rangle \cap \langle 3, 4, 5, 6 \rangle = 1.0 \\ \omega(CF) &= 1.0 \\ \gamma(d_3) &= \inf(\gamma(EP), \gamma(SEV), \omega(CF)) = 1.0\end{aligned}$$

Portanto, no exemplo dado as condições de controle do caso são compatíveis com as condições de controle do modelo de uma desordem d_3 . Agora, o índice de consistência das condições de controle $\gamma(d_i)$ no escopo da TCN pode ser definido como descrito na definição 5.8.

Definição 5.8 *Um caso particular CA é consistente com as condições de controle do modelo de uma desordem d_i se:*

- *se seu grau de consistência de intensidade das condições de controle for maior que zero, isto é, se $\gamma(d_i) > 0$.*

5.3.2 Solução para o Diagnóstico

No escopo da TCN, uma explicação para o diagnóstico assumida é uma desordem que tem consistência temporal e categórica com todos os sintomas e explica todos os sintomas presentes no caso com certo grau de verossimilhança, como definido a seguir.

Definição 5.9 *Um conjunto $SOL_d \subseteq D$ é uma explicação plausível para um problema $P = \langle KB_d, CA \rangle$ onde $CA = \langle M^+, M^-, EV^+, TIME^+, CC^+ \rangle$ se*

- *SOL_d cobre M^+, e*
- *SOL_d satisfaz algum critério de parcimônia, e*

- $\forall d_i \in SOL_d, \alpha(d_i) > 0$, isto é, d_i é consistente temporalmente e categoricamente, e
- $\forall D_l \subseteq SOL_d, \alpha(D_l) = \min(\alpha(d_i))|\{d_i \in D_l\}$, isto é, o grau de certeza de uma hipótese D_l é igual ao menor grau de consistência de uma desordem $d_i \in D_l$ tal que $D_l \in SOL_d$.

A melhor explicação para um caso particular CA é uma uma cobertura D_l com maior grau de certeza, isto é, $\forall D_l \in SOL_d$ existe uma classificação onde as hipóteses com os mais altos graus de certeza constituem as hipóteses mais plausíveis. O grau de certeza das hipóteses mais plausíveis é calculado pelo menor grau de consistência temporal e categórica das desordens que a compõem, como definido acima. Entretanto, quando existir duas desordens com mesmo grau de certeza ($\alpha(D_l)$), é considerada a hipótese mais plausível é a hipótese com maior grau de compatibilidade com as condições de controle ($\gamma(D_l)$) como definido a seguir.

Definição 5.10 *Dado que $\gamma(D_l) = \min(\gamma(d_i))|\{d_i \in D_l\}$, uma cobertura $D_x \in SOL_d$ é uma explicação mais plausível para um caso particular CA se $\forall D_l \subseteq SOL_d | D_l \neq D_x, \alpha(D_l) = \alpha(D_x)$ e $\gamma(D_x) > \gamma(D_l)$.*

No escopo da TCN, quando uma hipótese D_l não atinge um certo grau de verossimilhança (i.e, $\alpha(D_l) < \Delta$) pode ser assumido que os efeitos causados por cada d_i tal que $d_i \in D_l$ devem ser melhor investigados (i.e, através de exames complementares ou testes laboratoriais) visando um refinamento das hipóteses. O ambiente abduutivo para investigação será detalhado no próximo capítulo.

CAPÍTULO 6

AMBIENTE ABDUTIVO PARA INVESTIGAÇÃO

Neste capítulo é apresentado detalhadamente o ambiente abduativo para investigação no escopo da TCN, que incorpora conceitos de inferência abduativa, lógica nebulosa e teoria de decisão. Primeiramente, nas seções 6.1 e 6.2 são apresentados detalhadamente os mecanismos para representação do conhecimento do problema de investigação e de um caso particular, bem como das incertezas inerentes a estes problemas, no escopo desta nova infraestrutura. Na seção 6.3 são apresentados os mecanismos de inferência e de tomada de decisão utilizados para investigação.

6.1 Representação do Conhecimento em Investigação

O ambiente abduativo para investigação é similar ao ambiente abduativo definido para diagnóstico no escopo da TCN. Na Figura 6.1 estão representados os dois conjuntos finitos e discretos para modelar o subproblema de investigação. O conjunto M representa todas as manifestações m_j (sintomas, falhas, etc.) que podem ocorrer quando uma ou mais desordens estiverem presentes. O conjunto E representa todos os exames possíveis e_l para verificar a existência de manifestações em um determinado domínio de aplicação. A relação $I : E \times M \rightarrow \{0, 1\}$ é utilizada para modelar a associação que parte dos exames para as manifestações; se $(e_l, m_j) \in I$ então e_l é um exame apropriado para a detecção da existência de m_j .

Por exemplo, quando um dado paciente tem uma dor no joelho, o médico pode solicitar um raio-X, uma ecografia, uma ressonância magnética ou uma tomografia para confirmar o sinal físico que está provocando o efeito de dor no joelho. Entretanto, estes exames não são necessariamente específicos a uma única desordem; eles podem ser utilizados para confirmar manifestações causadas por várias desordens. A solicitação de exames não depende apenas da manifestação em si mas também de qual é a hipótese dominante (conjunto de desordens), obtida por exemplo através de anamnese.

Durante o processo de diagnóstico, é importante selecionar quais as manifestações prioritárias para completar o diagnóstico e, conseqüentemente, quais os exames a serem solicitados para confirmar a existência destas manifestações. Este ciclo diagnóstico/investigação pode ser repetido até que um diagnóstico final seja encontrado. A próxima definição formaliza a base de conhecimento para exames em um dado momento no tempo.

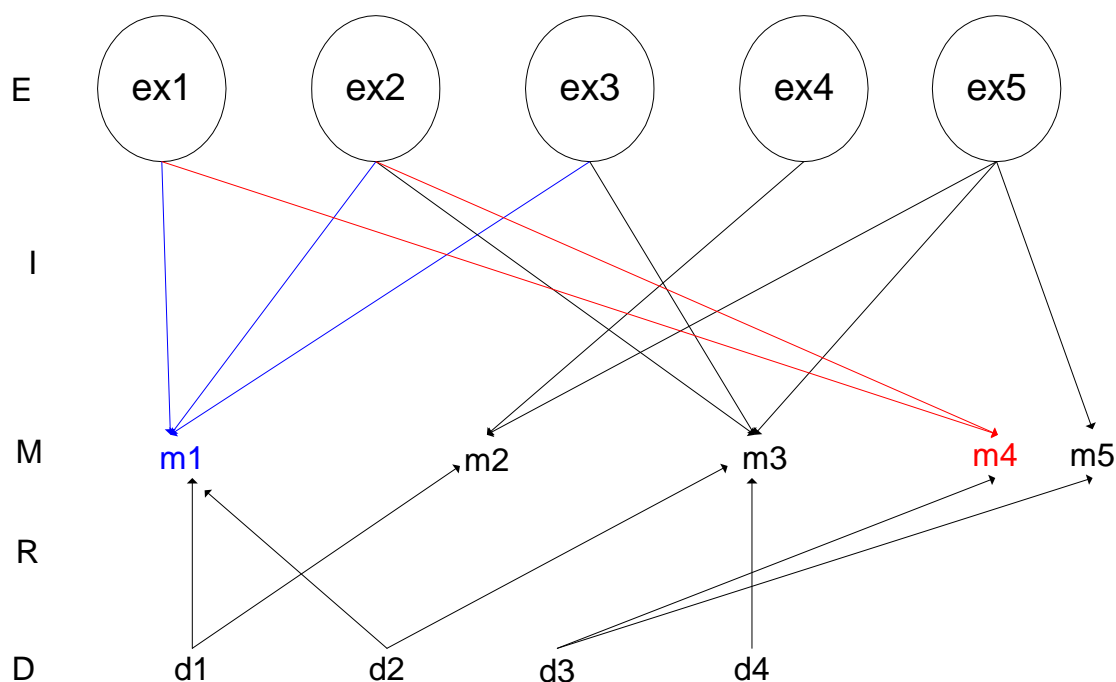


FIGURA 6.1 – Um ambiente abduativo para investigação.

Definição 6.1 A base de conhecimento para investigação é representada pela quintupla $KB_e = \langle E, D, M, I, R \rangle$, onde:

- E é o conjunto de exames ou testes laboratoriais;
- D é o conjunto de desordens;
- M é o conjunto de manifestações ou efeitos observados;
- $I : E \times M \rightarrow \{0, 1\}$ é uma relação binária que modela a associação entre exames e manifestações;
- $R : D \times M \rightarrow [0, 1]$ é uma relação nebulosa que modela a associação gradual entre desordens e manifestações como apresentado na tabela 5.1.

Para um problema de investigação representado por KB_e , é conveniente e útil definir os seguintes conjuntos obtidos através da relação I:

Definição 6.2 Para todo $e_i \in E$ e $m_j \in M$ em uma base de conhecimento KB_e , temos

- $resultados(e_i) = \{m_j \mid \langle e_i, m_j \rangle \in I\}$ é o conjunto das manifestações diretamente identificados pelo exame e_i ;

- $exames(m_j) = \{e_l \mid \langle e_l, m_j \rangle \in I\}$ é o conjunto dos exames que podem diretamente identificar a manifestação m_j .

Os conjuntos definidos acima são facilmente generalizados de exames e manifestações individuais para conjuntos de exames e manifestações, conforme apresentado a seguir.

Definição 6.3 Para todo $E_L \subseteq E$ e $M_J \subseteq M$ em uma base de conhecimento KB_e , temos

- $resultados(E_L) = \bigcup_{e_l \in E_L} resultados(e_l)$,
- $exames(M_J) = \bigcup_{m_j \in M_J} exames(m_j)$.

Por exemplo, no problema de investigação apresentado na Figura 6.1 tem-se $resultados(e_1) = \{m_1, m_4\}$, $resultados(e_2) = \{m_1, m_3, m_4\}$, $exames(m_1) = \{e_1, e_2, e_3\}$ e $exames(m_2) = \{e_4, e_5\}$. Consequentemente, $resultados(\{e_1, e_2\}) = \{m_1, m_3, m_4\}$ e $exames(\{m_1, m_2\}) = \{e_1, e_2, e_3, e_4, e_5\}$.

6.2 Especificação do Caso particular

Para completar a formulação do problema de investigação é necessário formalizar os casos particulares. No escopo de investigação, a informação sobre um caso particular deve contemplar apenas observações sobre manifestações presentes e ausentes.

Definição 6.4 A informação sobre um caso particular na etapa de investigação é modelada por uma tripla $CA = \langle D^+, M^+, M^- \rangle$ onde:

- $D^+ \subseteq \wp(D)$ é o conjunto de hipóteses para o caso;
- M^+ é o conjunto de manifestações certamente presentes no caso;
- M^- é o conjunto de manifestações certamente ausentes no caso.

O conjunto D^+ pode ter sido derivado a partir da aplicação do módulo de diagnóstico, e neste caso, portanto, $D^+ = SOL_d$. O conjunto de desordens presentes nas coberturas em D^+ é dado por $D^* = \bigcup_{D_k \in D^+} D_k$. Por exemplo, para $D^+ = \{\{d_1\}, \{d_2, d_3\}\}$, temos $D^* = \{d_1, d_2, d_3\}$. O conjunto de manifestações cuja investigação é pertinente ao caso apresentado é dado por $M^* = efeitos(D^*) - (M^+ \cup M^-)$. Por exemplo, para $M^+ \cup M^- = \{m_3\}$, $D^+ = \{\{d_1\}, \{d_2, d_3\}\}$, $efeitos(d_1) = \{m_1, m_3\}$, $efeitos(d_2) = \{m_2, m_3\}$ e $efeitos(d_3) = \{m_4\}$ temos $M^* = \{m_1, m_2, m_3, m_4\} - \{m_3\} = \{m_1, m_2, m_4\}$. O conjunto de coberturas em D^+ que podem causar uma manifestação m_j é dado por $D_j^+ = \{D_k \mid D_k \in D^+, D_k \cap causas(m_j) \neq \emptyset\}$. Por exemplo, no exemplo anterior, se $causas(m_2) = \{d_2, d_5\}$ então temos $D_2^+ = \{\{d_2, d_3\}\}$.

6.3 Tomada de Decisão para Investigação

Nesta seção é descrita a solução $SOL_e(P)$ para um dado problema de investigação. Inicialmente, define-se a noção de consistência entre um modelo de exames (seção 6.1) e as informações de um caso particular (seção 6.2). Na sequência, uma solução plausível para um problema de investigação é definida formalmente.

6.3.1 Consistência entre o Modelo de Exames e o Caso

Neste trabalho, uma classificação em termos de prioridade para o conjunto de exames plausíveis E (exames ou testes laboratoriais) é obtida a partir da verificação da consistência entre a base de conhecimento KB_e e a informação disponível num dado momento sobre um caso particular CA . A classificação dos exames é feita da seguinte maneira:

- Para cada manifestação $m_j \in M^*$, encontra-se uma distribuição $\delta_j : D_j^+ \rightarrow [0, 1]$, que indica o quanto a confirmação da manifestação m_j é importante para a confirmação de cada hipótese $D_k \in D_j^+$.
- Para cada manifestação $m_j \in M^*$ calcula-se um índice $f(m_j)$, obtido em função da distribuição δ_j , que mede a importância geral de m_j em termos do quanto a comprovação de sua existência (ou ausência) diminui a dificuldade de se escolher, dentre as hipóteses em D^+ , qual a mais adequada para explicar o caso.
- A partir da classificação das manifestações, obtida através do índice f , encontram-se quais os exames são os mais prioritários para o caso. Os exames que verificam a existência de m_j serão tanto mais prioritários quanto melhor for a classificação de m_j .

Eventualmente, em uma última etapa de classificação poderiam ser considerados fatores como custo e risco (no caso de exames intrusivos) que poderiam ajudar a refinar a escolha do melhor conjunto de exames a serem feitos.

Por exemplo, suponhamos que $D^+ = \{\{d_1\}, \{d_2, d_3\}\}$, $M^+ = \{m_3\}$, $efeitos(d_1) = \{m_1, m_3\}$, $efeitos(d_2) = \{m_2, m_3\}$ e $efeitos(d_3) = \{m_4\}$ temos $M^* \cup M^- = \{m_1, m_2, m_3, m_4\} - \{m_3\} = \{m_1, m_2, m_4\}$. Suponhamos que os índices calculados para $\{m_1, m_2, m_4\}$ sejam $\{f(m_1) = .2, f(m_2) = .4, f(m_4) = .5\}$. Supondo que se deseje minimizar f , os exames devem ser solicitados na seguinte ordem $\langle exames(m_1), exames(m_2), exames(m_4) \rangle$. Suponhamos que $exames(m_1) =$

$\{e_1, e_2\}, \text{exames}(m_2) = \{e_2, e_3\}$ e $\text{exames}(m_4) = \{e_2, e_3\}$. Se considerarmos a manifestação m_1 como prioritária, poderíamos executar os exames e_1 ou e_2 para confirmar sua presença. Se considerássemos $\{m_1, m_2\}$ como prioritárias, somente com o exame e_2 já cobriríamos as duas manifestações.

Entretanto, se considerarmos, por exemplo, o custo associado aos exames, poderíamos obter outro resultado. Na tabela 6.1, e_2 está associado ao custo de R\$500,00; já o custo associado com os outros dois exames $\{e_1, e_3\}$ que poderiam identificar $\{m_1, m_2\}$ juntos somariam R\$200,00. Portanto, neste caso, o especialista poderia optar por fazer os exames e_1 e e_3 ao invés do exame e_2 . Este aspecto, embora simples de ser tratado, ainda não foi formalmente considerado neste trabalho.

TABELA 6.1 – Exemplos de custos associados aos exames.

Manifestações m_j	$f(m_j)$	$\text{Exames}(m_j)$	$\text{Custo}(e_i)$
m_1	.2	e_1	100,00
		e_2	500,00
m_2	.3	e_2	500,00
		e_3	100,00
m_4	.5	e_2	500,00
		e_3	100,00

Na próxima subseção serão discutidas duas funções que implementam o índice de entropia f , e na seção seguinte será discutido como as distribuições δ_j podem ser obtidas. Na última subseção serão apresentados exemplos do uso desta abordagem.

Em relação à questão do raciocínio abduutivo, parte-se aqui do princípio que são os exames que são adequados à determinação da existência de uma dada manifestação (exames \rightarrow manifestações). Como o que se tem (ou supõe-se ter) são manifestações, o raciocínio é pois abduutivo. Para tal, é utilizada a noção de coberturas irredundantes da TCP apresentada na seção 3.2.2. O critério de parcimônia irredundante garante a noção de minimalidade e simplicidade explicitada na seção 3.1. No entanto, na tarefa de investigação quando se considerar somente coberturas unitárias de exames em E , o critério de parcimônia não é significativo.

6.3.2 Funções de Entropia

A Figura 6.2 traz duas distribuições de plausibilidade de hipóteses considerando-se m_1 e m_2 respectivamente. Observa-se nessa Figura que, mesmo que m_1 seja confirmada, isso não nos auxiliará a distinguir qual a mais plausível dentre as diversas hipóteses

possíveis. No entanto, se m_2 for confirmada existe uma diferenciação maior entre as diversas hipóteses. Esta diferenciação entre as hipóteses estamos chamando de “entropia” em um abuso de linguagem.

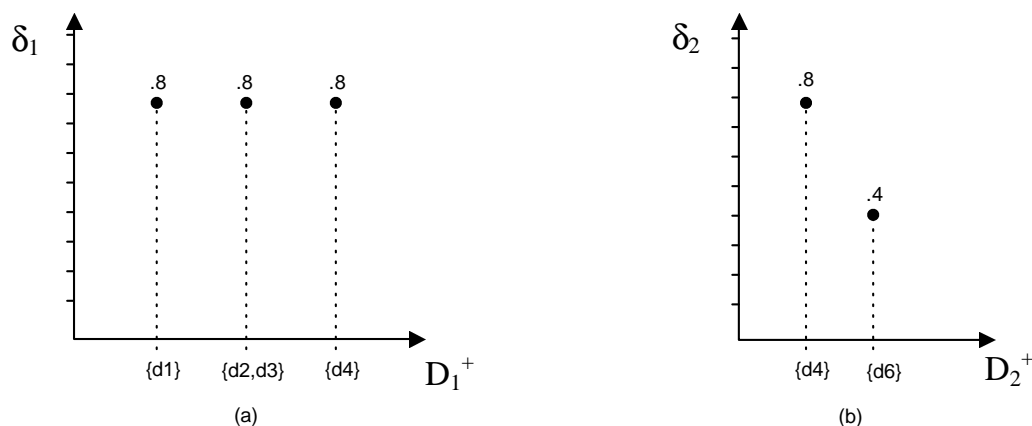


FIGURA 6.2 – Distribuição da plausibilidade das hipóteses, supondo-se a confirmação de (a) m_1 e de (b) m_2 .

Neste trabalho estudamos duas funções (f_{owa} e f_U) para medir a entropia de uma distribuição de valores $\delta_j : \Omega \rightarrow [0, 1]$, sem quaisquer restrições. Antes de apresentar as funções, definimos uma propriedade que consideramos desejáveis no contexto da tarefa de investigação. Utilizamos a notação $m_A \prec m_B$ para indicar que a investigação de m_a é prioritária em relação a m_b . Sejam as distribuições δ_a e δ_b associadas às manifestações m_a e m_b .

- Suponhamos que $|D_a^+| = |D_b^+| = q$, $\delta_a(D_k) = \delta_b(D_k)$, $1 \leq k \leq q-1$, $\delta_a(D_q) = r$, e $\delta_b(D_q) = s$:
 - se $\max(\delta_a) < r < s$ então $m_B \prec m_A$;
 - se $\max(\delta_a) < s < r$ então $m_A \prec m_B$;
 - se $\max(\delta_a) > r > s$ então $m_B \prec m_A$;
 - se $\max(\delta_a) > s < r$ então $m_A \prec m_B$.

Ao longo deste capítulo, uma dada distribuição δ será também representada por um vetor A .

6.3.2.1 Função Baseada na Família de Operadores OWA

A primeira função (f_{owa}) proposta é baseada na família dos operadores de médias ponderadas OWA (*Ordered Weighted Averaging*) (Yager, 1993) como definida a seguir.

Definição 6.5 *Seja $A = [a_1, \dots, a_n]$ um vetor com valores em R . Seja $W = [w_1, \dots, w_n]$ um vetor de pesos tal que $\forall k, w_k \in [0, 1]$ e $\sum_{k=1}^n w_k = 1$. Cada vetor W determina um operador OWA, dado por*

$$\psi_W(A) = \sum_{k=1}^n w_k u_k,$$

onde u_k é o k -ésimo maior elemento de $[a_1, \dots, a_n]$.

Sejam $A = [a_1, \dots, a_n]$ e $B = [b_1, \dots, b_n]$ dois vetores. Os operadores OWA têm as seguintes propriedades de operadores de média:

- Comutatividade: Se o vetor B for formado por uma permutação dos elementos de A , então $\psi_W(A) = \psi_W(B)$;
- Monotocidade: Se $\forall k, a_k \geq b_k$, então $\psi_W(A) \geq \psi_W(B)$;
- Idempotência: Se $\forall k, a_k = a$, então $\psi_W(A) = a$.

Três casos especiais de operadores OWA são:

Max: $\psi_{max}(A) = Max(a_1, \dots, a_n)$ para $W = [10\dots0]$,

Min: $\psi_{min}(A) = Min(a_1, \dots, a_n)$ para $W = [00\dots1]$,

Média aritmética: $\psi_{ma}(A) = (a_1 + \dots + a_n)/n$ para $W = [1/n1/n\dots1/n]$.

Suponhamos que queiramos utilizar um operador OWA no cálculo da força de m_j . Suponhamos que cada $a_k \in]0, 1]$ em A_j é tal que $a_k = \delta_j(D_k)$, $D_k \in D_j^+$ e $n = |D_j^+|$. Para isso, primeiramente, transformamos a distribuição de valores em A_j em uma distribuição normalizada $A'_j = [a_1/s, \dots, a_n/s]$, com $s = \sum_{1 \leq k \leq n} a_k$ e depois acrescentamos o valor $(1 - a_{max})$, onde $a_{max} = \max\{a_1, \dots, a_n\}$.

$$f_{owa}(A_j) = \psi_{W^{(n)}}(A'_j) * n + (1 - a_{max}), \quad (6.1)$$

onde, por definição, $\psi_{W^{(1)}}(A) = 0$

O vetor de pesos $W^{(n)}$ é o reverso do vetor $P^{(n)}$, dado por

$$p_i^{(n)} = \frac{g(i)}{k_n}, \quad k_n = \sum_{i=1}^n g(i),$$

onde $g(\cdot)$ é qualquer função estritamente monotônica crescente com $g(1) > 0$. Portanto, como os elementos de $P^{(n)}$ estão em ordem crescente, o vetor $W^{(n)}$, sendo seu reverso, tem seus elementos em ordem decrescente. Por exemplo, para $g(i) = i$, tem-se

$$\begin{aligned} W^{(2)} &= \text{reverso}([1/3, 2/3]) = [.67, .33] \\ W^{(3)} &= \text{reverso}([1/6, 2/6, 3/6]) = [.5, .34, .16] \\ W^{(4)} &= \text{reverso}([1/10, 2/10, 3/10, 4/10]) = [.4, .3, .2, .1] \end{aligned}$$

e para $f_i = i^2$ tem-se

$$\begin{aligned} W^{(2)} &= \text{reverso}([1/5, 4/5]) = [.8, .2] \\ W^{(3)} &= \text{reverso}([1/14, 4/14, 9/14]) = [.6, .3, .1] \\ W^{(4)} &= \text{reverso}([1/30, 4/30, 9/30, 16/30]) = [.53, .3, .13, .03] \end{aligned}$$

A Figura 6.3 ilustra o uso das funções $g(i) = i$ e $g(i) = i^2$.

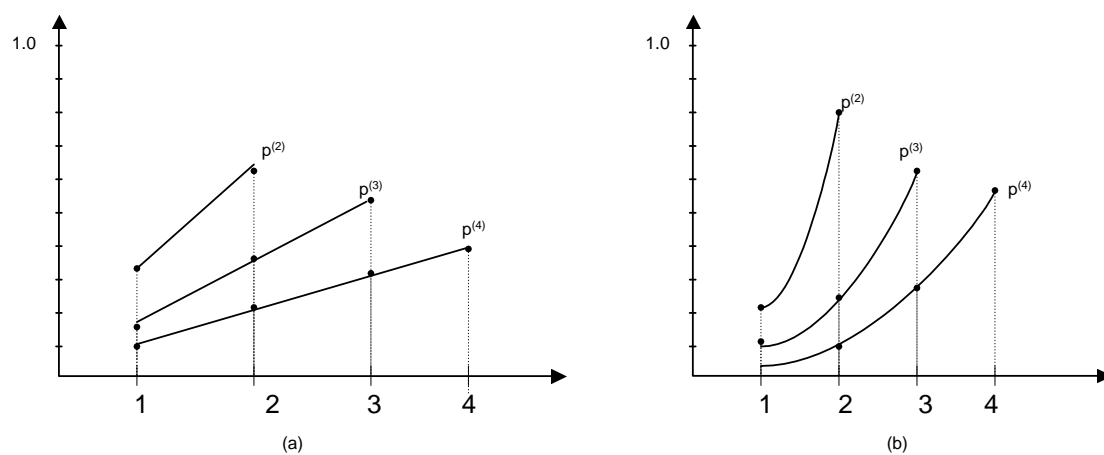


FIGURA 6.3 – Vetores OWA criados usando: (a) $g(i) = i$ (b) $g(i) = i^2$.

] **Vetores OWA criados usando: (a) $g(i) = i$ (b) $g(i) = i^2$.**

Pode-se também definir a função g analisando-se a relação entre os elementos p_i e p_j de um dado vetor P . Por exemplo, $g(i) = i$ pode ser obtido com a restrição $p_i/p_j = i/j$; conseqüentemente, neste caso a relação entre o primeiro e o último elemento do vetor é dada por $p_1/p_n = 1/n$. A função $\psi_{W^{(n)}}(A)$ comporta-se de maneira similar à entropia de

Shannon, pois ela representa o produto de 2 vetores, sendo um com valores crescentes (A ordenado) e o outro com valores decrescentes ($W^{(n)}$)¹.

A função g apresentaria alguns inconvenientes caso fosse aplicada ao vetor A (que representa a distribuição δ):

- Suponhamos que os vetores A e B relativos a m_a e m_b tenham respectivamente n_a e n_b elementos, tal que $n_a \leq n_b$. Se $a_1 = \dots = a_{n_a} = 1 = b_1 = \dots = b_{n_b}$ então $\psi_{W^{(n)}}(A) = \psi_{W^{(n)}}(B)$. Como o vetor A é menor que B , a investigação de m_a deve ser mais prioritária que m_b . Por esta razão estes vetores são distinguidos na fórmula 6.1 multiplicando-se $f_{W^{(n)}}$ por n .
- Suponhamos que os vetores A e B relativos a m_a e m_b tenham $n_a = n_b = n$ elementos, que $\psi_{W^{(n)}}(A) = \psi_{W^{(n)}}(B)$, que $\max\{a_1, \dots, a_n\} = a_{max} = 1$ e que $\max\{b_1, \dots, b_n\} = b_{max} < 1$. Como a distribuição A é normalizada e a distribuição B é subnormalizada, a investigação de A deveria ter a prioridade sobre a de B . Por esta razão, a prioridade de m_b é penalizada na fórmula 6.1 com o acréscimo de $1 - b_{max}$ ao valor $\psi_{W^{(n)}}(B) * n$. Isso também faz com que seja possível distinguir dois vetores A e B relativos a m_a e m_b quando $n_a = n_b = 1$ com $a_{max} = 1$ e $b_{max} < 1$, já que sem este acréscimo, a força de m_a seria considerada igual à de m_b .

Para ilustrar a aplicação da fórmula 6.1, consideremos as distribuições $A = [.5, 1, .2, .2]$ e $B = [1, .9, .8, .7]$ associadas às manifestações m_a e m_b (Figura 6.4). Seja $W^{(4)} = [.1, .2, .3, .4]$. A partir da definição 6.5, tem-se que $\psi_{W^{(4)}}(A) = .1 \times 1 + .2 \times .5 + .3 \times .2 + .4 \times .2 = .34$ e $\psi_{W^{(4)}}(B) = .1 \times 1 + .2 \times .9 + .3 \times .8 + .4 \times .7 = .8$. Finalmente, ao aplicarmos a fórmula 6.1 obtém-se $f_{owa}(m_a) = .34 * 4 + 0 = 1.36$ e $f_{owa}(m_b) = .8 * 4 + 0 = 3.2$. Como $f_{owa}(m_a) < f_{owa}(m_b)$, então, a investigação de m_a é considerada prioritária em relação à m_b ($m_A \prec m_B$).

6.3.2.2 Função Baseada na Medida de Não-especificidade

A segunda função (f_U) proposta é baseada na medida de incerteza U (eq. 6.2), apresentada em Klir e Folger (1988).

Definição 6.6 *Seja Ω um domínio qualquer e $A = [a_1, \dots, a_n]$ um vetor com n posições, onde $a_i \in \Omega$. Suponhamos que $\exists a_k \in A, a_k = 1$, ou seja, que a distribuição modelada*

¹A entropia de Shannon é definida como $-\sum_{i=1}^n p_i \log(p_i)$, onde $\sum_{i=1, n} p_i = 1$; se os valores de p_i forem colocados em um vetor em ordem crescente, o vetor de $-\log(p_i)$ correspondente terá seus valores em ordem decrescente.

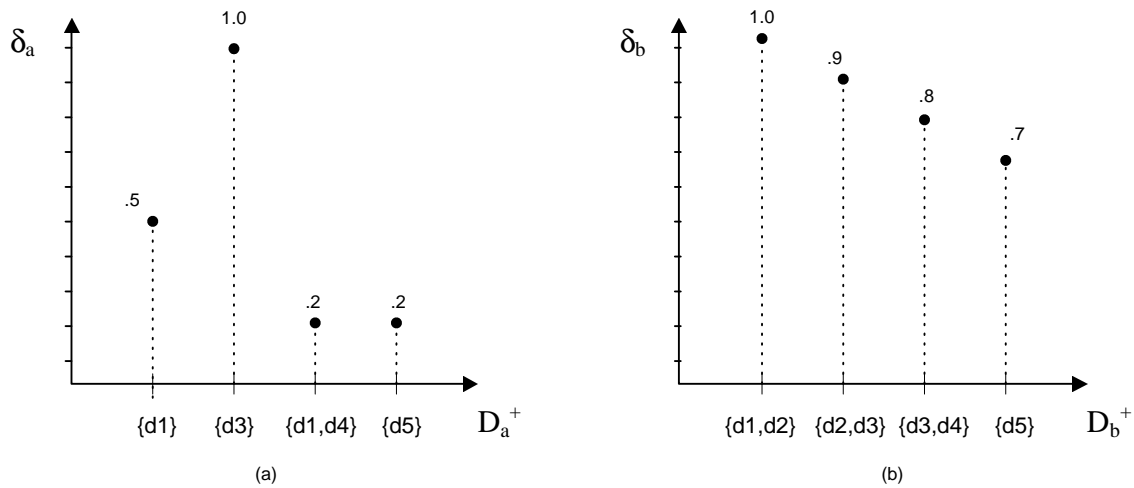


FIGURA 6.4 – Exemplos de distribuições δ .

por A é normalizada em termos possibilísticos. Seja $T(A) = \{\rho \mid \exists a_k = \rho, a_k \in A\}$ o conjunto dos valores dos elementos de A (imagem de A). Seja $\langle \sigma_1, \dots, \sigma_q \rangle$, $q = |T(A)|$ a seqüência obtida através da ordenação decrescente dos valores em $T(A)$; portanto, $\sigma_1 = 1$. Seja $\sigma_{q+1} = 0$. Então, temos

$$U(A) = \sum_{1 \leq i \leq n} (\sigma_i - \sigma_{i+1}) \log |A_i|, \quad (6.2)$$

onde $A_i = \{a_k \mid a_k \geq \sigma_i\}$.

Klir e Folger (1988) mostram que a equação 6.2 atende algumas propriedades básicas como: expansibilidade, subaditividade, aditividade, continuidade, monotocidade, mínimo, máximo, *branching* e normalização. Para ilustrar a aplicação da fórmula 6.2, consideremos novamente as distribuições de valores apresentadas na Figura 6.4. Para $A = [.5, 1, .2, .2]$ temos $\sigma = \langle 1, .5, .2, 0 \rangle$ e $A_1 = \{\{d_3\}\}$, $A_2 = \{\{d_3\}, \{d_1\}\}$, $A_3 = \{\{d_3\}, \{d_1\}, \{d_1, d_4\}, \{d_5\}\}$. Para $B = [1, .9, .8, .7]$ temos $\sigma = \langle 1, .9, .8, .7, 0 \rangle$ e $A_1 = \{\{d_1, d_2\}\}$, $A_2 = \{\{d_1, d_2\}, \{d_2, d_3\}\}$, $A_3 = \{\{d_1, d_2\}, \{d_2, d_3\}, \{d_3, d_4\}\}$, $A_4 = \{\{d_1, d_2\}, \{d_2, d_3\}, \{d_3, d_4\}, \{d_5\}\}$.

A partir da definição 6.6, aplicando-se a fórmula 6.2 obtemos $U(A) = .5 \log 1 + .3 \log 2 + .2 \log 4 = 0 + .3 + .4 = .7$ e $U(B) = .1 \log 1 + .1 \log 2 + .1 \log 3 + .7 \log 4 = 0 + .1 + .17 + 1.4 = 1.67$. Portanto, temos $m_a \prec m_b$. No exemplo acima, todas as distribuições es ão normalizadas. Como isso nem sempre ocorre, como será visto adiante, não é possível a utilização de $U(A)$ obtido como mostrado acima diretamente para obter a força de uma manifestação m_j qualquer, sendo necessária uma adaptação.

Suponhamos que queiramos utilizar a medida de não especificidade acima no cálculo da força de m_j . Suponhamos que cada $a_k \in]0, 1]$ em A_j é tal que $a_k = \delta_j(D_k)$, $D_k \in D_j^+$ e $n = |D_j^+|$. Para isso, primeiramente, transformamos a distribuição de valores em A_j em uma distribuição normalizada $A'_j = [a_1/s, \dots, a_n/s]$, com $s = \sum_{1 \leq k \leq n} a_k$ e depois acrescentamos o valor $(1 - \max\{a_1, \dots, a_n\})$.

Definição 6.7 A função $f_U(A_j)$ especifica a força de m_j para investigação através da fórmula a seguir:

$$f_U(A_j) = U(A'_j) + (1 - \max\{a_1, \dots, a_n\}) \quad (6.3)$$

Três casos especiais na aplicação da fórmula 6.3 são:

- Se $\sigma_1 = \dots = \sigma_q = 1$ então $f_U(A_j) = U(A'_j) = \log q$;
- Se $\sigma_1 = 1, \sigma_2 = \dots = \sigma_q = \alpha$ então $f_U(A_j) = U(A'_j) = \alpha \log q$;
- Se $\sigma_1 = \dots = \sigma_q = \alpha$ então $U(A'_j) = \log q$ e $f_U(A_j) = (1 - \alpha) + \log q$.

6.3.2.3 Comparação das Funções

Consideremos a aplicação das fórmulas 6.1 (f_{owa}) e 6.3 (f_U) para as distribuições apresentadas na Figura 6.5. No caso desta Figura, todas as distribuições $A_i, 1 \leq i \leq 7$, estão normalizadas, no sentido possibilístico do termo ($\exists \omega \in \Omega, A(\omega) = 1$). Pode-se observar na tabela 6.2 que as duas funções geram praticamente a mesma classificação; para ambas funções temos $m_i \preceq m_{i+1}$, onde $m_i \preceq m_{i+1}$ se a entropia de A_i for menor que a entropia de A_{i+1} . No entanto, a função f_U classifica igualmente m_4 e m_5 , enquanto que f_{owa} é capaz de distingui-las.

TABELA 6.2 – Comparação das medidas f_{owa} e f_U para distribuições normalizadas.

A_k	f_{owa}	f_U
A_1	0	0
A_2	.76	.2
A_3	1.65	.75
A_4	2.0	1.0
A_5	2.2	1.0
A_6	3.6	1.8
A_7	4.0	2.0

Na Figura 6.6 são apresentados exemplos de distribuições não normalizadas. Nas colunas 2 e 3 da tabela 6.3 são apresentados os resultados de f_{owa} e f_U para cada distribuição

da Figura 6.6. Também neste caso, pode-se observar que as duas funções têm o mesmo comportamento para distribuições não normalizadas com $m_i \preceq m_{i+1}$, $8 \leq i \leq 9$.

TABELA 6.3 – Comparação das medidas f_{owa} e f_U para distribuições não normalizadas.

A_k	f_{owa}	f_U
A_8	1.3	.5
A_9	1.4	.6
A_{10}	1.8	1.0

6.3.3 Modelo das Distribuições

Apresentadas as funções de entropia a serem aplicadas em distribuições de valores, resta ainda definir o que representam exatamente estas distribuições. As definições seguintes apresentam duas possíveis alternativas.

Definição 6.8 *Uma distribuição de valores $\delta_j : D_j^+ \rightarrow [0, 1]$, que indica o quanto $m_j \in \text{efeitos}(D^+)$, tal que $m_j \notin M^+ \cup M^-$, é importante para uma dada cobertura $D_k \in D_j^+$, pode ser calculada como:*

$$(a) \delta_j(D_k) = \perp_{d_i \in D_k} R(d_i, m_j),$$

$$(b) \delta_j^+(D_k) = \perp_{d_i \in D_k} \top(R(d_i, m_j), \alpha_j^+(d_i, m_j)),$$

$$\delta_j^-(D_k) = \perp_{d_i \in D_k} \top(R(d_i, m_j), \alpha_j^-(d_i, m_j))$$

onde $R(d_i, m_j)$ refere-se ao grau de associação causal nebulosa de cada desordem d_i para cada manifestação m_j (tabela 5.1), $\alpha_j^+(d_i)$ refere-se ao grau de consistência temporal de d_i dado que a presença da manifestação m_j é confirmada e $\alpha_j^-(d_i)$ refere-se ao grau de consistência temporal de d_i dado que a ausência da manifestação m_j é confirmada, \top é uma t-norma e \perp é uma t-conorma.

No caso (b), as entropias de δ_j^+ e δ_j^- são posteriormente agregadas através da fórmula 6.4:

$$f_j = \top(f(\delta_j^+), f(\delta_j^-)) \quad (6.4)$$

onde \top é uma t-norma e $f(\cdot)$ refere-se a uma das funções de entropia f_{owa} ou f_U .

Exemplos de t-conormas que podem ser utilizadas para modelar \perp são, por exemplo, $\perp'(a, b) = \max(a, b)$ e $\perp''(a, b) = (a + b - ab)$ que são as duais das t-normas \min e produto $(*)$, respectivamente. Por exemplo, sejam as distribuições $A_1 = [.5, .7]$ e

$A_2 = [.3, .7]$. Para A_1 temos então $\perp'(.5, .7) = .7$ e $\perp''(.5, .7) = .85$. Para A_2 temos $\perp'(.5, .7) = .7$ e $\perp''(.5, .7) = .79$. Neste trabalho, foi adotada a t-conorma \perp'' porque ela produz resultados mais diferenciados do que a t-conorma max usualmente empregada em aplicações envolvendo conjuntos nebulosos. Pelo mesmo motivo, neste trabalho, utiliza-se a t-norma produto (*) para modelar \top ao invés da t-norma usual min, tanto na definição 6.8(b) quanto na fórmula 6.4. Maiores detalhes das definições de t-norma e t-conorma podem ser encontradas no apêndice A.

Suponhamos que durante o processo de diagnóstico, o fitopatologista tenha o seguinte conjunto de hipóteses iniciais $D^+ = \{\{d_1, d_2\}\}$ para um caso no qual $M^+ = \{m_1, m_2, m_3\}$, $TIME^+(m_1^b) = \langle 6, 7, 8, 9 \rangle$, $TIME^+(m_2^b) = \langle 8, 9, 10, 11 \rangle$, e $TIME^+(m_3^b) = \langle 16, 17, 18, 19 \rangle$.

