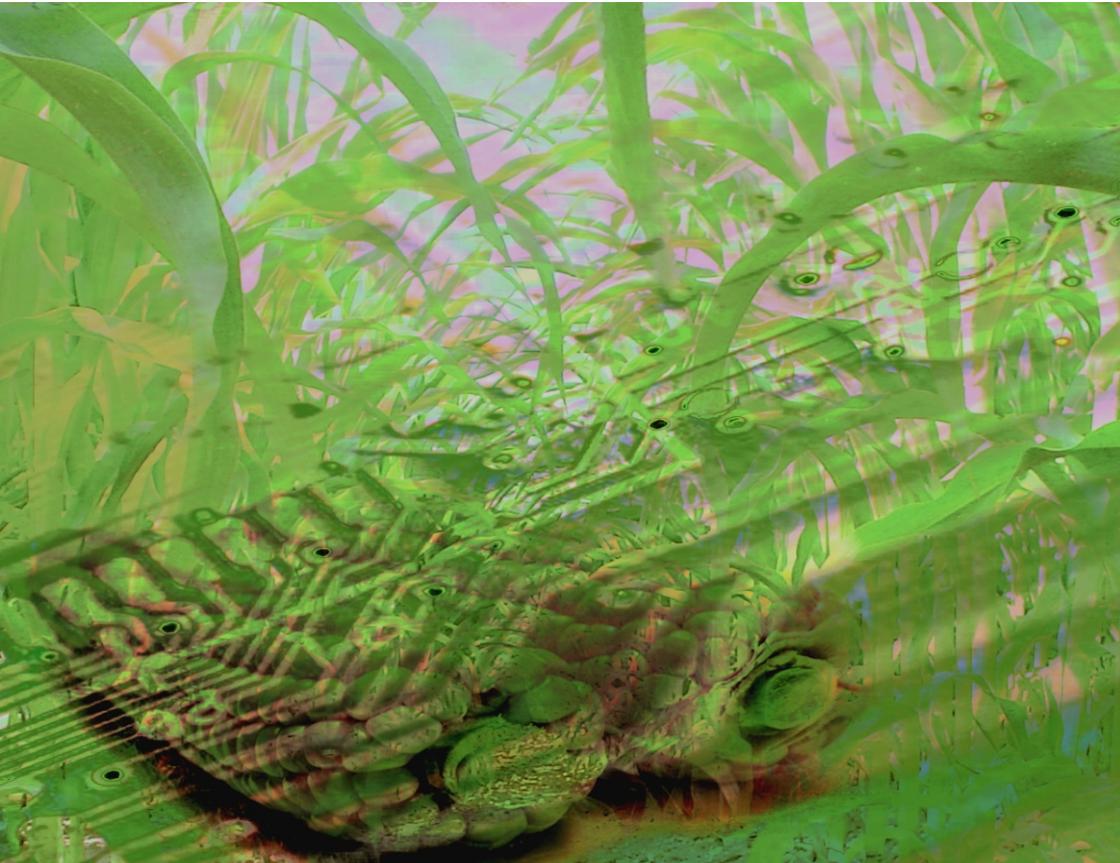


ISSN 1677-9274

Um Sistema Especialista via *Web* para Diagnóstico e Tratamento de Doenças de Milho baseado em um *Framework* orientado a Objetos





Empresa Brasileira de Pesquisa Agropecuária
Embrapa Informática Agropecuária
Ministério da Agricultura, Pecuária e Abastecimento

ISSN 1677-9274
Setembro, 2007

Documentos 69

Um Sistema Especialista via *Web* para Diagnóstico e Tratamento de Doenças de Milho baseado em um *Framework* orientado a Objetos

Silvia Maria Fonseca Silveira Massruhá
Juliano Pastorelli Dutra
Sérgio Aparecido Braga Cruz

Embrapa Informática Agropecuária
Campinas, SP
2007

Embrapa Informática Agropecuária
Área de Comunicação e Negócios (ACN)

Av. André Tosello, 209

Cidade Universitária "Zeferino Vaz" – Barão Geraldo

Caixa Postal 6041

13083-970 – Campinas, SP

Telefone (19) 3789-5743 – Fax (19) 3289-9594

URL: <http://www.cnptia.embrapa.br>

e-mail: sac@cnptia.embrapa.br

Comitê de Publicações

Adriana Farah Gonzalez (secretária)

Ivanilde Dispatto

Kleber Xavier Sampaio de Souza (presidente)

Marcia Izabel Fugisawa Souza

Martha Delphino Bambini

Silvia Maria Fonseca Massruhá

Stanley Robson de Medeiros Oliveira

Suplentes

Goran Neshich

Leandro Henrique Mendonça de Oliveira

Luiz Manuel Silva Cunha

Maria Goretti Gurgel Praxedes

Supervisão editorial: *Ivanilde Dispatto*

Normalização bibliográfica: *Marcia Izabel Fugisawa Souza*

Editoração eletrônica: *Área de Comunicação e Negócios (ACN)*

1ª. edição on-line - 2007

Todos os direitos reservados.

Massruhá, Silvia Maria Fonseca Silveira.

Um sistema especialista via *web* para diagnóstico e tratamento de doenças de milho baseado em um *framework* orientado a objetos / Silvia Maria Fonseca Silveira Massruhá, Juliano Pastorelli Dutra, Sérgio Aparecido Braga Cruz. – Campinas : Embrapa Informática Agropecuária, 2007.

37 p. : il. – (Documentos / Embrapa Informática Agropecuária ; 69).

ISSN 1677-9274

1. Sistema especialista. 2. Inferência abduativa. 3. Lógica nebulosa. 4. Orientação a objeto. 5. Inteligência artificial. 6. Representação do conhecimento. 7. Doença da planta. I. Dutra, Juliano Pastorelli. II. Cruz, Sérgio Aparecido Braga. III. Título. IV. Série.

CDD – 21st ed. 006.3

© Embrapa 2007

Autores

Silvia Maria Fonseca Silveira Massruhá

Doutora em Computação Aplicada,
Pesquisadora da Embrapa Informática
Agropecuária, Caixa Postal 6041,
Barão Geraldo
13083-970 - Campinas, SP
Telefone: 3789-5814
e-mail: silvia@cnptia.embrapa.br

Juliano Pastorelli Dutra

Bacharel em Ciência da Computação,
Consultor da Embrapa Informática
Agropecuária, Caixa Postal 6041,
Barão Geraldo
13083-970 - Campinas, SP
Telefone: 3789-5700
e-mail: juliano@brainweb.com.br

Sérgio Aparecido Braga Cruz

Mestre em Engenharia Elétrica,
Pesquisador da Embrapa Informática
Agropecuária, Caixa Postal 6041,
Barão Geraldo
13083-970 - Campinas, SP
Telefone: 3789-5700
e-mail: sergio@cnptia.embrapa.br

Apresentação

No acervo bibliográfico da Embrapa pode-se encontrar vários documentos sobre “Diagnóstico de Doenças e Métodos de Controle”, tanto na área animal como vegetal passíveis de serem incorporados em *software*. Entretanto, o conhecimento se apresenta disperso nesses documentos. Estes registros acumulam-se com o tempo, tornando difícil a manipulação e também a recuperação do conhecimento embutido nos mesmos. Os especialistas, por sua vez, detêm o conhecimento e possuem atalhos que aprendem com a experiência e que os tornam ágeis em suas decisões. Sistemas especialistas (SE) apresentam-se como uma abordagem eficiente para assegurar que o conhecimento permaneça na empresa após a saída do especialista e seja utilizado amplamente.

