
**Uma Abordagem Híbrida Baseada em Redes
Bayesianas e Ontologias para Modelagem do
Estudante em Sistemas Adaptativos e
Inteligentes para Educação**

Hiran Nonato Macedo Ferreira



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Hiran Nonato Macedo Ferreira

**Uma Abordagem Híbrida Baseada em Redes
Bayesianas e Ontologias para Modelagem do
Estudante em Sistemas Adaptativos e
Inteligentes para Educação**

Tese de doutorado apresentada ao Programa de
Pós-graduação da Faculdade de Computação
da Universidade Federal de Uberlândia como
parte dos requisitos para a obtenção do título
de Doutor em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Renan Gonçalves Cattelan

Uberlândia
2018

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

F383a
2018

Ferreira, Hiran Nonato Macedo, 1988-

Uma abordagem híbrida baseada em redes bayesianas e ontologias para modelagem do estudante em sistemas adaptativos e inteligentes para educação / Hiran Nonato Macedo Ferreira. - 2018.

168 f. : il.

Orientador: Renan Gonçalves Cattelan.

Tese (Doutorado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Ciência da Computação.

Disponível em: <http://dx.doi.org/10.14393/ufu.te.2018.786>

Inclui bibliografia.

1. Computação - Teses. 2. Estudantes - Avaliação - Teses. 3. Testes e medidas educacionais - Teses. 4. Ensino auxiliado por computador - Teses. I. Cattelan, Renan Gonçalves. II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Ciência da Computação. III. Título.

CDU: 621.3

Maria Salete de Freitas Pinheiro – CRB6/1262

*Dedico este trabalho a toda a minha família,
em especial, a meus pais, Juracy e Dalva.*

Agradecimentos

Agradeço ...

Primeiramente a Deus, pelo dom da vida, por estar sempre ao meu lado, principalmente nos momentos mais difíceis, e por permitir a realização de mais este sonho.

Ao meu orientador, Prof. Dr. Renan Gonçalves Cattelan, pela oportunidade, conselhos, apoio, confiança e, acima de tudo, pelos valiosos ensinamentos transmitidos durante essa etapa.

A minha esposa, Mônica, pelo seu amor e companheirismo durante todo esse trabalho. Agradeço também pela paciência, compreensão e por estar do meu lado, sempre me motivando a seguir em frente.

A Meus pais, que mesmo distantes fisicamente, sempre participaram de toda a minha formação. Obrigado pelos ensinamentos, conselhos, orações, e por me ensinar os valores essenciais da vida.

A Minha família, em especial minhas irmãs Wilma e Isis, e meus sobrinhos Iago e Pedro Augusto. Muito obrigado pelo apoio e torcida.

Aos colegas de laboratório pela parceria e por todas as conversas, discussões e, acima de tudo, a troca de experiências.

Aos professores e funcionários da FACOM/UFU. Especialmente, aos professores Fabiano de Azevedo Dorça e Márcia Aparecida Fernandes pelos valiosos ensinamentos e recomendações durante a realização desta pesquisa.

Ao Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais (IF-SULDEMINAS) pelo apoio e pela liberação para realização deste trabalho.

Às agências de pesquisa CNPq, CAPES, FAPEMIG, PROAP/UFU e PROPP/UFU pelo apoio financeiro. Agradeço também ao PPGCO/FACOM/UFU pelo estrutura necessária para realização desta pesquisa.

Aos professores Edson Pinheiro Pimentel, Renato de Freitas Bulcão Neto, Carlos Roberto Lopes e Márcia Aparecida Fernandes pelas valiosas contribuições sugeridas na avaliação deste trabalho.

Enfim, a todos que contribuíram de alguma maneira para a realização deste trabalho.

*“Comece fazendo o que é necessário, depois o que é possível,
e de repente você estará fazendo o impossível.”*
(São Francisco de Assis)

UNIVERSIDADE FEDERAL DE UBERLÂNDIA – UFU
FACULDADE DE COMPUTAÇÃO – FACOM
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO – PPGCO

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação da tese intitulada **“Uma Abordagem Híbrida Baseada em Redes Bayesianas e Ontologias para Modelagem do Estudante em Sistemas Adaptativos e Inteligentes para Educação”** por **Hiran Nonato Macedo Ferreira** como parte dos requisitos exigidos para a obtenção do título de **Doutor em Ciência da Computação**.

Uberlândia, 18 de Junho de 2018.

Orientador: _____

Prof. Dr. Renan Gonçalves Cattelan
Universidade Federal de Uberlândia

Banca Examinadora:

Prof. Dr. Edson Pinheiro Pimentel
Universidade Federal do ABC

Prof. Dr. Renato de Freitas Bulcão Neto
Universidade Federal de Goiás

Prof. Dr. Carlos Roberto Lopes
Universidade Federal de Uberlândia

Prof^a. Dr^a. Márcia Aparecida Fernandes
Universidade Federal de Uberlândia

Resumo

Novas tecnologias surgem diariamente com o intuito de auxiliar na criação de ferramentas cada vez mais adaptáveis às particularidades de cada usuário. Quando associadas a ambientes educacionais, tais adaptações permitem melhorar o processo de ensino. Um dos recursos computacionais que busca melhorar esses processos são os Sistemas Adaptativos e Inteligentes para Educação (SAIEs), que possuem a capacidade de alterar suas características para atender as necessidades dos usuários. Esses sistemas utilizam Modelos de Estudantes, que são capazes de monitorar características importantes do estudante e fazer ajustes apropriados a fim de apoiar e aprimorar a aprendizagem. Alguns trabalhos disponíveis na literatura propõem modelos capazes de avaliar o desempenho do estudante. No entanto, questões importantes como conhecimento e comportamento raramente são analisadas conjuntamente. Dessa forma, este trabalho apresenta a definição de uma abordagem híbrida para modelagem de estudante capaz de identificar o seu desempenho em um SAIE. A abordagem proposta baseia-se em informações sobre o nível de conhecimento e comportamento dos estudantes para definir o seu desempenho. Para isso, a abordagem combina a utilização de Redes Bayesianas e ontologias, com o objetivo de criar um modelo que seja dinâmico, probabilístico, independente de domínio, extensível e reutilizável. É também proposta uma extensão para tornar o modelo um Modelo Aberto do Estudante, o qual permite a visualização das principais capacidades e limitações dos envolvidos no processo de aprendizagem. Foi realizado, ainda, um estudo de caso por meio da implementação e integração da proposta a um ambiente real de ensino. A fim de avaliar a eficácia do modelo, foram conduzidos quatro experimentos com estudantes matriculados em cursos presenciais, para verificar a correlação entre os valores inferidos pelo modelo e o conhecimento do estudante. Também foram verificadas questões comportamentais, por meio de testes estatísticos, sobre os dados obtidos dos experimentos. Os resultados mostraram que o modelo proposto é capaz de determinar um nível de desempenho que seja correspondente às capacidades e limitações dos estudantes. Sendo assim, a abordagem proposta mostrou-se eficaz e adequada para a modelagem do estudante em SAIEs.

Palavras-chave: Sistemas Adaptativos e Inteligentes para Educação; Modelagem de Estudante; Avaliação de Desempenho; Ontologias; Redes Bayesianas.

Abstract

New technologies emerge everyday with the aim of supporting the development of tools that are more adaptable to the particular needs of each user. When these technologies are associated with the educational environment, such adaptations allow for the improvement of the teaching process. One of the computational resources that strives to improve these processes are Adaptive and Intelligent Educational Systems (AIESs), which possess the capacity to alter its characteristics in order to attend to the users' needs. These systems use Student Models that are capable of monitoring important student characteristics and make appropriate adjustments, which are directed toward supporting and improving learning. Studies available in the literature propose models capable of evaluating the student's performance. However, important questions such as knowledge and behavior are rarely analyzed together. In this manner, this study presents the definition of a hybrid approach for student modeling, which is capable of identifying their performance on an AIES. The proposed approach is based on information concerning the level of knowledge and behavior of the students for defining their real cognitive state. In order to realize this task, the approach combines the use of Bayesian Networks and ontologies, with the aim of creating a model that is dynamic, probabilistic, independent of domain, extensible and reusable. It is also proposed an extension to make the model an Open Student Model, which allows for the visualization of the main capacities and limitations of those involved in the learning process. In addition, a case study was performed by means of the implementation and integration of the proposal into a real teaching environment. Under the intent of evaluating the efficiency of the model, four experiments were carried out with students enrolled in face-to-face courses, to verify the correlation between the values inferred by the model and the student's knowledge. Behavioral issues were also verified by means of statistical tests on the data obtained from the experiments. Results showed that the proposed model is sufficient to determining a level of performance that corresponds to the skills and limitations of the students. Therefore, the proposed approach proved to be effective and adequate for modeling the students in AIESs.

Keywords: Adaptive and Intelligent Educational Systems; Student Modeling; Performance Evaluation; Ontology; Bayesian Network.

Lista de ilustrações

Figura 1 – Exemplo de nós (<i>k</i>) e (<i>e</i>) em uma RB.	48
Figura 2 – Probabilidades condicionais associadas a uma RB.	49
Figura 3 – Exemplos de diferentes disponibilizações dos instrumentos avaliativos. .	51
Figura 4 – Exemplo de instâncias representando os níveis de aprendizado.	58
Figura 5 – Exemplos de instâncias (estudantes e níveis).	59
Figura 6 – Três formas diferentes de visualização de desempenho: (a) Gráfico de Linha; (b) Gráfico de Barra; (c) Gráfico de Radar	65
Figura 7 – Fluxo de interações entre professores, estudantes e modelo	66
Figura 8 – Arquitetura do modelo.	68
Figura 9 – Incorporação do modelo proposto na arquitetura dos AEUs.	73
Figura 10 – Salas de aula instrumentadas utilizadas pela plataforma CX.	74
Figura 11 – Criação de um instrumento avaliativo na plataforma CX.	76
Figura 12 – Interações entre os módulos desenvolvidos no modelo proposto.	78
Figura 13 – Interface para visualização disponível no CX: (a) Visualização do estudante, (b) Visualização do professor.	79
Figura 14 – Arquitetura para integração do modelo proposto a um ambiente educacional genérico.	80
Figura 15 – Gráficos de dispersão para as variáveis: (a) <i>Teste de Usabilidade</i> – T3, (b) <i>Análise de Tarefas</i> – T2, (c) <i>MIPS</i> – T5, and (d) <i>RISC</i> – T5.	87
Figura 16 – Gráfico de radar para dois estudantes em T5.	88
Figura 17 – RB utilizada no experimento 2 (T1).	90
Figura 18 – RB utilizada no experimento 2 (T2).	91
Figura 19 – RB utilizada no experimento 2 (T3).	91
Figura 20 – RB utilizada no experimento 2 (T4).	92
Figura 21 – RB utilizada no experimento 2 (T5).	92
Figura 22 – Trecho de uma RB com componentes mais conexos.	94
Figura 23 – Trecho de uma RB com diferentes instrumentos avaliativos.	94

Figura 24 – Trecho de uma RB com poucos nós e instrumentos avaliativos distri- buídos.	95
Figura 25 – Trecho de uma RB com nós sem evidências.	95
Figura 26 – Trecho de uma RB com instrumentos avaliativos do tipo 1...1	95
Figura 27 – Trecho de uma RB com nós folhas.	96
Figura 28 – Gráficos de dispersão para a distribuição dos clusters	101
Figura 29 – Comparação entre os três <i>clusters</i>	103
Figura 30 – Grau de concordância sobre a percepção de utilidade.	107
Figura 31 – Grau de concordância sobre a alteração no comportamento.	109
Figura 32 – Respostas à questão três.	109
Figura 33 – Respostas às questões quatro e cinco.	110
Figura 34 – Grau de concordância sobre a questão seis.	111
Figura 35 – RB utilizada no Experimento 1 (T1)	148
Figura 36 – RB utilizada no Experimento 1 (T2)	149
Figura 37 – RB utilizada no Experimento 1 (T3)	150
Figura 38 – RB utilizada no Experimento 1 (T4)	151
Figura 39 – RB utilizada no Experimento 1 (T5)	152
Figura 40 – RB utilizada no Experimento 2 - IHC 2016-1 (Execuções 1 e 2)	156
Figura 41 – RB utilizada no Experimento 2 - IHC 2016-1 (Execuções 3 e 4)	157
Figura 42 – RB utilizada no Experimento 2 - AOC 2016-2 (Execuções 1 e 2)	158
Figura 43 – RB utilizada no Experimento 2 - AOC 2016-2 (Execuções 3 e 4)	159
Figura 44 – RB utilizada no Experimento 2 - IHC 2016-2 (Execuções 1 e 2)	160
Figura 45 – RB utilizada no Experimento 2 - IHC 2016-2 (Execuções 3 e 4)	161
Figura 46 – RB utilizada no Experimento 2 - AOC 2017-1 (Execuções 1 e 2)	162
Figura 47 – RB utilizada no Experimento 2 - AOC 2017-1 (Execuções 3 e 4)	163
Figura 48 – RB utilizada no Experimento 2 - IHC 2017-1 (Execuções 1 e 2)	164
Figura 49 – RB utilizada no Experimento 2 - IHC 2017-1 (Execuções 3 e 4)	165

Lista de tabelas

Tabela 1 – Trabalhos correlatos	43
Tabela 2 – Principais classes representadas na ontologia.	57
Tabela 3 – Propriedades das classes <i>Personal Information</i> , <i>Gamification</i> e <i>Trace</i> . .	58
Tabela 4 – Turmas participantes do Experimento 1.	85
Tabela 5 – Coeficiente de correlação de postos de <i>Spearman</i> entre notas reais dos estudantes e as notas inferidas pela RB.	86
Tabela 6 – Turmas participantes do Experimento 2.	89
Tabela 7 – Turmas participantes do Experimento 3.	98
Tabela 8 – Total de interações por turma analisada.	99
Tabela 9 – Distribuição média das interações em cada <i>cluster</i>	100
Tabela 10 – Médias de desempenhos e resultado do teste de <i>Mann-Whitney</i> em cada <i>cluster</i> na primeira estratégia.	102
Tabela 11 – Médias de desempenhos e resultado do teste de <i>Kruskal-Wallis</i> em cada <i>cluster</i> na segunda estratégia.	102
Tabela 12 – Resultados de testes de comparação par a par entre os <i>clusters</i>	103
Tabela 13 – Turmas participantes do Experimento 4.	105
Tabela 14 – Resultados sobre a percepção de utilidade.	107
Tabela 15 – Resultados agregados.	108
Tabela 16 – Resultados sobre a alteração no comportamento.	108
Tabela 17 – Alfa de Cronbach para as respostas ao questionário.	111
Tabela 18 – Alfa de Cronbach separado por turma.	112
Tabela 19 – Resultado da análise de correlação entre os valores reais e os valores inferidos sobre o nível de conhecimento dos estudantes (Parte 1).	154
Tabela 20 – Resultado da análise de correlação entre os valores reais e os valores inferidos sobre o nível de conhecimento dos estudantes (Parte 2).	155

Lista de siglas

AEU Ambiente Educacional Ubíquo

BKT *Bayesian Knowledge Tracing*

C&A Captura e Acesso

CX Classroom eXperience

EA Estilo de Aprendizagem

FSLSM *Felder-Silverman Learning Style Model*

ME Modelo de Estudante

MAE Modelo Aberto do Estudante

OA Objeto de Aprendizagem

RB Rede Bayesiana

RBD Rede Bayesiana Dinâmica

SAIE Sistema Adaptativo e Inteligente para Educação

STI Sistema Tutor Inteligente

SWRL *Semantic Web Rule Language*

Sumário

1	INTRODUÇÃO	25
1.1	Motivação	26
1.2	Questões de Pesquisa	28
1.3	Objetivos	29
1.4	Organização desta Tese	30
2	REVISÃO DE LITERATURA	31
2.1	Considerações Iniciais	31
2.2	Modelagem do Estudante	31
2.3	Redes Bayesianas	34
2.4	Ontologias	37
2.5	Modelo Aberto do Estudante	40
2.6	Síntese dos Trabalhos Correlatos	41
2.7	Considerações Finais	43
3	ABORDAGEM PROPOSTA	45
3.1	Considerações Iniciais	45
3.2	Avaliação do Nível de Aquisição de Conhecimento do Estudante	46
3.2.1	Independência de Domínio	47
3.2.2	Representação dos Conceitos	48
3.2.3	Definição das Probabilidades	49
3.2.4	Definição dos Instrumentos Avaliativos	50
3.2.5	Definição Formal	51
3.2.6	Inferência	52
3.3	Avaliação de Desempenho do Estudante	53
3.3.1	Ontologia	54
3.3.2	Decomposição Modular da Ontologia	58
3.3.3	Regras para Inferência de Conhecimento	61

3.4	Modelo Aberto do Estudante	62
3.5	Funcionamento e Arquitetura do Modelo	64
3.6	Considerações Finais	68
4	ESTUDO DE CASO - INTEGRAÇÃO DA PROPOSTA A UM AMBIENTE REAL DE ENSINO	71
4.1	Considerações Iniciais	71
4.2	Ambiente de Integração	72
4.3	Integração do Modelo ao Ambiente	74
4.3.1	1 ^a Fase - Módulo Probabilístico	75
4.3.2	2 ^a Fase - Módulo Semântico	77
4.3.3	3 ^a Fase - Módulo de Visualização	78
4.4	Modelo Genérico de Integração	80
4.5	Considerações Finais	81
5	EXPERIMENTOS E ANÁLISE DOS RESULTADOS	83
5.1	Considerações Iniciais	83
5.2	Experimento 1: Análise dos Valores Inferidos pelo Módulo Probabilístico	84
5.2.1	Método Utilizado	84
5.2.2	Resultados e Discussão	86
5.3	Experimento 2: Análise da Estrutura da Rede Bayesiana	88
5.3.1	Método Utilizado	89
5.3.2	Resultados e Discussão	93
5.4	Experimento 3: Análise do Desempenho dos Estudantes com base em Perfis de Interação	96
5.4.1	Método Utilizado	97
5.4.2	Resultados e Discussão	100
5.5	Experimento 4: Análise sobre o Impacto na Utilização do Modelo Aberto do Estudante (MAE)	104
5.5.1	Método Utilizado	105
5.5.2	Resultados e Discussão	106
5.6	Considerações Finais	114
6	CONCLUSÕES	117
6.1	Contribuições	118
6.2	Produção Bibliográfica	120
6.3	Limitações	122
6.4	Trabalhos Futuros	123
	REFERÊNCIAS	127

APÊNDICES

141

APÊNDICE A	—	EXEMPLOS DE CATEGORIAS E REGRAS PARA INFERÊNCIA DO NÍVEL DE DESEMPENHO	143
APÊNDICE B	—	REDES BAYESIANAS UTILIZADAS NO EX- PERIMENTO 1	147
APÊNDICE C	—	REDES BAYESIANAS UTILIZADAS NO EX- PERIMENTO 2	153
APÊNDICE D	—	QUESTIONÁRIOS UTILIZADOS NO EXPE- RIMENTO 4	167

CAPÍTULO 1

Introdução

O avanço da tecnologia tem trazido diversos recursos que visam melhorar a forma de comunicação entre os computadores e os usuários. Em particular, nos últimos anos, muito se discute sobre mecanismos que facilitam a integração de recursos inteligentes, a fim de adaptar o conteúdo mais adequado às características e expectativas de cada usuário. Essa personalização individualizada de conteúdo tem sido um tema recorrente de estudo e atraído a atenção de muitos pesquisadores nos últimos anos (KLAŠNJA-MILIĆEVIĆ et al., 2017; BEEŁ et al., 2016).

Adaptabilidade é um dos principais requisitos nos ambientes computacionais da atualidade. De acordo com Barra, Negro e Scarano (1999), a adaptabilidade visa identificar o comportamento e o nível de conhecimento do usuário, a fim de prover o material que se deseja, na forma que se espera. Novas tecnologias surgem diariamente com o intuito de auxiliar na criação de ferramentas cada vez mais adaptáveis às particularidades dos usuários. Quando associadas a ambientes educacionais, tais adaptações permitem melhorar a dinâmica e, consequentemente, o processo de ensino/aprendizagem. Esse conceito está relacionado diretamente com a proposta dos Sistemas Adaptativos e Inteligentes para Educação (SAIEs).

Os SAIEs possuem a capacidade de alterar suas características para atender às necessidades dos usuários. Esses ambientes monitoram características importantes do estudante e fazem ajustes apropriados ao ambiente para apoiar e aprimorar a aprendizagem (DURLACH, 2011). O objetivo desses sistemas é criar um ambiente sólido e flexível que auxilie na melhoria do desempenho e aprendizado dos estudantes.

Para que essa adaptação ocorra de forma eficiente, o sistema deve possuir a capacidade de identificar algumas características importantes dos estudantes. Uma das principais características a ser tratada é a capacidade intelectual e cognitiva do estudante ante o conteúdo que está sendo estudado (WENGER, 1987; LAMB et al., 2014). A chave para esses ambientes é conseguir identificar o nível de conhecimento do estudante, para, a partir daí, ajustar suas características ao perfil dos aprendizes.

1.1 Motivação

Predizer o nível de conhecimento do estudante não é uma tarefa simples (RAMÍREZ-NORIEGA; JUÁREZ-RAMÍREZ; MARTÍNEZ-RAMÍREZ, 2017), até mesmo na educação presencial a identificação do nível de conhecimento é um dos principais desafios enfrentados pelos professores. Essa tarefa se torna ainda mais difícil em um ambiente virtual. Por essa razão, alguns autores propõem o uso de sistemas especialistas para satisfazer tal necessidade (MOURI et al., 2016; SEFFRIN et al., 2016).

Diversos trabalhos têm apresentado abordagens para a identificação do nível de conhecimento dos estudantes enquanto utilizam um sistema educacional. Grande parte desses trabalhos se baseia em técnicas de Inteligência Artificial (TORABI; MORADI; KHANTAIMOORI, 2012; GRUBIŠIĆ; STANKOV; PERAIĆ, 2013) para predizer e inferir o estado atual de conhecimento, outras, por sua vez, recorrem a teorias probabilísticas (MOURI et al., 2016; SEFFRIN et al., 2016) com o intuito de tratar incertezas nesse processo.

Um dos componentes fundamentais na identificação do conhecimento do estudante e na personalização do processo de ensino/aprendizagem em SAIEs é o Modelo de Estudante (ME). De acordo com Li et al. (2011), a modelagem do estudante é um dos fatores-chave que afetam sistemas automatizados de tutoria na tomada de decisões. Essa modelagem pode ser definida como o processo de coleta de informações importantes para inferir o atual estado cognitivo do estudante (MITROVIC; THOMSON, 2009). Um ME deve representar não somente informações sobre o nível de conhecimento do estudante, mas também refletir, o mais fiel possível, o seu processo de raciocínio (CLEMENTE; RAMÍREZ; ANTONIO, 2011).

Segundo Self (1990), a modelagem de estudante é uma tarefa que envolve a representação de diversas questões cognitivas, como: análise do desempenho do estudante, identificação de equívocos, representação de objetivos e planejamentos, observação do conhecimento a priori, descrição de características pessoais, dentre outras. O ME oferece descrições do que realmente é relevante sobre o comportamento, conhecimento e aptidões dos estudantes. Baseado no ME, um SAIE ajusta o seu comportamento a fim de fornecer conteúdo adequado às reais necessidades dos estudantes¹.

Além da modelagem de estudante, técnicas computacionais têm surgido com o intuito de auxiliar neste processo de análise de dados sobre as características dos estudantes.

¹ No decorrer deste trabalho, diversos termos que remetem à análise de conhecimento do estudante serão apresentados. Tendo isso em vista, para o contexto desta monografia, os termos conhecimento, análise de conhecimento ou nível de aquisição de conhecimento se referem a uma medida aferida, ou métrica, de conhecimento instantâneo, medida naquele determinado instante (momento), utilizando, principalmente, algum instrumento avaliativo para medição, como: quizzes, questionários, questões de múltipla escolha, entre outros. Já o termo análise de desempenho corresponde a uma medida mais global, que envolve não somente instrumentos avaliativos, mas também questões cognitivas e comportamentais do estudante. Ainda no contexto deste trabalho, entende-se que desempenho é uma medida composta pelo conhecimento mais o comportamento do estudante.

Uma que vem ganhando destaque nas últimas décadas é a *Learning Analytics* (SIEMENS; BAKER, 2012). Definida como a medição, coleta e análise de dados sobre os estudantes em ambientes educacionais, possui o objetivo de realizar uma interpretação individualizada do percurso, das dificuldades e das características de estudos dos estudantes (SIEMENS; BAKER, 2012).

A técnica de análise de dados por meio da *Learning Analytics* vem ganhando força e consistência, permitindo que novas abordagens para interpretação dessas informações sejam sugeridas. Uma importante faceta que explora essa técnica foi introduzida por Shum e Crick (2012) e é conhecida como *Dispositional Learning Analytics*. Essa estratégia combina os dados gerais de aprendizagem com elementos “dispositionais” próprios dos estudantes, incluindo seus comportamentos, suas atitudes e seus valores (TEMPE-LAAR et al., 2018). A *Dispositional Learning Analytics* vai além da análise de dados sobre o estudante, pois combina também facetas comportamentais e cognitivas dos indivíduos (RIENTIES; CROSS; ZDRAHAL, 2016).

Pesquisas relacionadas com ME têm voltado o foco de estudo para a forma como os dados tratados pelo ME serão interpretados pelos usuários envolvidos nesse processo. Essas pesquisas propõem uma maneira diferente de visualização das informações tratadas pelo modelo, pelo qual estudantes e professores podem interagir diretamente com os valores inferidos para o conhecimento durante o processo de aprendizagem. Esse modelo de apresentação de conteúdo é conhecido como Modelo Aberto do Estudante (MAE)².

O MAE é um importante tipo de interface educacional que surgiu na área de sistemas de aprendizagem personalizada (BULL; KAY, 2007). Enquanto nos sistemas tradicionais, os MEs, que representam o estado atual de conhecimento, estão escondidos dos estudantes e são utilizados apenas para personalizar o processo educacional, no modelo aberto de aprendizado existe a capacidade de visualizar, explorar e até mesmo modificar o estado de conhecimento do próprio estudante. Utilizar um MAE significa tornar o modelo explícito ao estudante, no qual ele tem a capacidade de visualizar o seu nível de conhecimento, permitindo autoconhecimento e possibilitando autorregulação do processo de aprendizagem (GUERRA et al., 2016).

O MAE explora a área de Visualização de Informação e utiliza uma variedade de visualizações, como medidores de habilidades (*skillmeters*) (MITROVIC; MARTIN, 2007; MARTINEZ-MALDONADO et al., 2015) e mapas de conceitos (*concept maps*) (LOBODA et al., 2014; GUERRA et al., 2016) para facilitar e melhorar o processo de aprendizagem. Long e Aleven (2017) apontaram que os MAEs têm grande capacidade para melhorar o processo de ensino/aprendizagem e grande potencial para facilitar processos metacognitivos envolvidos na aprendizagem autorregulada, como autoavaliação, planejamento e reflexão.

Um dos pontos mais relevantes na definição de modelos que sejam capazes de avaliar

² Do Inglês: Open Learner Models (OLMs) ou Open Student Models (OSMs)

o nível de desempenho do estudante é considerar questões comportamentais no processo de inferência. Diversos trabalhos disponíveis na literatura apresentam abordagens que se baseiam exclusivamente no conhecimento e nas suas relações para definir qual o estado cognitivo do estudante (CHRYSAFIADI; VIRVOU, 2013b; HAWKINS; HEFFERNAN; BAKER, 2014; SEFFRIN et al., 2016). Mesmo que sejam abordagens bem discutidas e muito empregadas, não levar em conta o comportamento no processo de identificação de desempenho é um ponto negativo.

A maioria dos trabalhos encontrados na literatura que tratam de questões sobre modelagem de estudante com o intuito de auxiliar na identificação de comportamento ou de conhecimento, raramente, aborda ambas características em um único ambiente. Outro fato interessante diz respeito ao domínio de conhecimento modelado por esses MEs. A maioria dos sistemas inteligentes que utilizam MEs é específica para um determinado domínio, não permitindo, assim, a sua aplicação em outros contextos. Uma questão que vem sendo pouco discutida na criação desses modelos é o fato de que muitos trabalhos não abordam a capacidade de reutilização e extensão dos modelos criados, não permitindo, com isso, a aplicação desses modelos em outros ambientes.

Para a criação desses MEs, diversas abordagens têm sido empregadas ao longo dos anos, como: sobreposição (*overlay*) (BRUSILOVSKY; MILLA'N, 2007; MAHNANE; LAS-KRI, 2012; GAUDIOSO; MONTERO; OLMO, 2012) e estereótipos (*stereotypes*) (CHRYSAFIADI; VIRVOU, 2008; DURRANI; DURRANI, 2010; TOURTOGLOU; VIRVOU, 2012). Outros trabalhos empregam recursos de Inteligência Artificial, como Aprendizado de Máquina (CETINTAS et al., 2010; LI et al., 2011; AL-HMOUZ et al., 2011). Algumas técnicas são focadas em modelos probabilísticos, como é o caso de Lógica Fuzzy (GOEL; LALLÉ; LUENGO, 2012; JEREMIĆ; JOVANOVIĆ; GAVSEVIĆ, 2012; CHRYSAFIADI; VIRVOU, 2015) ou Redes Bayesianas (RBs) (CHIEU et al., 2010; GOGUADZE et al., 2011; MILLÁN et al., 2015).

Pesquisas mais recentes têm se pautado em modelos semânticos, como é o caso da utilização de ontologias (CLEMENTE; RAMÍREZ; ANTONIO, 2011; GRUBIŠIĆ; STANKOV; PERAIĆ, 2013; KUSUMAWARDANI; PRAKOSO; SANTOSA, 2014) para a descrição e a representação das características dos estudantes. Mesmo sendo um tema de pesquisa recorrente e bem estudado, são poucas as propostas que abordam a criação de um ME que combine mais de uma técnica para inferir sobre o nível de desempenho do estudante, tomando como base seu conhecimento e seu comportamento durante a utilização de um SAIE (CHRYSAFIADI; VIRVOU, 2013b).

1.2 Questões de Pesquisa

Como pode ser observado com o que foi discutido até o momento, existem algumas lacunas de pesquisa sobre os processos de modelagem de estudante. Essas estão, justamente,

no estudo de modelos que permitem a análise sobre desempenho do estudante de forma a considerar o seu comportamento com a criação de modelos que sejam independentes de domínio, extensíveis e reutilizáveis.

A fim de delimitar o escopo e direcionar o andamento deste estudo, foi definida uma questão de pesquisa principal e duas questões de pesquisa acessórias para auxiliar no desenvolvimento do trabalho.

- **Questão de Pesquisa Principal.** Como definir e utilizar uma abordagem híbrida para modelagem de estudante, que se baseia no nível de aquisição de conhecimento e no comportamento para auxiliar na determinação do seu nível de desempenho em um SAIE?
 - **Questão de Pesquisa Acessória 1.** Qual impacto de utilizar métricas de comportamento em um SAIE para determinar o nível de desempenho do estudante?
 - **Questão de Pesquisa Acessória 2.** Redes Bayesianas podem ser utilizadas para auxiliar na determinação do conhecimento do estudante?

1.3 Objetivos

Este trabalho tem como principal objetivo a definição de uma abordagem híbrida para modelagem de estudante com condição para realizar avaliação e identificação de desempenho em SAIEs. A abordagem proposta baseia-se em informações sobre o nível de conhecimento e o comportamento dos estudantes para definir o seu desempenho. Para isso, a abordagem proposta combina a utilização de RBs e ontologias, a fim de criar um modelo que seja dinâmico, probabilístico, independente de domínio, extensível e reutilizável. É também proposta uma extensão para tornar o modelo um MAE, o qual permite a visualização das principais capacidades e limitações dos envolvidos no processo de ensino/aprendizagem. Pretende-se, ainda, realizar um estudo de caso por meio da integração da proposta em um ambiente real de ensino, com o intuito de realizar experimentos mensurando diversos indicadores para avaliar o impacto da abordagem.

Além desse objetivo geral, a proposta ora apresentada contempla, ainda, os seguintes objetivos específicos:

- Definição de uma abordagem probabilística e dinâmica para a identificação de conhecimento do estudante, utilizando instrumentos avaliativos e RBs.
- Definição de uma abordagem semântica para modelagem de estudante com capacidade de identificação de desempenho, tomando como base o conhecimento e o comportamento durante a utilização de um SAIE.

- Aplicação de recursos de Visualização de Informação ao ME criado, fornecendo capacidade para atuar como um MAE, o qual, dentre outras funções, permite o monitoramento e o acompanhamento do progresso do estudante.
- Implementação da abordagem proposta através de módulos fracamente acoplados e sua integração em um ambiente de aprendizagem real.

1.4 Organização desta Tese

Esta tese está organizada da seguinte maneira: no Capítulo 2, são apresentados os conceitos-chave para entendimento da proposta, e é realizado um levantamento sobre os principais trabalhos correlatos disponíveis na literatura. No Capítulo 3, são expostos os detalhes sobre a proposta, destacando, principalmente, as características e funcionalidades de cada parte do trabalho. No Capítulo 4, é apresentado um estudo de caso, o qual aborda a integração do modelo proposto a um ambiente real de ensino. No Capítulo 5, são descritos os experimentos e os resultados sobre a avaliação da proposta; e, por fim, no Capítulo 6, são discutidas as contribuições, limitações, trabalhos futuros e, adicionalmente, uma lista com as publicações resultantes deste trabalho.

CAPÍTULO 2

Revisão de Literatura

2.1 Considerações Iniciais

Neste capítulo, é apresentada uma revisão de literatura sobre os principais conceitos que fundamentam a abordagem proposta. Os conceitos são apresentados de forma aprofundada, com foco nas especificidades dos contextos abordados no trabalho.

Além desses conceitos, ao longo de todo o capítulo, foi feita uma revisão crítica e comparativa frente a trabalhos correlatos, sintetizados em uma tabela ao final.

2.2 Modelagem do Estudante

Com o avanço das tecnologias que permitem maior representação e contextualização da informação, os ambientes educacionais ganham, a cada dia, mais recursos para melhorar na tarefa de ensino individualizado. Um Sistema Adaptativo e Inteligente para Educação (SAIE) é um exemplo de ferramenta que auxilia no processo de adaptação de conteúdo de acordo com as preferências, as capacidades e as limitações de cada estudante.

Tais ambientes fazem o uso de técnicas inteligentes com o intuito de se adaptar às reais necessidades dos estudantes, buscando, com isso, prover um ensino mais personalizado e individualizado. Esses sistemas devem ser capazes de gerenciar os percursos de aprendizagem de cada estudante, baseando-se nos seus conhecimentos, atividades, necessidades e preferências. Em outras palavras, um SAIE deve oferecer personalização para as necessidades específicas, conhecimento e experiência de cada estudante (BOTICARIO; SANTOS; ROSMALEN, 2010).

O principal objetivo de um SAIE é adequar sua comunicação com o estudante – navegação e interface – de acordo com um modelo predefinido que reflete suas características, objetivos, preferências, estado cognitivo e competências. De acordo com Iglesias et al. (2009), dentre as várias vantagens de se utilizar um SAIE, podem-se destacar: *(i)* um SAIE pode fornecer conteúdo individualizado permitindo que todos os estudantes tenham acesso ao mesmo material com diferentes apresentações e atividades; *(ii)* um SAIE pode

permitir que os estudantes alcancem níveis de proficiência de forma mais eficiente; e (iii) um SAIE pode permitir que os professores se concentrem em um pequeno subconjunto de estudantes que realmente precisam de ajuda extra em determinado tópico do conteúdo.

Um dos principais aspectos que devem ser observados durante construção de SAIEs é a forma como acontece a adaptação individualizada de conteúdo (GILMAN et al., 2015). Para que essa adaptação funcione corretamente, o sistema deve gerenciar um conjunto de informações sobre o estudante – informações pessoais, contexto de acesso, conhecimento, experiência, interesses, preferências, objetivos, entre outras informações. Esse conjunto de informações é definido e representado mediante um modelo de dados, conhecido como Modelo de Estudante (ME).

A modelagem de estudante é uma tarefa que envolve a representação de diversas questões cognitivas, como: análise de desempenho, identificação de equívocos, representação de objetivos e planejamentos, observação do conhecimento *a priori*, descrição de características pessoais, dentre outras (SELF, 1990). Estas características são todas dinâmicas, sendo necessário um ajuste constante dos seus valores a fim de garantir sua consistência em relação à realidade do estudante. O ME oferece descrições do que realmente é relevante sobre o comportamento, o conhecimento e as aptidões dos estudantes. Baseado no ME, um SAIE ajusta o seu comportamento para atender às necessidades dos estudantes, com o intuito de fornecer conteúdo adequado às suas reais características.

Um ME deve ser capaz de representar toda informação sobre o estudante, inclusive, aquela aparentemente irrelevante. De acordo com Self (1988), o ME é uma representação explícita do que se acredita a respeito do conhecimento, do progresso no conteúdo, preferências, Estilos de Aprendizagem (EAs), objetivos de aprendizagem, motivação, crenças e características pessoais. Neste caso, ele deve funcionar como um espelho, refletindo as características essenciais, em um dado instante, sobre o real estado do estudante.

A construção de um ME não é uma tarefa simples. Informações e dados que representam o estudante devem ser levantadas cuidadosamente, o processo de atualização das informações também deve ser claro, e, não menos importante, deve-se definir como o ambiente irá prover a adaptação de conteúdo (MILLÁN; LOBODA; CRUZ, 2010). Segundo Chrysafiadi e Virvou (2013b), o processo de modelagem do estudante requer a resposta para as seguintes questões: (i) quais características serão representadas? (ii) como serão modeladas? E, por fim, (iii) como o modelo será usado?

