

Modelagem Multinível de Conceitos Clínicos para Apoio à Decisão ao Diagnóstico de Demência, Doença de Alzheimer e Transtorno Cognitivo Leve

Flavio L. Seixas¹, Luciana T. Cavalini², Timothy W. Cook²,
Débora C. Muchaluat-Saade¹, Aura Conci¹

¹Instituto de Computação – Universidade Federal Fluminense (UFF)
Rua Passo da Pátria, 156 – 24.210-240 – Niterói – RJ – Brasil

²Faculdade de Ciências Médicas – Universidade do Estado do Rio de Janeiro (UERJ)
Av. 28 de setembro, 87 – 20.551-030 – Rio de Janeiro – RJ – Brasil

fseixas@ic.uff.br, lutricav@lampada.uerj.br, tim@mlhim.org,
debora@midia.com.uff.br, aconci@ic.uff.br

Abstract. *This paper describes the clinical concepts modeling for a decision support system applied to clinical diagnosis of Dementia, Alzheimer's Disease and Mild Cognitive Impairment. The decision model is based on a Bayesian network, where each clinical concept is represented by a node on causal diagram. The semantic interoperability issue is addressed using a multilevel approach described by MLHIM (Multilevel Healthcare Information Model). Specifications are presented and the results are discussed, focusing on the usage of multilevel information principles to support data models designed for clinical decision support systems.*

Resumo. *Este artigo tem por objetivo descrever a modelagem dos conceitos clínicos utilizados em um sistema de apoio à decisão ao diagnóstico de Demência, Doença de Alzheimer e Transtorno Cognitivo Leve. O modelo de decisão é baseado em uma rede Bayesiana, onde cada conceito clínico é codificado por um nodo em um diagrama causal. O problema da interoperabilidade semântica foi considerado através de uma abordagem multinível descrita pelo modelo de referência MLHIM (Multilevel Healthcare Information Model). As especificações são discutidas, realçando de que forma os princípios da modelagem multinível podem apoiar a construção do modelo de dados em um sistema de apoio à decisão.*

1. Introdução

As preocupações crescentes com a segurança do paciente e com o gerenciamento de riscos na área clínica têm motivado inúmeras pesquisas e desenvolvimento de produtos relacionados a sistemas de apoio à decisão [Miller, 2009]. Sistemas de apoio à decisão clínica, conhecidos pelo acrônimo CDSS (*Clinical Decision Support System*), são sistemas computacionais utilizados para ajudar os médicos na tomada de decisão no momento do atendimento. Há evidências da redução da quantidade de erros médicos no processo clínico, após a implantação de CDSS [Haynes e Wilczynski, 2010].

Além disso, o processo de envelhecimento populacional, observado nos países desenvolvidos e também no Brasil, tem elevado a prevalência de doenças mais comuns em idosos, como a Doença de Alzheimer (DA) [Luzardo *et al.*, 2006]. A DA é

caracterizada por uma afecção neurodegenerativa e irreversível, causando lesões no cérebro [Jalbert *et al.*, 2008]. A DA é considerada um subtipo de Demência, e seus estágios iniciais envolvem dificuldades de recuperação e retenção de novas informações, seguidas de outras deficiências cognitivas. Assim, o Transtorno Cognitivo Leve (TCL) é de interesse da comunidade médica, pois pode representar um estágio pré-clínico de DA [Petersen, 2003].

Uma preocupação importante quanto ao desenvolvimento de sistemas de apoio à decisão é o uso de uma arquitetura informacional comum, possibilitando a sua interoperação semântica [Berner, 2007]. Sendo assim, este trabalho apresenta o processo de modelagem dos conceitos clínicos utilizados em um sistema de apoio à decisão aplicado ao diagnóstico de Demência, DA e TCL [Seixas, 2012]. O modelo de decisão empregado é baseado em redes Bayesianas e diagrama de influência [Howard e Matheson, 2005]. As redes Bayesianas têm recebido considerável atenção de pesquisadores em vários campos, incluindo ciência da computação, ciência cognitiva e estatística. Em ciência da computação, o desenvolvimento de redes Bayesianas é impulsionado por pesquisas em inteligência artificial, cuja meta é propor um *framework* de representação e inferência do raciocínio comum [Friedman *et al.*, 2000].

Este artigo apresenta a estratégia de adoção da modelagem multinível para a representação dos conceitos clínicos utilizados em um sistema de apoio à decisão [Seixas, 2012], de acordo com as especificações *Multilevel Healthcare Information Model* (MLHIM) [Cavalini e Cook, 2012]. Optou-se por adotar a modelagem multinível em função de sua capacidade de prover interoperabilidade semântica entre sistemas, além de facilitar futuras atualizações do modelo de dados [Garde *et al.*, 2007]. A interoperabilidade semântica é definida pela capacidade de sistemas diferentes trocarem informações semanticamente válidas entre si.