A utilização de SE na área de Diagnósticos em geral na agricultura já vem sendo explorada em vários trabalhos. Os grandes desafios da construção destes sistemas estão relacionados com o entendimento, modelagem e representação do conhecimento. A Embrapa Informática Agropecuária tem experiência no desenvolvimento de um sistema para diagnóstico de doenças de milho, que está atualmente em uso e foi desenvolvido em parceria com a Embrapa Milho e Sorgo.

Este trabalho, mais uma contribuição da Embrapa Informática Agropecuária, apresenta os resultados obtidos na especificação, implementação e validação de uma nova versão do sistema especialista para diagnóstico e controle de doenças de milho via *web*. No desenvolvimento deste sistema especialista foi utilizada uma infra-estrutura orientada a objetos que contempla dois módulos principais: Especialista e Produtor. O módulo especialista permite aos fitopatologistas configurarem e alimentarem o sistema com as doenças da cultura, sua sintomatologia bem como as medidas de controle mais recomendadas. O módulo Produtor visa atender produtores, extensionistas, pesquisadores e estudantes interessados na cultura do milho.

Eduardo Delgado Assad
Chefe-Geral

Sumário

Introdução	9
Metodologia	10
Montagem da Arquitetura e Modelagem da Aplicação.....	10
Ambiente de Desenvolvimento.....	15
Estrutura da Aplicação.....	16
Módulo Especialista.....	17
Módulo Produtor.....	24
Resultados e Discussão	30
Conclusões	31
Referências	33
Apêndice	34

Um Sistema Especialista via *Web* para Diagnóstico e Tratamento de Doenças de Milho baseado em um *Framework* orientado a Objetos

*Silvia Maria Fonseca Silveira Massruhá
Juliano Pastorelli Dutra
Sérgio Aparecido Braga Cruz*

Introdução

Na Embrapa Informática Agropecuária foi desenvolvido, em parceria com a Embrapa Milho e Sorgo, um sistema para diagnóstico de doenças de milho, que está atualmente em uso (<http://diagnose.cnptia.embrapa.br>). Os grandes desafios da construção destes sistemas estão relacionados com o entendimento, modelagem e representação do conhecimento. No sistema para diagnóstico de doenças do milho foram utilizadas técnicas manuais para aquisição do conhecimento baseadas em entrevistas, em acompanhamento ou modelos. Quanto à representação do conhecimento foram utilizadas regras de produção que são baseadas nos mecanismos de inferência dedutiva (efeitos -> causa). Pode-se argumentar que o problema principal com este tipo de abordagem é que os fenômenos ocorrem na natureza de forma contrária a este raciocínio.

Na década de 90, o conhecimento passou a ser modelado na sua forma natural: causas -> efeitos, com o raciocínio do tipo abduutivo (Peng & Reggia, 1990). Além do diagnóstico correto, o tratamento efetivo é um outro fator que interfere no sucesso de um processo de raciocínio clínico. Em complementação a estas tarefas, quando a investigação é bem feita, obtém-se maior eficácia e eficiência no diagnóstico, além de reduzir os custos e riscos.

Uma abordagem integrada para o diagnóstico, investigação e tratamento de distúrbios foi proposta por Massruhá (2003) e, posteriormente detalhada em Massruhá et al. (2003, 2004, 2005a, 2005b).

Nesta nova abordagem, denominada Teoria das Coberturas Nebulosas (TCN), o conhecimento é basicamente modelado através de associações

causais e a inferência é abduativa (desordens -> manifestações). Conceitos da Teoria das Coberturas Parcimoniosas (TCP) (Parsimonious Covering Theory - PCT) (Peng & Regia, 1990), lógica nebulosa e teoria de decisão são também integrados, de maneira a tratar os vários aspectos inerentes aos processos envolvidos em raciocínio clínico. A TCN foi validada para diagnóstico, investigação e tratamentos de doenças de milho.

O objetivo deste trabalho é apresentar um sistema especialista para *web* que utilizou um *framework* orientado a objetos que suporta a abordagem integrada para diagnóstico, investigação e tratamento de desordens proposta em Massruhá (2003). No contexto deste trabalho, está sendo validada a flexibilidade e a facilidade de expandir as características do sistema, bem como sua usabilidade, utilizando o *framework* proposto para o desenvolvimento de outros SE para diagnóstico, investigação e tratamentos de desordens tanto na área vegetal quanto animal. Uma base de conhecimento com 41 doenças de milho foi utilizada como estudo de caso do presente trabalho (Massruhá et al., 1999, 2000).

Metodologia

Este trabalho foi desenvolvido em três etapas principais: Modelagem (da arquitetura e da aplicação), especificação do ambiente de desenvolvimento e testes da aplicação. Cada etapa está descrita detalhadamente a seguir.

Montagem da Arquitetura e Modelagem da Aplicação

Primeiramente, foi necessário investigar a Teoria das Coberturas Nebulosas - TCN – proposta em Massruhá (2003) e em Massruhá et al. (2005a, 2005b).

A arquitetura proposta para suportar o modelo TCN é composta por três módulos principais: diagnóstico, investigação e tratamento (Fig. 1). Em todos os três módulos, o conhecimento envolvido é do tipo causal e todo o raciocínio é abduativo.

Na tarefa de diagnóstico, as associações que modelam o conhecimento são do tipo *desordem*¹ *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 abduativo.

¹ No decorrer do trabalho, o termo *desordem* é utilizado no lugar de doença, por ser mais genérico. Nem sempre um agente infeccioso gera uma doença, mas os sintomas tornam possível acusar uma desordem. Além disso, em certas ocasiões mais de um agente pode estar presente e o tratamento deve contemplar isto.

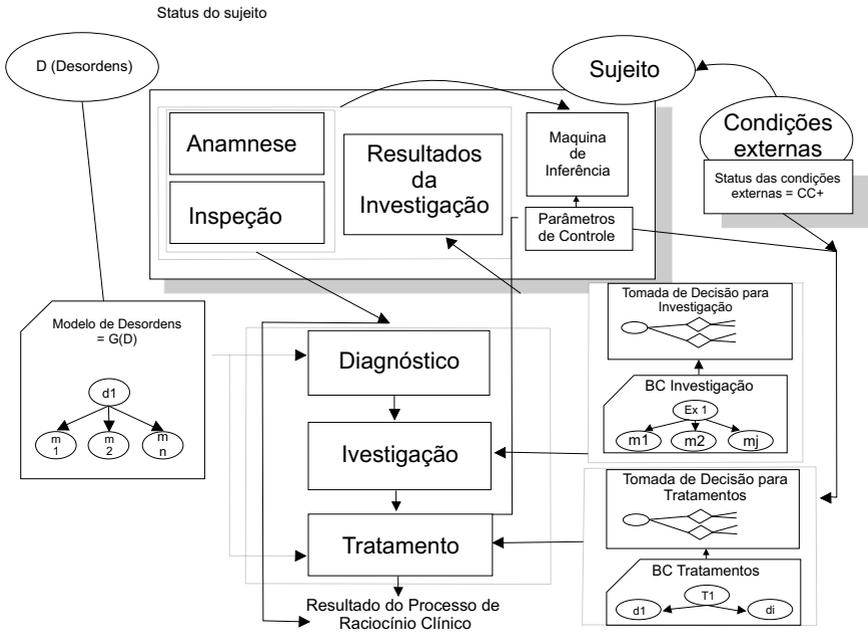


Fig. 1. Um ambiente integrado para diagnose, investigação e tratamento no modelo TCN.

Na tarefa de investigação, as associações são do tipo *exames* → *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* → *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 abdutivo.