É possível encontrar na literatura diversos trabalhos que abordam modelagem do estudante. A maioria deles foi desenvolvida com o objetivo de melhorar a capacidade de representação das informações cognitivas e educacionais visando um ensino mais adaptável e personalizado. Para o desenvolvimento dessa modelagem, diversas abordagens podem ser utilizadas, como: *Sobreposição (overlay)*, *Estereótipos (stereotypes)*, *Aprendizado de Máquina*, *Lógica Fuzzy*, *Redes Bayesianas* e *Ontologias*.

Na modelagem por *sobreposição* (CARBONELL, 1970), o conhecimento do estudante

é considerado um subconjunto do conhecimento de todo o domínio. As diferenças no comportamento do estudante em relação ao comportamento de alguém com conhecimento completo são tratadas como uma indicação de lacunas de conhecimento. De acordo com Millán, Loboda e Cruz (2010), esse modelo funciona razoavelmente bem sempre que o objetivo principal do sistema é transmitir conhecimento do sistema para o estudante.

Já a modelagem por *estereótipos* baseia-se em características comuns a grupos de pessoas. Segundo Rich (1999), um estereótipo é um conjunto de características, por meio do qual é possível inferir o comportamento provável de uma pessoa de características semelhantes, conhecido o comportamento das outras pessoas do mesmo estereótipo.

As observações do comportamento do estudante podem fornecer dados nos quais um sistema de aprendizado de máquina pode usar para inferir ações futuras. Essa é a principal função dos MEs fundamentados em técnicas de *aprendizagem de máquina*. De acordo com Sison e Shimura (1998), técnicas de aprendizado de máquinas têm sido empregadas na modelagem de estudante com duas principais razões: (i) definir um ME único e consistente a partir de múltiplos comportamentos observados, e (ii) estender ou construir automaticamente uma biblioteca de erros (*bugs*) sobre os comportamentos observados.

Determinar o conhecimento de um estudante não é uma tarefa direta, às vezes, é refletida através de coisas que não podem ser diretamente observadas e medidas. Essa é uma das razões de se utilizar *Lógica Fuzzy* (ZADEH, 1965) no processo de modelagem de estudante. Segundo Chrysafiadi e Virvou (2012), técnicas Fuzzy são utilizadas, principalmente, com o intuito de tratar incertezas no processo de identificação e predição de comportamento de estudante.

Outra técnica muito empregada para o tratamento das incertezas no processo de modelagem de estudante é a Rede Bayesiana (RB) (CONATI; GERTNER; VANLEHN, 2002). Na modelagem de estudante, os nós de uma RB podem representar os diferentes componentes/dimensões de um estudante, tais como conhecimento, equívocos, emoções, EAs, motivação, metas, etc. Segundo Desmarais e Baker (2012), as RBs atraíram muita atenção de teóricos e desenvolvedores de sistemas devido às suas sólidas bases matemáticas e também por uma maneira natural de representar a incerteza usando probabilidades.

Recentemente, muitas pesquisas têm explorado as capacidades das *ontologias* para representação e modelagem de estudante (SOSNOVSKY; DICHEVA, 2010). O fato de uma ontologia suportar a representação de conceitos e propriedades para ser facilmente reutilizada e, se necessário, estendida em diferentes contextos de aplicação, permite o raciocínio sobre a informação descrita na ontologia. Essa característica propicia um avanço substancial na modelagem de estudante, possibilitando inferências mais consistentes e capacidades antes limitadas aos modelos criados.

As duas últimas técnicas citadas (RBs e ontologias) foram escolhidas como mecanismos para modelagem da abordagem proposta neste trabalho. A escolha se deu principalmente pela necessidade de criação de uma abordagem que fosse probabilística e que também

oferecesse recursos para reúso e extensão do modelo criado. Por isso, essas técnicas serão explicadas mais detalhadamente a seguir.

2.3 Redes Bayesianas

Dentre os diversos métodos estatísticos utilizados em *Inteligência Artificial*, é possível destacar o aprendizado Bayesiano, que recorre a um modelo probabilístico baseado no conhecimento prévio com o intuito de definir a probabilidade final de uma determinada hipótese. Para trabalhar com aprendizado Bayesiano, geralmente, é empregado o modelo de RBs.

De acordo com Russell e Norvig (2013), RBs são modelos probabilísticos que descrevem a relação de dependência entre variáveis de um determinado domínio, a fim de realizar inferências sobre dados e informações incertas. As RBs podem representar, essencialmente, qualquer distribuição de probabilidade conjunta e completa.

Ainda segundo Russell e Norvig (2013), a especificação de uma RB pode ser dada por:

- Cada nó corresponde a uma variável aleatória, que pode ser discreta ou contínua;
- Um conjunto de arestas que conecta pares de nós. Uma aresta de X até Y mostra que X é pai de Y. Não existem ciclos.
- Cada nó X possui uma distribuição de probabilidade condicional $P(X|Pais(X))$, que quantifica o efeito dos pais sobre o nó.

De acordo com Pearl (1988), essa estrutura representada pelas RBs possui todas as informações necessárias para responder a qualquer questão sobre um conjunto de variáveis. Utilizando uma Tabela de Probabilidade Condicional, é possível determinar todas as probabilidades associadas a cada nó. Cada nó na rede possui a sua tabela, e, com isso, é possível determinar a probabilidade de acontecer um evento dada a ocorrência dos eventos associados.

As RBs são representadas por meio de Grafo Direcionado Acíclico¹, no qual os vértices representam as variáveis do domínio e as arestas representam as relações condicionais entre as variáveis. Cada variável possui um ou mais estados, os quais podem receber um conjunto de valores que representa a probabilidade de ocorrência desse estado para essa variável.

Quando uma variável X está diretamente ligada a uma variável Y ($X \rightarrow Y$), isso aponta que Y é dependente de X e a probabilidade de ocorrência de Y é dada pela probabilidade condicional $P(Y|X)$. Essa probabilidade pode ser interpretada como “*a probabilidade do evento Y ocorrer dada a ocorrência do evento X*”. Um outro caso acontece quando uma determinada variável Y não possui nenhuma aresta incidindo nela, neste caso, esta

¹ Do Inglês *Directed Acyclic Graph - DAG*

variável é considerada um nó raiz e possui somente as probabilidades *a priori*, ou seja, $P(X)$ (PEARL, 1988).

O processo de inferência sobre as RBs acontece por meio da apresentação de evidências e observação das alterações ocorridas nos valores das probabilidades representadas para as variáveis. Evidências são variáveis nas quais existe a certeza de que ela irá assumir um determinado valor. Por exemplo, quando é possível dizer que determinado estudante sabe determinado conceito, essa informação é usada como uma evidência, pois tem-se a certeza de que ele conhece sobre esse conceito.

Quando aplicadas para modelagem do estudante, as RBs desempenham um papel bem definido, capaz de determinar o domínio de conhecimento por meio da sua estrutura de relação entre conceitos e realizar inferências sobre o estado de conhecimento dos estudantes. Geralmente, a estruturação das RBs segue a seguinte abordagem: (i) alguns nós são utilizados para representar os conhecimentos (tópicos, assuntos, conteúdo, etc); (ii) outros nós são utilizados como evidências, ou seja, recebem informações sobre o conhecimento do estudante, e (iii) a inferência é realizada observando as mudanças ocorridas nas probabilidades, a fim de determinar qual o nível de cada conhecimento, representado pelos nós.

Na literatura, é possível encontrar diversos trabalhos que propõem a utilização de RBs em SAIEs. Um dos principais foi proposto por Corbett e Anderson (1994) e define um método de modelagem conhecido como *Bayesian Knowledge Tracing* (BKT), pelo qual se avalia o nível de conhecimento do estudante por meio de cálculos de probabilidade. O conhecimento é avaliado por meio da aplicação de regras, sendo que, caso o conhecimento atinja um determinado percentual, a regra é considerada aprendida. A abordagem utiliza inferência bayesiana para auxiliar na determinação do nível de conhecimento.

Um dos trabalhos mais conhecidos que trata sobre a modelagem de estudante utilizando RB foi proposto por Millán, Loboda e Cruz (2010). Esse trabalho fornece diretrizes e exemplos para entender as RBs e usá-las para projetar e implementar MEs. O trabalho aborda a forma como os modelos devem ser propostos, indicando, por exemplo, diferentes modelos para a topologia da rede, como os nós evidências devem ser alocados, bem como técnicas para avaliação dos valores inferidos pelo modelo.

Yudelson, Koedinger e Gordon (2013) revisitam a proposta apresentada por Corbett e Anderson (1994) e propõem outra abordagem para configuração das variáveis, pela qual se determina que parametrizar a velocidade de aprendizado dos estudantes é mais benéfico do que parametrizar o conhecimento *a priori*. Com isso, é possível conseguir individualizar a inferência do modelo, tornando-o mais condizente com as características individuais de cada estudante.

Torabi, Moradi e Khantaimoori (2012) propõem uma abordagem, baseada em RBs, para a previsão da pontuação dos estudantes com base no seu histórico educacional. A rede é utilizada tanto para estruturação dos conceitos quanto para auxiliar no processo de

inferência. Os resultados mostram que a aplicação do método proposto tem efeitos positivo na qualidade do aprendizado dos estudantes e pode ser usado como uma ferramenta útil para análise de conhecimento.

Outra abordagem também pode ser vista em Ramírez-Noriega, Juárez-Ramírez e Martínez-Ramírez (2017), em que é apresentado um módulo para avaliação do conhecimento do estudante baseado em RBs. O módulo se adapta dinamicamente a forma como os exercícios são apresentados aos estudantes com base nas suas respostas. Para isso, uma rede de conceitos foi utilizada com o intuito de criar um relacionamento entre os conceitos e receber as inferências sobre o nível de conhecimento do estudante.

Grubišić, Stankov e Peraić (2013) propõem uma abordagem probabilista baseada em RBs, que permite tirar conclusões sobre o estado de conhecimento do usuário, e assim, melhorar o processo de ensino/aprendizagem. Nesse trabalho, é proposto um método não empírico para determinar as probabilidades condicionais da RB. Isso proporciona maior autonomia do algoritmo, pois dispensa a difícil tarefa de adequar pesos à rede.

Em Klerk, Eggen e Veldkamp (2016), é descrito um estudo que emprega RBs para avaliação baseada em simulação. Nesse trabalho, é apresentada uma metodologia para o acompanhamento do comportamento interativo e complexo dos estudantes em um tipo específico de avaliação baseada em simulação². O método auxilia na determinação do conhecimento e habilidades dos estudantes em ambientes experimentais. No trabalho apresentado por Käser, Hallinen e Schwartz (2017), é definido um modelo que baseia, principalmente, nas escolhas e estratégias tomadas pelos estudantes durante o aprendizado. Nesse trabalho, foi investigado o impacto das estratégias de exploração sobre a aprendizagem e foi proposto o uso de um modelo probabilístico, baseado em RBs, que representa o conhecimento e as estratégias dos estudantes.

Alguns trabalhos empregam as RBs com um objetivo bem específico, como é o caso do trabalho desenvolvido por Taraghi et al. (2016), que busca a modelagem e a representação de equívocos em multiplicação de um dígito³. Esse trabalho visa mapear os erros dos estudantes a fim de manter atualizado o estado cognitivo destes.

Em Lin, Shen e Chi (2016), é apresentada uma abordagem que estende a proposta da modelagem baseada em BKT adicionando recursos externos sobre o comportamento do estudante. Na abordagem, uma nova técnica é definida (*Intervention Bayesian Knowledge Tracing - Intervention-BKT*) a qual incorpora vários tipos de intervenções instrucionais e tempo de resposta ao modelo tradicional BKT. Resultados mostram que a modelagem *Intervention-BKT* é mais eficiente que a BKT.

Algumas abordagens utilizam extensões das RBs, como Redes Bayesianas Dinâmicas (RBDs) para auxiliar no processo de descoberta e acompanhamento do nível de conhecimento do estudante. Em Seffrin et al. (2016), é apresentada uma modelagem de RBD

² Do Inglês Simulation-Based Assessment (SBA)

³ Do Inglês one-digit Multiplication

para a avaliação do conhecimento algébrico do estudante em Sistemas Tutores Inteligentes (STIs). Segundo os autores, o principal diferencial desse trabalho é a sua capacidade de avaliar tanto o conhecimento conceitual, quanto procedural do estudante, utilizando um modelo de inferência que é independente de problemas (equações). Já em Kaser et al. (2017) é apresentada uma abordagem para diferenciação do nível de conhecimento representado pelo modelo, para isso, é utilizado um algoritmo para aprendizagem de parâmetros que se mostrou eficiente nos testes realizados. Um estudo realizado por Millán, Cruz e García (2003) fez uma comparação entre os valores de probabilidades inferidos pelas RBs estáticas e pelas RBDs, no qual foi possível constatar que, mesmo as duas redes apresentando valores bem similares, as RBs estáticas tiveram um desempenho um pouco melhor que as RBDs (MILLÁN; LOBODA; CRUZ, 2010).

2.4 Ontologias

A utilização de ontologias no contexto da *Ciência da Computação* tem evoluído rapidamente com o passar dos anos. Diversas propostas têm empregado esse conceito com o intuito de padronizar a representação de informações, de forma que estas possam ser entendidas e manipuladas tanto por usuários quanto por sistemas computacionais.

Na literatura é possível encontrar diversas definições para as ontologias. Uma bem aceita e empregada pela comunidade acadêmica foi proposta por Gruber (1993), na qual ontologia é definida como “*uma especificação explícita de uma conceitualização*”, ou seja, uma formalização dos conceitos e relacionamentos em um domínio específico. Outra definição mais específica pode ser encontrada em Borst (1997), na qual uma ontologia é definida como uma especificação formal e explícita de uma conceitualização compartilhada, em que o termo “especificação formal” diz respeito a algo que é interpretável pelos computadores, o termo “explícita” são os conceitos, propriedades, relações, funções, restrições e axiomas explicitamente definidos, o termo “conceitualização” representa um modelo abstrato de algum fenômeno do mundo real e, por fim, o termo “compartilhada” significa conhecimento consensual.

Na prática, uma ontologia define um “vocabulário”, conjunto de termos, que será utilizada para formular consultas (ALMEIDA; BAX, 2003). De acordo com Morais e Ambrósio (2007), a ontologia define as regras de combinação entre os termos e seus relacionamentos, estes são criados por especialistas, e os usuários elaboram consultas usando os conceitos especificados. Mesmo com essas características, Gruber (1993) assegura que o principal propósito da construção de ontologias é possibilitar o compartilhamento e a reutilização de conhecimento.

Segundo Yu (2011), os principais benefícios na utilização das ontologias são: *(i)* permitir interpretação e compartilhamento sobre conceitos de um domínio específico; *(ii)* permitir reutilização do domínio de conhecimento; *(iii)* permitir descrever o conhecimento

e a semântica para que computadores possam compreender o domínio descrito.

A OWL é considerada a linguagem de ontologias da Web e é bastante empregada para o desenvolvimento de aplicações baseadas na Web Semântica (ISOTANI; BITTENCOURT, 2015). Uma ontologia OWL pode representar um domínio específico, criando classes e propriedades, definindo indivíduos e afirmações sobre eles e especificando fatos e conhecimento. Segundo Welty, McGuinness e Smith (2004), a OWL possui três sublinguagens que podem ser usadas por diferentes propósitos: (i) *OWL Lite*, é uma sublinguagem que usa somente algumas características da linguagem OWL; (ii) *OWL DL*, permite o máximo de expressividade, com completude (todas as conclusões são computáveis) e decidibilidade computacional (todas as operações terminarão em tempo finito); e por fim, (iii) *OWL Full*, proporciona o máximo de expressividade e independência sintática de RDFs (*Resource Description Framework*), sem nenhuma garantia computacional.

As ontologias descritas com o OWL-DL possuem a capacidade de realizar processamento por meio de um *Reasoner*. O *Reasoner* possui a capacidade de inferir novos conhecimentos a partir das descrições e relações presentes na ontologia. Além dessas relações é necessário um conjunto de regras descritivas, o qual permite descrever logicamente um novo conhecimento a ser obtido pela ontologia. Os *Reasoners* mapeiam uma base de conhecimento existente, com conceitos, relações e fatos, e, a partir disso, inferem conhecimento adicional apresentando informações implícitas. Os principais motores de inferência disponíveis na literatura são o Pellet⁴, Racer⁵ e o Fact++⁶.

O uso de ontologias para representar modelos de usuários tem atraído, cada dia mais, a atenção de pesquisadores. Chrysafiadi e Virvou (2013b) mostram um crescente número de trabalhos que abordam essa técnica para modelagem de informações educacionais. Entre eles, é possível citar SoNITS (NGUYEN et al., 2011), um ME fundamentado em ontologias para um ambiente educacional baseado em redes sociais com o intuito de ajudar na organização do conhecimento e do raciocínio sobre as relações de habilidade. O ambiente é focado no ensino de conceitos relevantes em Tecnologia da Informação (TI). Para a representação do estudante, três dimensões são tratadas: (i) dados demográficos e características mentais, (ii) competências e interesses e (iii) política e conectividade em redes sociais. Mesmo expondo um bom *feedback* por parte dos estudantes, o modelo proposto é bem simplificado, deixando de representar algumas informações necessárias, como contexto e atividades de colaboração.

Clemente, Ramírez e Antonio (2011) apresentam o MAEVIF, uma plataforma de software voltada ao desenvolvimento de ambientes educacionais para treinamentos. Esse ambiente proporciona conteúdo personalizado e fornece o *feedback* mais adequado em cada instante da interação por meio de um ME, que é baseado em ontologias e regras de diagnóstico. São utilizados vários agentes de software que interagem com uma rede

⁴ <https://www.w3.org/2001/sw/wiki/Pellet>

⁵ <https://www.w3.org/2001/sw/wiki/RacerPro>

⁶ <http://owl.man.ac.uk/factplusplus/>

de ontologia capaz de representar informações sobre o perfil, estado atual e atividades desempenhadas pelos estudantes. Também são discutidas algumas regras que podem ser empregadas para realizar inferência sobre diagnósticos.

No trabalho de Panagiotopoulos et al. (2012), é proposto um ME baseado em ontologias para ambiente educacional voltado à educação a distância. A ontologia é capaz de representar informações sobre localização, domínio de conhecimento, interesses, objetivos, estilos de aprendizado e afeto. A ontologia desenvolvida foi combinada com abordagens de sobreposição e estereótipos com o objetivo de garantir uma caracterização mais adequada aos usuários. Regras semânticas também são propostas, as quais, combinadas com mecanismos de inferência, classificam o estudante em perfis, permitindo, assim, a produção de conhecimento adicional.

Em Rezende et al. (2015), é descrita uma ontologia para a modelagem do perfil e do contexto do estudante em um ambiente *e-learning*, a qual visa padronizar as informações e facilitar o processo de recomendação de objetos de aprendizagem. A ontologia foi especificada de forma que possa ser adaptada a qualquer ambiente de aprendizagem. Resultados preliminares mostraram boa adequação em cenários reais.

No trabalho apresentado por Nafea et al. (2016), é proposto um novo modelo para definição automática do perfil de estudante em um ambiente *e-learning*. A abordagem baseia-se em padrões reais de comportamento dos estudantes durante a interação com um portal educacional (AAST) e emprega ontologias e mecanismos de inferências para identificar EAs automaticamente de acordo com o modelo *Felder-Silverman Learning Style Model* (FSLSM). A partir disso, a personalização pode ser alcançada coordenando o perfil do estudante com os perfis dos cursos oferecidos na faculdade.

Pukkhem (2013) propõe um *framework* baseado em ontologias e regras de mapeamento semântico para personalização em *e-learning* baseado no estilo de aprendizagem do estudante. Esse *framework* emprega ontologias para que computadores sejam capazes de interpretar e processar objetos de aprendizagem em um sistema de recomendação. Em seguida, a recomendação baseia-se no mapeamento semântico entre o ME e os objetos de aprendizagem.

Ouf et al. (2017) propõem e implementam um *framework* para *smart e-learning* que utiliza ontologias e regras *Semantic Web Rule Language* (SWRL). A abordagem proposta cria quatro ontologias com o intuito de acompanhar todo o processo de personalização de conteúdo educacional aos estudantes do ambiente. Uma ontologia é utilizada para representar um modelo de aprendizagem e outras três para representar todos os componentes do processo de aprendizagem (objetos de aprendizagem, atividades de aprendizagem e métodos de ensino). Regras SWRL são utilizadas para auxiliar no processo de inferências.

No trabalho proposto por Casals e Brandão (2017), são definidas ontologias para representar e integrar dados de contexto móvel e objetivos de aprendizagem. Uma ontologia

é utilizada para representar as dimensões envolvidas em um ambiente de aprendizagem virtual acessado por um dispositivo móvel e uma outra é empregada para representar os objetivos de aprendizagem como parte de metodologias pedagógicas. O principal foco do ambiente é representar três dimensões: física, social e temporal.

Outros trabalhos definem ME de forma híbrida, ou seja, exploram as melhores características de diferentes abordagens para produzir um modelo que seja mais adaptável ao contexto e domínio do estudante (CHRYSAFIADI; VIRVOU, 2015). Um modelo híbrido foi proposto por Sani e Aris (2014), que recorre à Lógica Fuzzy e as técnicas de ontologias para modelar o comportamento de aprendizagem do estudante com o objetivo de melhorar a aprendizagem e aumentar a adaptabilidade do sistema. A Lógica Fuzzy possibilita a análise computacional das características dos estudantes a fim de lidar com as questões de incertezas relacionadas com o design do ME. Outro exemplo de ambiente híbrido é o *PeRSIVA* (CHRYSAFIADI; VIRVOU, 2013a). Esse apresenta uma abordagem híbrida, que combina as técnicas de modelagem por sobreposição, estereótipos e Lógica Fuzzy, a fim de criar um ambiente capaz de avaliar resultados na satisfação, no desempenho, no progresso, no comportamento e no estado dos estudantes.

Mesmo com diversas abordagens de modelagens utilizando modelos híbridos, a união de RB com ontologias ainda é pouco explorada (CHRYSAFIADI; VIRVOU, 2013b), sendo possível perceber uma lacuna de estudos que exploram a capacidade de representação das ontologias trabalhando junto com a grande capacidade de processamento de incertezas providas pelas RBs. Esse é justamente um dos diferenciais da proposta apresentada neste trabalho, a qual faz o uso conjunto dessas técnicas para prover um modelo para a representação de estudantes em SAIEs.

2.5 Modelo Aberto do Estudante

Tradicionalmente, os dados tratados pelo ME são “fechados” para os usuários, fornecendo informação somente para o próprio sistema criar mecanismos específicos para a personalização individualizada. Contudo, diversas pesquisas, ao longo dos anos, têm proposto ferramentas para “externalizar” essas informações tratadas pelo ME, ou seja, deixar os dados “abertos” para todos os usuários envolvidos no processo educacional (LONG; ALEVEN, 2013; BULL; KAY, 2013). Essa capacidade de “abrir” as informações inferidas pelo ME é conhecida como Modelo Aberto do Estudante (MAE). Os MAEs exploram a área de Visualização de Informação para produzir ferramentas que sejam capazes de prover mecanismos para estudantes e professores visualizarem, explorarem e até mesmo modificarem a forma como o conhecimento é construído e tratado pelo ME. De acordo com Brusilovsky et al. (2016), permitir que os estudantes visualizem aspectos relacionados a seu modelo pode melhorar a auto-reflexão e o aprendizado auto-regulado, possibilitando, com isso, a transparência da personalização e da motivação do estudantes.

Esses modelos considerados “abertos” tem por finalidade apresentar as informações processadas pelo ME de forma clara e concisa. Como o processamento sobre o nível de desempenho do estudante envolve uma série de variáveis, muitas vezes, fica difícil a interpretação desses resultados, assim, estes modelos buscam auxiliar na apresentação, bem como, na interpretações de tais valores.

Alguns trabalhos disponíveis na literatura exploram a capacidade dos MAEs. Dimitrova, Self e Brna (2001) exploram uma abordagem interativa de modelagem aberta do estudante, no qual, o diagnóstico do aluno é considerado como um processo interativo envolvendo um sistema adaptativo e um aprendiz que desempenham papéis simétricos (até certo ponto) e constroem juntos o modelo do estudante. Chen et al. (2007) investigam a capacidade dos MAEs na motivação dos estudantes para melhorar seus desempenhos acadêmicos. Nesse trabalho, tanto os modelos abertos individuais como os em grupo foram estudados; ambos demonstraram um aumento na reflexão e interações úteis entre colegas de equipe.

Ainda sobre o ponto de vista da motivação, no trabalho proposto por Hsiao et al. (2011), MAEs foram incorporados na anotação de links adaptáveis, no qual foi possível mostrar que essa combinação pode aumentar notavelmente a motivação dos estudantes a trabalharem com conteúdo educacional não obrigatório.

Em Mazza e Dimitrova (2007), é proposto um dos poucos ambientes que fornece visualização gráfica para vários grupos de usuários, tanto professores quanto estudantes. Neste ambiente, o modelo proposto ajuda os professores a identificar os problemas desde o início do curso, evitando-assim, alguns problemas comuns na educação a distância.

De acordo com Long e Aleven (2017), poucos são os trabalhos que verificam a capacidade dos MAEs melhorarem o desempenho dos estudantes. Eles propõem um estudo que verifica a capacidade dos estudantes se auto-avaliarem em relação a um domínio de resolução de equações matemáticas. Nesse caso, o MAE foi integrado ao tutor *Lynnette*, no qual foi possível avaliar o ganho de conhecimento dos que utilizaram a plataforma.

Como é possível observar, são poucos os trabalhos que exploram a capacidade de integração de MAEs a ambientes que permitem monitorar constantemente o nível de desempenho dos estudantes. Muitas das abordagens propostas investigam somente questões de auto-regulação e questões de motivação, não explorando capacidades para monitoramento sobre o nível de desempenho (LONG; ALEVEN, 2017). Um dos diferenciais deste trabalho é justamente a capacidade do MAE poder trabalhar diretamente tanto com questões comportamentais quanto sobre o nível de desempenho dos estudantes.

2.6 Síntese dos Trabalhos Correlatos

Esta seção apresenta uma síntese dos principais trabalhos citados, ao longo do capítulo, que exploram a criação de mecanismos e abordagens para modelagem do estudante.

A partir da Tabela 1, é possível ter um panorama geral sobre as principais características desses trabalhos. O levantamento foi realizado com o intuito de especificar seis características, a saber: (1) o trabalho apresenta características probabilísticas; (2) o estudo explora a utilização de abordagens semânticas; (3) o trabalho implementa modelagem aberta do estudante; (4) o modelo proposto pela pesquisa é considerado dinâmico; (5) o trabalho pode ser aplicado em diferentes domínios de conhecimento (independente de domínio); e por fim, (6) a pesquisa apresenta um modelo que é extensível e pode ser reutilizável.

É importante destacar que os trabalhos foram agrupados em quatro frentes: trabalhos que envolvem a modelagem do estudante, modelos que utilizam RBs, abordagens construídas por meio de ontologias e, trabalhos que exploram capacidades de modelagem aberta do estudante. Esses são, justamente, os principais pontos do trabalho proposto.

Para o levantamento realizado, foram consideradas três dimensões de comparação: (●) o trabalho implementa completamente a funcionalidade/característica; (◐) o trabalho implementa parcialmente a funcionalidade/característica e (○) o trabalho não implementa a funcionalidade/característica. Na última linha da Tabela 1 é apresentada a abordagem ora proposta.

A primeira observação que pode ser feita a partir da Tabela 1 é que nenhum dos trabalhos comparados implementa todas as características levantadas. Pode-se observar que, no máximo, os trabalhos conseguem explorar quatro características; ao contrário da abordagem proposta que implementa todas essas características estudadas. É possível observar também que a característica de permitir a modelagem aberta do estudante é a que apresentou uma menor concentração de estudos. Por outro lado, a dinamicidade foi observada em quase todas as pesquisas.

Outra característica interessante que pôde ser observada com este levantamento diz respeito ao tipo de abordagem para inferência sobre as informações manipuladas pelo modelo. Dos 28 trabalhos apresentados na tabela, somente um (além da abordagem proposta), implementa as inferências de forma probabilística e semântica, que é o caso do trabalho proposto por Sani e Aris (2014). Nos demais trabalhos, nenhum explora a capacidade de criação de modelos híbridos, baseados em métodos probabilísticos e semânticos para inferência de informações.

Uma última análise recai sobre os trabalhos classificados como extensíveis e reutilizáveis. É possível observar que a maioria deles diz permitir a extensão e a reutilização, no entanto, a maior parte permite isso de forma parcial. Foi possível verificar que muitos desses utilizam mecanismos que permitem a extensão e reutilização, mas não implementam ou disponibilizam suporte para esse reúso. Poucos desses trabalhos apresentam uma arquitetura de integração e reutilização para outros ambientes.

Tabela 1 – Trabalhos correlatos

Trabalho	PB. ¹	SM. ²	MAE ³	DN. ⁴	ID. ⁵	ER. ⁶
Casals e Brandão (2017)	○	●	○	○	○	●
Ouf et al. (2017)	○	●	○	●	●	○
Kaser et al. (2017)	●	○	○	●	○	○
Käser, Hallinen e Schwartz (2017)	●	○	○	○	○	○
Long e Aleven (2017)	○	○	●	●	○	●
Ramírez-Noriega, Juárez-Ramírez e Martínez-Ramírez (2017)	●	○	○	●	●	○
Nafea et al. (2016)	○	●	○	●	●	○
Seffrin et al. (2016)	●	○	○	●	○	○
Intervention-BKT (LIN; SHEN; CHI, 2016)	●	○	○	●	●	○
Taraghi et al. (2016)	●	○	○	●	○	●
MBPA (KLERK; EGGEN; VELDKAMP, 2016)	●	○	○	●	○	○
Personna (REZENDE et al., 2015)	○	●	○	●	●	●
Sani e Aris (2014)	●	●	○	○	●	○
PeRSIVA (CHRYSAFIADI; VIRVOU, 2013a)	●	○	○	●	●	○
Yudelson, Koedinger e Gordon (2013)	●	○	○	●	●	○
Pukkhem (2013)	○	●	○	○	●	●
Grubišić, Stankov e Peraić (2013)	●	○	○	○	●	○
Panagiotopoulos et al. (2012)	○	●	○	●	○	●
Torabi, Moradi e Khantaimoori (2012)	●	○	○	○	●	○
MAEVIF (CLEMENTE; RAMÍREZ; ANTONIO, 2011)	○	●	○	●	○	●
SoNITS (NGUYEN et al., 2011)	○	●	○	○	○	●
Sosnovsky e Dicheva (2010)	○	●	○	○	○	●
Millán, Loboda e Cruz (2010)	●	○	○	●	○	○
Mazza e Dimitrova (2007)	○	○	●	○	○	○
Andes (CONATI; GERTNER; VANLEHN, 2002)	●	○	○	○	○	○
APT (CORBETT; ANDERSON, 1994)	●	○	○	●	○	○
Modelo Proposto	●	●	●	●	●	●

●: Presente

○: Presente parcialmente

○: Ausente

¹Probabilístico. ²Semântico. ³Modelo Aberto do Estudante. ⁴Dinâmico. ⁵Independente de domínio.⁶Extensível e reutilizável.

2.7 Considerações Finais

Neste capítulo, foi apresentada uma revisão de literatura acerca dos principais conceitos que norteiam este trabalho. Foi discutido o conceito de modelagem do estudante, no qual foram abordadas questões sobre as diferentes características de um ME, bem como técnicas utilizadas na literatura para tal modelagem. Conceitos sobre RBs e ontologias também foram vistos, no qual foi possível discorrer sobre a aplicação dessas técnicas computacionais no escopo da modelagem de estudantes. Por fim, foram discutidas questões

sobre a modelagem aberta do estudante, como suas características, utilização e implicações.

Junto com todos esses conceitos abordados, foram apresentados os principais trabalhos disponíveis na literatura passíveis de comparação com a abordagem aqui proposta. Utilizando esse paralelo, foi possível construir uma seção de síntese e uma tabela com as principais características presentes em cada trabalho.

CAPÍTULO 3

Abordagem Proposta

3.1 Considerações Iniciais

Neste capítulo, é descrita uma abordagem híbrida para modelagem do estudante, a qual baseia-se em Redes Bayesianas (RBs) e ontologias, para análise, acompanhamento e predição do desempenho do estudante. Para isso, métricas comportamentais e de conhecimento são utilizadas como parâmetros de entrada para a inferência dos valores tratados pelo modelo. O modelo proposto integra diferentes técnicas de representação de informação e análise de incertezas com o intuito de maximizar a acurácia na identificação do nível de desempenho do estudante e, assim, melhorar o processo de ensino/aprendizagem.

Este modelo é capaz de identificar as capacidades e limitações dos estudantes de forma automática e dinâmica, sempre levando em consideração a forma como ele se comporta durante a utilização do ambiente educacional. Na abordagem ora proposta, uma RB foi empregada para inferir sobre o nível de conhecimento baseado em atividades avaliativas e uma ontologia foi empregada para a representação das características dos estudantes e a inferência sobre o nível de desempenho baseado no comportamento. Uma importante característica desse modelo é o fato de que a representação do domínio de conhecimento está separada das informações tratadas na ontologia, possibilitando, assim, que o modelo seja independente de domínio.

Outra característica importante deste trabalho é que são propostas abordagens para que o Modelo de Estudante (ME) aqui apresentado seja considerado um Modelo Aberto do Estudante (MAE). O modelo explora capacidades de visualização de informações a fim de criar uma ferramenta capaz de apresentar, de forma estruturada, as informações inferidas, permitindo, portanto, um conhecimento prévio e o monitoramento constante da evolução na aprendizagem.

Ademais, a arquitetura proposta contempla:

- ❑ A definição de ontologias para representação das principais informações que caracterizam o comportamento, o estado atual de conhecimento, as habilidades e as

restrições dos estudantes. (Resultados desta abordagem foram publicados em: Ferreira et al., 2015 e Ferreira et al., 2017c disponíveis na lista de publicações deste trabalho.)

- a criação de um ME Híbrido, independente de domínio, que combina ontologias e RBs com o objetivo de tratar incertezas e determinar o nível de desempenho dos estudantes. (Resultados desta abordagem foram publicados em: Ferreira et al., 2016a, Ferreira et al., 2016b e Ferreira et al., 2017a disponíveis na lista de publicações deste trabalho.)
- A incorporação de recursos de Visualização de Informação para estender a abordagem proposta, criando, dessa forma, um MAE. (Resultado desta abordagem foi publicado em: Ferreira et al., 2017b disponível na lista de publicações deste trabalho.)

3.2 Avaliação do Nível de Aquisição de Conhecimento do Estudante

A Modelagem do Estudante é um dos componentes fundamentais na personalização do processo de ensino/aprendizagem em Sistemas Adaptativos e Inteligentes para Educação (SAIEs) e Sistemas Tutores Inteligentes (STIs) (MITROVIC; THOMSON, 2009). De acordo com Li et al. (2011), essa modelagem faz parte de um dos fatores-chave que afetam sistemas automatizados de tutoria na tomada de decisões e pode ser definida como o processo de coleta de informações importantes para inferir o atual estado cognitivo do estudante. Um ME deve representar não somente informações sobre conhecimento do estudante, mas também, refletir, o mais fiel possível, o seu processo de raciocínio (CLEMENTE; RAMÍREZ; ANTONIO, 2011).

Trabalhar com o processo de raciocínio do estudante exige a manipulação e o tratamento de diversas questões cognitivas. Um dos grandes desafios na modelagem do estudante é justamente a forma como serão tratadas as incertezas no processo de aprendizagem. Teorias probabilísticas têm sido propostas por alguns autores para lidar com a incerteza no diagnóstico do nível de aquisição de conhecimento do estudante (CONATI; GERTNER; VANLEHN, 2002; MILLÁN; LOBODA; CRUZ, 2010; CHRYSAFIADI; VIRVOU, 2012; CHRYSAFIADI; VIRVOU, 2015).

Uma teoria bastante difundida para a modelagem do estudante são as RBs, modelos matemáticos representados por meio de um grafo direcionado acíclico, cujos vértices representam as variáveis do domínio e as arestas representam as relações de dependências entre essas variáveis. Essa teoria é empregada, dentro de uma estrutura de inteligência probabilística e artificial, com o intuito de modelar como um sistema inteligente deve inferir questões referentes à causalidade (TORABI; MORADI; KHANTAIMOORI, 2012). Além

disso, possui representação e comportamento semelhante à mente humana (NAVARRO, 2008). Segundo Millán, Loboda e Cruz (2010), RBs são uma das melhores opções para construção de MEs.

O modelo proposto neste trabalho emprega RBs com o objetivo de inferir a probabilidade do nível de aquisição do conhecimento do estudante baseando-se em mecanismos avaliativos disponibilizados no ambiente educacional. A rede é utilizada com o intuito de prover um enfoque probabilístico ao modelo, a partir do qual é possível trabalhar com dados e informações incertas. Como exposto por Clemente, Ramírez e Antonio (2011), modelos baseados em RBs permitem realizar conclusões sobre o estado de conhecimento do estudante e, assim, melhorar o processo de ensino/aprendizagem.

3.2.1 Independência de Domínio

Além da capacidade de tratamento e realização de inferências sobre a probabilidade de conhecimento de determinado estudante em determinado tópico educacional, as RBs também foram utilizadas como mecanismos para representação do domínio de conhecimento no modelo. Como as RBs são, por natureza, modelos que representam hierarquias de estruturação e precedência de conceitos, esses modelos tornam-se interessantes para a representação de um mapa conceitual, descrevendo os tópicos e suas relações em um determinado domínio de conhecimento.