A partir da KB_e temos $efeitos(d_1) = \{m_1, m_2, m_4, m_5\}$ e $efeitos(d_2) = \{m_1, m_3, m_4, m_5\}$. Neste caso, as manifestações $M^* = \{m_4, m_5\}$, precisam ser investigadas (por exemplo através de testes laboratoriais) para que um diagnóstico correto possa ser formulado. Suponhamos que:

$$D_4^+ = \{\{d_1, d_2\}\} \text{ com } R(d_1, m_4) = .75, R(d_2, m_4) = .75,$$

$$D_5^+ = \{\{d_1, d_2\}\} \text{ com } R(d_1, m_5) = .5, R(d_2, m_5) = .5.$$

No caso (a), considera-se somente o grau de associação causal $R(d_i, m_j)$ para calcular δ_j e, em seguida, calcula-se a entropia de m_j , através de $f_{owa}(\delta_j)$ ou de $f_U(\delta_j)$.

Para utilizar a definição 6.8(a) no exemplo, temos pois que considerar somente as relações causais $R(d_i, m_j)$ para a cobertura $\{d_1, d_2\}$. Obtemos então a seguinte distribuição: $\delta_4(\{d_1, d_2\}) = 0.75+0.75-(0.75*0.75) = .9325$ e $\delta_5(\{d_1, d_2\}) = 0.5+0.5-(0.5*0.5) = .75$. Finalmente, poderíamos calcular a entropia final para m_4 e m_5 aplicando uma função de entropia (f_{owa} ou f_U) a δ_4 e δ_5 .

Para utilizar a definição 6.8(b) no exemplo, não se considera apenas a associação causal $R(d_i, m_j)$, mas também a consistência temporal entre os modelos e o caso. A consistência hipotética α_j^+ é obtida supondo-se a confirmação da presença de m_j enquanto a consistência hipotética α_j^- é obtida supondo-se a confirmação da ausência de m_j . Conforme descrito na definição 6.8(b), a partir da agregação de $R(d_i, m_j)$ com α_j^+ e α_j^- gera-se as distribuições δ_j^+ e δ_j^- , respectivamente.

Por exemplo, caso a presença ou ausência de m_4 e m_5 sejam confirmadas, obtém-se a seguinte consistência temporal: $\alpha_4^+(\{d_1, d_2\}) = \min(\alpha_4^+(d_1), \alpha_4^+(d_2)) = .7$, $\alpha_4^-(\{d_1, d_2\}) =$

$$\min(\alpha_4^-(d_1), \alpha_4^-(d_2)) = .6, \alpha_5^+(\{d_1, d_2\}) = \min(\alpha_5^+(d_1), \alpha_5^+(d_2)) = .5, \text{ e } \alpha_5^-(\{d_1, d_2\}) = \min(\alpha_5^-(d_1), \alpha_5^-(d_2)) = .6.$$

Na definição 6.8(b), temos que utilizar a t-norma \top para agregar as relações causais com a consistência temporal. Seja $\top_{i,j}^+ = (R(d_i, m_j) * \alpha_j^+(d_i))$ e $\top_{i,j}^- = (R(d_i, m_j) * \alpha_j^-(d_i))$. Assim, $\top_{1,4}^+ = .75 * .7 = .52$, $\top_{2,4}^+ = .75 * .7 = .52$, $\top_{1,4}^- = .75 * .6 = .45$, $\top_{2,4}^- = .75 * .6 = .45$, $\top_{1,5}^+ = .5 * .5 = .25$, $\top_{2,5}^+ = .5 * .5 = .25$, $\top_{1,5}^- = .5 * .6 = .3$ e $\top_{2,5}^- = .5 * .6 = .3$.

Em seguida, como a cobertura é $\{d_1, d_2\}$ utilizamos a t-conorma \perp para agregar d_1 e d_2 . Temos então $\delta_4^+(\{d_1, d_2\}) = 0.52 + 0.52 - (0.52 * 0.52) = 1.04 - .27 = .77$, $\delta_4^-(\{d_1, d_2\}) = 0.45 + 0.45 - (0.45 * 0.45) = .90 - .20 = .7$, $\delta_5^+(\{d_1, d_2\}) = 0.25 + 0.25 - (0.25 * 0.25) = .44$, $\delta_5^-(\{d_1, d_2\}) = 0.3 + 0.3 - (0.3 * 0.3) = .51$.

Finalmente, pode-se calcular a entropia final para m_4 e m_5 através da fórmula 6.4 (seção 6.3.3):

$$f(m_4) = (f(\delta_4^+) * f(\delta_4^-))$$

$$f(m_5) = (f(\delta_5^+) * f(\delta_5^-))$$

onde f refere-se as funções de entropia f_{owa} ou f_U .

6.3.4 Estudo de Caso

Nesta seção está descrito um exemplo de hipótese diagnóstica no domínio de doenças de milho onde serão aplicadas as fórmulas 6.1 e 6.3 apresentadas nas seções 6.3.2.1 e 6.3.2.2.

Para ilustrar as funções que calculam a importância de uma manifestação no contexto de investigação, consideremos novamente como exemplo as manifestações causadas pelo fungo *Ferrugem Comum*. Estas manifestações são muito similares às manifestações causadas por outros fungos como *Ferrugem Polysora* (d_4) e *Ferrugem Branca* (d_5).

Suponhamos que durante o processo de diagnóstico, o fitopatologista tenha o seguinte conjunto de hipóteses iniciais $D^+ = \{\{d_3\}, \{d_4\}\}$, onde $\alpha(\{d_3\}) = \alpha(\{d_4\}) = 0.75$ para um caso no qual $M^+ \cup M^- = \{m_1, m_3\}$, $TIME^+(m_1^b) = \langle 6, 7, 8, 9 \rangle$ e $TIME^+(m_3^b) = \langle 16, 17, 18, 19 \rangle$ (o modelo de d_3 e d_4 estão na Figura 5.2). Dado que $efeitos(d_3) = \{m_1, m_2, m_3, m_4, m_5\}$ e $efeitos(d_4) = \{m_1, m_3, m_4, m_5, m_6\}$, as manifestações $M^* = \{m_2, m_4, m_5, m_6\}$, precisam ser investigadas (por exemplo através de testes laboratoriais) para que um diagnóstico correto possa ser formulado. A partir da KB_e temos que:

- $D_2^+ = \{d_3\}$ com $R(d_3, m_2) = .5$.
- $D_4^+ = \{d_3, d_4\}$ com $R(d_3, m_4) = .75$, $R(d_4, m_4) = .75$.
- $D_5^+ = \{d_3, d_4\}$ com $R(d_3, m_5) = .5$, $R(d_4, m_5) = .5$.
- $D_6^+ = \{d_4\}$ com $R(d_4, m_6) = .5$.

Para cada fórmula serão apresentados dois casos, considerando-se as definições 6.8(a) e 6.8(b). No caso (a) considera-se somente o grau de associação causal $R(d_i, m_j)$ enquanto que no caso (b) considera-se este grau mas também o grau de consistência temporal.

No caso(a), utilizando-se a definição 6.8(a), obtemos então a seguinte distribuição de valores: $\delta_2(\{d_3\}) = 0.5$, $\delta_4(\{d_3\}) = 0.75$, $\delta_4(\{d_4\}) = 0.75$, $\delta_5(\{d_3\}) = 0.5$, $\delta_5(\{d_4\}) = 0.5$ e $\delta_6(\{d_4\}) = 0.5$.

Para utilizar a definição 6.8(b), necessitamos também levar em conta a consistência temporal entre os modelos e o caso, dado que a presença ou ausência de uma manifestação m_j é confirmada (m_j^+ , m_j^-).

Por exemplo, seja $\alpha_2^+(d_3) = .5$, $\alpha_4^+(d_3) = \alpha_4^+(d_4) = .75$, $\alpha_5^+(d_3) = \alpha_5^+(d_4) = .75$, $\alpha_6^+(d_4) = .5$, $\alpha_4^-(d_3) = \alpha_4^-(d_4) = .75$, $\alpha_5^-(d_3) = \alpha_5^-(d_4) = .75$, e $\alpha_6^-(d_4) = .5$.

No caso (b), a partir da definição 6.8(b) obtemos então as seguintes distribuições para δ_j^+ (i.e. presença de m_j confirmada): $\delta_2^+(\{d_3\}) = 0.5 * 0.5 = 0.25$, $\delta_4^+(\{d_3\}) = 0.75 * 0.75 = 0.5625$, $\delta_4^+(\{d_4\}) = 0.75 * 0.75 = 0.5625$, $\delta_5^+(\{d_3\}) = 0.5 * 0.75 = 0.375$, $\delta_5^+(\{d_4\}) = 0.5 * 0.75 = 0.375$ e $\delta_6^+(\{d_4\}) = 0.5 * 0.5 = 0.25$.

Ainda no caso (b), a partir da definição 6.8(b) temos então as seguintes distribuições para δ_j^- (i.e. ausência de m_j confirmada): $\delta_2^-(\{d_3\}) = 0.5*0.75 = 0.375$, $\delta_2^-(\{d_4\}) = 0*0.75 = 0$, $\delta_4^-(\{d_3\}) = 0.75*0.75 = 0.5625$, $\delta_4^-(\{d_4\}) = 0.75*0.75 = 0.5625$, $\delta_5^-(\{d_3\}) = 0.5*0.75 = 0.375$, $\delta_5^-(\{d_4\}) = 0.5*0.75 = 0.375$, $\delta_6^-(\{d_4\}) = 0.5*0.75 = 0.375$ e $\delta_6^-(\{d_3\}) = 0*0.75 = 0$.

Neste exemplo, como as coberturas são conjuntos unitários ($\{d_3\}$ e $\{d_4\}$), não foi necessária a utilização de uma t-conorma para agregar os resultados. A partir das distribuições acima, calcula-se a entropia de $\{m_2, m_4, m_5, m_6\}$ através das funções f_{owa} e f_U .

6.3.4.1 Método f_{owa}

Neste exemplo será apresentado como selecionar a manifestação prioritária utilizando a função f_{owa} apresentada na seção 6.3.2.1, onde a manifestação com menor grau deve ser investigada prioritariamente. A fórmula 6.1 pode ser utilizada como exemplificado abaixo para selecionar as manifestações prioritárias a serem investigadas, para os casos (a) e (b), considerando-se as definições 6.8(a) e 6.8(b), respectivamente.

Sejam os vetores de pesos dados por $W^{(2)} = [.33, .67]$ e $W^{(3)} = [.15, .35, .5]$. Aplicando-se a definição 6.5, obtém-se que $\psi_{W^{(1)}}(A_2) = \psi_{W^{(1)}}([.5]) = \psi_{W^{(1)}}([1]) = 0$, $\psi_{W^{(2)}}(A_4) = \psi_{W^{(2)}}([.75, .75]) = \psi_{W^{(2)}}([1, 1]) = (1 * .33) + (1 * .67) = .33 + .67 = 1.0$, $\psi_{W^{(2)}}(A_5) = \psi_{W^{(2)}}([.5, .5]) = \psi_{W^{(2)}}([1, 1]) = 1.0$ e $\psi_{W^{(1)}}(A_6) = \psi_{W^{(1)}}([.5]) = 0$.

Finalmente, ao aplicar-se a fórmula 6.1 tem-se que $f_{owa}(A_2) = 0 + .5 = .5$, $f_{owa}(A_4) = (1.0) * 2 + .25 = 2.0 + .25 = 2.25$, $f_{owa}(A_5) = (1.0) * 2 + .5 = 2.0 + .5 = 2.5$ e $f_{owa}(A_6) = 0 + .5 = .5$.

Assim, os exames que identificam as causas das manifestações em M^* na KB_e , considerando-se somente as associações causais, devem ser solicitados na ordem $\langle m_2, m_6, m_4, m_5 \rangle$, visando confirmar ou refutar as hipóteses que causam estes efeitos.

Neste caso, portanto, a investigação deve começar pelas manifestações m_2 ou m_6 . Nos modelos de $SOL_d = \{d_3, d_4\}$ (vide Figura 5.2), observa-se que, embora m_2 e m_6 tenha uma baixa associação causal (0.5) com as desordens d_3 e d_4 , m_2 e m_6 são muito específicas ($|D_2^+| = 1, |D_6^+| = 1$), fazendo com que a força de m_2 e m_6 seja maior que a força de m_4 e m_5 . Por outro lado, a força de m_4 é maior que a força de m_5 , pois embora m_5 e m_4 apareçam no mesmo número de desordens da hipótese, m_4 tem uma associação causal mais alta para as desordens.

No caso (b), supondo-se a presença da manifestação, obtém-se $\psi_{W^{(1)}}(\delta_2^+) = \psi_{W^{(1)}}([.25]) = \psi_{W^{(1)}}([1]) = 0$, $\psi_{W^{(2)}}(\delta_4^+) = \psi_{W^{(2)}}([.5625, .5625]) = \psi_{W^{(2)}}([1, 1]) = 1.0$, $\psi_{W^{(2)}}(\delta_5^+) = \psi_{W^{(2)}}([.375, .375]) = \psi_{W^{(2)}}([1, 1]) = 1.0$ e $\psi_{W^{(1)}}(\delta_6^+) = \psi_{W^{(1)}}([.25]) = \psi_{W^{(1)}}([1]) = 0$.

Em seguida, ao aplicar-se a fórmula 6.1 tem-se que $f_{owa}(A_2) = 0 + .75 = .75$, $f_{owa}(\delta_4^+) = (1.0) * 2 + .4375 = 2.4375$, $f_{owa}(\delta_5^+) = (1.0) * 2 + .625 = 2.625$ e $f_{owa}(\delta_6^+) = 0 + .75 = .75$.

Ainda no caso(b), supondo-se a ausência das manifestações, obtém-se: $\psi_{W^{(1)}}(\delta_2^-) = \psi_{W^{(1)}}([.375]) = \psi_{W^{(1)}}([1]) = 0$, $\psi_{W^{(2)}}(\delta_4^-) = \psi_{W^{(2)}}([.5625, .5625]) = \psi_{W^{(2)}}([1, 1]) = 1.0$, $\psi_{W^{(2)}}(\delta_5^-) = \psi_{W^{(2)}}([.375, .375]) = \psi_{W^{(2)}}([1, 1]) = 1.0$ e $\psi_{W^{(1)}}(\delta_6^-) = \psi_{W^{(1)}}([.375]) = \psi_{W^{(1)}}([1]) = 0$.

Em seguida, ao aplicar-se a fórmula 6.1 tem-se que $f_{owa}(\delta_2^-) = 0 + .625 = .625$, $f_{owa}(\delta_4^-) = (1.0) * 2 + .4375 = 2.4375$, $f_{owa}(\delta_5^-) = (1.0) * 2 + .625 = 2.625$ e $f_{owa}(\delta_6^-) = 0 + .75 = .625$.

Finalmente, a entropia final de m_j é calculada através da fórmula 6.4 como apresentada na tabela 6.4.

TABELA 6.4 – Cálculo da entropia f_{owa} para manifestações m_j .

m_j	$f_{owa}(\delta_j^+)$	$f_{owa}(\delta_j^-)$	$\top(f_{owa}(\delta_j^+), f_{owa}(\delta_j^-))$
m_2	.75	.625	.468
m_4	2.4375	2.4375	5.94
m_5	2.625	2.625	6.89
m_6	.75	.625	.468

Assim, os exames que identificam as causas das manifestações em M^* na KB_e , considerando-se tanto as associações causais quanto a consistência temporal, devem ser solicitados na ordem $\langle m_2, m_6, m_4, m_5 \rangle$, visando confirmar ou refutar as hipóteses que causam estes efeitos. Observa-se, neste caso, que não houve alteração nos resultados obtidos na utilização das definições 6.8(a) e 6.8(b), considerando-se o f_{owa} .

6.3.4.2 Método f_U

Neste exemplo será apresentado como selecionar a manifestação prioritária utilizando a função f_U apresentada na seção 6.3.2.2, onde a manifestação com menor grau deve ser investigada prioritariamente. Conforme apresentado na seção 6.3.4, tem-se que as manifestações $\{m_2, m_4, m_5, m_6\}$ devem ser investigadas de modo a auxiliar no diagnóstico.

Para ilustrar a aplicação da fórmula 6.2 para o mesmo exemplo apresentado na seção anterior, consideremos novamente as distribuições de valores do caso(a) (veja seção 6.3.4): $\delta_2(\{d_3\}) = .5$, $\delta_4(\{d_3\}) = .75$, $\delta_4(\{d_4\}) = .75$, $\delta_5(\{d_3\}) = .5$, $\delta_5(\{d_4\}) = .5$ e $\delta_6(\{d_4\}) = .5$.

A partir da definição 6.6, para $A_2 = [.5]$, $A_4 = [.75, .75]$, $A_5 = [.5, .5]$ e $A_6 = [.5]$. Após a normalização temos $A'_2 = [1]$, $A'_4 = [1, 1]$, $A'_5 = [1, 1]$ e $A'_6 = [1]$. Aplicando-se a fórmula 6.2 obtemos $U(A'_2) = 1 \log 1 = 0$, $U(A'_4) = 1 \log 2 = 1$, $U(A'_5) = 1 \log 2 = 1$ e $U(A'_6) = 1 \log 1 = 0$. Finalmente, a partir da definição 6.3 temos $f_U(A_2) = 0 + (1 - .5) = .5$, $f_U(A_4) = 1 + (1 - .75) = 1.25$, $f_U(A_5) = 1 + (1 - .5) = 1.5$ e $f_U(A_6) = 0 + (1 - .5) = .5$. Finalmente, temos que $m_2 \preceq m_6 \prec m_4 \prec m_5$.

No caso (b) temos então a seguinte distribuição de valores para m_j^+ (veja seção 6.3.4): $\delta_2^+(\{d_3\}) = .25$, $\delta_4^+(\{d_3\}) = .5625$, $\delta_4^+(\{d_4\}) = .5625$, $\delta_5^+(\{d_3\}) = .375$, $\delta_5^+(\{d_4\}) = .375$ e $\delta_6^+(\{d_4\}) = .25$.

Para $A_2 = [.25]$, $A_4 = [.5625, .5625]$, $A_5 = [.375, .375]$ e $A_6 = [.25]$, após a normalização $A'_2 = [1]$, $A'_4 = [1, 1]$, $A'_5 = [1, 1]$ e $A'_6 = [1]$, aplicando-se a fórmula 6.2, obtemos $U(A'_2) = 1 * \log 1 = 0$, $U(A'_4) = 1 * \log 2 = 1$, $U(A'_5) = 1 * \log 2 = 1$ e $U(A'_6) = 1 * \log 1 = 0$. Finalmente, a partir da definição 6.3 temos $f_U(A_2) = 0 + (1 - .25) = .75$, $f_U(A_4) = 1 + (1 - .5625) = 1.4375$, $f_U(A_5) = 1 + (1 - .375) = 1.625$ e $f_U(A_6) = 0 + (1 - .25) = .75$.

Caso a ausência de m_j seja confirmada temos então a seguinte distribuição (veja seção 6.3.4): $\delta_2^-(\{d_3\}) = .375$, $\delta_2^-(\{d_4\}) = 0$, $\delta_4^-(\{d_3\}) = .5625$, $\delta_4^-(\{d_4\}) = .5625$, $\delta_5^-(\{d_3\}) = .375$, $\delta_5^-(\{d_4\}) = .375$ e $\delta_6^-(\{d_3\}) = 0$.

Para $A_2 = [.375]$, $A_4 = [.5625, .5625]$, $A_5 = [.375, .375]$ e $A_6 = [.375]$, após a normalização $A_2 = [1]$, $A_4 = [1, 1]$, $A_5 = [1, 1]$ e $A_6 = [1]$, aplicando-se a fórmula 6.2 obtemos $U(A'_2) = 1 * \log 1 = 0$, $U(A'_4) = 1 * \log 2 = 1$, $U(A'_5) = 1 * \log 2 = 1$ e $U(A'_6) = 1 * \log 1 = 0$. Finalmente, a partir da definição 6.3 temos $f_U(A_2) = 0 + (1 - .375) = .625$, $f_U(A_4) = 1 + (1 - .5625) = 1.4375$, $f_U(A_5) = 1 + (1 - .375) = 1.625$ e $f_U(A_6) = 0 + (1 - .375) = .625$. Assim, a entropia final de m_j é calculada através da fórmula 6.4 como apresentada na tabela 6.5.

TABELA 6.5 – Cálculo da entropia f_U para manifestações m_j .

m_j	$f_U(\delta_j^+)$	$f_U(\delta_j^-)$	$\top(f_U(\delta_j^+), f_U(\delta_j^-))$
m_2	.75	.625	.468
m_4	1.4375	1.4375	2.066
m_5	1.625	1.625	2.68
m_6	.75	.625	.468

Finalmente, temos que $m_2 \preceq m_6 \prec m_4 \prec m_5$. Assim, os exames laboratoriais (E) devem ser solicitados na ordem $\{exames(m_2), exames(m_6), exames(m_4), exames(m_5)\}$ na KB_e visando confirmar ou refutar as hipóteses que causam estes efeitos. Portanto, a investigação deve começar pelas manifestações m_2 ou m_6 assim como apresentado no método f_{owa} .

6.3.4.3 Análise Comparativa dos Resultados Obtidos

Pode-se observar que as duas funções f_{owa} e f_U geraram a mesma classificação no caso (a) e no caso (b); para ambas funções temos $m_i \preceq m_{i+1}$, onde $m_i \preceq m_{i+1}$ se a entropia de A_i for menor que a entropia de A_{i+1} . No entanto, a função f_U pode classificar igualmente algumas manifestações, como m_4 e m_5 na tabela 6.2, enquanto que f_{owa} foi capaz de distingui-las.

Quanto a distribuição de valores (δ_j) , no caso (a) em que se considerou somente o grau de associação causal $(R(d_i, m_j))$, os cálculos são mais simples. Enquanto que no caso (b) em que se considerou o grau de associação causal e o grau de consistência temporal dado que a presença ou ausência de uma manifestação m_j é confirmada (m_j^+, m_j^-) é mais trabalhoso. Além disso, o grau da consistência temporal apresentado no caso (b) é hipotético e os intervalos temporais assumidos para cada manifestação são aleatórios.

Dado que uma manifestação m_j é confirmada, os valores temporais considerados estão no intervalo do melhor caso. Adicionalmente, poderia ser necessário calcular para cada manifestação a ser investigada todos os intervalos temporais possíveis. Diante dos resultados obtidos, optou-se por utilizar a função de entropia f_{owa} no contexto de múltiplas desordens como apresentada na seção 6.3.2.1 e pela distribuição de valores (δ) como apresentada no caso (a) pela maior simplicidade nos cálculos (veja no apêndice B).

6.3.5 Solução para Investigação

A partir das idéias acima, pode-se definir formalmente a seleção de exames ou testes laboratoriais para investigação.

Definição 6.9 (Solução abdutiva) *Dado uma base de conhecimento $KB_e = \langle E, I, D, M, R \rangle$ e uma hipótese $D_l \in SOL_d$ para um conjunto de observações M^+ tal que $\alpha(D_l) < \Delta$, SOL_e , é uma solução abdutiva a partir de E para M^+ **ordenadas** sse $I \cup SOL_e$ cobre M^+ e $I \cup SOL_e$ é consistente.*

A partir de SOL_e um conjunto de manifestações são confirmadas ou refutadas, consequentemente, atualizando os conjuntos de manifestações de M^+ e M^- que pertencem a CA gerando um novo caso C_{atual} . Dado a base de conhecimento KB_d e o caso particular C_{atual} , infere-se um nova explicação abdutiva SOL_d a partir da definição 5.9. Para toda $D_l \in SOL_d$ onde $\alpha(D_l) > \Delta$ pode-se inferir os melhores tratamentos através de um ambiente abduativo para tratamentos como será apresentado no próximo capítulo.

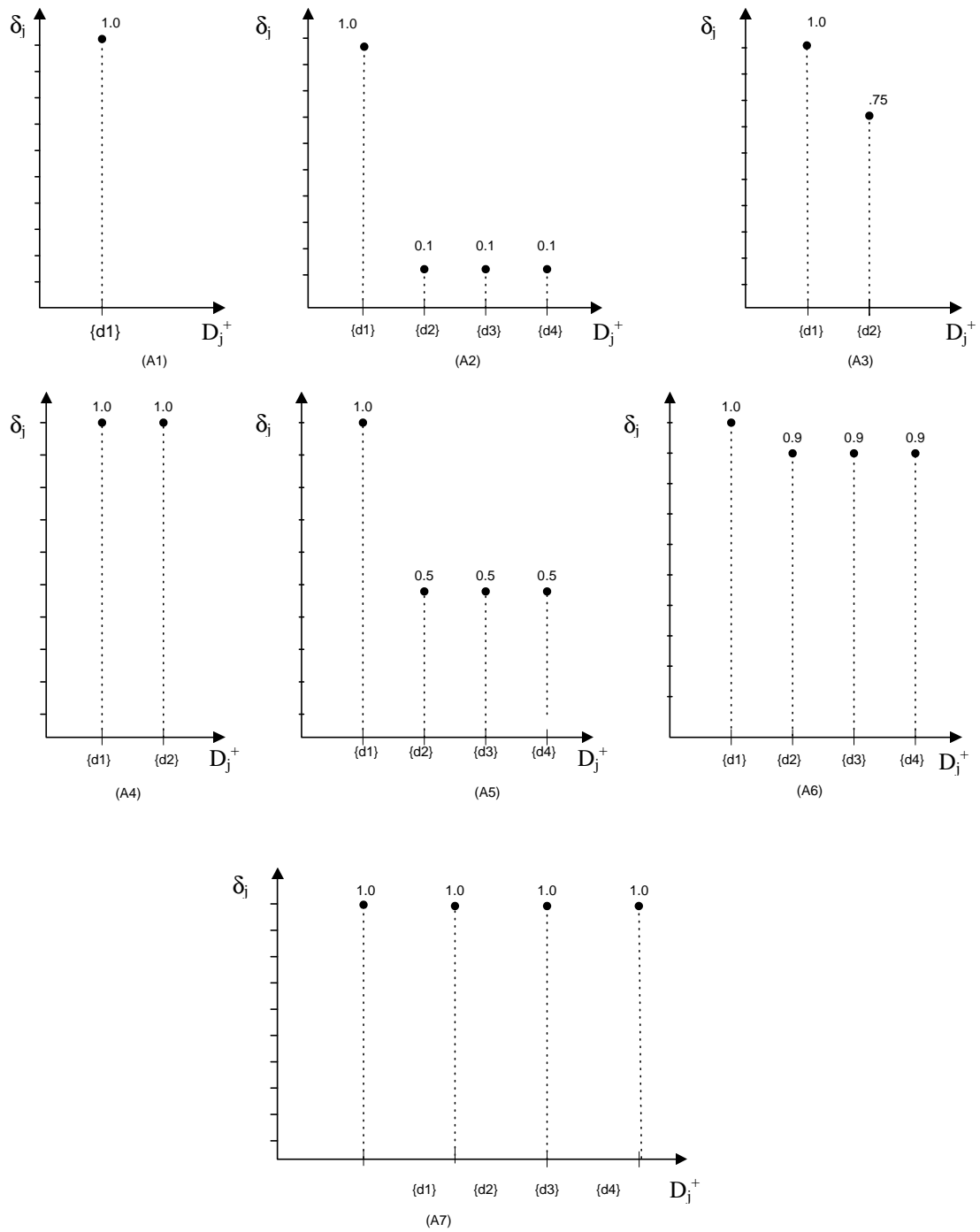


FIGURA 6.5 – Distribuições δ normalizadas.

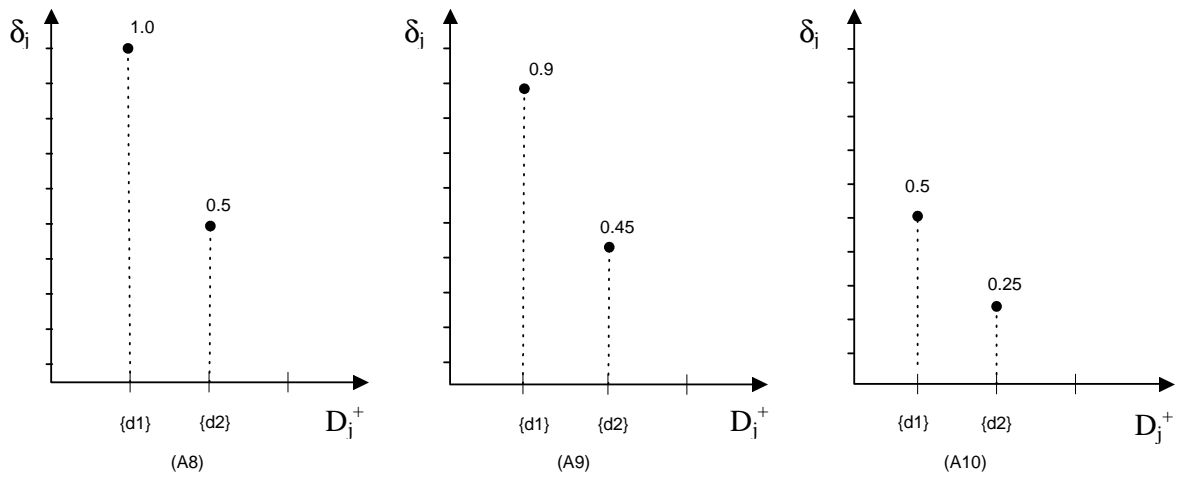


FIGURA 6.6 – Distribuições δ não normalizadas.

CAPÍTULO 7

AMBIENTE ABDUTIVO PARA TRATAMENTOS

Neste capítulo é apresentado detalhadamente o ambiente abduativo de tratamentos da TCN que incorpora conceitos de inferência abduativa, lógica nebulosa e teoria de decisão no escopo da TCN. Primeiramente, na seções 7.1 e 7.2 são apresentados respectivamente a representação do conhecimento dos módulos de tratamentos e dos casos particulares. Nas seção 7.3 são apresentados os mecanismos de inferência e tomada de decisão utilizados para inferir os tratamentos mais adequados para um dado caso particular.

7.1 Representação do Conhecimento para Tratamentos

O conhecimento do especialista ou diagnosticador em um dado domínio não se resume apenas em fazer classificações corretas mas também se baseia no mapeamento correto do diagnóstico em tratamentos efetivos.

No modelo proposto no escopo da TCN, modela-se a relação entre os conjuntos partindo dos tratamentos para as desordens porque é o tratamento que influi na desordem (ex. *tratamentos* \rightarrow *desordens*). Como o que se tem em um caso particular é uma desordem, a especificação de um tratamento adequado caracteriza então um raciocínio do tipo abduativo. (veja Figura 7.1)

O subproblema de tratamentos pode ser modelado por uma quintupla $\langle T, D, Z, CC, V \rangle$ onde: T é um conjunto de tratamentos; D, é o conjunto de desordens que são gerados como hipóteses; e $Z \subseteq T \times D$, é a teoria que associa o conjunto de desordens com os tratamentos que as controlam. Uma dupla $\langle t_k, d_i \rangle$ em Z significa que t_k pode tratar diretamente d_i . Adicionalmente, cada dupla $\langle t_k, d_i \rangle$ em Z está associada à um conjunto de condições de controle CC e um conjunto de variáveis V. As condições de controle CC devem ser satisfeitas pelo caso para que um tratamento t_k para uma desordem d_i possa ser garantido. O conjunto V auxiliará os especialistas (médicos, fitopatologistas ou técnicos) na tomada de decisão para escolher o melhor tratamento para uma desordem.

No escopo da TCN é proposto representar as condições de controle de cada dupla $Z = \langle t_k, d_i \rangle$ através de um conjunto $CC = \{EP, SEV, CF\}$, como descrito na definição 7.1.

Definição 7.1 *As condições de controle de uma dupla $\langle t_k, d_i \rangle$ é representada pelo conjunto $CC = \{CF, SEV, EP\}$ onde:*

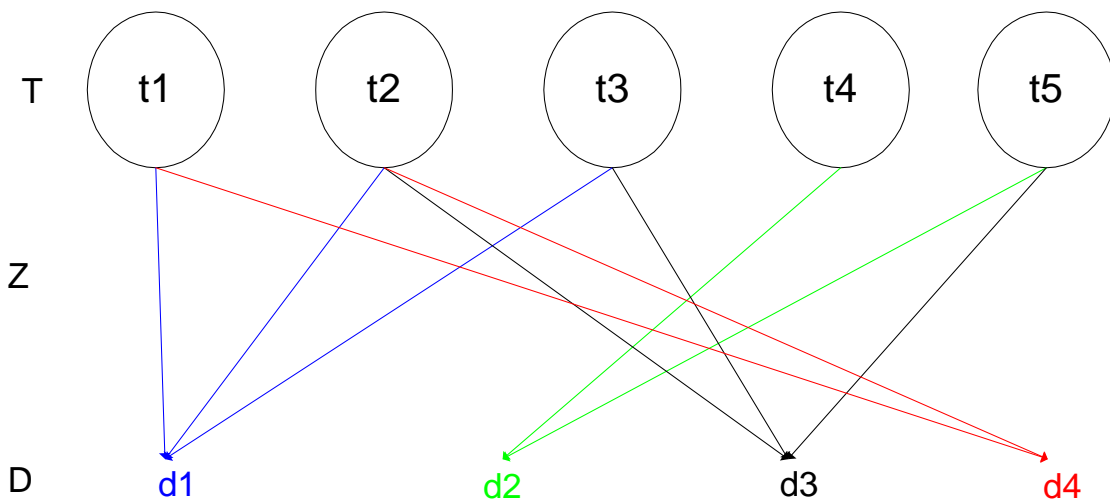


FIGURA 7.1 – Um ambiente abdutivo para tratamentos.

- *CF* corresponde às condições que favorecem a eficácia de um tratamento t_k para uma desordem d_i , no contexto de fitopatologia $CF = (TEMP, UMID)$ são intervalos nebulosos relativos à temperatura e umidade favoráveis para tratar uma desordem d_i usando tratamento t_k ;
- *SEV* corresponde a um grau nebuloso trapezoidal de intensidade ou severidade de uma desordem d_i que favorece a eficácia de um tratamento t_k para uma desordem d_i ;
- *EP* corresponde a um intervalo nebuloso que representa a fase de desenvolvimento ou idade de um determinado paciente ou planta que favorecem a eficácia de um tratamento t_k para uma desordem d_i .

Similarmente às condições de controle, neste trabalho é proposto representar as variáveis auxiliares de cada dupla $Z = \langle t_k, d_i \rangle$ através de um conjunto $V = \{GCONT, GEVOL, CUSTO, RISCO\}$, como definido a seguir.

Definição 7.2 O conjunto V de uma dupla $Z = \langle t_k, d_i \rangle$ é formado pelo conjunto de variáveis $V = \{GCONT, GEVOL, CUSTO, RISCO\}$, onde:

- $GCONT \in [0, 1]$ corresponde ao grau efetivo de controle de um tratamento t_k para uma desordem d_i caso as condições em CC sejam satisfeitas;
- $GEVOL \in [0, 1]$ corresponde ao grau de evolução de uma determinada desordem d_i quando aplicado um tratamento t_k em determinadas condições de controle CC ;

- *CUSTO* é o custo associado ao tratamento t_k para uma desordem d_i ;
- *RISCO* corresponde aos riscos associados a um determinado tratamento (por ex. efeitos colaterais).

Todas as variáveis do conjunto CC , conforme apresentado no capítulo 5, além do *CUSTO* e do *RISCO*, são representadas por intervalos nebulosos trapezoidais. Observemos ainda que estamos supondo que os intervalos nebulosos das variáveis *RISCO* e *CUSTO* estão no mesmo universo de discurso (ex. sacas/ha). Entretanto, os valores dos intervalos nebulosos devem ser mapeados para $[0,1]$ quando estes não estiverem no mesmo universo de discurso. As variáveis *GCONT* e *GEVOL* correspondem a possibilidade de um tratamento t_k ser efetivo (Π_e) ou a possibilidade de um tratamento t_k não ser efetivo ($\Pi_{\bar{e}}$).

Formalmente, uma base de conhecimento KB_t pode ser definida como descrito na definição 7.3.

Definição 7.3 A base de conhecimento KB_t é representada por uma quintupla $\langle T, D, Z, CC, V \rangle$, onde:

- T é o conjunto de tratamentos ou medidas de controle;
- D é o conjunto de desordens (doenças, falhas);
- Z é uma relação para modelar associações entre tratamentos e desordens como apresentadas nos catálogos farmacêuticos ou fitopatológicos;
- CC é representada pelo conjunto $CC = \{CF, SEV, EP\}$ (definição 7.1)
- V é o conjunto de variáveis $\{GCONT, GEVOL, CUSTO, RISCO\}$ (definição 7.2).

Para um problema de tratamentos representado por KB_t , é conveniente e útil definir os conjuntos ou funções baseadas na relação Z :

Definição 7.4 Para todo $t_k \in T$ e $d_i \in D$ em uma base de conhecimento KB_t ,

- $indicações(t_k) = \{d_i \mid \langle t_k, d_i \rangle \in Z\}$, o conjunto dos elementos diretamente controlados por t_k ;
- $tratamentos(d_i) = \{t_k \mid \langle t_k, d_i \rangle \in Z\}$, o conjunto dos elementos que podem diretamente controlar ou tratar d_i .

Os conjuntos definidos acima podem facilmente ser generalizados de tratamentos e desordens individualmente para conjuntos de tratamentos e desordens conforme apresentado a seguir.

Definição 7.5 Para todo $T_L \subseteq T$ e $D_I \subseteq D$ em uma base de conhecimento KB_t ,

- $indicações(T_L) = \bigcup_{t_k \in T_L} controles(t_k), e$
- $tratamentos(D_I) = \bigcup_{d_i \in D_I} tratamentos(d_i).$

Assim, os resultados de um conjunto de tratamentos são apenas a união dos resultados individuais de cada tratamento do conjunto. Por exemplo, no problema de tratamento apresentado na Figura 7.1 tem-se:

$$indicações(t_1) = \{d_1, d_4\},$$

$$indicações(t_2) = \{d_1, d_3, d_4\},$$

$$indicações(t_3) = \{d_1, d_3\},$$

$$indicações(t_4) = \{d_2\},$$

$$indicações(t_5) = \{d_2, d_3\},$$

$$tratamentos(d_1) = \{t_1, t_2, t_3\},$$

$$tratamentos(d_2) = \{t_4, t_5\}.$$

Utilizando estes dados tem-se que:

$$indicações(t_1, t_2) = \{d_1, d_3, d_4\}.$$

$$tratamentos(d_1, d_2) = \{t_1, t_2, t_3, t_4, t_5\}.$$

7.2 Especificação do Caso Particular

Para complementar a formulação do problema de tratamentos é necessário formalizar os casos particulares. No escopo de tratamentos, a informação sobre um caso particular deve contemplar apenas observações sobre as hipóteses e as condições de controle em que estão ocorrendo as desordens (temperatura, umidade, severidade e época).

Formalmente, a informação sobre um caso particular é modelada por uma 5-tupla $CA = \langle M^+, M^-, EV^+, TIME^+, CC^+ \rangle = (CF^+, SEV^+, EP^+)$, como apresentada na definição 5.5 (seção 5.2). A partir da definição 5.5, pode-se extrair a informação sobre um caso particular na etapa de tratamentos.

Definição 7.6 A informação sobre um caso particular na etapa de tratamentos é modelada por uma dupla $CA = \langle D^+, CC^+ \rangle$ onde:

- D^+ é o conjunto de desordens presentes conhecidas no caso;
- $CC^+ = (CF^+, SEV^+, EP^+)$ são as condições externas em que ocorreram os sintomas, onde CF^+ é uma função que associa as condições de temperatura

e umidade ($TEMP^+, UMID^+$) no momento do diagnóstico, SEV^+ e EP^+ são as funções que associam a severidade de cada $m_j \in M^+$ e a fase em que elas ocorreram, respectivamente.

Uma vez caracterizado os problemas de diagnóstico, investigação e tratamentos ($P = \langle KB_d, KB_e, KB_t, CA \rangle$), caracteriza-se agora as soluções para o problema de tratamentos, denominado $SOL_t(P)$. Para isto, além da noção de "cobertura" (seção 3.2.2), utiliza-se outros índices de consistência e coberturas nebulosas que serão descritos a seguir.

7.3 Tomada de Decisão em Tratamentos

Nesta seção descreve-se as soluções para os problemas de tratamentos definidos na seção 7.1. Inicialmente, define-se a noção de consistência entre o modelo de tratamentos (seção 7.1) e um caso particular (seção 7.2). Na sequência, uma explicação plausível para um problema de tratamentos é definida formalmente.