A etapa de pré-diagnóstico, 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* → *desordens*. 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.

De forma complementar a análise da TCN, também foram realizadas entrevistas com a própria pesquisadora e analisados os códigos fontes da aplicação utilizados na prova de conceito do modelo. A linguagem utilizada na implementação dos algoritmos da TCN foi o AMZI PROLOG devido a sua sintaxe e aderência ao problema (Amzi, 2004). Adicionalmente, a implementação em PROLOG facilitou a comparação da TCN com o TCP original (Peng & Reggia, 1990) e TCP temporal através das ferramentas que também foram desenvolvidas em PROLOG no escopo do trabalho de Wainer & Rezende (1997).

Os programas em PROLOG que compõem a TCN são claramente divididos em três partes: base de conhecimento, mecanismo de inferência e especificação do caso particular. O mecanismo de inferência abduutivo, apresentado em Massruhá et al. (2005b), foi implementado através dos algoritmos: *bipartido_nebuloso*, *exames_nebulosos* e *tratamentos_nebulosos* (ver trechos do bipartido nebuloso no Apêndice).

A arquitetura da aplicação (Fig. 1) foi revisada e desenhada no modelo três camadas, conforme as especificações do J2EE. Isto permite que se tenha uma simplificação no processo de manutenção e expansões futuras do sistema. O modelo J2EE em três camadas contempla:

- **Camada de Apresentação:** responsável pelo *layout*, navegação e tratamento das ações;
- **Camada de Negócios:** responsável por todas as lógicas e regras que definem a aplicação;
- **Camada de Persistência:** responsável pelo armazenamento e consulta ao Banco de Dados.

Um sistema feito em camadas, com a estrutura apresentada na Fig. 2, proporciona maior dinamismo para expansões futuras.

Apresentação

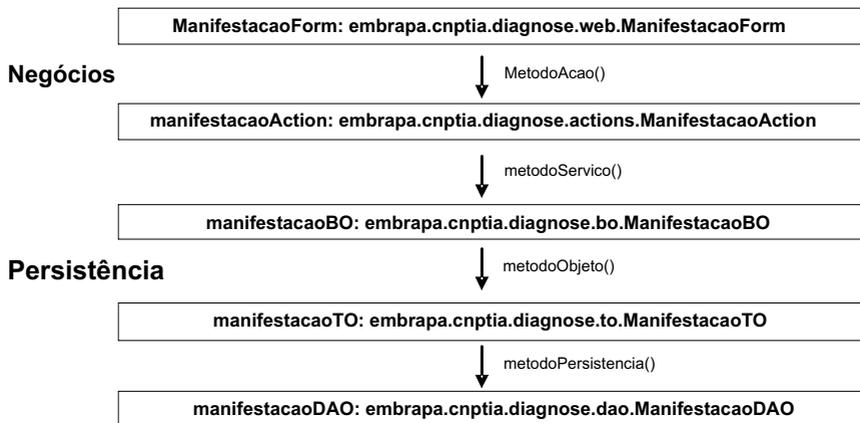


Fig. 2. Diagrama com Arquitetura do Sistema em Camadas.

O segundo passo foi montar uma infra-estrutura do projeto (arquitetura) baseada em definições de padrões de projeto e *frameworks*. Nesta etapa, foram utilizados os mais importantes *frameworks* (em código livre), para tornar mais produtivo o processo de desenvolvimento, a saber:

- Struts: facilita e padroniza o controle das ações vindas da camada *web* e sua apresentação;
- Spring: injeção de dependências, instanciando objetos antecipadamente;
- Hibernate: facilita a persistência e cria uma abstração para maior Independência da base.

Para que o projeto pudesse ser implementado em tecnologia J2EE foi necessário inicialmente modelar o problema sob uma perspectiva de Orientação a Objetos (OO), bem como desenhar toda a base de dados, que no sistema original era apenas uma base simplificada de conhecimento textual em Prolog, contendo apenas a chave de identificação (id) e nome da entidade.

Na Fig. 3 está representado o modelo de classes gerado nesta etapa a partir da revisão da arquitetura da TCN. O modelo é composto por 6 classes principais (Desordens, Manifestações, Exames, Tratamentos, Problema e Caso) e 12 classes auxiliares (Grafo, Nó, Aresta, Nebulosos, CondiçõesControle, Categoria, Fase, Severidade, Pergunta, Cidade, Evento e Usuário).

Na Fig. 4 é mostrado o diagrama final com o banco de dados projetado que contempla os três módulos: diagnóstico, investigação e tratamentos.

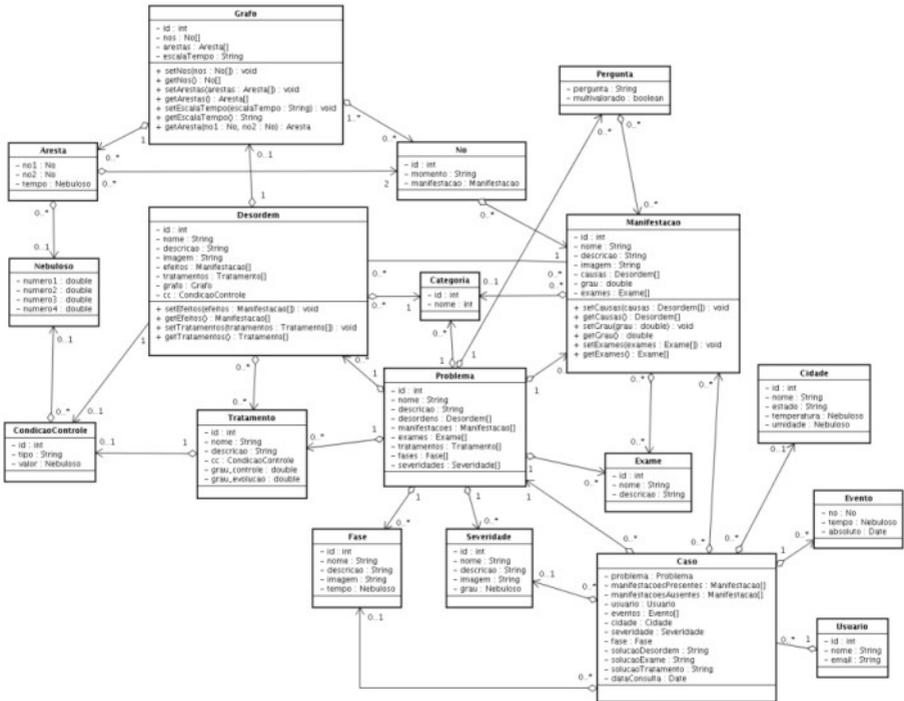


Fig. 3. Modelo de Classes - Orientação a Objetos.

O diagrama da Fig. 4 está representado no Modelo de Entidades e Relacionamentos (MER) que cobre todas as necessidades e requisitos do projeto, em conformidade com a tese apresentada em Massruhá (2003).

Observe que a quantidade de entidades é bastante grande, e que estas possuem muitas relações, expandindo assim as possibilidades e velocidade em relação ao que existe atualmente em forma textual.

Os principais objetivos da primeira etapa foram alcançados com as definições anteriores. O ambiente, modelagem e toda a documentação para a codificação do projeto foram criados, abrindo caminho para o desenvolvimento.

O Eclipse também conta com inúmeros plugins que facilitam a construção de páginas JSP (Java Server Pages) e de XML (eXtensible Markup Language).