O fato de recorrer a uma RB como um mapa conceitual determina um dos principais fundamentos do modelo ora proposto, que é independente de domínio. Para um modelo ser considerado independente de domínio, ele não deve estar associado a um determinado domínio de conhecimento, ou seja, o modelo é genérico o suficiente para representar qualquer domínio educacional.

Todo o conhecimento representado pelo modelo proposto neste trabalho está descrito na RB, isso quer dizer que a base de conhecimento é fundamentada exclusivamente pela rede. Por ser um modelo híbrido, que utiliza RBs e ontologias, as atribuições de cada técnica estão bem definidas e delimitadas. A RB permite a representação do domínio e a inferência da probabilidade de conhecimento, ao passo que a ontologia fica responsável pela representação das características e informações referentes ao comportamento do estudante (o funcionamento e integração da ontologia ao modelo serão explicados na Subseção 3.3).

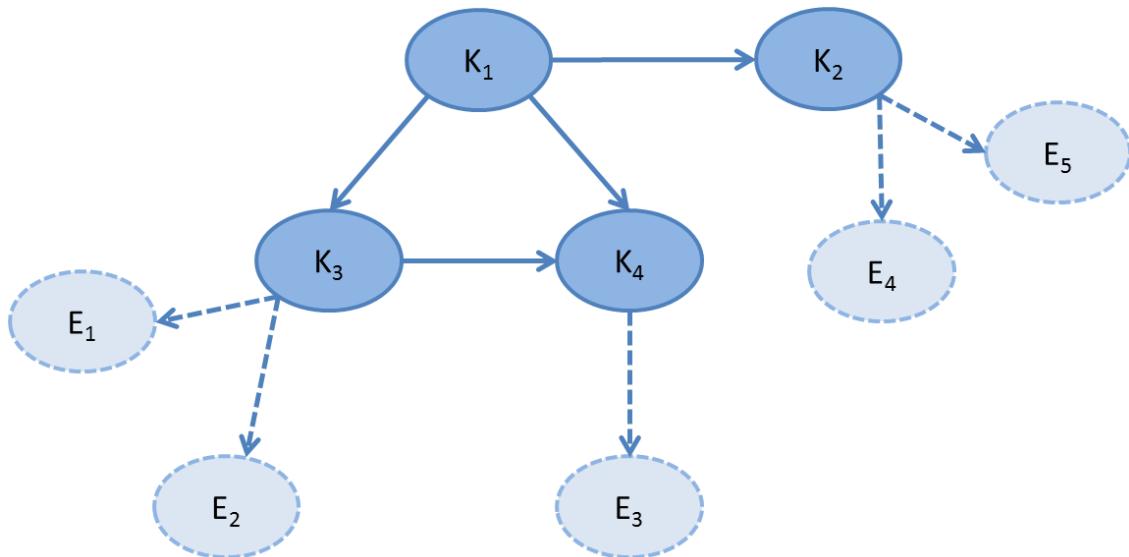
Na literatura, é possível encontrar diversos trabalhos que abordem a criação de MEs baseados unicamente em RBs. Esses modelos são bem aceitos pela comunidade, no entanto, apresentam grande limitação ao representar um único domínio de conhecimento, não sendo possível sua extensão a outros contextos. Essa abordagem de integrar RBs ao ME com o intuito de descrever um modelo conceitual é viável de ser implementada, como defendido por Millán, Loboda e Cruz (2010).

3.2.2 Representação dos Conceitos

A RB utilizada neste modelo segue, em parte, a abordagem de modelagem de estudante orientada por Millán, Loboda e Cruz (2010). A abordagem proposta define que a RB a ser integrada ao modelo deve ser composta por dois tipos de nós: **conceitos** (k) e **evidências** (e). Nós conceitos são aqueles que devem representar um conhecimento associado ao domínio modelado (domínio de interesse). Por exemplo, para o domínio de Introdução à Programação, um nó conceito pode ser “*IF*” ou “*Estruturas Condicionais*”. Por outro lado, é possível um nó ser definido como nó evidência, ou seja, um nó utilizado para observação do desempenho do estudante em determinado conceito, por exemplo, uma questão de múltipla escolha.

Os nós conceitos relacionam-se entre si, formando uma rede de conceitos inter-relacionados. Esses conceitos são relacionados de acordo com sua estrutura semântica e demonstram uma relação de dependência. Já os nós evidências, por possuírem uma semântica diferente, estão associados a um determinado conceito representando a entrada para o nível de conhecimento do estudante em um determinado assunto. Na Figura 1, é possível observar um trecho de uma RB com quatro nós conceitos e cinco nós evidências.

Figura 1 – Exemplo de nós (k) e (e) em uma RB.



Os nós $K_1 \dots K_n$ são nós conceitos e estão representando, cada um, um assunto a ser abordado em uma determinada disciplina. As arestas no grafo representam a relação de dependência entre dois nós conceitos. Por exemplo, a aresta (K_1, K_3) determina que K_1 é pré-requisito de K_3 . As arestas (K_1, K_4) e (K_3, K_4) determinam que K_1 e K_3 são pré-requisitos para K_4 .

Para esse modelo, não é imposto um nível de abstração para cada nó K , ou seja, esses nós podem representar conceitos mais genéricos ou conceitos mais específicos. O modelo não distingue o nível de densidade semântica do nó. Por exemplo, um nó K_i

pode representar um conceito amplo (ex., estruturas de repetição, tipos de dados) ou um conceito restrito (ex., *while*, *inteiro*).

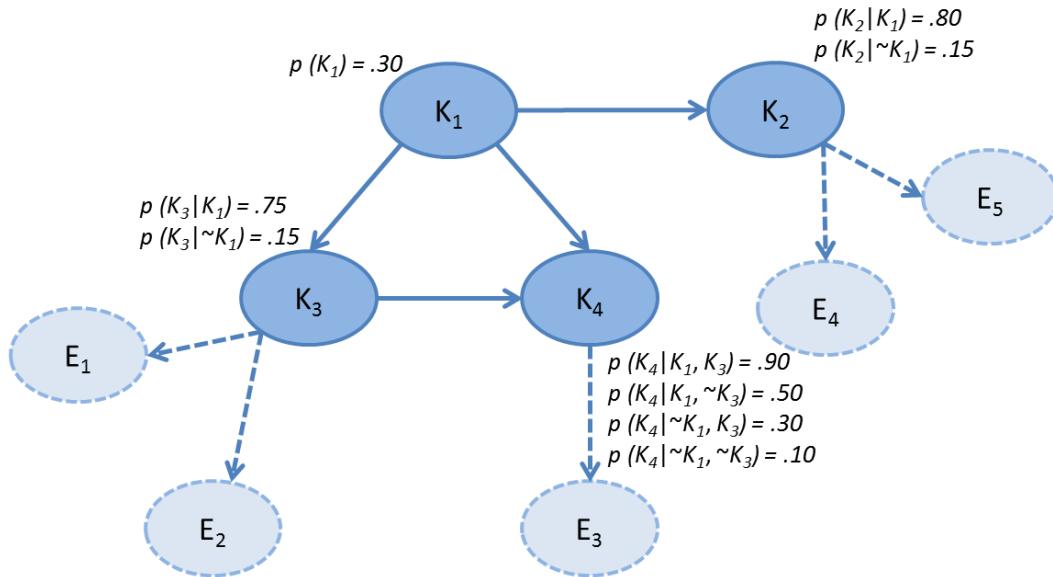
Os nós $E_1 \dots E_n$ são nós evidências e estão representando, cada um, um instrumento avaliativo associado a um ou mais nós K_s . Esse nó é utilizado como entrada de informação sobre o conhecimento do estudante para um determinado conceito. Informações sobre os nós evidências serão explicados com mais detalhes na Subseção 3.2.4.

3.2.3 Definição das Probabilidades

No modelo proposto, as probabilidades condicionais associadas aos nós conceitos seguem a semântica originalmente definida por Pearl (1985) e Pearl (1988), pela qual a probabilidade de um determinado evento (neste caso, o conhecimento em um determinado assunto) ocorrer depende da ocorrência de um outro evento anterior.

Na Figura 2, é possível observar as probabilidades associadas a cada nó K na rede. Para um K_i , a sua probabilidade de ocorrência é condicionada à probabilidade da ocorrência dos seus pré-requisitos. Por exemplo, a probabilidade de ocorrência de K_2 está associada à ocorrência de K_1 , que é representada por $P(K_2|K_1)$: “A probabilidade da ocorrência do evento K_2 , dada a ocorrência do evento K_1 ”.

Figura 2 – Probabilidades condicionais associadas a uma RB.



A definição das probabilidades condicionais para a rede constitui uma das tarefas mais importantes quando se trabalha com métodos baseados em RBs. Neste modelo, são propostos dois métodos para a definição das probabilidades. O primeiro método é, prioritariamente, baseado em informações empíricas. Neste método, o professor (instrutor ministrante da disciplina) estabelece as probabilidades para cada nó K baseado no seu próprio conhecimento. O próprio professor tem a capacidade de planejar a topolo-

gia (estrutura) da rede e suas probabilidades, tomando como base a sua experiência na disciplina.

O segundo método para definição da tabela de probabilidade baseia-se em uma abordagem não empírica e automatizada, a qual não depende da experiência do professor. Nesta abordagem é considerada somente a estrutura da rede para determinação dessas probabilidades. Esse método emprega como principal informação a quantidade de nós pré-requisitos para realizar o cálculo da probabilidade condicional relativa a cada nó. Por exemplo, se um determinado nó K possui três pré-requisitos, a probabilidade deste evento K ocorrer é de 0,3 para cada nó pré-requisito. A abordagem em questão foi estendida de um trabalho bastante explorado na literatura, proposto por Grubišić e colaboradores (GRUBIŠIĆ; STANKOV; PERAIĆ, 2013).

3.2.3.1 Probabilidade Inicial

Um detalhe relevante neste modelo é a definição das probabilidades iniciais de conhecimento. Essa informação é importante para que os resultados inferidos pelas RBs apresentem melhor correspondência com os resultados reais de conhecimento. Neste modelo, as probabilidades iniciais podem ser baseadas em informações de pré-testes. Nesse caso, cada estudante possui suas habilidades medidas antes de qualquer inferência pela rede. Por outro lado, é possível definir uma probabilidade inicial igual para todos os estudantes.

Os valores das probabilidades de conhecimento representados neste modelo sempre variam na escala de 0,1 a 0,9, onde 0,1 define extrema dificuldade do estudante e 0,9 define extrema facilidade pelo estudante. Valores menores que 0,1 e maiores que 0,9 não são representados pelo modelo. De acordo com Grubišić, Stankov e Peraić (2013), valores extremos como 0,0 e 1,0 não devem ser considerados em modelos de estudante, visto que devem-se levar em conta ocorrências de acerto por meio de chutes ou erros por causa de algum deslize. Essa abordagem tem sido considerada em diversos modelos probabilísticos.

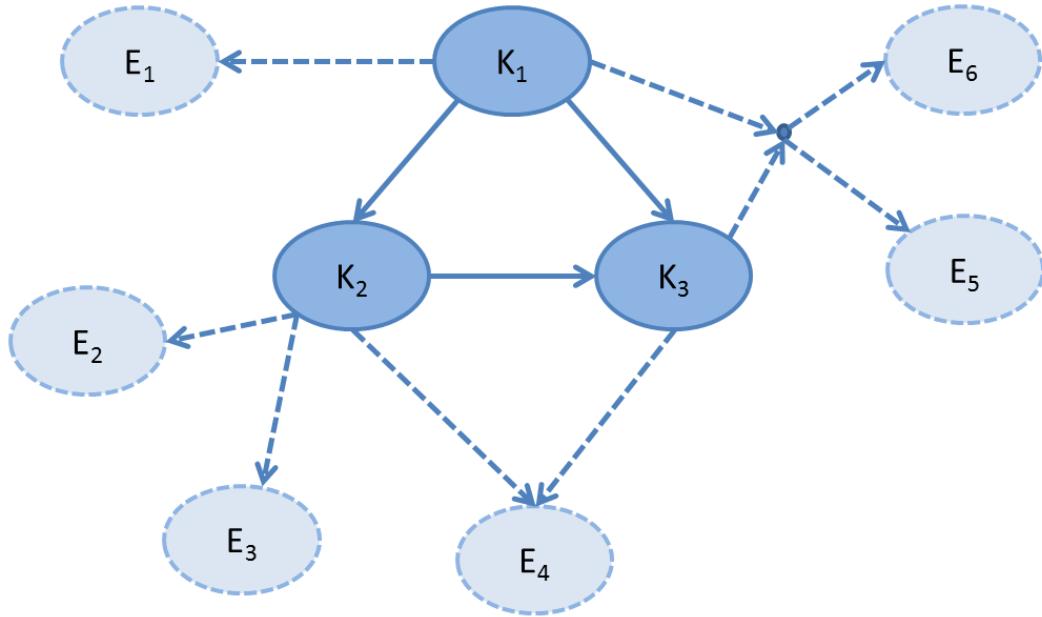
3.2.4 Definição dos Instrumentos Avaliativos

Instrumentos avaliativos são mecanismos para avaliar o nível de conhecimento do estudante em determinado tópico no domínio de conhecimento e são representados, neste modelo, por meio dos nós evidências (E). Esses instrumentos podem ser de diferentes tipos, devendo, contudo, representar o nível de conhecimento do estudante com um valor entre 0,1 e 0,9 (segundo o mesmo padrão apresentado anteriormente).

Cada nó E_n está associado a um ou mais nós K . Na Figura 3, é possível verificar quatro maneiras diferentes de associação dos nós E aos nós K . A primeira maneira é quando um instrumento avaliativo E está relacionado somente a um conhecimento K (E_1 associado a K_1); a segunda maneira é quando um instrumento avaliativo E está associado

a mais de um conhecimento K (E_4 associado a K_2 e K_3); a terceira maneira é quando mais de um instrumento avaliativo E está associado a um conhecimento K (E_2 e E_3 associado a K_2); e, por fim, a última maneira é quando mais de um instrumento avaliativo E está associado a mais de um conhecimento K (E_5 e E_6 associados a K_1 e K_3).

Figura 3 – Exemplos de diferentes disponibilizações dos instrumentos avaliativos.



Alguns trabalhos disponíveis na literatura discutem diferentes abordagens para disponibilização de instrumentos avaliativos (HAWKINS; HEFFERNAN; BAKER, 2014; CHEN; WUILLEMIN; LABAT, 2014; VIER; GLUZ; JAQUES, 2015). A proposta deste modelo estende a técnica proposta por Hawkins, Heffernan e Baker (2014) no sentido de organização e disponibilização dos instrumentos avaliativos.

3.2.5 Definição Formal

Buscando uma melhor representação conceitual, foi criado, para o contexto deste trabalho, uma definição formal para o modelo proposto. Neste sentido, a representação do ME por meio de uma RB é composta por um conjunto de conceitos, evidências e relações (que podem ser relações entre conceitos ou relações entre evidências e conceitos). Para indicar a direção do relacionamento entre conceitos, o termo pai e filho é utilizado. Por exemplo, dizer que K_1 é pai de K_2 é equivalente a dizer que K_1 é pré-requisito de K_2 . Com o intuito de indicar os relacionamentos entre conceitos e evidências, pode ser introduzida a seguinte definição:

Definição 1: Considere $K = \{k_1, \dots, k_n\}$, $n \geq 0$, um conjunto de conceitos; $E = \{e_1, \dots, e_m\}$, $m \geq 0$, um conjunto de evidências; $Rk = \{(k_i, k_j) \mid k_i \in K \text{ e } k_j \in K \text{ e } k_i \neq k_j\}$, $|Rk| \geq 0$, um conjunto de pares ordenados. Se $(k_i, k_j) \in Rk$, k_i é pré-requisito de k_j ;

$Re = \{(k_i, e_j) \mid k_i \in K \text{ e } e_j \in E\}$, $|Re| \geq 0$, um conjunto de pares ordenados. Se $(k_i, e_j) \in Re$, e_j é uma avaliação associada a k_i ;

Definição 2: O domínio de conhecimento (DC) são pares ordenados constituídos de conceitos (K) e evidências (E) que determinam o conteúdo, ou disciplina, para o qual se deseja representar tal informação. Destaca-se que esse conteúdo pode ser representado em diferentes granularidades, sendo um assunto específico, ou até mesmo, uma disciplina completa.

Foi definido um grafo do domínio de conhecimento (GDC), que pode ser representado pela seguinte definição: $GDC = (V, A)$, onde o conjunto de vértices é formado por $V = (K \cup E)$ e o conjunto de arestas é formado por $A = (Rk \cup Re)$.

Definição 3: O conjunto de vértices que é vizinho de um conceito pode determinar o seu grau, que pode ser definido como *grauConceito*, indicando a quantidade de vértices conceitos adjacentes a um determinado conceito, ou *grauEvidencia*, indicando a quantidade de vértices evidências adjacentes a um determinado conceito.

Para um vértice conceito $k \in V$, seus conceitos vizinhos $Nk(v)$ podem ser definidos por: $Nk(v) = \{w/(w, v), \text{ onde, } w \in K \text{ e } v \in K\}$. Portanto, o *grauConceito* ($Dk(v)$) do vértice v em GDC é o número de conceitos adjacentes a v , isto é: $Dk(v) = |Nk(v)|$.

Para um vértice conceito $v \in V$, suas evidências vizinhas $Ne(v)$ podem ser definidas por: $Ne(v) = \{e/(v, e), \text{ onde, } v \in K \text{ e } e \in E\}$. Portanto, o *grauEvidencia* ($De(v)$) do vértice v em GDC é o número de evidências adjacentes a v , isto é: $De(v) = |Ne(v)|$.

3.2.6 Inferência

Os dados processados pela RB representam o nível de conhecimento do estudante em determinado instante. Esse método de tratamento de incertezas permite detectar o conhecimento “global” a partir da identificação de partes específicas desse conhecimento. Com isso, é possível perceber que não é necessária a inserção de nós evidências em todos os conceitos para se ter uma visão do estado de conhecimento global. A rede permite que, com apenas algumas evidências, inferências sejam processadas para predizer, com certa consistência, o nível de conhecimento global do estudante. Para o processamento de inferências nesta abordagem foi utilizado o algoritmo de formação de agrupamentos (também conhecidos como algoritmos de árvore de junção).

No entanto, somente os valores inferidos pela RB não são capazes de predizer o nível de conhecimento dos estudantes; é preciso mais que isso. Deve-se, também, levar em consideração o seu comportamento, a fim de auxiliar na definição desse desempenho. Para uma análise mais completa e consistente sobre o desempenho do estudante, a abordagem proposta neste trabalho emprega uma ontologia com o intuito de refinar o desempenho final do estudante. Esta ontologia é capaz de representar características e comportamentos do estudante, para, a partir disso, realizar inferência sobre o desempenho global, levando em consideração o conhecimento e o comportamento durante a utilização de um SAIE.

3.3 Avaliação de Desempenho do Estudante

A abordagem proposta neste trabalho apresenta características para a descrição dos principais aspectos educacionais que representam o comportamento do estudante em um sistema adaptativo. Estes ambientes são providos de particularidades que, muitas vezes, não são tratadas na maioria dos ambientes convencionais de ensino. Informações como Estilo de Aprendizagem (EA), contexto de acesso, colaboração e atividades de gamificação são exemplos de características que devem ser tratadas em uma modelagem de estudante.

A modelagem de estudante não é uma tarefa simples. Sua construção exige que as informações representadas sejam cuidadosamente levantadas. O modelo deve ser simples, uma vez que informações importantes não sejam deixadas de fora. Por outro lado, deve ser também completo, representando as principais informações que caracterizam o estado atual de conhecimento, habilidades e restrições de cada estudante.

Na literatura, é possível encontrar diversas propostas para a especificação e a padronização de dados sobre o perfil do estudante. Dentre elas, podem ser citadas: PAPI Learners¹, IMS LIP² e ULF³. Alguns trabalhos também propõem ontologias para essa representação, como: Rezende et al. (2015), Ehimwenma, Beer e Crowther (2015), Chrysafidi e Virvou (2015). Mesmo sendo padrões bem completos, eles não são capazes de representar a grande variedade de informações necessárias para adaptação e personalização de conteúdo em SAIEs. As informações estão dispersas, ou seja, às vezes é necessária a utilização de vários padrões para representar aquilo que realmente é adequado para o ambiente.

A partir do exposto, foi conduzido um intenso estudo com o propósito de levantar as principais características e dados necessários para a descrição de um estudante, para, assim, criar um modelo que fosse conciso e que representasse as principais informações para esses ambientes. As características escolhidas podem ser subdivididas em quatro grupos:

- **Informações Pessoais:** Dados pessoais do estudante, como estilos de aprendizagem segundo a abordagem de Felder-Silverman (FSLSM)⁴, preferências e restrições.
- **Comportamento:** Informações sobre atividades e comportamentos desempenhados no ambiente, como atividades desenvolvidas, atividades de colaboração e atividades de gamificação.
- **Contexto:** Informações sobre o contexto atual de utilização, como dispositivo, localização e motivo de acesso.

¹ Standard for Learning Technology - Public and Private Information for Learners. Tech. report, IEEE
² IMS Learner Information Package Specification. Tech. report, IMS Global Learning Consortium
³ Universal Learning Format Technical Specification, Version 1.0. Tech. report, Saba Software Inc
⁴ Felder-Silverman Learning Styles Model - (FELDER; SILVERMAN et al., 1988)

- **Progresso/Conhecimento:** Informações sobre o estado cognitivo atual do estudante, descrevendo seus conhecimentos, erros e equívocos, capacidades e limitações.

A partir dessas características, foi definida uma ontologia para representar todas essas informações.

3.3.1 Ontologia

A escolha de um formalismo mediante ontologias para modelar o estudante é justificável principalmente por duas razões: *(i)* uma ontologia viabiliza a representação de conceitos e propriedades de modo a serem facilmente reutilizados e, se necessário, estendidos em diferentes contextos da aplicação; *(ii)* o uso de ontologias permite o *reasoning* das informações representadas. Com isso, é possível a implementação de inferências a partir do desenvolvimento de regras (CLEMENTE; RAMÍREZ; ANTONIO, 2011).

A ontologia proposta apresenta características suficientes para abstração de um modelo que seja capaz de representar o estado atual de conhecimento dos estudantes, possibilitando melhor adaptação e personalização de conteúdo educacional. Um detalhe interessante desta abordagem é que o domínio de conhecimento não está integrado à ontologia e sim a um modelo probabilístico, garantindo que a ontologia pode ser empregada em outros domínios sem a necessidade de alterações. Mais detalhes sobre esse recurso podem ser vistos no decorrer desta seção.

3.3.1.1 Construção da Ontologia

A construção de modelos que representem as características de determinados usuários é uma tarefa difícil e complexa. Em particular, criar tais modelos utilizando ontologias traz alguns desafios a mais. Identificar quais características devem ser modeladas, como permitir que as informações representadas possam ser reusadas e estendidas, e ainda como desenvolver um modelo que seja consistente e capaz de inferir sobre as questões importantes para o contexto específico são algumas das particularidades que devem ser observadas durante a construção e especificação de um modelo baseado em ontologias.

Na literatura é possível encontrar diversas metodologias que orientam o processo de construção de ontologias. Essas metodologias são disponibilizadas com o intuito de sistematizar a construção e manipulação dessas ontologias. De acordo com Bautista-Zambrana (2015), esses métodos envolvem um conjunto de atividades, como conceitualização, formalização, implementação e manutenção. Dentre as metodologias mais utilizadas, encontram-se: *(i)* Methontology (LOPEZ; PEREZ; JURISTO, 1997), *(ii)* On-To-Knowledge (STAAB et al., 2001), *(iii)* metodologia de Gruninger e Fox (GRÜNINGER; FOX, 1995), *(iv)* método de Uschold e King (USCHOLD; GRUNINGER, 1995), e por fim, *(v)* método *Ontology Development 101* (NOY; MCGUINNESS, 2001).

Para a construção da ontologia proposta neste trabalho, foi utilizado o método *Ontology Development 101*. Esse tem se mostrado um método bem consistente e amplamente aplicado pela comunidade científica. Além de considerar as principais fases na modelagem de uma ontologia, é capaz de utilizar ferramentas especializadas (como o *Protégé*⁵) para auxiliar no processo de desenvolvimento. Os principais processos representados por essa metodologia são: (i) determinar o domínio e o escopo da ontologia; (ii) considerar a reutilização de ontologias existentes; (iii) enumerar termos importantes na ontologia; (iv) definir classes e hierarquia de classes; (v) definir as propriedades de classes/slots; (vi) definir as restrições; e (vii) criar as instâncias.

A seguir serão detalhadas as atividades desenvolvidas em cada uma das sete etapas:

1. Determinar escopo: O primeiro passo para o desenvolvimento da ontologia é a determinação do domínio e escopo. Nesta etapa, foram respondidas questões básicas que serviram para compreender o propósito da ontologia.

❑ Qual o domínio a ontologia deve cobrir?

A ontologia proposta possui características suficientes para descrever um modelo abstrato de estudante. Informações pessoais, comportamento, contexto e estado cognitivo são exemplos de características que devem ser representadas pelo modelo.

❑ Quais serão os usos da ontologia?

A ontologia foi construída para aplicações que permitam avaliar o nível de desempenho acadêmico do estudante baseando-se em seu comportamento. É importante ressaltar que a ontologia tem comunicação direta com um modelo probabilístico baseado em RBs, que atua como um processador de inferências sobre atividades desenvolvidas pelos estudantes.

❑ Quem irá usar e manter a ontologia?

A ontologia poderá ser mantida (reusada e estendida) por qualquer pessoa ou organização que queira empregar tal modelo em sua aplicação.

❑ Exemplos de questões que a ontologia deverá responder?

a) Qual o desempenho do estudante que participa ativamente das atividades de colaboração?

b) Qual o desempenho do estudante que se enquadra como jogador do tipo *Socializer*?

c) Os estudantes com maior acesso ao ambiente também obtiveram melhores desempenhos na disciplina?

⁵ <https://protege.stanford.edu/>

- d) Os estudantes que ficam muito tempo online na plataforma e realizam poucas atividades possuem desempenho melhor que os estudantes que ficam pouco tempo online?

É importante frisar que o último item visa estabelecer as questões de competência da ontologia (GRÜNINGER; FOX, 1995). De acordo com Isotani e Bittencourt (2015), estas questões são empregadas tanto para estabelecer um limite do escopo da ontologia, como servem de base para a validação da ontologia.

- 2. Considerar reúso:** Esta etapa está relacionada ao reúso de outras ontologias para auxiliar na construção da ontologia de interesse.

A ontologia proposta faz parte de um domínio bastante específico e com diversas características muito peculiares. Além disso, considera diferentes aspectos pessoais, comportamentais e de contexto. Sendo assim, foi difícil encontrar vocabulários já existentes que auxiliariam na construção da ontologia. Para que a ontologia ficasse concisa e com informações claras sobre cada termo utilizado, preferiu-se definir os próprios termos a serem usados.

- 3. Enumerar termos:** Nesta etapa, foram definidos os termos que serão utilizados para a definição da ontologia. É importante frisar que nela não acontece a distinção entre quais termos serão classes, propriedades e elementos. Isso é realizado nas etapas seguintes.

- 4. Definir classes:** Após a definição dos termos a serem utilizados na ontologia, definiu-se a relação de classes e hierarquia de classes. A Tabela 2 apresenta as principais classes criadas para a ontologia proposta. É possível observar também na tabela algumas referências encontradas na literatura sobre cada classe. Para a definição da hierarquia de classes, seguiu-se o modelo proposto por Uschold e Gruninger (1995), que faz uma combinação entre a abordagem *Top-Down* (de cima para baixo) e *Bottom-Up* (de baixo para cima).

- 5. Definir propriedades:** Nesta etapa foram definidas as propriedades que cada classe deve possuir. Após um extenso levantamento, foi possível determinar quais as principais características e propriedades que cada classe deve representar.

A Tabela 3 expõe as propriedades associadas a três classes utilizadas na ontologia.

- 6. Definir restrições:** Nesta etapa foram definidas as restrições associadas às classes e propriedades. Definir restrições é uma importante etapa na construção de ontologias, pois permite maior consistência e representabilidade para o domínio definido na ontologia.

Tabela 2 – Principais classes representadas na ontologia.

N ¹	Classe	Descrição	Referência
<i>i</i>	Student	Representa o estudante no ambiente educacional	(DOLOG; NEJDL, 2007; MAHNANE; LASKRI, 2012; REZENDE et al., 2015)
<i>ii</i>	Personal Information	Informações como nome, email, idade, sexo, etc.	(DOLOG; NEJDL, 2007; REZENDE et al., 2015)
<i>ii</i>	Learning Style	Estilo de Aprendizagem baseado no modelo de Felder e Silverman	(MAHNANE; LASKRI, 2012)
<i>ii</i>	Preferences	Classe genérica para representar a preferência de um estudante	(DOLOG; NEJDL, 2007; REZENDE et al., 2015)
<i>iii</i>	Context	Representação do contexto de acesso	(REZENDE et al., 2015)
<i>iii</i>	Device	Dispositivo usado pelo estudante	(VERBERT et al., 2012; REZENDE et al., 2015)
<i>iii</i>	Location	Localização do estudante durante o acesso	(VERBERT et al., 2012; REZENDE et al., 2015)
<i>iv</i>	Knowledge State	Estado de conhecimento do estudante	(TING; PHON-AMNUAISUK, 2012)
<i>iv</i>	Knowledge	Representa o conhecimento em um conceito particular	(DOLOG; NEJDL, 2007; CLEMENTE; RAMÍREZ; ANTONIO, 2011)
<i>iv</i>	Course	Unidade educacional	(VERBERT et al., 2012)
<i>iv</i>	Content	Conteúdo de um curso	(VERBERT et al., 2012)
<i>v</i>	Behavior	Comportamento do estudante durante a utilização do ambiente educacional	(CLEMENTE; RAMÍREZ; ANTONIO, 2011)
<i>v</i>	Collaboration	Atividades de colaboração, como, comentários, <i>ratings</i> , etc.	(HSIAO et al., 2013)
<i>v</i>	Gamification	Pontuação em atividades de gamificação	(SHUTE et al., 2015; MI-NOVIĆ et al., 2015)
<i>v</i>	Game	Mecanismo de game disponível para pontuação	(SHUTE et al., 2015; MI-NOVIĆ et al., 2015)
<i>vi</i>	Trace	Rastro de atividades executadas durante um acesso	(CLEMENTE; RAMÍREZ; ANTONIO, 2011)
<i>vi</i>	Activity	Atividade executada durante um acesso	(SHUTE et al., 2015)

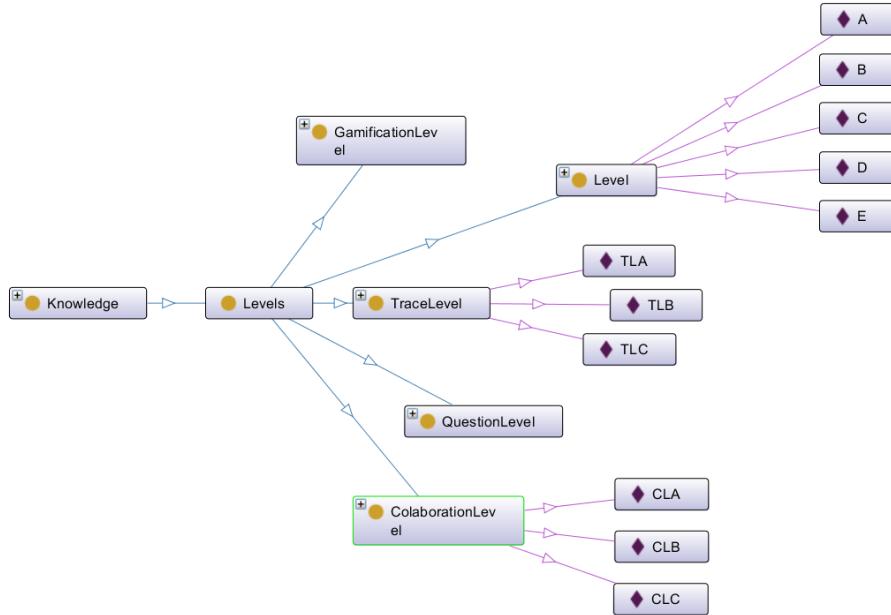
¹Decomposição Modular (Seção 3.3.2).

7. Criar instâncias: A última, e não menos importante etapa na construção da ontologia proposta, diz respeito à criação das instâncias. Inicialmente, foram criadas instâncias obrigatórias que farão parte da ontologia, como pode ser visto na Figura 4. Após a criação destas, foram definidas instâncias de testes com o intuito de verificar a consistência e responder às questões de competências definidas na primeira etapa do processo. Na Figura 5 podem ser vistas algumas instâncias de testes.

Tabela 3 – Propriedades das classes *Personal Information*, *Gamification* e *Trace*.

Classe	Propriedade	Tipo	Descrição
Personal Information	Name	String	Nome
	Surname	String	Sobrenome
	Gender	Male/Female	Sexo
	Telephone	String	Telefone
	Email	String	Email
	Address	String	Endereço
Gamification	Birth_date	Date	Data de Nascimento
	Type	String	Tipo da atividade
	hasGame	Instância de Game	Mecanismo de Game utilizado
	hasBadges	Instância de Badge	Medalhas conquistadas
	Punctuation	Integer	Pontuação conquistada
Trace	hasContent	Instância de Content	Conteúdo associado a atividade
	Description	String	Descrição
	hasActivity	Instância de Activity	Lista de atividades desenvolvidas
	Start	Time	Início
	End	Time	Término

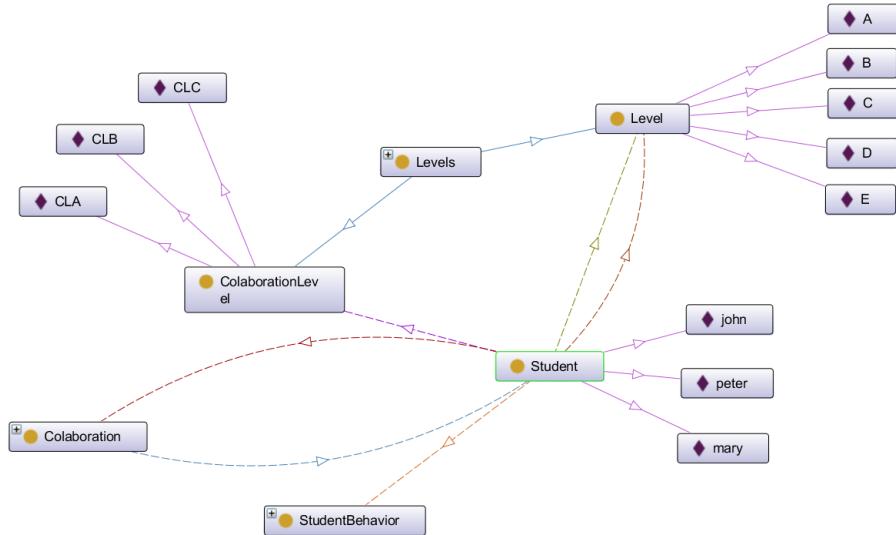
Figura 4 – Exemplo de instâncias representando os níveis de aprendizado.



3.3.2 Decomposição Modular da Ontologia

Durante a construção da ontologia foi empregada uma recomendação proposta por Suárez-Figueroa et al. (2008), na qual as informações modeladas são estruturadas por meio de uma rede de ontologias, também conhecida como rede ontológica, onde várias ontologias menores são interligadas por meio de relações de dependências de modularização. Vale ressaltar que, graças a essa decomposição modular, o impacto de alterações em uma determinada ontologia não afetará as outras. Isso também garante que a ontologia tenha

Figura 5 – Exemplos de instâncias (estudantes e níveis).



maior capacidade de extensão e reutilização.

A partir da etapa de modelagem apresentada anteriormente, foi decidido que, para a correta representação das informações do estudante, seria necessária a construção de seis ontologias: ontologia de perfil, ontologia de informações pessoais, ontologia de contexto, ontologia de estado, ontologia de comportamento, e por fim, ontologia de rastro de atividades.

Nas próximas subseções, além de apresentar os principais elementos que compõem cada ontologia, serão descritas as decisões por trás da sua construção.

3.3.2.1 Ontologia do Perfil

Esta é a ontologia base para o modelo proposto. Nela, são representadas informações sobre o estudante e relações que auxiliam na determinação do seu nível de conhecimento. Esta ontologia possui duas classes principais: *Student* e *Knowledge*. A classe *Student* é uma classe que possui relação direta com todas as outras ontologias do modelo e a classe *Knowledge* permite representar o nível de desempenho do estudante, baseando no seu conhecimento e no seu comportamento.

3.3.2.2 Ontologia das Informações Pessoais

Esta ontologia é utilizada para modelar as características pessoais de cada estudante, ou seja, suas informações exclusivas, independente do ambiente ou contexto da aprendizagem. Esta ontologia é estruturada através de três classes principais – *Personal Information*, *Learning Style* e *Preferences*.

A primeira classe, *Personal Information*, permite representar as características pessoais, como: nome, sobrenome, endereço, idade, sexo, etc. A classe *Learning Style* é decomposta em outras subclasses que possibilitam a representação de informações so-

bre EAs baseados no *Felder-Silverman Learning Style Model* (FSLSM); por fim, a classe *Preferences* possui detalhes para representar as preferências pessoais de cada estudante.