7.3.1 Consistência entre o Modelo de Tratamentos e o Caso

Os tratamentos mais plausíveis T (remédios, fungicidas), para um dado caso, devem ser inferidas abduativamente a partir da consistência entre a informação sobre as hipóteses mais plausíveis $D_l \in SOL_d \subseteq D$ (desordens potenciais) e uma base de conhecimento $KB_t = \langle T, D, Z, CC, V \rangle$. Para tal, é utilizada a noção de coberturas irredundantes da TCP apresentada na seção 3.2.2. O critério de parcimônia irredundante garante a noção de minimalidade e simplicidade explicitada na seção 3.1.

Primeiramente, é necessário verificar a consistência das hipóteses em relação ao modelo de tratamentos utilizando as condições de controle CC , além dos critérios de parcimônia de irredundância. No escopo deste trabalho, são utilizados os índices de consistência de intensidade propostos em Wainer e Sandri (1998). O grau de consistência das condições de controle de um tratamento $t_k \in KB_t$ e um caso particular CA é definido de maneira similar ao diagnóstico (definição 5.7).

Definição 7.7 O grau de consistência das condições de controle de uma dupla $\langle t_k, d_i \rangle \in KB_t$ e um caso particular CA é calculado através da fórmula $\gamma(t_k, d_i) = \inf_{c_k \in CC} \gamma(c_k)$ de modo que $\gamma(c_k) = h(INT(c_k) \cap INT^+(c_k))$, $INT(c_k)$ corresponde a intensidade de $c_k \in CC$ em $KB_t = \langle T, D, Z, CC, V \rangle$ e $INT^+(c_k)$ corresponde a intensidade de cada $c_k \in CC^+$.

Adicionalmente, no escopo da TCN é possível estreitar as possíveis decisões terapêuticas através do conjunto de variáveis $V = \{GCONT, GEVOL, CUSTO, RISCO\}$ como

definido a seguir.

Definição 7.8 Um determinado tratamento t_k para uma determinada desordem d_i que explica um caso particular CA é mais plausível se:

- a consistência de CF, SEV, EP é máxima,
- o $GONT(t_k)$ é máximo,
- o $GEVOL(t_k)$ é mínimo, e
- o $Retorno(t_k)$ é máximo onde $Retorno(t_k)$ é calculado em função das variáveis $CUSTO(t_k)$ e $RISCO(t_k)$ como definido na equação 7.1.

$$Retorno(t_k) = RISCO_{produtividade}(t_k) \ominus CUSTO(t_k) \quad (7.1)$$

A seguir é apresentado um exemplo visando ilustrar as definições 7.7 e 7.8. Suponhamos o exemplo em que se tem a seguinte quintupla da base de conhecimento $KB_t = \langle T, D, Z, CC, V \rangle$, onde $T = \{t_1, t_2\}$ e $D = \{d_3\}$. Para cada Z tem-se os seguintes valores para CC e V :

- $Z = \langle t_1, d_3 \rangle$ tem-se $CC = (EP = \langle 2, 12 \rangle, SEV = \langle 1, 3, 5, 7 \rangle, CF = \langle 0, 1 \rangle)$ e $V = (GCONT = \langle 1.0 \rangle, GEVOL = \langle 0.5 \rangle, CUSTO = \langle 10, 20, 30, 40 \rangle, RISCO = \langle 100, 112, 118, 120 \rangle)$
- $Z = \langle t_2, d_3 \rangle$ tem-se $CC = (EP = \langle 10, 22 \rangle, SEV = \langle 7, 8, 9, 10 \rangle, CF = \langle 0, 1 \rangle)$ e $V = (GCONT = \langle 1.0 \rangle, GEVOL = \langle 0 \rangle, CUSTO = \langle 10, 10, 20, 20 \rangle, RISCO = \langle 100, 112, 118, 120 \rangle)$

Consideremos o caso particular $CA_1 = \langle M^+, M^-, EV^+, TIME^+, CC^+ \rangle$ com $CC^+ = \{CF^+, SEV^+, DS^+\}$ onde $DS^+ = \langle 8, 9 \rangle$, $SEV^+ = \langle 3, 4, 5, 6 \rangle$, $CF^+ = TEMP^+ = \langle 22, 23, 25, 26 \rangle$ e $UMID^+ = \langle 80, 85, 90, 95 \rangle$.

Suponhamos que a partir da inferência abdutiva (seção 5.3), tem-se o diagnóstico para o caso particular CA_1 , onde $\alpha_{ct}(\{d_3\}) > 0.5$, tal que $\{d_3\} \subseteq D_l \subseteq SOL_d$. A partir da KB_d , tem-se que as condições favoráveis para evolução de d_3 são: $CF(d_3) = (TEMP = \langle 25, 30, 35, 45 \rangle, UMID = \langle 85, 90, 95, 100 \rangle)$.

Agora suponhamos que a partir da inferência abdutiva (seção 7.3) em KB_t tem-se os possíveis tratamentos para d_3 representados por $tratamentos(d_3) = \{t_1, t_2\}$. O processo de inferência para a seleção do tratamento mais adequado deve seguir alguns passos.

Primeiramente, aplica-se a fórmula para consistência das condições favoráveis:

$$\omega_{CF}(\{d_3\}) = \sup(\gamma(TEMP), \gamma(UMID)) \quad (7.2)$$

Aplicando a fórmula 3.3 encontra-se o grau de consistência de intensidade de temperatura e umidade como descrito abaixo.

$$\gamma(TEMP) = h(< 25, 30, 35, 45 > \cap < 22, 23, 25, 26 >) = 1.0$$

$$\gamma(UMID) = h(< 85, 90, 95, 100 > \cap < 80, 85, 90, 95 >) = 1.0$$

Ao aplicar a fórmula 7.2 tem-se o grau de consistência das condições favoráveis:
 $\omega_{CF}(d_3) = 1.0$.

Observemos que o domínio de intensidade da variável $CF \in V$ é $\Omega_{INT(CF)} = [0, 1]$ tal que $\Omega_{INT(CF)} = [0, 0.5] = \neg CF$ (quando CF é falsa) e $\Omega_{INT(CF)} = [0.5, 1.0] = CF$ (quando CF é verdadeira).

No exemplo acima, tanto o tratamento t_1 quanto t_2 são aplicáveis em quaisquer condições de temperatura e umidade. Em um segundo passo, conforme descrito na definição 7.7, calcula-se o grau de consistência entre um tratamento $t_k \in KB_t$ e um caso particular CA é através da fórmula $\gamma(t_k) = \inf_{c_k \in CC} \gamma(c_k)$. Primeiramente, calcula-se o grau de consistência para t_1 :

$$\gamma(EP) = h < 2, 12 > \cap < 8, 9 > = 1.0$$

$$\gamma(SEV) = h < 1, 3, 5, 7 > \cap < 3, 4, 5, 6 > = 1.0$$

$$\omega(CF) = 1.0$$

$$\gamma(t_1) = \inf(\gamma(EP), \gamma(SEV), \omega(CF)) = 1.0$$

Similarmente, calcula-se o grau de consistência para t_2 :

$$\gamma(EP) = h < 10, 22 > \cap < 8, 9 > = 0$$

$$\gamma(SEV) = h < 7, 8, 9, 10 > \cap < 3, 4, 5, 6 > = 0$$

$$\omega(CF) = 1.0$$

$$\gamma(t_2) = \inf(\gamma(EP), \gamma(SEV), \omega(CF)) = 0$$

No exemplo acima, t_1 é o tratamento mais plausível para d_3 em relação as condições de controle. Entretanto, neste caso não é necessário utilizar as variáveis V pois encontrou-se somente um tratamento como solução para o caso CA.

Suponhamos agora o caso particular $CA_2 = \langle M^+, M^-, EV^+, TIME^+, CC^+ \rangle$ tal que $CC^+ = CF^+, SEV^+, DS^+$ onde $DS^+ = \langle 10, 12 \rangle$, $SEV^+ = \langle 6, 8 \rangle$ e $CF^+ = TEMP^+ = \langle 23, 25 \rangle$, $UMID^+ = \langle 85, 90 \rangle$. Neste caso $\gamma(t_1) = \gamma(t_2) = 1.0$. No entanto, como o grau de controle do tratamento de t_1 é 1.0 (GCONT= $\langle 1.0 \rangle$) e o grau de controle de t_2 é 0 (GCONT= $\langle 1.0 \rangle$), t_1 ou t_2 são plausíveis no caso CA_2 .

Assim, o diagnosticador deve considerar na sua tomada de decisão outras variáveis como o custo e o risco de perda da produtividade. Para o exemplo acima, o custo de t_1 é $\langle 20, 40 \rangle$ e o de t_2 é $\langle 10, 20 \rangle$ e a produtividade esperada para os dois é em torno de $\langle 100, 120 \rangle$. Assim sendo, deve-se escolher aquele tratamento que dá um maior retorno que é calculado através da fórmula 7.1:

$$Retorno(t_1) = \langle 100, 112, 118, 120 \rangle \ominus \langle 10, 20, 30, 40 \rangle = \langle 60, 82, 98, 110 \rangle$$

$$Retorno(t_2) = \langle 100, 12, 118, 120 \rangle \ominus \langle 10, 10, 20, 20 \rangle = \langle 80, 92, 108, 110 \rangle$$

Assim, $\max(Retorno(t_1), Retorno(t_2)) = t_2$.

Neste último caso, t_2 é o tratamento mais plausível.

Para obter $\max(Retorno(t_1), Retorno(t_2))$ é necessário utilizar um método para ordenação de intervalos nebulosos. Neste trabalho está sendo utilizado os índices globais de *exceedance* propostos por Dubois e Prade (1988).

Um índice global de *exceedance* pode ser definido pela avaliação do grau que um intervalo M_i domina todos os outros intervalos M_j para $j \neq i$. Dado um intervalo M_i e um intervalo M_j , a ordem de M_i e M_j pode ser calculada através de 4 índices que são calculados pela comparação de M_i e $\widetilde{\max}_{j \neq i} M_j$ como definido a seguir:

- $PSE(M_i) = \prod_{M_i}([\widetilde{\max}_{j \neq i} M_j, +\infty))$ (*exceedance possibility*)
- $PS(M_i) = \prod_{M_i}([\widetilde{\max}_{j \neq i} M_j, +\infty))$ (*strict exceedance possibility*)
- $NSE(M_i) = N_{M_i}([\widetilde{\max}_{j \neq i} M_j, +\infty))$ (*exceedance necessity*)
- $NS(M_i) = N_{M_i}([\widetilde{\max}_{j \neq i} M_j, +\infty))$ (*strict exceedance necessity*)

onde, Π_{M_i} e N_{M_i} são as medidas de possibilidade e necessidade (veja apêndice A) definidas em termos da distribuição μ_{M_i} , respectivamente;

$\widetilde{\max}(M_j)$ significa uma extensão do operador max no contexto de princípio de extensão de conjuntos nebulosos e M_j um intervalo nebuloso, dado por

$$\max_j M_j(U) = \sup\{\alpha \in]0, 1[/ U \in \max_j (M_j)_\alpha\}$$

onde $\max([a, b], [c, d]) = [\max(a, c), \max(b, d)]$;

$[\widetilde{\max}_{j \neq i} M_j, +\infty)$ e $] \widetilde{\max}_{j \neq i} M_j, +\infty)$ referem-se ao conjunto de números que *pode* e ao conjunto que é *necessariamente* maior ou igual ao maior valor do intervalo nebuloso M_j , respectivamente.

Dado dois intervalos nebulosos do tipo LR (Dubois e Prade, 1988), isto é, $P = (p, \bar{p}, \alpha, \beta)_{LR}$ e $Q = (q, \bar{q}, \gamma, \delta)_{LR}$, o cálculo dos 4 índices se reduzem a encontrar os intervalos e os valores modais dos intervalos nebulosos como descrito a seguir:

$$\begin{aligned} \text{(PSE): } Pos(\bar{X} \geq \underline{Y}) &= \max(0, \min(1, 1 + \frac{(\bar{p}-q)}{(\beta+\gamma)})) \\ \text{(NSE): } Nes(\underline{X} \geq \underline{Y}) &= \max(0, \min(1, \frac{(p-q+\gamma)}{(\alpha+\gamma)})) \\ \text{(PS): } Pos(\bar{X} \geq \bar{Y}) &= \max(0, \min(1, 1 + \frac{(\bar{p}-\bar{q}-\beta)}{(\beta+\delta)})) \\ \text{(NS): } Nes(\underline{X} \geq \bar{Y}) &= \max(0, \min(1, \frac{(p-\bar{q})}{(\alpha+\delta)})). \end{aligned}$$

Para exemplificar como os índices PSE, PS, NSE e NS podem ser utilizados para classificar intervalos nebulosos, consideremos os intervalos nebulosos trapezoidais $\langle 60, 82, 98, 110 \rangle$ e $\langle 80, 92, 108, 110 \rangle$ referentes à $Retorno(t_1)$ e $Retorno(t_2)$ respectivamente (veja Figura 7.2).

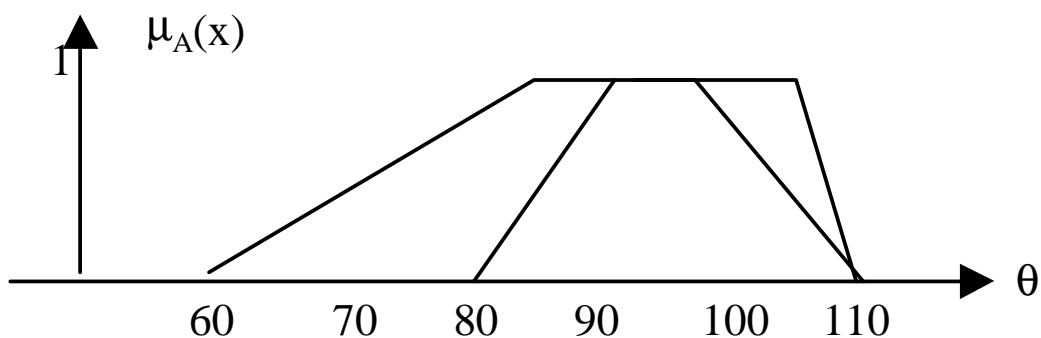


FIGURA 7.2 – Intervalos nebulosos referentes à $Retorno(t_1)$ e $Retorno(t_2)$.

Dado que um intervalo trapezoidal é representado pela quadrúpla $\langle a, b, c, d \rangle$, tem-se que $P = (p, \bar{p}, \alpha, \beta)_{LR} = \langle b, c, b - a, d - c \rangle = \langle 82, 98, 22, 12 \rangle$ e $Q = (q, \bar{q}, \gamma, \delta)_{LR} = \langle b, c, b - a, d - c \rangle = \langle 92, 108, 12, 2 \rangle$ referentes aos intervalos nebulosos $Retorno(t_1) = \langle 60, 82, 98, 110 \rangle$ e $Retorno(t_2) = \langle 80, 92, 108, 110 \rangle$ e aqui denominados R_1 e R_2 , respectivamente.

Os valores dos 4 índices são calculados como apresentado na tabela 7.1. Pelos índices PSE e PS não é possível diferenciar os dois intervalos, entretanto, pelos índices NSE e NS torna-se claro que R_2 é maior que R_1 .

TABELA 7.1 – Resultados do exemplo.

	PSE	NSE	PS	NS
R_1	1	2/34	12/14	-26/24
R_2	1	32/34	12/14	1/4

7.3.2 Solução para Tratamentos

A partir das idéias acima, pode-se definir formalmente a seleção de tratamentos através da inferência abdutiva.

Definição 7.9 (Solução abdutiva) *Dado uma base de conhecimento $KB_t = \langle T, D, Z, CC, V \rangle$ e uma hipótese $D_l \in SOL_d \subseteq D$, SOL_t é uma solução abdutiva para D_l a partir de T sse $Z \cup SOL_t$ cobre D_l , $Z \cup SOL_t$, $CC \cup SOL_t$ sejam consistente e V é máximo.*

A partir das definições 5.9, 6.9 e 7.9, completa-se a formalização da definição 4.2, onde dada uma base de conhecimento $\Theta = \{KB_d, KB_e, KB_t\}$ para o problema de raciocínio clínico P e um caso particular CA , $SOL(P) = \{SOL_d, SOL_e, SOL_t\}$ é uma solução abdutiva para caso particular CA .

No apêndice B estão apresentados os algoritmos desenvolvidos para implementar os mecanismos de inferência dos ambientes abduativos para diagnóstico, investigação e tratamentos apresentados nos capítulos 5, 6 e 7, respectivamente. No próximo capítulo serão apresentados alguns resultados obtidos na validação da TCN.

CAPÍTULO 8

VALIDAÇÃO DA ABORDAGEM PROPOSTA

Neste capítulo está descrita a experiência da aplicação da Teoria das Coberturas Nebulosas (TCN) proposta nesta dissertação a um problema real de diagnóstico e controle de doenças de plantas. O problema de raciocínio clínico investigado neste trabalho está relacionado às doenças de milho.

Conforme descrito no capítulo 1, a base de conhecimento de um sistema especialista para diagnóstico de doenças do milho, que foi desenvolvido na Embrapa Informática Agropecuária em parceria com a Embrapa Milho e Sorgo, está sendo utilizada como estudo de caso no presente trabalho (disponível em <http://diagnose.cnptia.embrapa.br>).

A base de conhecimento original contém 41 doenças de milho com suas respectivas manifestações ou efeitos. No escopo deste trabalho foi necessário coletar informação adicional das 41 doenças de milho para representar o conhecimento da base $KB = \{KB_d, KB_e, KB_t\}$. O conhecimento utilizado para complementar a base de conhecimento foi obtido através de entrevistas com fitopatologistas da Embrapa e médicos, além de literatura correspondente na área (Reis e Casa, 1996; Fernandes e Oliveira, 1997; Malcolm, 1980; Agroofit-OnLine, 2003). A especificação completa da base de conhecimento está apresentada no apêndice C e os algoritmos que implementam a TCN no apêndice B.

Os programas desenvolvidos para a solução dos problemas de raciocínio clínico correspondem a implementação dos algoritmos apresentados no apêndice B. A linguagem utilizada na implementação dos algoritmos da TCN foi o AMZI PROLOG devido a sua sintaxe e aderência ao problema (<http://www.amzi.com>). Adicionalmente, a implementação em PROLOG facilitou a comparação da TCN com o TCP original e TCP temporal através das ferramentas que também foram desenvolvidas em Prolog no escopo do trabalho de Wainer e Rezende (1997).

Este capítulo está organizado da seguinte forma. Na seção 8.1 está apresentado um exemplo completo para diagnóstico, investigação e tratamentos na área de doenças de plantas, mais especificamente, doenças de milho. Embora o exemplo utilizado no decorrer deste capítulo seja para diagnose e tratamento de doenças de milho, espera-se que os desenvolvimentos teóricos obtidos possam ser aplicados para outros problemas de diagnóstico dinâmico. Na seção 8.2 é apresentado um pequeno exemplo na área de

diagnóstico e controle de falhas de satélite. Finalmente, na seção 8.3 está descrita uma análise dos resultados obtidos nas etapas de diagnóstico, investigação e tratamentos.

8.1 Diagnóstico e Controle de Desordens em Cultura de Milho

Nesta seção está descrita a representação da base de conhecimento da cultura de milho contemplando os modelos de representação das desordens bem como de suas manifestações, da evolução temporal e das condições favoráveis ao desenvolvimento de cada desordem. Além disso, são descritos como representar os testes laboratoriais e as medidas de controle que devem ser adotadas para cada desordem.

Os programas em Prolog que compõem a TCN são claramente divididos em 3 partes: base de conhecimento, mecanismo de inferência e especificação do caso particular. O mecanismo de inferência abduutivo foi implementado através dos algoritmos do apêndice B: *bipartido_nebuloso*, *exames_nebulosos* e *tratamentos_nebulosos*. A base de conhecimento e o caso particular estão apresentados a seguir.

Antes de apresentá-los, na seção 8.1.1 tem-se uma breve descrição do desenvolvimento da cultura de milho desde a germinação até a fase de maturação fisiológica, pois, a evolução temporal das manifestações de cada desordem são descritas referindo-se a escala de desenvolvimento da cultura.

8.1.1 Descrição do Problema

A Fenologia é o estudo dos eventos periódicos da vida da planta em função da sua reação às condições do ambiente. Engloba todas as etapas do crescimento e desenvolvimento da vida vegetal como a germinação, emergência, elaboração do aparato fotossintético, florescimento, aparecimento de estruturas reprodutivas e maturação de sementes.

O conhecimento da fenologia da cultura possibilita a determinação das relações e do grau de influência dos fatores envolvidos no processo produtivo, favorecendo a previsão de problemas, estabelecimento de estratégias de manejo e tomada de decisão. A escala fenológica de uma plantação de milho está representada na tabela 8.2 e ilustrada na Figura 8.1.

Observemos na Figura 8.1 que os intervalos de dias ou semanas correspondente a cada estágio de desenvolvimento são intervalos aproximados, portanto, eles podem ser representados como intervalos nebulosos trapezoidais (coluna 4 da tabela 8.2).

TABELA 8.1 – Etapas de desenvolvimento da cultura de milho.

Estágios	Descrição	Intervalos (em dias)	Intervalo Nebuloso (em semanas)
e_0	Germinação	0 a 10	<0,1,1,2>
e_1	Crescimento Vegetativo	10 a 55	<2,4,6,8>
e_2	Pendoamento	55 a 70	<8,9,9,10>
e_3	Florescimento	70 a 80	<9,10,11,12>
e_4	Frutificação	80 a 120	<12,13,16,17>
e_5	Maturação	120 a 150	<17,18,21,22>

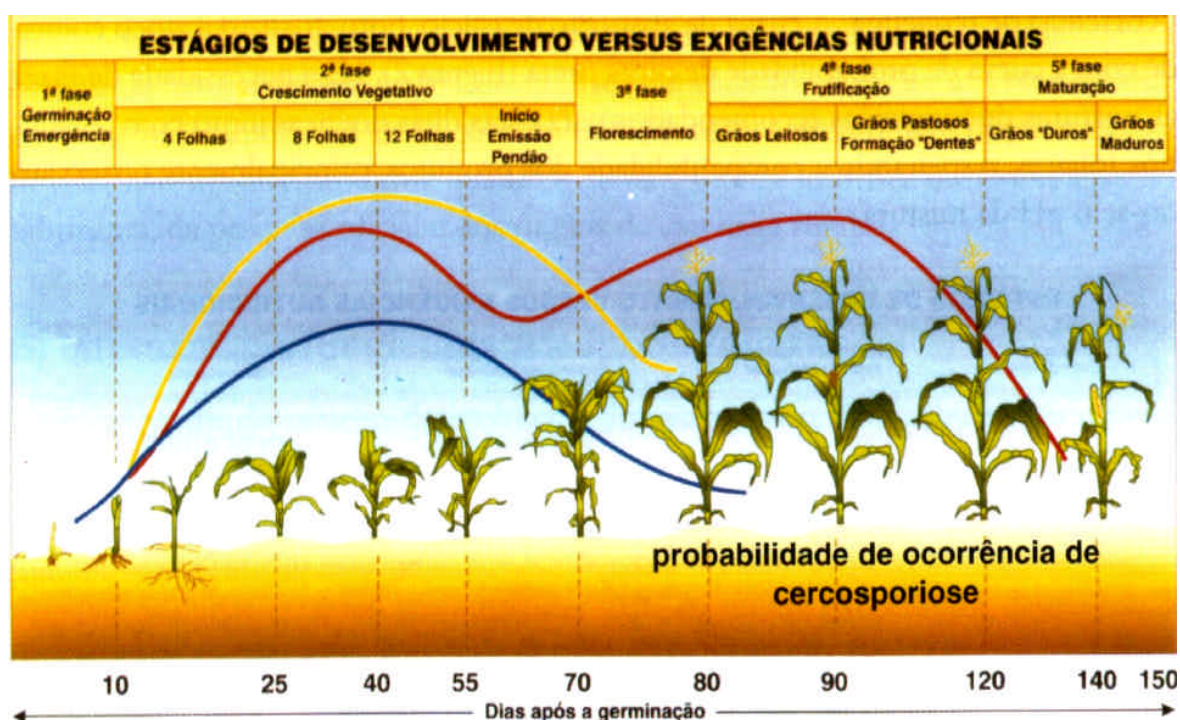


FIGURA 8.1 – Escala fenológica da cultura do milho.

Os estágios de desenvolvimento de uma cultura podem ser também representados por um grafo onde os estágios correspondem aos nós e os intervalos nebulosos entre os estágios correspondem aos arcos como apresentado na Figura 8.2.

Além do estágio de desenvolvimento da cultura, a intensidade ou a severidade da desordem no momento do diagnóstico interfere na tomada de decisão diagnóstica. Vários autores tem proposto modelos e escalas para medir estas variáveis (Bergamin e L., 1996). A escala diagramática a ser utilizada nos exemplos de doenças de milho está representada na tabela 8.3 adaptada de Morandi (2002).

8.1.2 Representação e Implementação da Base de conhecimento KB_d

Nesta seção será exemplificado como representar o conhecimento sobre um problema de diagnóstico no domínio de doenças de plantas sob o escopo da TCN. Um subconjunto

TABELA 8.2 – Etapas de desenvolvimento da cultura de milho.

Estágios	Descrição	Intervalos (em dias)	Intervalo Nebuloso (em semanas)
e_0	Germinação	0 a 10	<0,1,1,2>
e_1	Crescimento Vegetativo	10 a 55	<2,4,6,8>
e_2	Pendoamento	55 a 70	<8,9,9,10>
e_3	Florescimento	70 a 80	<9,10,11,12>
e_4	Frutificação	80 a 120	<12,13,16,17>
e_5	Maturação	120 a 150	<17,18,21,22>

TABELA 8.3 – Níveis de severidade da cultura do milho.

Nível de severidade	Descrição do nível	Grau nebuloso
s_1	Sem lesões (antes do aparecimento dos sintomas)	-
s_2	Lesões esparsas na planta, restritas ao terço inferior	<1,2,3,4>
s_3	Até 50% das folhas com lesões; severas nos 25% das folhas inferiores	<3,4,5,6>
s_4	Até 75% das folhas com lesões; severas nos 50% das folhas inferiores	<5,6,7,8>
s_5	100% das folhas com lesões; lesões severas nos 75% das folhas inferiores	<7,8,9,10>
s_6	Planta morta	-

da base de conhecimento de doenças de milho que contém 10 desordens foliares como listadas na tabela 8.4 será utilizado como exemplo. A especificação completa da base de conhecimento de doenças de milho KB_d está descrita no apêndice C.

TABELA 8.4 – Exemplo de doenças foliares de milho.

Referência	Desordem	Agente Causal
d_1	PODRIDÃO do CARTUCHO por Erwinia	<i>Erwinia chrysanthemi</i>
d_2	QUEIMA POR Pseudomonas	<i>Pseudomonas albopropitans</i>
d_3	FERRUGEM COMUM	<i>Puccinia sorghi</i>
d_4	FERRUGEM POLYSORA	<i>Puccinia polysora</i>
d_5	FERRUGEM BRANCA OU TROPICAL	<i>Physopella zae</i>
d_6	MANCHA POR Physoderma	<i>Physoderma maydis</i>
d_7	ANTRACNOSE	<i>Colletotrichum graminicola</i>
d_8	MANCHA POR Helminthosporium maydis - Raça O	<i>Helminthosporium maydis</i>
d_9	MANCHA POR Helminthosporium maydis - Raça T	<i>Helminthosporium maydis</i>
d_{10}	MANCHA POR Cercospora	<i>Cercospora zae-maydis</i>

Primeiramente, um resumo das principais características de uma das doenças da tabela 8.4, denominada Mancha por Cercospora (d_{10}), está apresentado na tabela 8.4. Estas características foram obtidas através de entrevistas com fitopatologistas (Morandi, 2002) e literatura na área (Fernandes e Oliveira, 1997; Malcolm, 1980). Em seguida, na tabela 8.6 está apresentada a 6-tupla da base de conhecimento $KB_d = \langle \theta, D, M, R, G, CC \rangle$ correspondente a desordem d_{10} que ilustra como estas características podem ser representadas em KB_d .

TABELA 8.5 – Principais características da Cercosporiose.

MANCHA POR Cercospora(d_{10})
<p>Sintomas -A reação de susceptibilidade é caracterizada por lesões retangulares, limitadas pelas nervuras, inicialmente de cor palha passando a cinza. Podem ocorrer na bainha, colmo e palhas das espigas.</p> <p>A Cercosporiose, sob condições favoráveis, pode causar seca prematura das plantas e geralmente causam tombamento. Essa doença pode ser encontrada em qualquer fase do desenvolvimento das plantas de milho, inicialmente nas folhas baixas.</p> <p>Condições favoráveis - temperatura entre 23 e 35°C, e água livre na superfície das folhas.</p> <p>Medidas de Controle - aplicação de fungicidas, utilização de cultivares resistentes; rotação de cultura; aração e gradagem.</p>

TABELA 8.6 – Ênupla correspondente a representação da desordem Cercospora (d_{10}) em KB_d .

θ	D	M	R= $\langle d_{10}, m_j \rangle$	G(V,T)	CC=(CF,SEV,EP)
$\langle 0,22 \rangle$ semanas	d_{10}	m_1 : folhas se coalescem m_2 : seca em grandes áreas foliares m_{13} : seca prematura das plantas m_{19} : lesões necróticas palha m_{23} : lesões necróticas cinzas m_{56} : geralmente tombam	0.5 0.75 0.75 1.0 1.0 0.75	grafo de eventos (Figura 8.4)	CF=(TEMP= $\langle 25,30,35,40 \rangle$,UMID= $\langle 85,90,95,100 \rangle$) SEV= $\langle 1,10 \rangle$ (Tabela 8.3) EP= $\langle 0,22 \rangle$ (Tabela 8.2)

A primeira coluna da tabela 8.6, θ refere-se a escala de tempo representada na tabela 8.2. A segunda coluna D refere-se as desordens que compõem a base de conhecimento (neste exemplo d_{10}). Na coluna M está representado o conjunto de manifestações causadas por d_{10} : ($efeitos(d_{10}) = \{m_1, m_2, m_{13}, m_{19}, m_{23}, m_{56}\}$). Na coluna R está representada a associação gradual entre desordens e manifestações como apresentado na Figura 8.3. Na coluna G está referenciado o grafo mínimo de eventos associado a desordem d_{10} (Figura 8.4). A última coluna, CC, representada pela tripla $CC = (CF, SEV, EP)$ onde o par $CF = (TEMP, UMID)$ corresponde aos intervalos nebulosos trapezoidais que representam as condições favoráveis de temperatura e umidade que favorecem a evolução da desordem d_{10} (veja Tabela 5.2). A variável SEV corresponde a um intervalo nebuloso trapezoidal de intensidade ou severidade em que d_{10} pode aparecer (Tabela 8.3) e a variável EP corresponde a um intervalo nebuloso que representa a fase de desenvolvimento ou idade de um determinado paciente ou planta que favorece a evolução de d_{10} (Tabela 8.2).

A Figura 8.3 apresenta um grafo simplificado que representa a desordem d_{10} com suas respectivas manifestações e relação causal $R(d_{10}, m_j)$.

8.1.3 Representação e Implementação da Base de conhecimento KB_e

Nesta seção será exemplificado como representar o conhecimento sobre um problema de investigação no domínio de doenças de plantas sob o escopo da TCN. Um subconjunto da base de conhecimento de testes laboratoriais para identificação de doenças de milho

que contém 9 testes laboratoriais listados na tabela 8.7 será utilizado como exemplo. A especificação completa da base de conhecimento de doenças de milho KB_d está descrita no apêndice C.

TABELA 8.7 – Exemplo de testes laboratoriais.

Referência	Exames	Tipos de desordens
e_1	Análise Macroscópica	<i>fúngicas</i>
e_2	Microscopia ótica	<i>fúngicas, bacterianas</i>
e_3	Isolamento em cultura pura	<i>fúngicas</i>
e_4	Testes bioquímicos	<i>bacterianas</i>
e_5	Análise Perfil ácidos Graxos	<i>bacterianas</i>
e_6	Análise de Ribossomo	<i>bacterianas</i>
e_7	Testes Sorológicos	<i>bacterianas, viroses, enfezamentos e fúngicas</i>
e_8	Análise de DNA	<i>viroses, enfezamentos</i>
e_9	Microscopia eletrônica	<i>viroses, enfezamentos</i>

Os principais sintomas e as condições favoráveis para a evolução da Cercosporiose (d_{10}) representados na tabela 8.6 exemplificam como as desordens podem ser representadas no escopo do ambiente abduativo (KB_d). Similarmente, será mostrado como os exames ou testes laboratoriais para desordens de milho podem ser representados no escopo da TCN.

Na tabela 8.8 está apresentada um exemplo da 6-tupla da base de conhecimento $KB_e = \langle \theta, E, I, D, M, R \rangle$ correspondente aos testes laboratoriais que podem identificar d_{10} .

TABELA 8.8 – Tupla correspondente a representação dos testes laboratoriais que identificam (d_{10}) em KB_e .

θ	E	I	D	M	R= $\langle d_{10}, m_j \rangle$
$\langle 0, 22 \rangle$ semanas	ex_1	$\langle ex_1, d_{10} \rangle$	d_{10}	m_1 : folhas se coalescem	0.5
	ex_2	$\langle ex_2, d_{10} \rangle$		m_2 : seca em grandes areas foliares	0.75
	ex_3	$\langle ex_3, d_{10} \rangle$		m_{13} : seca prematura das plantas	0.75
	ex_7	$\langle ex_7, d_{10} \rangle$		m_{19} : lesoes necroticas palha	1.0
	ex_8	$\langle ex_8, d_{10} \rangle$		m_{23} : lesoes necroticas cinzas	1.0
					m_{56} : geralmente tombam

A primeira coluna da tabela 8.8, θ refere-se a escala de tempo representada na tabela 8.2. A segunda coluna E refere-se aos testes laboratoriais que compõem a base de conhecimento (tabela 8.7). Na coluna I está representada a associação entre os testes laboratoriais e as desordens. Na coluna D está representado as desordens em que aparecem

as manifestações (neste exemplo somente d_{10}). Na coluna M está representado o conjunto de manifestações causadas por d_{10} : ($efeitos(d_{10}) = \{m_1, m_2, m_{13}, m_{19}, m_{23}, m_{56}\}$). Na coluna R está representada a associação gradual entre desordens e manifestações como apresentado na Figura 8.3.

8.1.4 Representação e Implementação da Base de conhecimento KB_t

Nesta seção será exemplificado como representar o conhecimento sobre um problema de tratamentos no domínio de doenças de plantas sob o escopo da TCN.

Os principais sintomas e as condições favoráveis para a evolução da Cercosporiose (d_{10}) representados na tabela 8.6 exemplificam como as desordens podem ser representadas no escopo do ambiente abduativo (KB_d). Similarmente, será mostrado como as medidas de controle para desordens de milho podem ser representadas no escopo da TCN.

Dentre as medidas de controle citadas na tabela 8.4, somente a aplicação de fungicidas é um tratamento corretivo enquanto as outras medidas são preventivas. Portanto, os exemplos de tratamentos considerados aqui serão, principalmente, aplicação de fungicidas. Na tabela 8.9 estão descritos os 10 principais fungicidas já registrados no Ministério da Agricultura. A base de conhecimento completa está apresentada no apêndice C.

TABELA 8.9 – Exemplo de fungicidas.

Referência	Tratamentos
t_1	Priori
t_2	Opera
t_3	Palisade
t_4	Juno (propicanazole)
t_5	Folicur
t_6	Folicur 200CE (tebuconazole)
t_7	Tilt (propicanazole)
t_8	Maxim (fluxiaxomil)
t_9	Tecto 100 (thiabendazole)
t_{10}	Tecto 600 (thiabendazole)

Durante o processo de raciocínio clínico, o diagnosticador tem que tomar a decisão de quando aplicar (A) ou não aplicar ($\neg A$) um fungicida. Na Figura 8.5 está apresentado um exemplo de uma árvore de decisão utilizada por um fitopatologista para decidir sobre a aplicação de um fungicida.

Os círculos representam as variáveis de estado (EP,SEV,CF), os arcos que saem de cada um destes nós representam os valores possíveis que estas variáveis podem assumir. Os nós losângos representam as variáveis de decisão (A) e os arcos que saem destes nós correspondem as ações que podem ser executadas (A - aplicar ou $\neg A$ - não aplicar). As folhas correspondem as possíveis consequências (c_1, c_2, \dots, c_{10}) de cada decisão. Neste exemplo, cada consequência c_i mostra que cada decisão envolve maximizar o controle da desordem e, conseqüentemente, a produtividade (tabela 8.10).

TABELA 8.10 – Consequências da árvore de decisão da Figura 8.5.

Consequência	Desordem	Produtividade
c_1	controle alto	normal
c_2	evolução rápida	redução baixa
c_3	evolução lenta	normal
c_4	controle médio-alto	redução baixa
c_5	evolução lenta-intermediária	redução baixa
c_6	controle baixo	redução média-alta
c_7	evolução rápida	redução alta
c_8	evolução lenta-intermediária	redução média-alta
c_9	controle nulo	redução alta
c_{10}	evolução lenta-intermediária	redução alta

Na Figura 8.5 tem-se apenas a representação dos eventos e ações para a decisão de um tratamento para uma determinada etapa de desenvolvimento da plantação de milho. Entretanto, a representação de todas as árvores de decisão para todos os tratamentos de todas desordens é muito trabalhosa e difícil.

No escopo da TCN é proposta uma solução alternativa na área de tomada de decisão em tratamentos. A TCN separa claramente o conhecimento do domínio (conjuntos T, D e relação Z), o papel do raciocínio abdutivo (critério de parcimônia e a definição da cobertura), e heurísticas do domínio (variáveis CC e V) diferentemente dos sistemas baseados em regras.

No escopo da TCN, pode-se a qualquer momento entrar com novos tratamentos e variáveis auxiliares na base de conhecimento incrementalmente a partir da experiência do diagnosticador, independente do mecanismo de inferência abdutivo.

Para ilustrar a representação de conhecimento em $KB_t = \langle T, D, Z, CC, V \rangle$ serão considerados 5 fungicidas que podem ser aplicados como tratamentos alternativos para Cercosporiose (d_{10})(tabela 8.11). Os intervalos nebulosos das variáveis auxiliares

correspondentes a cada um destes fungicidas foram obtidos de um experimento com estes fungicidas (Morandi, 2002). Na tabela 8.11 estão representadas as tuplas referentes a cada tratamento em KB_t .

TABELA 8.11 – Tupla correspondente a representação dos tratamentos para a desordem Cercospora (d_{10}) em KB_t .

T	D	Z= $\langle t_k, d_{10} \rangle$	CC=(EP,SEV,CF)	V=(GCONT,GEVOL,CUSTO,RISCO)
T1: PRIORI	d_{10}	$\langle t_1, d_{10} \rangle$	$=(0,17],[1,4],[0,1])$	$(1.0,0,[10,40],[145,150])$
T2: OPERA		$\langle t_2, d_{10} \rangle$	$(0,17],[1,4],[0,1])$	$(1.0,0,[10,40],[140,150])$
T3: PALISADE		$\langle t_3, d_{10} \rangle$	$(0,17],[1,4],[0,1])$	$(1.0,0.5,[10,40],[130,140])$
T4: JUNO		$\langle t_4, d_{10} \rangle$	$(0,17],[1,4],[0,1])$	$(1.0,0.5,[10,40],[130,140])$
T5: FOLICUR		$\langle t_5, d_{10} \rangle$	$(0,17],[1,4],[0,1])$	$(1.0,0.5,[10,40],[140,150])$
T15: TESTEMUNHA		$\langle t_{15}, d_{10} \rangle$	$(0,17],[1,4],[0,0.5])$	$(0.5,1.0,[0,0],[100,120])$

Para complementar a formulação do problema necessita-se da representação do caso particular que será apresentada na próxima seção.

8.1.5 Representação de um Caso Particular

Suponhamos um caso hipotético em que um produtor de milho procura um fitopatologista na décima oitava semana de plantio, pois sua lavoura apresentou algumas manifestações. Após uma entrevista com o produtor, o fitopatologista coletou algumas observações adicionais:

- por volta da oitava à nona semana começaram aparecer os sintomas;
- por volta da oitava à nona semana apareceram nas folhas, algumas lesões necróticas palha (m_{19} da tabela 8.6);
- 50% das folhas apresentam lesões e 25% das folhas apresentam em maior intensidade (s_3 da tabela 8.3);
- por volta da décima primeira e décima segunda semana aproximadamente após a germinação as folhas começaram a coalescer (m_1) e algumas tornaram-se secas em grandes áreas foliares (m_2 da tabela 8.6);
- as condições climáticas da região oscilam com uma temperatura entre 23 e 25 °C e com uma alta umidade relativa do ar;
- foi notado que as lesões não estavam circundadas por halo amarelado (corresponde a m_{20} na KB_d no apêndice C).