Esta primeira versão do sistema roda na forma de uma aplicação *web*, com Banco de Dados em PostgreSQL. Sendo de interesse pode-se mudar o Banco de Dados para outro qualquer, apenas alterando as configurações da camada de persistência. Outra simplicidade proporcionada pelo modelo é a possibilidade de se construir uma versão da aplicação para ser distribuída em CD-ROM, neste caso basta alterar a camada de apresentação, mantendo as outras duas intactas.

O PostgreSQL, além de estar de acordo com as diretrizes governamentais relacionadas ao uso de software livre, é robusto e tem uma alta performance, eliminando assim a necessidade de aquisição de licenças de uso de banco de dados proprietário tanto para o ambiente de desenvolvimento como para o ambiente de produção.

Estrutura da Aplicação

A aplicação pode ser dividida em dois módulos principais conforme apresentado na Fig. 5 – especialista e produtor.

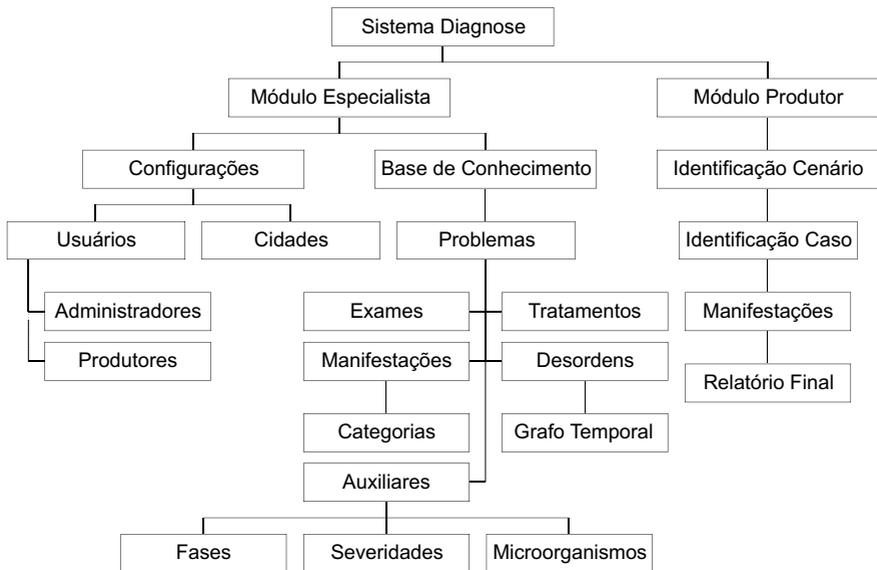


Fig. 5. Árvore de navegação do fluxo de informação do sistema.

Neste ponto fica clara a distinção de papéis entre os especialistas e os produtores rurais. Enquanto os primeiros possuem a responsabilidade de configurar e alimentar o sistema com a base de conhecimento, o segundo grupo entra com as informações específicas do caso, obtendo como resultado uma indicação de quais as prováveis desordens encontradas e recomendações de tratamentos.

A partir da estrutura de organização do diagrama, serão apresentadas algumas telas capturadas dos dois módulos principais do sistema (Fig. 5): Especialista e Produtor. Junto às telas serão tecidos comentários importantes para a compreensão mais ampla do funcionamento do sistema. Os dados apresentados nas capturas foram utilizados apenas nos testes e não representam o conhecimento do fitopatologista ou outro especialista qualquer.

Módulo Especialista

Através de um *login* e senha o especialista entra na área restrita a partir da qual poderá efetuar configurações no ambiente e manter a base de conhecimentos atualizada. Na tela inicial do módulo especialista (Fig. 6), também chamado de administrador do sistema, existem quatro opções: a primeira é para cadastrar toda a base de conhecimento (os problemas), a segunda para cadastrar as cidades, e as duas últimas servem para manter a base de usuários, administradores e produtores rurais, os que utilizaram o sistema para diagnosticar seu caso.



Fig. 6. Tela Inicial do Administrador (Especialista).

De forma análoga existe o cadastro de administradores, os que possuem permissão para cadastrar e manter a base de usuários e conhecimento. Para criar, editar ou remover um usuário especialista, basta clicar em 'Manter Adms' na tela inicial.

O processo de diagnóstico solicita ao produtor rural que informe a cidade na qual se encontra a produção para que se possa fazer recomendações de tratamento mais adequadas ao clima da região. Isto é possível a partir do gerenciamento do cadastro de cidades da tela inicial (opção "CIDADEs"). As informações de uma cidade são utilizadas em todos os problemas, e simplificam o diagnóstico, pois o produtor não precisará saber a temperatura e umidade mínima e máxima da região, o que é essencial para um melhor diagnóstico e assertividade no tratamento. Na Fig. 7 são cadastrados os detalhes das Cidades, tais como, temperatura e umidade. Entretanto, estes dados climáticos, na continuação do projeto, deverão ser importados do Agritempo (www.agritempo.gov.br).



Detalhes de CIDADE

* Nome:

Estado:

Temperatura:

Umidade:

Fig. 7. Tela com dados de uma Cidade.

Nas Fig. 8 e 9, seguem as telas relacionadas à manutenção da base de conhecimento de Problemas.



Listagem de PROBLEMAS

Código	Nome
1	Milho
2	Feijão
3	Soja
4	Gado
5	Frango
6	Porco

Export options: [CSV](#) | [Excel](#) | [XML](#) | [PDF](#)

Fig. 8. Tela de Listagem dos Problemas.

Fig. 9. Tela com Dados e Menu de um Problema.

Cada um dos problemas listados na Fig. 8 simboliza toda uma base de conhecimento específica da cultura ou criação em questão. Como mostrado na Fig. 9, a partir do problema 'Milho', tem-se acesso às suas desordens, manifestações, categorias, fases, severidades, exames, tratamentos e microorganismos.

Em 'Desordens', tem-se cadastro e manutenção das doenças que acometem a cultura. Todas as listagens são exportáveis para os formatos CSV, Excel, XML e PDF como mostrado na Fig. 10. Isto permite gerar cópias da base de conhecimento em outros formatos. Na Fig. 11 é apresentado um exemplo de desordem com suas relações.

Código	Doença	Descrição
2	Podridão da Espiga	
3	Mosaico Comum do Milho	
4	Virose Rayado Fino	
5	Queima por Herbicida	
6	Antracnose do Colmo	
7	Deficiência de Fósforo	

Export options: CSV | Excel | XML | PDF

Fig. 10. Tela de Listagem de Desordens.

Detalhes de DESORDEM

Data Criação: 28/06/2006 09:17:48 Data Atualização: 22/08/2006 01:10:38

* Desordem: 

Descrição:

Agente Causal:

Tipo de Agente:

Cadastrar MANIFESTAÇÕES

NECESSARIAMENTE CAUSA

Listagem de MANIFESTAÇÕES

Manifestação	Grau		
m1	1.0		
m2	1.0		
m3	0.5		

Export options: [CSV](#) | [Excel](#) | [XML](#) | [PDF](#)

Cadastrar CONDIÇÕES CONTROLES

15.0 25.0 Temperatura (°C)

60.0 80.0 Umidade (%)

50% do caule e 25% das folhas estão comprometidos 50% do caule e 25% das folhas estão comprometidos

Severidade

Germinação Fase

Cadastrar TRATAMENTOS

t1

t2

t3

t4

t5

t6

t7

t8

t9

t10

t11

t12

t13

t14

t15

Fig. 11. Tela com Dados de uma Desordem.

Uma desordem, em conformidade com o modelo possui relação com manifestações, tratamentos e condições de controle e pode também ter um grafo temporal nele associado que indica a quantidade de semanas que separa duas manifestações (sintomas) (Fig. 12). Estas informações são utilizadas no momento em que um caso é submetido à consistência temporal.