3.3.2.3 Ontologia de Estado

Esta ontologia permite representar informações sobre o estado atual de conhecimento do estudante e, por meio dela, é possível mapear e acompanhar todo o seu processo de desenvolvimento. Essa ontologia possui uma classe principal *Knowledge*, que representa um conhecimento específico do estudante. Essa classe *Knowledge* tem relação com outras classes, como *Knowledge_level*, que representa o nível de conhecimento (que, nesse caso é um valor inferido por um modelo probabilístico), e três classes que representam níveis diferentes de agrupamento de conteúdo: *Course*, *Discipline* e *Content*. No qual, *Content* é mais específico que *Discipline*, que, por sua vez, é mais específico que *Course*.

3.3.2.4 Ontologia de Comportamento

Esta ontologia é empregada para representar informações sobre o comportamento do estudante durante a utilização do ambiente educacional. Essa é uma das ontologias mais importantes para o processo de aquisição do nível de conhecimento deste modelo. Ela possui duas classes principais, *Colaboration* e *Gamification*. *Colaboration* representa informações colaborativas que o estudante desempenha durante a sua permanência no ambiente. Atividades de colaboração são atividades em que o estudante pode interagir com outros estudantes, professores ou até mesmo com o material educacional. A classe *Colaboration* relaciona-se com outras três subclasses, *Colaboration_level* que representa o nível de colaboração do estudante em determinado conteúdo; *Media* e *Type*, que representa a mídia e o tipo de colaboração, respectivamente. A classe *Gamification* representa informações acerca do comportamento do estudante ante a utilização de atividades de gamificação.

3.3.2.5 Ontologia de Contexto

Esta ontologia permite representar informações sobre o contexto de acesso do estudante. Essas informações são importantes quando se procura adaptabilidade por meio de aplicações que estejam inseridas em ambientes de acesso heterogêneos, como é o caso de ambientes ubíquos e presenciais. Um dos motivos para a utilização desta ontologia é o fato de muitas aplicações estarem direcionadas a esse contexto. Além da classe *Context*, a ontologia possui outras classes, como *Device*, *Location* e *Language*.

3.3.2.6 Ontologia de Rastro de Atividades

Esta ontologia também está diretamente relacionada com o comportamento do estudante. Ela é utilizada para representar suas atividades durante o acesso à plataforma

educacional. Todas as atividades desenvolvidas pelo estudante – abrir conteúdo, responder *quiz*, colaborar com outro estudante, fechar uma aula, etc – são registradas. Além da classe *Trace*, essa ontologia possui outras duas classes, *Activity*, que representa uma atividade desempenhada no ambiente e *Trace_level*, que representa o nível de atividades desenvolvidas. É importante destacar que a classe *Trace* possui relação com o contexto de utilização do ambiente por meio da classe *Context* na ontologia de contexto.

3.3.3 Regras para Inferência de Conhecimento

Uma das grandes vantagens de utilizar ontologias está justamente na sua capacidade de processar inferências. O modelo proposto neste trabalho possui a capacidade de realizar inferências sobre o desempenho do estudante com base nas informações descritas na ontologia. Para processar essas inferências, regras foram criadas, a fim de proporcionar um resultado mais condizente com as reais capacidades dos estudantes, pois além de analisarem o conhecimento, também consideram aspectos comportamentais (nível de colaboração, nível de gamificação, atividades desenvolvidas, etc.).

As regras definidas na ontologia são empregadas para inferir o nível de conhecimento que um estudante possui em um determinado conteúdo. O conhecimento inferido pelas regras é classificado em cinco níveis (A, B, C, D e E), onde A representa um conhecimento mais alto (*Expert*) e E um conhecimento limitado (*Limited*). Essa métrica de classificação de conhecimento em níveis tem sido uma abordagem utilizada de forma recorrente na literatura (NGUYEN et al., 2011; TING; PHON-AMNUAISUK, 2012; CHEN; WUILLEMIN; LABAT, 2014). Vale destacar que as regras foram criadas com o auxílio de especialistas no domínio de modelagem de estudantes e por professores com experiência na utilização de SAIEs.

As regras propostas para esse modelo baseiam-se nas informações representadas pela ontologia combinadas com os valores de inferência já realizados pelo modelo probabilístico (ver Subseção 3.2) e produzem um resultado que indica o nível de desempenho do estudante. Um interessante detalhe do modelo proposto é justamente a capacidade de extensão e inclusão de novas regras, visto que cada domínio traz suas próprias características e particularidades.

Para este modelo foram propostas quatro categorias de regras: *(i)* Desempenho do estudante a partir das atividades de socialização/colaboração; *(ii)* Desempenho do estudante a partir das atividades de gamificação; *(iii)* Desempenho do estudante a partir de instrumentos avaliativos, e por fim, *(iv)* Desempenho do estudante a partir das atividades desempenhadas durante a utilização do ambiente “rastro”.

A seguir, serão descritas duas regras, em *Semantic Web Rule Language* (SWRL), que estão disponíveis no modelo. A regra N.1 está disponível na primeira categoria, e a regra N.2 está disponível na terceira categoria. Todas as categorias e subcategorias de regras podem ser vistas no Apêndice A.

Regra N.1:

```
integer[>LLow,<=LHigh] (?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
colabVal(?Co, low) → hasLevel(?S, D)
```

A Regra N.1 baseia-se em apenas duas métricas para inferir o desempenho do estudante: *hasKnowledge* e *hasColaboration*. Esta regra presume que, se o estudante possui um nível de colaboração baixo e seu nível de conhecimento em um determinado tópico (*Content*) é considerado bom, ele possui um nível (*hasLevel*) igual a D. É importante frisar que a escala para representar o nível de conhecimento é flexível, determinada pelas variáveis LLow e LHigh.

Regra N.2:

```
integer[>LLow,<=LHigh] (?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K)
∧ hasContent(?K, ?C) ∧ hasContent(?G, ?C) ∧ hasContent(?Co, ?C) ∧
hasGamification(?S, ?G) ∧ gamVal(?G, high) ∧ hasColaboration(?S, ?Co) ∧
colabVal(?Co, high) → hasLevel(?S, B)
```

A Regra N.2 baseia-se em três métricas para descrever o nível de conhecimento do estudante: *hasKnowledge*, *hasColaboration* e *hasGamification*. Esta regra presume que, se o estudante possui um alto nível de interação com o ambiente, alto nível de participação em atividades de gamificação e um nível de conhecimento regular, ele possui um nível (*hasLevel*) igual a B. Percebe-se que, nesta regra, a participação em atividades de colaboração e gamificação influenciam positivamente a avaliação do desempenho do estudante.

Vale ressaltar que esses são somente alguns exemplos de utilização de regras na ontologia. Outras regras podem ser facilmente criadas dependendo dos reais interesses de onde o modelo proposto for aplicado.

Um detalhe interessante desta proposta é justamente sobre a apresentação das informações obtidas sobre o conhecimento do estudante. Essa fase de apresentação de conteúdo não é, necessariamente, uma função de um ME tradicional. A maioria das aplicações educacionais que empregam MEs utiliza um módulo de interface (independente do modelo) que é utilizado pontualmente para essa apresentação. A proposta apresentada neste trabalho estende a abordagem padrão, criando um MAE, o qual pode ser visualizado e manipulado por estudantes e professores. A próxima seção descreve detalhes que caracterizam o modelo proposto como um MAE.

3.4 Modelo Aberto do Estudante

Além da concepção do modelo previamente mostrado, esse trabalho também apresenta uma extensão, permitindo que o modelo seja “aberto” para visualização e manipulação

pelos usuários do ambiente de aprendizagem. Essa característica traz diversas vantagens para o ambiente, onde professores e estudantes poderão interagir de forma mais efetiva com o ME, podendo, com isso, autorregular suas ações com base no conteúdo adaptado pelo ambiente.

De acordo com Guerra et al. (2016), interfaces inteligentes permitem que a aprendizagem seja mais clara e facilitada, possibilitando uma maior interação com as próprias capacidades cognitivas dos envolvidos diretamente no processo, neste caso, os estudantes. A integração de ferramentas de Visualização de Informação ao contexto de adaptação de conteúdo educacional proporciona novas características aos MEs, permitindo assim, o surgimento, além de modelos abertos, também de modelos inteligentes.

Neste trabalho, o modelo proposto é capaz de fornecer subsídios para que estudantes e professores interajam diretamente com as informações processadas e inferidas pelo modelo. São diversas as possibilidades de interação. Professores conseguem acompanhar de forma clara e simplificada a evolução do estudante durante um curso. A ferramenta também permite aos professores visualizarem, de forma agrupada, os estudantes que possuem melhores e piores desempenhos. Aos estudantes, é possível observar os conteúdos em que possuem mais dificuldades e comparar o seu desempenho ao da turma. Também é permitido acompanharem a sua evolução acadêmica ao longo do tempo em que estão cursando a disciplina. É importante destacar que todas as formas de visualização podem ser personalizadas e adaptadas ao perfil do usuário.

As principais características fornecidas por essa extensão são:

- ❑ Permitir que o professor **identifique**, de forma simplificada, as características educacionais dos estudantes;
- ❑ Permitir que o professor **acompanhe** as capacidades e limitações dos estudantes ao longo do tempo;
- ❑ Permitir que o professor detecte quais são os estudantes com maior chance de **sucesso e fracasso**;
- ❑ Permitir que o estudante **conheça** suas principais capacidades e limitações em relação a determinada disciplina;
- ❑ Permitir que o estudante **ajuste** seus estudos para sanar os assuntos em que tem mais dificuldade;
- ❑ Permitir que o estudante **acompanhe** seu desempenho ao longo do tempo.

Este MAE não se limita a produzir informações que orientem professores e estudantes no intuito de maximizar o sucesso e minimizar o fracasso, mas sim, constitui uma importante ferramenta de auxílio na identificação das reais capacidades e limitações dos

sujeitos envolvidos no processo de ensino e aprendizagem. A utilização do modelo aberto propicia um melhor rendimento aos estudantes, pois estão sempre concentrados nas suas limitações, melhorando, assim, a eficácia no processo de aquisição de conhecimento.

O modelo aberto disponibiliza recursos suficientes capazes de abstrair a geração de diferentes formas de visualização. Para os experimentos propostos, foram implementadas três formas de visualização: gráfico de linha, gráfico de barra e gráfico de radar. Essas visualizações podem ser disponibilizadas para os diferentes usuários do ambiente. A Figura 6 ilustra essas três formas de visualização para uma disciplina de Arquitetura e Organização de Computadores. Na Figura 6(a) é utilizado um gráfico de linha; na Figura 6(b) é utilizado um gráfico de barra; e finalmente, na Figura 6(c) é utilizado um gráfico de radar. Em todas essas visualizações o eixo X corresponde aos tópicos de conhecimento e o eixo Y corresponde ao desempenho inferido. A partir destas visualizações é possível também que o estudante compare seu desempenho com o desempenho global da turma.

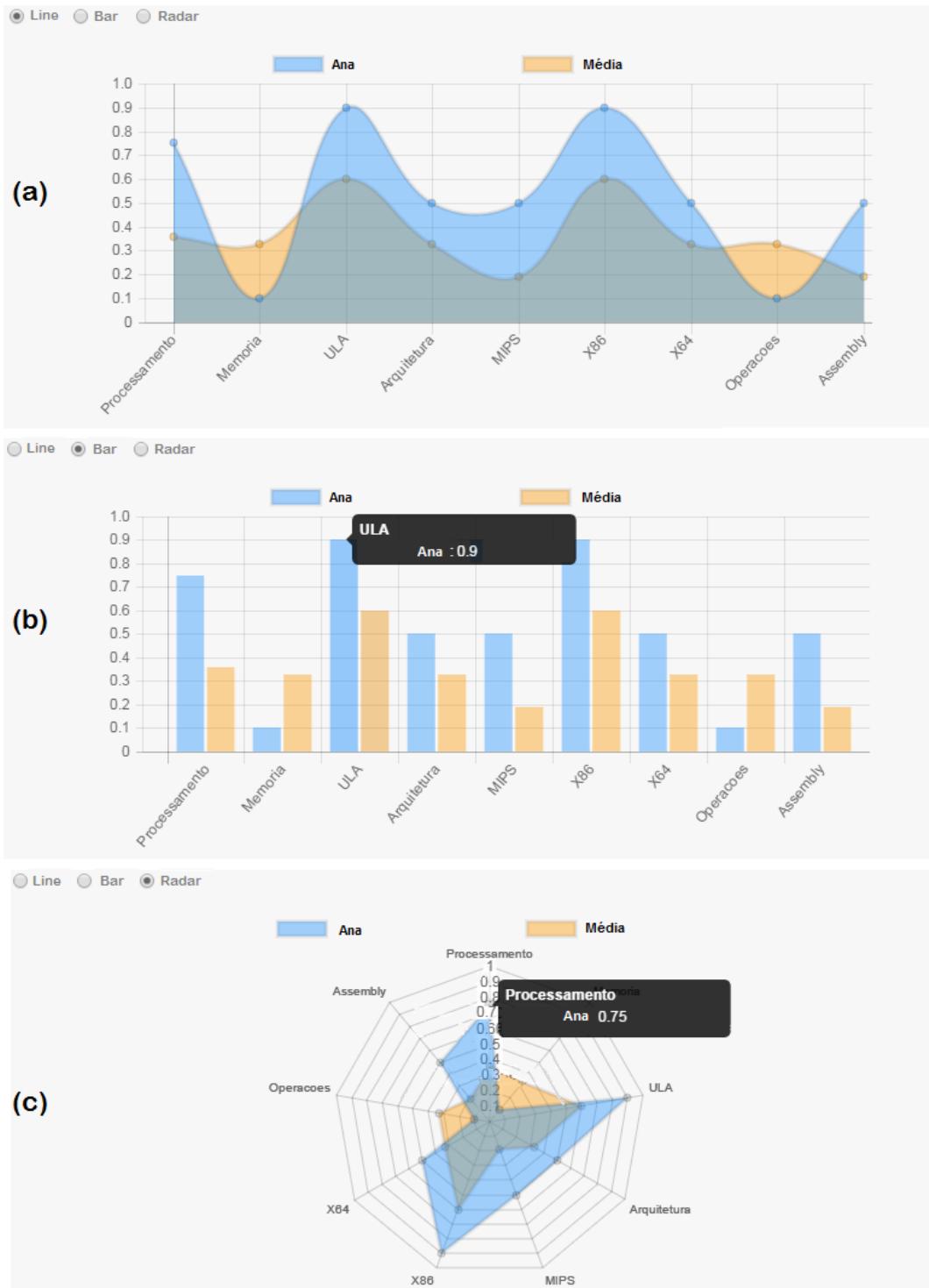
Para o ambiente do estudante, foram adicionados recursos que permitem a visualização do nível de desempenho em diferentes pontos do sistema. O modelo também fornece subsídios para que o estudante possa comparar o seu desempenho ao dos demais estudantes matriculados na disciplina. Para o ambiente do professor, foram disponibilizadas ferramentas que o auxiliam no acompanhamento do processo de aprendizagem dos estudantes matriculados em sua disciplina. Além da visualização individual, o professor pode verificar o desempenho de determinado estudante ante o desempenho da turma. Além disso, o professor consegue também verificar os estudantes com melhor e pior desempenhos, assuntos que apresentam maiores dificuldades para os estudantes, entre outras informações sobre o rendimento da turma.

3.5 Funcionamento e Arquitetura do Modelo

Uma das características da proposta apresentada neste trabalho é a integração de diferentes técnicas de Inteligência Artificial, com o intuito de criar um modelo que seja capaz de identificar as capacidades e limitações de estudantes, baseado no seu nível de aquisição de conhecimento e comportamento durante a utilização de um ambiente educacional. Outra importante característica versa sobre a capacidade de representação do ME de forma independente de domínio. O conhecimento é representado por meio das RBs, ou seja, independente da ontologia. Caso seja necessário aplicar o modelo a outro domínio, basta substituir a rede e o modelo funcionará corretamente.

Quando essas características citadas são disponibilizadas em ambientes educacionais, eles se tornam mais adaptáveis e inteligentes, capazes de fornecer adaptabilidade e conteúdo personalizado às características e individualidades de cada estudante. Consequentemente, esses ambientes inteligentes possuem a capacidade de maximizar o sucesso e

Figura 6 – Três formas diferentes de visualização de desempenho: (a) Gráfico de Linha; (b) Gráfico de Barra; (c) Gráfico de Radar



minimizar o fracasso nas atividades desenvolvidas pelos estudantes.

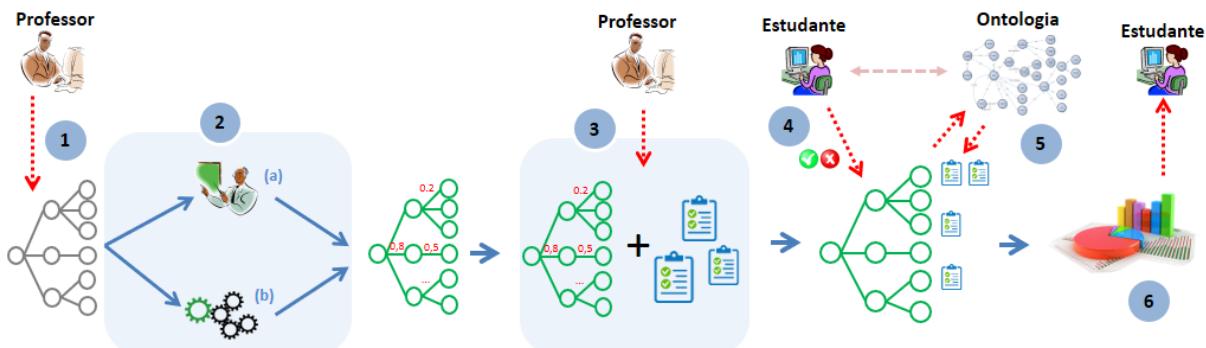
Por trabalhar com diferentes técnicas e recursos inteligentes, o processo de representação e inferência de informações se torna uma tarefa complicada. Com o intuito de facilitar o entendimento geral da proposta, serão apresentados detalhes de como ocorre o processo de descoberta do desempenho do estudante, desde a configuração do ambiente

pelo professor até a inferência e apresentação do conhecimento aos usuários.

O processo de inferência sobre o desempenho do estudante se inicia com a criação de uma RB para um domínio de conhecimento. Depois disso, instrumentos avaliativos são associados para medir o nível de aquisição de conhecimento de cada estudante em determinado tópico da rede. A partir das inferências realizadas pela RB, essas informações são repassadas para a ontologia, a qual irá interpretá-las juntamente com as outras já obtidas pelo modelo. Com todas essas informações, a ontologia utilizará as regras criadas para inferir o desempenho global do estudante. Vale ressaltar que esse modelo é completamente dinâmico, ou seja, a cada interação do estudante com o ambiente, as variáveis são atualizadas e repassadas para a ontologia, que, através de um conjunto de regras SWRL, se encarregará de refazer as inferências com base nos dados atualizados.

A Figura 7 expõe o fluxo de interações realizadas entre professores, estudantes e modelo desde a definição da rede até a inferência final sobre nível de desempenho do estudante.

Figura 7 – Fluxo de interações entre professores, estudantes e modelo



- **Fase 1:** A Fase 1 é o procedimento inicial para inserção de um novo domínio de conhecimento ao modelo. É nessa fase que o professor determina as relações e dependências da RB. Aqui também podem ser informados os valores de conhecimento *a priori* de cada estudante. Vale ressaltar que o modelo não é restrito a um único domínio de conhecimento. Caso seja necessário, vários domínios podem ser inseridos e, mesmo assim, o processo ocorrerá corretamente.
- **Fase 2:** A Fase 2 se inicia logo após a finalização da primeira fase. Nesse ponto, é preciso definir as tabelas de probabilidades para cada vértice na rede. Existem duas possibilidades para isso: (a) o próprio professor determina os valores com base em seu conhecimento empírico, ou, (b) um sistema especialista define, automaticamente, esses valores com base na configuração da rede (Discutido na Seção 3.2.3).
- **Fase 3:** Após toda a configuração inicial de inserção do domínio de conhecimento no ME, é preciso definir os mecanismos para aquisição de nível de conhecimento. Como mecanismos, podem ser utilizados testes, avaliações, exercícios, entre outros.

O tipo de mecanismo a ser utilizado fica a cargo da implementação (ambiente) e não do modelo (Discutido na Seção 3.2.4).

- **Fase 4:** Essa é a primeira fase em que os estudantes entram em contato com o modelo. Aqui o estudante precisa interagir com os instrumentos de aquisição de conhecimento para que esses valores sejam repassados para a RB. Esses instrumentos são associados com os vértices evidências na rede, que, por sua vez, geram valores que representam o nível de conhecimento do estudante nos diversos tópicos na rede.
- **Fase 5:** Essa é a fase mais minuciosa no processo de descoberta do nível de desempenho. Essa fase está fortemente relacionada com a Fase 4, onde após cada interação com os instrumentos avaliativos, o modelo reajusta os pesos na RB e atualiza todos os atributos das classes dependentes na ontologia. Essa fase é responsável por fazer o cruzamento das informações entre a RB, as interações no ambiente e a ontologia.
- **Fase 6:** Após todo o processamento das informações na Fase 5, a Fase 6 fica encarregada de gerenciar a apresentação dos resultados para estudantes e professores. Nessa fase, teorias relacionadas à Visualização de Informação são processadas com o intuito de apresentar o conteúdo que melhor se adapte ao perfil do usuário.

Esse é o principal fluxo de interações para conseguir uma visão global do funcionamento do modelo.

Sobre a arquitetura do modelo, é possível destacar alguns pontos-chave para entendimento da abordagem. A Figura 8 destaca os principais componentes que participam de todas as interações no processo de identificação de desempenho dos estudantes. O modelo é composto por três módulos principais (Módulo Probabilístico, Módulo Semântico e Módulo de Visualização).

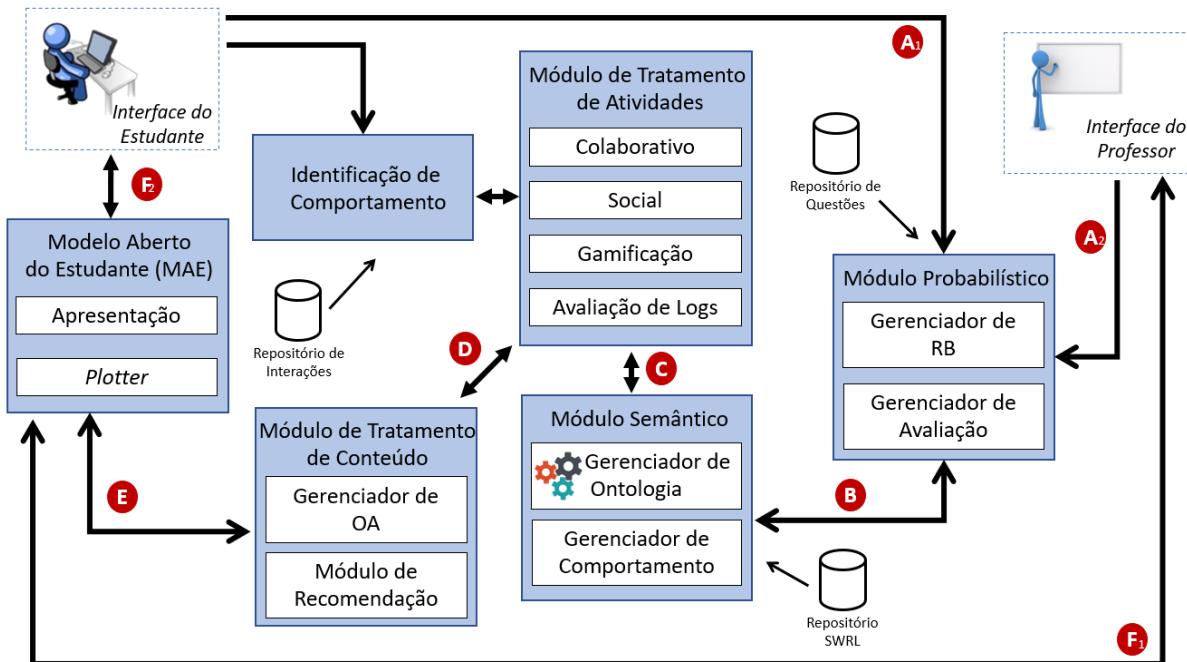
O Módulo Probabilístico é responsável por avaliar o nível de conhecimento do estudante com base em conteúdos avaliativos disponibilizados por professores. Este módulo recebe informações do estudante (A_1) e do professor (A_2) para alimentar as informações na RB e sobre o nível de aquisição de conhecimento.

Após o processamento por meio das RBs, as informações são enviadas para o Módulo Semântico (B), responsável por coletar as informações de comportamento – atividades de colaboração e socialização, gamificação e logs do sistema (C) e inferir, com base nas regras SWRL, o nível de desempenho do estudante. Neste ponto é interessante destacar que o Módulo de Tratamento de Atividades não faz parte do modelo, é uma abordagem para apoio ao funcionamento do mesmo.

Após a inferência pelo Módulo Semântico, as informações processadas são repassadas para um Módulo de Tratamento de Conteúdo (D – parte integrante do ambiente educacional e não do modelo proposto). Este módulo faz o tratamento das informações

educacionais a serem apresentadas aos estudantes – como, objetos de aprendizagem, recomendação, etc – e transfere (*E*) tais informações para o Módulo de Visualização (MAE). Essas informações são passadas, por meio de um documento *JSON* (*JavaScript Object Notation*) produzido dinamicamente. Essa informação personalizada é repassada para o “*Plotter*” responsável por fazer a plotagem dos gráficos e informações para os usuários (*F*₁ e *F*₂).

Figura 8 – Arquitetura do modelo.



O funcionamento do componente é dinâmico e segue uma abordagem que abstrai as características de implementação do ambiente onde será integrado, propiciando assim, fácil inclusão em qualquer ambiente educacional.

3.6 Considerações Finais

Este capítulo apresentou uma abordagem híbrida para modelagem de estudante em SAIEs. A proposta integra recursos de Inteligência Artificial com o intuito de criar um modelo que seja adequado para avaliar o desempenho de estudantes baseando-se em nível de conhecimento e comportamento durante a utilização de um ambiente virtual de ensino.

O modelo proposto possui as seguintes características: (*i*) é *dinâmico*, o modelo é atualizado a cada interação do usuário com o ambiente, permitindo, com isso, acompanhamento constante das características dos estudantes; (*ii*) é *probabilístico*, o modelo baseia-se em forte fundamento probabilístico (RBs) para avaliar a probabilidade de conhecimento em cada tópico do conteúdo estudado; (*iii*) é *extensível* e *reutilizável*, o modelo

utiliza ontologias, as quais permitem reúso e extensão das informações representadas; *(iv)* é *independente de domínio*, o domínio de conhecimento, representado por meio de RBs, atua separadamente das informações a respeito do estudante, representadas por meio das ontologias. Isso possibilita que o modelo seja integrado a qualquer domínio de conhecimento; e por fim, *(v)* *fornecer um analisador de desempenho*, ou seja, o modelo integra métricas comportamentais para avaliar, além do conhecimento, o desempenho (conhecimento + comportamento) dos estudantes.

Este capítulo apresentou também uma extensão do modelo criado, dando-lhe características de MAEs, capaz de tratar as informações inferidas e apresentar um *feedback* da atuação dos estudantes para os usuários do ambiente. O modelo baseia-se em teorias de Visualização de Informação com o intuito de “abrir” as informações inferidas para interpretação e manipulação dos estudantes e professores. A extensão permite que o professor acompanhe mais de perto o processo de aprendizagem do estudante, ao passo que o próprio estudante pode autorregular seu comportamento de estudo, buscando um desempenho melhor nas disciplinas que utilizam o modelo proposto.

CAPÍTULO 4

Estudo de Caso - Integração da Proposta a um Ambiente Real de Ensino

4.1 Considerações Iniciais

Com o objetivo de avaliar a proposta apresentada junto a um ambiente real, a mesma foi implementada e integrada a uma plataforma de ensino utilizada na Faculdade de Computação da Universidade Federal de Uberlândia. O Classroom eXperience (CX) (FERREIRA et al., 2012; ARAÚJO et al., 2013), ambiente utilizado na integração, se baseia em recursos de Computação Ubíqua e visa dar suporte a professores e estudantes em suas diversas tarefas de ensino/aprendizagem.

O ambiente foi escolhido devido a sua capacidade de permitir interações entre professores e estudantes por ser um ambiente que possui capacidades colaborativas, sociais e de gamificação, também pelo fato de serem raros os casos de integração com recursos de tutoria inteligente a Ambientes Educacionais Ubíquos (AEUs). Dessa forma, essa integração permitiu verificar o correto funcionamento do Modelo de Estudante (ME) proposto, bem como a sua capacidade de integração em ambientes com características ubíquas.

Como o modelo proposto atua separadamente da plataforma a ser integrada, seus recursos podem ser implementados por meio de módulos fracamente acoplados. Esses módulos possuem interdependência baixa e podem ser adicionados à plataforma de maneira mais simples e mais fácil. A integração aconteceu em três fases: inicialmente, um módulo de avaliação de conhecimento (baseado em Redes Bayesianas (RBs) para definição do domínio e avaliação das probabilidades de conhecimento) foi criado e integrado; posteriormente, aconteceu a integração de um segundo módulo, responsável pela inferência do nível de desempenho (utilizando ontologias para modelagem das características dos estudantes e definição do nível de desempenho com base na análise de conhecimento e comportamento do estudante). A terceira fase diz respeito à integração de recursos que

permitem que o modelo atue como um Modelo Aberto do Estudante (MAE). Todos os detalhes de implementação e integração serão expostos neste capítulo.

4.2 Ambiente de Integração

Um fato que auxiliou na melhoria das atividades educacionais foi a integração de recursos computacionais ao ambiente de ensino com o intuito de auxiliar professores durante sua atividade de docência. A integração de recursos computacionais ao ambiente de forma não intrusiva e favorecendo a heterogeneidade dos dispositivos recebe o nome de Computação Ubíqua. A Computação Ubíqua foi proposta por Weiser (1991) e tem como principal premissa a integração, de forma transparente, de recursos computacionais no cotidiano das pessoas com o intuito de auxiliar na realização de suas atividades. A Computação Ubíqua pode ser subdividida em algumas áreas, sendo uma delas a Captura e Acesso (C&A), um tema de pesquisa que tem como foco utilizar ferramentas especializadas para capturar informações do ambiente e prover recursos para os usuários acessarem esse conteúdo posteriormente.

Os ambientes educacionais são um dos cenários propícios para a aplicação de técnicas de C&A. Diversos trabalhos propõem ferramentas para apoiarem professores e estudantes na realização de suas atividades (TRUONG; HAYES, 2009; FERREIRA et al., 2012; DICKSON et al., 2012). Nesses ambientes, geralmente, são empregados mecanismos para capturar diferentes fluxos de mídias e disponibilizar essas informações, posteriormente, aos estudantes, com o intuito de auxiliá-los no processo de aprendizagem.

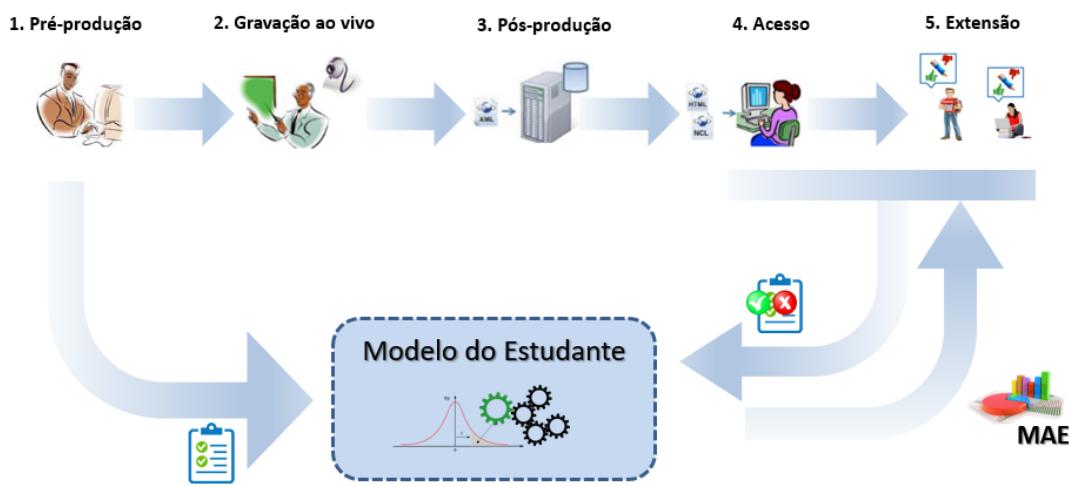
Uma das propostas para apoiar atividades em sala de aula consiste no uso de AEUs (SETTLE; DETTORI; DAVIDSON, 2011) – espaços físicos equipados com recursos computacionais (sensores, lousas eletrônicas, câmeras, microfones, entre outros), tendo o potencial de incorporar a visão de Computação Ubíqua à área de educação. Quando associados a sistemas automatizados de captura multimídia, os AEUs proporcionam o registro e o armazenamento de artefatos de mídia que, uma vez combinados e sincronizados, são capazes de reproduzir, no futuro, as experiências vivenciadas em sala de aula. Nesses ambientes, os estudantes podem concentrar a atenção na experiência da aula, com a confiança de que os detalhes estão sendo registrados e estarão disponíveis para o posterior acesso.

Quando empregadas a ambientes educacionais, as aplicações de C&A tendem a seguir uma sequência bem estruturada, composta de quatro fases, proposta por Abowd et al. (1997): pré-produção, gravação ao vivo, pós-produção e acesso. Algumas extensões para essa sequência têm sido propostas ao longo dos anos. Em especial, uma extensão que merece atenção foi definida por Pimentel et al. (2001), na qual uma nova fase é incluída, a fase de extensão, permitindo maior interação entre estudantes e conteúdo educacional a fim de enriquecer o material previamente capturado.

Com a incorporação do ME proposto neste trabalho, as arquiteturas voltadas para

os AEUs que utilizam a proposta de Abowd ganham mais dinamicidade e inteligência no processo de ensino, pois passam a ter mais conhecimento e controle sobre as características e o comportamento dos estudantes. O modelo adiciona uma importante característica, na qual todo o processo de apresentação do conteúdo passa a ser guiado por meio de mecanismos inteligentes fornecidos pelo ME. A Figura 9 exibe um esboço das interações entre as fases da arquitetura e o modelo proposto. É interessante notar que o ME interage com as fases propostas por Abowd trazendo uma visão mais inteligente do processo de captura, permitindo com isso, uma infraestrutura mais adequada para auxiliar no processo de ensino/aprendizagem.

Figura 9 – Incorporação do modelo proposto na arquitetura dos AEUs.



Um exemplo de AEU é o CX¹ (FERREIRA et al., 2012; ARAÚJO et al., 2013), uma plataforma multimídia para captura de aulas em uma sala instrumentada com dispositivos computacionais ubíquos, tais como lousa eletrônica, microfones, câmeras de vídeo e projetores. Desenvolvido para registrar, armazenar, sincronizar e disponibilizar as diferentes mídias capturadas, o CX gera documentos hipermídia em vários formatos de apresentação e que podem ser empregados para reconstituir a experiência vivida em sala de aula.

Por se tratar de uma plataforma baseada em conceitos de AEUs, o CX realiza a captura da aula de maneira automática, transparente e não intrusiva, empregando recursos de C&A para auxiliar tanto professores durante o processo de captura, quanto estudantes no acesso ao conteúdo capturado. O sistema conta com um arcabouço de hardware e software especializado para realizar a gravação das atividades educacionais, sincronizar os fluxos de mídia e disponibilizar o conteúdo por meio de uma interface de acesso amigável e intuitiva. A Figura 10 exibe duas salas de aula instrumentadas, disponibilizadas para uso na Faculdade de Computação (FACOM) da Universidade Federal de Uberlândia (UFU).

Desde a sua versão inicial (FERREIRA et al., 2012), novas extensões vêm sendo criadas a fim de aprimorar os recursos para melhoria da plataforma. Uma extensão foi

¹ <http://cx.facom.ufu.br/>

Figura 10 – Salas de aula instrumentadas utilizadas pela plataforma CX (FERREIRA et al., 2012; ARAÚJO et al., 2014).



proposta por Araújo et al. (2013), em que o CX utiliza o contexto de acesso do estudante, suas preferências pessoais e as restrições de apresentação do dispositivo de acesso para personalizar o conteúdo individualmente para cada estudante. Outra extensão foi proposta por Mendonça et al. (2014), quando foram adicionadas características de colaboração e interação social ao ambiente de acesso. Isso permite aos usuários realizarem comentários sobre o conteúdo capturado e classificação (*rating*) das informações que acharem mais relevantes. Essas interações sociais propiciam uma integração entre os estudantes do ponto de vista colaborativo, na qual um estudante pode enriquecer um determinado conteúdo, contribuindo, assim, para o aprendizado dos demais.

A última extensão incorporada pelo CX foi proposta por Ferreira et al. (2015), na qual recursos de gamificação foram incorporados ao ambiente com o intuito de melhorar o engajamento e a dinâmica de utilização pelos estudantes. Diversos mecanismos de *games* foram adicionados à plataforma proporcionando aos usuários acumular pontos, conquistar medalhas (*badges*) e criar competições entre seus colegas de turma.

O CX é hoje uma plataforma multimídia fundamentada em Computação Ubíqua que está em pleno uso em ambientes reais de ensino. Desde 2012, em sua primeira versão, o CX tem se mostrado muito eficiente e eficaz para auxiliar atividades de ensino em ambiente instrumentado. Referências sobre essa plataforma podem ser encontradas no site do grupo de pesquisa UbiMedia².