Conforme apresentado na seção 5.2, o caso particular acima pode ser representado através do formalismo proposto na definição 5.5 pela 5-tupla $C_a = \langle M^+, M^-, EV^+, TIME^+, CC^+ \rangle$ onde,

$$M^+ = \{m_1, m_2, m_{19}\}, M^- = m_{20},$$

$$EV^+ = \{m_1^b, m_{13}^b, m_{19}^b, m_{56}^b\},$$

$$TIME^+(m_1^b) = \langle 10, 11, 12, 14 \rangle, TIME^+(m_2^b) = \langle 10, 11, 12, 15 \rangle,$$

$$\text{e } TIME^+(m_{19}^b) = \langle 7, 8, 9, 10 \rangle.$$

$$CC^+ = \{CF^+, SEV^+, DS^+\} \text{ onde}$$

$$DS^+ = \langle 8, 9 \rangle, SEV^+ = \langle 3, 4, 5, 6 \rangle \text{ e } CF^+ = \{TEMP^+ = \langle 22, 23, 25, 26 \rangle, UMID^+ = \langle 80, 85, 90, 95 \rangle\}.$$

Após a definição da representação do caso, pode-se definir como se obter a solução para um caso particular que será apresentada na próxima seção.

8.1.6 Solução para o Problema de Raciocínio Clínico

Nesta seção são apresentados os resultados obtidos durante a solução do caso descrito na seção 8.1.5. No escopo da TCN, o diagnosticador busca a solução do problema de raciocínio clínico apresentada na Figura 4.1 através da execução dos módulos `bipartido_nebuloso`, `exames_nebulosos` e `tratamentos_nebulosos` como representado na Figura 4.6.

EXEMPLO 1: Conforme apresentado na seção 8.1.5, a lista abaixo mostra a forma com que foi fornecida a informação do caso ao programa em prolog `bipartido_nebuloso.pro` (apêndice B).

```
bipartido([m1, m13, m19, m56, [m20]], [[m1b, [10, 11, 12, 14]],
[m13b, [10, 11, 12, 15]], [m19b, [7, 8, 9, 10]], [m56b, [12, 13, 14, 15]]],
[[[x, [0, 0, 0, 0]], [], [], [], [], [], Hipoteses, Indices)
```

Observemos que a sintaxe desta lista corresponde a $([M^+, [M^-]], [[ev_i^b, [TIME^+]], [ev_j^b, [TIME^+]]], [[[e_0, [0, 0, 0, 0]], [], [], [], [], [], Hipoteses, Indices)$.

Note que embora a pseudo-manifestação 'x' não pertença á M^+ , o instante 0 é associado ao seu início para representar o momento da sementeira (e_0).

Após processar a informação acima, conforme descrito na seção 5.3, obtém-se o seguinte resultado:

```
[10]
[18,2]
```

```
[10 - mancha_por_cercospora, 0.75]\\
```

Grau de Certeza da Hipotese:0.75

```
[18 - mancha_por_phaeosphaeria, 0.33]\\
```

```
[2 - queima_por_pseudomonas, 0.6666666666666667]\\
```

Grau de Certeza da Hipotese: 0.33

Dado que o fator de certeza da hipótese $[d_{10}]$ é igual a 0.75 e o grau de certeza da outra hipótese é igual a 0.33, então a explicação mais plausível para este caso é $[d_{10}]$.

Observemos que na segunda hipótese acima ($[18, 2]$), a vírgula representa um \wedge , portanto, a hipótese $[18, 2]$ representa d_{18} e d_2 ocorrem simultaneamente. Como d_{18} é uma desordem fúngica e d_2 é uma doença bacteriana, é muito baixa a possibilidade das duas desordens estarem ocorrendo simultaneamente. Entretanto, somente através de testes laboratoriais as hipóteses poderiam ser confirmadas.

Para selecionar as manifestações prioritárias a serem investigadas utiliza-se a fórmula 6.1 implementada no programa *exames - nebuloso.pro*, conforme apresentado no capítulo 6.

```
investigacao([m1, m2, m19], [2, 10, 18], [], Resexpd).
```

Ao executar *exames_nebuloso.pro* para o caso acima obtém-se o seguinte resultado:

```
[m23,m27,m56,m28,m13,m3]
```

```
[analise_macroscopica]
```

```
[microscopia_otica]
```

```
[isolamento_em_cultura_pura]
```

```
[testes_sorologicos]
```

```
[analise_DNA]
```

Observemos que m_{23} e $\{m_{27}\}$ são manifestações necessárias para confirmação das desordens d_{10} e d_{18} respectivamente. Portanto, durante a investigação, a confirmação da presença ou ausência de m_{23} auxilia na confirmação ou na refutação da desordem d_{10} na hipótese diagnóstica. Similarmente, a confirmação da presença ou ausência de

m_{27} auxilia na confirmação ou na refutação da desordem d_{18} na hipótese diagnóstica. Enquanto as manifestações $\{m_{56}, m_{28}, m_{13}, m_3\}$ não necessariamente contribuirão para otimização da hipótese diagnóstica, pois, são manifestações possíveis de acontecer com diferentes graus de importância.

Assim, os exames laboratoriais (E) que identificam as manifestações $\{m_3, m_{13}, m_{23}, m_{27}, m_{28}, m_{56}\}$ na KB_e devem ser solicitados na ordem $exames(m_{23}) \prec exames(m_{27}) \prec exames(m_2) \prec exames(m_1) \prec exames(m_{56}) \prec exames(m_{13})$ visando confirmar ou refutar as hipóteses que causam estes efeitos. Portanto, a investigação deve começar pela m_{23} que é uma manifestação específica da desordem d_{10} , pois a confirmação de sua presença garante que a hipótese d_{10} é a desordem que cobre o caso particular apresentado na seção 8.1.5. Como apresentado no resultado acima, a confirmação desta manifestação pode ser obtida através de uma análise-macroscópica. Caso não seja possível identificar a manifestação 'in loco', pode-se executar outros exames descritos na lista acima, como microscopia_óptica, testes_sorológicos para tentar identificar a causa dos sintomas apresentados no caso.

Após a confirmação de m_{23} , com a informação completa do caso particular, pode-se executar novamente o programa *bipartido_nebuloso.pro*:

```
bipartite([lesoes_necroticas_cinzas,lesoes_necroticas_palha,
seca_em_grandes_areas_foliar,folhas_se_coalescem],
[[m23b,[10,11,13,14],[m19b,[7,8,9,10],[m2b,[10,11,12,13]],
[m1b,[10,11,12,14]]],[[x,[0,0,0,0]],[],[],[],[],[],Resultado,Indices)
```

Após processar a informação acima, conforme descrito na seção 5.3, obtém-se o seguinte resultado:

[10]

```
10 - mancha_por_cercospora Grau = 0.5\
Grau de Certeza da Hipotese: 0.5
```

Observe que o grau de certeza da hipótese d_{10} abaixou de 0.75 para 0.5 após o refinamento da informação e o grau de certeza da hipótese $\{d_{18}, d_2\}$ tornou-se incompatível (grau de certeza = 0) como consequência da diminuição do grau de entropia diagnóstica (incerteza). Por outro lado, obteve-se uma maior diferenciação das hipóteses que pertencem a solução abduativa SOL_d .

Supondo que d_{10} é a explicação para o caso particular acima, isto é, $d_{10} \in SOL_d$, calcula-se abduktivamente os tratamentos mais plausíveis para d_{10} . No caso do programa *tratamentos_nebuloso.pro*, este caso particular é fornecido como a seguir.

$([d_{10}], [8, 9], [3, 6], [[23, 25], [85, 90]], Resultado)$.

Observemos que a sintaxe desta lista corresponde a

$([SOL_d], [EP^+], [SEV^+], [[TEMP^+], [UMID^+]], Resultado)$.

Ao executar *tratamentos_nebuloso.pro* para o caso acima tem-se:

[PRIORI]

[OPERA]

t1 - 1 - priori Grau = 0.5 \\

t1 - 2 - priori Grau = 1\\

t1 - 3 - priori Grau = 0.5

Grau de Certeza da Hipotese: 1

Grau de Controle:1

Grau de Evolucao da desordem: 0.5

Retorno esperado: [110,122,128,140] sa/ha

t2 - 1 - opera Grau = 0.5\\

t2 - 2 - opera Grau = 1\\

t2 - 3 - opera Grau = 0.5

Grau de Certeza da Hipotese: 1

Grau de Controle:1

Grau de Evolucao da desordem:0.5

Retorno esperado: [110,122,128,140] sa/ha

Observemos que foram apresentados como resultado 2 fungicidas dos 5 aplicáveis à Cercosporiose (d_{10}).

EXEMPLO 2: Suponhamos um outro exemplo que corresponde ao exemplo descrito no capítulo 5. Observemos que $\{m_6, m_7, m_{13}\}$ correspondem aos índices de referência da base de conhecimento (apêndice C) para as manifestações $\{m_1, m_3, m_4\}$ do exemplo apresentado no capítulo 5.

```
bipartite([pustulas_individualizadas_marrom_claro,
pustulas_individualizadas_marrom_escuro,
rompimento_da_epiderme, seca_prematura_das_plantas],
[[m6b, [6, 7, 8, 9]], [m7b, [16, 17, 18, 19]], [m8b, [7, 8, 9, 10]],
[[x, [0, 0, 0, 0]], [], [], [], [], [], Hipoteses, Indices),
```

Após rodar o *bipartido_nebuloso.pro*, obteve-se o seguinte resultado:

```
[ferrugem_comum]
[ferrugem_polysora]
```

```
[3 - ferrugem_comum, 0.75]
```

Grau de Certeza da Hipotese: 0.75

```
[4 - ferrugem_polysora, 0.75]
```

Grau de Certeza da Hipotese: 0.75

Embora não se tenha certeza se o diagnóstico é d_3 ou d_4 , isto é, se a plantação foi atacada pelo fungo *Puccinia sorghi* ou *Puccinia polysora*, o diagnosticador pode optar por identificar diretamente os tratamentos mais plausíveis neste caso C_a pois $\alpha_{ct}([d_3], [d_4]) > 0.5$.

Consideremos o caso abaixo, caso ele opte por investigar novas manifestações de $\{d_3, d_4\}$:

```
investigacao([m6, m7, m8], [3, 4], [], Resexpd).
```

Ao executar o programa *exames_nebuloso.pro* para o caso acima obtém-se que:

```
efeitos(d3) = [m6, m7, m8, m9, m13, m14] e
efeitos(d4) = [m6, m7, m8, m10, m13, m14].
```

A ordem das manifestações a serem investigadas e os exames mais plausíveis para este caso são:

```
[m9, m10, m14, m13]
```


[analise_macroscopica]
[microscopia_otica]
[isolamento_em_cultura_pura]
[testes_sorologicos]
[analise_DNA]

As manifestações m_9 e m_{10} que têm força maior são específicas de d_3 e d_4 , respectivamente. Entretanto, d_3 e d_4 não necessariamente precisam causar m_9 e m_{10} pois a associação causal $R(d_3, m_9) = .5$ e $R(d_4, m_{10}) = .5$. Assim, o diagnosticador pode optar por tentar identificar diretamente os tratamentos mais plausíveis ao invés de continuar na investigação.

Pela KB, tem-se que as condições favoráveis para evolução de d_3 e d_4 são: $\{TEMP = < 15, 20, 25, 30 >, UMID = < 85, 90, 95, 100 >\}$ e $\{TEMP = < 25, 30, 35, 45 > e UMID = < 85, 90, 95, 100 >\}$, respectivamente. Sabe-se também que $CC^+ = \{CF^+, SEV^+, DS^+\}$ onde $DS^+ = < 12, 17 >$, $SEV^+ = < 7, 8 >$ e $CF^+ = \{TEMP^+ = < 15, 16 >, UMID^+ = < 85, 90 >\}$.

No caso do programa *tratamentos_nebuloso.pro*, este caso particular é fornecido como a seguir.

main : -decisao([d3, d4], [12, 17], [7, 8], [[15, 16], [85, 90]]).

Ao executar o programa *tratamentos_nebuloso.pro* para o caso acima tem-se:

[folicur_200_ce_tebuconazole]

t6 - 1 - folicur_200_ce_tebuconazole

Grau = 0.714285714285714

Grau de Certeza da Hipotese: 0.714285714285714

Grau de Controle:1

Grau de Evolucao da desordem:0.5

Retorno esperado: [100, 112, 128, 140] sa/ha

Embora não se tenha certeza se o diagnóstico é d_3 ou d_4 , pode-se usar o tratamento T_6 - FOLICUR 200 CE pois este fungicida serve para controlar as duas desordens.

8.2 Diagnóstico e Controle em Falhas de Satélite

Nesta seção será dado um pequeno exemplo na área de diagnóstico e controle de falhas de satélite de modo a ilustrar como a TCN pode ser utilizada em outros problemas de diagnóstico dinâmico. Primeiramente, no item 8.2.1 é apresentada uma breve descrição do problema, retirada de A822007-ANL-61/D1 (2001). Em seguida, é apresentado como este problema pode ser representado no escopo da TCN. Finalmente, no item 8.2.3 é apresentada a solução do problema no escopo da TCN.

8.2.1 Descrição do Problema

Conforme descrito no capítulo 4.1, o diagnóstico de falhas de satélite é um ponto crucial na área espacial. Os satélites em órbita comunicam-se com seus centros de controle através de telemetria (satélite para terra) e telecomandos (terra para satélite). Parte da telemetria é dedicada ao monitoramento do satélite. Quando uma falha é detectada pelos serviços de monitoramento em tempo-real dos centros de controle, é necessário localizar o mais rápido possível a origem da falha de modo a tomar as ações corretivas para garantir a continuidade da missão.

A preocupação por parte de engenheiros e técnicos é tão grande que existe uma metodologia detalhada de análise de risco e gerenciamento para a plataforma multi-missão (MMP) do INPE (A822007-ANL-61/D1, 2001). Esta análise de risco do sistema é desenvolvida na fase de especificação do projeto.

Estas falhas são detectadas através de telemetria e para cada uma delas existe um formulário onde é preenchido a falha, suas prováveis causas e as ações corretivas que devem ser tomadas para cada falha através de telecomandos.

Para ilustrar como a TCN pode ser aplicada neste domínio é apresentado um pequeno exemplo retirado de (A822007-ANL-61/D1, 2001), onde os nomes dos equipamentos, das falhas, dos métodos de detecção e das ações corretivas foram mantidos em inglês em conformidade com o documento original.

8.2.2 Representação do Conhecimento

Nesta seção será exemplificado como representar o conhecimento sobre um problema de diagnóstico e controle de falhas de satélite sob o escopo da TCN. Conforme apresentado nas seções 8.1.2, 8.1.3 e 8.1.4, a representação do conhecimento é suportada pela base de conhecimento $KB = \{KB_d, KB_e, KB_t\}$ em PROLOG.

A base de conhecimento Kb_d contém 28 itens críticos (equipamentos) e 51 falhas que estes equipamentos podem causar. Um subconjunto da Kb_d contendo 8 componentes e 16 falhas será utilizado como exemplo (tabela 8.12 e 8.13).

TABELA 8.12 – Exemplo de itens críticos que podem causar falhas em satélite.

Referência	Equipamentos	Agente Causal
d_9	Rodas de reação	Falhas mecânicas e eletrônicas
d_{12}	Tanque de Hidrazina	Reação química, Falha Mecânica, Falha do Controle Térmico e Decomposição do Catalisador
d_{13}	Propulsores do satélite (Thrusters)	Operação em baixa temperatura, Partículas de poeira nos tubos
d_{14}	Válvulas de enchimento e drenagem	Aberta
d_{16}	Filtros	Excesso de partículas no fluido ou entupimento parcial
d_{18}	Parte Elétricas dos Painéis Solares - EPSA	Manufatura, Curto Circuito ou Circuito aberto
d_{19}	Bateria	Manufatura, Curto Circuito ou Circuito aberto
d_{20}	Unidade de Condicionamento e Distribuição de Potência - PCDU	Falha eletrônica da interface do PCDU

TABELA 8.13 – Exemplo de falhas em um missão de satélite.

Referência	Falhas	Causas
m_1	perda dos serviços de telecomunicações	[1,2,3,4]
m_4	perda de redundância	[1,4,5,6,8,9,26]
m_{16}	perda de determinação de altitude	[5,6,7,8,9,11]
m_{18}	perda da dessaturação das rodas de reação	[10]
m_{19}	perda da função de propulsão	[12,14]
m_{20}	ruptura da membrana	[12]
m_{21}	perda de pressão	[12,14]
m_{22}	aumento da pressão	[12]
m_{23}	congelamento de propelente	[12,14]
m_{27}	perda do incremento da velocidade mínima	[13,16]
m_{29}	redução da capacidade de geração de energia	[18]
m_{30}	redução da capacidade de armazenamento de energia	[19]
m_{31}	queda do cabo principal voltagem	[19]
m_{32}	perda de controle e condicionamento de potência	[20]
m_{33}	perda de distribuição de potência e funções de proteção	[20]

Fazendo o mapeamento de um subconjunto de itens críticos acima para a TCN, tem-se o conjunto $D = \{d_{12}, d_{13}, d_{14}, d_{16}, d_{18}, d_{19}, d_{20}\}$, o conjunto de efeitos ou falhas M tal que $M = \{m_{19}, m_{27}, m_{29}, m_{30}, m_{32}, m_{33}\}$ e a relação causal C denotada pela rede causal apresentada na Figura 8.6.

A representação das bases de conhecimento Kb_e e Kb_t são similares a base de conhecimento Kb_d . A base de conhecimento Kb_e contém 16 métodos de detecção de falhas que podem ser utilizados durante o processo de diagnóstico e controle de falhas. A base de conhecimento Kb_t contém 15 ações corretivas que podem ser executadas durante

o processo de controle de falhas. Um resumo dos principais métodos de detecção e das principais ações corretivas estão descritos nas tabelas 8.14 e 8.15, respectivamente.

TABELA 8.14 – Exemplo de métodos de detecção de falhas em satélite.

Referência	Métodos de Detecção
e_1	telemetria
e_2	estação recebendo sinais degradados ou sinais perdidos totalmente
e_3	detecção de circuito aberto e alarme operacional no painel do computador
e_4	status do equipamento por telemetria
e_5	análise de telemetria da terra
e_6	telemetria em modo de segurança
e_7	monitoramento por computador
e_8	alarme
e_9	teste de consistência lógica
e_{10}	falha de alarme do modo de segurança

TABELA 8.15 – Exemplo de ações corretivas para controle de falhas em satélite.

Referência	Ações Corretivas
t_1	comutação para unidade redundante através de telecomando da terra
t_2	modo de segurança
t_3	telecomando enviado da estação da terra
t_4	chaveamento para equipamento redundante por telecomando
t_5	análise da terra
t_6	telecomando por modo de segurança
t_7	recalibração dos propulsores
t_8	mudança do modo de operação mmp
t_9	degradação dos requisitos operacionais de mmp
t_{10}	chaveamento automático
t_{15}	não tem ação corretiva

Para complementar a formulação do problema $P = \langle KB_d, KB_e, KB_t, CA \rangle$, consideremos um caso particular CA onde obteve-se por telemetria que ocorreu uma perda da função de propulsão do satélite por volta da quarta semana do satélite em órbita.

8.2.3 Solução do Problema

Nesta seção são apresentados os resultados obtidos para alguns casos de falha do satélite.

EXEMPLO 3: Consideremos o caso da perda da função de propulsão como descrito abaixo.

bipartite([*perda_funcao_propulsao*], [[*m19b*, [0, 12, 12, 24]]],
[[*x*, [0, 0, 0, 0]]], [], [], [], [], [], *Hipoteses*, *Indices*).

Após processar o *bipartido – nebuloso.pro* obtém-se o seguinte resultado:

[12 - tanque_de_hidrazina, 0.5]
Grau de Certeza da Hipotese: 0.5

[14 - valvulas_de_enchimento_e_drenagem, 0.5]\\
Grau de Certeza da Hipotese:0.5

Então o próximo passo do investigador é tentar identificar a causa da falha ou outras falhas que possam estar acontecendo. Na fase de investigação obtém-se o seguintes resultados:

investigacao([*m19*], [12, 14], [], *Resexpd*).

Ao executar *exames – nebuloso.pro* para o caso acima tem-se que os *efeitos*(*d12*) = [*m19*, *m20*, *m21*, *m22*, *m23*] e *efeitos*(*d14*) = [*m19*, *m21*, *m23*].

Assim, obtém-se que a ordem das manifestações a serem investigadas e os exames mais plausíveis para este caso são:

[*m20*, *m22*, *m19*, *m21*, *m23*]

[*telemetria*]

[*monitoramento_por_computador*]

Observemos que a manifestação m_{20}, m_{21}, m_{22} e m_{23} não estão contempladas em M^+ . Assim, os exames laboratoriais (E) que identificam as manifestações $\{m_{19}, m_{20}, m_{21}, m_{22}, m_{23}\}$ na KB_e devem ser solicitados na ordem $exames(m_{20}) \preceq exames(m_{22}) \prec exames(m_{19}) = exames(m_{21}) = exames(m_{23})$ visando confirmar ou refutar as hipóteses que causam estes efeitos. Portanto, a investigação deve começar pelas manifestações m_{20} ou m_{22} . Observemos que m_{20} é uma manifestação específica da desordem d_{12} , portanto, a confirmação de sua presença confirma a hipótese que d_{12} é a desordem que cobre o caso particular apresentado acima.

A confirmação desta manifestação pode ser obtida através de telemetria. Após a confirmação de m_{20} e m_{22} , com a informação do caso completa, pode-se executar novamente o programa `bipartido-nebuloso.pro`:

```
bipartite([perda_funcao_propulsao, ruptura_da_mebrana, aumento_da_pressao],
[[m19b, [0, 12, 12, 24]], [m20b, [12, 14, 18, 24]], [m21b, [12, 14, 18, 24]]],
[[x, [0, 0, 0, 0]], [], [], [], [], [], R, R2)
```

Após processar a informação acima, conforme descrito na seção 5.3, obtém-se o seguinte resultado:

```
[10 - tanque_de_hidrazina, 0.75]\\
Grau de Certeza da Hipotese: 0.75
```

Supondo que d_{12} é a explicação para o caso particular acima, isto é, $d_{12} \in SOL_d$, calcula-se abduktivamente os tratamentos mais plausíveis para d_{12} . Ao executar `tratamentos_nebuloso.pro` para o caso acima tem-se:

```
[nao_tem_acao_corretiva]\\
```

Portanto, para este item crítico não há ação corretiva e é um risco de severidade muita alta que pode acabar com a missão do satélite.

8.3 Considerações Finais sobre a Abordagem Proposta

Na seção 8.1 são descritos os resultados obtidos durante a solução de um problema de diagnóstico, investigação e tratamentos na área de doenças de milho no escopo da TCN. Na seção 8.2 são descritos os resultados obtidos em um outro exemplo na área de diagnóstico e controle de falhas de satélite. Nesta seção está descrita uma análise dos resultados obtidos na área de diagnóstico, investigação e tratamentos.

No escopo da etapa de diagnóstico, pode-se comparar e comentar os resultados obtidos pelo algoritmo *Bipartido–Nebuloso* que é baseado no algoritmo *bipartite* apresentado na seção 3.2. Entretanto, no *bipartido–nebuloso* antes de aplicar os critérios de parcimônia de irredundância através da função *revise* verifica-se a consistência temporal e categórica do caso particular em relação ao modelo de desordens que é implementado através da função VNCD - Verificador Nebuloso de Consistência para Diagnóstico.

TABELA 8.16 – Comparação de algoritmos na fase de diagnóstico.

Bipartido	Bipartido-temporal	Bipartido-nebuloso
[10]	∅	[10, [0.75]]
[8, 1]		[18, 2, [0.33]]
[8, 2]		
[10, 1]		
[10, 2]		
[18, 1]		
[18, 2]		

Na tabela 8.16 são apresentados os resultados obtidos com os algoritmos *bipartido-nebuloso*, *bipartido* e *bipartite-temporal* proposto em Wainer e Rezende (1997). Na terceira coluna da tabela 8.16 são descritos os resultados obtidos após processar a informação do caso descrita na seção 8.1.5 pelo programa *bipartido – nebuloso.pro*. Na primeira coluna estão descritos os resultados obtidos ao rodarmos o mesmo caso particular sem considerar informação temporal e nebulosa no programa (*bipartido.pro*). Na segunda coluna são descritos os resultados obtidos ao rodar o caso no programa *temporal – pct.pro* que utiliza intervalos bem definidos.

Observemos na tabela 8.16 que no escopo do TCP original (bipartido) foram geradas 7 explicações plausíveis para este caso particular e debaixo da ferramenta que implementa o bipartido-temporal.pro nenhuma hipótese foi encontrada dentre as 7 apresentadas para o mesmo caso no bipartido.pro. Já no contexto do TCP-nebuloso foram geradas apenas 2 explicações plausíveis dentre as 7 apresentadas inicialmente no TCP original. Portanto, no escopo do *bipartido_nebuloso.pro* obteve-se 86% de redução das hipóteses em relação ao *bipartido.pro*.

Adicionalmente, o grau de certeza associado a cada hipótese no bipartido_nebuloso.pro também auxilia o diagnosticador na tomada de decisão durante o diagnóstico. Nos resultados obtidos, observa-se que d_{10} é a explicação mais plausível. Portanto, o diagnosticador investiga outras manifestações de d_{10} . Ao confirmar novas manifestações presentes o diagnosticador pode rodar novamente o bipartido-nebuloso e obter os resultados apresentados na tabela 8.17.

TABELA 8.17 – Comparação de algoritmos na fase de diagnóstico.

Bipartido	Bipartido-temporal	Bipartido-nebuloso
[10]	∅	[10, [0.5]]
[10, 1]		
[10, 2]		

Observemos na tabela 8.17 que no escopo do TCP original foram geradas 3 explicações plausíveis para especificação completa do caso particular e debaixo da ferramenta que implementa o TCP-temporal nenhuma explicação foi encontrada para o mesmo caso. Já no contexto do TCP-nebuloso foi gerada apenas 1 explicação plausível dentre as 2 apresentadas inicialmente no TCP original (redução de 50% das hipóteses). Portanto, a utilização de intervalos nebulosos para modelar informação temporal mostrou-se mais eficiente do que o modelo TCP-temporal que utiliza intervalos bem definidos além de auxiliar na investigação e na redução do conjunto de hipóteses.

No escopo da investigação, pode-se comentar os resultados obtidos pelo algoritmo *exames_nebulosos* que é baseado no algoritmo bipartido apresentado na seção 3.2. Entretanto, o algoritmo *exames_nebulosos* antes de aplicar os critérios de parcimônia de irredundância através da função *revise* permite selecionar as manifestações prioritárias a serem investigadas dentro do contexto das hipóteses consistentes. Esta funcionalidade é implementada através da função VNCCI - Verificador Nebuloso de Consistência para Investigação. Na tabela 8.18 pode-se observar os resultados obtidos com o algoritmo *exames_nebulosos* nos exemplos 1 e 2 apresentados na seção 8.1.6.

TABELA 8.18 – Análise dos algoritmos na fase de investigação.

Exemplos	M^+ ordenadas	SOL_d	SOL_e
EXEMPLO 1	$[m23, m27, m28, m56, m13, m3]$	$[d10]$	$[analise - macroscopica]$ $[microscopia - otica]$ $[isolamento - em - cultura - pura]$ $[testes - sorologicos]$
EXEMPLO 2	$[m9, m10, m14, m13]$	$[d3, d4]$	$[analise - DNA]$ $[analise - macroscopica]$ $[microscopia - otica]$ $[isolamento - em - cultura - pura]$ $[testes - sorologicos]$ $[analise - DNA]$

Observemos que a classificação de manifestações auxilia na tomada de decisão para diagnóstico como apresentado no EXEMPLO 1. Entretanto, a classificação de hipóteses não contribuiu muito para a redução de hipóteses no EXEMPLO 2. Tais resultados são coerentes com o raciocínio do diagnosticador na fase de investigação e mostram como esta etapa pode contribuir para a tomada de decisão na etapa de diagnóstico.

No escopo da etapa de tratamentos, pode-se comentar os resultados obtidos pelo algoritmo *tratamentos_nebulosos* que é baseado no algoritmo bipartido apresentado na seção 3.2. Entretanto, o algoritmo *tratamentos_nebulosos* antes de aplicar os critérios de

parcimônia de irredundância através da função *revise* permite selecionar os tratamentos mais plausíveis dentro do contexto das hipóteses consistentes. Esta funcionalidade é implementada através da função VNCT - Verificador Nebuloso de Consistência para Tratamentos. Na tabela 8.19 pode-se observar os resultados obtidos com o algoritmo *tratamentos_nebulosos* nos exemplos 1 e 2 apresentados na seção 8.1.6.

TABELA 8.19 – Análise dos algoritmos na fase de tratamentos.

Exemplos	SOL_d	SOL_t
<i>EXEMPLO1</i>	$[d_{10}]$	$[t_1, t_2]$
<i>EXEMPLO2</i>	$[d_3, d_4]$	$[t_6]$

Observemos que no primeiro exemplo, os fungicidas OPERA E PRIORI são mais indicados para controle da Cercosporiose quando se considera variáveis como época do aparecimento da desordem, condições climáticas e a severidade da desordem. Inicialmente, no escopo do programa *bipartido.pro* aparecem 5 tipos de fungicidas indicados para controlar d_{10} : OPERA, PRIORI, PALISADE, JUNO e FOLICUR. Portanto, no escopo do *tratamentos_nebulosos* conseguiu-se reduzir 3 hipóteses.

Estes resultados correspondem aos resultados obtidos através de um experimento em campo (Morandi, 2002) onde foram aplicados os 5 tipos de fungicidas para combater o fungo *Cercospora zea-maydis* e os fungicidas OPERA E PRIORI foram mais eficientes ao considerar as variáveis época do aparecimento da desordem, condições climáticas e a severidade da desordem.

No segundo exemplo, também os resultados são coerentes com o raciocínio do diagnosticador. Embora ele não tenha certeza se o diagnóstico é d_3 ou d_4 , pode-se usar tratamento T_6 - FOLICUR 200 CE, pois é um fungicida que serve para controlar as duas desordens.

Além dos dois exemplos apresentados neste capítulo, os algoritmos da TCN foram aplicados para as outras 38 desordens de milho e os resultados obtidos foram coerentes com o raciocínio do diagnosticador.

Finalmente, pode-se concluir que foram satisfatórios os resultados obtidos na validação da TCN no domínio de doenças de milho. A TCN atende os requisitos de inferência, incertezas e tomada de decisão necessários para suportar o processo de raciocínio clínico. Adicionalmente, a TCN se mostra adequada à solução dos problemas de diagnóstico, investigação e tratamentos em outros domínios de diagnóstico dinâmico, como por exemplo, no diagnóstico e controle de falhas de satélite.

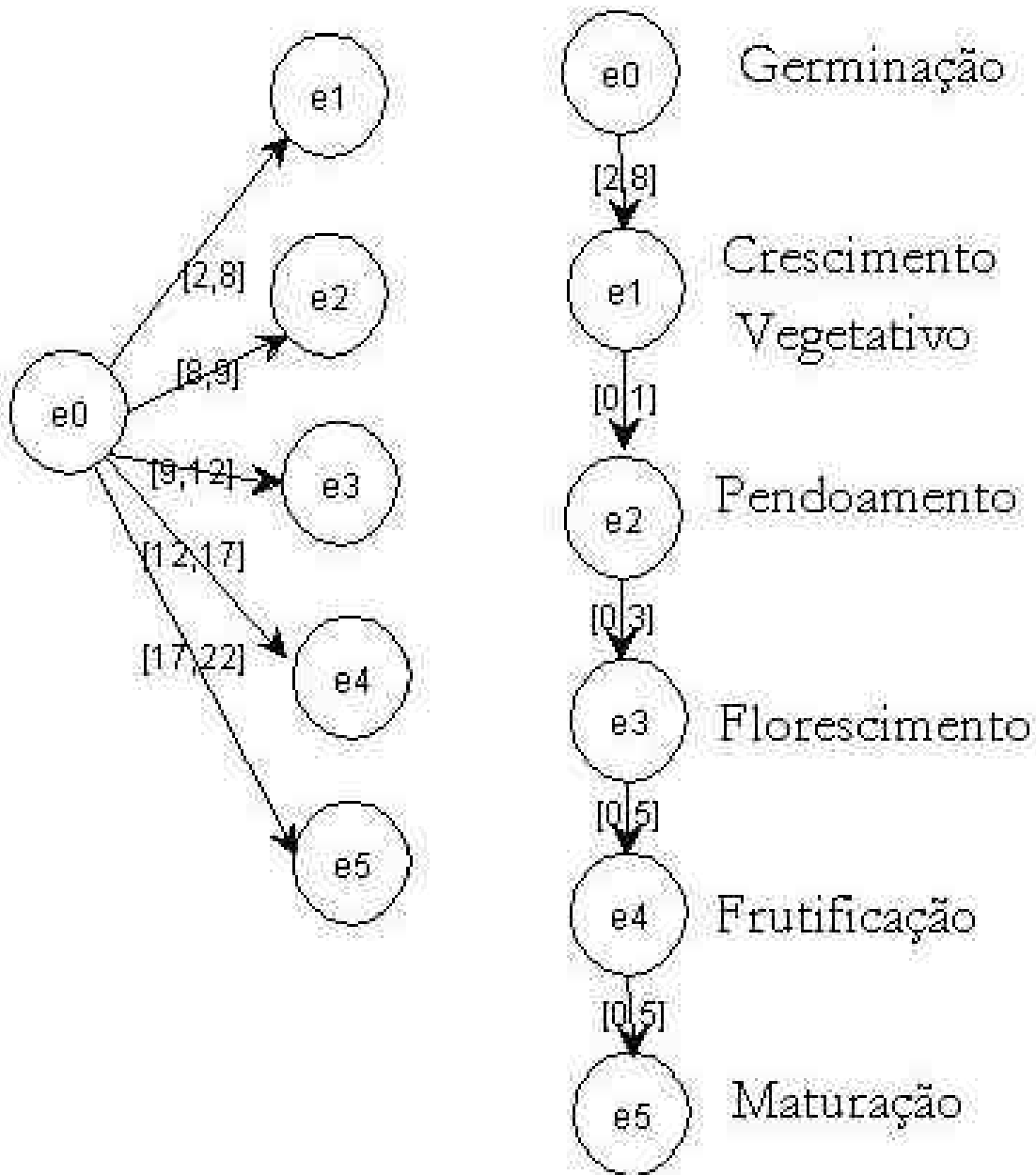


FIGURA 8.2 – Grafo temporal da cultura do milho.

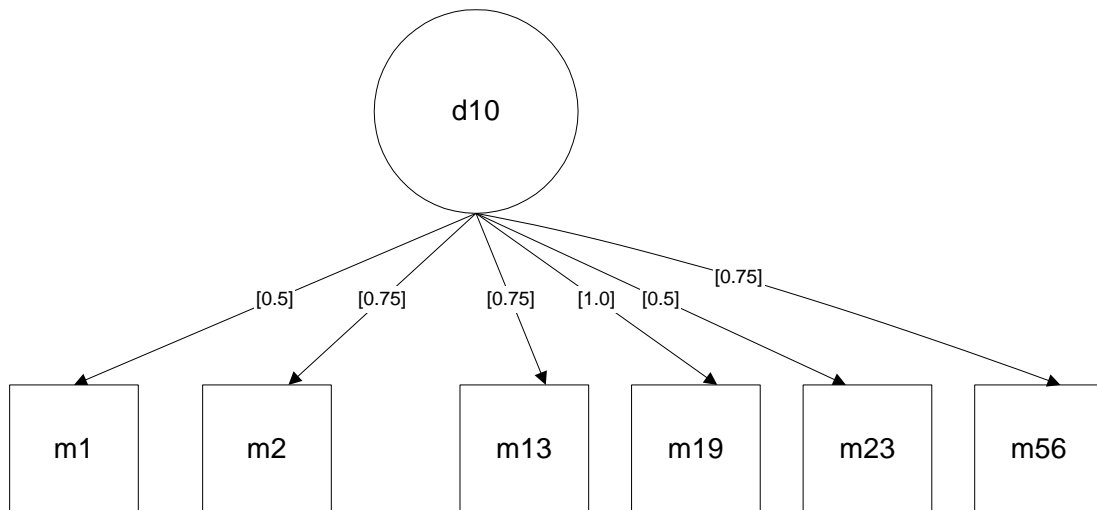


FIGURA 8.3 – Associação gradual $R = \langle d_{10}, m_j \rangle$.

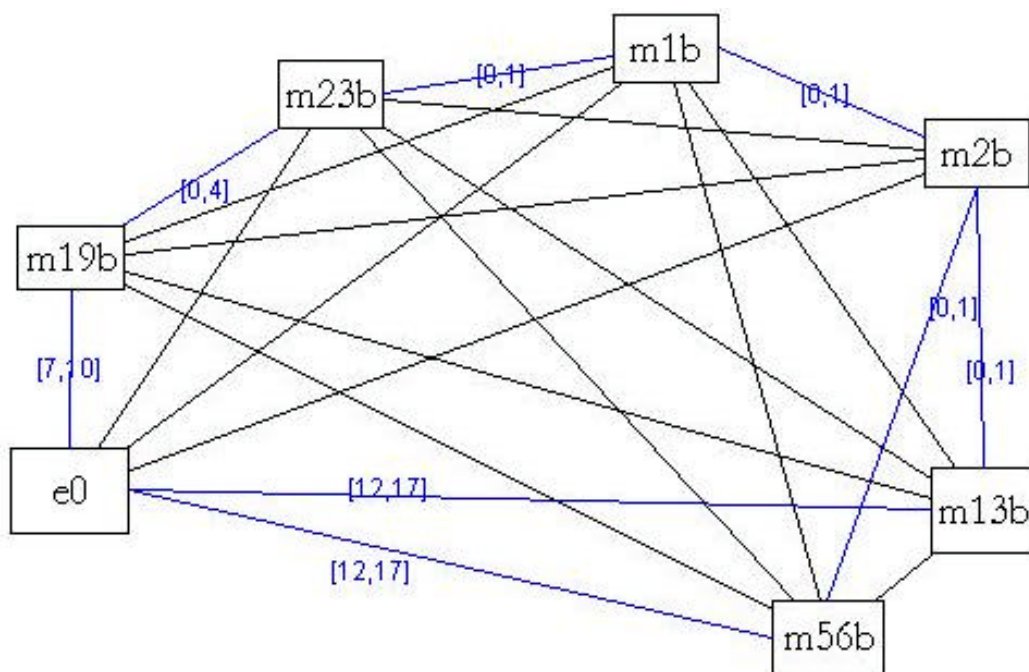


FIGURA 8.4 – Grafo mínimo de eventos de d_{10} .