As próximas seções estão organizadas da seguinte maneira. A Seção 2 comenta alguns trabalhos relacionados. A Seção 3 apresenta a modelagem da rede Bayesiana utilizada no sistema de apoio à decisão. A Seção 4 descreve a modelagem multinível dos dados e a especificação de conceitos utilizados na rede Bayesiana. Por fim, a conclusão e trabalhos futuros são descritos na Seção 5.

2. Trabalhos Relacionados

A modelagem de conceitos utilizados na área de saúde apresenta complexidades relacionadas a diferentes dimensões. Na dimensão espacial, um conceito clínico pode ser específico a uma região geográfica, ou instituição de saúde, ou mesmo, parte anatômica ou sistema biológico de um organismo [Zusman, 2010]. Na dimensão temporal, os dados de saúde devem persistir no longo prazo, independente da evolução tecnológica ou de conhecimentos etiológicos das doenças [Zywietz, 2004]. Por exemplo, a *Systematized Nomenclature of Medicine - Clinical Terms* (SNOMED-CT), terminologia de referência em medicina, apresenta mais de 311.000 conceitos da área de saúde, relacionados entre si por milhões de conexões, o que demonstra a complexidade ontológica nesse domínio do conhecimento.

Existem algumas arquiteturas de informação que se propõem a definir estruturas de relacionamentos, assim como procedimentos de modelagem, para os conceitos da área de saúde. Os principais modelos de arquitetura informacional utilizados na área de saúde são: o padrão HL7 [Dolin *et al.*, 2006], a Norma ISO 13606 [Santos *et al.*, 2010],

as especificações *openEHR* [Kalra *et al.*, 2005] e as especificações MLHIM (*Multilevel Healthcare Information Modeling*) [Cavalini e Cook, 2011]. Em HL7, os dados são modelados utilizando uma estrutura referencial denominada *Reference Information Model* (RIM), de forma semelhante a um modelo relacional encontrado em banco de dados baseado em *Structured Query Language* (SQL). Já a Norma ISO 13606 e as especificações *openEHR* e MLHIM seguem os princípios de modelagem multinível, visando atingir interoperabilidade semântica entre sistemas distribuídos [Cavalini e Cook, 2011].

Na modelagem multinível, o primeiro nível é denominado Modelo de Referência ou *Reference Model* (RM), no qual se descrevem as unidades básicas de informação, através da combinação de tipos de dados, relacionamentos e regras de restrições. O conceito clínico é descrito no segundo nível, o Modelo de Domínio ou *Domain Model* (DM), utilizando de forma concreta o modelo de restrição especificado abstratamente no primeiro nível. Em uma perspectiva de interoperabilidade semântica sustentável, o primeiro nível deve representar uma estrutura estável, sem a necessidade de intervenções técnicas, gerada pela revisão ou a criação de novos conceitos. Por outro lado, como os conceitos clínicos estão descritos no segundo nível do modelo, eles podem evoluir sempre que necessário, sem que isso afete os mecanismos de persistência dos aplicativos [Sundvall, 2013].

As especificações *openEHR* são geridas por uma Fundação responsável por licenciar o Modelo de Referência e as ferramentas de modelagem correspondentes, e inspirou a elaboração da Norma ISO 13606. Em ambas, as definições dos conceitos clínicos são denominadas arquétipos. Há exemplos de implementação das especificações *openEHR*, tais como o trabalho de Atalag e Yang (2010), no qual se desenvolveu uma aplicação para registro do laudo de endoscopias, o trabalho de Dias *et al.* (2011), que adotaram *openEHR* para remodelar os dados presentes no sistema de Troca de Informação em Saúde Suplementar (TISS), ou ainda Ratamero *et al.* (2010) que utilizaram *openEHR* para representar laudos de exames de ressonância magnética da coluna lombar. Gomes *et al.* (2012) propuseram uma arquitetura e modelo de engenharia de software, construído com base nas especificações *openEHR*, denominado SPLiCE (*Software Product Line for healthCarE*). Embora a qualidade técnica das especificações *openEHR* seja indiscutível, o modelo de governança dos arquétipos é centralizado, o que exige o consenso de toda a comunidade de especialistas do domínio antes da publicação de um conceito, tornando o processo consideravelmente lento e oneroso [Shalom *et al.*, 2009].