4.3 Integração do Modelo ao Ambiente

Esta seção descreve detalhes sobre a integração do modelo proposto ao ambiente CX. Como o modelo possui algumas atividades bem definidas, foi possível dividir a integração em três fases: integração dos recursos probabilísticos, integração dos recursos semânticos e integração dos recursos de visualização.

² <http://ubimedia.facom.ufu.br/>

4.3.1 1^a Fase - Módulo Probabilístico

Esta primeira fase diz respeito à integração dos recursos capazes de realizar inferência acerca do nível de conhecimento do estudante. Essa inferência é realizada por meio de uma RB que recebe como entrada informações sobre o nível de conhecimento do estudante sobre alguns tópicos da rede e, por meio de cálculos de probabilidade condicional, é capaz de inferir o nível nos demais tópicos de conhecimento. A implementação dessa fase acontece mediante dois módulos: o Módulo de Inserção da RB e o Módulo de Inserção dos Instrumentos Avaliativos.

Para trabalhar com as RBs foi utilizada a biblioteca *SMILE Engine*³, uma plataforma de inferência para modelos gráficos, também utilizada para diagramas de influência e modelos de equações estruturais. Esta é a principal biblioteca disponível na literatura para manipulação de RBs. O projeto é mantido pela Escola de Ciência da Informação e o Programa de Sistemas Inteligentes da Universidade de Pittsburgh, Estados Unidos.

4.3.1.1 Rede Bayesiana

Inicialmente, foi criado um mecanismo que permitisse aos professores associarem uma RB a uma disciplina. É possível perceber que o modelo não se limita a um determinado domínio de conhecimento, pois é possível adicionar quantas disciplinas forem necessárias, e cada uma, com uma rede específica. Para este estudo de caso foi implementado um mecanismo que interpreta a RB no formato *GeNIe Network (.xdsl)*⁴.

Uma vez inserida a RB, o módulo interpreta todos os tópicos do domínio de conhecimento e insere-os em uma base de dados relacional, que possibilita um rápido e fácil acesso posteriormente.

4.3.1.2 Instrumentos Avaliativos

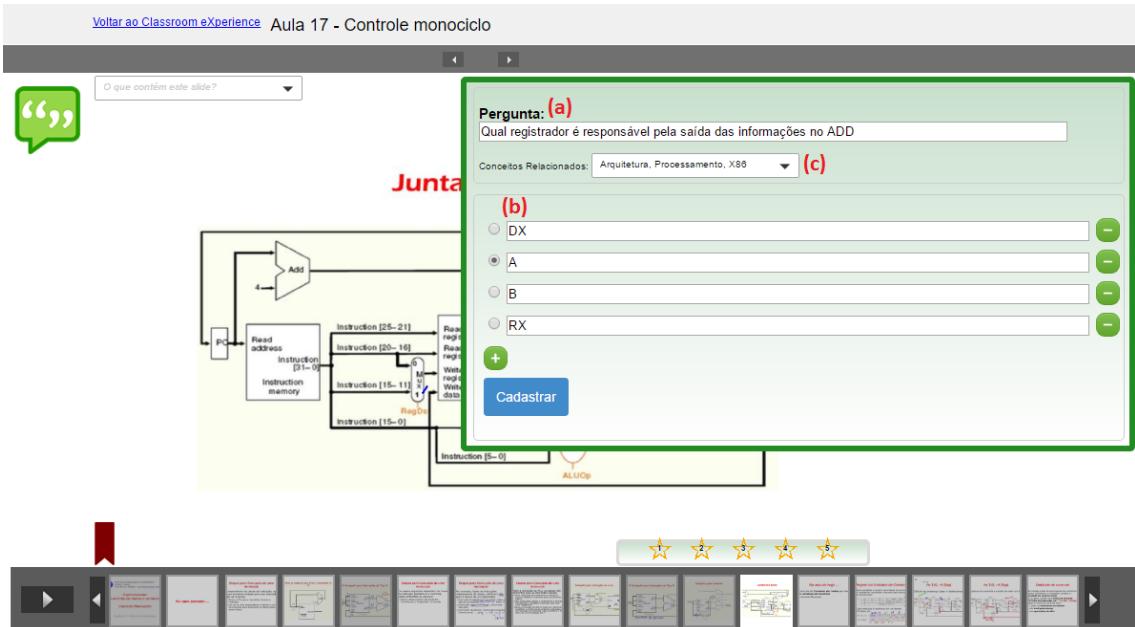
Após a criação do mecanismo que auxilia na inserção da RB ao ambiente, foi necessário criar um mecanismo que permitisse a definição de instrumentos avaliativos. Esses instrumentos são utilizados para avaliar o conhecimento do estudante em determinado tópico do conteúdo e repassar essas informações para a RB.

Neste estudo de caso, foi implementado um instrumento avaliativo baseado em questões de múltipla escolha por meio de *quizzes* que são cadastrados pelos professores em determinados pontos da aula. Tanto os *quizzes* quanto as respostas dos estudantes são armazenados em uma base de dados relacional. A Figura 11 apresenta uma tela para cadastro de um *quiz* no ambiente CX. Durante o cadastro de um *quiz* o professor determina a pergunta (a), as alternativas (b) e quais são os tópicos da RB associados a este *quiz* (c).

³ Projeto BayesFusion, LCC - <https://www.bayesfusion.com/smile-engine>

⁴ Projeto BayesFusion, LCC - <https://www.bayesfusion.com/genie-modeler>

Figura 11 – Criação de um instrumento avaliativo na plataforma CX.



Para essa implementação também foi determinado um cálculo para quantificar a pontuação de um *quiz*, que se baseia na quantidade de tentativas que o estudante realizou até o acerto para definir o seu nível de conhecimento. O cálculo é dado pela Equação (1).

$$PC = \frac{Max}{Q - 1} * (Q - N) \quad (1)$$

Onde:

PC é a probabilidade de conhecimento;

Max é a probabilidade máxima de conhecimento;

Q é a quantidade de alternativas para a questão; e

N é a tentativa em que o estudante obteve sucesso (acertou).

É possível notar na Equação (1) que a Probabilidade de Conhecimento (PC) está relacionada diretamente com as tentativas que o estudante utilizou para acertar a questão (N). Por exemplo, o instrumento avaliativo apresentado na Figura 11 possui quatro alternativas (Q), e o valor Max para PC , definido pelo modelo, é de 0,9. Assim, se o estudante acertar a questão na primeira tentativa o valor de PC será igual a 0,9. Caso o estudante acerte na segunda tentativa, o valor de PC será igual a 0,6. Caso ele acerte na terceira ou quarta tentativa, seu PC será de 0,3 e 0,1, respectivamente. Neste último caso, nota-se que o estudante não adquire 0,0, pois, como determinação do modelo, o valor para PC deve sempre variar de 0,1 a 0,9.

Vale ressaltar que o módulo proposto não restringe o tipo de instrumento a ser utilizado. Para este estudo de caso foram utilizados *quizzes*, no entanto, outros instrumentos poderiam ser implementados. É importante destacar também que o cálculo para a avaliação do *quiz* pode ser definido na implementação.

Com os módulos de inserção de RBs e instrumentos avaliativos já é possível realizar as inferências sobre o nível de conhecimento do estudante. A partir do momento em que os estudantes se matriculam na disciplina, é criada uma rede abstrata para cada estudante, a qual irá representar o seu conhecimento global. Essa rede é atualizada a cada interação com os instrumentos avaliativos. Para saber a probabilidade de conhecimento de um determinado estudante, basta verificar qual o valor representado no nó da rede abstrata do estudante em questão.

4.3.2 2^a Fase - Módulo Semântico

Após a integração do Módulo Probabilístico, foi possível iniciar a integração da segunda parte do modelo, o Módulo Semântico. A parte semântica é utilizada para representar as características dos estudantes e inferir sobre o seu nível de desempenho global, tomando como base as informações inferidas pelo modelo probabilístico juntamente com as informações sobre o comportamento do estudante.

Para o desenvolvimento desse módulo, foi utilizado o *framework Apache Jena*⁵. Ambiente de código fonte aberto construído em Java que permite trabalhar com aplicações de Web Semântica e dados conectados. Este é o *framework* de construção de ambientes semânticos mais utilizado pela comunidade acadêmica. Também foi utilizada a OWL API⁶, uma API JAVA que permite criação, manipulação e serialização de ontologias OWL. O Módulo Semântico é capaz de trocar informações com o CX possibilitando que as informações necessárias para representação das características dos estudantes sejam armazenadas, tanto na base original do CX quanto na estrutura ontológica representada pelo modelo.

Além da representação das características dos estudantes, este módulo também é composto por um componente *reasoner*, utilizado para realizar inferências a partir de regras *Semantic Web Rule Language* (SWRL). Após alguns testes para verificar a capacidade de processamento das inferências e a capacidade de comunicação com o *framework Jena*, optou-se pela utilização do *reasoner Pellet*⁷. Esse tem se mostrado muito eficiente e conciso para o tratamento de inferências baseando-se nas regras SWRL.

O Módulo Semântico recebe como entrada para o processamento dois conjuntos de informações: um decorrente do Módulo Probabilístico e outro vindo do ambiente CX, e produz informações que são repassadas para o terceiro módulo, o Módulo de Visualização (que será descrito na próxima seção). A Figura 12 ilustra o fluxo de interações em torno do Módulo Semântico. Sempre que o estudante interage com o CX (1), suas atividades são capturadas e repassadas para o Módulo Probabilístico (2), este se encarrega de processar informações sobre o nível de conhecimento do estudante e repassar para o Módulo Semântico (3) que, por sua vez, também busca informações sobre o comportamento do

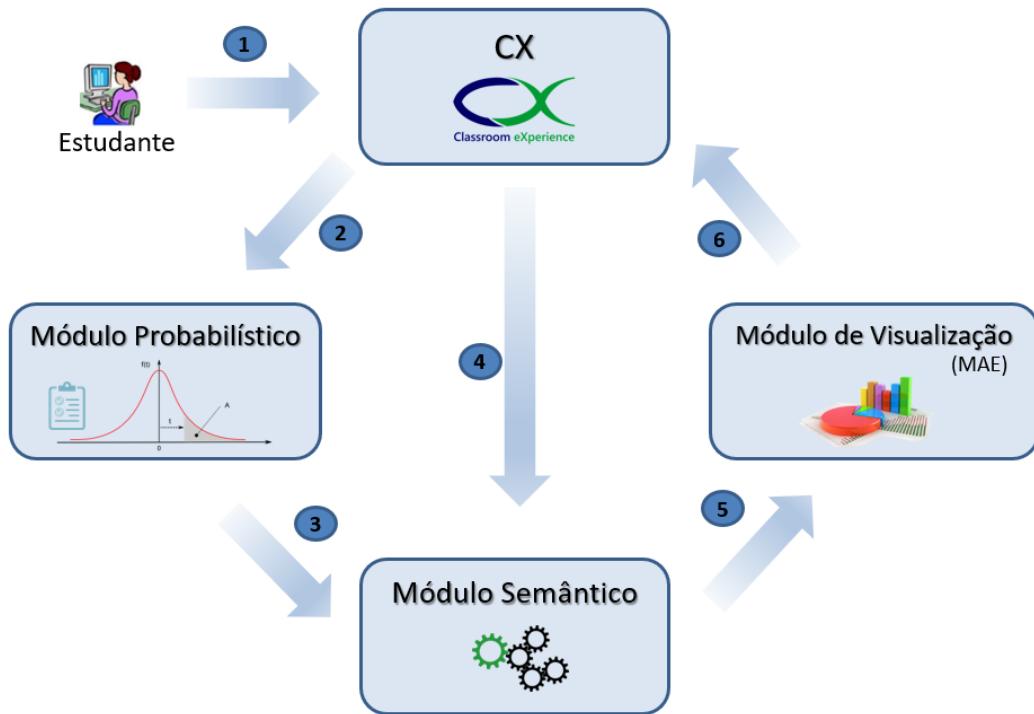
⁵ <https://jena.apache.org/>

⁶ <https://github.com/owlcs/owlapi/wiki>

⁷ <https://www.w3.org/2001/sw/wiki/Pellet>

estudante no CX (4). Com todas essas informações o Módulo Semântico utiliza as regras SWRL para realizar inferências sobre o nível de desempenho do estudante e repassar essas informações para o Módulo de Visualização (5), responsável por adaptar o conteúdo a ser apresentado no CX (6).

Figura 12 – Interações entre os módulos desenvolvidos no modelo proposto.



Integrado à ontologia que representa as características dos estudantes está um banco de regras SWRL, o qual permite que as inferências sejam processadas pelo modelo. É importante destacar que tais regras podem ser manipuladas de acordo com a necessidade e particularidade de cada ambiente.

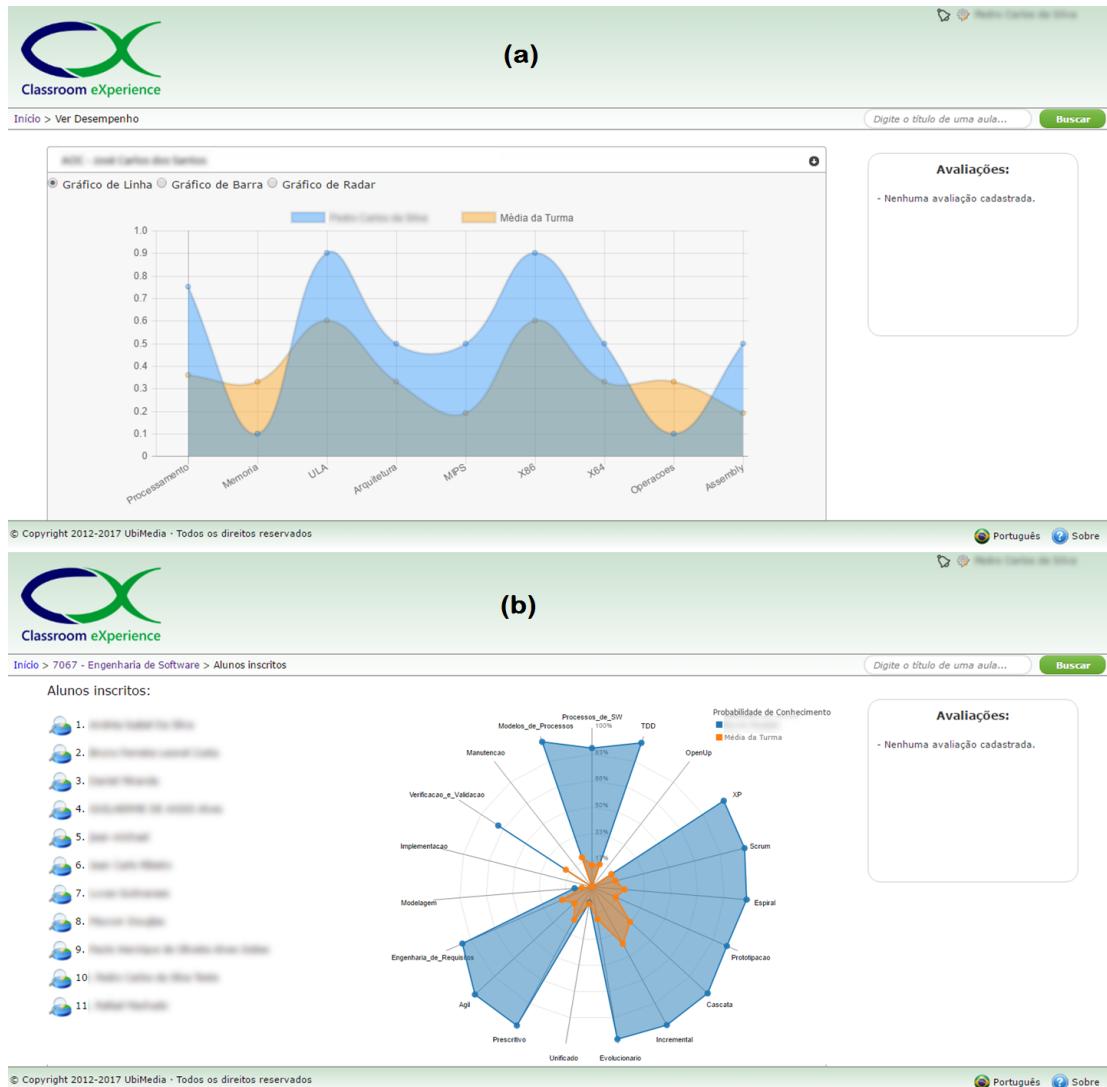
4.3.3 3^a Fase - Módulo de Visualização

Esta seção apresenta a parte final da integração do modelo proposto ao ambiente CX, o Módulo de Visualização de informações. Este é um importante módulo, pois agrega a capacidade de visualização de conhecimento ao ME proposto, permitindo que o modelo em questão seja tratado como um MAE.

Este módulo foi construído com o intuito de “externalizar” as informações inferidas pelo ME, proporcionando que estudantes e professores visualizem tais informações. Na Figura 13 é possível observar duas interfaces disponibilizadas pelo CX para a visualização do desempenho do estudante. Na parte superior da imagem é apresentada a visualização disponível para um estudante; é possível observar os tópicos de avaliação (eixo X) e sua média (eixo Y). Também é possível visualizar a sua nota em relação à média da turma. A

parte inferior da imagem apresenta a visualização disponível para o professor, pela qual este tem acesso a todos os estudantes matriculados e seus respectivos desempenhos. Ele também pode visualizar a comparação da nota do estudante com a média da turma. É oportuno frisar que esses recursos de visualização são personalizáveis.

Figura 13 – Interface para visualização disponível no CX: (a) Visualização do estudante, (b) Visualização do professor.



Para a construção desse módulo, utilizou-se de diversas tecnologias de *WEB 2.0* e *HTML 5*. A principal biblioteca utilizada para apresentação e plotagem das informações por meio de gráficos foi *Chart.js*⁸. Por ser baseada em *JavaScript*, a biblioteca permite fácil integração e comunicação com os mais diversos recursos Web disponíveis atualmente, além da geração de uma grande variedade de gráficos estáticos e dinâmicos.

Como pode ser visto na Figura 12, esse módulo (MAE) recebe como entrada as informações sobre o nível de desempenho inferidos pelo modelo ontológico (5), processa essas

⁸ <http://www.chartjs.org/>

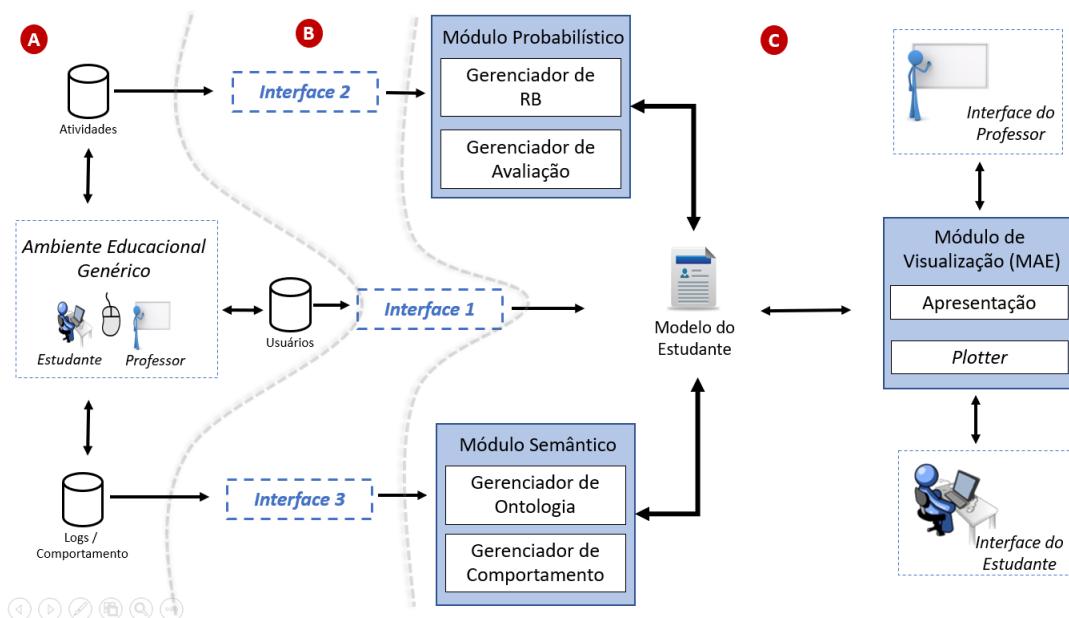
informações com base no perfil do estudante e gera um documento no padrão *JSON*⁹ para ser interpretado pela plataforma CX (6). Após isto, as informações estarão disponíveis para visualização.

É interessante notar que todo o processo de inferência sobre o nível de desempenho é concluído com o Módulo Semântico, no entanto, até este momento as informações ainda estão “fechadas” e só podem ser interpretadas internamente pelo modelo. Esse recurso de “abrir” essas informações permite que os usuários do sistema tenham acesso a essas informações, permitindo, assim, maior interação com o processo de inferência e definição do desempenho.

4.4 Modelo Genérico de Integração

Nesta seção, será apresentada uma arquitetura conceitual genérica que permite a integração do modelo proposto a outros ambientes. Esse modelo de integração é composto por três partes principais: (A) Componentes do Ambiente Educacional Genérico; (B) Interfaces de Comunicação; e (C), Modelo Híbrido para Sistemas Inteligentes (modelo proposto neste trabalho). Na Figura 14 é possível ver tal divisão.

Figura 14 – Arquitetura para integração do modelo proposto a um ambiente educacional genérico.



A arquitetura genérica de integração deve conter, pelo menos, essas três partes. A parte A da imagem (mais à esquerda) apresenta um Ambiente Educacional Genérico. Esse ambiente deve possuir, pelo menos, um repositório de usuários onde estão armazenados todos os estudantes do ambiente; um repositório de Atividades que contenha os instrumen-

⁹ <http://www.json.org/json-pt.html>

tos avaliativos, por exemplo, *quizzes*; e por fim, um repositório de Logs/Comportamento, onde estão armazenadas todas as interações dos estudantes com o material didático.

Na parte C da imagem (mais à direita) são apresentados os módulos principais do modelo: Módulo Probabilístico, Módulo Semântico e Módulo de Visualização¹⁰. Esses módulos são responsáveis por todas as atividades realizadas pelo modelo aqui proposto.

Por fim, na parte B da imagem (centro) é possível observar as interfaces de comunicação entre o ambiente genérico e o ME proposto. A Interface 1 está relacionada com toda a comunicação dos usuários do sistema e à parte central do modelo. A Interface 2 ilustra uma ponte de comunicação entre as atividades avaliativas e o Módulo Probabilístico e, a Interface 3 é responsável pela comunicação entre os repositórios de interações (Logs/Comportamentos) e o Módulo Semântico.

Como apresentado, caso seja necessário realizar a alteração do ambiente educacional, basta que novas interfaces de comunicação sejam criadas com vistas às particularidades do novo ambiente a ser integrado. Vale destacar que essa é uma arquitetura genérica e foi apresentada como uma ilustração para o processo de integração. A complexidade de implementação das interfaces depende diretamente da organização das informações do ambiente genérico.

4.5 Considerações Finais

Este capítulo apresentou um estudo de caso com o intuito de avaliar a viabilidade de integração do modelo proposto a um ambiente real de ensino. Para isso, foi escolhido o CX, um ambiente de aprendizagem fundamentado em recursos e métodos da Computação Ubíqua que visa a captura, integração e apresentação de conteúdo educacional.

A integração aconteceu por meio da construção de três módulos – Módulo Probabilístico, Módulo Semântico e Módulo de Visualização de informação – nos quais foi possível implementar as características propostas pelo modelo. Estes módulos foram fracamente acoplados ao ambiente CX, permitindo assim, poucas alterações na estrutura já existente.

A integração permitiu atribuir características inteligentes ao ambiente CX, possibilitando identificar as reais capacidades dos estudantes tomando como base o seu conhecimento e o seu comportamento. A integração permitiu também que professores acompanhassem mais de perto o progresso dos seus estudantes, proporcionando, dessa forma, um conhecimento mais global e constante sobre a evolução no aprendizado da sua turma.

¹⁰ Responsável pelas atividades do MAE

CAPÍTULO 5

Experimentos e Análise dos Resultados

5.1 Considerações Iniciais

Esta seção apresenta experimentos cujo intuito é verificar a eficiência do modelo proposto¹. Os experimentos versam sobre diferentes análises, buscando, com isso, uma ampla verificação da corretude acerca dos valores determinados pelo modelo, bem como sobre o impacto da proposta no comportamento dos usuários que o utilizam. Um diferencial desses experimentos é o fato de que todos foram realizados com estudantes reais e aplicados a diferentes disciplinas em diferentes cursos.

O primeiro experimento teve como objetivo verificar a correlação entre os valores inferidos pelo Módulo Probabilístico e os níveis de conhecimento dos estudantes. Este experimento contou com a participação de 129 estudantes de cinco turmas durante três semestres (2015-2, 2016-1 e 2016-2). Parte dos resultados deste experimento foi publicada em: Ferreira et al., 2016a e Ferreira et al., 2016b, disponíveis na lista de publicações deste trabalho.

O segundo experimento buscou identificar quais são as propriedades ou características de uma Rede Bayesiana (RB) que influenciam positivamente na inferência de valores sobre o conhecimento do estudante. Este experimento foi realizado durante três semestres (2016-1, 2016-2 e 2017-1) e contou com a participação de 166 estudantes.

O terceiro experimento teve como objetivo apresentar uma análise do desempenho dos estudantes com base em perfis de comportamento. Este experimento analisou se os estudantes que se comportam de maneira diferente durante a utilização de um Sistema Adaptativo e Inteligente para Educação (SAIE), alcançaram médias de desempenho também diferentes. Participaram deste experimento um total de 119 estudantes, divididos em cinco turmas ao longo de três semestres (2016-2, 2017-1 e 2017-2). Parte dos resultados deste experimento foi publicada em: Ferreira et al., 2017c, disponível na lista de publicações deste trabalho e outra parte foi submetida recentemente para a revista *Transaction on Education (TOE)*, a qual, ainda se encontra em avaliação.

¹ Projeto aprovado no Comitê de Ética da UFU: CAAE 46909515.4.0000.5152

O quarto experimento teve como objetivo analisar o impacto na utilização de um Modelo Aberto do Estudante (MAE) associado a uma plataforma ubíqua. Este foi um experimento qualitativo, que buscou identificar, entre outras, questões associadas à satisfação e alteração no comportamento dos estudantes. Este experimento contou com a participação de 139 estudantes ao longo de dois semestres (2017-1 e 2017-2). Parte dos resultados deste experimento foi publicada em: Ferreira et al., 2017b, disponível na lista de publicações deste trabalho.

A descrição completa dos experimentos, métodos e resultados é apresentada a seguir.

5.2 Experimento 1: Análise dos Valores Inferidos pelo Módulo Probabilístico

Este primeiro experimento visa avaliar a eficiência do Módulo Probabilístico, parte essencial para o Modelo de Estudante (ME) proposto. O Módulo Probabilístico é responsável pela inferência da probabilidade de conhecimento do estudante a partir de instrumentos avaliativos disponibilizados na plataforma educacional. A tarefa desenvolvida por esse módulo é essencial para que o resultado final corresponda, realmente, às capacidades e limitações dos estudantes. Caso essa tarefa não funcione adequadamente, todo o processo posterior pode estar comprometido.

Com base nisso, esta seção apresenta experimentos para avaliar a correlação entre os valores de probabilidade de conhecimento inferidos pelo Módulo Probabilístico e os valores reais de conhecimento dos estudantes.

5.2.1 Método Utilizado

O experimento proposto foi conduzido ao longo de três semestres (2015-2, 2016-1 e 2016-2) para cinco diferentes turmas: quatro de graduação (curso de Ciência da Computação e Sistemas de Informação) e uma de pós-graduação em Ciência da Computação. No total, 129 estudantes participaram do estudo. Uma RB foi criada e configurada pelo professor de cada turma do experimento, com o intuito de cobrir os principais tópicos abordados. É interessante observar que o professor de cada disciplina possuía conhecimentos suficientes para a criação das redes. Destaca-se também que os valores “reais” ditos aqui correspondem às notas obtidas por meio de avaliações periódicas, ou seja, as provas aplicadas pelo professor da disciplina. A Tabela 4 traz detalhes sobre as turmas utilizadas no experimento.

O principal foco deste experimento é verificar a correlação entre as notas reais e as probabilidades de conhecimento inferidas pelo modelo proposto. Para isso, o processo de avaliação do conhecimento do estudante foi cuidadosamente elaborado, a fim de obter um resultado condizente com suas capacidades. Tanto as RBs quanto as avaliações fo-

Tabela 4 – Turmas participantes do Experimento 1.

Semestre	Disciplina	Curso	Qtd ¹
T1	2015-2	Programação Procedimental	20
T2	2016-1	Interação Humano-Computador	35
T3	2016-1	Interação Humano-Computador	11
T4	2016-2	Interação Humano-Computador	33
T5	2016-2	Arq. e Org. de Computadores	30

¹Quantidade de estudantes na turma.

ram elaboradas buscando uma diferenciação entre as turmas. Algumas RBs foram mais detalhadas, apresentando conceitos de alto e baixo nível, e outras foram mais sucintas, pautando-se apenas em conceitos de alto nível. Quanto às avaliações, algumas foram mais extensas e outras mais curtas. Algumas avaliações abrangeram todos os tópicos da rede, outras, por sua vez, foram mais limitadas. Todos esses detalhes foram implementados para buscar maior heterogeneidade nos dados.

As avaliações foram compostas por questões ($Q_1, Q_2, Q_3, \dots, Q_n$) que foram relacionadas aos tópicos ($T_1, T_2, T_3, \dots, T_m$) das RBs. Alguns tópicos (T) foram classificados como *Tópicos de entrada (IN)* e outros como *Tópicos de saída (OUT)*. *INs* foram usados para receber informações sobre o conhecimento dos estudantes, ou seja, como entrada para a RB. Por outro lado, *OUTs* foram utilizados para visualizar os resultados produzidos pela RB. No Apêndice B é possível observar as RBs criadas para cada uma das disciplinas com seus respectivos tópicos *INs* e *OUTs*.

Após a aplicação das avaliações, as notas reais dos estudantes foram inseridas nos *INs*. Então, a RB processou a informação e gerou os valores inferidos para os *OUTs*. Após o processamento das inferências, os valores inferidos pelo modelo para os *OUTs* foram comparados com os valores reais alcançados por meio das avaliações. Desta forma, foi possível avaliar a correlação entre as notas reais dos estudantes e as notas inferidas pelo Módulo Probabilístico.

Antes de analisar a correlação entre os valores, o teste de *Shapiro-Wilk* foi realizado para verificar se as amostras seguiam uma distribuição normal. Dentre os testes de normalidade existentes, este foi escolhido por ser o mais adequado para pequenas amostras (RAZALI; WAH et al., 2011).

Quatro tópicos foram analisados para cada grupo e verificou-se que a maioria das amostras analisadas não apresentava normalidade de resíduos, como pode ser visto na Tabela 5. Como resultado, a correlação de posto de *Spearman* foi realizada para analisar a relação entre os valores reais obtidos pelos estudantes e os inferidos pela RB. Desta forma, o objetivo era verificar qualquer correlação entre as variáveis.

A Tabela 5 exibe o coeficiente de correlação de postos de *Spearman* entre os valores reais para as notas dos estudantes e os valores inferidos pela RB. De acordo com Davis

Tabela 5 – Coeficiente de correlação de postos de *Spearman* entre notas reais dos estudantes e as notas inferidas pela RB.

T ¹	Variáveis	W(P) ²	W(P) ³	r ⁴
T1	Expressões Condicionais	0.920 (0.097)	0.901 (0.043)	0.809
T1	Break	0.540 (0.000)	0.749 (0.000)	0.328
T1	Loops	0.885 (0.021)	0.900 (0.041)	0.544
T1	Caracteres	0.850 (0.005)	0.716 (0.000)	0.582
T2	Análise de Tarefas	0.000 (0.001)	0.887 (0.002)	0.973
T2	Protótipo de Baixa Fidelidade	0.591 (0.000)	0.000 (0.000)	0.797
T2	Teste de Usabilidade	0.000 (0.000)	0.869 (0.001)	0.859
T2	Entendimento	0.848 (0.000)	0.875 (0.001)	0.514
T3	Design Humano	0.868 (0.073)	0.742 (0.002)	0.963
T3	Teste de Usabilidade	0.804 (0.011)	0.823 (0.019)	0.608
T3	Antecipaçāo	0.863 (0.063)	0.741 (0.002)	0.977
T3	Design	0.896 (0.163)	0.742 (0.002)	0.933
T4	Computação Móvel	0.727 (0.000)	0.369 (0.000)	0.731
T4	Design Humano	0.000 (0.000)	0.876 (0.001)	0.321
T4	Protótipo de Baixa Fidelidade	0.479 (0.000)	0.903 (0.005)	0.437
T4	WIMP	0.000 (0.000)	0.430 (0.000)	0.437
T5	Caminho de Dados	0.901 (0.004)	0.894 (0.003)	0.642
T5	MIPS	0.914 (0.009)	0.943 (0.069)	0.757
T5	RISC	0.891 (0.002)	0.974 (0.568)	0.625
T5	Tempo de Resposta	0.857 (0.000)	0.955 (0.160)	0.505

¹Turma; ²W: Teste de Shapiro-Wilk para as notas reais; ³W: Teste de Shapiro-Wilk para as notas inferidas pela RB; ⁴r: Coeficiente de correlação de posto de Spearman.

(1971), certos adjetivos podem ser usados para interpretar a análise de correlação. Neste sentido, $r = 1$ indica uma correlação perfeita; $0.70 < r < 0.99$ representa uma correlação muito alta; $0.50 < r < 0.69$ indica uma correlação substancial; $0.30 < r < 0.49$ aponta uma correlação moderada; $0.10 < r < 0.29$ representa uma baixa correlação; e, $0.01 < r < 0.09$ indica uma correlação insignificante.

5.2.2 Resultados e Discussão

Os resultados apresentados a seguir estão agrupados por turma, no qual, as variáveis correspondem aos tópicos de cada uma das disciplinas ministradas. É possível notar que correlações muito altas foram encontradas para quase todas as amostras analisadas:

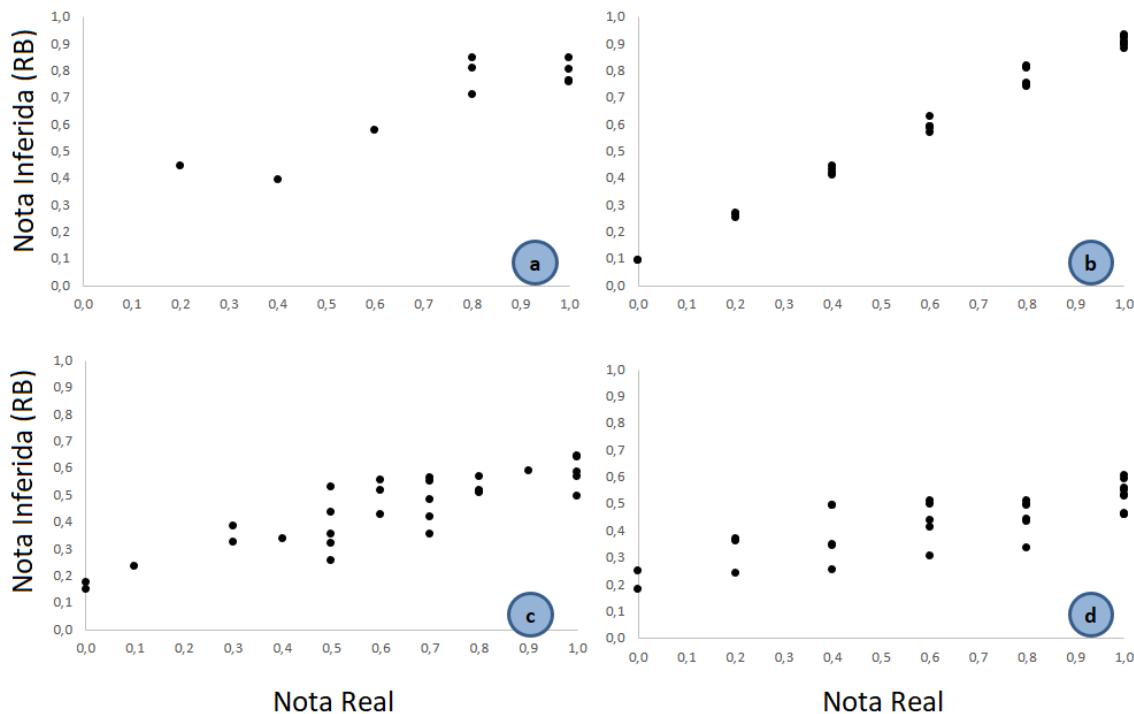
- ❑ Em T1, houve uma correlação muito alta relacionada à variável *Expressões Condicionais*, correlação substancial para *Loops* and *Caracteres* e correlação moderada para *Break*;
- ❑ Em T2, uma correlação muito alta foi obtida para as variáveis *Análise de Tarefas*, *Protótipo de Baixa Fidelidade* e *Teste de Usabilidade* e uma correlação substancial

para a variável *Entendimento*;

- Em T3, foi possível observar uma correlação muito alta para as variáveis *Design Humano*, *Antecipação* e *Design* e correlação substancial para *Teste de Usabilidade*;
- Em T4, a amostra da variável *Computação Móvel* apresentou alta correlação e correlação moderada para as variáveis *Design Humano*, *Protótipo de Baixa Fidelidade* e *WIMP*.
- Finalmente, para T5, foi obtida uma correlação muito alta para a variável *MIPS* e uma correlação substancial para *Caminho de Dados*, *RISC* e *Tempo de Resposta*.