Detalhes de GRAFO

Cadastrar ARESTAS

Manifestação :	Coloração Avermelhada	Momento :	Início		Salvar Aresta
Manifestação :	Excesso de Brotações	Momento :	Início		Cancelar Aresta
Tempo :	2.0 4.0	Semanas			

Listagem de ARESTAS

Momento	Manifestação	Momento	Manifestação	Tempo		
inicio	Coloração Avermelhada	fim	Excesso de Brotações	[2,0,2.75,4.25,5,0]		
inicio	Folhas se Coalecem	inicio	Lesões Alongadas	[3,0,3.75,5,25,6,0]		
inicio	Seca em Grandes Áreas Foliare	inicio	Folha se Desprendem Facilmente	[1,0,2,0,4,0,5,0]		
fim	Excesso de Brotações	fim	Folhas se Coalecem	[5,0,5.75,7,25,8,0]		
inicio	Folha se Desprendem Facilmente	inicio	Coloração Avermelhada	[4,0,4,5,5,5,6,0]		
fim	Lesões Alongadas	fim	Seca em Grandes Áreas Foliare	[8,0,8,5,9,5,10,0]		

Export options: [CSV](#) | [Excel](#) | [XML](#) | [PDF](#)

Fig. 12. Tela com Dados de um Grafo Temporal.

Uma das relações mais importantes é desordem -> manifestações e o grau de relação que a desordem possui com cada uma das manifestações, conforme Fig. 11. As manifestações são os possíveis sintomas apresentados na cultura e que são utilizados na tentativa de diagnosticar a doença e seu possível agente causador. Na Fig. 13 é apresentado um exemplo de manifestação com suas relações.

Fig. 13. Tela com Dados de uma Manifestação.

Observe a relação estabelecida entre a manifestação e os exames propostos. A pergunta que aparece no cadastro é a que será feita ao produtor no caso desta manifestação ser relevante na investigação. As possíveis respostas do produtor são sempre: Sim, Não e Não Sei.

A categoria que classifica uma manifestação também pode ser cadastrada para servir de filtro inicial para o produtor. A primeira pergunta que ele responde junto com o cenário é o tipo (categoria) de sintomas que encontrou em sua plantação ou criação. Na Fig. 14 é mostrada a tela na qual pode-se cadastrar, alterar ou remover uma categoria.

Listagem de Categorias	
Código	Nome
1	Folha
2	Colmo
3	Raiz
4	Espiga

Export options: [CSV](#) | [Excel](#) | [XML](#) | [PDF](#)

Fig. 14. Tela de Listagem de Categorias.

Observe que se trata de um cadastro secundário, bastante simples, mas que torna possível filtrar bastante as possibilidades de perguntas (manifestações) já numa primeira interação do produtor rural com o sistema de diagnóstico. A partir desta categorização é possível também gerar futuramente estatísticas que informem o tipo de manifestação mais facilmente observada, as que são normalmente ignoradas, etc.

Na desordem Fig. 11, também existe uma caixa de seleção para determinação do intervalo de fases e severidades (início e fim), nas quais esta desordem costuma aparecer. As Fig. 15 e 16 mostram imagens do cadastro e manutenção destas entidades, que também são informadas pelo produtor rural, no ato em que submete seu caso para consulta.

Listagem de FASES		
Código	Nome	Descrição
1	Germinação	
2	Crescimento Vegetativo	
3	Pendoamento	
4	Florescimento	
5	Frutificação	
6	Maturação	

Export options: [CSV](#) | [Excel](#) | [XML](#) | [PDF](#)

Fig. 15. Tela de Listagem de Fases.

As fases e severidades cadastradas estão relacionadas com um intervalo nebuloso, que será utilizado nos cálculos de hipóteses de desordens.

Listagem de SEVERIDADES		
Código	Nome	Descrição
1	Sem Lesões	Antes do aparecimento dos sintomas
2	Lesões Esparsas	Restritas ao terço inferior da planta
3	Até 50% das Folhas com Lesões	Severas nos 25% das folhas inferiores
4	Até 75% das Folhas com Lesões	Severas nos 50% das folhas inferiores
5	100% das Folhas com Lesões	Lesões severas nos 75% das folhas inferiores
6	Planta Morta	Ausência dos sintomas

Export options: [CSV](#) | [Excel](#) | [XML](#) | [PDF](#)

Fig. 16. Tela de Listagem de Severidades.

As entidades exames e tratamentos apresentam uma importância muito grande no funcionamento final do algoritmo, já que a proposta é justamente propor exames para sintomas não identificados (Fig. 17), e tratamentos, a partir do momento em que se atinja determinado nível de confiabilidade de diagnóstico (Fig. 18).

Listagem de EXAMES		
Código	Nome	Descrição
1	Análise Microscópica	Microscopia Ótica
2	Análise de Ribossomo	Doenças Bacterianas
3	Identificação Sorológica	Doenças Fúngicas
4	Isolamento em Cultura Pura	Doenças Fúngicas
5	Testes Bioquímicos	Doenças Bacterianas
6	Testes Sorológicos	Viroses e Enfezamentos

Export options: [CSV](#) | [Excel](#) | [XML](#) | [PDF](#)

Fig. 17. Tela de Listagem de Exames.

A seguir são apresentados os tratamentos cadastrados e na seqüência os detalhes de um tratamento. Como se pode ver na Fig. 11 é possível se associar mais de um tratamento para a mesma desordem. O interesse em se propor mais de um possível tratamento é poder sugerir o mais adequado para o cenário apresentado pelo produtor (clima, fase, etc.) visando subsidiar sua tomada de decisão. Na Fig. 18 é apresentado o exemplo de uma lista de tratamentos cadastrados para controle de doenças do milho e registrados pelo Ministério da Agricultura, Pecuária e Abastecimento (Agrofit, 2003). Na Fig. 19 são apresentados os detalhes para cadastramento dos tratamentos.

Listagem de TRATAMENTOS	
Código	Nome
1	Bromo Fersol
2	Captan 750 TS
3	Folicur 200 CE
4	Maxim XL
5	Tecto 600
6	Triade

Export options: [CSV](#) | [Excel](#) | [XML](#) | [PDF](#)

Fig. 18. Tela de Listagem dos Tratamentos.

Detalhes de TRATAMENTO		
* Nome:	Bromo Fersol	
Descrição:	Methyl bromide	
Grau Controle:	1.0	
Grau Evolução:	5.0	
Agente infeccioso:	Bactéria	
Cadastrar CONDIÇÕES CONTROLES		
15.0	25.0	Temperatura
70.0	85.0	Umidade
Lesões Esparsas	Até 75% das Folhas com Lesões	Severidade
Crescimento Vegetativo	Florescimento	Fase
10.0	15.0	Custo
1.0	3.0	Risco

Fig. 19. Tela com Dados de um Tratamento.

Ao terminar o módulo especialista, foram feitas validações por parte da Embrapa, para que fosse possível contemplar todo o modelo proposto em Massruhá (2003) e a necessidade de servir de auxílio para o módulo do produtor rural a ser descrito a seguir.

Módulo Produtor

Esta parte do sistema é acessada publicamente, sem a necessidade de ter um *login* e senha. Basta que o produtor preencha a tela de registro apresentada na Fig. 20, com nome e *e-mail* e ele passará a usufruir do sistema de diagnóstico e investigação. É possível também fazer uma consulta anônima (não preenchendo os dois campos abaixo). Entretanto, se for necessário, é fácil reproduzir o sistema de autenticação do administrador para o produtor, permitindo, por exemplo, que este retorne ao diagnóstico

exatamente do ponto em que parou. Esta é uma grande novidade a ser contemplada nos trabalhos futuros.