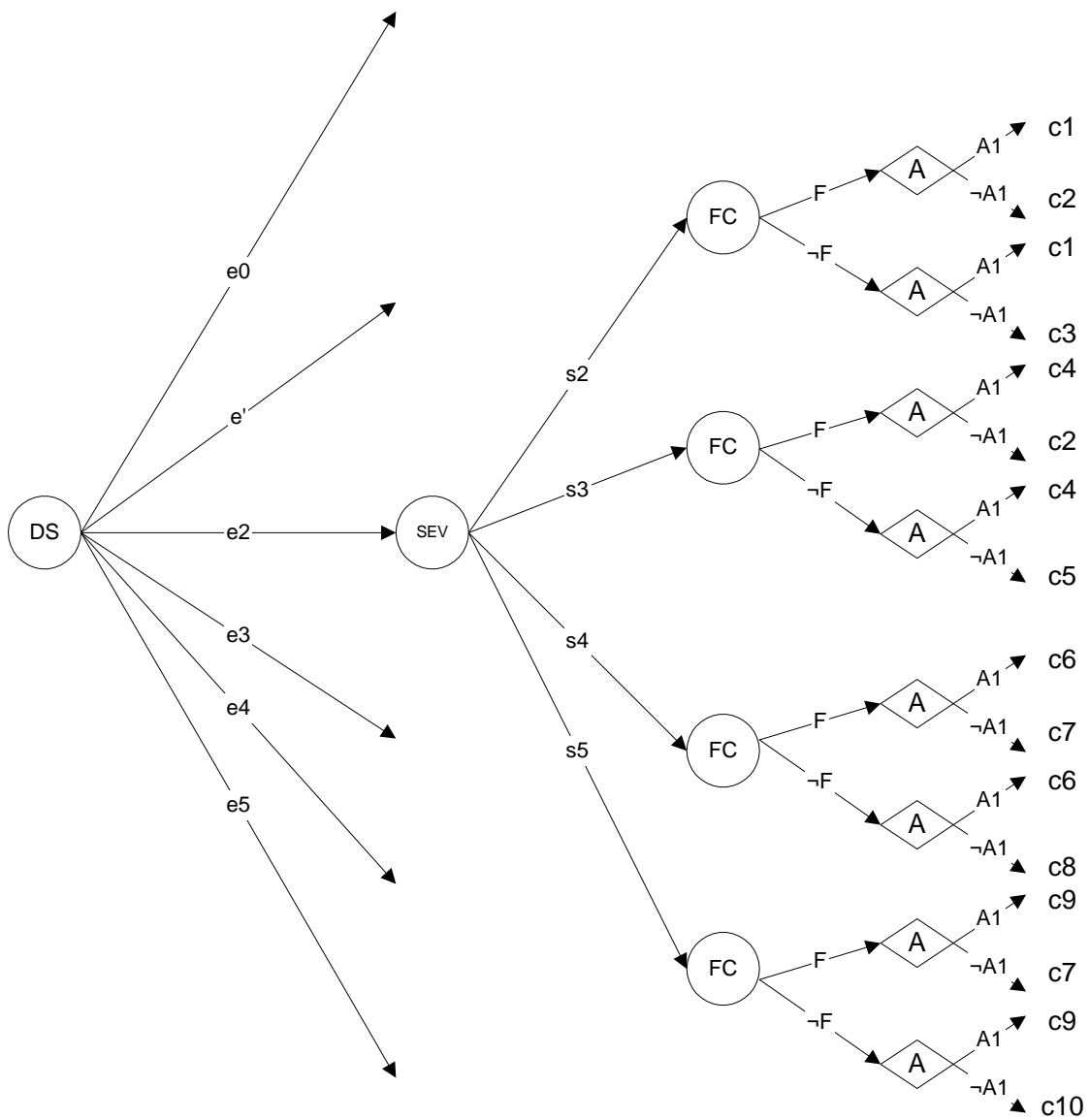


FIGURA 8.5 – Exemplo de uma árvore de decisão no tratamento da Cercosporiose.

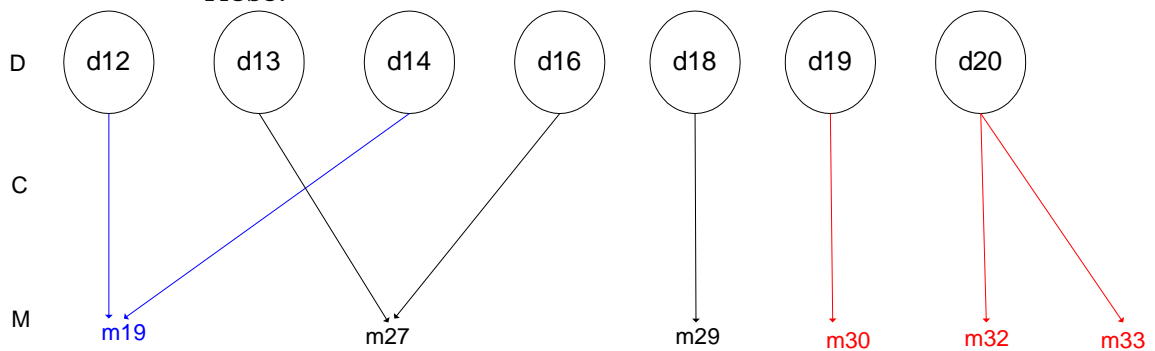


FIGURA 8.6 – Rede causal de um problema de diagnóstico de falhas em satélite <D,M,C>.

CAPÍTULO 9

CONCLUSÕES

Primeiramente, neste capítulo são apresentadas as principais contribuições do modelo proposto e suas limitações. Na sequência são apresentadas as extensões futuras que devem ser suportadas pela TCN.

9.1 Contribuições

Neste trabalho foi proposta a Teoria das Coberturas Nebulosas (TCN) que utiliza inferência abdutiva e lógica nebulosa para solucionar um problema de raciocínio clínico envolvendo tanto diagnóstico e investigação quanto tratamentos. Através da TCN foi possível mostrar que o processo de raciocínio clínico como um todo pode ser automatizado utilizando-se uma abordagem abdutiva causal. De fato, todo o conhecimento do diagnosticador pode ser representado através de associações causais e os diagnósticos, exames e tratamentos mais plausíveis para o processo de raciocínio clínico podem ser inferidos abductivamente.

Embora tarefas para diagnóstico e tratamentos estejam fortemente relacionadas, elas não são frequentemente investigadas juntas. Consequentemente, aspectos relevantes para diagnóstico e tratamentos não são considerados. Mecanismos de tomada de decisão necessários na etapa de tratamentos têm sido estudados mais detalhadamente na área de Teoria de Decisão enquanto mecanismos de inferência em diagnóstico têm sido mais pesquisados na área de Inteligência Artificial.

Neste trabalho foi possível mostrar que todos estes aspectos podem ser abordados de maneira integrada, permitindo-se a organização dos vários tipos de informação gerados nas várias tarefas.

A incorporação de conceitos de lógica nebulosa e teoria de decisão à TCN, minimizou uma das principais limitações da teoria TCP, que é a possibilidade de incluir um grande número de hipóteses.

Neste trabalho, também foi feita a validação da abordagem proposta utilizando-se um exemplo real em fitopatologia. Os resultados obtidos foram satisfatórios tanto no escopo de diagnóstico e investigação quanto no escopo de tratamentos.

No escopo de diagnóstico, a TCN permitiu a justificativa de respostas na solução do problema além da redução de hipóteses. A justificativa de resposta refere-se à

habilidade de um sistema baseado em conhecimento em explicar como e porque chegou às conclusões como diagnose. É muito importante desenvolver métodos para justificativa de respostas, pois, estudos demonstram que, por exemplo, na medicina e na fitopatologia, há necessidade de boas explicações por sistemas de diagnóstico para aceitação pelos diagnosticadores. Na TCN, por exemplo, o grau de consistência entre uma hipótese (conjunto de desordens) e um caso, permitem aos usuários quantificar o quanto aquela hipótese é plausível para explicar o caso.

No escopo de investigação, pode-se destacar as funções de entropia desenvolvidas na TCN que permitem selecionar as manifestações prioritárias mesmo que não se tenha informações completas sobre elas. Esta funcionalidade que está sendo incorporada através das funções de entropia propostas no capítulo 6, além de contribuir na seleção de manifestações prioritárias e redução de hipóteses, pode ser útil em um módulo de geração de perguntas. A habilidade de gerar boas perguntas a partir da informação disponível pode ser crucial na eficiência do problema.

Na literatura são encontrados muitos trabalhos na área de teoria de decisão para problemas de tomada de decisão como seleção de tratamentos. Entretanto, a abordagem abdutiva integrada a conjuntos nebulosos gera um método alternativo para solução de problemas na área de tratamentos mais vantajoso porque aproveita os resultados gerados pelo módulo de diagnóstico para reduzir a complexidade de determinar um tratamento.

No escopo geral da TCN, a integração dos processos de diagnóstico, investigação e tratamentos em uma abordagem abdutiva integrada a lógica nebulosa consolida uma nova abordagem em IA para o processo de raciocínio clínico.

É importante destacar como o formalismo proposto na TCN para representar o conhecimento e o caso considera aspectos importantes como o conhecimento incerto e incompleto inerentes ao processo de raciocínio clínico, além do formalismo proposto para suportar o mecanismo de inferência cujo o fundamento teórico está bem sedimentado na lógica abdutiva.

Finalmente, é válido ressaltar que este trabalho trouxe contribuições tanto para o avanço científico na área da ciência computação (Massruhá et al., 2001, 2002a, 2002b), quanto em áreas de diagnóstico dinâmico como por exemplo, na área agrícola (Massruhá et al., 2003).

Na área agrícola, espera-se que este trabalho de pesquisa contribua para melhorar o processo de aquisição, formalização e representação do conhecimento dos fitopatologistas

e veterinários, e, conseqüentemente, o desenvolvimento de sistemas de suporte à decisão em duas áreas prioritárias na agricultura, a saber: fitossanidade (Agronomia) e sanidade animal (Veterinária).

9.2 Limitações

Embora a TCN seja uma ferramenta simples e poderosa para raciocínio de diagnóstico, investigação e tratamentos, o modelo proposto tem algumas limitações.

Uma de suas limitações está no escopo do problema apresentado na figura 4.1. Conforme descrito no capítulo 4.1, em complementação à decisão terapêutica deve-se ter um plano educacional do paciente e um procedimento de monitoramento dos tratamentos (efeitos do tratamento na progressão da desordem). Como se trata de um processo cíclico e dinâmico, os resultados do monitoramento devem modificar constantemente a síntese do problema. Entretanto, a retroalimentação do processo de raciocínio clínico através da automatização do procedimento de monitoramento não está incluído no escopo deste trabalho.

Outra limitação encontra-se na incapacidade de representar formas mais complexas de relacionamento causal tanto no diagnóstico e na investigação quanto nos tratamentos. Por exemplo, na TCN assume-se que as manifestações de duas desordens não interferem uma na outra. Não é possível representar a interferência da presença de uma desordem nas manifestações de outra desordem, ou que se duas desordens ocorrem simultaneamente pode acontecer uma manifestação não prevista em uma dessas desordens isoladamente. Similar a etapa de diagnóstico, a TCN assume que os testes laboratoriais e os tratamentos de duas desordens não interferem um no outro, isto é, não trata incompatibilidades entre desordens, exames ou tratamentos.

Os outros problemas da TCN estão relacionados com a sua implementação. A geração da base de conhecimento $KB = \{KB_d, KB_e, KB_t\}$ demanda muito tempo e atenção. Portanto, seria interessante uma ferramenta que pudesse auxiliar na geração automática da base de conhecimento. Outro ponto que poderia ser melhor tratado seria a integração bem como a interface dos 3 módulos em Prolog que compõem a TCN (*bipartido – nebuloso.pro, exames – nebulosos.pro* e *tratamentos – nebulosos.pro*) que são executados separadamente.

9.3 Trabalhos Futuros

As futuras extensões deste trabalho poderiam ser divididas em três níveis de atuação: especificação, implementação e validação.

No nível de especificação é importante estender a TCN para incorporar o procedimento de monitoramento, isto é, os efeitos do tratamento na progressão da desordem. Outra extensão importante ao modelo da TCN seria incorporar incompatibilidades entre desordens, exames e tratamentos.

Em relação á implementação seria importante desenvolver uma interface que integre os 3 módulos que compõem a TCN. No escopo deste trabalho já foram realizados testes iniciais para integração de uma interface JAVA aos programas em PROLOG, sem necessidade de alterar os módulos já implementados. Um outra extensão interessante seria a geração automática da base de conhecimento a partir de entrevistas com usuários. Esta ferramenta poderia ser desenvolvida, por exemplo, em uma dissertação de mestrado.

Quanto á validação, seria interessante utilizar a TCN na solução de problemas de raciocínio clínico em outros domínios, como por exemplo em medicina, veterinária e mais profundamente na área de satélites. Neste sentido pode-se formalizar projetos entre instituições como Embrapa, Unicamp e INPE de modo a viabilizar estas aplicações.

Um outro projeto interessante que poderia ser apresentado futuramente seria um trabalho da Embrapa junto ao Ministério da Agricultura e Abastecimento de modo a utilizar a TCN para integrar as bases de conhecimento de desordens e testes laboratoriais da Embrapa com a base de fungicidas, acaricidas, bactericidas disponíveis no Ministério.

REFERÊNCIAS BIBLIOGRÁFICAS

- A822007-ANL-61/D1. **Risk Analysis for the Multi-Mission Platform.** Instituto Nacional de Pesquisas Espaciais - INPE, 2001. [31](#), [160](#)
- Agrofit-OnLine. **Relatório de Acaricidas, Bactericidas, Fungicidas, Inseticidas ...** Ministério da Agricultura, Pecuária e Abastecimento, 2003. [145](#)
- Bergamin, A.; L., A. **Doenças de Plantas Tropicais: Epidemiologia e Controle Econômico.** SãoPaulo: Editora Agronômica Ceres Ltda., 1996. [147](#)
- Cruz, S.; Leite, M. A.; Moura, M.; Souza, E. ServCLIPS: A Tool to Develop WWW Expert Systems. In: World Congress on Computer in Agriculture and Natural Resources, 2001, Iguazu Falls. **Proceedings...Iguazu Falls: ASAE.**, p. 636-641, 2001. [31](#)
- Dubois, D.; Prade, H. **Possibility Theory: An Approach to Computerized Processing of Uncertainty.** 1. ed. New York: Plenum Press Inc., 1988. [37](#), [142](#), [143](#), [179](#), [182](#)
- . Fuzzy Relation Equations and Causal Reasoning. **Fuzzy Sets and Systems**, v. 75, p. 417–134, 1995. [62](#), [73](#), [82](#)
- Fernandes, F.; Oliveira, E. **Principais Doenças na Cultura do Milho.** 1. ed. Brasília: EMBRAPA-CNPMS / EMBRAPA-SPI, 90p, 1997. [29](#), [30](#), [145](#), [148](#), [191](#)
- Henrion, M. Breeze, J. S.; Horvitz, E. J. Decision Analysis and Expert Systems. **AP Magazine**, v. 13, n. 8, p. 64–91, 1992. [40](#), [48](#)
- Klir, G. J.; Folger, T. A. **Fuzzy Sets, Uncertainty, and Information.** 1. ed. Prentice Hall, Englewood Cliffs, New Jersey 07632, 1988. [121](#), [122](#), [181](#)
- Lucas, P. Modelling Interactions for Diagnosis. In: CESA '96 IMACS Multiconference: Symposium and Modelling, Analysis and Simulation. **Proceedings ...**, v. 1, p. 541-546, 1996. [28](#), [62](#), [82](#)
- Malcolm, C. **Compendium of Corn Diseases.** 2. ed. Minnesota (USA):The American Phytopathological Society, 1980. [145](#), [148](#), [191](#)
- Massruhá, S.; Cruz, S.; Romani, L.; Souza, E. Remote Diagnosis: An Expert System for Diagnosis of the Corn Diseases by Web. In: The Inaugural Australian Workshop on the Application of Artificial Intelligence to Plant and Animal Production - AWAPAP'99, 1., 1999, Sidney. **Proceedings...** Sidney: Australia, 1999. [30](#)

Massruhá, S.; Sandri, S.; Wainer, J. Uma Extensão Nebulosa para Teoria das Coberturas Parcimoniosas Aplicada ao Diagnóstico de Doenças de Plantas. In: Workshop dos Cursos de Computação Aplicada do INPE, 1., 2001, São José dos Campos. **Anais do I WORCAP**. São José dos Campos: INPE, 2001. p. 97-99. [172](#)

———. Teoria das Coberturas Nebulosas: Uma Abordagem Abdutiva Difusa para Solução de Problemas de Diagnóstico. In: Workshop dos Cursos de Computação Aplicada do INPE, 2., 2002, São José dos Campos. **Anais do II WORCAP**. São José dos Campos: INPE, 2002a. p. 11-16. [172](#)

———. Um Modelo Alternativo Abduativo Nebuloso para Solução de Problemas de Diagnóstico. In: Workshop de Teses e Dissertações em Inteligência Artificial - WTDIA'02, 1., Brazilian Symposium on Artificial Intelligence - SBIA'02, 16., 2002, Recife. **Proceedings... . 1 CD-ROM**. Recife: Sociedade Brasileira de Computação, 2002b. p. 61-70. [172](#)

———. Fuzzy Covering Theory: An Alternative Approach for Diagnostic Problem-Solving. In: European Conference for Information Technology in Agriculture - EFITA 2003, 4th., 2003, Budapeste. **Proceedings...** Budapeste: HAAI, 2003. p. 768–775. [172](#)

McIlraith, S. Logic-Based Abductive Inference. **Revision of an earlier draft written while the author was doctoral candidate at the University of Toronto, Stanford, CA: Stanford University**, p. 29p, 1998. [28](#), [38](#), [51](#), [52](#), [53](#)

Morandi, M. B. **Monitoramento das Principais Doenças Foliares do Milho e Controle Químico da Cercosporiose (Cercospora Zeae-Maydis) em Função do Nível de Severidade da Doença**. Jaguariúna: Embrapa Meio Ambiente, 2002. [147](#), [148](#), [153](#), [167](#)

Peng, Y.; Reggia, J. A. **Abductive Inference Models for Diagnostic Problem-Solving**. 1. ed. New York: SpringerVerlag New York Inc., 1990. [27](#), [28](#), [35](#), [36](#), [38](#), [42](#), [43](#), [53](#), [54](#), [58](#), [59](#), [61](#), [62](#), [82](#), [86](#), [100](#)

Raiffa, H. **Decision Analysis: Introductory Lectures on Choices under Uncertainty**. 1. ed. Addison-Wesley Publishing Co. Inc., 1970. [46](#)

Réa-Neto, A. O Raciocínio Clínico - O Processo de Decisão Diagnóstica e Terapêutica. **Rev Ass Med Brasil**, p. 301–311, 1998. [75](#), [76](#), [78](#), [79](#), [80](#)

Reis, E.; Casa, R. T. **Manual de Identificação e Controle de Doenças de Milho**. 1. ed. Passo Fundo: Aldeia Norte Editora, 80p, 1996. [145](#)

Savage, L. **The Foundations of Statistics**. 1. ed. New York: Dover Publications., 1972. [44](#), [45](#)

Shenoy, P. Valuation-Based Systems for Bayesian Decision Analysis. **Operations Research.**, v. 10, n. 3, p. 864-784, may-june 1992. [47](#), [48](#)

Shortliffe, E. **Computer-Based Medical Consultation; MYCIN**. Elsevier, 1976. [40](#), [81](#)

Wainer, J.; Rezende, A. A Temporal Extension to the Parsimonius Covering Theory. **Artificial Intelligence in Medicine**, v. 10, p. 235-255, 1997. [62](#), [63](#), [72](#), [82](#), [83](#), [145](#), [165](#)

Wainer, J.; Sandri, S. A Fuzzy Temporal/Categorical Information in Diagonis. **Special Issue on Inteligent Temporal Information Systems in Medicine**, p. 1–19, 1998. [62](#), [63](#), [64](#), [66](#), [67](#), [68](#), [69](#), [70](#), [71](#), [72](#), [73](#), [83](#), [84](#), [86](#), [88](#), [102](#), [107](#), [108](#), [109](#), [139](#)

Yager, R. Families of OWA Operators. **Fuzzy Sets and Systems**, v. 59, p. 925–148, 1993. [119](#)

APÊNDICE A

CONCEITOS BÁSICOS DA TEORIA DE CONJUNTOS NEBULOSOS E TEORIA DE POSSIBILIDADES

A.1 Teoria de Conjuntos Nebulosos

Primeiramente, um conjunto nebuloso é definido formalmente conforme descrito em Dubois e Prade (1988):

Definição A.1 *Um conjunto nebuloso (normalizado) A em Θ é caracterizado pela função de pertinência $\mu_A : \theta \rightarrow [0, 1]$, tal que $\exists x \in \Theta, \mu_A(x) = 1$.*

Sejam A e B conjuntos nebulosos em θ , onde θ é um domínio para o qual as operações $+$, \min e $-$ são definidas.

Definição A.2 *A soma $A \oplus B$, a subtração $A \ominus B$, a negação $-A$, e a intersecção $A \cap B$ são respectivamente caracterizadas pelas funções de pertinência (Dubois e Prade, 1988):*

$$m_{A \oplus B}(z) = \sup_{\{(x,y)/z=x+y\}} \min(\mu_A(x), \mu_B(y))$$

$$m_{A \ominus B}(z) = \sup_{\{(x,y)/z=x-y\}} \min(\mu_A(x), \mu_B(y))$$

$$\mu_{-A}(z) = \mu_A(-z)$$

$$m_{A \cap B}(z) = \min(\mu_A(z), \mu_B(z)).$$

Assim, a altura de um conjunto nebuloso pode ser definida como a seguir.

Definição A.3 *A altura de um conjunto nebuloso A não-normalizado é calculado como: $h(A) = \sup_{x \in \theta} \mu_A(x)$.*

Outra definição importante no contexto deste trabalho, um conjunto nebuloso A tal que μ_A é convexo, é chamado de intervalo nebuloso. A definição de convexo e positivo retirada de Dubois e Prade (1988) estão descritas a seguir .

Definição A.4 *A é convexo se somente se:*

$$\forall x_1 \in X, \forall x_2 \in X, \forall \lambda \in [0, 1]$$

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2) \geq \min(\mu_A(x_1), \mu_A(x_2))$$

Definição A.5 *Um intervalo será positivo se θ é a linha real, e $\forall x < 0, \mu(x) = 0$.*

Em alguns casos será assumido que o intervalo nebuloso é trapezoidal. Neste caso o intervalo será representado por uma quadrúpla $\langle a, b, c, d \rangle$, como representado na Figura A.1.

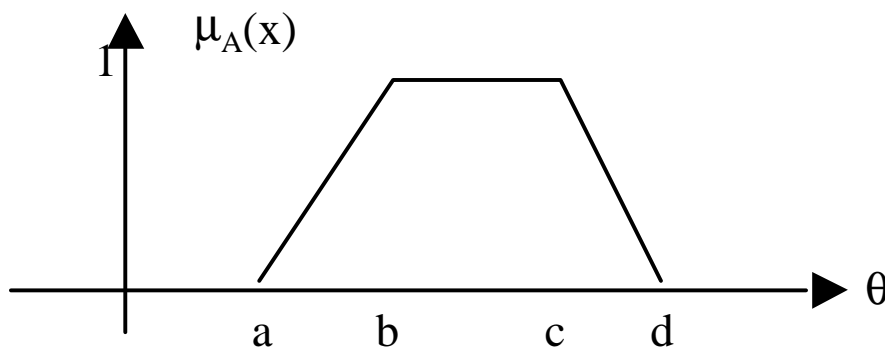


FIGURA A.1 – Intervalo nebuloso trapezoidal.

Definição A.6 Em um intervalo trapezoidal $A = \langle a_1, a_2, a_3, a_4 \rangle$ o intervalo $[a_2, a_3]$, onde $\mu_A(x) = 1$, é chamado de núcleo (core) de A . O intervalo $[a_1, a_4]$, onde $\mu_A(x) > 0$, é chamado de suporte de A . Um intervalo $\langle a, a, b, b \rangle$ é dito ser bem definido e é representado por $\langle a, b \rangle$.

Definição A.7 Sejam A e B dois intervalos trapezoidais definidos por $A = \langle a_1, a_2, a_3, a_4 \rangle$ e $B = \langle b_1, b_2, b_3, b_4 \rangle$. As operações \oplus e \ominus correspondem a $A \oplus B = \langle a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4 \rangle$ e $A \ominus B = \langle a_1 - b_4, a_2 - b_3, a_3 - b_2, a_4 - b_1 \rangle$, respectivamente.

Finalmente, um intervalo A é definido ser mais específico que um intervalo B , se $\mu_A(x) \leq \mu_B(x)$ para todo $x \in \Theta$. Se A e B são trapezoidais então $A = \langle a_1, a_2, a_3, a_4 \rangle$ é mais específico que $B = \langle b_1, b_2, b_3, b_4 \rangle$, se somente se $a_1 \geq b_1, a_2 \geq b_2, a_3 \leq b_3, a_4 \leq b_4$.

A seguir, são apresentadas algumas definições baseadas nas operações em conjuntos nebulosos.

Definição A.8 Sejam A e B conjuntos nebulosos definidos em Ω . A intersecção destes conjuntos, $E = A \cap B$, é implementada por uma família de operadores denominados de t – normas e a união $F = A \cup B$ é implementada por uma família de

operadores denominados de t -conormas ou S -normas.

Uma função $\nabla : [0, 1]^2 \rightarrow [0, 1]$, é dita ser comutativa, associativa e monotônica se ∇ satisfaz as seguintes propriedades, respectivamente, para $\forall a, b \in [0, 1]$:

Comutatividade: $\nabla(a, b) = \nabla(b, a)$;

Associatividade: $\nabla(a, \nabla(b, c)) = \nabla(\nabla(a, b), c)$;

Monotonicidade: $\nabla(a, b) \leq \nabla(c, d)$ se $a \leq c$ e $b \leq d$.

Um operador $\top : [0, 1]^2 \rightarrow [0, 1]$ é denominado de t -norma se \top é comutativo, associativo e monotônico e verifica a seguinte propriedade, para $\forall a \in [0, 1]$:

Elemento neutro = 1: $\top(a, 1) = a$.

Da mesma maneira, uma t -conorma \perp é uma função $\perp : [0, 1]^2 \rightarrow [0, 1]$ que é comutativa, associativa e monotônica e verifica a propriedade, para $\forall a \in [0, 1]$:

Elemento neutro = 0: $\perp(a, 0) = a$.

Uma t -norma \top e uma t -conorma \perp são duais em relação a uma operação de negação $\neg : [0, 1] \rightarrow [0, 1]$ se elas satisfazem as relações de De Morgan, dadas por, para $\forall a, b \in [0, 1]$:

Leis de De Morgan:

$$\neg(\top(a, b)) = \perp(\neg a, \neg b);$$

$$\neg(\perp(a, b)) = \top(\neg a, \neg b).$$

O principal operador de negação é dado por $\neg a = 1 - a$, mas outros operadores podem ser encontrados na literatura (Klir e Folger, 1988).

É importante notar que as t -normas e t -conormas se reduzem aos operadores clássicos de união e intersecção quando os conjuntos são “crisp”.

A Tabela A.1 indica as t -normas e t -conormas mais utilizadas em relação a dois conjuntos nebulosos A e B . Estes operadores satisfazem as leis de De Morgan em relação ao operador de negação $\neg a = 1 - a$.

TABELA A.1 – Principais \top -normas e \top -conormas duais.

t – norma	t – conorma	nome
$\min(a, b)$	$\max(a, b)$	Zadeh
$a \cdot b$	$a + b - ab$	probabilista
$\max(a + b - 1, 0)$	$\min(a + b, 1)$	Lukasiewicz
$\begin{cases} a, & \text{se } b = 1 \\ b, & \text{se } a = 1 \\ 0, & \text{senão} \end{cases}$	$\begin{cases} a, & \text{se } b = 0 \\ b, & \text{se } a = 0 \\ 1, & \text{senão} \end{cases}$	Weber

A maior (respec. a menor) t -norma é o \min (respec. a t -norma de Weber); a maior (respec. a menor) t -conorma é a t -conorma de Weber (respec. o \max). Os operadores mais usualmente utilizados são os operadores de Zadeh e os probabilistas.

No decorrer deste trabalho, além de alguns conceitos de conjuntos nebulosos, é necessário alguma noção de teoria de possibilidades para representar incertezas. A seguir, serão descritas as definições das medidas de possibilidade e necessidade para conjuntos bem-definidos e para conjuntos nebulosos segundo Dubois e Prade (1988).

A.2 Teoria de Possibilidades

Definição A.9 Dado um conjunto finito (crisp) Ω , a medida de possibilidades Π pode ser definida em termos de uma distribuição de possibilidades $\pi : \Omega \rightarrow [0, 1]$, onde

$$\forall A, \Pi(A) = \sup\{\pi(x) | x \in A\}$$

e $\Pi(A)$ é normalizado, isto é,

$$\exists x \in \Omega, \pi(x) = 1 | \Pi(\Omega) = 1$$

Definição A.10 Dado um conjunto finito (crisp) Ω , a medida de necessidade N é o dual de Π e pode ser definida em termos de uma distribuição de possibilidades $\pi : \Omega \rightarrow [0, 1]$, onde

$$\forall A, N(A) = \inf\{1 - \pi(x) | x \notin A\}$$

Definição A.11 Dado um conjunto nebuloso (normalizado) A em Ω caracterizado pela função de pertinência $\mu_A : \Omega \rightarrow [0, 1]$, tal que $\exists x \in \Theta, \mu_A(x) = 1$, a medida de possibilidade Π pode ser definida em termos de uma distribuição de possibilidades tal que:

$$\forall A, \Pi(A) = \max_{x \in A} \min(\mu_A(x), \pi(x))$$

.

Definição A.12 Dado um conjunto nebuloso (normalizado) A em Ω caracterizado pela função de pertinência $\mu_A : \Omega \rightarrow [0, 1]$, tal que $\exists x \in \Theta, \mu_A(x) = 1$, a medida de necessidade N pode ser definida em termos de uma distribuição de possibilidades tal que:

$$\forall A, N(A) = 1 - \max_{x \in A} \min(\mu_{\bar{A}}(x), \pi(x)) = \min_{x \in A} \max(\mu_A(x), 1 - \pi(x))$$

.

APÊNDICE B

IMPLEMENTAÇÃO DO MODELO PROPOSTO

Nesta capítulo estão descritos os algoritmos para a solução do processo de raciocínio clínico. Primeiramente, é descrito o algoritmo principal TCN que implementa a execução apresentada na figura 4.6 e executa os outros algoritmos. Na sequência é apresentado o algoritmo *Bipartido – Nebuloso* que é baseado no algoritmo bipartido apresentado na seção 3.2 e o algoritmo *VNCD– Verificador Nebuloso de Consistência para Diagnóstico* que implementa a solução para problemas de diagnóstico apresentada na seção 5.3. Em seguida, são apresentados os algoritmos *Exames – Nebulosos* e *VNCI– Verificador Nebuloso de Consistência para Investigação* que implementam a solução para Investigação apresentada na seção 6.3. Finalmente, são apresentados os algoritmos *Tratamentos – Nebulosos* e *VNCT– Verificador Nebuloso de Consistência para Tratamentos* que implementam a solução para tomada de decisão em tratamentos apresentada na seção 7.3.

B.1 Algoritmos

O primeiro algoritmo para implementação do problema de raciocínio clínico descrito é a função TCN que executa todos os outros algoritmos necessários para solução do problema.

```
function TCN( $KB = (KB_d, KB_e, KB_t), CA$ )
```

```
variables
```

```
     $SOL_d$ : generator-set;
```

```
     $SOL_e$ : generator-set;
```

```
     $SOL_t$ : generator-set;
```

```
begin
```

```
     $SOL_d = \{ \}$ ;
```

```
     $SOL_e = \{ \}$ ;
```

```
     $SOL_t = \{ \}$ ;
```

```
     $SOL_d = \text{Bipartido – Nebuloso}(\theta, D, M, R, G(V, T), CC(CF, SEV, EP), CA)$ 
```

```
    while prompt  $\neq$  "Tratamento" do
```

```
        if  $INVEST \neq \emptyset$ 
```

```
             $SOL_e = \text{exames – nebulosos}(Ex, I, INVEST, M, R)$ 
```

```
             $C_{atual} = \text{atualizar}_{\text{caso}}(SOL_e, result)$ 
```

```
             $SOL_d = \text{Bipartido – Nebuloso}(\theta, D, M, R, G(V, T), CC(CF, SEV, EP), C_{atual})$ 
```

```
        endif;
```

```
    endwhile;
```

```

    SOLt = tratamentos – nebulosos(T, D, Z, V, SOLd)
    return SOLd, SOLt;
end.

```

A rotina acima corresponde a implementação da figura 4.6. Observe que o usuário tem a opção de executar diretamente o módulo de tratamentos após obter as hipóteses iniciais na execução do módulo de diagnóstico ou ele pode investigar melhor as manifestações no módulo de investigação se houver alguma desordem onde $\alpha(d_i) < \Delta$.

B.1.1 Algoritmos para Diagnóstico

O próximo algoritmo a ser apresentado é o *Bipartido – Nebuloso* que é baseado no algoritmo bipartido apresentado na seção 3.2. Entretanto, antes de aplicar os critérios de parcimônia de irredundância através da função *revise* é necessário verificar a consistência temporal e categórica do caso particular em relação ao modelo de desordens que é implementada através da função VNCD apresentada a seguir.

```

function Bipartido – Nebuloso( $\theta$ , D, M, R, G(V, T), CC(CF, SEV, EP), CA)

```

```

variables

```

```

    mnova: manifestation;

```

```

    Hipoteses: generator-set;

```

```

begin

```

```

    Hipoteses = { };

```

```

    while Maismanifs do

```

```

        mnova=Proxman;

```

```

        H = causas(mnova);

```

```

        Eventos (mnova) = (Ev+, Time+);

```

```

        VNCD(H, G, Ev+, Time+, mj);

```

```

        hipoteses=revise(Hipoteses, (H - DI));

```

```

    endwhile;

```

```

return Hipoteses, IND, INVEST;

```

```

end.

```

O algoritmo VNCD que implementa a solução para problema de diagnóstico está descrita abaixo.

```

function VNCD(D, G, Ev+, Time+, mj)

```

```

variables

```

```

DI: set – of – disorders;

```

```

eventos – antecessores, eventos – sucessores, eventos – vizinhos: set – of – arcs;

```

```

IND, INVEST: set – of – disorders – index;

```

```

inconsistente: boolean;
begin
   $DI = \emptyset$ ;
   $Ind = \emptyset$ ;
  while  $D \neq \emptyset$  do
     $d \in D$ ;
     $ev_j \in Ev^+$ ;
     $G_d = (V(d), T(d)) \in G$ 
     $\alpha(d)_{anterior}, \alpha(d)_{atual}, \alpha(d)_t = 0$ ;
     $eventos - antecessores = (evi, evj) | (evi, evj) \in T(d) \text{ e } Time + (evi) \text{ existe}$ ;
     $eventos - sucessores = (evj, evk) | (evj, evk) \in T(d) \text{ e } Time + (Evk) \text{ existe}$ ;
     $eventos - vizinhos = eventos - antecessores \cup eventos - sucessores$ ;
    inconsistente = false;
    while  $eventos - vizinhos \neq \emptyset$  and not inconsistente do
       $(ev_1, ev_2) \in eventos_{vizinhos}$ ;
       $\pi(ev_1, ev_2) \in T(d)$ ;
       $Dist^+(ev_1, ev_2) = Time^+(ev_2) \ominus Time^+(ev_1)$ ;
      if  $((Dist^+(ev_1, ev_2) \cap \pi(ev_1, ev_2)) = \emptyset$  then
        inconsistente = false;
         $DI = DI \cup \{d\}$ ;
      else
         $\alpha(d)_{anterior} = \alpha(d)_t$ 
         $\alpha(d)_t = (h(Dist^+(ev_1, ev_2) \cap \pi(ev_1, ev_2)))$ 
         $\alpha(d)_{atual} = \min(\alpha(d)_{anterior}, \alpha(d)_t)$ 
      endif
       $eventos - vizinhos = eventos - vizinhos - (ev_1, ev_2)$ ;
    endwhile;
     $D = D - d$ ;
    If  $\alpha(d)_{atual} > 0$  then  $IND = IND \cup \alpha(d)_{atual}$ ;
    If  $\alpha(d)_{atual} < \Delta$  then  $INVEST = INVEST \cup d$ ;
  endwhile
return  $DI, IND, INVEST$ ;
end.

```

Na sequência são apresentados os algoritmos *Exames – Nebulosos* e *VNCI* que implementam a solução para problema de investigação.

B.1.2 Algoritmos para Investigação

function *Exames – Nebulosos*((*Ex*, *I*, *D*, *M*, *R*))

variables

mnova: manifestation;

Exames: generator-set;

begin

Exames = { };

Manifs = *efeitos*(*d_i*) | *d_i* ∈ *D*;

Ordmanifs = *VNCI*(*Manifs*, *D*, *R*);

 while *Ordmanifs* do

m_{nova} = *Proxman*;

Ex = *exames*(*m_{nova}*);

Exames = *revise*(*Exames*, *Ex*);

 endwhile;

return *Exames*;

end.

function *VNCI*(*Manifs*, *D*, *R*)

variables

mforca: set-of-manifestations;

begin

mforca = { };

 while *Manifs* ≠ ∅ do

forca_{ant} = *forca*(*m_j*);

forca(*m_j*) = [*f*(*p₁*, ..., *p_n*) * *n_j*(*p*)] + (1 - *max*(*p_{i=1,.., n}*));

forca_{maior} = *min*(*forca_{ant}*, *forca*(*m_j*)) ;

mforca = *mforca* ∪ *m_j* ← *forca_{maior}*;

 endwhile;

return *mforca*;

end.

Finalmente, são apresentados os algoritmos *Tratamentos – Nebulosos* e *VNCT* que implementam a solução para problema de tratamentos.

B.1.3 Algoritmos para Tratamentos

function *Tratamentos – Nebulosos*((*T*, *D*, *Z*, *V*, *Hipoteses*, *CA*))

variables

```

     $d_{nova}$ : disorder;
    Tratamentos: generator-set;
begin
    Tratamentos = { };
    while Hipoteses do
         $d_{nova}$ =Proxhip;
        trats = tratamentos( $d_{nova}$ );
         $VNCT(d_{nova}, trats, Z, V, CA)$ ;
        Tratamentos = revise(Tratamentos, (Tratamentos - TI));
    endwhile;
    while Tratamentos do
         $t_{max}$ =Proxtrat;
         $\gamma(t_{max})$ =Proxind;
         $Retorno(t_{max}) = RISCO(t_{max}) \ominus CUSTO(t_{max})$ ;
    endwhile;
    return (max(Retorno(Tratamentos),max(GCONT(Tratamentos)));
end.

```

```

function  $VNCT(d_{nova}, Trats, Z, V, CA)$ 
variables
    TI: set-of-treatments;
begin
    TI = { };
     $\gamma(TEMP) = h(INT(TEMP) \cap INT^+(TEMP))$ ;
     $\gamma(UMID) = h(INT(UMID) \cap INT^+(UMID))$ ;
     $\omega_{CF}(d_{nova}) = sup(\gamma(TEMP), \gamma(UMID))$ ;
    while Trats do
         $t_{nova}$ =Proxtrat;
         $v_{nova}$ =ProxV;
         $\forall v_{nova} \in V do$ 
             $\gamma(v_{nova}) = h(INT(v_{nova}) \cap INT^+(v_{nova}))$ ;
             $\gamma(t_{nova}) = inf_{v_{nova} \in V}$ ;
            if  $\gamma(t_{nova}) < 0$  then
                 $TI = TI \cup t_{nova}$ ;
            else
                 $Tratamentos = Tratamentos \cup \gamma(t_{nova})$ ;
            endwhile;
    endwhile;
end.

```

```
return Tratamentos, TI ;  
end.
```

Estes algoritmos estão sendo implementados em PROLOG devido a sua fácil sintaxe e aderência ao problema. No próximo capítulo tem um exemplo de como uma base de conhecimento bem como um caso particular são representados em Prolog. Também é apresentado o resultado de diagnóstico para o caso particular apresentado. É importante lembrar que este exemplo foi executado através dos programas *bipartido – nebuloso.pro*, *exames – nebuloso.pro* e *tratamentos – nebuloso.pro* que contemplam a implementação da TCN.

APÊNDICE C

ESPECIFICAÇÃO DA BASE DE CONHECIMENTO

Neste apêndice são apresentadas as informações utilizadas na implementação do problema de raciocínio clínico para desordens de milho do capítulo 8 retiradas de Fernandes e Oliveira (1997) e (Malcolm, 1980).