Esses resultados mostram que todas as variáveis estudadas evidenciaram algum nível de correlação entre os dados reais e os valores inferidos pela RB. A correlação foi ainda mais expressiva para os grupos T2 e T3. Observa-se que pelo menos uma correlação muito alta foi encontrada em todos os grupos analisados e, entre todas as probabilidades, nenhuma indicou correlação baixa ou desprezível.

Figura 15 – Gráficos de dispersão para as variáveis: (a) *Teste de Usabilidade* – T3, (b) *Análise de Tarefas* – T2, (c) *MIPS* – T5, and (d) *RISC* – T5.



A Figura 15 exibe gráficos de dispersão para quatro das vinte variáveis analisadas: (a) *Teste de Usabilidade* - T3, (b) *Análise de Tarefas* - T2, (c) *MIPS* - T5, e (d) *RISC* - T5. Para cada uma, o eixo X ilustra a escala de pontuação obtida pelos estudantes, e o eixo Y indica a escala de valores inferidos pela RB.

Nesta figura é possível identificar o comportamento de algumas variáveis de acordo com o seu nível de correlação. Observa-se um crescimento linear dos valores analisados no

Figura 16 – Gráfico de radar para dois estudantes em T5.

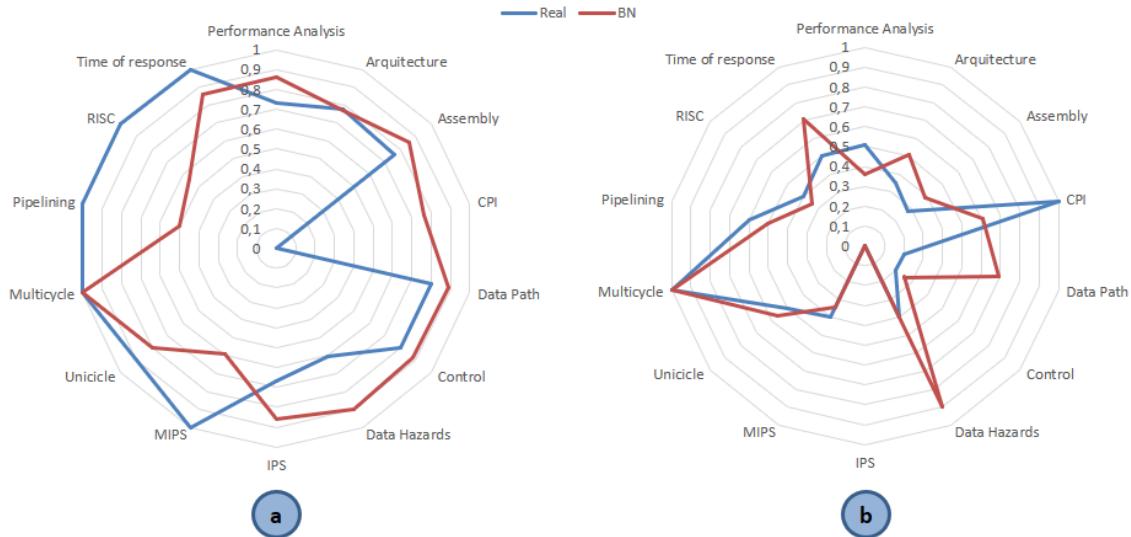


gráfico de dispersão (b), à medida que os valores obtidos a partir das notas dos estudantes crescem, os valores inferidos pela RB também crescem. O gráfico de dispersão (c) não mostra este comportamento tão claramente, no entanto, ele agrupa maior quantidade de pontos de dados em seu centro, o que indica uma boa correspondência entre os valores médios de notas reais e as notas inferidas pela RB. Finalmente, (a) e (d) apresentam pontos de dados mais dispersos, o que confirma os valores encontrados de correlação substancial.

A Figura 16 exibe gráficos de radar individualizados para dois estudantes do grupo T5. É possível observar que as notas inferidas pela RB podem ser muito próximas às reais, mesmo que algumas variáveis indiquem uma correlação substancial, como no caso da variável *Tempo de Resposta* no gráfico (a). Isso também pode acontecer com outras variáveis que não foram analisadas no teste de correlação, por exemplo, *Multiciclo*, *Controle* e *Arquitetura*. Este comportamento também é notado nas variáveis *RISC*, *Pipelining*, *Multiciclo*, *Monociclo*, *MIPS*, *IPS*, *Controle* e *Assembly* no gráfico (b).

Dessa forma, é possível observar que, na maioria das variáveis analisadas, a correlação apontou resultados muito satisfatórios, em que 80% das amostras analisadas obtiveram correlações muito altas ou substanciais e em nenhuma das amostras observaram-se correlações baixas ou insignificantes. Com isso, é possível concluir que o Módulo Probabilístico é adequado para analisar a probabilidade de conhecimento dos estudantes.

5.3 Experimento 2: Análise da Estrutura da Rede Bayesiana

A RB é a estrutura básica para determinar a probabilidade de conhecimento pelo modelo. Para que esse modelo tenha sucesso, é imprescindível que a estrutura da rede

seja condizente com a real semântica a ser abordada na disciplina, caso contrário, o processo de inferência das probabilidades pode estar comprometido. Dessa forma, será que é possível determinar diretrizes (*guidelines*) que possam orientar o processo de construção de uma RB? Será que inserir mais instrumentos avaliativos melhora a acurácia nos valores inferidos? Ou será melhor inserir os instrumentos avaliativos de forma mais distribuída? É melhor ter uma rede maior, com mais abrangência ou uma rede mais sucinta? Essas são algumas perguntas que podem orientar o processo de criação das RBs.

Com base nisso, este experimento teve como objetivo analisar a estrutura topológica de diversas RBs a fim de identificar quais organizações evidenciam melhores resultados quando comparados com o conhecimento do estudante, para, assim, ser capaz de definir *guidelines* que orientem o processo de criação das RBs a serem utilizadas para identificação de conhecimento. Ao final desse experimento, serão discutidas algumas orientações que foram alcançadas, a partir dos resultados obtidos, com o intuito de auxiliar o professor no momento da construção das redes.

5.3.1 Método Utilizado

Este experimento foi aplicado a cinco turmas ministradas ao longo de três semestres (2016-1, 2016-2 e 2017-1) para o curso de Sistemas de Informação na Universidade Federal de Uberlândia. No total, foram analisados 71 nós distribuídos em cinco redes diferentes (uma para cada disciplina).

O professor de cada turma foi orientado a criar uma RB com algumas características que a diferenciasse das redes anteriormente criadas, como: *(i)* números de nós; *(ii)* número de instrumentos avaliativos, *(iii)* número de relações entre os nós e, por fim, *(iv)* expressividade dos conceitos – se mais genéricos ou mais específicos. Todas essas configurações foram utilizadas com o intuito de diversificar ao máximo a estrutura da rede, a fim de se obter resultados que fossem realmente significativos.

A Tabela 6 ilustra detalhes sobre as cinco turmas participantes do experimento.

Tabela 6 – Turmas participantes do Experimento 2.

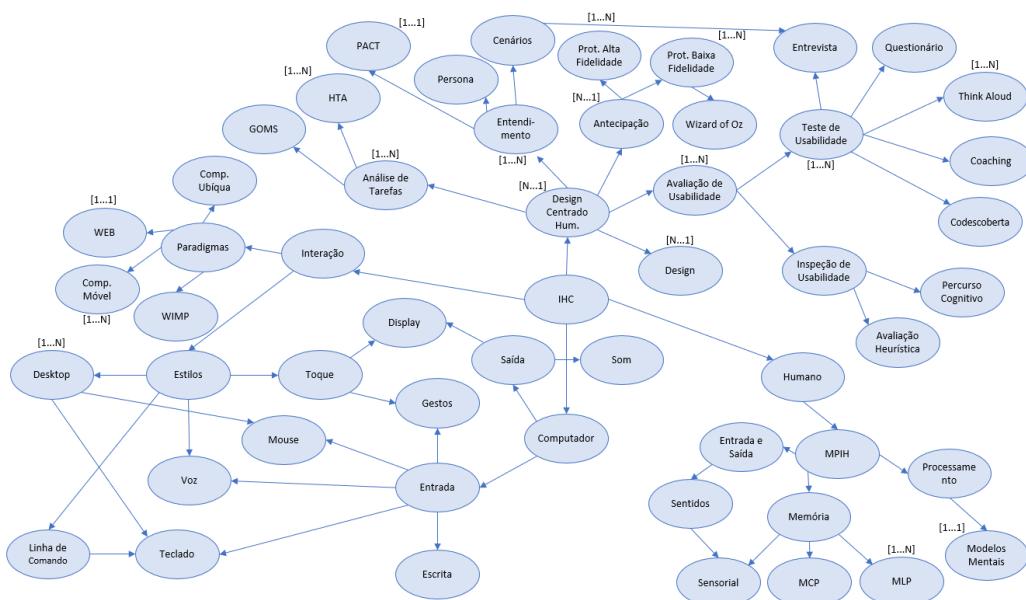
Sem.	Disciplina	Qtd Est. ¹	Qtd Tóp. ²	Ent/Sai ³	Qtd Comb. ⁴
T1 2016-1	Interação Humano-Computador	35	17	4/14	130084
T2 2016-2	Arq. e Org. de Computadores	30	14	4/11	15808
T3 2016-2	Interação Humano-Computador	33	12	3/10	4004
T4 2017-1	Arq. e Org. de Computadores	31	15	4/12	32071
T5 2017-1	Interação Humano-Computador	37	13	4/10	7772

¹Quantidade de estudantes matriculados na turma.²Quantidade de tópicos utilizados como evidência ou observação.³Limite de tópicos para combinação (apresentado em detalhes na Etapa 4).⁴Quantidade de combinações realizadas (apresentado em detalhes na Etapa 4).

Com o intuito de melhorar a apresentação e execução do método, escolheu-se distribuí-lo em cinco etapas, como segue:

- **1ª Etapa - Determinar a estrutura da rede:** Nesta fase, foram definidas as diferentes características de cada RB utilizada no experimento. Como mencionado anteriormente, foi solicitado a cada professor criar uma rede distinta das demais, a fim de permitir diferentes análises ao comportamento das probabilidades inferidas. As Figuras 17, 18, 19, 20, 21 apresentam as redes das turmas T1, T2, T3, T4 e T5, respectivamente. Observando as redes das turmas T1, T3 e T5, é possível perceber que, mesmo sendo redes da mesma disciplina, elas possuem diferenças significativas. Por exemplo, a rede da turma T5 apresenta bem menos tópicos que as demais. É possível perceber também que, mesmo bem parecidas, as redes da disciplina T1 e T3 apresentam diferenças sutis, principalmente, na disposição dos nós evidências (representado por [X...Y]). Já nas redes das disciplinas T2 e T4, é possível perceber diferenças, tanto na quantidade e abrangência dos tópicos quanto na distribuição dos nós evidências.

Figura 17 – RB utilizada no experimento 2 (T1).



- **2^a Etapa: Período letivo:** Esta é a etapa do período letivo em que os professores ministraram suas disciplinas. Aqui, foram adicionados os instrumentos avaliativos à RB e as avaliações periódicas de conhecimento foram aplicadas. Nesta etapa, buscou-se interferir o mínimo possível na dinâmica como o professor ministra seu conteúdo.
 - **3^a Etapa: Associação das avaliações à RB:** Após a conclusão da disciplina, os professores fizeram a correspondência (ligação) entre os conteúdos abordados

Figura 18 – RB utilizada no experimento 2 (T2).

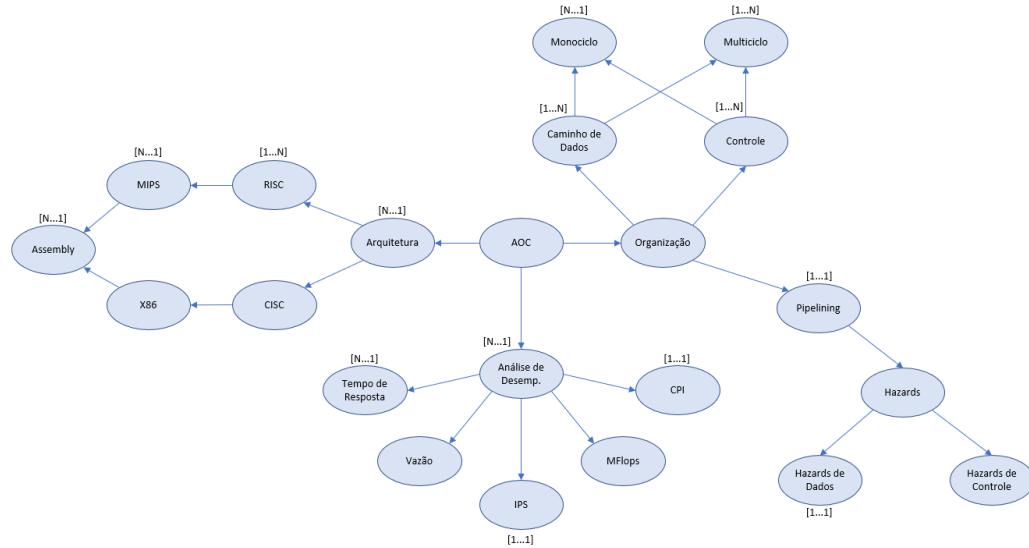
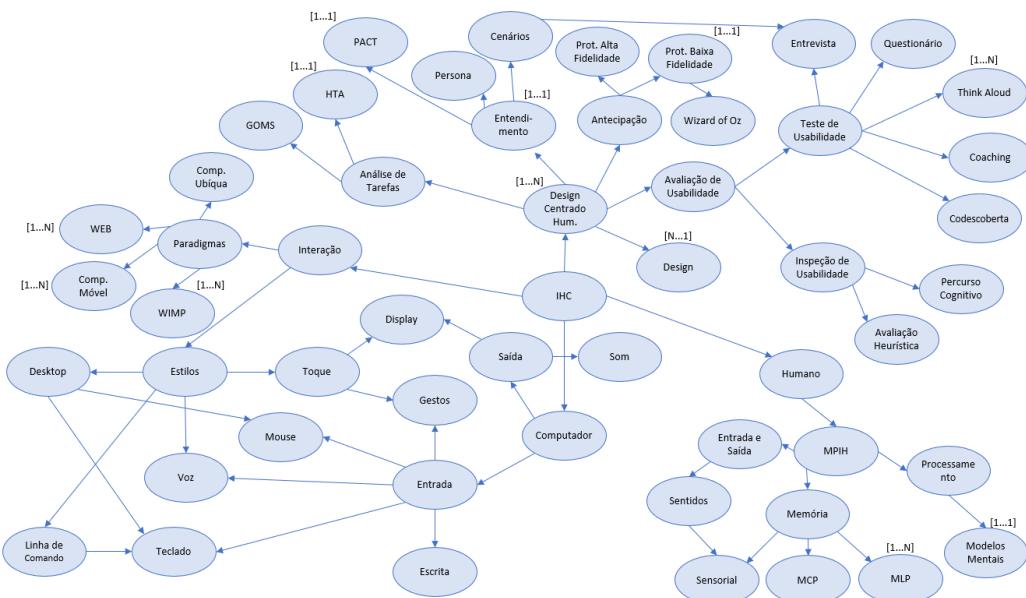


Figura 19 – RB utilizada no experimento 2 (T3).



nas avaliações e os tópicos disponíveis na rede. Aqui, foi possível relacionar, com alto grau de correspondência, quais os assuntos abordados em cada questão nas avaliações correspondem a determinados tópicos na rede. Nas Figuras 17, 18, 19, 20, 21 pôde-se observar uma notação de cardinalidade ([1...1], [1...N], [N...1]), que representa como foi realizada a associação de tal tópico a determinadas questões nas avaliações. A notação [1...1] aponta que existe uma questão avaliativa relacionada com este (um) conceito. A notação [1...N] mostra que existe uma questão avaliativa relacionada com vários conceitos. E por fim, a notação [N...1] indica que existem várias questões avaliativas relacionadas com um conceito.

Figura 20 – RB utilizada no experimento 2 (T4).

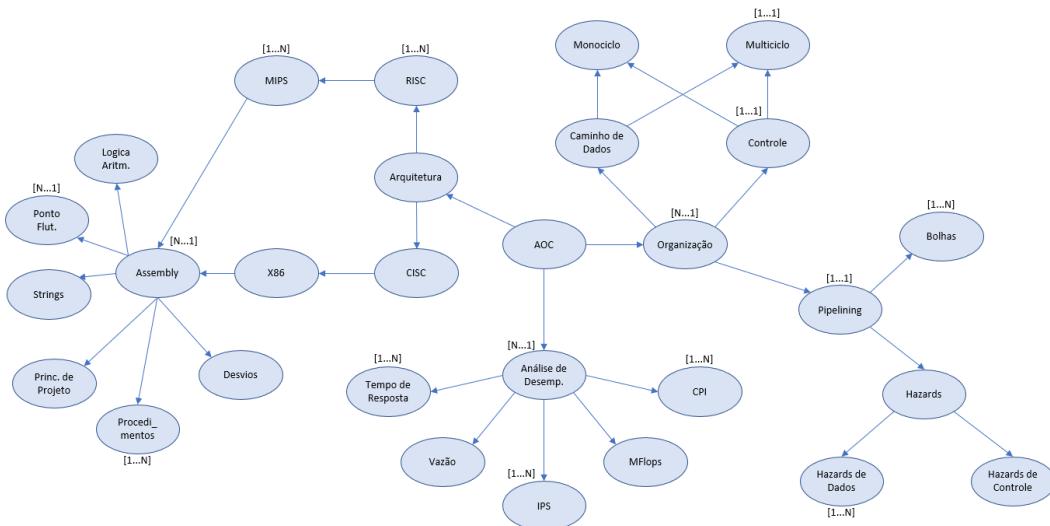
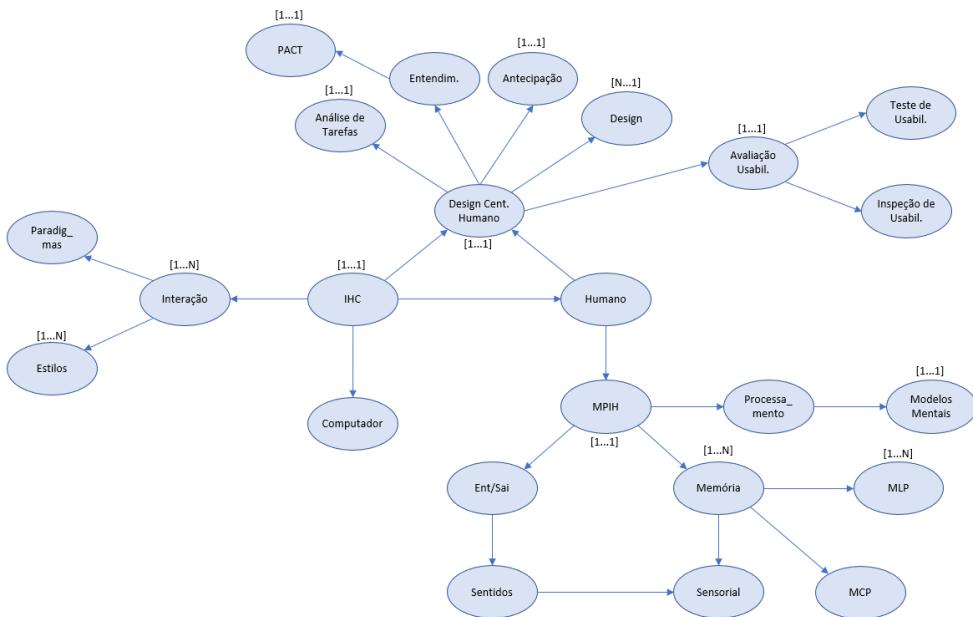


Figura 21 – RB utilizada no experimento 2 (T5).



Após a fase de associação, foi realizado um levantamento que mapeou as avaliações de todos os estudantes, a fim de determinar quais questões os estudantes acertaram ou erraram e, consequentemente, qual o nível de conhecimento em cada tópico da rede. A partir disso, foi possível obter, para cada tópico associado a questões avaliativas, qual o nível de conhecimento do estudante.

□ **4^a Etapa: Seleção dos tópicos de entrada e saída:** Esta é uma das etapas mais minuciosas deste experimento. Esta etapa teve como objetivo determinar, dentre os tópicos associados às questões nas avaliações, quais seriam utilizados como entrada

e saída. Para determinar quais tópicos foram utilizados como entrada e saída, foi necessário realizar todas as combinações possíveis dentro de um limite e verificar quais combinações apresentam melhores resultados. É importante destacar que foi definido um limite superior e inferior de 20% (vinte por cento) do número de tópicos associados para determinar o número de combinações que foram executadas. Este valor se justifica, pois não faz sentido verificar as combinações para uma quantidade muito baixa de entradas, tendo em vista que não seria possível obter resultados positivos para os demais tópicos; também não faz sentido verificar as combinações para quantidades muito altas de entradas, pois quase todos os tópicos já estariam com os valores reais de conhecimento. O valor de intervalo e o número total de combinações podem ser vistos na 6^a e 7^a colunas, respectivamente, da Tabela 6.

Vale destacar que foi construído um procedimento para realizar as combinações necessárias e verificar quais abordagens evidenciaram melhores resultados de correlação. Para analisar a correlação entre os valores reais e os valores inferidos, foi utilizada a correlação de posto de *Spearman*. Como já mencionado no método do experimento anterior, recorreu-se aos adjetivos propostos por Davis (1971) para interpretação desses valores. Todas as combinações selecionadas para a próxima etapa, bem como os valores da correlação de *Spearman* podem ser vistos no Apêndice C.

- **5^a Etapa: Análise das propriedades da rede:** Esta é a última etapa da análise, por meio da qual foram analisadas as quatro configurações de cada rede selecionada na etapa anterior com o intuito de analisar quais características estão presentes e se elas são comprovadas em alguma outra rede. Após essa etapa, foi determinado quais configurações são essenciais para um alto nível de correlação entre os valores inferidos pela rede e os valores reais de conhecimento.

5.3.2 Resultados e Discussão

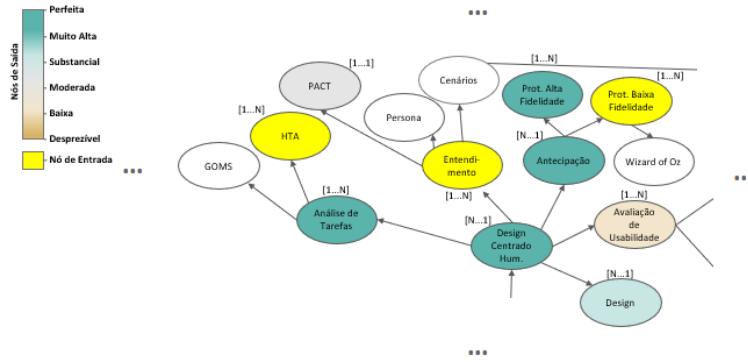
O experimento foi conduzido seguindo todos os passos propostos e, ao final, foi possível obter resultados importantes para entendimento e auxílio na orientação para criação de novas RBs. No total, foram realizadas mais de 180.000 combinações entre entrada e saída da rede, nas quais foi possível selecionar as 20 melhores combinações (quatro para cada disciplina). Para essas 20 combinações, 63 instrumentos avaliativos e mais 150 estudantes foram analisados. Dentre todos os tópicos de saída, foi possível constatar 3 tópicos com correlação desprezível, 13 tópicos com correlação baixa, 31 tópicos com correlação moderada, 57 tópicos com correlação substancial e 85 tópicos com correlação muito alta. Nota-se clara predominância dos nós com correlações mais altas (substancial e muito alta).

Após a realização de todas as etapas propostas na metodologia, foi possível levantar algumas características que influenciaram positivamente na obtenção de melhores valores

de correlação:

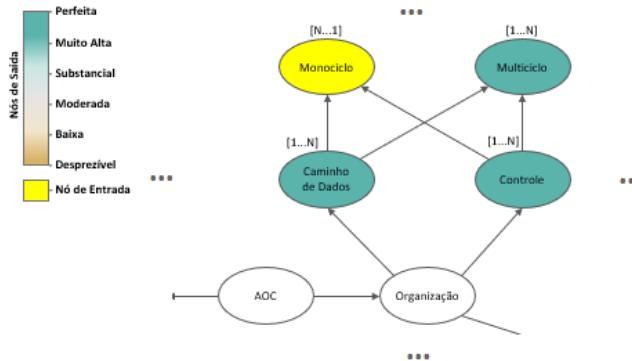
- Componentes mais conectados com maior número de nós evidências (entrada) apresentam maior probabilidade de alcançar nós de saída com correlações mais altas. Essa característica pode ser vista na Figura 22.

Figura 22 – Trecho de uma RB com componentes mais conexos.



- Redes que contêm instrumentos avaliativos associados a vários nós apontam maior probabilidade de alcançar correlações mais altas. Essa característica pode ser vista na Figura 23.

Figura 23 – Trecho de uma RB com diferentes instrumentos avaliativos.



- Redes mais sucintas e com instrumentos avaliativos mais distribuídos (em diferentes pontos da rede) apresentam correlações maiores. Essa característica pode ser vista na Figura 24.
- Nós que não possuem vizinhos que sejam evidências tendem a apresentar correlações mais baixas. Essa característica pode ser vista na Figura 25.
- Nós que contêm instrumentos avaliativos associados de forma um para um, ou seja, instrumentos avaliativos associados somente a um nó, tendem a apresentar correlações mais baixas. Essa característica pode ser vista na Figura 26.
- Nós folhas tendem a apresentar correlações mais baixas. Característica pode ser vista na Figura 27.

Figura 24 – Trecho de uma RB com poucos nós e instrumentos avaliativos distribuídos.

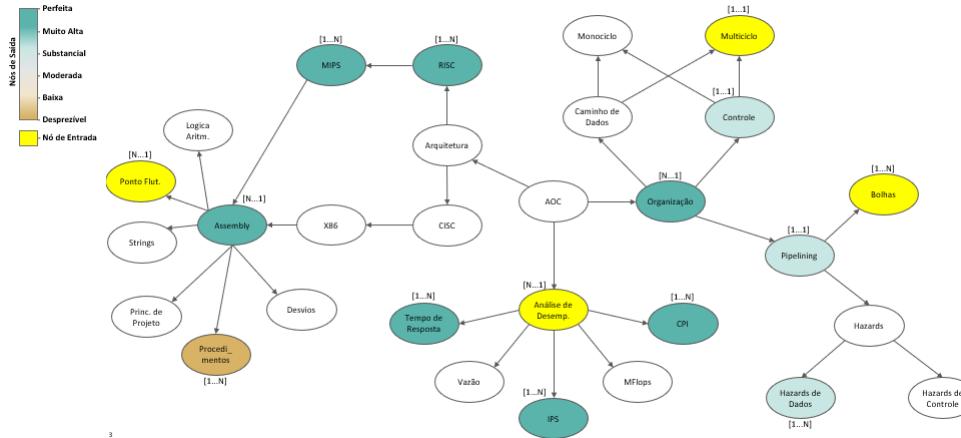


Figura 25 – Trecho de uma RB com nós sem evidências.

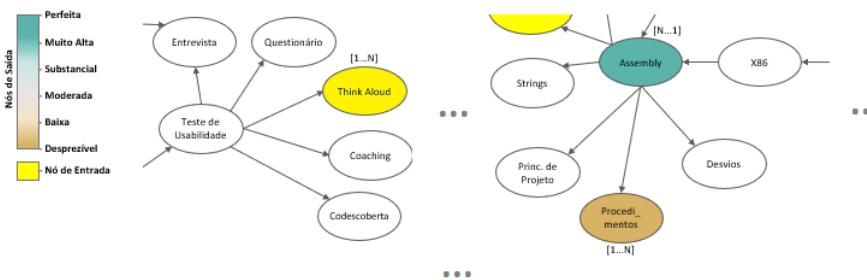
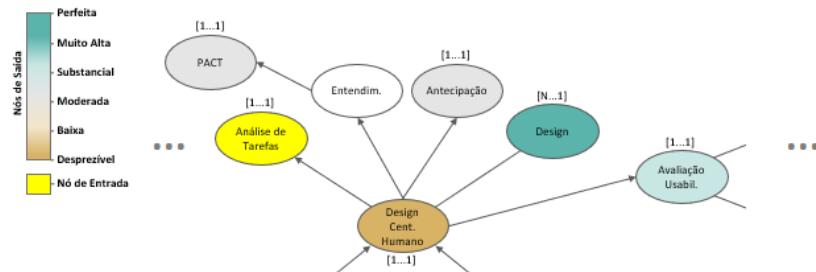


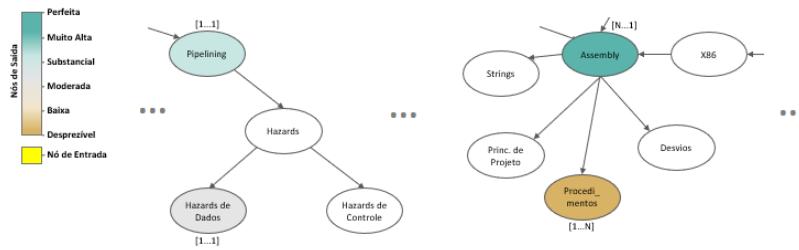
Figura 26 – Trecho de uma RB com instrumentos avaliativos do tipo 1...1 .



É importante ressaltar que tais resultados foram obtidos a partir de uma análise minuciosa das melhores alternativas obtidas em um conjunto de mais de 180.000 combinações realizadas. Para cada uma das combinações foi verificada a correlação e, a partir disso, com base na qualidade das correlações foram escolhidas as melhores redes. A análise contou, também, com a participação dos professores das disciplinas, que puderam opinar sobre as informações encontradas e deixar suas considerações.

Dentre as informações encontradas que merecem destaque, está a forma de disponibilização dos instrumentos avaliativos. A partir dos resultados encontrados, é possível definir que instrumentos avaliativos mais abrangentes, ou seja, instrumentos que abordam mais tópicos de conhecimento, possibilitam que as redes apresentem resultados mais positivos. Por outro lado, instrumentos mais restritos, ou seja, aqueles que atendem somente a um tópico de conhecimento tendem a apresentar resultados não tão positivos. Acredita-se que esse fato acontece pela característica probabilística das RBs, na qual percebe-se que, quando as informações de entrada são repassadas para mais nós, os nós interconectados

Figura 27 – Trecho de uma RB com nós folhas.



tendem a obter maior nível de confiabilidade.

Outra característica que merece atenção diz respeito ao quanto os nós estão conectados e ao número de nós evidências nesses componentes mais conexos. É possível considerar que componentes conexos tendem a apresentar resultados mais satisfatórios, mas não basta que eles sejam conexos, é necessário que haja a presença de instrumentos avaliativos neste componente. Uma abordagem interessante para explorar tais resultados é definir redes que tenham componentes mais fortemente conectados e seja possível distribuir alguns instrumentos internamente nestes componentes. Um resultado importante diz que o fato de uma rede possuir mais nós evidências não traz a certeza de que terá mais resultados positivos, no entanto, se esses nós evidências estiverem em pontos estratégicos, por exemplo, componentes mais conexos, essa informação pode ser verdadeira. Assim, é mais importante ter poucos instrumentos avaliativos em pontos estratégicos do que ter muitos instrumentos de forma aleatória na rede.

Por fim, uma última consideração que merece destaque é sobre a quantidade de nós disponibilizados. É possível perceber com tais resultados que redes mais sucintas apresentam resultados mais positivos. Acredita-se que essa característica está relacionada à facilidade na disponibilização dos instrumentos avaliativos. Segundo um professor, na rede menor fica mais fácil definir e associar os instrumentos avaliativos que nas redes maiores. Acredita-se que esse fato pode influenciar diretamente nos resultados de correlação encontrados. Mesmo que a associação foi realizada pelo professor, se ele sentir dificuldade nessa tarefa, o processo de inferência probabilística perde em confiabilidade.

5.4 Experimento 3: Análise do Desempenho dos Estudantes com base em Perfis de Intereração

A identificação do desempenho dos estudante com base em seu nível de comportamento é uma abordagem interessante para auxiliar nos processos tradicionais de avaliação. Quando aplicados a SAIEs, tais abordagens trazem ainda melhores resultados, pois, além de avaliar o nível de conhecimento através de instrumentos avaliativos sobre o conteúdo educacional transmitido pelo professor, esses ambientes ainda conseguem monitorar como é o comportamento desses estudantes ao acessar o conteúdo educacional. A quantidade

de informações advindas do processo de análise de comportamento é enorme, com isso, é possível realizar diversos estudos com o intuito de explorar o perfil de acesso e o nível de interação de cada estudante.

Dessa forma, este experimento teve como objetivo verificar se é possível classificar os estudantes em grupos de acordo com o seu nível de comportamento e, a partir disso, verificar se tais grupos evidenciam ou não desempenho significativamente diferentes.

5.4.1 Método Utilizado

Este experimento visa explorar conceitos de agrupamento de dados, a fim de identificar se existe relação entre o nível de comportamento do estudante com o nível de desempenho alcançado na disciplina. Como pontos específicos, verificou se diferentes comportamentos ao utilizarem um SAIE correspondem a diferentes níveis de aprendizagem. Dessa forma, foram definidas duas perguntas de pesquisa que guiaram este experimento:

- a) *A performance do estudante pode estar associada ao seu nível de comportamento em um SAIE?*
- b) *Quais perfis de acesso apresentam melhores e piores desempenhos quando associados à utilização de um SAIE?*

Participaram deste estudo um total de 119 estudantes, distribuídos em seis turmas ao longo de três semestres (2016-2, 2017-1, 2017-2). Foram escolhidas três turmas da disciplina Interação Humano-Computador (IHC) e três turmas da disciplina Arquitetura e Organização de Computadores (AOC), todas ministradas de forma presencial para o ensino superior na Faculdade de Computação da Universidade Federal de Uberlândia. Durante os três semestres em que o estudo foi realizado, os professores das disciplinas apresentaram aos estudantes uma nova ferramenta para auxiliá-los nos estudos, o CX². Uma primeira demonstração da utilização da plataforma foi realizada pelo professor no início de cada turma e, a partir daí, todas as aulas foram ministradas utilizando a plataforma, permitindo assim, que os conteúdos ministrados em aula estivessem disponíveis posteriormente para acesso no ambiente ubíquo. Vale ressaltar que não foi necessária alteração alguma na dinâmica como o professor ministrava sua disciplina, possibilitando assim, que o sistema não fosse intrusivo na forma rotineira como o professor costumava apresentar o conteúdo.

Para responder às perguntas de pesquisa, um procedimento padrão foi adotado para todas as turmas. Para não influenciar diretamente na dinâmica da disciplina, o experimento foi realizado de forma a não alterar a maneira como o professor ministrava suas disciplinas. A utilização da plataforma não era obrigatória, era apenas opcional, e servia como um complemento para os estudos em sala. Os procedimentos para realização

² Mesma plataforma apresentada no Capítulo 4

deste experimento foram divididos em três fases: aquisição dos dados, quantificação das interações e agrupamento dos dados.

Fase 1: Aquisição dos Dados

Durante as aulas, todo o conteúdo apresentado pelo professor era capturado e armazenado pelo ambiente educacional, permitindo assim, acesso futuro a esse conteúdo. Com esse recurso, os estudantes poderiam centrar sua atenção ao material ministrado pelo professor, pois teriam a certeza de que todas as informações estariam disponíveis posteriormente para acesso. O material educacional apresentado pelo professor era dividido em diversas partes e juntas eram transformadas em Objetos de Aprendizagem (OAs). A Tabela 7 apresenta alguns números sobre as turmas participantes deste experimento.

Tabela 7 – Turmas participantes do Experimento 3.

Sem.	Disciplina	Qtd Est. ¹	Qtd Aulas ²	Qtd OAs ³	
T1	2016-2	Interação Humano-Computador	25	14	835
T2	2016-2	Arq. e Org. de Computadores	22	21	928
T3	2017-1	Interação Humano-Computador	14	15	865
T4	2017-1	Arq. e Org. de Computadores	20	22	882
T5	2017-2	Interação Humano-Computador	24	15	867
T6	2017-2	Arq. e Org. de Computadores	14	23	843

¹Quantidade de estudantes. Neste experimento, foram utilizados somente os estudantes que concluíram a disciplina e que realizaram todas as avaliações. ²Quantidade de aulas ministradas. ³Quantidade de Objetos de Aprendizagem gerados para a turma.

Durante o semestre letivo, os estudantes utilizavam a plataforma para estudos regulares. É importante ressaltar que todas as interações dos estudantes com o material educacional estavam sendo registradas pela plataforma educacional, com isso, era possível saber como foi o comportamento desses estudantes e quais atividades eles realizaram durante seus estudos. Durante todos os semestres letivos, essas interações foram capturadas e armazenadas em forma de *logs* de acesso.

Fase 2: Quantificação das interações

Para ter uma medida mais quantitativa sobre o acesso dos estudantes, foi realizado um estudo com o intuito de definir limites para o tempo de acesso a cada material educacional. Dessa forma, pôde-se definir três níveis de tempo de interações (interações curtas – intA com tempos entre 2 e 10 segundos; interações médias – intB com tempos entre 11 e 30 segundos; e interações longas – intC com tempos acima de 31 segundos) a cada parte do material. Após a definição desses limiares, foram executados procedimentos para quantificar tais interações, permitindo assim, uma visão global de como foram os acessos

pelos estudantes em cada turma analisada. A Tabela 8 apresenta o número de interações em cada categoria analisada para cada uma das turmas.