The image shows a web browser window displaying the registration page for rural producers on the Embrapa website. The header features a banner with a tractor and the URL <http://www.cnpia.embrapa.br>. The Embrapa logo and the text "Embrapa Informática Agropecuária" are visible. The main content area contains the text "Bem-vindo ao Diagnose !" followed by two input fields labeled "nome:" and "e-mail:". Below these fields is a button labeled "ENTRAR..". At the bottom of the page, there is a copyright notice: "Copyright © 2004. Todos os direitos reservados Embrapa Informática Agropecuária".

Fig. 20. Tela para Registro do Produtor Rural.

O produtor precisa identificar o cenário de sua cultura. Através do cenário (Fig. 21) é possível identificar a produção que vai ser diagnosticada. São variáveis que definem o cenário: a localização, o tipo de cultura e a fase em que se encontra a plantação ou criação. É importante destacar que os campos são interligados utilizando a moderna estrutura do AJAX (<http://ajaxtags.sourceforge.net/>). Isto significa que ao escolher o Estado, as opções disponíveis para Cidade passam a ser limitadas a ele. O mesmo ocorre quando escolhe-se o ramo de atividade, que pré-seleciona os tipos de atividade rural, que por fim limitam as possíveis fases de crescimento. Este encadeamento de respostas elimina distorções ou incoerências na entrada de informações por parte do produtor rural.

Na seqüência são solicitadas informações sobre as partes da planta ou animal que estão afetadas, ou melhor, os pontos nos quais estão visíveis os sintomas da doença (Fig. 22). Além disso, pode ser fornecido o grau de severidade observado, conforme uma lista de opções. Estas informações juntamente com as fornecidas no cenário serão processadas para formar a primeira lista de possíveis doenças. A este processo dá-se o nome de pré-diagnóstico, que como pode-se ver na Fig. 23 é de bastante apelo visual. O objetivo neste caso é dar início ao processo de investigação. Se o produtor errar a imagem que mais se parece com a doença isto retarda o encontro da solução, mas não impede o diagnóstico.

The screenshot shows a web browser window with a header banner for 'Embrapa Informática Agropecuária' and the URL 'http://www.cnpia.embrapa.br'. Below the banner are 'INÍCIO' and 'SAIR' buttons. The main content area is a form titled 'Identificação do Cenário' with the following fields:

- Localização Geográfica :
 - * Estado : São Paulo (dropdown)
 - * Cidade : Campinas (dropdown)
 - * Ramo de Atividade : Agricultura (dropdown)
 - * Tipo : Milho (dropdown)
 - * Fase de Desenvolvimento : Germinação (dropdown)
- Tamanho da Produção : 10 (text input)

A 'Continuar' button is located at the bottom right of the form. At the bottom of the page, there is a copyright notice: 'Copyright © 2004. Todos os direitos reservados Embrapa Informática Agropecuária'.

Fig. 21. Tela de Identificação do Cenário do Produtor Rural.

The screenshot shows a web browser window with the same header banner as Figure 21. Below the banner are 'INÍCIO' and 'SAIR' buttons. The main content area is a form titled 'Identificação da SEVERIDADE' with the following fields:

- Você observou sintomas em quais partes da planta?
 - folhas colmo espigas pendão
- Grau de severidade : 50% do caule e 25% das folhas estão comprometidos (dropdown)

A 'Continuar' button is located at the bottom right of the form. At the bottom of the page, there is a copyright notice: 'Copyright © 2004. Todos os direitos reservados Embrapa Informática Agropecuária'.

Fig. 22. Tela de Identificação do Problema do Produtor Rural.

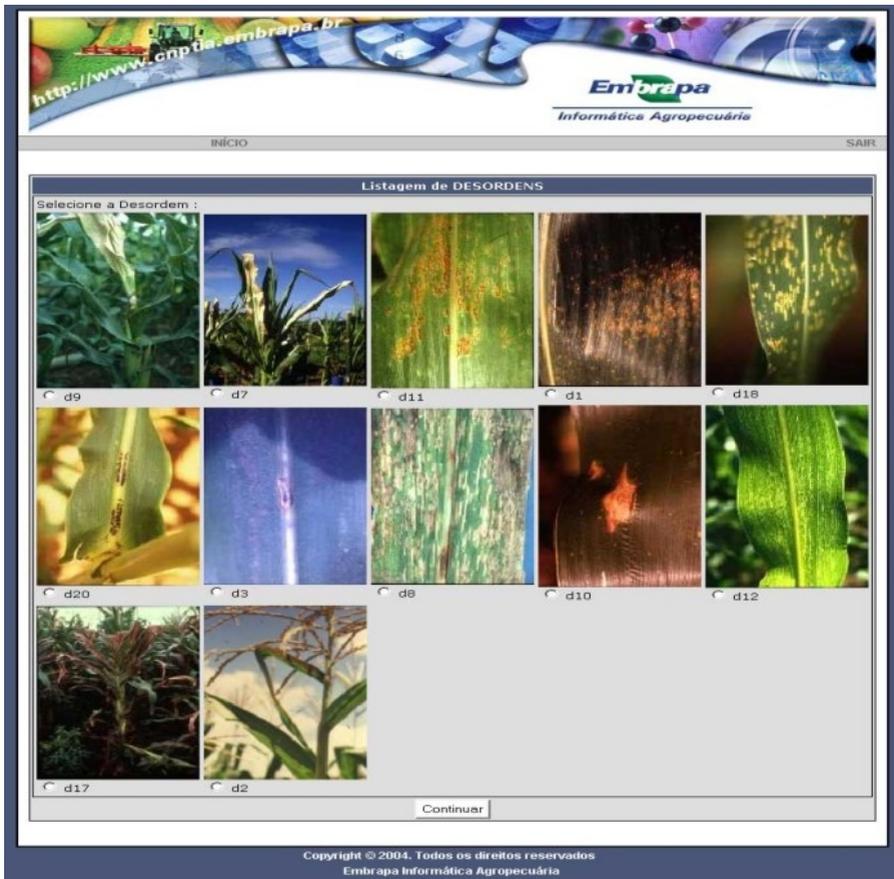


Fig. 23. Tela para Sugestão de Problema.

O sistema usando o método de investigação já consegue informar quais são as perguntas mais relevantes para a rápida redução do conjunto de hipóteses. Isto não é nada simples, pois se a pergunta for para manifestações que necessariamente causam a doença ela pode ser tão específica, que em caso afirmativo reduz o tamanho do conjunto para, por exemplo, 2%, praticamente definindo qual é a doença, o que seria excelente. Mas para o caso de uma resposta negativa, a quebra não terá auxiliado muito na resolução do problema, mantendo 98% das possibilidades.

No outro extremo, se o algoritmo escolher apenas os sintomas mais fracos (aqueles que raramente definem uma doença, mas que estão sempre presentes) a filtragem também seria muito lenta, pois poucas hipóteses seriam descartadas por vez. Para tanto o sistema utiliza um algoritmo que

considera todas as manifestações filtradas e, através de uma combinação, indica as perguntas mais relevantes. Mais detalhes podem ser encontrados em Massruhá (2003) ou em Massruhá et al. (2005b).

http://www.cnpia.embrapa.br

Embrapa
Informática Agropecuária

INÍCIO SAIR

Identificação da PERGUNTA



m15?

Sim
 Não
 Não Sei

Início :
Fim :

Continuar

Copyright © 2004. Todos os direitos reservados
Embrapa Informática Agropecuária

Fig. 24. Tela de Identificação dos Sintomas. Exemplo de Pergunta.