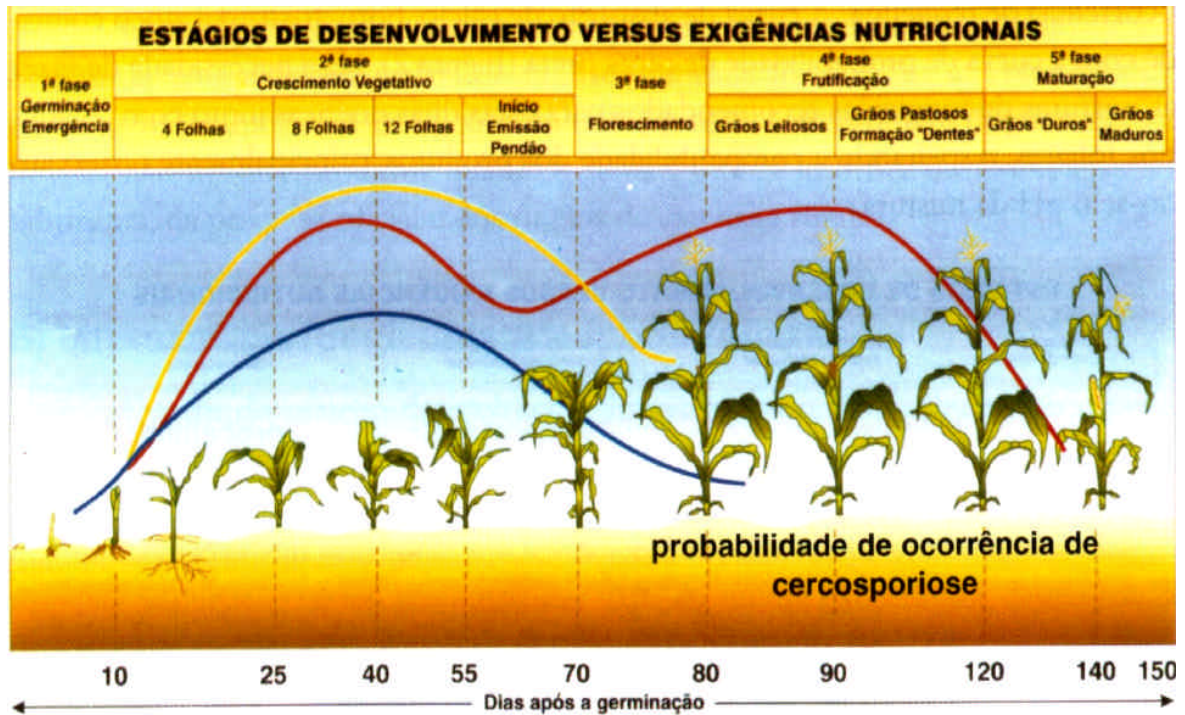


FIGURA C.1 – Escala Fenológica da Cultura do Milho.

TABELA C.1 – Etapas de desenvolvimento da cultura de milho.

Estágios	Descrição	Intervalos (em dias)	Intervalo Nebuloso (em semanas)
e_0	Germinação	0 a 10	<0,1,1,2>
e_1	Crescimento Vegetativo	10 a 55	<2,4,6,8>
e_2	Pendoamento	55 a 70	<8,9,9,10>
e_3	Florescimento	70 a 80	<9,10,11,12>
e_4	Frutificação	80 a 120	<12,13,16,17>
e_5	Maturação	120 a 150	<17,18,21,22>

TABELA C.2 – Níveis de severidade da cultura do milho.

Nível de severidade	Descrição do nível	Grau nebuloso
s_1	Sem lesões (antes do aparecimento dos sintomas)	-
s_2	Lesões esparsas na planta, restritas ao terço inferior	<1,2,3,4>
s_3	Até 50% das folhas com lesões; severas nos 25% das folhas inferiores	<3,4,5,6>
s_4	Até 75% das folhas com lesões; severas nos 50% das folhas inferiores	<5,6,7,8>
s_5	100% das folhas com lesões; lesões severas nos 75% das folhas inferiores	<7,8,9,10>
d_6	Planta morta	-

TABELA C.3 – Graus de temperatura e umidade.

Graus	Temperatura(T)	Umidade(U)
alta	$T > 25^{\circ}C$	$U > 85\%$
média	$15^{\circ}C < T < 25^{\circ}C$	$70\% < U < 80\%$
baixa	$T < 15^{\circ}C$	$U < 70\%$

TABELA C.4 – Tabela das principais desordens da cultura de milho.

Índice	Descrição
d_1	podridao_cartucho_por_erwinia
d_2	queima_por_pseudomonas
d_3	ferrugem_comum
d_4	ferrugem_polysora
d_5	ferrugem_branca_ou_tropical
d_6	mancha_por_physoderma
d_7	antracnose
d_8	mancha_por_helminthosporium_maydis(raca _O)
d_9	mancha_por_helminthosporium_maydis(raca _T)
d_{10}	mancha_por_cercospora
d_{11}	mancha_por_diplodia_macrospora
d_{12}	mancha_por_helminthosporium_turcicum
d_{13}	mancha_por_physoderma
d_{14}	mancha_por_Curvularia
d_{15}	mancha_por_helminthosporium_carbonum(raca ₁)
d_{16}	mancha_por_helminthosporium_carbonum(raca ₂)
d_{17}	mancha_por_helminthosporium_carbonum(raca ₃)
d_{18}	mancha_por_phaeosphaeria
d_{19}	queima_por_herbicida
d_{20}	mildio_do_sorgo
d_{21}	virose_rayado_fino
d_{22}	mosaico_comum_do_milho
d_{23}	efeito_de_frio
d_{24}	deficiencia_de_fosforo
d_{25}	enfezamento_vermelho)
d_{26}	enfezamento_palido
d_{27}	carvao_comum_da_espiga
d_{28}	causas_diversas(mildio_do_sorgo_ou_carvao_do_topo_ou_problemas_nutricionais)
d_{29}	podridao_rosada_da_espiga(fusarium)
d_{30}	podridao_da_espiga_por_gibberella)
d_{31}	podridao_branca_da_espiga(diplodia_maydis)
d_{32}	podridao_da_espiga(helminthosporium_maydis)
d_{33}	podridao_do_colmo_por_pythium
d_{34}	podridao_do_colmo_por_erwinia_carotovora
d_{35}	antracnose_do_colmo
d_{36}	podridao_do_colmo_por_rhizoctonia
d_{37}	podridao_do_colmo_por_fusarium(gibberella))
d_{38}	podridao_do_colmo_por_diplodia
d_{39}	podridao_do_colmo_por_macrophomina
d_{40}	carvao_do_pendao
d_{41}	mildio_do_sorgo_no_pendao

TABELA C.5 – Tabela das principais manifestações da cultura de milho.

Índice	Descrição	Índice	Descrição
<i>m</i> ₁	folhas_se_coalescem	<i>m</i> ₃₉	reducao_crescimento_plantas
<i>m</i> ₂	seca_em_grandes_areas_foliares	<i>m</i> ₄₀	manchas_amareladas
<i>m</i> ₃	podridoes_aquosas	<i>m</i> ₄₁	manchas_avermelhadas
<i>m</i> ₄	odor_desagradavel	<i>m</i> ₄₂	manchas_arroxeadas
<i>m</i> ₅	folhas_se_desprendem_facilmente	<i>m</i> ₄₃	nao_ha_formacao_graos
<i>m</i> ₆	pustulas_individualizadas_marrom_claro	<i>m</i> ₄₄	espigas_raquiticas
<i>m</i> ₇	pustulas_individualizadas_marrom_escuro	<i>m</i> ₄₅	graos_frouxos
<i>m</i> ₈	rompimento_da_epiderme	<i>m</i> ₄₆	encurtamento_internodios
<i>m</i> ₉	formacao_de_fendas	<i>m</i> ₄₇	proliferacao_espigas
<i>m</i> ₁₀	adquirem_aspecto_pulvulento	<i>m</i> ₄₈	presenca_galhas
<i>m</i> ₁₁	pustulas_individualizadas_branca	<i>m</i> ₄₉	massa_de esporos_negros
<i>m</i> ₁₂	pustulas_circundadas_com_halo_escuro_avermelhado	<i>m</i> ₅₀	ausencia_de_galhas
<i>m</i> ₁₃	seca_prematura_das_plantas	<i>m</i> ₅₁	graos_bem_formados
<i>m</i> ₁₄	reducao_da_producao	<i>m</i> ₅₂	recobertos_com_micelio_cotonoso_rosa
<i>m</i> ₁₅	lesoes_necroticas_amareladas	<i>m</i> ₅₃	presenca_micelio_branco_entre_graos
<i>m</i> ₁₆	lesoes_necroticas_marrom	<i>m</i> ₅₄	presenca_de_numerosos_pontinhos_negros
<i>m</i> ₁₇	lesoes_necroticas_marrom_avermelhada	<i>m</i> ₅₅	graos_mal_formados
<i>m</i> ₁₈	murchamento_folhas_apicais	<i>m</i> ₅₆	tombamento
<i>m</i> ₁₉	lesoes_necroticas_palha	<i>m</i> ₅₇	podridao_seca
<i>m</i> ₂₀	lesoes_necroticas_circundadas_com_halo_amarelado	<i>m</i> ₅₈	sintomas_internos
<i>m</i> ₂₁	lesoes_necroticas_circundadas_com_halo_avermelhado	<i>m</i> ₅₉	sintomas_externos
<i>m</i> ₂₂	lesoes_necroticas_escuras(pretas)	<i>m</i> ₆₀	presenca_de_folhas_rudimentares
<i>m</i> ₂₃	lesoes_necroticas_cinzas	<i>m</i> ₆₁	excesso_de_brotacoes
<i>m</i> ₂₄	ponto_de_infeccao_visivel	<i>m</i> ₆₂	coloracao_avermelhada
<i>m</i> ₂₅	lesoes_necroticas_circundadas_com_halo_dourado	<i>m</i> ₆₃	coloracao_marrom_claro
<i>m</i> ₂₆	presenca_micelio_negro_sobre_graos	<i>m</i> ₆₄	coloracao_marrom_escuro
<i>m</i> ₂₇	reducao_ciclo_da_planta	<i>m</i> ₆₅	sintomas_esparsos
<i>m</i> ₂₈	reducao_tamanho_dos_graos	<i>m</i> ₆₆	sintomas_iniciais_ponta_da_espiga
<i>m</i> ₂₉	lesoes_necroticas_branca	<i>m</i> ₆₇	sintomas_primeiro_entreno
<i>m</i> ₃₀	deformacao_no_pendao	<i>m</i> ₆₈	sintomas_varios_entreno
<i>m</i> ₃₁	nao_ha_formacao_do_pendao	<i>m</i> ₆₉	lesoes_alongadas
<i>m</i> ₃₂	folhas_pequenas	<i>m</i> ₇₀	lesoes_irregulares
<i>m</i> ₃₃	folhas_eretas_estreitas	<i>m</i> ₇₁	lesoes_arredondadas
<i>m</i> ₃₄	estrias_esbranquicadas	<i>m</i> ₇₂	lesoes_estreitas
<i>m</i> ₃₅	nao_ha_formacao_de_espigas	<i>m</i> ₇₃	lesoes_retangulares
<i>m</i> ₃₆	folhas_com_pontuacoes_cloroticas	<i>m</i> ₇₄	lesoes_centro_escuro
<i>m</i> ₃₇	manchas_verde_claro_escuro	<i>m</i> ₇₅	coloracao_cinza
<i>m</i> ₃₈	reducao_tamanho_espigas		

TABELA C.6 – Principais testes laboratoriais.

Testes	Doenças
Doenças bacterianas	d_1, d_2, d_{34}
Testes bioquímicos	
Testes sorológico	
Análise de perfil de ácidos graxos	
Análise de ribossomo	
Doenças Fúngicas	$d_3, d_4, d_5,$
Análise macroscópica (sintomas)	$d_6, d_7, d_8, d_9, d_{10},$
Análise microscópica (microscopia ótica)	$d_{11}, d_{12}, d_{13},$
isolamento em cultura pura	$d_{14}, d_{15}, d_{16}, d_{17},$
identificação sorológica	$d_{18}, d_{20}, d_{27}, d_{29}$
Análise de DNA	$d_{30}, d_{31}, d_{32},$
	$d_{33}, d_{35}, d_{36}, d_{37},$
	$d_{38}, d_{39}, d_{40}, d_{41}$
	$d_{21}, d_{22}, d_{25}, d_{26}$
Viroses e enfezamentos	
Testes sorológicos, especialmente ELISA	
Análise de DNA/RNA	
Microscopia eletrônica	
pode ser usado para complementar, porém não é usado como rotina de diagnose	

TABELA C.7 – Fungicidas para milho fornecidos pelo Ministério da Agricultura.

Nome comercial	Ingrediente ativo	Registrante
Brometila	Methyl bromide	BROMISA
Bromex	Methyl bromide	CASA BERNARDO
Bromo Fersol	Methyl bromide	FERSOL
Bromo Flora	Methyl bromide	FERSOL
Captan 200 Fungicida Agroceres	captan	AGRICUR
Captan 750 TS	captan	HOKKO
Comet	pyraclotrobin	BASF
Constant	tebuconazole	BAYER
Elite	tebuconazole	BAYER
Euparen M 500 PM	tolyfluanid	BAYER
Folicur 200 CE	tebuconazole	BAYER
Kumulus DF	enxofre	BASF
Maxim	fluxioxonil	SYNGENTA
Maxim XL	fluxioxonil	SYNGENTA
Mayran	thiram	ENRO
Opera	epoxiconazole	BASF
Tecto 100	thiabendazole	SYNGENTA
Tecto 600	thiabendazole	SYNGENTA
Thiram 480 TS	Thiram	CROMPTON
Tilt	Propiconazole	SYNGENTA
Triade	tebuconazole	BAYER
Vitavax-Thiram 200 SC	Carboxim+thiram	CROMPTOM
Vitavax-Thiram PM Uniroyal	Carboxim+thiram	CROMPTOM

TABELA C.8 – Fungicidas sistêmicos X desordens.

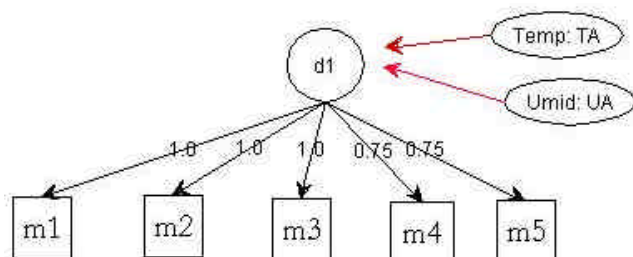
Referência	Tratamentos	Desordens
t_1	Priori	d_{10}, d_{11}, d_{12}
t_2	Opera	d_{10}, d_{11}, d_{12}
t_3	Palisade	d_{10}, d_{11}, d_{12}
t_4	Juno (propicanazole)	d_{10}, d_{11}, d_{12}
t_5	Folicur	d_{10}, d_{11}, d_{12}
t_6	Folicur 200CE (tebuconazole)	d_3, d_4, d_{12}
t_7	Tilt (propicanazole)	d_5, d_{12}
t_8	Maxim (fluxiaxomil)	d_{29}, d_{37}
t_9	Tecto 100 (thiabendazole)	d_{29}, d_{37}
t_{10}	Tecto 600 (thiabendazole)	d_{29}, d_{37}
t_{11}	fungicida com mancozeb	d_{18}
t_{12}	fungicida com metalaxil	d_{29}, d_{41}
t_{13}	fungicida com marbofuran	d_{25}, d_{26}
t_{14}	fungicida com imidacloprid	d_{25}, d_{26}
t_{15}	não aplicar fungicida)	todas

Grafos associados a cada desordem d_i (Associação causal, Temporal, Variáveis de Decisão)

TABELA C.9 – Consequências da árvore de decisão de cada desordem d_i .

Consequência	Desordem	Produtividade
c_1	controle alto	normal
c_2	evolução rápida	redução baixa
c_3	evolução lenta	normal
c_4	controle médio-alto	redução baixa
c_5	evolução lenta-intermediária	redução baixa
c_6	controle baixo	redução média-alta
c_7	evolução rápida	redução alta
c_8	evolução lenta-intermediária	redução média-alta
c_9	controle nulo	redução alta
c_{10}	evolução lenta-intermediária	redução alta

Modelo da desordem D1



Grafo temporal de D1 - G(V,T)

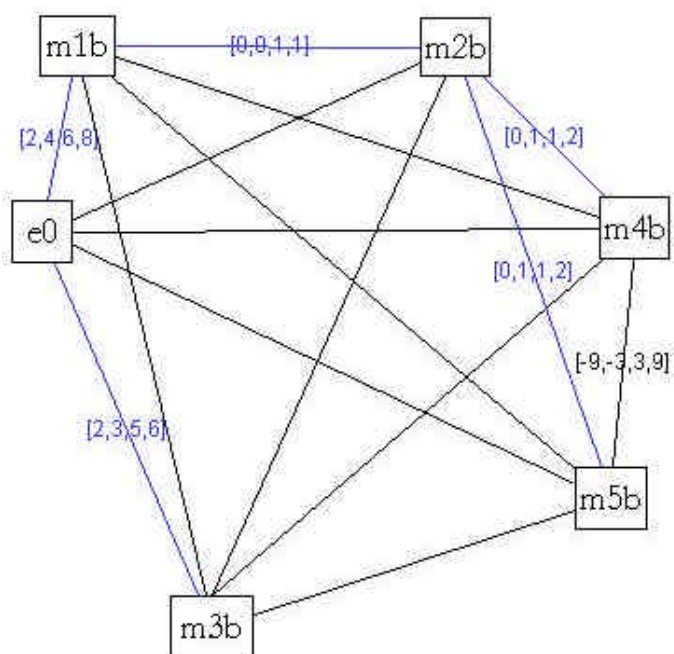
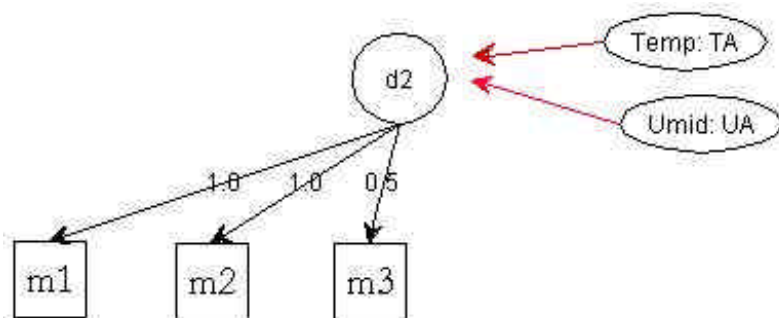


FIGURA C.2 – PODRIDÃO DO CARTUCHO POR *Erwinia* (d_1).

Modelo da desordem D2



Grafo temporal de D2 - G(V,T)

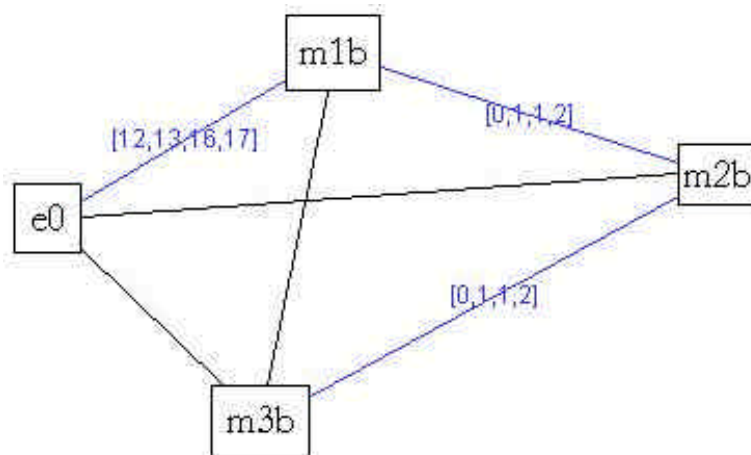
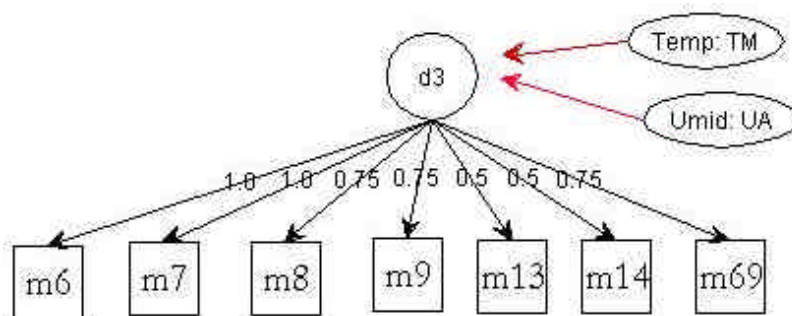


FIGURA C.3 – QUEIMA POR *Pseudomonas*(d_2).

Modelo da desordem D3



Grafo temporal de D3 - $G(V,T)$

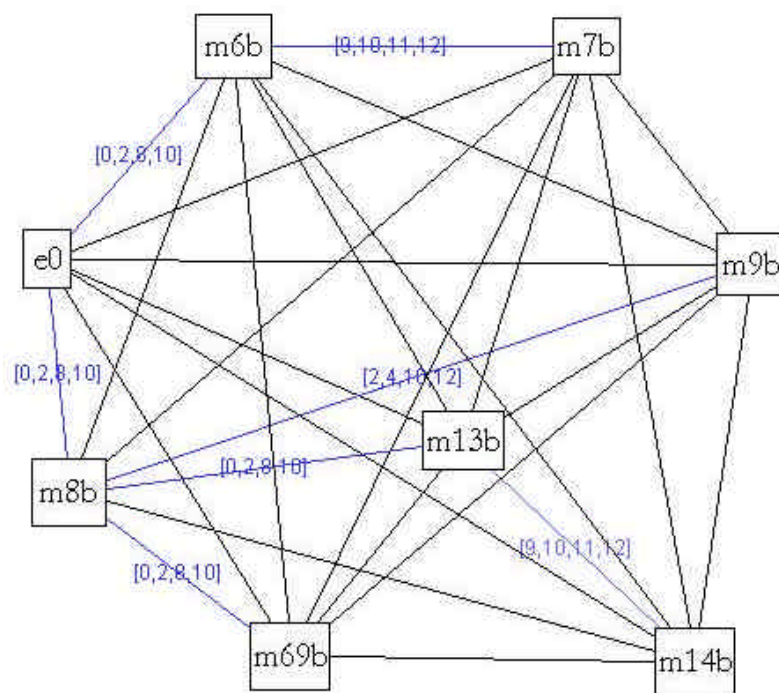


FIGURA C.4 – FERRUGEM COMUM (d_3).

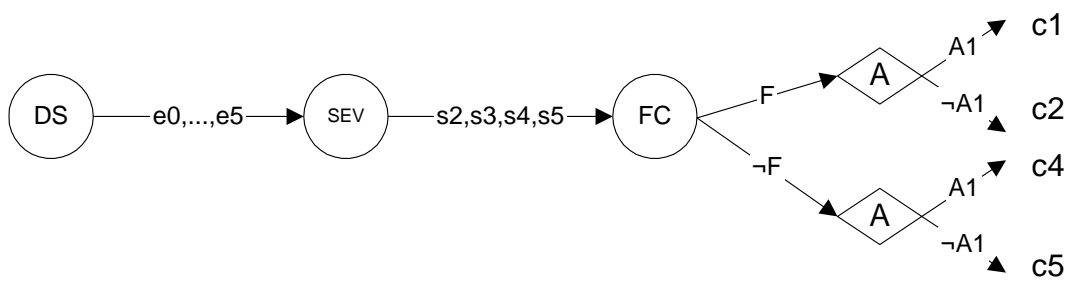
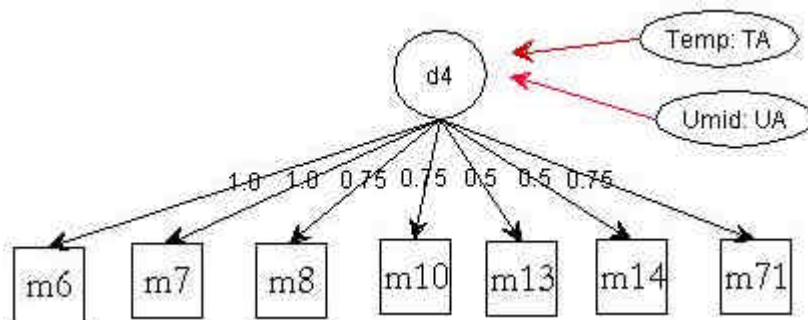


FIGURA C.5 – FERRUGEM COMUM - Variáveis de Decisão (d_3).

Modelo da desordem D4



Grafo temporal de D4 - G(V,T)

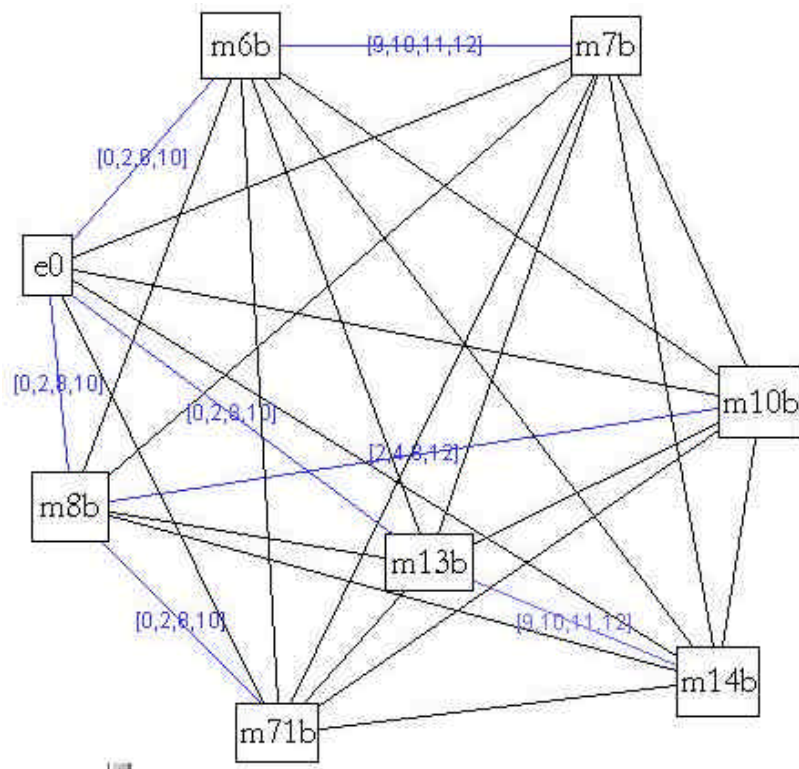


FIGURA C.6 – FERRUGEM *polysora* (d_4).

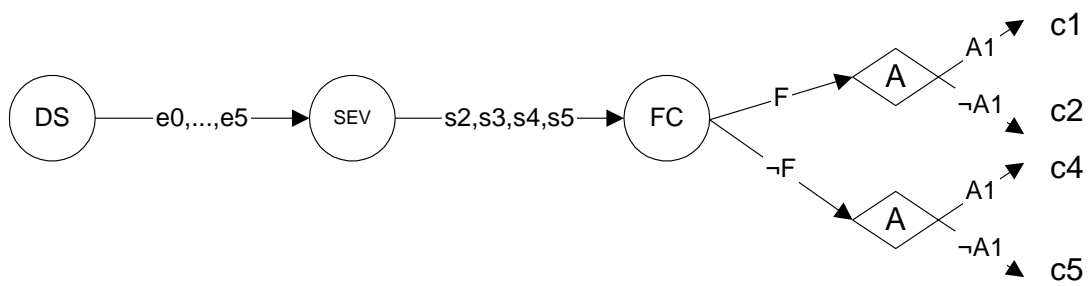
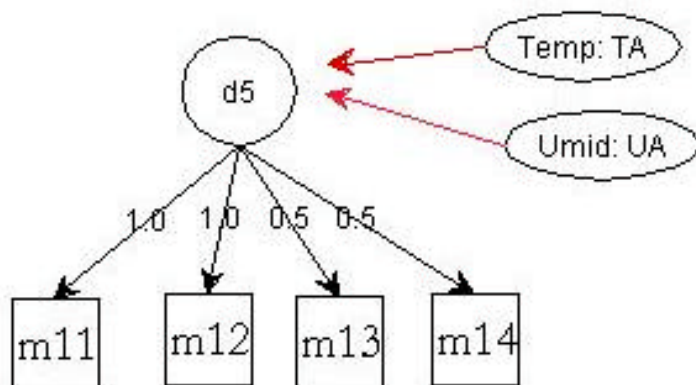


FIGURA C.7 – FERRUGEM *polysora* - Variáveis de Decisão (d_4).

Modelo da desordem D5



Grafo temporal de D5 - $G(V,T)$

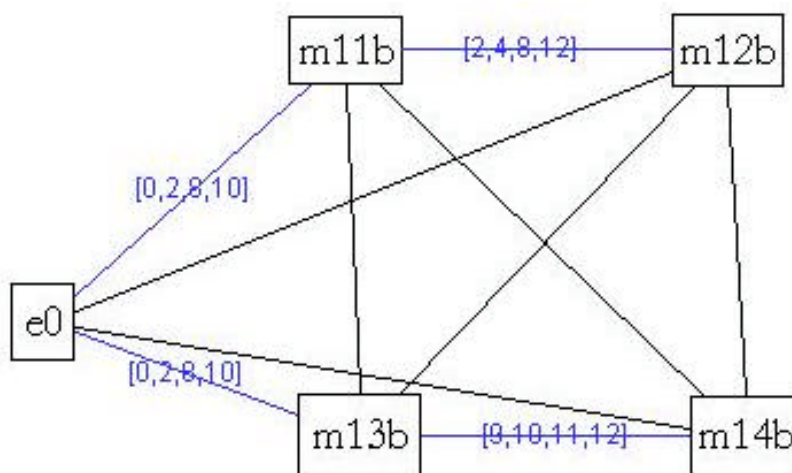


FIGURA C.8 – FERRUGEM BRANCA OU TROPICAL (d_5).

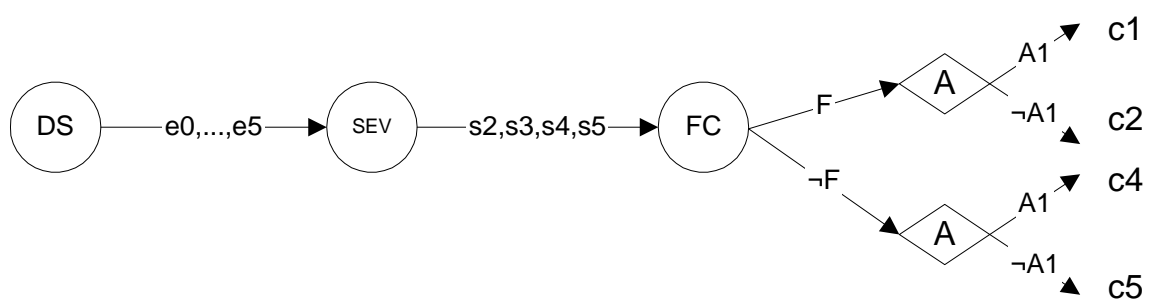
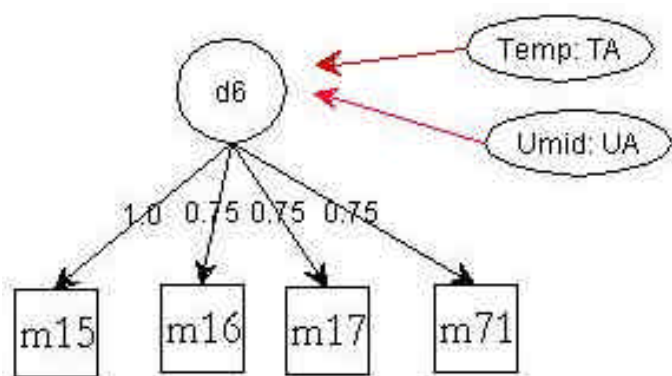


FIGURA C.9 – FERRUGEM BRANCA OU TROPICAL - Variáveis de Decisão (d_5).

Modelo da desordem D6



Grafo temporal de D6 - G(V,T)

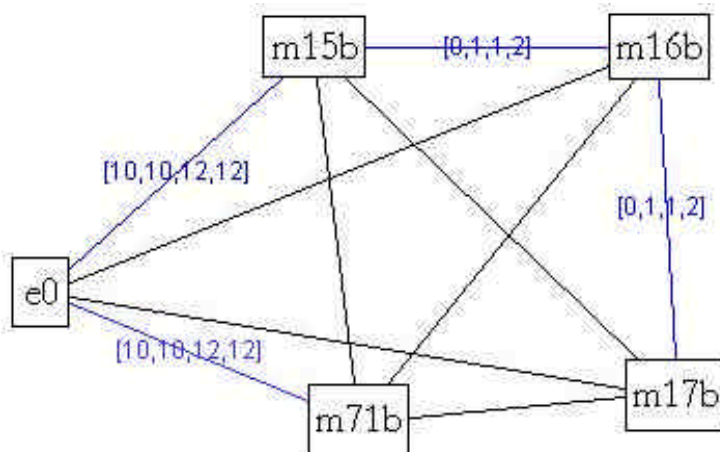
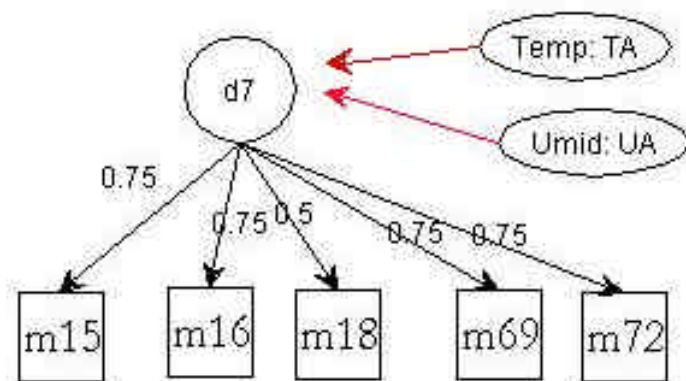


FIGURA C.10 – MANCHA POR *Physoderma* (d_6).

Modelo da desordem D7



Grafo temporal de D7 - $G(V,T)$

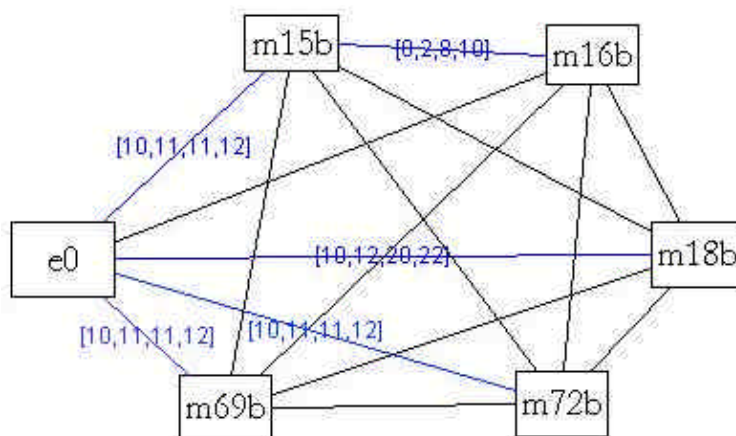
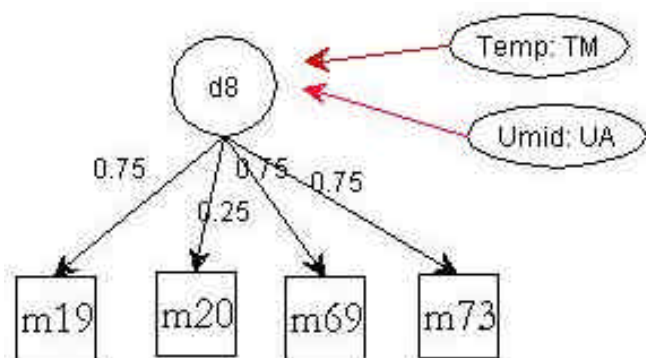


FIGURA C.11 – ANTRACNOSE (d_7).

Modelo da desordem D8



Grafo temporal de D8 - G(V,T)

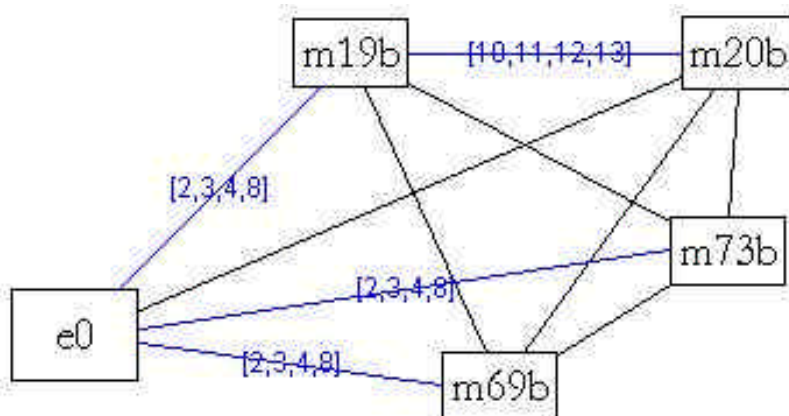
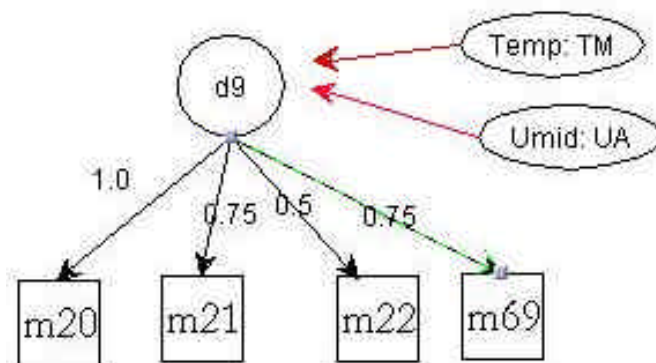


FIGURA C.12 – MANCHA POR *Helminthosporium maydis* Raça O (d_8).

Modelo da desordem D9



Grafo temporal de D9- G(V,T)

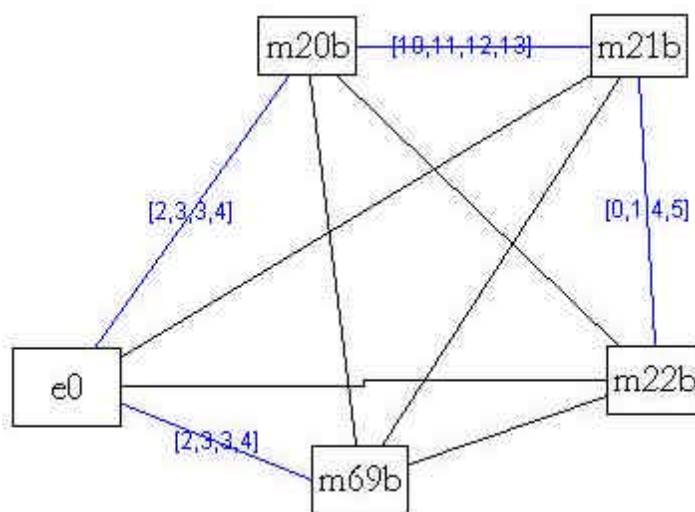
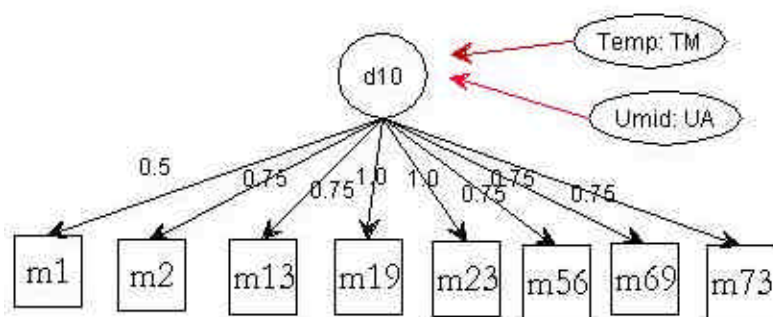


FIGURA C.13 – MANCHA POR *Helminthosporium maydis* Raça T (d_9).

Modelo da desordem D10



Grafo temporal de D10 - G(V,T)

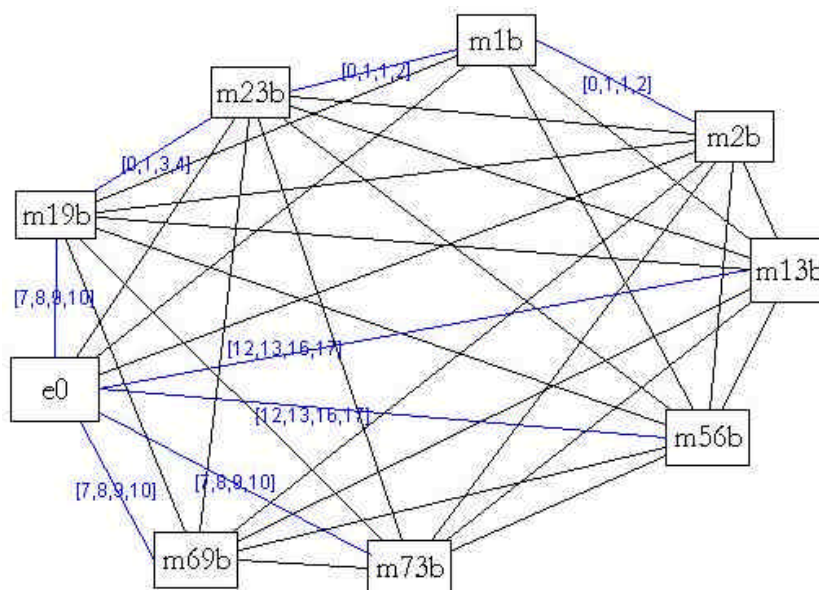


FIGURA C.14 – MANCHA POR *Cercospora* (d_{10}).