Em MLHIM, as definições de conceito são denominadas *Concept Constraint Definition* (CCD), que são expressos em *XML Schema Definition* (XSD), padrão W3C baseado em XML para descrição de dados, incluindo a estrutura do documento, os tipos de dados e as regras de restrição. Esse formato vem sendo amplamente utilizado em sistemas computacionais que necessitam assegurar a interoperabilidade semântica dos dados [Sokolowski e Dudeck, 1999].

CCDs expressam restrições ao Modelo de Referência, sendo que o modelo de governança dos CCDs é comunitário e descentralizado, dado que CCDs são identificados de forma unívoca por *Unique Universal Identifiers* (UUIDs), o que permite a publicação de um CCD sem que haja um consenso global. Desta forma, é possível a geração de mais de uma descrição para um mesmo conceito, garantindo-se a

interoperabilidade semântica através da cadeia de validação que parte da instância de dados, para o CCD, para o Modelo de Referência, e finalmente para os padrões W3C. Contudo, com a prática, espera-se que os melhores CCDs relacionados a um dado conceito clínico se sobressaiam e sejam naturalmente selecionados para uso nos sistemas [Cavalini e Cook, 2011].

3. Modelagem da Rede Bayesiana

Uma rede Bayesiana é descrita por $B = (G, P)$ como uma representação de uma distribuição de probabilidade conjunta [Friedman *et al.*, 2000]. O primeiro componente G é um grafo acíclico dirigido, consistindo de vértices V , e arestas E , ou seja, $G = (V, E)$. Os vértices V representam o domínio das variáveis aleatórias X_1, \dots, X_n , também denominados de nodos. Cada nodo X_i contém um valor discreto x_i de um domínio finito $Val(X_i)$. As arestas representam as dependências causais entre os nodos. Uma aresta partindo do nodo X_i em direção ao nodo X_j , significa que X_i é pai de X_j . O segundo componente, P , da rede Bayesiana B , representa as distribuições de probabilidades condicionadas para cada nodo $P(X_i)$, dado o conjunto de pais $Pa(X_i)$ em G . Assim, é possível especificar uma distribuição de probabilidades conjuntas do domínio X_1, \dots, X_n , como forma do produto mostrado na Equação 1.

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i)) \quad (1)$$

Para obter a distribuição conjunta, as respectivas probabilidades condicionais utilizadas no produto devem estar especificadas. O componente P deve descrever a distribuição de valores possíveis x_i de X_i , em todas as combinações de valores possíveis $pa(X_i)$ em $Pa(X_i)$, onde $pa(X_i)$ representa o conjunto de valores de $Pa(X_i)$. Estas probabilidades condicionais são representadas por tabelas, denominadas Tabelas de Distribuição de Probabilidades (CPT – *Conditional Probability Tables*). Utilizando um mecanismo de inferência, é possível calcular a probabilidade *a posteriori* da variável aleatória X_i assumir um estado x_{ij} , dada as evidências de outras variáveis aleatórias. Nos próximos parágrafos, a modelagem da rede Bayesiana utilizada para o diagnóstico de DA será descrita.

Neste trabalho, a estrutura da rede Bayesiana G foi definida manualmente utilizando a estrutura genérica de três níveis descrita por Pradhan (1994). As variáveis aleatórias foram especificadas utilizando características presentes nos critérios de diagnóstico das doenças de interesse, e os atributos da base de casos clínicos. A base de casos clínicos é composta pelos registros de sintomas, sinais, resultados de testes neuropsicológicos, diagnóstico do médico e dados demográficos de pacientes assistidos pelo Centro de Tratamento da Doença de Alzheimer (CDA) do Instituto de Psiquiatria da Universidade Federal do Rio de Janeiro (IPUB-UFRJ), conforme projeto nº 284/2010, aprovado pelos Conselhos de Ética em Pesquisa (CEP) do IPUB-UFRJ e da Universidade Federal Fluminense (UFF).

Pelo processo de diagnóstico utilizado no CDA, o diagnóstico de DA é avaliado para os pacientes que tiveram o diagnóstico positivo de demência. Na base de casos clínicos, 180 pacientes foram classificados com diagnóstico positivo de demência. Dentre esses, há um subconjunto com 45 casos negativos de DA, e outro com 135 casos positivos de DA. Foi aplicado um método de balanceamento da base de pacientes de DA por superamostragem, conhecida por SMOTE (*Synthetic Minority Over-sampling*

Technique) [Chawla *et al.*, 2002], totalizando 135 casos positivos e 90 casos negativos. Os parâmetros P da rede Bayesiana foram estimados utilizando um método de aprendizagem supervisionada denominado EM (*Expectation-Maximization*) [Dempster *et al.*, 1977]. Os resultados obtidos indicaram uma área da curva ROC (*Receiver Operating Characteristic*) de 0,86, com taxa de verdadeiros-positivos de 0,94 e verdadeiros-negativos de 0,70. Comparando com outros classificadores, os melhores resultados para área da curva ROC foram obtidos pelas redes neurais multicamadas Perceptron (0,90) [Witten *et al.*, 2011]. Contudo, um ponto negativo do modelo neural é a ausência de semântica dos coeficientes obtidos da aprendizagem.