Tabela 8 – Total de interações por turma analisada.

Sem.	Disciplina	intA ¹	intB ²	intC ³	Total
T1	2016-2 Interação Humano-Computador	12802	6131	41982	60915
T2	2016-2 Arq. e Org. de Computadores	5973	2823	21988	30784
T3	2017-1 Interação Humano-Computador	4925	2450	15625	23000
T4	2017-1 Arq. e Org. de Computadores	8025	4435	29443	41903
T5	2017-2 Interação Humano-Computador	8152	3792	26257	38201
T6	2017-2 Arq. e Org. de Computadores	3802	1674	16711	22187

¹Total de interações curtas. ²Total de interações médias. ³Total de interações longas.

Depois da quantificação das interações, foram realizados procedimentos para verificar e associar os acessos a seus respectivos usuários. Para isso, foi analisada a sessão de *login* de cada usuário e verificado qual foi a proporção de interações curtas, médias e longas que ele realizou durante cada acesso. Além disso, foi verificada a sequência de estudo para tentar identificar estudantes que sempre seguiam um comportamento linear ou estudantes que eram mais interativos durante o estudo. Após isso, todas essas informações foram agrupadas por usuário, o que proporcionou ter uma medida precisa do nível de comportamento para cada estudante. Neste caso, definiu-se uma estrutura para representação de um conjunto de interações no âmbito deste estudo, a qual pode ser vista a seguir:

$$IS = \{Us, pA(\%), pB(\%), pC(\%) \}$$

Onde,

IS = Conjunto de interação;

Us = Usuário;

pA = Proporcionalidade de interações curtas;

pB = Proporcionalidade de interações médias;

pC = Proporcionalidade de interações longas.

Fase 3: Agrupamento dos Dados

Após a quantificação das interações por usuário, foi realizado um procedimento de agrupamento de dados (*clustering*) com o objetivo de agrupar os estudantes com características semelhantes em um mesmo grupo. Em geral, a técnica de agrupamento classifica entidades de modo que cada objeto é semelhante aos outros no agrupamento com base em um conjunto de características (no caso deste estudo, os diferentes níveis de interação).

Os grupos resultantes desse procedimento devem apresentar elevada homogeneidade interna (dentro dos agrupamentos) e elevada heterogeneidade externa (entre grupos) (HAIR; ANDERSON; BABBIN, 2009).

O agrupamento foi realizado por meio do algoritmo *K-Means*, um dos mais objetivos e populares algoritmos de clusterização disponíveis na literatura (JAIN, 2010). O princípio deste algoritmo é encontrar K grupos nos dados fornecidos. O algoritmo funciona iterativamente para atribuir cada instância a um dos grupos K com base nos recursos fornecidos. As instâncias são agrupadas com base na similaridade das características.

Neste procedimento, foram realizadas duas abordagens para criação dos grupos: na primeira, foram criados dois grupos e, na segunda, foram criados três grupos. Essa medida foi necessária para verificar se existe diferença significativa em relação ao desempenho dos estudantes quando classificados em diferentes grupos.

A Tabela 9 apresenta os dados centrais obtidos a partir do processo de agrupamento. É possível perceber que, na primeira abordagem, estão alocados no *cluster0* (*CL0*) os usuários que possuem maiores valores de médias para pA (30.43) e pB (17.08), já no *cluster1* (*CL1*) estão os usuários que possuem maiores valores de médias para pC (74.44). Na segunda abordagem, é possível perceber que o *cluster2* (*CL2*) possui maiores valores para pA (33.17) e pB (19.73) e, por sua vez, menores valores para pC (46.67). Já o *cluster1* (*CL1*) possui os maiores valores de pC (79.37) e menores de pA (15.18) e pB (5.06). Por fim, o *cluster0* (*CL0*) possui valores medianos para pA (24.03), pB (11.23) e pC (64.46).

Tabela 9 – Distribuição média das interações em cada *cluster*.

Atributo	Abordagem 1		Abordagem 2		
	<i>CL0</i> (%)	<i>CL1</i> (%)	<i>CL0</i> (%)	<i>CL1</i> (%)	<i>CL2</i> (%)
pA	30.43	18.19	24.03	15.18	33.17
pB	17.08	7.07	11.23	5.06	19.73
pC	52.06	74.44	64.46	79.37	46.67

Após o processo de clusterização, testes estatísticos foram executados com o intuito de verificar se existem ou não diferenças significativas no desempenho dos estudantes em cada *cluster*.

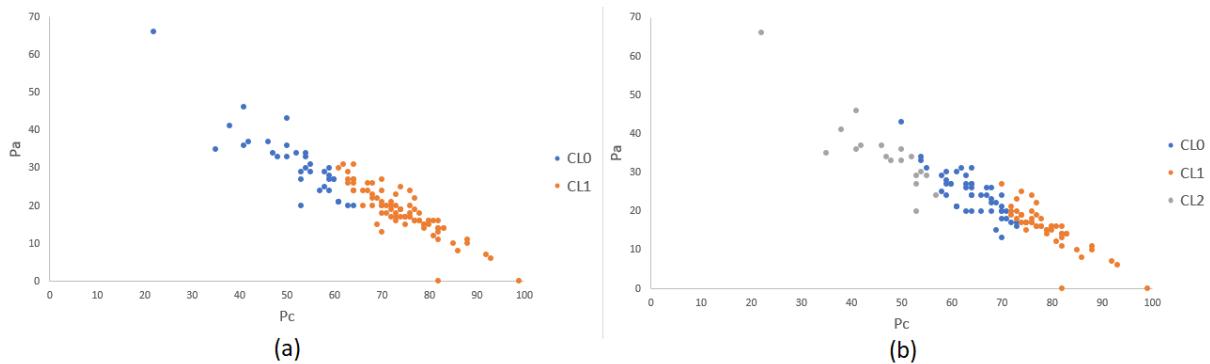
5.4.2 Resultados e Discussão

O principal objetivo desta avaliação foi verificar se existe diferença significativa entre as médias do desempenho dos estudantes que tiveram diferentes comportamentos de utilização do SAIE. Para realizar esta avaliação, os dados obtidos na quantificação das interações foram submetidos a um algoritmo de clusterização com o intuito de agrupar os

estudantes com comportamento semelhante e, a partir disso, verificar se existe diferença entre o desempenho final na disciplina dos estudantes nos diferentes grupos.

A divisão dos estudantes em grupos foi realizada em duas estratégias diferentes. Na primeira estratégia, todos os estudantes foram agrupados em dois grupos (*CL0* com 35 estudantes e *CL1* com 84 estudantes). Na segunda estratégia, todos os estudantes foram agrupados em três grupos (*CL0* com 49 estudantes, *CL1* com 52 estudantes e, por fim, *CL3* com 18 estudantes). A Figura 28 traz os gráficos de dispersão das duas estratégias, no qual é possível ver a distribuição dos estudantes em cada grupo. A Figura 28(a) apresenta a clusterização da primeira estratégia e a Figura 28(b) a clusterização da segunda estratégia. Em ambos os gráficos, o eixo X é representado valores para *pC* e o eixo Y representa valores de *pA*.

Figura 28 – Gráficos de dispersão para a distribuição dos clusters



Para verificar se as notas dos estudantes apresentaram alguma diferença entre os grupos, inicialmente, foram realizados dois testes com o intuito de analisar a homogeneidade de variância e a normalidade das amostras. Para a homogeneidade, foi utilizado o teste de *Levene*, e para verificar se os resíduos das amostras seguem a distribuição normal, foi utilizado o teste de *Shapiro-Wilk*. Com os testes, foi possível concluir que as amostras são homogêneas (para dois *clusters* *p-value* = 0,427 e para três *clusters* *p-value* = 0,118), no entanto, o teste de *Shapiro-Wilk* mostrou que as amostras analisadas não evidenciaram normalidade residual ($W(p) = 0,908 (0,00)$). Esses resultados orientaram a escolha dos próximos testes para verificar a diferença entre as médias. Neste caso, foram escolhidos os testes não paramétricos de *Mann-Whitney* para a primeira estratégia e o teste de *Kruskal-Wallis* para a segunda estratégia.

O teste de *Mann-Whitney* mostrou que o nível de utilização do ambiente interfere diretamente sobre a média do desempenho dos estudantes ($U = 1103,00$; $p < 0,05$). Neste caso, pôde-se identificar que existe uma diferença estatisticamente significativa entre as notas dos estudantes que estão agrupados nos diferentes *clusters*. Como pode ser visto na Tabela 10, os estudantes alocados no *CL1* possuem médias de *rank* maiores que os estudantes alocados no *CL0*.

Tabela 10 – Médias de desempenhos e resultado do teste de *Mann-Whitney* em cada *cluster* na primeira estratégia.

Cluster	N ¹	Média de Rank	Mann-Whitney U (p) ²
CL0	35	49,51	
CL1	84	64,37	1103,00 (0,032)

¹Número de Estudantes no *cluster*. ²p-value < 0,05 indica que existe diferença significativa entre as médias dos grupos analisados.

Analizando as características dos *clusters* a partir da Tabela 9 e os resultados do teste de *Mann-Whitney* na Tabela 10, percebe-se que os estudantes alocados no *CL1* alcançaram maiores médias para *pC* e menores médias para *pA* e *pB*, ou seja, aqueles estudantes que ficam mais tempo no ambiente, consequentemente, acessa por mais tempo o conteúdo educacional. O perfil desses estudantes é aquele que entra no ambiente para estudar, que fica mais tempo acessando o conteúdo da aula e menos tempo navegando pelo ambiente. Esses, conforme pode ser visto na Tabela 10, são justamente os estudantes que apresentaram melhores desempenhos. Por outro lado, os estudantes que entram menos no ambiente educacional e ficam por menos tempo, são justamente os estudantes que apresentam notas menores.

O teste de *Kruskal-Wallis* empregado para analisar a estratégia com três *clusters* também mostrou que o comportamento dos estudantes interfere no seu desempenho ($H(2) = 7,063$; $p < 0,05$). Neste caso, foi possível identificar que existe uma diferença estatisticamente significativa entre as notas dos estudantes que estão agrupados nos diferentes *clusters*. Na Tabela 11, é possível observar que os estudantes do *CL1* possuem as maiores médias de *rank*, e os estudantes do *CL2* as menores.

Tabela 11 – Médias de desempenhos e resultado do teste de *Kruskal-Wallis* em cada *cluster* na segunda estratégia.

Cluster	N ¹	Média de Rank	Chi-square (p) ²	df
Cl0	49	60,74		
Cl1	52	65,90	1103,00 (0,032)	2
Cl2	18	40,92		

¹Número de Estudantes no *cluster*. ²p-value < 0,05 indica que existe diferença significativa entre as médias dos grupos analisados.

A partir desses resultados, pode-se observar que os estudantes classificados no *CL1* tiveram as maiores notas. Esse grupo é justamente o grupo que possui acessos mais longos (*pC*) e poucos acessos curtos (*pA*). Ele é semelhante ao *CL1* da abordagem anterior. Já os estudantes classificados no *CL2* são os estudantes com menores notas. Esses estudantes possuem o perfil mais mediano, pois apresentam comportamento não tão longo nem tão

curto (Tabela 9). Os estudantes classificados no *CL0* apresentam comportamento mais dinâmico, fazendo o uso do ambiente menos frequente e com muitos acessos rápidos.

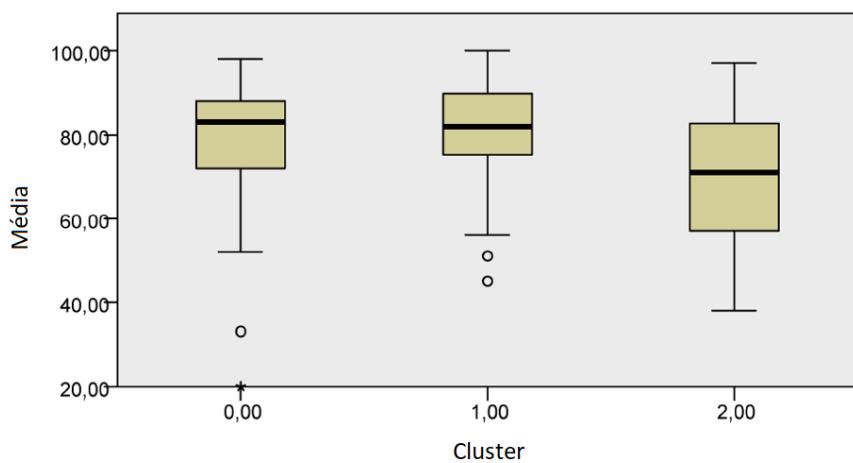
No caso do agrupamento em três *clusters*, é importante identificar onde realmente está a diferença estatisticamente significante. Para isso, recorreu-se à abordagem de *Dunn-Bonferroni* para realização da comparação par a par. Na Tabela 12, é possível verificar a diferença que existe entre os três grupos. É possível constatar que existe uma diferença significante entre o *CL2* e o *CL1* ($p = 0,024$), neste caso, o grupo que possui maior média (estudantes que acessam o ambiente prioritariamente por longos períodos) e o grupo que possui menor média (estudantes que costumam não acessar o sistema por longos períodos). O gráfico de Blox-plot na Figura 29 ilustra, de forma mais clara, a visualização dessas diferenças entre as médias. É possível perceber que a diferença entre o *CL0* e *CL1* e o *CL0* e *CL2* não foi tão expressiva quanto à diferença entre o *CL1* e *CL2*.

Tabela 12 – Resultados de testes de comparação par a par entre os *clusters*.

Cluster	Teste estatístico	Erro Padrão	Padrão de Teste	Sig	Ajuste de Sig ¹
Cl2 - Cl0	19,82	9,50	2,08	.037	.111
Cl2 - Cl1	24,98	9,42	2,65	.008	.024
Cl0 - Cl1	-5,15	6,86	-.752	.452	1.00

¹Valor de ajuste $< 0,05$ indica diferença significativa entre os grupos analisados.

Figura 29 – Comparação entre os três *clusters*.



Analizando o comportamento dos quartis e das medianas representadas em cada *cluster*, nota-se maior influência da variável que representa o acesso longo (*pC*) no resultado final da performance dos estudantes. Essa é a variável que mais influenciou no desempenho de cada *cluster*, ou seja, quanto maior é a ocorrência de *pC* maior é a média no desempenho no *cluster*, e quanto menor é a ocorrência de *pC* menor é a média no desempenho. Existe também uma influência da variável de acesso curto (*pA*), não tão significativa quanto *pC*, mas capaz de representar uma certa influência sobre os *clusters*.

No caso da variável pA , quanto menor são os acessos curtos, maiores são as médias no desempenhos dos estudantes. Com isso, verifica-se que a variável pC influencia positivamente no desempenho do estudante, ao passo que a variável pA influencia, de certa forma, negativamente, no desempenho do estudante.

Respondendo à primeira questão proposta neste experimento, provou-se estatisticamente que existe, sim, relação entre o comportamento dos estudantes durante a utilização de um SAIE e o seu desempenho. O teste de *Mann-Whitney* mostrou que quando a amostra é dividida em dois grupos, o grupo que acessa o sistema por períodos mais longos apresenta o desempenho maior que o grupo que acessa o sistema prioritariamente por períodos curtos. Fortalecendo ainda mais esse resultado, o teste de *Kruskal-Wallis* mostrou que, quando a amostra é dividida em três grupos, também é possível perceber um maior desempenho no mesmo grupo de estudantes (os que acessam por períodos mais longos) e um menor desempenho para aqueles que não acessam por períodos longos.

Respondendo à segunda questão proposta, é possível dividir o grupo de estudantes que faz o uso de um SAIE em até três perfis: *Perfil A* são os estudantes que possuem mais acessos longos e menor acesso curto e médio; o segundo perfil (*Perfil B*) são os estudantes que acessam o sistema com menos acessos longos e mais acessos curtos e médios; e por fim, o *Perfil C* são os estudantes que possuem um comportamento mais mediano, estando na média para acessos curtos, médios e longos. É possível concluir estatisticamente que os estudantes do *Perfil A* possuem as melhores notas ao passo que os estudantes do *Perfil B* possuem as menores. Os estudantes do *Perfil C* se enquadram em uma região de mais incerteza, no qual não é possível fazer afirmações com fortes fundamentos estatísticos.

5.5 Experimento 4: Análise sobre o Impacto na Utilização do MAE

Um dos pontos importantes deste trabalho é a integração de um recurso para visualização e acompanhamento da aprendizagem, tanto pelos professores das disciplinas quanto pelos próprios estudantes. Esse modelo de visualização está fundamentado em um conceito conhecido como MAE, o qual permite “abrir” as informações sobre o nível de conhecimento dos estudantes da turma com o intuito de proporcionar auto-regulação (do inglês *self regulated*) para que os próprios estudantes consigam guiar seus estudos nos conteúdos em que estão com maior dificuldade. Por outro lado, esse recurso possibilita que professores possam acompanhar com maior frequência o desempenho dos estudantes na sua disciplina.

Com isso em vista, este experimento tem como objetivo analisar o impacto causado pela utilização do MAE na satisfação e no comportamento dos estudantes. Este é um experimento qualitativo e permite verificar se o fato dos estudantes conseguirem acompanhar o seu desenvolvimento influencia positiva ou negativamente no seu comportamento

(motivação, competitividade, interesse pelos estudos, etc). Para essa análise, foram aplicados questionários e realizadas entrevistas com o intuito de verificar a satisfação dos estudantes com a utilização das ferramentas de visualização.

5.5.1 Método Utilizado

Este experimento foi aplicado a quatro turmas ministradas ao longo de dois semestres (2017-1 e 2017-2) para o curso de Sistemas de Informação na Universidade Federal de Uberlândia. No total, foram respondidos 139 questionários e foram realizadas 13 entrevistas. Os questionários foram aplicados aos estudantes das quatro turmas, duas de Interação Humano-Computador e duas de Arquitetura e Organização de Computadores. Já as entrevistas foram realizadas somente com os estudantes do segundo semestre de 2017, em duas turmas, Interação Humano-Computador e Arquitetura e Organização de Computadores. A Tabela 13 apresenta os números de participantes deste experimento.

Tabela 13 – Turmas participantes do Experimento 4.

Sem.	Disciplina	Qtd Questionários. ¹	Qtd Entrevistas ²
T1	2017-1 Interação Humano-Computador	33	–
T2	2017-1 Arq. e Org. de Computadores	30	–
T3	2017-2 Interação Humano-Computador	38	9
T4	2017-2 Arq. e Org. de Computadores	38	4

¹Quantidade de estudantes que responderam os questionários. ²Quantidade de estudantes que participaram das entrevistas.

A aplicação dos questionários seguiu o seguinte procedimento: no início da disciplina o professor apresentou a ferramenta que seria utilizada para apoiar as atividades educacionais (o CX). Neste momento, o professor demonstrou como seria a utilização e descreveu as funcionalidades desta. Durante o semestre, os alunos fizeram o uso da plataforma, e ao final, o professor aplicou um questionário para avaliação da satisfação dos usuários.

O questionário aplicado utilizou, na maioria das perguntas, uma escala do tipo *Likert* de sete pontos (discordo totalmente a concordo totalmente). As perguntas deste questionário podem ser classificadas em três grupos: o primeiro é responsável por avaliar a percepção de utilidade das funcionalidades relacionadas com o MAE; o segundo grupo foi responsável por avaliar se houve ou não alteração na forma como o estudante estudava; e por fim, o terceiro grupo verificava a satisfação na utilização da ferramenta de visualização e quais gráficos permitem uma visualização mais adequada para diferentes situações. O questionário completo pode ser visto no Apêndice D.

As entrevistas foram utilizadas neste experimento com o intuito de oferecer uma avaliação mais subjetiva, no qual foi possível observar não somente a percepção sobre utilidade e satisfação na utilização da ferramenta, como também foi possível identificar pontos fra-

cos e fortes da funcionalidade do ponto de vista dos estudantes. Para guiar a entrevista, os entrevistadores seguiram o seguinte roteiro de perguntas: **(a)** Você utilizou a plataforma? Achou útil? Qual foi a sua percepção? Como era o seu acesso ao sistema? **(b)** Você percebeu que existia uma funcionalidade de *quiz*? O que você achou? Qual era o seu comportamento quando não sabia a resposta? **(c)** Você utilizou a funcionalidade de visualização de desempenho? O que você achou? Com que frequência você costumava acessar essa funcionalidade? Essa funcionalidade te motivou? Ela alterou o seu comportamento ao utilizar a plataforma? **(d)** O que você fazia quando percebia a existência de um *quiz* naquele ponto do conteúdo? Você respondeu a um mesmo *quiz* mais de uma vez? **(e)** Gostaria de propor alguma sugestão de melhoria para a plataforma? É importante destacar que as entrevistas foram realizadas com certo caráter informal com o intuito de não exercer uma determinada “pressão” sobre os estudantes³.

5.5.2 Resultados e Discussão

Como esse experimento se propôs a analisar questões descritivas sobre a percepção dos estudantes, uma avaliação qualitativa fez-se necessária para uma melhor interpretação sobre os resultados obtidos. Para isso, inicialmente, serão discutidos os resultados dos questionários e, logo após, serão discutidos os resultados das entrevistas.

5.5.2.1 Avaliação sobre os Questionários

Para avaliar as respostas obtidas por meio dos questionários, inicialmente, foi realizada uma análise descritiva para obter a mediana e a moda das respostas. Esses testes são padrões quando são analisados questionários qualitativos, como é o caso deste experimento. Para uma análise mais objetiva, as perguntas dos questionários foram agrupadas e serão discutidas separadamente, inicialmente serão discutidos os resultados sobre a percepção de utilidade (Questões 1.1, 1.2, 1.3 e 1.4), depois sobre as possíveis mudanças de comportamento (Questões 2.1, 2.2 e 2.3), questões relativas à satisfação do usuário (questão 3), e por fim, preferência sobre a visualização (Questões 4, 5 e 6).

A Tabela 14 apresenta valores referentes às respostas sobre a percepção de utilidade em cada nível da escala Likert bem como a mediana e a moda para cada questão.

É possível identificar que em nenhuma das questões analisadas houve maiores concentrações de respostas entre os valores de discordância, para todas as questões os estudantes avaliaram com resultados positivos. No gráfico da Figura 30 é possível perceber a maior concentração, para todas as questões, de valores positivos, ou seja, maior concordância sobre as questões levantadas.

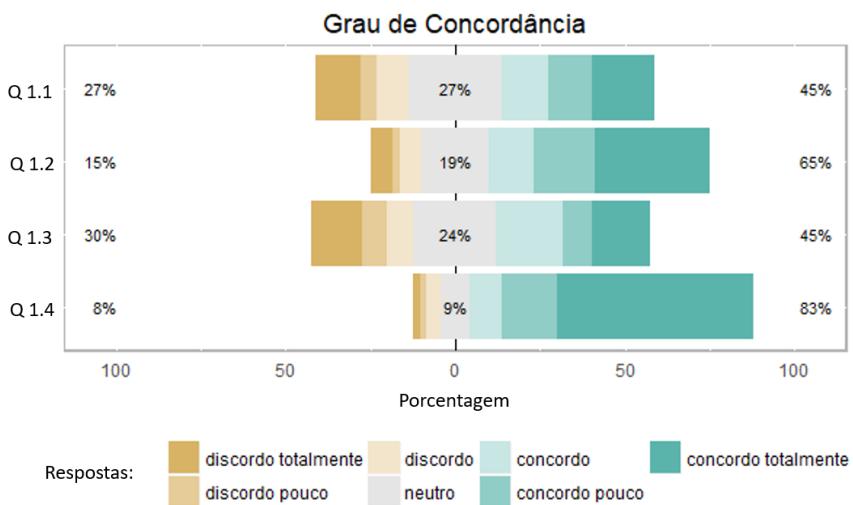
³ É importante destacar que as entrevistas foram feitas com os estudantes de forma anônima, ou seja, eles sabiam que nenhuma informação seria passada para o professor ou qualquer outro integrante do projeto.

Tabela 14 – Resultados sobre a percepção de utilidade.

Q ¹	DT ²	DP ³	D ⁴	N ⁵	C ⁶	CP ⁷	CT ⁸	Med ⁹	Mod ¹⁰
1.1	18(12,9%)	7(5,0%)	13(9,4%)	38(27,3%)	19(13,7%)	18(12,9%)	26(18,7%)	4	4
1.2	9(6,5%)	3(2,2%)	9(6,5%)	27(19,4%)	19(13,7%)	25(18,0%)	47(33,8%)	6	7
1.3	21(15,1%)	10(7,2%)	11(7,9%)	34(24,5%)	27(19,4%)	12(8,6%)	24(17,3%)	4	4
1.4	3(2,2%)	2(1,4%)	6(4,3%)	12(8,6%)	13(9,4%)	23(16,5%)	80(57,6%)	7	7

¹ Questão do Questionário. ² Discordo Totalmente. ³ Discordo Pouco. ⁴ Discordo. ⁵ Neutro. ⁶ Concordo. ⁷ Concordo Pouco. ⁸ Concordo Totalmente. ⁹ Mediana. ¹⁰ Moda.

Figura 30 – Grau de concordância sobre a percepção de utilidade.



Para a Questão 1.1, a qual considera o nível de satisfação sobre a funcionalidade da visualização de desempenho, é possível perceber claramente maior concentração de respostas positivas, concordando que a funcionalidade influenciou positivamente no aprendizado. A mesma interpretação pode ser obtida ao analisar as Questões 1.2 e 1.3 que verificam a utilidade da funcionalidade de *quiz* e gamificação, respectivamente. Quando perguntado se gostariam que a plataforma fosse utilizada em outras disciplinas, a resposta é muito positiva (mais de 80% das respostas).

Ainda analisando as questões referentes à utilidade e ao impacto positivo na aprendizagem, é possível verificar que os resultados agregados de todas as respostas também mostram uma tendência positiva para a utilização de uma plataforma educacional com suporte ao MAE. Na Tabela 15, verifica-se uma clara maior concentração de valores para as respostas de concordo totalmente (31,8%) e também maior concentração quando analisadas todas as respostas positivas (concordo, concordo pouco e concordo totalmente) com 59,9 % das respostas, bem maior que as respostas de discordância (discordo, discordo pouco e discordo totalmente) com 20% das respostas.

A Tabela 16 expõe valores referentes às respostas sobre a mudança de comportamento a partir da utilização da ferramenta de visualização.

Com esses resultados da Tabela 16, observa-se um maior balanceamento entre os

Tabela 15 – Resultados agregados.

	N ¹	Porcentagem ²
Discordo Totalmente	51	9,2%
Discordo Pouco	22	4,0%
Discordo	39	7,0%
Neutro	111	20,0%
Concordo	78	14,0%
Concordo Pouco	78	14,0%
Concordo Totalmente	177	31,8%
Total	556	100,0%

¹Numero Total de Respostas. ²Porcentagem das Respostas.

Tabela 16 – Resultados sobre a alteração no comportamento.

Q ¹	DT ²	DP ³	D ⁴	N ⁵	C ⁶	CP ⁷	CT ⁸	Med ⁹	Mod ¹⁰
2.1	31(22,3%)	12(8,6%)	11(7,9%)	31(22,3%)	17(12,2%)	13(9,4%)	24(17,3%)	4	1
2.2	32(23,0%)	7(5,0%)	12(8,6%)	21(15,1%)	17(12,2%)	23(16,5%)	27(19,4%)	4	1
2.3	40(28,8%)	13(9,4%)	18(12,9%)	29(20,9%)	14(10,1%)	12(8,6%)	13(9,4%)	3	1

¹Questão do Questionário. ²Discordo Totalmente. ³Discordo Pouco. ⁴Discordo. ⁵Neutro. ⁶Concordo.

⁷Concordo Pouco. ⁸Concordo Totalmente. ⁹Mediana. ¹⁰Moda.

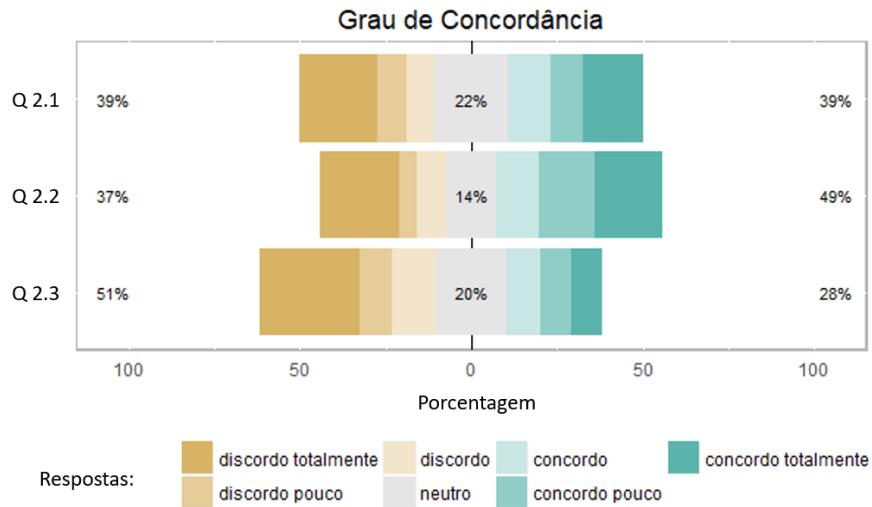
valores de concordância e discordância. Houve maior concentração de valores referentes à discordância sobre a alteração no comportamento (pode ser visto pela moda obtida nas questões), o que pode indicar que muitos dos estudantes não concordam que a visualização de desempenho possa alterar o seu comportamento de forma geral. No entanto, nota-se que, para algumas dimensões de comportamento, a funcionalidade de visualização pode sim influenciar, como é o caso da necessidade de melhoria no desempenho (Questão 2.2).

No gráfico da Figura 31, constata-se maior concentração de estudantes que concordam que os gráficos de desempenho podem influenciar na sua alteração de comportamento do ponto de vista de melhoria de desempenho (Questão 2.2). O fato do estudante poder conhecer os conteúdos em que possuem maior e menor conhecimento, pode sim, de certa forma, auxiliá-lo a identificar seus pontos fracos e alterar de forma significativa o seu nível de estudo.

Na Figura 31, também percebe-se um balanceamento para a Questão 2.1 (motivação para estudo), e maior tendência de discordância para a Questão 2.3 (competitividade). De forma geral, entende-se que os estudantes veem a funcionalidade de visualização de desempenho mais como um auxílio para identificar seus pontos fortes e fracos em cada disciplina, e não como uma ferramenta de competitividade e que explora questões motivacionais.

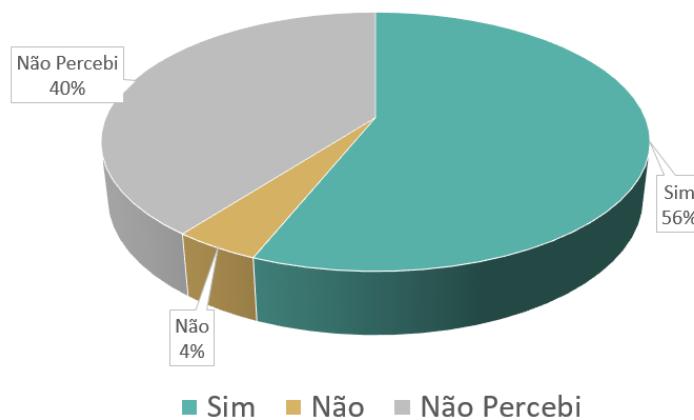
A Figura 32 exibe os resultados para a terceira pergunta do questionário, que avalia a satisfação do usuário ao utilizar a ferramenta de visualização. Nesta pergunta foi ques-

Figura 31 – Grau de concordância sobre a alteração no comportamento.



tionado se eles gostaram ou não desta ferramenta. A grande maioria (56%) respondeu que sim, gostou da funcionalidade, ao passo que somente 4% responderam que não gostaram da funcionalidade. É possível perceber que 40% responderam que não perceberam a funcionalidade integrada na plataforma. Isso pode ter acontecido por alguma inconsistência no design que dificultou a visualização da funcionalidade. É interessante destacar, também, que estes 40% pode ter acontecido pelo fato da ferramenta não disponibilizar diálogo entre os estudantes e o sistema, uma característica bem utilizada em ferramentas com suporte a auto-regulação.

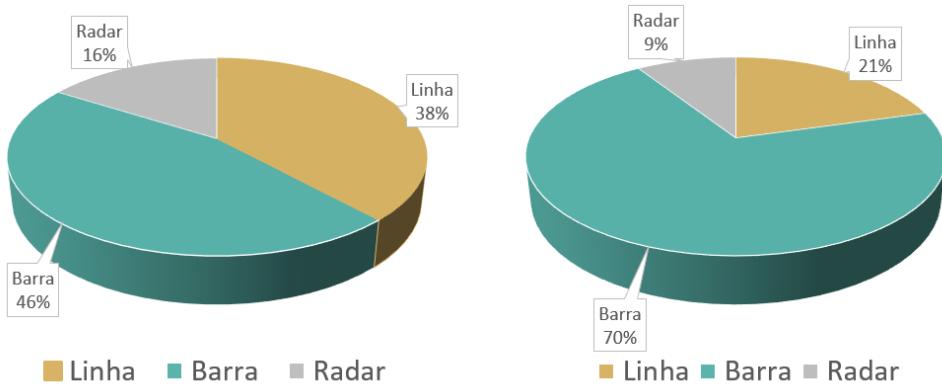
Figura 32 – Respostas à questão três.



A quarta e quinta pergunta buscam identificar qual a forma de visualização que os estudantes preferem. O MAE permite que o nível de desempenho do estudante seja apresentado por três tipos diferentes de gráficos: gráfico de linha, gráfico de barra e gráfico de radar. Neste sentido, a pergunta quatro questionou qual o tipo de gráfico o estudante prefere para visualizar o seu nível de desempenho, e a pergunta cinco queria saber qual tipo de gráfico o estudante prefere para visualizar seu rendimento com relação

à nota média da turma. A Figura 33 traz dois gráficos com dados dessas perguntas (Figura 33(a) corresponde à pergunta quatro e a Figura 33(b) corresponde à pergunta cinco).

Figura 33 – Respostas às questões quatro e cinco.

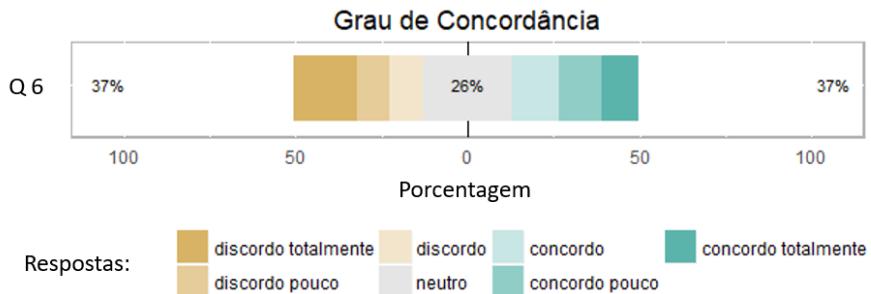


É possível verificar que tanto para a visualização individualizada do desempenho quanto para a visualização em relação à média da disciplina, os estudantes preferiram o gráfico de barra, seguido pelo gráfico de linha e, por fim, o gráfico de radar. É interessante notar que para a comparação entre as notas e as médias da turma, o gráfico de barra foi ainda maior, alcançando um total de 70% das respostas. Esses dados são interessantes e podem guiar novos processos para construção de ferramentas de visualização.

A última pergunta versava sobre o quanto a comparação da média do estudante com a turma influenciou positivamente na sua aprendizagem. Neste caso, nota-se certo balanceamento entre as respostas que concordam e discordam dessa funcionalidade. Observando a Figura 34, verifica-se de forma mais clara tais valores. Nota-se que 37% dos estudantes concordaram que a comparação da sua nota com a média da turma pode sim influenciar positivamente no seu aprendizado e outros 37% consideraram que não existe influência sobre o seu aprendizado. Esse é um resultado interessante e está diretamente relacionado com a Questão 2.3, no qual é possível perceber também certa compatibilidade entre as respostas com objetivos semelhantes, ou seja, verificar um recurso de visualização fundamentado em MAE pode influenciar na competitividade dos estudantes. Neste caso, foi possível perceber que existe sim uma influência, mas ela não necessariamente domina os estudantes e traz resultados para a turma em seu todo.

Com o intuito de avaliar a confiabilidade das respostas dos questionários, o teste de Alfa de Cronbach foi realizado. Este teste visa analisar a consistência interna das respostas baseada na correlação entre diferentes itens para a mesma escala. Para este experimento, o questionário foi dividido em três categorias – **CAT1**: Influência das funcionalidades na aprendizagem da disciplina; **CAT2**: Alteração na forma de estudo; **CAT3**: Preferências na visualização.

Figura 34 – Grau de concordância sobre a questão seis.



Para interpretação dos valores do Alfa de Cronbach, utilizaram-se os adjetivos propostos por Landis e Koch (1977), que definem os seguintes intervalos: $\alpha > 0.80$ indica uma consistência interna quase perfeita; $0.61 < \alpha < 0.80$ representa uma consistência interna substancial; $0.41 < \alpha < 0.60$ indica uma consistência interna moderada; $0.21 < \alpha < 0.40$ indica uma consistência interna razoável; e por fim, $\alpha < 0.21$ representa uma consistência interna pequena.

A Tabela 17 registra os valores de Alfa de Cronbach obtidos para as três categorias.

Tabela 17 – Alfa de Cronbach para as respostas ao questionário.