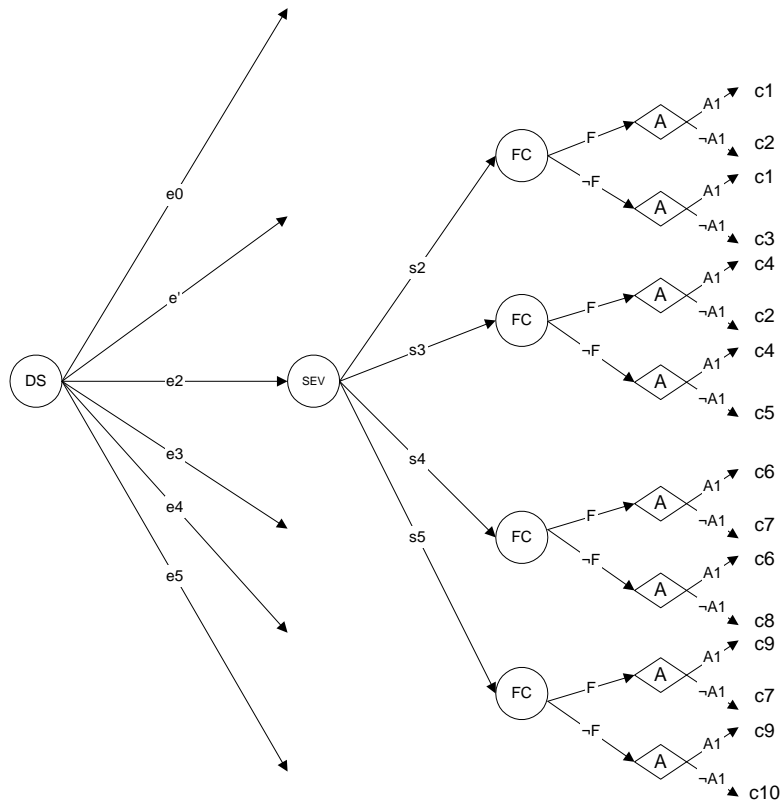
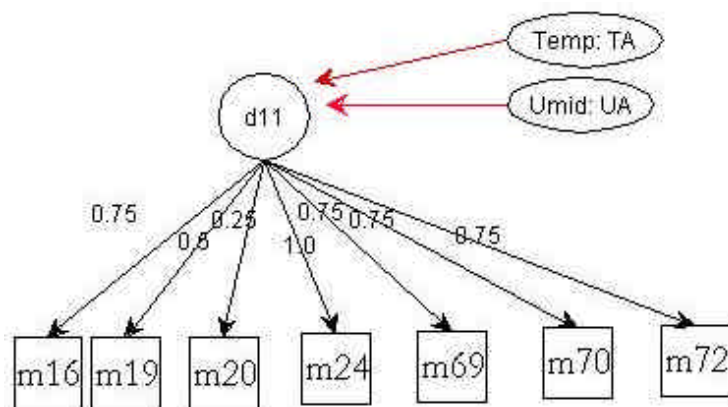


FIGURA C.15 – MANCHA POR Cercospora - Variáveis de Decisão (d_{10}).

Modelo da desordem 11



Grafo temporal de D11 - $G(V,T)$

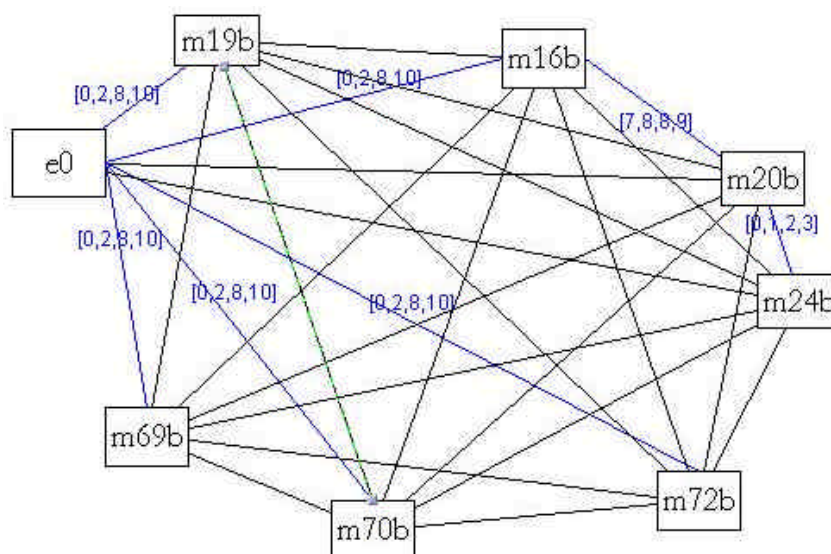


FIGURA C.16 – MANCHA POR *Diplodia Macrospora* (d_{11}).

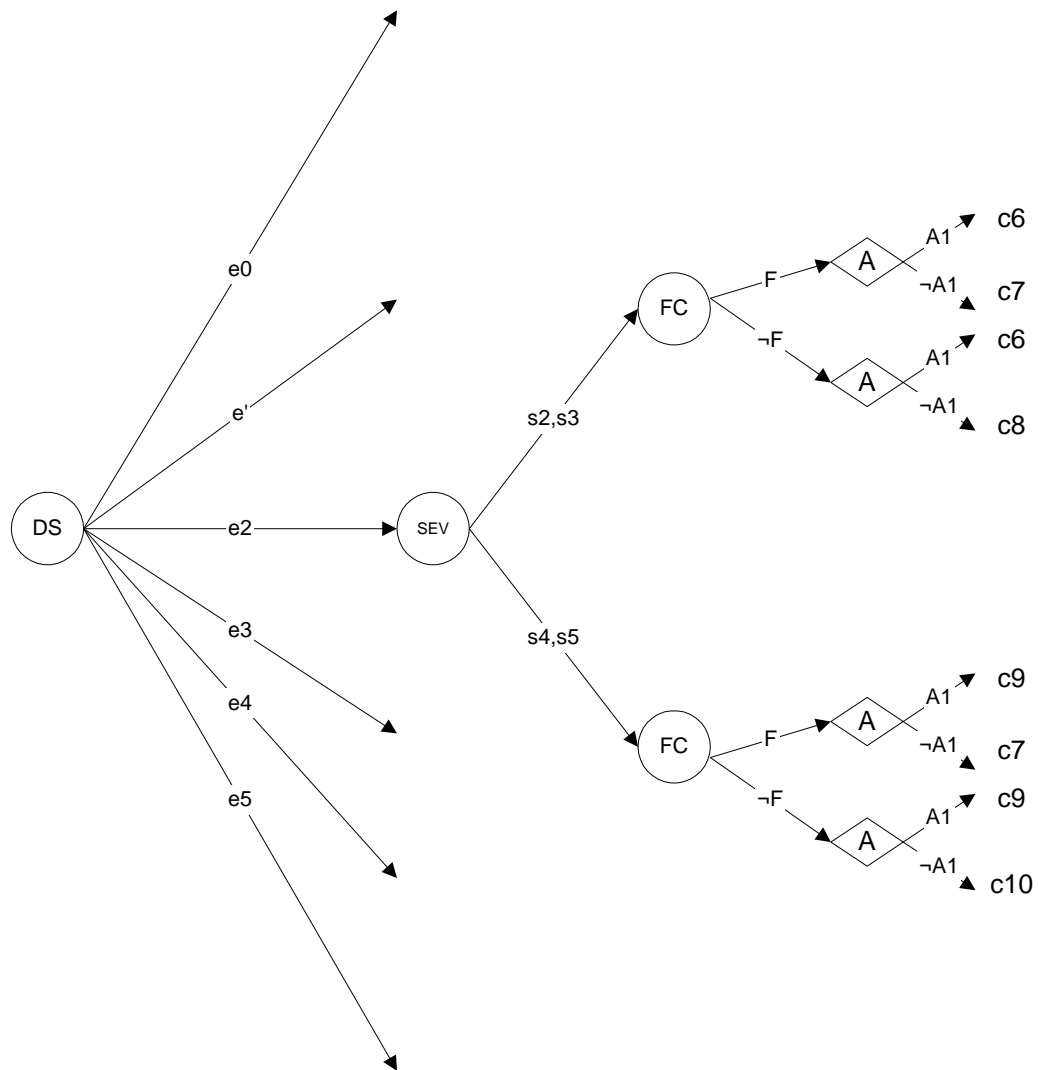
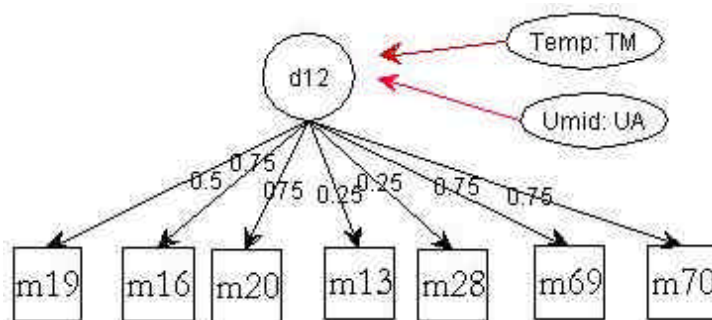


FIGURA C.17 – MANCHA POR *Diplodia Macrospora* - Variáveis de Decisão (d_{11}).

Modelo da desordem D12



Grafo temporal de D12 - G(V,T)

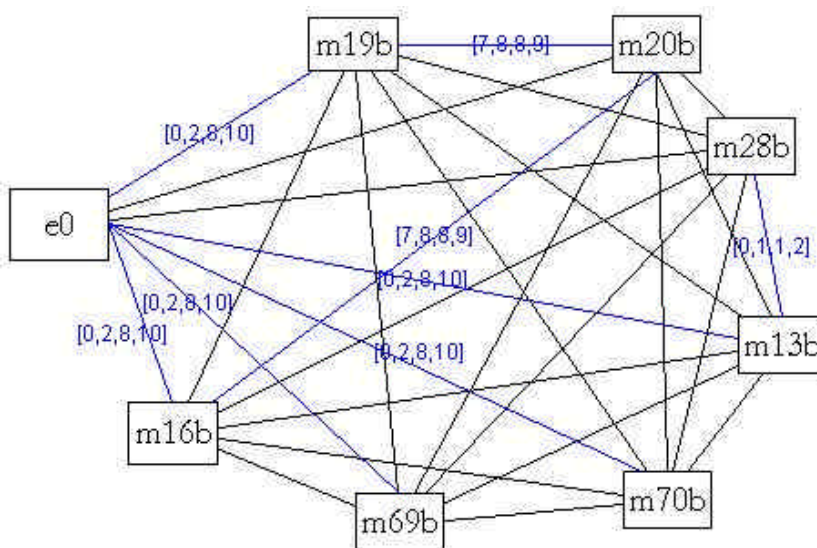


FIGURA C.18 – MANCHA POR *Helminthosporium Turcicum* (d_{12}).

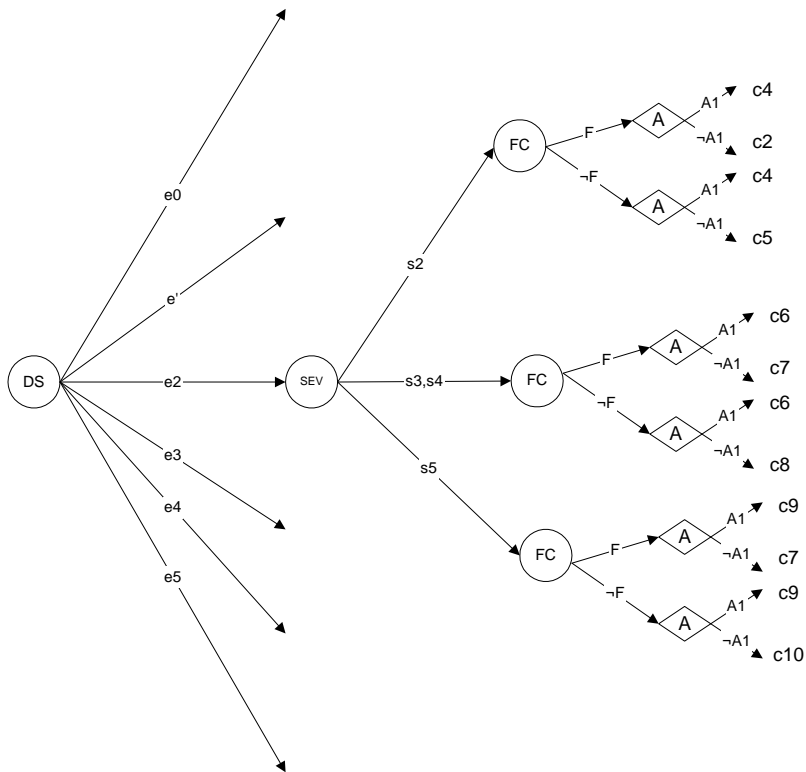
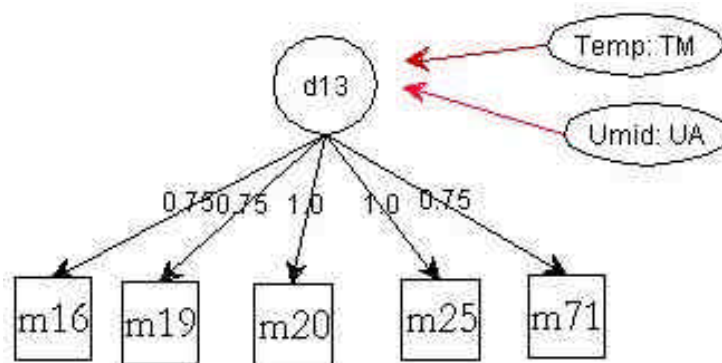


FIGURA C.19 – MANCHA POR *Helminthosporium Turcicum* - Variáveis de Decisão(d_{12}).

Modelo da desordem 13



Grafo temporal de D13 - $G(V,T)$

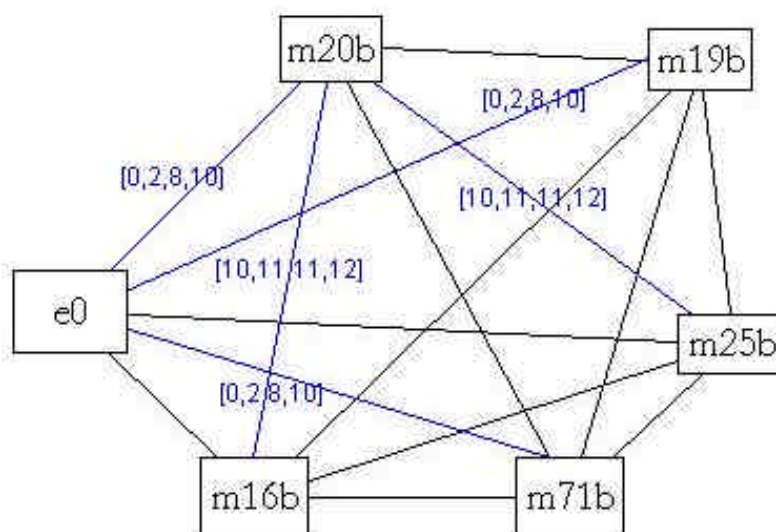
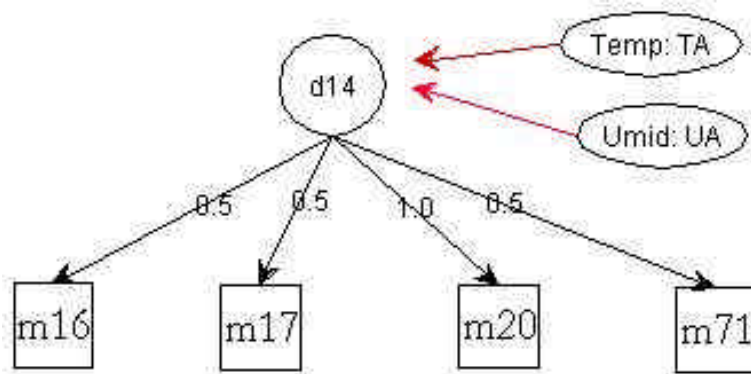


FIGURA C.20 – MANCHA POR Physoderma (d_{13}).

Modelo da desordem 14



Grafo temporal de D14 - $G(V,T)$

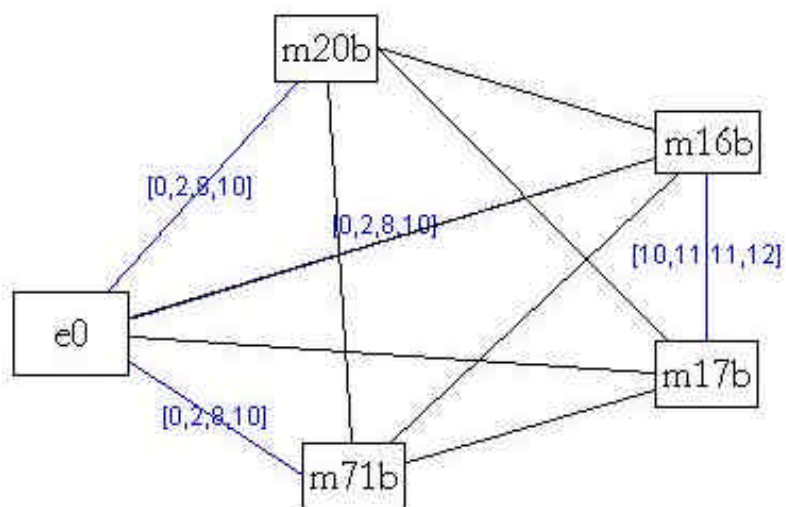
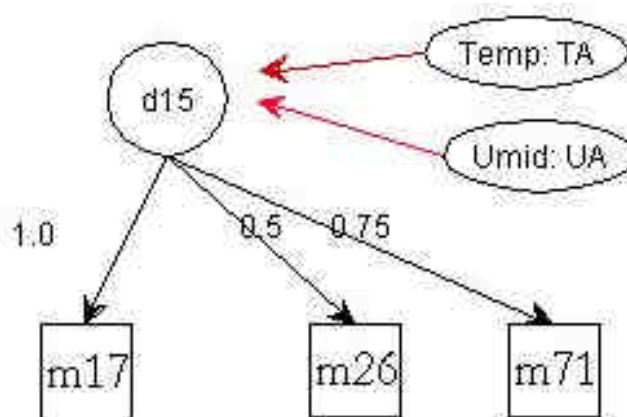


FIGURA C.21 – MANCHA POR $Curvularia(d_{14})$.

Modelo da desordem 15



Grafo temporal de D15 - G(V,T)

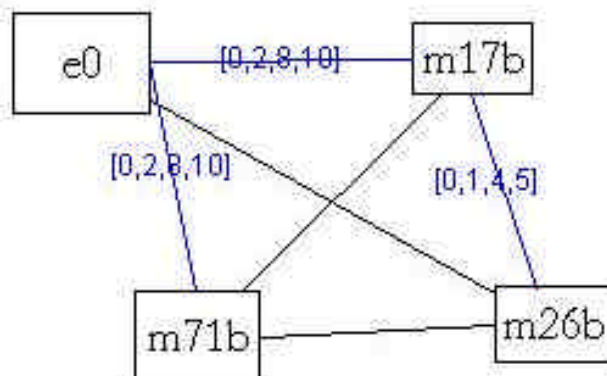
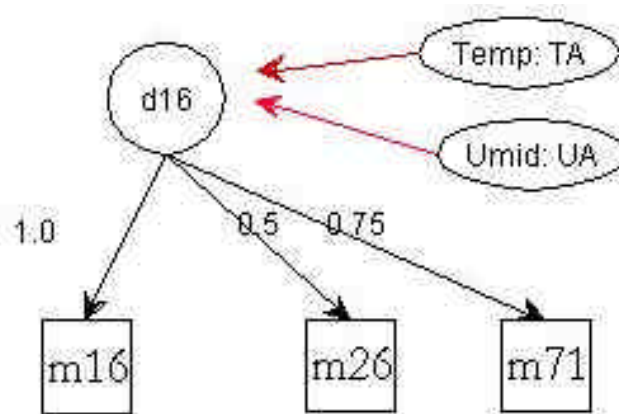


FIGURA C.22 – MANCHA POR *Helminthosporium Carbonum*(raça 1) (d_{15}).

Modelo da desordem 16



Grafo temporal de D16 - G(V,T)

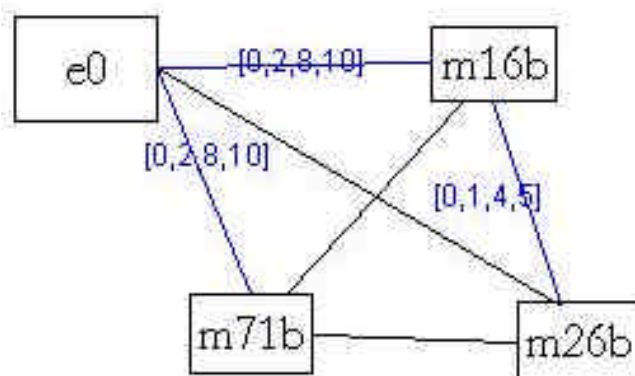
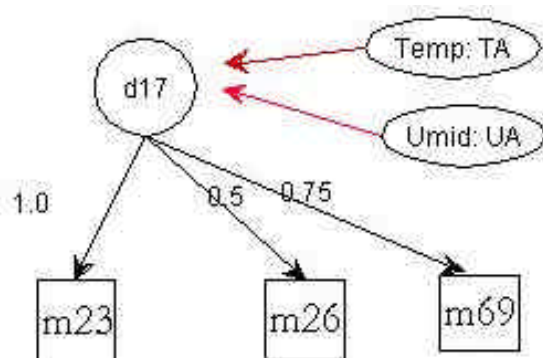


FIGURA C.23 – MANCHA POR *Helminthosporium Carbonum*(raça 2) (d_{16}).

Modelo da desordem 17



Grafo temporal de D17 - G(V,T)

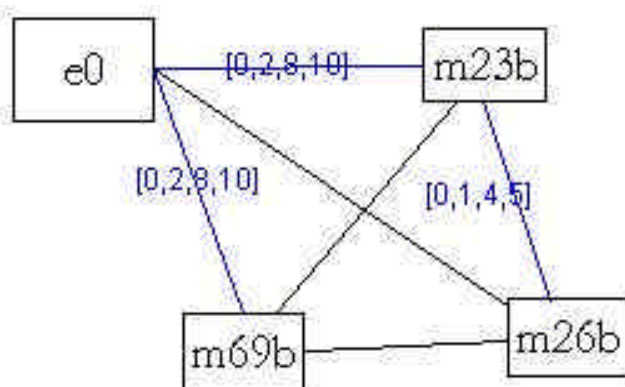
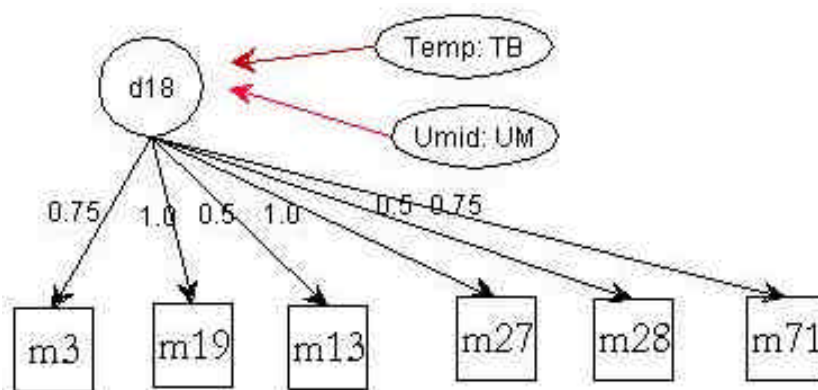


FIGURA C.24 – MANCHA POR *Helminthosporium Carbonum*(raça 3) (d_{17}).

Modelo da desordem 18



Grafo temporal de D18 - G(V,T)

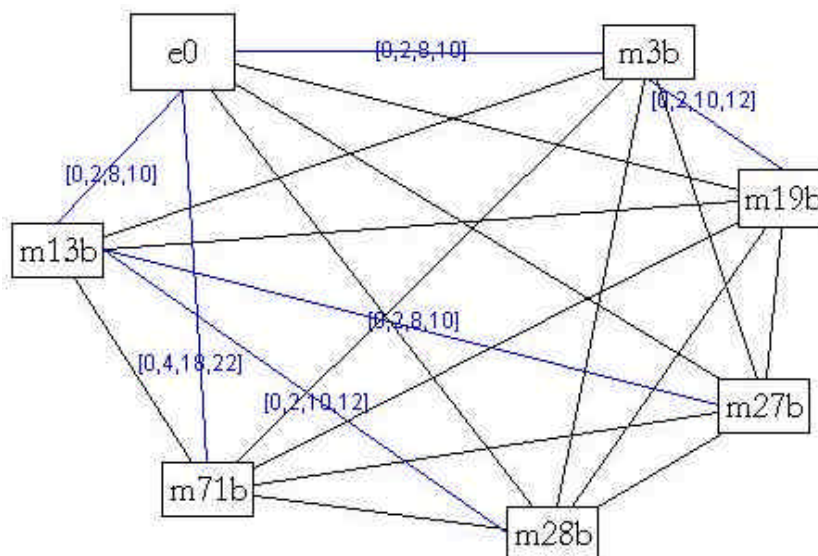


FIGURA C.25 – MANCHA POR *Phaeosphaeria* (d_{18}).

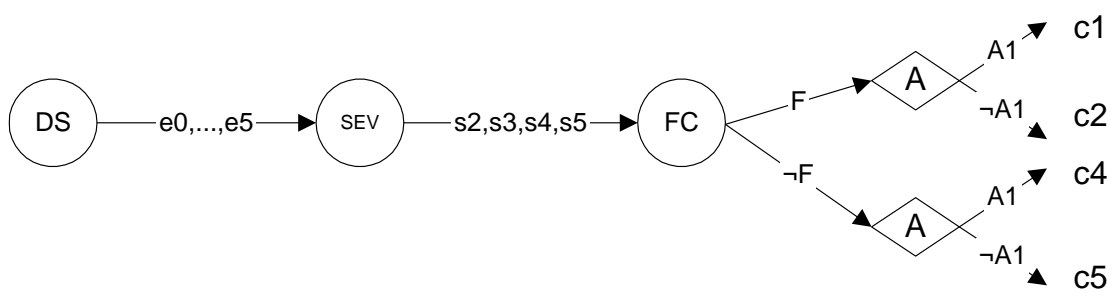
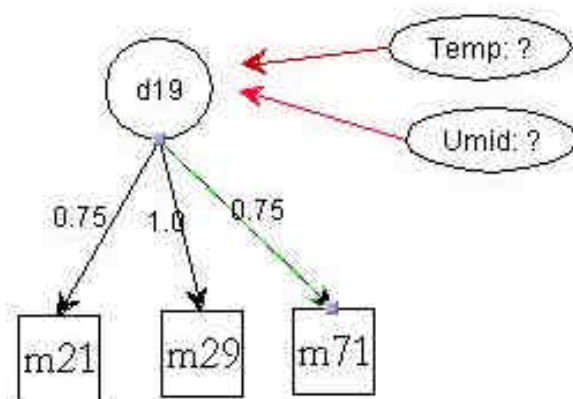


FIGURA C.26 – MANCHA POR *Phaeosphaeria* (d_{18}) - Variáveis de Decisão.

Modelo da desordem 19



Grafo temporal de D19 - G(V,T)

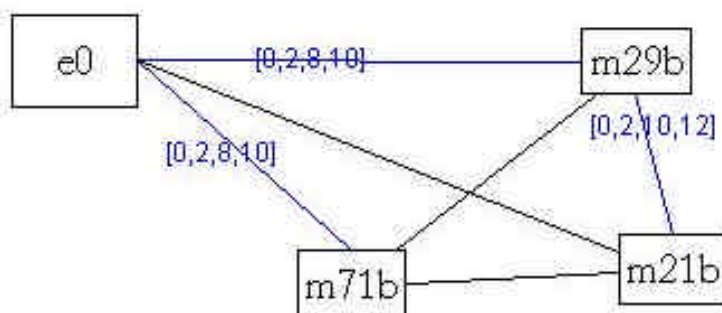
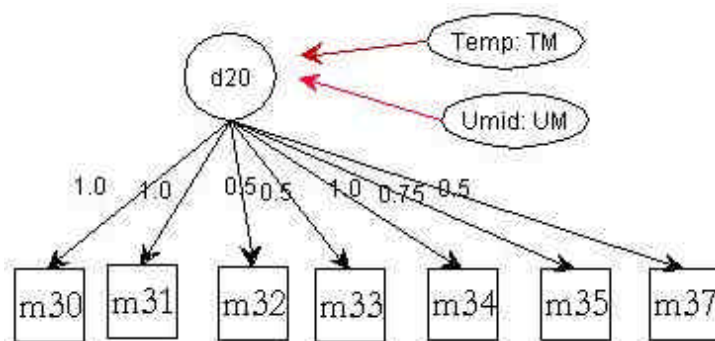


FIGURA C.27 – Queima por Herbicida (d_{19}).

Modelo da desordem 20



Grafo temporal de D20 - G(V,T)

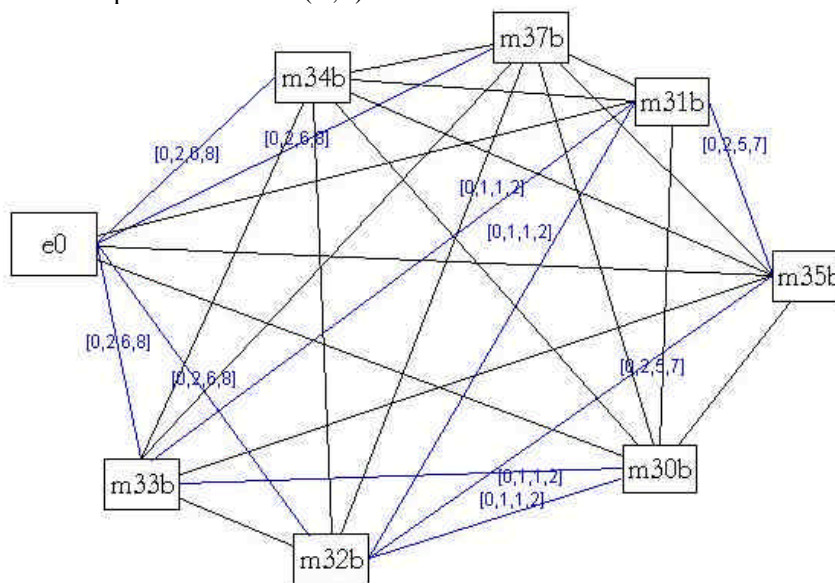
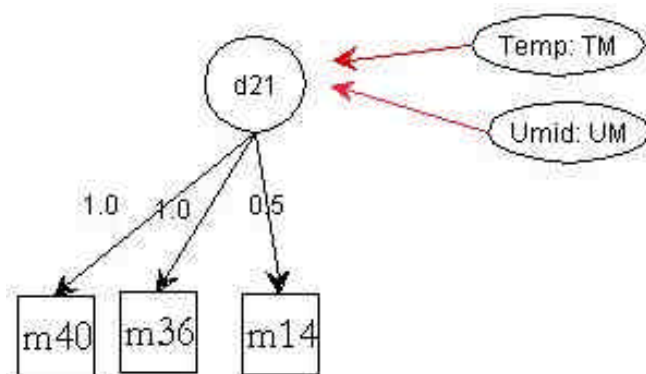


FIGURA C.28 – Míldio do Sorgo (d_{20}).

Modelo da desordem 21



Grafo temporal de D21 - G(V,T)

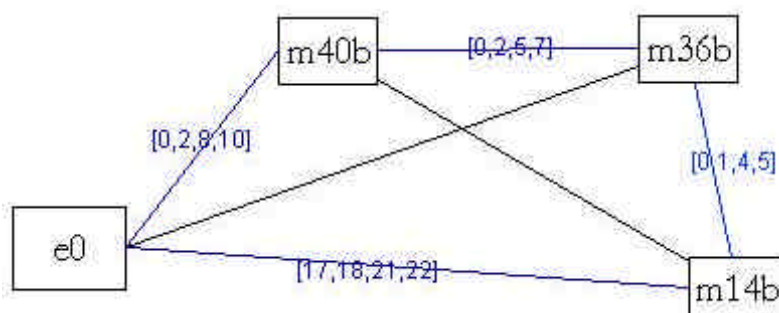
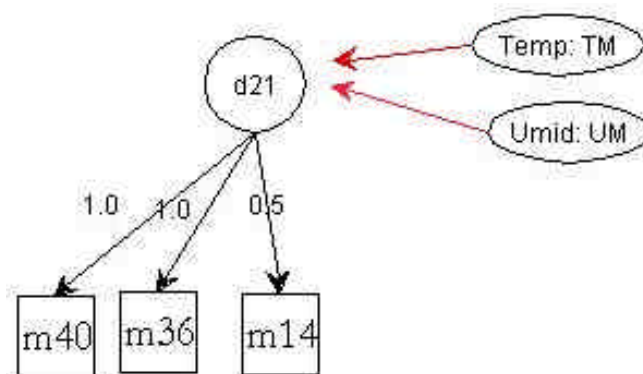


FIGURA C.29 – Virose Rayado Fino(d_{21}).

Modelo da desordem 21



Grafo temporal de D21 - G(V,T)

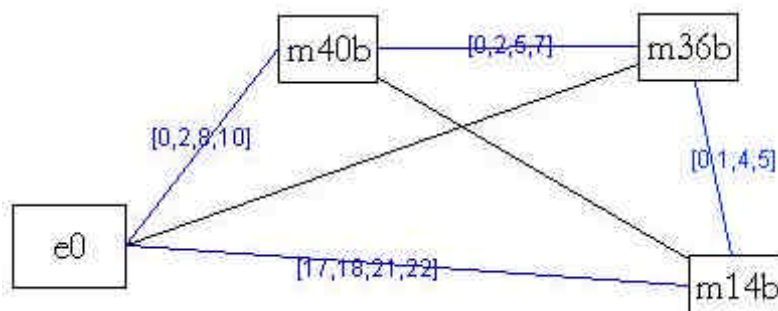
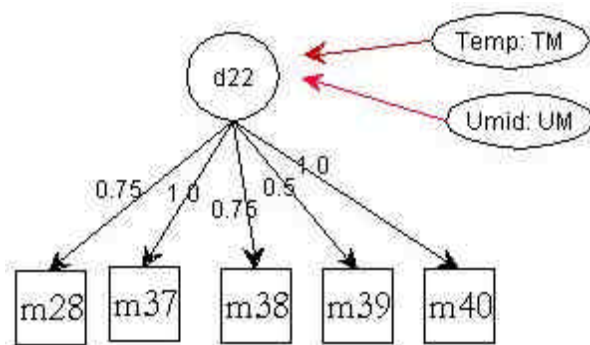


FIGURA C.30 – **Virose Rayado Fino**(d_{21}).

Modelo da desordem 22



Grafo temporal de D22 - $G(V,T)$

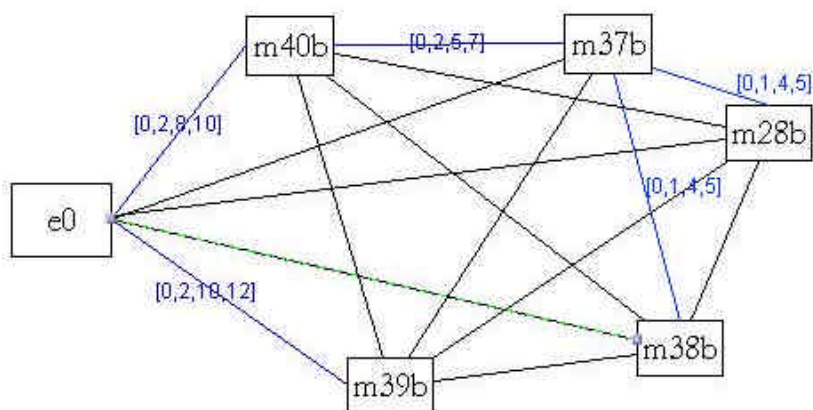
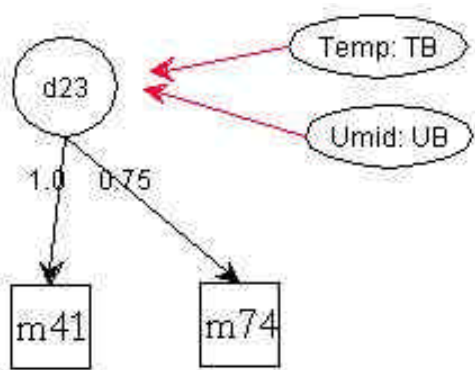


FIGURA C.31 – Mosaico Comum do Milho(d_{22}).

Modelo da desordem D23



Grafo temporal de D23 - $G(V,T)$

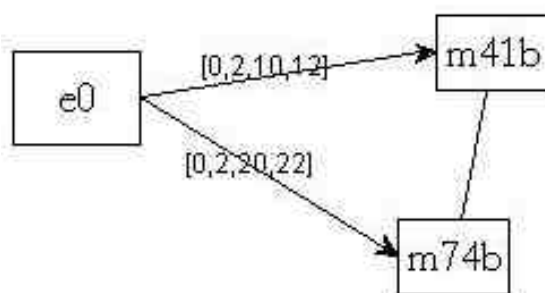
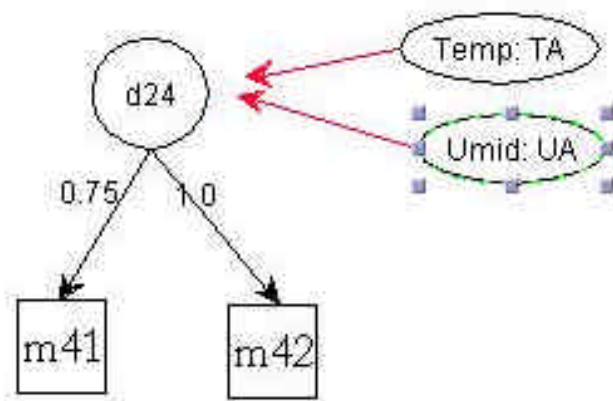


FIGURA C.32 – Efeito de Frio(d_{23}).

Modelo da desordem D24



Grafo temporal de D24 - G(V,T)

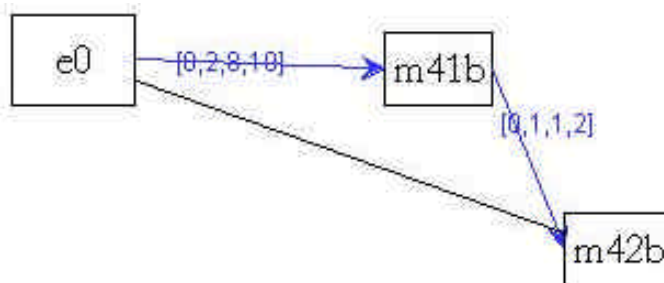
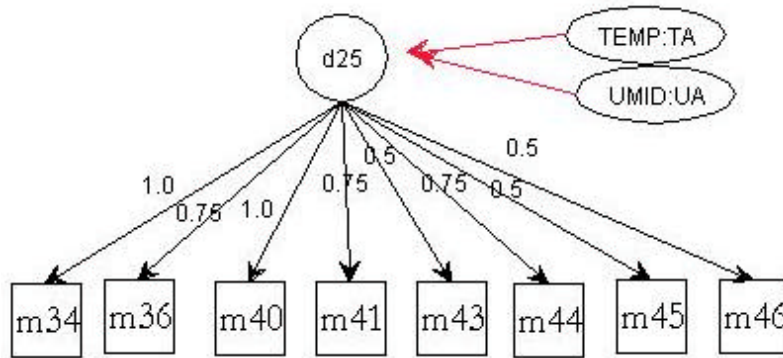


FIGURA C.33 – Deficiência de Fósforo (d_{24}).