A Figura 1 mostra a representação gráfica da rede Bayesiana modelada para o diagnóstico de DA. O modelo de decisão inclui a função utilidade e caixas de decisão. Como comentado, a rede Bayesiana foi estruturada em três níveis: no primeiro foram colocados os fatores predisponentes da DA, disponíveis na base de casos clínicos. No segundo, a variável aleatória consultada que representa a probabilidade de diagnóstico de DA. E no terceiro nível estão os sintomas e resultados de testes neuropsicológicos. A função utilidade foi modelada de forma simétrica, ou seja, um verdadeiro positivo tem um mesmo peso do verdadeiro negativo. As caixas de decisão estão conectadas de forma que a decisão para diagnóstico de Demência preceda o diagnóstico de DA.

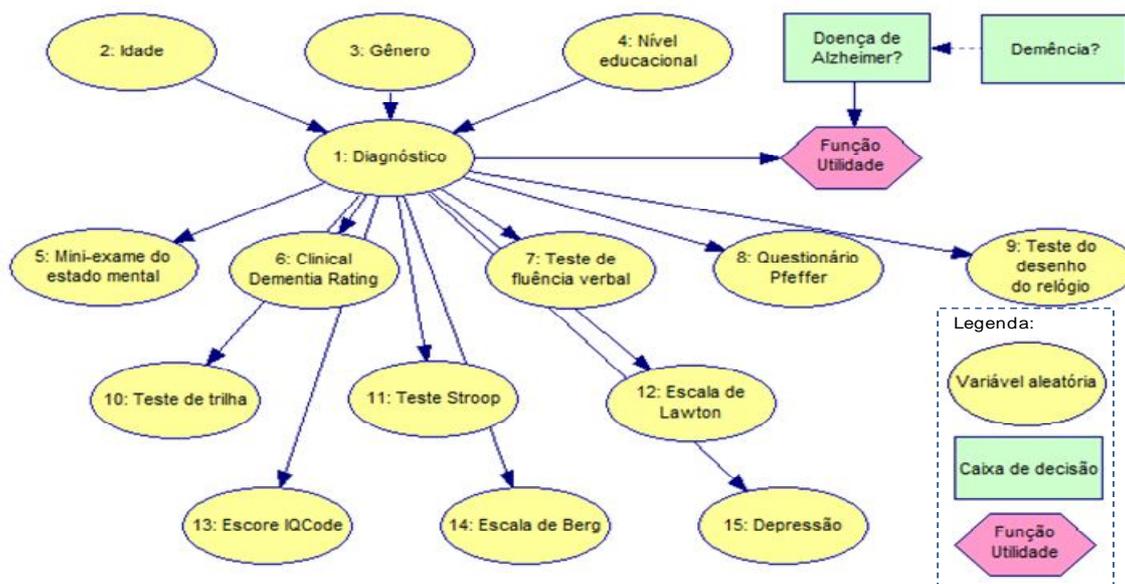


Figura 1. Rede Bayesiana e Diagrama de Influência, modelados para diagnóstico de Doença de Alzheimer.

4. Modelagem dos Conceitos Mapeados em uma Rede Bayesiana

A utilização de um modelo de dados interoperável é importante para a correta correspondência das observações do paciente com as variáveis aleatórias e níveis discretos da rede Bayesiana. Uma evidência da rede Bayesiana é gerada sempre que houver correspondência semântica entre esses dois domínios de dados (observações do paciente e variáveis aleatórias da rede Bayesiana). A interoperabilidade semântica é assegurada pela validação do CCD pela sua distribuição na web, ou através do repositório público de CCDs denominado *Healthcare Knowledge Component Repository* (HKCR).

A Tabela 1 apresenta a relação de conceitos presentes na rede Bayesiana para diagnóstico de DA. O modelo de dados clínicos foi especificado para expressar apenas a pontuação final de cada avaliação neuropsicológica, que foi utilizada na modelagem da rede Bayesiana.

Tabela 1. Relação de conceitos utilizados na rede Bayesiana para diagnóstico da Doença de Alzheimer.