Com as perguntas sendo feitas em ordem de relevância, para cada resposta positiva uma nova manifestação é adicionada à coleção das presentes, enquanto as negativas compõem as ausentes. As perguntas são feitas até que se atinja um nível de confiança adequada determinada pelo usuário.

Após cada resposta positiva, o método de diagnóstico é executado formando um lista das manifestações mais relevantes a serem indagadas. Estas manifestações são passadas uma a uma na forma de pergunta ao usuário. Sempre que a resposta for negativa ou desconhecida, o sistema avança para a próxima da fila. Apenas no caso onde uma nova manifestação presente é agregada que se roda novamente o método de diagnóstico, com possibilidade de se refinar ainda mais as possibilidades.

Na Fig. 24 é apresentado um exemplo do que ocorre quando não há imagem cadastrada no sistema. Como pode-se observar, através desta tela, é possível capturar não apenas se o sintoma ocorre como também quais seriam os eventos associados: início e fim deste sintoma (se é que já teve

fim). Portanto, além do 'Sim', 'Não' e 'Não Sei' outra coisa que se captura em caso afirmativo é quando o sintoma começou e/ou terminou.

Após cada pergunta e sua correspondente resposta o sistema analisa, utilizando os métodos de investigação e diagnóstico, qual é a pergunta seguinte que pode ser mais relevante na determinação da doença em questão. Este é o principal papel destes dois algoritmos.

Na seqüência, o sistema parte para uma nova iteração, até que se atinja determinado nível de satisfação (delta). Este delta simboliza a probabilidade de um determinado problema estar ocorrendo. Se atingir o limite estabelecido no cadastro do problema então é interrompido o processo de investigação e diagnóstico para chamar o tratamento e o relatório final.

Neste relatório estão descritas as possíveis doenças, manifestações desconhecidas (mesmo após o questionário) e as mais importantes: as sugestões de exames e tratamentos para guiar o produtor na sua decisão de investimento (Fig. 25).



Fig. 25. Tela de Relatório do Diagnóstico.

O sistema armazena todos os casos consultados em uma base o que permitirá, futuramente, avaliar quais são as desordens que mais aparecem, em quais regiões, quais condições climáticas, etc. visando auxiliar na detecção ágil de epidemias, ou na identificação de problemas sazonais, para trabalhos de prevenção.

Resultados e Discussão

De forma resumida, foram alcançados os seguintes resultados:

- Utilização da mesma Interface Gráfica (com a introdução do AJAX para produtividade).
- Implementação da Estrutura de Classes (segundo o modelo de objetos desenhado).
- Tratamento de Erros e Validação das Entradas (proporcionando confiabilidade).
- Desenvolvimento dos CRUDs (cadastro, remoção, edição e listagem das entidades).
- Configuração do Ambiente (para que tudo rodasse adequadamente). Realização de Testes de Lógica (em todas os métodos de negócio).

Para testar o núcleo do sistema, foi necessário realizar testes modulares *bottom-up*. No início foram estudadas as entradas e saídas de métodos auxiliares, tais como: nebulosos, intervalos, etc. Todo o princípio dos testes era o de executar a mesma chamada em Prolog e Java e comparar a saída, que deveria ser idêntica. Na seqüência foram realizados também testes similares com os métodos específicos com cada uma das três etapas do processo: diagnóstico, investigação e tratamento. A partir da conclusão destes testes e da manutenção da base de conhecimento, pode-se afirmar que tem-se um sistema completo e que atende às necessidades e expectativas de seus usuários, tanto especialistas como produtores.

Após a conclusão do sistema, o projeto foi validado pela equipe responsável dentro da Embrapa, como forma de cancelar os requisitos desenvolvidos versus o que havia sido planejado.

Para validar o sistema, 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, foi

utilizada como estudo de caso no presente trabalho (<http://diagnose.cnptia.embrapa.br>).

Para testar o módulo produtor, foram criados casos idênticos na versão Java e Prolog para na seqüência os resultados serem comparados. Como no Prolog todas as entradas são em lote e via linha simples de texto é mais fácil alterá-los nos testes do que o contrário.

A partir da detecção de qualquer anomalia, testes de execução passo-a-passo são realizados para que se possa identificar o momento exato em que os resultados divergem, para isto é utilizada a ferramenta de *debug* que vem no Eclipse, que como mencionado na primeira parte do projeto é o ambiente no qual o sistema fora desenvolvido.

Foram gerados três procedimentos de testes principais:

- Testes do processo de investigação.
- Testes do processo de diagnóstico.
- Testes do processo de tratamento.

Conclusões

O resultado final alcançado, neste projeto de pesquisa, foi bastante satisfatório tanto do ponto de vista de arquitetura da aplicação quanto do ponto de vista da aplicação final. Pode-se dizer que o projeto foi um grande sucesso. O sistema está pronto para trabalhar com múltiplas culturas, é possível trabalhar com um número infinito de desordens, manifestações, exames e tratamentos. Além disso, o sistema está extremamente intuitivo tanto para o produtor rural como para os administradores que irão alimentá-lo com dados. A preocupação com o visual foi recompensada com o uso apropriado das imagens, o que vai facilitar a identificação das desordens. E por fim, é fundamental informar que a parte central do sistema, composta pelos algoritmos de investigação, diagnóstico e tratamento foi testada à exaustão e eventuais erros foram identificados e corrigidos.

No escopo do aplicação, pode-se que concluir que os dois módulos principais do sistema, especialista e produtor, mostraram-se eficientes e robustos. O módulo especialista mostrou-se robusto para o gerenciamento da base de conhecimentos e o módulo produtor bastante eficiente do ponto de vista da lógica para a tomada de decisões.

O sistema apresenta apenas duas limitações: a apresentação visual do grafo temporal e a parte de monitoramento. A primeira delas não impacta em nada no processo de diagnóstico em si, apenas colaboraria para uma visão mais simples e lúdica do que o grafo temporal representa dentro do sistema. O monitoramento por outro lado pode ser muito útil para que produtores possam voltar e dar continuidade em um processo investigativo, que fora interrompido para se fazer um exame ou mais observações.

Portanto, o sistema pode ainda ser expandido para várias frentes, tanto para o lado de facilitar o monitoramento, como para reaproveitar os dados históricos na detecção de epidemias ou ainda para apontar o grau de eficiência das soluções que estão sendo propostas.