Modelo da desordem D25



Grafo temporal de D25 - G(V,T)

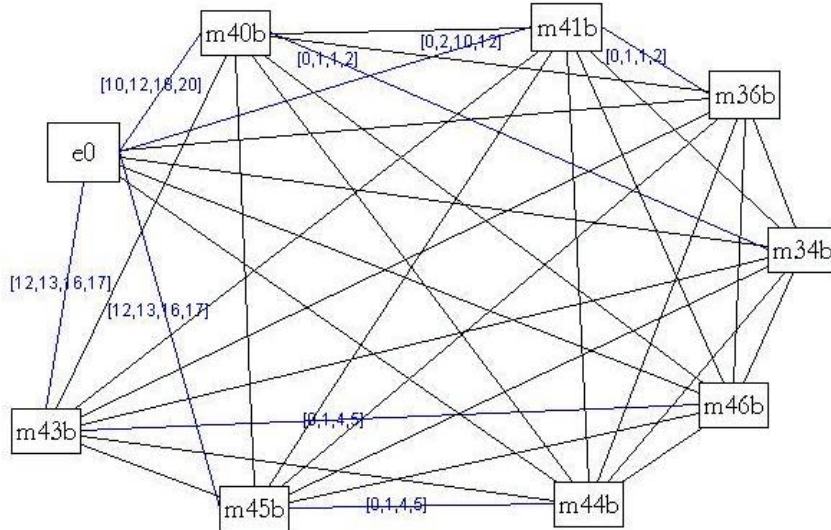


FIGURA C.34 – Enfezamento Vermelho(d_{25}).

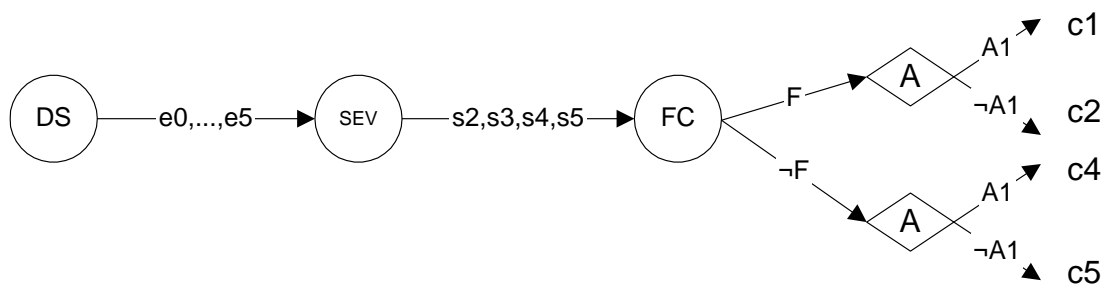
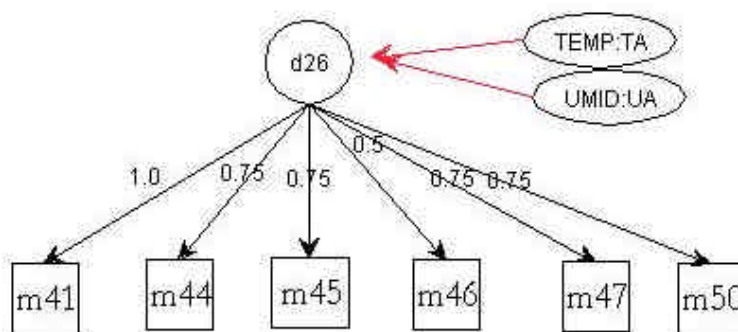


FIGURA C.35 – Enfezamento Vermelho(d_{25}) - Variáveis de Decisão.

Modelo da desordem D26



Grafo temporal de D26 - $G(V,T)$

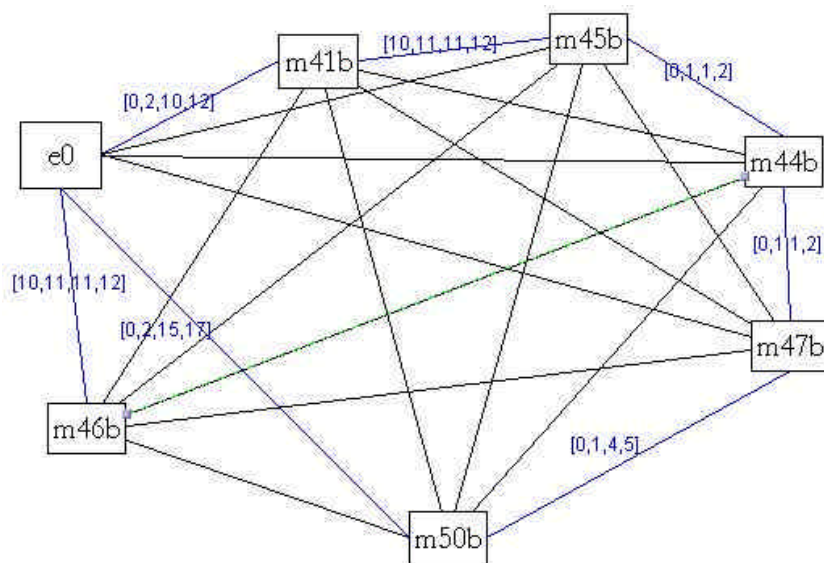


FIGURA C.36 – Enfezamento Pálido(d_{26}).

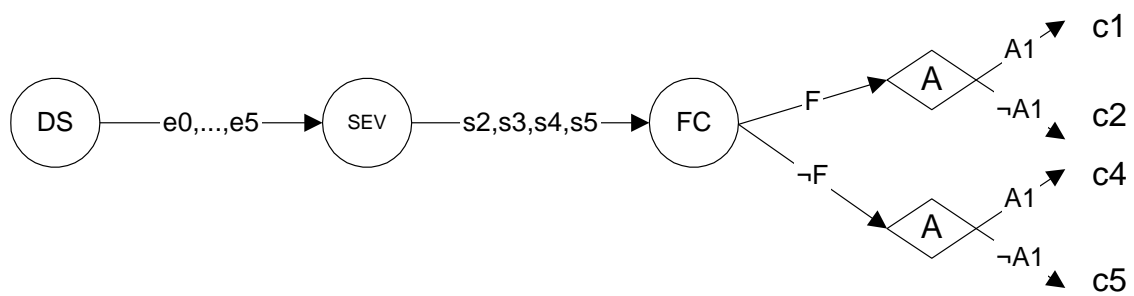
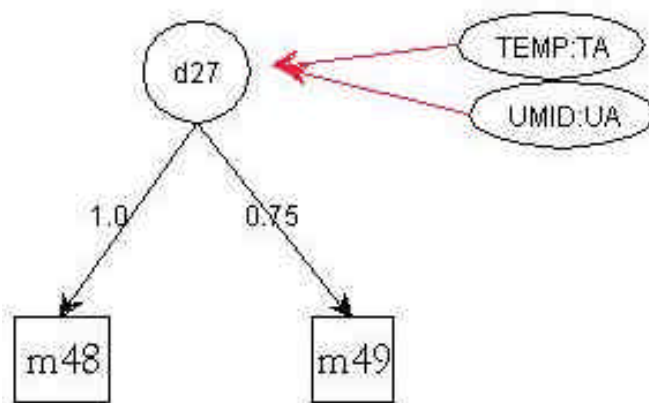


FIGURA C.37 – Enfezamento Pálido(d_{26}) - Variáveis de Decisão.

Modelo da desordem D27



Grafo temporal de D27 - G(V,T)

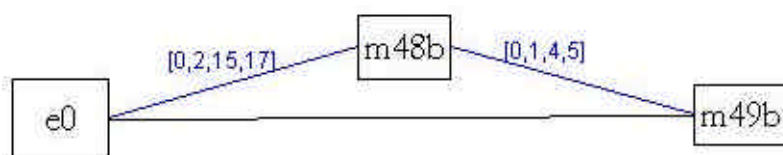
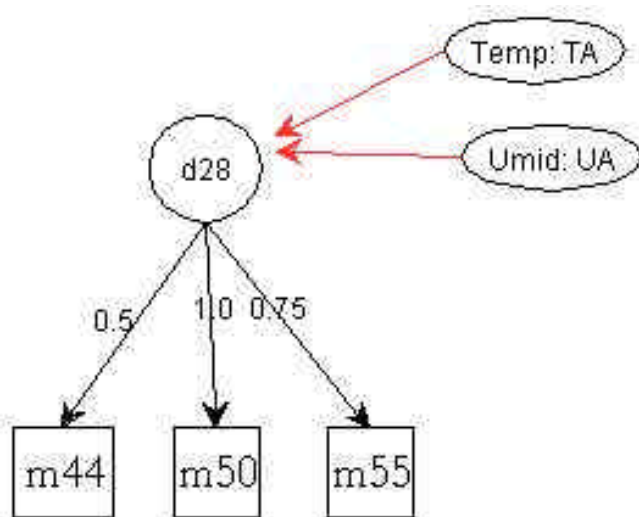


FIGURA C.38 – Carvão Comum da Espiga(d_{27}).

Modelo da desordem D28



Grafo temporal de D28 - G(V,T)

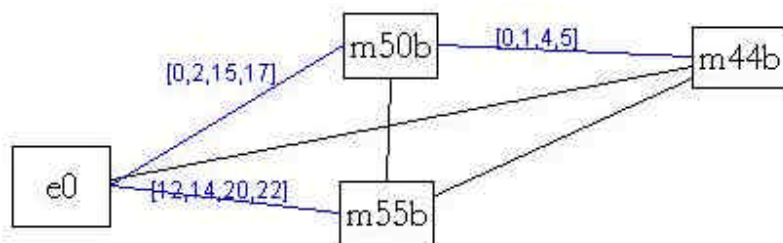
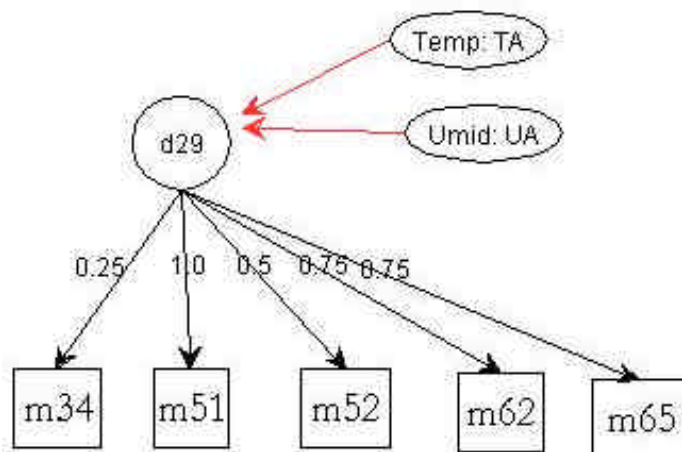


FIGURA C.39 – Causas Diversas(d_{28}).

Modelo da desordem D29



Grafo temporal de D29 - G(V,T)

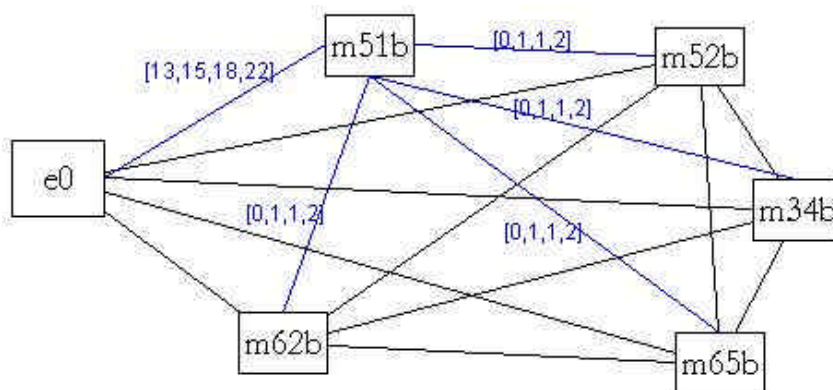


FIGURA C.40 – Podridão Rosada da Espiga (*Fusarium*) (d_{29}).

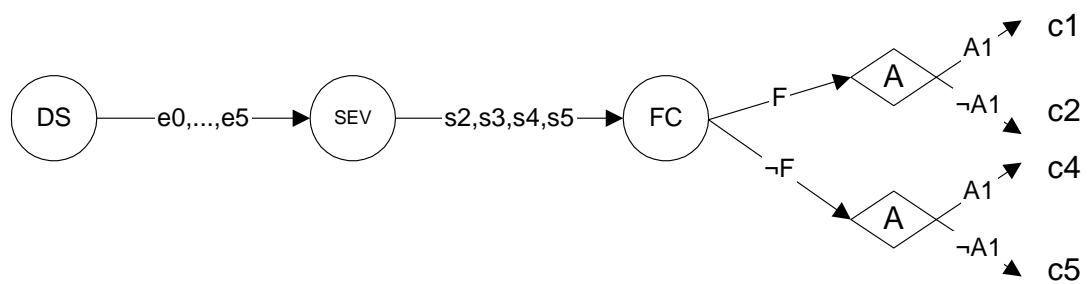
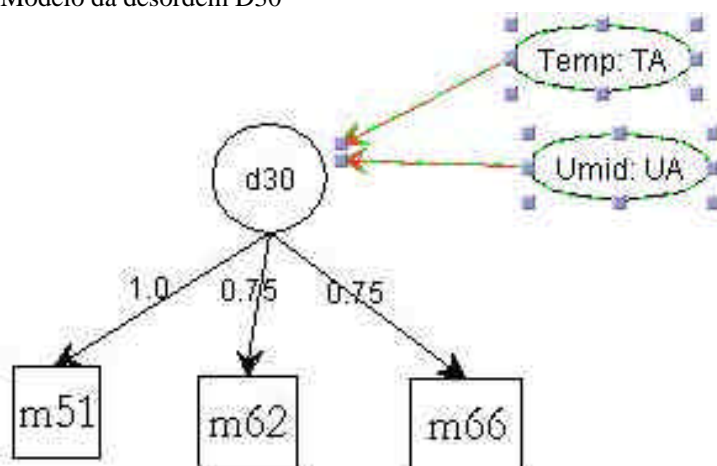


FIGURA C.41 – Podridão Rosada da Espiga (Fusarium) (d_{29}) - Variáveis de Decisão.

Modelo da desordem D30



Grafo temporal de D30 - G(V,T)

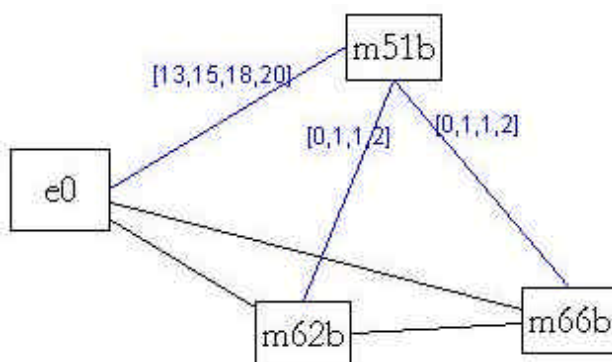
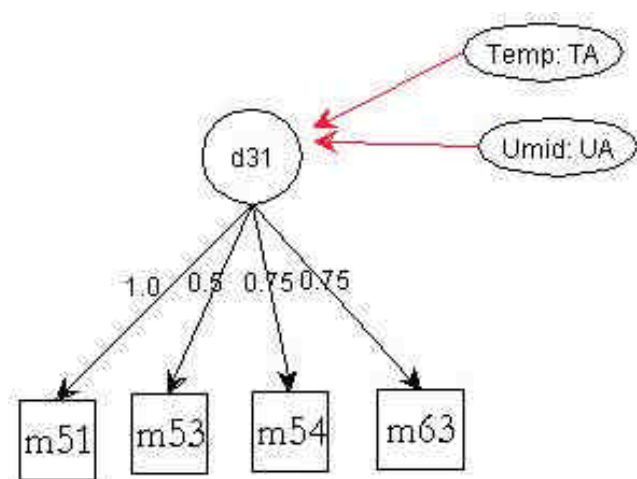


FIGURA C.42 – Podridão da Espiga por Gibberella (d_{30}).

Modelo da desordem D31



Grafo temporal de D31 - G(V,T)

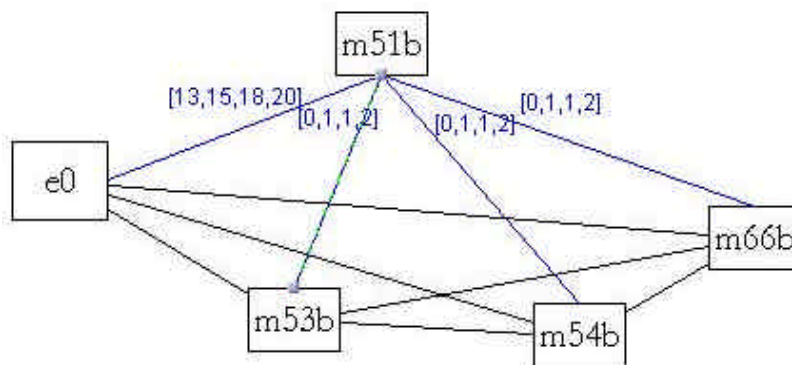
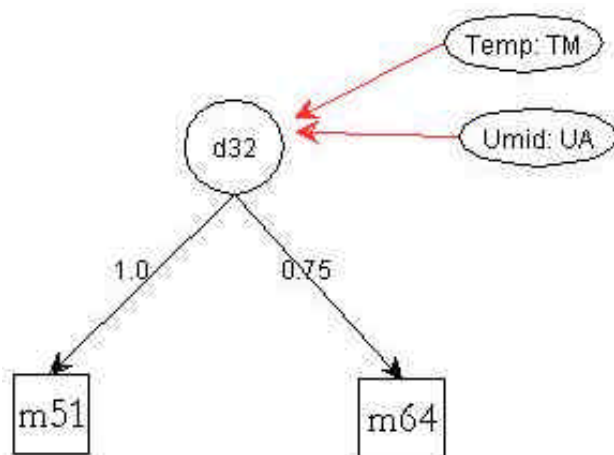


FIGURA C.43 – Podridão Branca da Espiga (*Diplodia Maydis*) (d_{31}).

Modelo da desordem D32



Grafo temporal de D32 - G(V,T)

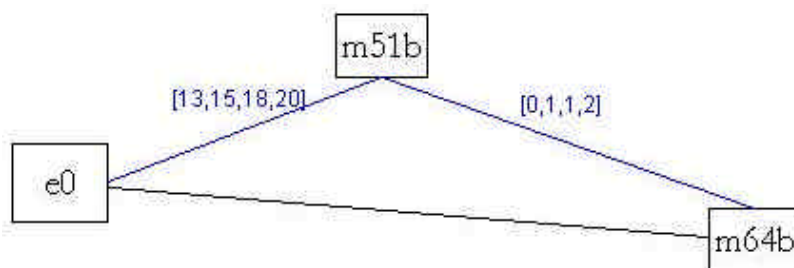
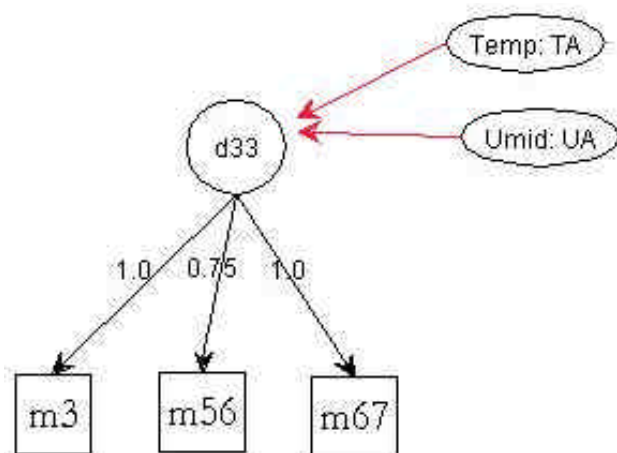


FIGURA C.44 – Podridão da Espiga (*Helminthosporium Maydis*) (d_{32}).

Modelo da desordem 33



Grafo temporal de D33 - G(V,T)

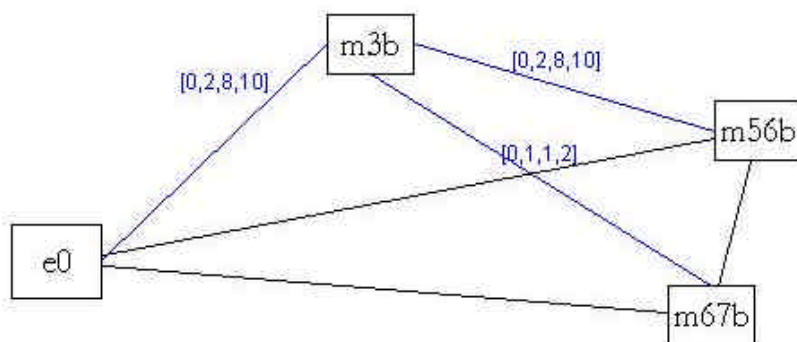
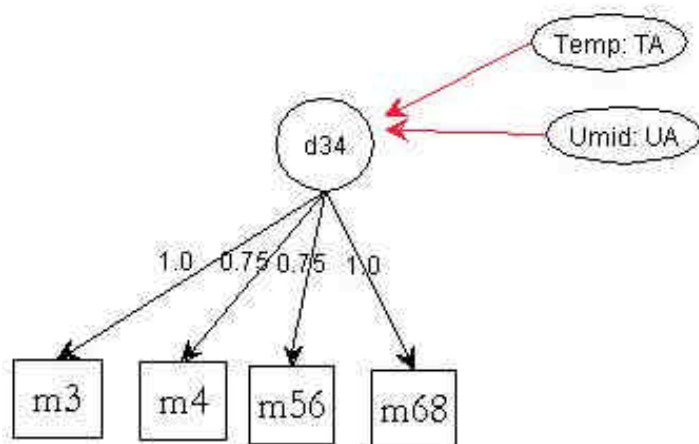


FIGURA C.45 – Podridão do Colmo por Pythium (d_{33}).

Modelo da desordem D34



Grafo temporal de D34 - G(V,T)

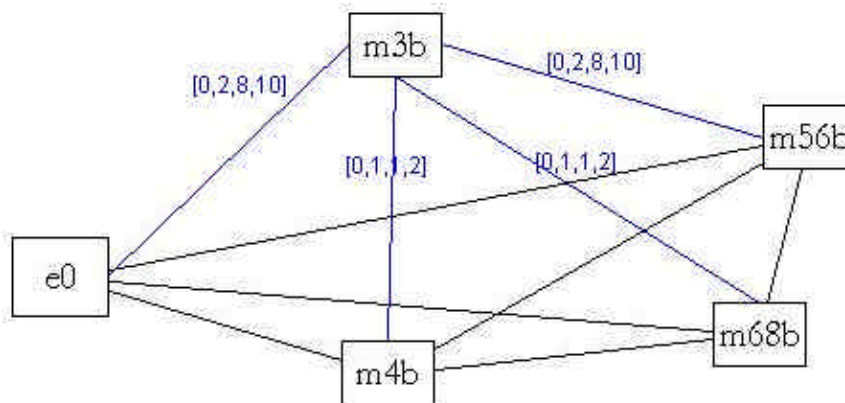
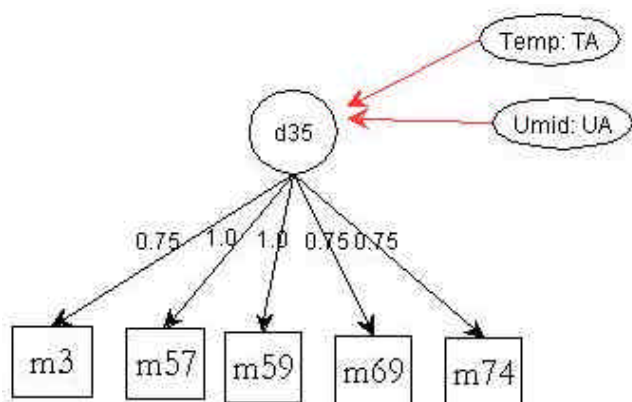


FIGURA C.46 – Podridão do Colmo por *Erwinia Carotovora* (d_{34}).

Modelo da desordem D35



Grafo temporal de D35 - $G(V,T)$

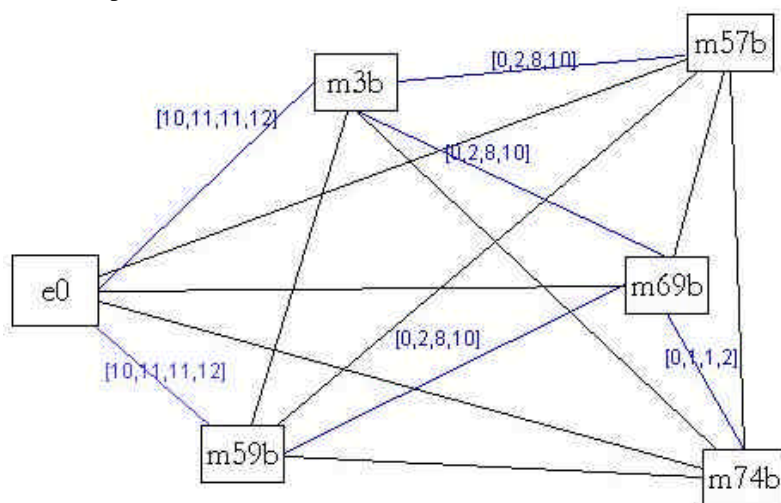
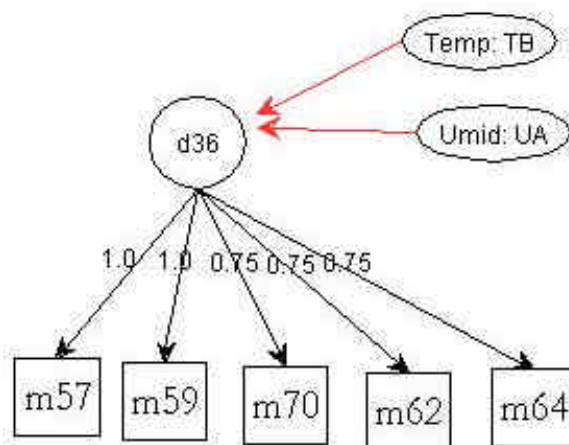


FIGURA C.47 – Antracnose do Colmo (d_{35}).

Modelo da desordem D36



Grafo temporal de D36 - G(V,T)

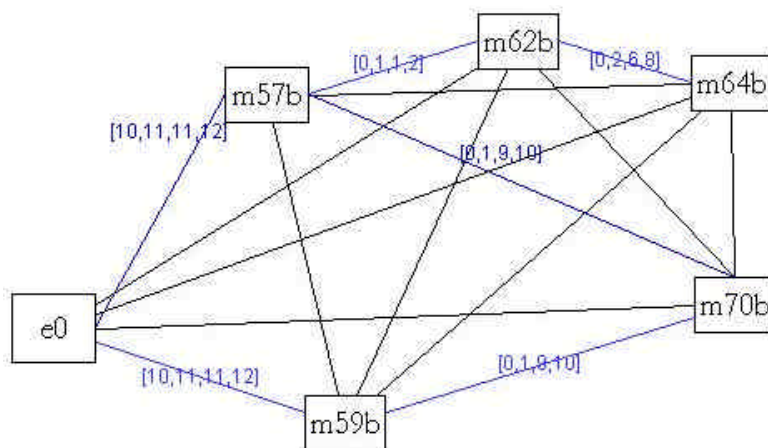
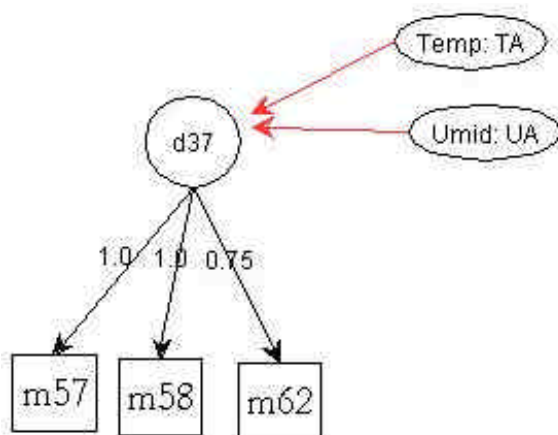


FIGURA C.48 – Podridão do Colmo por *Rhizoctonia* (d_{36}).

Modelo da desordem D37



Grafo temporal de D37 - G(V,T)

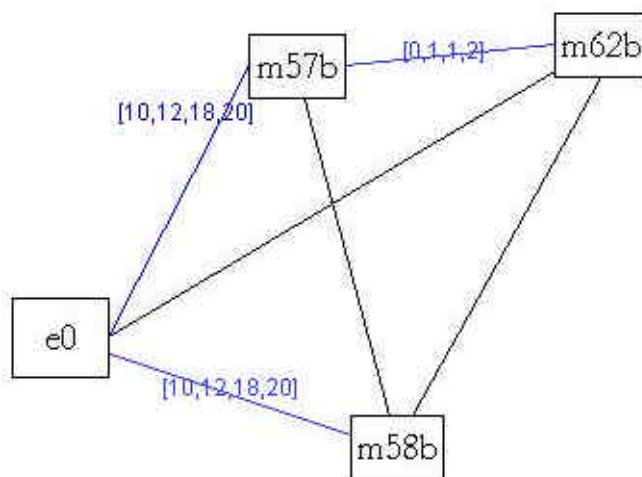


FIGURA C.49 – Podridão do Colmo por Fusarium(Gibberella) (d_{37}).

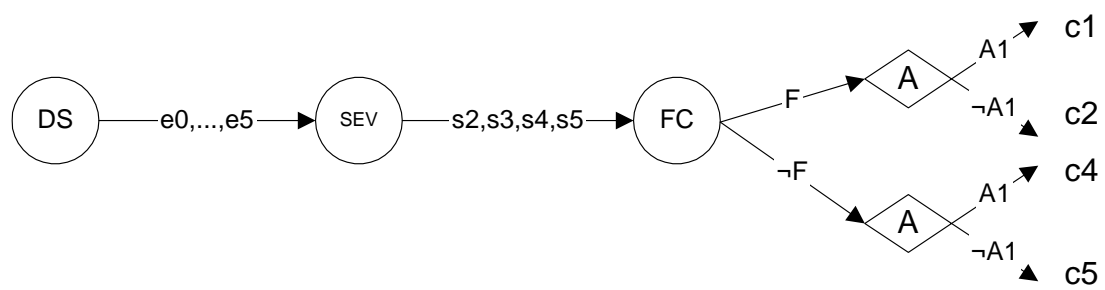
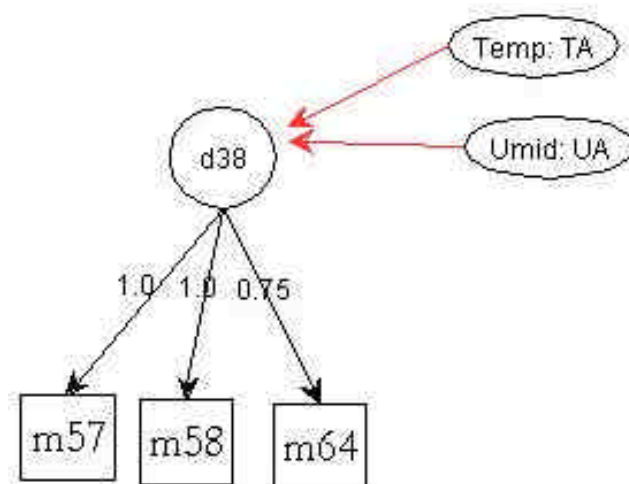


FIGURA C.50 – Podridão do Colmo por Fusarium(Gibberella) (d_{37}) - Variáveis de Decisão.

Modelo da desordem D38



Grafo temporal de D38 - G(V,T)

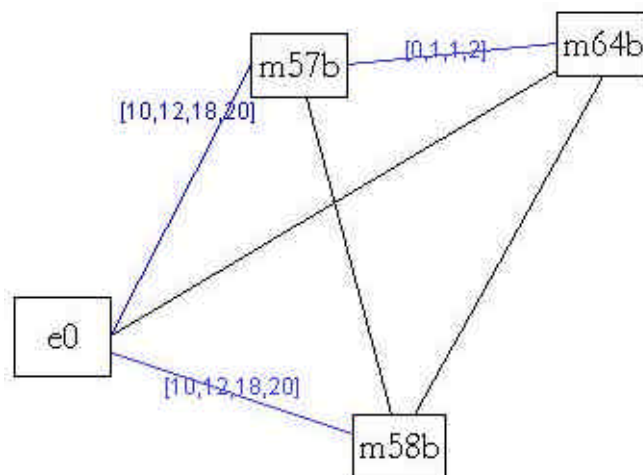
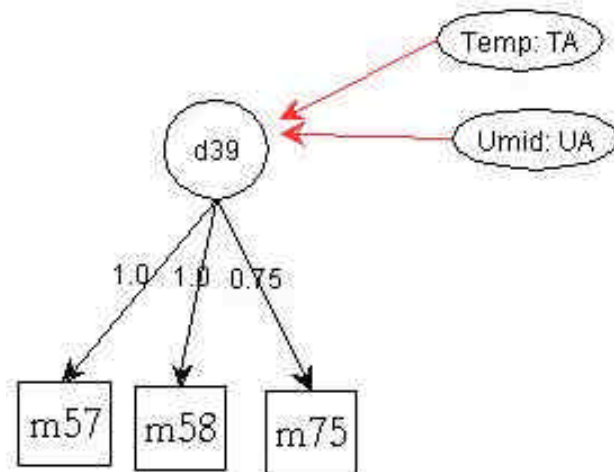


FIGURA C.51 – Podridão do Colmo por Diplodia (d_{38}).

Modelo da desordem D39



Grafo temporal de D39 - G(V,T)

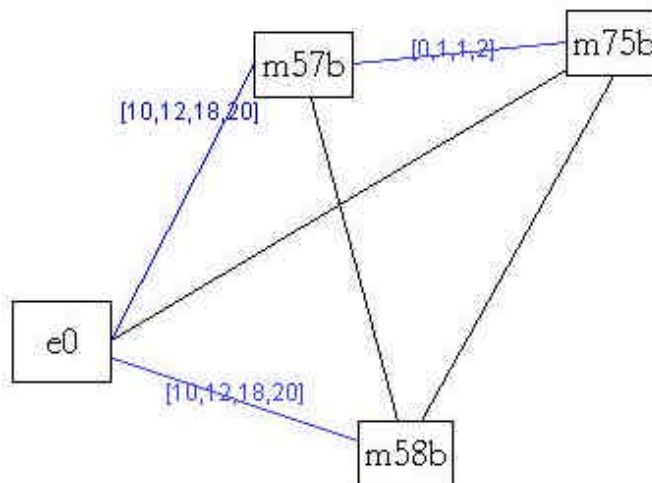
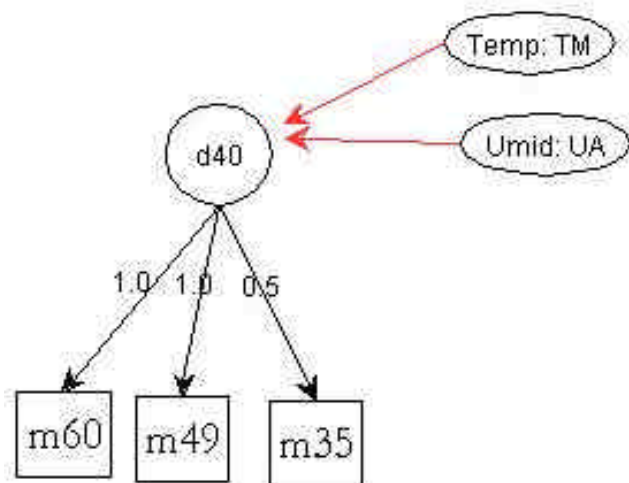


FIGURA C.52 – Podridão do Colmo por *Macrophomina* (d_{39}).

Modelo da desordem D40



Grafo temporal de D40 - G(V,T)

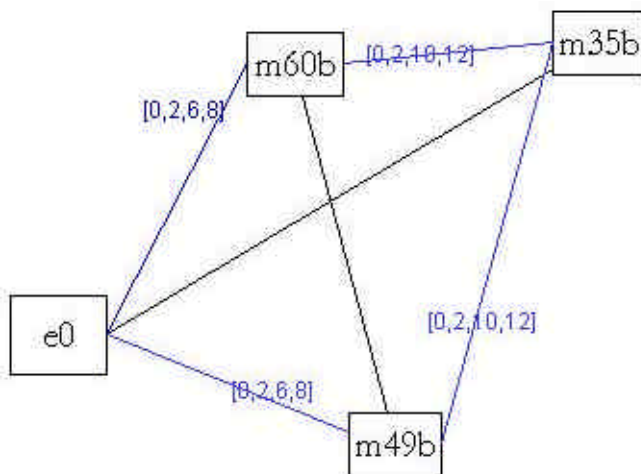
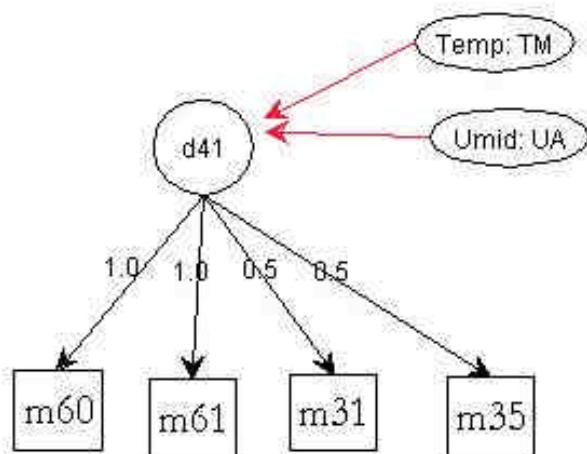


FIGURA C.53 – Carvão do Pendão (d_{40}).

Modelo da desordem D41



Grafo temporal de D41 - G(V,T)

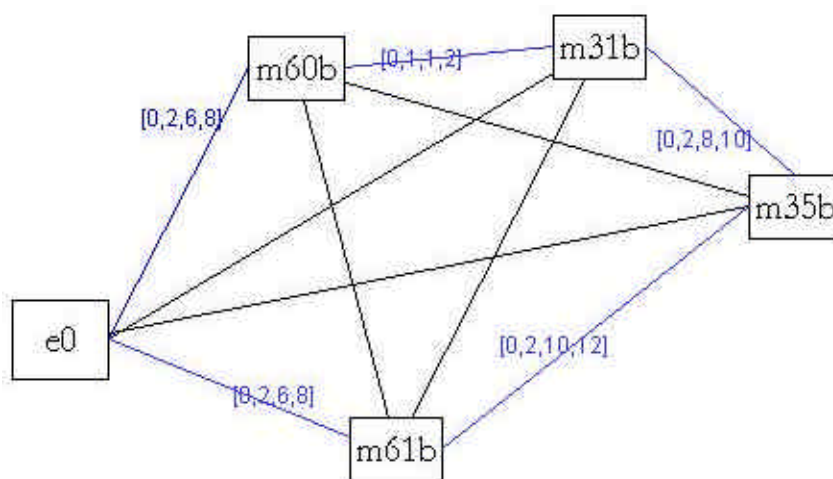


FIGURA C.54 – Míldio do Sorgo no Pendão (d_{41}).

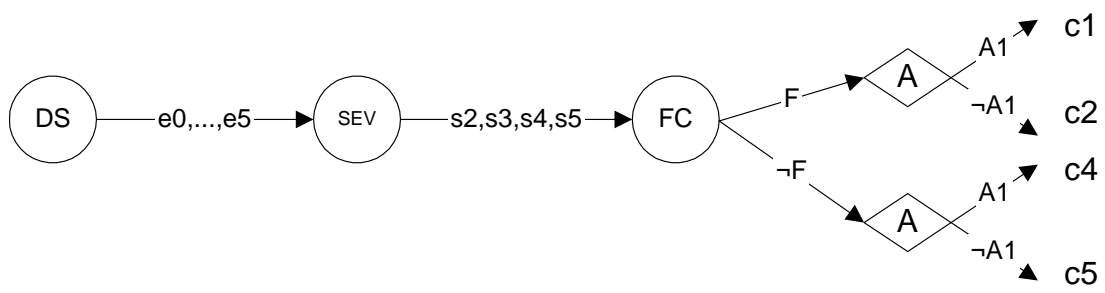


FIGURA C.55 – Míldio do Sorgo no Pendão (d_{41}) - Variáveis de Decisão.