Descrição da variável aleatória	Descrição da variável aleatória
1. Diagnóstico: indicação de Doença de Alzheimer.	9. Teste do desenho do relógio.
2. Faixa etária.	10. Teste de trilha: instrumento para avaliar características visomotoras.
3. Gênero: masculino ou feminino.	11. Teste Stroop desempenho de cor-palavra.
4. Nível educacional (em anos).	12. Escala de Lawton, usada como medida de qualidade de vida para demência.
5. Mini-exame do estado mental.	13. Escore IQCode: questionário do Informante sobre Declínio Cognitivo em Idosos.
6. Escala do <i>Clinical Dementia Rating</i> .	14. Escala de Berg para equilíbrio funcional.
7. Teste de fluência verbal.	15. Presença ou ausência de sintomas relacionados à depressão.
8. Questionário Pfeffer ou a escala Bayer de Atividades da Vida Diária.	

A Figura 2 mostra dois trechos de código XSD. O código da esquerda descreve o tipo complexo (*ComplexType*) que representa o *Clinical Dementia Rating* (CDR) no CCD “*Assessment Scales*”. O CDR é um teste neuropsicológico utilizado para avaliar o estágio de demência, e possui uma alta sensibilidade para o diagnóstico de DA [Morris, 1997]. O resultado é um número de 0 a 3, sendo: 0 (nenhuma alteração), 0,5 (demência questionável), 1 (demência leve), 2 (demência moderada) e 3 (demência grave). O código à direita da Figura 2 descreve o *ComplexType* que representa o *Mini-Mental State Examination* (MMSE) no CCD “*Assessment Scales*”. O MMSE é um teste neuropsicológico de avaliação cognitiva e instrumento de rastreamento, no qual o indivíduo recebe uma pontuação de 0 (pior escore) a 30 (melhor escore) [Folstein *et al.*, 1975].

Os resultados dos testes neuropsicológicos foram modelados em CCDs, cada variável sendo expressa em XSD como tipos complexos (*ComplexTypes*), para os quais foi atribuído o tipo de dados ordinal (*DvOrdinalType*) especificado no Modelo de Referência MLHIM. Este tipo foi utilizado porque as variáveis aleatórias associadas aos testes neuropsicológicos são de distribuição discreta. No CDR, cada item do tipo ordinal está associado a uma pontuação do teste (linhas 32 a 36 do código à esquerda da Figura 2). No caso do MMSE, o seu domínio de valores foi subdividido em três faixas numéricas consecutivas, representadas pela sintaxe *de – para*, e cada faixa associada a um nível discreto da variável aleatória (linhas 30 a 32 do código à direita da Figura 2). Os limites das faixas numéricas foram estimados através de um método de discretização supervisionada durante a modelagem da rede Bayesiana. O objetivo do método de discretização foi transformar atributos numéricos em tipos multinominais adequados à rede Bayesiana [Seixas, 2012]. Na linha 1, observa-se o conjunto de caracteres que identificam o conceito clínico, na forma de UUIDs atribuídos aos *ComplexTypes* (linha 1). Na linha 6, nota-se que os conceitos estão associados à terminologia médica de referência *Systematized Nomenclature of Medicine - Clinical Terms* (SNOMED-CT).

A rede Bayesiana foi modelada utilizando a ferramenta de autoria e aprendizagem inferência Bayesiana Genie/SMILE (*Structural Modeling, Inference, and Learning Engine*) da Universidade de Pittsburg [Druzdzet, 1999]. O diagrama de influência e a rede Bayesiana estão descritos em XDSL, também baseado em XML e suportado pela ferramenta. As variáveis aleatórias e os estados discretos correspondentes possuem uma identificação única que deve ser idêntica à do conceito descrito pelo CCD nos atributos *data-name* e *symbol*, respectivamente. Assim, no domínio do HIS (*Health Information System*), a instância do CCD (descrita em XML) é comparada com as variáveis aleatórias da rede Bayesiana (descrita em XDSL), no domínio do CDSS. Uma evidência é registrada quando é verificada alguma correspondência entre os dois domínios. A Figura 3 ilustra a correspondência entre os dois domínios de dados: os conceitos utilizados na rede Bayesiana e instâncias clínicas encaminhadas pelo HIS. Depois de registradas todas as evidências, a distribuição de probabilidades do nodo central é atualizada pelo mecanismo de inferência e a probabilidade de diagnóstico é encaminhada ao HIS.