Referências

- AGROFIT: relatório de acaricidas, bactericidas, fungicidas, inseticidas ... Brasília, DF: Ministério da Agricultura, Pecuária e Abastecimento, 2003. 145 p. Relatório técnico.
- AMZI! *Prolog+Logic Server*: tutorials. Disponível em: <<http://www.amzi.com>>. Acesso em: 24 nov. 2004.
- MASSRUHÁ, S. M. F. S. *Incorporação de ferramentas inteligentes na Agência de Informação Embrapa*. [Campinas: Embrapa Informática Agropecuária, 2004]. 30 p. (Embrapa. Macroprograma 3 – Desenvolvimento Tecnológico Incremental. Projeto).
- MASSRUHÁ, S. M. F. S. *Uma teoria de coberturas nebulosas para diagnóstico, investigação e tratamento*. 2003. 251 p. Tese (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- MASSRUHÁ, S. M. F. S.; CRUZ, S. A. B. da; SOUZA, E. de. Diagnose virtual: um sistema para diagnóstico de doenças do milho via web. In: CONGRESSO DA SBI-AGRO, 2. = AGROSOFT 99, 1999, Campinas. *Anais ...* [S.l.: s. n.], 2000. Não paginado.
- MASSRUHÁ, S. M. F. S.; Sandri, S. A.; Wainer, J. *Fuzzy covering theory: an alternative approach for diagnostic problem-solving*. In: EUROPEAN CONFERENCE FOR INFORMATION TECHNOLOGY IN AGRICULTURE – EFITA, 4th, 2003, Budapest. *Proceedings*. Budapest: HAAI, 2003. p. 768-775.
- MASSRUHÁ, S. M. F. S.; SANDRI, S. A.; WAINER, J. Ordering manifestations for investigation in incomplete diagnosis. In: INTERNATIONAL CONFERENCE IPMU, 10., 2004, Perugia, Italy. *Proceedings...* Perugia, 2004. p. 1153-1160.
- MASSRUHÁ, S. M. F. S.; SANDRI, S.; WAINER, J.; MORANDI, M. An integrated framework for clinical problem solving in agriculture. In: EFITA /WCCA 2005 JOINT CONGRESS ON IT IN AGRICULTURE, 2005, Vila Real, Portugal. *Proceedings...* [Vila Real: Efitas, 2005a]. p. 1400-1407.
- MASSRUHÁ, S. M. F. S.; SANDRI, S.; WAINER, J.; MORANDI, M. *Uma abordagem nebulosa para solução de problemas de diagnóstico, investigação e tratamento de desordens*. Campinas: Embrapa Informática Agropecuária, 2005b. 41 p. (Embrapa Informática Agropecuária. Boletim de Pesquisa e Desenvolvimento, 12).
- MASSRUHÁ, S. M. F. S.; SOUZA, E. de; ROMANI, L. A. S.; CRUZ, S. A. B. Virtual services for agricultural technology transfer. In: EUROPEAN CONFERENCE OF THE EUROPEAN FEDERATION FOR INFORMATION TECHNOLOGY IN AGRICULTURE, FOOD AND THE ENVIRONMENT, 2., 1999, Bonn, Germany. *Role and potential of IT, Intranet and Internet for advisory services*. Bonn: Universität Bonn-ILB, 1999. p. 53-62.
- PENG, Y.; REGGIA, J. A. *Abductive inference models for diagnostic problem-solving*. New York: Springer Verlag, 1990. 285 p.
- SUN Microsystems. *Java 2 Platform, Enterprise Edition (J2EE)*. Disponível em: <<http://java.sun.com/j2ee/index.jsp>>. Acesso em: 10 jun. 2005.
- WAINER, J.; REZENDE, A. A temporal extension to the parsimonius covering theory. *Artificial Intelligence in Medicine*, v. 10, n. 3, p. 235-255, July, 1997.

Apêndice

Exemplo de trechos do código em PROLOG que compõem a TCN que são claramente divididos em três partes: base de conhecimento, mecanismo de inferência e especificação do caso particular.

```
%#####
#%          Base de Conhecimento (KB)
%#####
#
% doenças foliares
doenca(1,podridao_cartucho_por_erwinia):-!.
doenca(2,queima_por_pseudomonas):-!.
doenca(3,ferrugem_comum):-!.
doenca(4,ferrugem_polysora):-!.
doenca(5,ferrugem_branca_ou_tropical):-!.
doenca(6,mancha_por_physoderma):-!.
doenca(7,antracnose):-!.
doenca(8,mancha_por_helminthosporium_maydis(raca_O)):-!.
doenca(9,mancha_por_helminthosporium_maydis(raca_T)):-!.
Doenca(10,mancha_por_cercospora):-!.
.
.
.
%manifestacoes
causes(m1,folhas_se_coalescem,[1,2,10]):-!.
causes(m2,seca_em_grandes_areas_foliares,[1,2,10]):-!.
causes(m3,podridoes_aquosas,[1,2,18,33,35]):-!.
causes(m4,odor_desagradavel,[1,34]):-!.
causes(m5,folhas_se_desprendem_facilmente,[1]):-!.
causes(m6,pustulas_individualizadas_marrom_claro,[3,4]):-!.
causes(m7,pustulas_individualizadas_marrom_escuro,[3,4]):-!.
Causes(m8,rompimento_da_epiderme,[3,4]):-!.

% grafo temporal

arcos(1,[x,m1b,[2,4,6,8]], [m1b,m2b,[0,1,1,2]], [m2b,m4b,[0,1,1,2]], [m2b,m5b,[0,1,1,2]], [x,m3b,[2,3,5,6]], [x,m2b,[2,4,7,9]], [x,m4b,[2,5,8,11]], [x,m5b,[2,5,8,11]], [m1b,m3b, [-6,-3,1,4]], [m1b,m4b,[-9,-4,1,6]], [m1b,m5b,[-9,-4,1,6]], [m2b,m3b,[-7,-4,1,4]], [m3b,m4b,[-4,0,5,9]], [m3b,m5b,[-4,0,5,9]], [m4b,m5b,[-9,-3,3,9]]):-!.
```



```
%-----Operacoes Com Intervalos nebulosos-----
```

```
Contido([I1,J1,K1,F1],[I2,J2,K2,F2]):-
```

```
  F1 =< F2,
```

```
  K1 =< K2,
```

```
  J2 =< J1,
```

```
  I2 =< I1.
```

```
soma_intervalos([I1,J1,K1,F1],[I2,J2,K2,F2],[I,J,K,F]):-
```

```
  write('soma_intervalos_fuzzy'),nl,
```

```
  I is I1 + I2,
```

```
  J is J1 + J2,
```

```
  K is K1 + K2,
```

```
  F is F1 + F2.
```

```
intersecao_intervalos([I1,J1,K1,F1],[I2,J2,K2,F2],[I,F]):-
```

```
  max(I1,I2,I),
```

```
  min(F1,F2,F).
```

```
sub_fuzzy([I1,J1,K1,F1],[I2,J2,K2,F2],[I,J,K,F]):-
```

```
  I is I1 - F2,
```

```
  J is J1 - K2,
```

```
  K is K1 - J2,
```

```
  F is F1 - I2.
```

```
.  
.
.
```

```
bipartite(Obs_Presentes,Eventos,Obs_Correntes,Hipoteses_Consistentes,DNI,DI_
Atual,Lcic1_Atual,Lcic2_Atual,Resultado,Lista_ind).
```

```
bipartite([Mj|Obs_Restantes],Eventos,Obs_Correntes,Hipoteses,DNI,DI,Lcic1,Lcic2,
Resultado,Lista_ind):-!,
```

```
causes(Indj,Mj,H1), nl,
```

```
write('Mj= '),write(Mj),nl,
```

```
write('H1= '),write(H1),nl,
```

```
bipartite([],_,_X,_,_X,X2):-!. %final do problema. Tentativas = Solucao Final
```

```
Bipartite(_,_,_,_,_,_):- write('Nao existe nenhuma hipotese consistente com o
caso !!!').
```

```
%#####  
#%          ESPECIFICAÇÃO DO CASO PARTICULAR  
%#####
```

#

Intervalo_diagnostico([0,0,22,22]).

MAIN:-

% Exemplo tese (d3 e d4) M+={m6,m7,m8,m13}

Bipartite([pustulas_individualizadas_marrom_claro,pustulas_individualizadas_marro
m_escuro,seca_prematura_das_plantas,rompimento_da_epiderme,[formacao_de_
fendas],[[m6b,[6,7,8,9]],m7b,[16,17,18,19]],m8b,[7,8,9,10]],m13b,[7,8,9,10]]],[x,[0
,0,0,0]],[],[],[],[],R,R2).



Informática Agropecuária

**Ministério da
Agricultura, Pecuária
e Abastecimento**