...	...
01 <xs:complexType name="ct-7cb5d29a-cd6b-4508-8296-abe889d8baa3">	01 <xs:complexType name="ct-1b5fb55f-035c-4590-a1e2-4b271e9289ba">
02 <xs:complexContent>	02 <xs:complexContent>
03 <xs:restriction base="mlhim2:ElementType">	03 <xs:restriction base="mlhim2:ElementType">
04 <xs:sequence>	04 <xs:sequence>
05 <xs:element maxOccurs="unbounded" minOccurs="0" ref="mlhim2:links"/>	05 <xs:element maxOccurs="unbounded" minOccurs="0" ref="mlhim2:links"/>
06 <xs:element maxOccurs="1" minOccurs="1" name="annotation" type="xs:string" fixed="Clinical dementia rating scale" rdf:resource="https://uts.nlm.nih.gov/snomedct/Browser.html?conceptId=273367002"/>	06 <xs:element maxOccurs="1" minOccurs="1" name="annotation" type="xs:string" fixed="Mini mental state score" rdf:resource="https://uts.nlm.nih.gov/snomedct/Browser.html?conceptId=414750000"/>
07 <xs:element minOccurs="1" maxOccurs="1" ref="mlhim2:el-1e4b3a59-2f1d-485d-b15b-0e9c9f7397c0"/>	07 <xs:element minOccurs="1" maxOccurs="1" ref="mlhim2:el-82277514-3bdf-4579-95f0-0ba3ea912fe4"/>
08 </xs:sequence>	08 </xs:sequence>
09 </xs:restriction>	09 </xs:restriction>
10 </xs:complexContent>	10 </xs:complexContent>
11 </xs:complexType>	11 </xs:complexType>
12 <xs:element name="el-1e4b3a59-2f1d-485d-b15b-0e9c9f7397c0" substitutionGroup="mlhim2:Element-dv" type="mlhim2:ct-1e4b3a59-2f1d-485d-b15b-0e9c9f7397c0"/>	12 <xs:element name="el-82277514-3bdf-4579-95f0-0ba3ea912fe4" substitutionGroup="mlhim2:Element-dv" type="mlhim2:ct-82277514-3bdf-4579-95f0-0ba3ea912fe4"/>
13 <xs:complexType name="ct-1e4b3a59-2f1d-485d-b15b-0e9c9f7397c0">	13 <xs:complexType name="ct-82277514-3bdf-4579-95f0-0ba3ea912fe4">
14 <xs:complexContent>	14 <xs:complexContent>
15 <xs:restriction base="mlhim2:DvOrdinalType">	15 <xs:restriction base="mlhim2:DvOrdinalType">
16 <xs:sequence>	16 <xs:sequence>
17 <xs:element name="data-name" type="xs:string" maxOccurs="1" minOccurs="1" fixed="Clinical dementia rating scale"/>	17 <xs:element name="data-name" type="xs:string" maxOccurs="1" minOccurs="1" fixed="Mini mental state score" />
18 <xs:element maxOccurs="1" minOccurs="1" name="DvOrdinal-dv">	18 <xs:element maxOccurs="1" minOccurs="1" name="DvOrdinal-dv">
19 <xs:simpleType>	19 <xs:simpleType>
20 <xs:restriction base="xs:decimal">	20 <xs:restriction base="xs:decimal">
21 <xs:enumeration value="0"/>	21 <xs:enumeration value="1"/>
22 <xs:enumeration value="0.5"/>	22 <xs:enumeration value="2"/>
23 <xs:enumeration value="1"/>	23 <xs:enumeration value="3"/>
24 <xs:enumeration value="2"/>	</xs:restriction>
25 <xs:enumeration value="3"/>	</xs:simpleType>
26 </xs:restriction>	26 </xs:element>
27 </xs:simpleType>	27 <xs:element maxOccurs="1" minOccurs="1" name="symbol">
28 </xs:element>	28 <xs:simpleType>
29 <xs:element maxOccurs="1" minOccurs="1" name="symbol">	29 <xs:restriction base="xs:string">
30 <xs:simpleType>	30 <xs:enumeration value="0-17"/>
31 <xs:restriction base="xs:string">	31 <xs:enumeration value="18-26"/>
32 <xs:enumeration value="0-normal control"/>	32 <xs:enumeration value="26-30"/>
33 <xs:enumeration value="0.5-very mild"/>	</xs:restriction>
34 <xs:enumeration value="1-mild"/>	</xs:simpleType>
35 <xs:enumeration value="2-moderate"/>	37 </xs:element>
36 <xs:enumeration value="3-severe"/>	38 </xs:sequence>
37 </xs:restriction>	39 </xs:restriction>
38 </xs:simpleType>	40 </xs:complexContent>
39 </xs:element>	41 </xs:complexType>
40 </xs:sequence>	...
41 </xs:restriction>	
42 </xs:complexContent>	
43 </xs:complexType>	
...	

Figura 2. Trecho do CCD utilizado para especificação dos exames neuropsicológicos CDR (código à esquerda) e MMSE (código à direita).

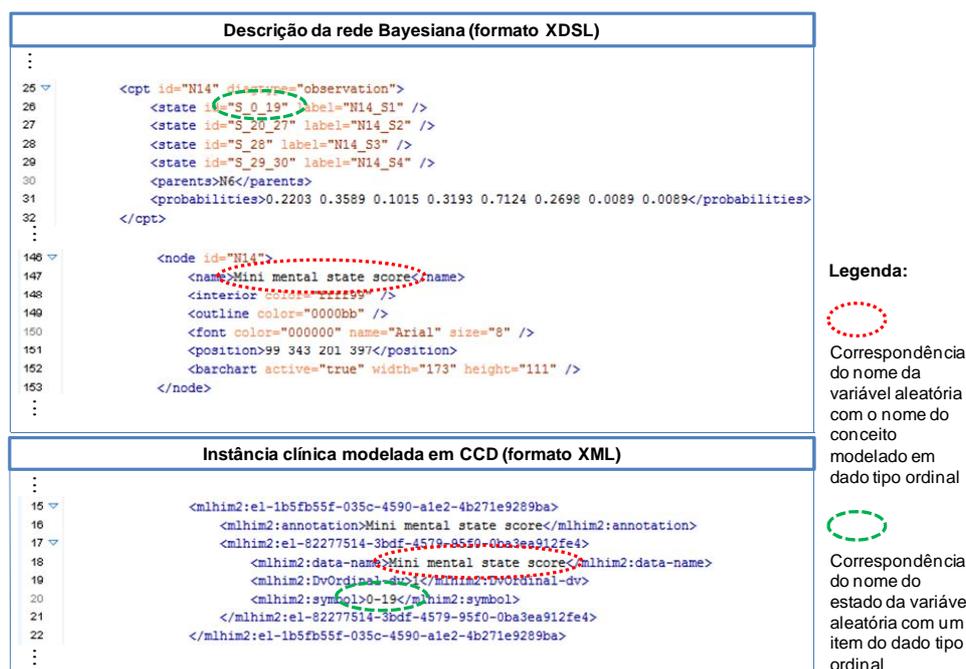


Figura 3. Correspondência semântica da descrição da rede Bayesiana em XDSL para uma instância clínica modelada em CCD.

5. Conclusão

Este artigo mostrou a modelagem de conceitos clínicos utilizados em um sistema de apoio à decisão para diagnóstico de Demência, DA e TCL. Foram apresentados os conceitos clínicos e a rede Bayesiana, modelados para diagnóstico de DA.

Com a disponibilização de um conjunto mais completo de casos clínicos, ou a necessidade de inclusão ou exclusão de algum critério no diagnóstico da doença de interesse, será necessário reexecutar a rotina de aprendizagem da rede Bayesiana. É possível que o método de discretização estime outros pontos de corte para os intervalos numéricos. Se for o caso, será necessário publicar um novo CCD com os novos pontos de corte. Assim, para cada atualização ou revisão da rede Bayesiana, haverá um CCD. Este ponto não representa um problema, pois, conforme mencionado anteriormente, é possível termos mais de um CCD para um mesmo conceito, mas com UUID distintos.

Como trabalho futuro, pretende-se estudar uma forma de representação da rede Bayesiana de acordo com os princípios de modelagem multinível. Assim, não seria necessário descrever a rede Bayesiana em XDSL, formato especificado pela ferramenta Genie/SMILE, dado que seria possível especificá-la diretamente em XML Schemas válidos de acordo com as especificações MLHIM. Com isso, seria possível descrever o conceito uma única vez no domínio do HIS, não sendo necessário replicá-lo no domínio do CDSS.

Referências

- Atalag, K. and H. Y. Yang (2010). From openEHR Domain Models to Advanced User Interfaces: A Case Study in Endoscopy. Health Informatics Conference, New Zealand
- Berner, E. S. (2007). Clinical decision support systems: theory and practice. Birmingham, England, Springer.

- Cavalini, L. T. and T. Cook (2011). "Health Informatics: The Relevance of Open Source and Multilevel Modeling." *Open Source Systems: Grounding Research* 365: 338-347.
- Cavalini, L. T. and T. W. Cook (2012). Knowledge engineering of healthcare applications based on minimalist multilevel models. *e-Health Networking, Applications and Services (Healthcom), 2012 IEEE 14th International Conference on, IEEE*.
- Chawla, N. V., K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer (2002). "SMOTE: Synthetic Minority Over-Sampling Technique." *Journal of Artificial Intelligence Research* 16: 321-357.
- Dempster, A. P., N. M. Laird and D. B. Rubin (1977). "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1): 1-38.
- Dias, R. D., T. W. Cook and S. M. Freire (2011). "Modeling healthcare authorization and claim submissions using the openEHR dual-model approach." *BMC Medical Informatics & Decision Making* 11(60): 60.
- Dolin, R. H., L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron and A. S. Shvo (2006). "HL7 clinical document architecture, release 2." *Journal of the American Medical Informatics Association* 13(1): 30-39.
- Druzdzel, M. J. (1999). *SMILE: Structural Modeling, Inference, and Learning Engine and GeNIe: a development environment for graphical decision-theoretic models*, John Wiley & Sons Ltd.
- Folstein, M. F., S. E. Folstein and P. R. McHugh (1975). "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician." *J Psychiatr Res* 12(3): 189-198.
- Friedman, N., M. Linial, I. Nachman and D. Pe'er (2000). "Using Bayesian Networks to Analyze Expression Data." *Journal of Computational Biology* 7(3-4): 601-620.
- Garde, S., P. Knaup, E. Hovenga and S. Heard (2007). "Towards semantic interoperability for electronic health records." *Methods of information in medicine* 46(3): 332-343.
- Gomes, A. T. A., A. Ziviani, B. S. P. M. Correa, I. M. Teixeira and V. M. Moreira (2012). *SPLiCE: a software product line for healthcare*. Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium. Miami, Florida, USA, ACM: 721-726.
- Haynes, R. B. and N. L. Wilczynski (2010). "Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: methods of a decision-maker-researcher partnership systematic review." *Implement Sci* 5(12): 12.
- Howard, R. A. and J. E. Matheson (2005). "Influence diagrams." *Decision Analysis* 2(3): 127-143.
- Jalbert, J. J., L. A. Daiello and K. L. Lapane (2008). "Dementia of the Alzheimer Type." *Epidemiologic Reviews* 30(1): 15-34.
- Kalra, D., T. Beale and S. Heard (2005). "The OpenEHR Foundation." *Studies in Health Technology and Informatics* 115: 153-173.

- Luzardo, A. R., M. Gorini and A. Silva (2006). "Características de idosos com doença de Alzheimer e seus cuidadores: uma série de casos em um serviço de neurogeriatria." *Texto & Contexto Enfermagem* 15(4): 587-594.
- Miller, R. A. (2009). "Computer-Assisted Diagnostic Decision Support: history, challenges, and possible paths forward." *Advances in Health Sciences Education* 14: 89-106.
- Morris, J. C. (1997). "Clinical Dementia Rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type." *International Psychogeriatrics* 9(S1): 173-176.
- Petersen, R. C. (2003). *Mild cognitive impairment: aging to Alzheimer's disease*, Oxford University Press, USA.
- Pradhan, M., G. Provan, B. Middleton and M. Henrion (1994). *Knowledge engineering for large belief networks. Uncertainty in Artificial Intelligence*, San Francisco, Proceedings of the Seventh Conference.
- Ratamero, E. M., T. Cook, L. T. Cavalini and D. b. C. M. Saade (2010). Modeling magnetic resonance clinical reports of the lumbar spine using openEHR archetypes. XII Congresso Brasileiro de Informática em Saúde, Porto de Galinhas.
- Santos, M. R., M. P. Bax and D. Kalra (2010). "Building a Logical EHR Architecture based on ISO 13606 Standard and Semantic Web Technologies." *Studies in Health Technology and Informatics* 160(Pt 1): 161.
- Seixas, F. L. (2012). *Sistema de Apoio à Decisão aplicado ao Diagnóstico de Demência, Doença de Alzheimer e Transtorno Cognitivo Leve*. Ciências de Computação. Niterói, Universidade Federal Fluminense (UFF). Doutorado em Computação: 155.
- Shalom, E., Y. Shahr, M. Taieb-Maimon, S. B. Martins, L. T. Vaszar, M. K. Goldstein, L. Gutnik and E. Lunenfeld (2009). "Ability of expert physicians to structure clinical guidelines: reality versus perception." *Journal of Evaluation in Clinical Practice* 15(6): 1043-1053.
- Sokolowski, R. and J. Dudeck (1999). *XML and its Impact on Content and Structure in Electronic Health Care Documents*. AMIA Annual Symposium, Washington, DC, The United States of America, American Medical Informatics Association.
- Sundvall, E. (2013). *Scalability and semantic sustainability in electronic health record systems*. Department of Biomedical Engineering. Linköping, Linköping University. Doctor: 95.
- Witten, I. H., E. Frank and M. A. Hall (2011). *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann.
- Zusman, E. E. (2010). "Form facilitates function: innovations in architecture and design drive quality and efficiency in healthcare." *Neurosurgery* 66(6): N24.
- Zywietz, C. (2004). "Communication and interoperability for serial comparison in continuous health care--the new challenges." *Studies in Health Technology and Informatics* 108: 172-180.