Categoria	Questão	Média	Erro desvio	α^1
CAT1	Q1.1	4.40	1.93	
	Q1.2	5.22	1.78	
	Q1.3	4.23	1.95	0.818
	Q1.4	6.02	1.48	
	Q6	3.87	1.92	
CAT2	Q2.1	3.91	2.12	
	Q2.2	4.16	2.21	0.895
	Q2.3	3.37	2.00	
CAT3	Q4	1.79	0.71	0.447
	Q5	1.89	0.54	

¹Consistência interna medida pelo Alfa de Cronbach.

Verifica-se que, para a primeira e segunda categoria (CAT1 e CAT2), foram obtidas consistências internas quase perfeitas, o que aponta que as respostas dos estudantes foram muito consistentes e seguiram um padrão confiável. Já para a terceira categoria (CAT3) foi obtida uma consistência interna moderada, o que indica que houve certas inconsistências entre as respostas.

Uma segunda análise sobre a consistência interna foi medida de forma diferente, agrupada por turma. Essa análise é interessante para identificar se existe alguma inconsistência individualizada na turma que pode alterar os valores gerais. A Tabela 18 apresenta os valores de Alfa de Cronbach obtidos para cada uma das turmas que participaram do experimento por meio de questionários.

Tabela 18 – Alfa de Cronbach separado por turma.

Categoría	Questão	$\alpha(T1)^1$	$\alpha(T2)^2$	$\alpha(T3)^3$	$\alpha(T4)^4$
CAT1	Q1.1				
	Q1.2				
	Q1.3	0.792	0.880	0.811	0.796
	Q1.4				
CAT2	Q6				
	Q2.1				
	Q2.2	0.876	0.778	0.956	0.917
CAT3	Q2.3				
	Q4	0.545	0.532	-0.207	0.657
	Q5				

¹Consistência interna medida pelo Alfa de Cronbach para a Turma 1. ²Consistência interna medida pelo Alfa de Cronbach para a Turma 2. ³Consistência interna medida pelo Alfa de Cronbach para a Turma 3.

⁴Consistência interna medida pelo Alfa de Cronbach para a Turma 4.

Constata-se, nesta segunda análise da consistência interna, que para as categorias 1 e 2 (CAT1 e CAT2) não houve uma turma que pudesse, individualmente, representar uma discrepância nos valores obtidos. Neste caso, observa-se uma normalidade nos valores. No entanto, para a categoria 3 (CAT3), é possível notar um comportamento bem atípico na Turma 3, o que indica que esse valor (-0.207) influenciou negativamente para que a consistência interna dessa categoria fosse classificada como moderada. Se a Turma 3 fosse excluída desta análise na categoria 3, certamente os valores de Alfa de Cronbach seriam maiores, podendo, com isso, indicar uma consistência interna maior.

No geral, é possível assegurar que os questionários foram respondidos de forma muito consistente, indicando, com isso, alta confiabilidade nos resultados obtidos.

5.5.2.2 Avaliação sobre as Entrevistas

Como já descrito anteriormente, o roteiro de perguntas para guiar as entrevistas foi dividido em cinco categorias: (a) uso geral da plataforma; (b) interação com a funcionalidade de *quiz*; (c) opinião sobre a funcionalidade de visualização de desempenho; (d) padrão de respostas sobre os *quizzes*; e por fim, (e) considerações finais sobre a plataforma.

Para a primeira categoria (uso geral da plataforma), foi possível a utilização da plataforma por todos os estudantes que participaram da entrevista, alguns estudantes utilizaram de forma mais contínua e outros preferiam utilizar mais em períodos próximos às avaliações. Percebeu-se, também, que mesmo os que utilizaram pouco, gostaram muito das funcionalidades. Quando perguntado se acharam a funcionalidade útil, as respostas foram bastante positivas, como pode ser visto nas transcrições abaixo:

"Foi muito útil. A ideia de registrar as anotações ajuda muito na hora de

estudar... "

"Achei mais útil que o moodle, porque é mais simples, prático, já te dá os tópicos organizados e já tem os slides das disciplinas, isso fica mais fácil pra estudar para as provas..."

Para a segunda categoria (interação com a funcionalidade de *quiz*), foi possível perceber que todos os estudantes entrevistados aprovaram a funcionalidade. Quando questionado sobre o comportamento, a maioria dos estudantes respondia o *quiz* tão logo quanto percebia. Isso ajuda a reiterar a utilidade de funcionalidade. As transcrições abaixo trazem respostas de diferentes estudantes para a pergunta sobre comportamento relacionado à resposta do *quiz*.

"Sim, decorei todos eles. Muito positivo, porque avalia o conhecimento do aluno. Sempre que eu via um quiz eu tentava responder, se eu errasse eu ia lá e olhava."

"Quando eu via o quiz eu já tentava responder. Primeiro eu via o slide e já tentava responder o quiz."

Quando questionado se eles viram com facilidade a funcionalidade, alguns estudantes relataram que a interface poderia ser melhorada, conforme transcrição abaixo:

"A funcionalidade ajuda muito... Às vezes o quiz ficava um pouco difícil de visualizar, a na barra inferior não mostrava onde tinha os quiz."

Na terceira categoria (opinião sobre a funcionalidade de visualização de desempenho) grande parte dos estudantes considerou a funcionalidade muito útil. Logo em seguida, são apresentadas algumas transcrições da resposta para a pergunta sobre utilidade e alteração no comportamento:

"... Sim. Essa é uma parte muito boa, dá pra você elaborar uma metodologia de estudo e observar o que você está falhando e acertando e consegue dedicar mais tempo ao conteúdo que está errando. Achei excelente. Isso ia ser muito proveitoso em outras matérias, às vezes você está estudando um assunto que já está dominando e fica perdendo tempo e não faz nas partes que não está tão bem..."

"Sim, uma funcionalidade interessante. Dá pra você ver em qual conteúdo você está melhor em relação a outros tópicos..."

"Sim, acho que seria bem útil. Com base nessa informação eu consigo saber onde eu preciso estudar mais..."

Analizando as entrevistas, também ficou claro para essa categoria que a funcionalidade pode não ter ficado tão clara na interface. Muitos estudantes relataram que não viram ou, às vezes, não perceberam ao ver o ícone. Isso, com certeza, é um problema no design que pode ser analisado e melhorado.

Para a quarta categoria (padrão de respostas sobre os *quizzes*), foi perguntado se eles responderam ao mesmo *quiz* mais de uma vez. Essa pergunta tinha o intuito de verificar se os estudantes praticavam ações relacionadas com *Gaming the System*⁴. Neste caso, alguns alunos responderam que sim para fixar melhor o conteúdo. Já outros estudantes afirmaram que só respondiam o mesmo *quiz* se tivessem errado nas tentativas anteriores. É interessante que alguns estudantes voltavam ao conteúdo tentando realmente encontrar a resposta, e outros simplesmente ficavam fazendo “chutes” até encontrar a resposta correta. Algumas transcrições para essa categoria podem ser vistas abaixo:

“...Sim, mesmo acertando eu tentava outras vezes. Mas pra revisar...”

“Na verdade eu tentava responder, até acertar...”

É interessante notar que a maioria dos estudantes não voltava no conteúdo para revisar, e, consequentemente, não respondeu ao mesmo *quiz* mais de uma vez. Foi possível notar, no comportamento de alguns estudantes, que eles tentavam burlar o sistema realizando “chutes” aleatórios nas respostas.

Para a última categoria, foi dada a oportunidade dos estudantes sugerirem alterações para melhoria da plataforma. Muitas sugestões interessantes foram obtidas, dentre elas: possibilidade de impressão do conteúdo apresentado; adicionar indicador de *quiz* na barra de visualização dos slides; melhorar a apresentação para a visualização por meio de gráficos; e retirar a barra de visualização em miniatura. Um estudo sobre essas alterações será viabilizado e adicionado nos trabalhos futuros.

Em geral, a parte da análise por meio das entrevistas trouxe uma visão mais qualitativa para este experimento. Considera-se que estas entrevistas foram positivas e contribuíram, significativamente, para um melhor entendimento da percepção dos estudantes acerca dos recursos de visualização, bem como, uma visão geral da plataforma ubíqua com recursos de MAE. Os achados neste experimento irão guiar o processo de consolidação das ferramentas testadas e também proporcionará que implementações futuras estejam adequadas às considerações dos estudantes.

5.6 Considerações Finais

Este capítulo apresentou experimentos e resultados acerca da avaliação do modelo proposto neste trabalho. Foram planejados e executados quatro experimentos. Todos eles

⁴ comportamento no qual o estudante tenta burlar o sistema apresentando respostas caracterizadas por “chute”.

foram aplicados a usuários reais e apontou um alto rigor estatístico, no qual se utilizou de conceituados métodos para validar os resultados. Resultados do primeiro experimento mostraram que o modelo proposto é eficiente e apresentaram valores com alta correlação quando comparados com as notas reais dos estudantes. Outra conclusão diz respeito ao procedimento de avaliação por meio do comportamento, como mostraram os resultados do terceiro experimento, o comportamento é uma variável que pode ser perfeitamente utilizada para auxiliar a avaliação do estudante. Por fim, o segundo experimento mostrou quais são as melhores estruturas para construção das RBs, e o quarto experimento trouxe uma análise qualitativa, mostrando a satisfação dos estudantes ao utilizarem tanto o MAE quanto à plataforma ubíqua.

CAPÍTULO 6

Conclusões

Este trabalho traz uma abordagem híbrida, baseada em ontologias e Redes Bayesianas (RBs), para modelagem de estudante, com capacidade para determinar o desempenho do estudante com base no nível de aquisição de conhecimento e no seu comportamento durante a utilização de um ambiente virtual de ensino. No modelo proposto, uma RB é utilizada para determinar as relações e as hierarquias no domínio de conhecimento e para inferir a probabilidade de conhecimento¹ em cada tópico desse domínio. Já a ontologia, é utilizada para representar as informações acerca do estudante, bem como para realizar inferências² sobre o nível de desempenho do estudante, levando em consideração seu conhecimento e comportamento³.

Faz parte da abordagem proposta, a criação de um Modelo de Estudante (ME) que possui recursos de Visualização de Informação, o qual permite estender a proposta do modelo criado, dando-lhe características de um Modelo Aberto do Estudante (MAE). Esse modelo é capaz de gerenciar diferentes formas de visualização com o intuito de apresentar o nível de conhecimento para estudantes e professores. Com isso, os estudantes são capazes de autorregularem seus estudos e os professores, de identificarem a evolução das capacidades e limitações dos seus estudantes de forma dinâmica e contínua.

Como prova de conceito, o modelo proposto foi integrado ao ambiente de ensino Classroom eXperience (CX), uma plataforma educacional ubíqua capaz de capturar, sincronizar e apresentar diferentes mídias educacionais em diversos formatos. Como a abordagem propõe que os recursos sejam criados por meio de módulos fricamente acoplados, a integração aconteceu sem muitas alterações na forma como o ambiente funciona. A abordagem criada permite também a integração e utilização do ME proposto em outros ambientes educacionais de ensino, por meio de um modelo genérico de integração.

Para validar a abordagem, quatro experimentos foram propostos: o primeiro, teve como objetivo verificar a correlação entre os valores inferidos pelo Módulo Probabilístico

¹ Por meio de instrumentos avaliativos.

² Por meio de regras *Semantic Web Rule Language* (SWRL).

³ Por meio de atividades comportamentais como: socialização, colaboração e gamificação.

e os valores reais de conhecimento dos estudantes. Foi possível verificar que, na maioria das variáveis analisadas, a correlação apontou resultados muito satisfatórios, sendo que, 80% das amostras analisadas obtiveram correlações muito altas ou substanciais e nenhuma das amostras apresentou correlações baixas ou insignificantes.

O segundo experimento buscou identificar quais são as propriedades ou características de uma RB que influenciam positivamente na inferência de valores sobre o conhecimento do estudante. Foi possível determinar seis diretrizes que podem auxiliar o processo de construção das RBs para o modelo proposto.

O terceiro experimento teve como objetivo apresentar uma análise do desempenho dos estudantes com base em perfis de comportamento. Foram utilizados algoritmos de agrupamento para dividir as amostras em diferentes grupos e analisar se o comportamento dos estudantes influencia seus respectivos desempenhos. Os resultados mostraram que existe relação significativa entre as notas dos estudantes e os grupos criados com base nos comportamentos. O teste de Mann-Whitney mostrou que o nível de utilização do ambiente interfere diretamente sobre a média do desempenho dos estudantes ($U = 1103,00$; $p < 0,05$) e o teste de Kruskal-Wallis também mostrou que o comportamento dos estudantes interfere no seu desempenho ($H(2) = 7,063$; $p < 0,05$).

O quarto e último experimento teve como objetivo analisar o impacto na utilização de um MAE associado a uma plataforma ubíqua. Foi possível observar que a ferramenta trouxe diversos resultados positivos aos estudantes que utilizaram a plataforma. A maioria deles respondeu que a ferramenta pode influenciar positivamente no processo de ensino/aprendizagem e 83% responderam que gostariam que a ferramenta fosse utilizada também em outras disciplinas.

6.1 Contribuições

A proposta apresentada neste trabalho traz a concepção de uma nova abordagem para modelagem do estudante, a qual é estruturada de forma híbrida, baseando-se em RBs e ontologias com capacidades para avaliação de desempenho em Sistemas Adaptativos e Inteligentes para Educação (SAIEs). Além disso, esta nova abordagem traz uma extensão, que permite que o modelo proposto atue como um MAE, permitindo que as informações inferidas sejam apresentadas para os envolvidos no processo de ensino. Mais especificamente, o trabalho traz as seguintes contribuições:

- A definição de uma abordagem probabilística (baseada em RBs) e dinâmica (capaz de atualizar as informações do modelo a cada interação do estudante) para a identificação do nível de aquisição do conhecimento do estudante. Para essa abordagem, foi criado um modelo que estende os conceitos das RBs permitindo a inserção de nós conceitos e nós evidências, os quais, por meio de uma estrutura de pré-requisitos e

associações de instrumentos avaliativos, tornam possível identificar o nível de aquisição de conhecimento do estudante em determinado domínio (Ferreira et al., 2016a; Ferreira et al., 2016b). Experimentações foram realizadas com o intuito de investigar se existem correlações entre os valores determinados pelo modelo e os valores reais de conhecimento do estudante (Ferreira et al., 2016a). Essa contribuição está relacionada ao primeiro objetivo específico proposto na Introdução deste trabalho (Seção 1.3).

- A definição de uma abordagem semântica (baseada em ontologias), extensível e reutilizável (permite sua integração em outros ambientes) para a identificação do desempenho global do estudante (Ferreira et al., 2015). A proposta identifica o desempenho, pois leva em consideração medidas de nível de aquisição de conhecimento (por meio da abordagem probabilística) e níveis de comportamento (por meio de atividades de socialização, colaboração e gamificação). Para a inferência do desempenho, ontologias foram modeladas e regras SWRL foram implementadas com o intuito de predizer, com certo nível de fidelidade, o conhecimento do estudante (Ferreira et al., 2017c). Experimentos foram executados com o intuito de verificar se é realmente possível utilizar medidas comportamentais para determinar o desempenho do estudante. Tal contribuição está relacionada ao segundo objetivo específico descrito na Introdução deste trabalho (Seção 1.3).
- A elaboração de uma extensão, baseada em recursos de Visualização de Informação, que fornece capacidade para a abordagem proposta atuar como MAE, possibilitando a apresentação das informações inferidas pelo modelo. Essa extensão permite que estudantes e professores façam o monitoramento e o acompanhamento constante sobre o progresso da aprendizagem (Ferreira et al., 2017b). Novamente, experimentações foram produzidas com o intuito de verificar o impacto da proposta na percepção e no comportamento dos estudantes (Ferreira et al., 2017b). Essa contribuição está diretamente relacionada ao terceiro objetivo específico proposto na Introdução deste trabalho (Seção 1.3).
- A integração da abordagem proposta a um ambiente real de aprendizagem. Neste caso, foram implementados módulos de softwares que abstraíam as capacidades do modelo para realizar as tarefas propostas (Ferreira et al., 2017a). Esses módulos foram implementados e integrados à plataforma (Ferreira et al., 2017a; Ferreira et al., 2017b). O funcionamento dos módulos integrados se mostrou efetivo e permitiu que diversas turmas fizessem o uso da abordagem proposta. Além disso, um Modelo Genérico de Integração foi criado com o intuito de permitir que a abordagem proposta pudesse ser integrada em outros ambientes. Tal contribuição está associada ao quarto objetivo específico descrito na Introdução deste trabalho (Seção 1.3).

Ademais, a proposta apresentada neste trabalho traz elucidações suficientes para responder às questões de pesquisa levantadas:

Respondendo a Principal Questão de Pesquisa, é possível sim, a definição e utilização de uma abordagem híbrida, que se baseia em RBs e ontologias para a modelagem de estudante, com capacidade de determinar o nível de desempenho dos estudantes em um SAIE. Essa abordagem foi proposta, construída e integrada a um ambiente real de ensino, o que comprova a sua viabilidade. Resultados que respondem a essa questão de pesquisa podem ser obtidos em Ferreira et al., 2017a; Ferreira et al., 2017b; Ferreira et al., 2016a; Ferreira et al., 2016b; entre outros artigos resultantes deste trabalho.

Respondendo a primeira questão de pesquisa acessória, é possível, sim, identificar que existe impacto positivo quanto à utilização de comportamento para determinar desempenho em um SAIE. Foi possível comprovar tal afirmativa a partir dos resultados obtidos no Experimento 3 (Seção 5.4), no qual foi possível comprovar que existem estudantes com comportamentos diferentes e com médias de desempenho diferentes. Resultados que respondem a essa questão de pesquisa podem ser vistos em Ferreira et al., 2017a e Ferreira et al., 2017c.

Para a segunda questão de Pesquisa acessória, também foi possível constatar que as RBs, podem, sim, ser utilizadas para auxiliar na determinação do nível de aquisição de conhecimento do estudante. A resposta a essa questão foi comprovada por meio do Experimento 1 (Seção 5.2), e indiretamente, por meio do Experimento 2 (Seção 5.3). Esses experimentos mostraram que houve altos índices de correlação entre as notas reais dos estudantes e as notas inferidas pelo modelo, e que, na maioria dos casos analisados, predominaram as correlações muito altas ou substanciais. Resultados que respondem a essa questão de pesquisa podem ser encontrados em Ferreira et al., 2016a e Ferreira et al., 2017a.

6.2 Produção Bibliográfica

As seguintes publicações estão diretamente relacionadas a este trabalho:

- **Ferreira, H. N. M.**; Brant-Ribeiro, T.; Araújo, R. D.; Dorça, F. A.; Cattelan, R. G. (2017). An Automatic and Dynamic Knowledge Assessment Module for Adaptive Educational Systems. In (ICALT 2017) 17th IEEE International Conference on Advanced Learning Technologies. pp. 517–521 (Ferreira et al., 2017a) - **Qualis B1**.
- **Ferreira, H. N. M.**; Araújo, R. D.; Dorça, F. A.; Cattelan, R. G. (2017). Open Student Modeling for Academic Performance Visualization in Ubiquitous Learning Environments. In (SMC 2017) IEEE International Conference on Systems, Man, and Cybernetics. pp. 641–646 (Ferreira et al., 2017b) - **Qualis A2**.

- **Ferreira, H. N. M.**; Araújo, R. D.; Dorça, F. A.; Cattelan, R. G. (2017). Uma Abordagem Baseada em Ontologias para Modelagem e Avaliação do Estudante em Sistemas Adaptativos e Inteligentes para Educação. In (SBIE 2017) XXVIII Simpósio Brasileiro de Informática na Educação. pp. 1197–1206. (Ferreira et al., 2017c) - **Qualis B1**.
- **Ferreira, H. N. M.**; Brant-Ribeiro, T.; Araújo, R. D.; Dorça, F. A.; Cattelan, R. G. (2016). An Automatic and Dynamic Student Modeling Approach for Adaptive and Intelligent Educational Systems Using Ontologies and Bayesian Networks. In (ICTAI 2016) 28th IEEE International Conference on Tools with Artificial Intelligence. pp. 738–745. (Ferreira et al., 2016a) - **Qualis B1**.
- **Ferreira, H. N. M.**; Araújo, R. D.; Dorça, F. A.; Cattelan, R. G. (2016). Uma Abordagem Híbrida para Acompanhamento da Aprendizagem do Estudante Baseada em Ontologias e Redes Bayesianas em Sistemas Adaptativos para Educação. In (WAvalia 2016) IX Workshop sobre Avaliação e Acompanhamento da Aprendizagem em Ambientes Virtuais. pp. 447–456. (Ferreira et al., 2016b).
- **Ferreira, H. N. M.**; Araújo, R. D.; Souza, P.; Chagas, S.; Dorça, F. A.; Cattelan, R. G. (2015). Gamificação em Ambientes Educacionais Ubíquos. In (SBIE 2015) XXVI Simpósio Brasileiro de Informática na Educação. pp. 509–518. (Ferreira et al., 2015) - **Qualis B1**.
- **Ferreira, H. N. M.**; Dorça, F. A.; Cattelan, R. G. (2018). Evaluating the Performance of Students Based on Behavior Profiles in Ubiquitous Learning Environments. In (ToE) IEEE Transactions on Education. ISSN: 0018-9359. (**Submetido em 16/04/2018 - Qualis A2**).

As seguintes publicações estão indiretamente relacionadas a este trabalho:

- Araújo, R. D.; **Ferreira, H. N. M.**; Dorça, F. A.; Cattelan, R. G. (2018). A Hybrid Architecture for Adaptive, Intelligent, and Ubiquitous Educational Systems. In R. Zheng (Ed.), Digital Technologies and Instructional Design for Personalized Learning. pp. 120–144. Hershey, PA: IGI Global. (Araújo et al., 2018).
- Dorça, F. A.; Carvalho, V.; Mendes, M.; Araújo, R. D.; **Ferreira, H. N. M.**; Cattelan, R. G. (2017). An Approach for Automatic and Dynamic Analysis of Learning Objects Repositories Through Ontologies and Data Mining Techniques for Supporting Personalized Recommendation of Content in Adaptive and Intelligent Educational Systems. In (ICALT 2017) 17th IEEE International Conference on Advanced Learning Technologies. pp. 514–516. (Dorça et al., 2017) - **Qualis B1**.

- Carvalho, V.; Araújo, R. D.; **Ferreira, H. N. M.**; Cattelan, R. G.; Dorça, F. A. (2017). OntAES: Uma Ontologia para Sistemas Adaptativos Educacionais Baseado em Objetos de Aprendizagem e Estilos de Aprendizagem. In (SBIE 2017) XXVIII Simpósio Brasileiro de Informática na Educação. pp. 1307–1316. (Carvalho et al., 2017) - **Qualis B1**.
- Araújo, R. D.; **Ferreira, H. N. M.**; Dorça, F. A.; Cattelan, R. G. (2016). Learning Objects Authoring Supported by Ubiquitous Learning Environments. In (IUI 2016) XXI International Conference on Intelligent User Interfaces. pp. 49–53. (Araújo et al. 2016a) - **Qualis A2**.
- Carvalho, V.; Mendes, M.; **Ferreira, H. N. M.**; Dorça, F. A.; Cattelan, R. G. (2016). Uma Ontologia para Apoio à Recomendação Automática e Personalizada de Conteúdo Considerando Estilos de Aprendizagem de Estudantes em Sistemas Adaptativos para Educação. In (SBIE 2016) XXVII Simpósio Brasileiro de Informática na Educação. pp. 1175–1184. (Carvalho et al., 2016) - **Qualis B1**.
- Araújo, R. D.; Brant-Ribeiro, T.; **Ferreira, H. N. M.**; Dorça, F. A.; Cattelan, R. G. (2016). Segmentação Colaborativa de Objetos de Aprendizagem Utilizando Bookmarks em Ambientes Educacionais Ubíquos. In (SBIE 2016) XXVII Simpósio Brasileiro de Informática na Educação. pp. 1205–1214. (Araújo et al., 2016b) - **Qualis B1**

6.3 Limitações

Algumas limitações foram observadas ao longo do desenvolvimento do trabalho, sejam por questões de escopo da proposta ou por questões mais relacionadas às implementações das abordagens.

A primeira limitação está especificamente integrada no processo de criação das RBs. Na abordagem proposta, estas devem ser criadas de forma manual pelos professores das disciplinas. Por ser um modelo matemático restrito a um domínio muito específico, faz-se necessária a presença de um especialista sobre o domínio a ser modelado. Mesmo que muito específico, este é um problema de difícil solução. Uma estratégia para contornar tal limitação pode ser obtida a partir do Experimento 2 (Seção 5.3), que busca definir diretrizes (*guidelines*) que podem auxiliar na criação dessas redes. Outra estratégia para contornar esta limitação seria a abordagem semi-automática descrita na Seção 6.4.

A segunda limitação deste trabalho está relacionada a questões de implementação à complexidade do processamento de informações por meio das RBs. Destaca-se também que este não é um problema do modelo e sim uma limitação devido à complexidade das RBs. Esta limitação é percebida durante a apresentação das informações pelo MAE, onde, no caso de redes muito grandes, o processamento sobre o nível de aquisição de

conhecimento pode demorar um tempo considerável (pelos testes realizados, no máximo 8 segundos, no caso do histórico do estudante). Para contornar tal situação, foi definida uma estratégia para se armazenar as informações previamente acessadas em Banco de Dados. Esta estratégia evita que a mesma informação seja processada várias vezes.

A terceira e última limitação do trabalho está relacionada às turmas participantes dos experimentos, uma vez que todas elas foram oriundas de cursos tecnológicos (Ciência da Computação ou Sistemas de Informação) ministradas por professores da área de Computação. Mesmo sendo uma limitação, o fato dos estudantes terem um conhecimento em assuntos computacionais, não invalida a proposta, pelo contrário, no caso do Experimento 4 (Seção 5.5) eles até contribuíram de forma muito positiva na identificação de alguns problemas de design de interfaces.

No contexto deste trabalho, é importante destacar que estas limitações são intrínsecas ao processo de desenvolvimento da abordagem proposta. Vale ressaltar que tais limitações não influenciam na qualidade real do trabalho, pois na sua maioria, estão relacionadas a questões de implementações e disponibilidade de realização dos experimentos, e não a uma abordagem estrutural do modelo proposto.

6.4 Trabalhos Futuros

Devido a sua abrangência, este trabalho abre diversas possibilidades para investigações futuras.

Um primeiro trabalho que pode ser investigado a partir da abordagem proposta diz respeito à capacidade de criação e ajustes das tabelas de probabilidade das RBs de forma semi-automática. Como apresentado na seção anterior, a criação da RB de forma manual pelos professores pode ser considerada uma limitação. Com isso em mente, uma possibilidade para contornar tal limitação seria uma abordagem de criação de RBs de forma semi-automática, explorando as capacidades sociais e colaborativas dos SAIEs. Nesta abordagem, seria possível que o professor criasse um domínio de conhecimento inicial, e, a partir das interações dos estudantes, o sistema realizaria perguntas com o intuito de ajustar a relação entre os nós (tópicos de conhecimento) da rede. Com as diversas respostas dos estudantes, seria possível realizar uma medida dos pesos de cada relação e determinar o grau de dependência entre esses nós. Assim, seria possível ajustar a probabilidade de ocorrência dos eventos para a tabela de probabilidade. Acredita-se que esta abordagem de criação de RBs de forma semi-automática, social e colaborativa traria diversas contribuições, não só para o contexto da modelagem de estudante, mas também para outras áreas que estão diretamente relacionadas com a modelagem de incertezas por meio das RBs.

Outro trabalho possível de ser explorado diz respeito à criação de mecanismos para recomendação automática de conteúdo com base nas informações processadas pelo MAE.

Uma das características do MAE é identificar e apresentar as informações sobre o nível de desempenho do estudante. Com isso, seria possível a criação de uma abordagem que identificasse as principais limitações dos estudantes e recomendasse Objetos de Aprendizagem (OAs) relacionados com os conceitos nos quais o estudante está com dificuldade. Como o nível de desempenho está relacionado com um determinado tópico, e este está relacionado diretamente com o material educacional, é possível identificar quais os OAs estão relacionados com o tópico de dificuldade do estudante. Acredita-se que essa funcionalidade poderia ser utilizada como uma abordagem pedagógica para auxiliar o estudante a encontrar, de forma mais fácil, uma maneira para remediar as suas dificuldades.

Uma terceira possibilidade de trabalho futuro está relacionada com a recomendação automática de abordagens pedagógicas para o professor da disciplina. Como na sugestão apresentada anteriormente, uma das capacidades do MAE é indicar ao professor quais os tópicos de conhecimento os estudantes estão com mais dificuldade. A partir disso, é possível relacionar os tópicos mais difíceis com o material educacional e sugerir ao professor que dê mais atenção a determinado ponto do material. Assim, o professor seria orientado a tomar atitudes pontuais em determinados tópicos de conhecimento, permitindo, com isso, estratégias mais eficientes do ponto de vista educacional.

Outra possibilidade de extensão da abordagem proposta com a realização de trabalhos futuros é o desenvolvimento de novas formas de apresentação das informações para professores e estudantes. Atualmente, esta versão do trabalho utiliza três gráficos (linha, barra e radar) e algumas tabelas para visualização das informações geradas pelo modelo. A ideia é que no futuro possam ser disponibilizadas mais algumas formas de visualização, como: *skillmeters* (Capacitadores), mapas de conhecimento, gráficos de área, gráficos de dispersão, entre outros. Como cada pessoa possui particularidades de preferências em determinados tipos de visualização e o desenvolvimento de ferramentas na área de Computação e Visualização de Informação está cada dia mais integrado, esta possibilidade de extensão surge como um importante trabalho para ser explorado e investigado no futuro.

Uma avaliação de usabilidade detalhada é outra investigação que pode ser explorada no futuro. Esse requisito foi levantado durante a realização do Experimento 4 (Seção 5.5), no qual foi possível perceber a existência de algumas inconsistências no design das interfaces, principalmente para os estudantes. Neste caso, um importante estudo diz respeito à realização de uma avaliação de usabilidade criteriosa, com a participação de especialistas em Interação Humano-Computador e usuários da plataforma. Em consonância com tal avaliação, poderão ser desenvolvidas adaptações necessárias com o intuito de melhorar a experiência dos usuários durante a realização de suas tarefas.

Por fim, e não menos importante, um estudo que merece atenção refere-se à proposta de estender a utilização do modelo proposto, aplicando-o em outros ambientes virtuais de aprendizagem, e consequentemente, a outras turmas e cursos. Esta proposta está em concordância com a última limitação levantada na seção anterior, a qual se refere à di-

versidade de turmas utilizadas nos experimentos. O objetivo desta abordagem seria uma maior heterogeneidade nas avaliações, o que permitiria, por exemplo, avaliar o comportamento dos estudantes em diferentes cursos ou diferentes níveis de educação (fundamental, médio, técnico e superior). Além disso, esse possível trabalho futuro abordaria questões relacionadas com a extensão do modelo proposto a outros domínios.

Referências

ABOWD, G. D. et al. Teaching and learning as multimedia authoring: the classroom 2000 project. In: **Proceedings of the 4th ACM International Conference on Multimedia (MM)**. New York, NY: ACM, 1997. p. 187–198. DOI: <https://doi.org/10.1145/244130.244191>.

AL-HMOUZ, A. et al. Modeling mobile learning system using ANFIS. In: **Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies (ICALT)**. Washington, DC: IEEE, 2011. p. 378–380. DOI: <https://doi.org/10.1109/ICALT.2011.119>.

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: Pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação, Brasília**, SciELO Brasil, v. 32, n. 3, p. 7–20, 2003. ISSN 1518-8353. DOI: <http://dx.doi.org/10.1590/S0100-19652003000300002>.

ARAÚJO, R. D. et al. Personalization of interactive digital media in ubiquitous educational environments. In: **Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC)**. Manchester, UK: IEEE, 2013. p. 3955–3960. DOI: <https://doi.org/10.1109/SMC.2013.675>.

_____. Autoria Automática de Objetos de Aprendizagem a Partir de Captura Multimídia e Associação a Estilos de Aprendizagem. In: **Anais do Simpósio Brasileiro de Informática na Educação**. Dourados, MS: SBC, 2014. p. 229–238. DOI: <http://dx.doi.org/10.5753/cbie.sbie.2014.229>.

BARRA, M.; NEGRO, A.; SCARANO, V. When the Teacher Learns: A Model for Symmetric Adaptivity. In: **Proceedings of the 2nd Workshop on Adaptive Systems and User Modeling on the World Wide Web**. Banff, CA: Springer, 1999. p. 21–28.

BAUTISTA-ZAMBRANA, M. R. Methodologies to Build Ontologies for Terminological Purposes. **Procedia - Social and Behavioral Sciences**, v. 173, p. 264–269, 2015. ISSN 1877-0428. DOI: <https://doi.org/10.1016/j.sbspro.2015.02.063>.

BEEL, J. et al. Research-paper Recommender Systems: a Literature Survey. **International Journal on Digital Libraries**, v. 17, n. 4, p. 305–338, 2016. ISSN 1432-1300. DOI: <https://doi.org/10.1007/s00799-015-0156-0>.

BORST, W. N. **Construction of Engineering Ontologies for Knowledge sharing and Reuse**. Centre for Telematics and Information Technology: Universiteit Twente, 1997.

BOTICARIO, J. G.; SANTOS, O.; ROSMALEN, P. V. Issues in developing standard-based adaptive learning management systems. In: **EADTU 2005 Working Conference: Towards Lisbon**. [S.l.: s.n.], 2010.

BRUSILOVSKY, P.; MILLA'N, E. The adaptive web: Methods and strategies of web personalization. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. cap. User Models for Adaptive Hypermedia and Adaptive Educational Systems, p. 3–53. DOI: https://doi.org/10.1007/978-3-540-72079-9_1.

BRUSILOVSKY, P. et al. Open social student modeling for personalized learning. **IEEE Transactions on Emerging Topics in Computing**, v. 4, n. 3, p. 450–461, 2016. ISSN 2168-6750. DOI: <https://doi.org/10.1109/TETC.2015.2501243>.

BULL, S.; KAY, J. Student Models That Invite the Learner In: The SMILI:() Open Learner Modelling Framework. **International Journal of Artificial Intelligence in Education**, v. 17, n. 2, p. 89–120, 2007. ISSN 1560-4292.

_____. Open Learner Models as Drivers for Metacognitive Processes. In: **International Handbook of Metacognition and Learning Technologies**. New York, NY: Springer, 2013. p. 349–365. ISBN 978-1-4419-5546-3. DOI: https://doi.org/10.1007/978-1-4419-5546-3_23.

CARBONELL, J. R. AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction. **IEEE Transactions on Man-Machine Systems**, v. 11, n. 4, p. 190–202, 1970. ISSN 0536-1540. DOI: <https://doi.org/10.1109/TMMS.1970.299942>.

CASALS, A.; BRANDÃO, A. A. F. Modeling a Mobile Learning Context Data Ontology. In: **Proceedings of the IEEE World Engineering Education Conference (EDUNINE)**. Santos, Brazil: IEEE, 2017. p. 71–75. DOI: <https://doi.org/10.1109/EDUNINE.2017.7918185>.

- CETINTAS, S. et al. Automatic Detection of Off-Task Behaviors in Intelligent Tutoring Systems with Machine Learning Techniques. **IEEE Transactions on Learning Technologies**, v. 3, n. 3, p. 228–236, 2010. ISSN 1939-1382. DOI: <https://doi.org/10.1109/TLT.2009.44>.
- CHEN, Y.; WUILLEMIN, P.-H.; LABAT, J.-M. Bayesian Student Modeling Improved by Diagnostic Items. In: **Proceedings of the 12th International Conference on Intelligent Tutoring Systems (ITS)**. Honolulu, HI: Springer, 2014. p. 144–149. DOI: https://doi.org/10.1007/978-3-319-07221-0_17.
- CHEN, Z.-H. et al. Active open learner models as animal companions: Motivating children to learn through interacting with my-pet and our-pet. **International Journal of Artificial Intelligence in Education**, v. 17, n. 2, p. 145–167, 2007. ISSN 1560-4292.
- CHIEU, V. M. et al. Student modeling in orthopedic surgery training: Exploiting symbiosis between temporal bayesian networks and fine-grained didactic analysis. **International Journal of Artificial Intelligence in Education**, v. 20, n. 3, p. 269–301, 2010. ISSN 1560-4292.
- CHRYSAFIADI, K.; VIRVOU, M. Personalized Teaching of a Programming language over the web: Stereotypes and rule-based mechanisms. In: **Proceedings of the 8th Joint Conference on Knowledge-Based Software Engineering**. Amsterdam: IOS, 2008. p. 484–492.
- _____. Evaluating the Integration of Fuzzy Logic into the Student Model of a Web-based Learning Environment. **Expert Systems with Applications**, v. 39, n. 18, p. 13127–13134, 2012. ISSN 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2012.05.089>.
- _____. PeRSIVA: An Empirical Evaluation Method of a Student Model of an Intelligent e-learning Environment for Computer Programming. **Computers & Education**, v. 68, p. 322–333, 2013. ISSN 0360-1315. DOI: <https://doi.org/10.1016/j.compedu.2013.05.020>.
- _____. Student modeling approaches: A literature review for the last decade. **Expert Systems with Applications**, v. 40, n. 11, p. 4715–4729, 2013. ISSN 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2013.02.007>.
- _____. Fuzzy Logic for Adaptive Instruction in an E-learning Environment for Computer Programming. **IEEE Transactions on Fuzzy Systems**, v. 23, n. 1, p. 164–177, 2015. ISSN 1063-6706. DOI: <https://doi.org/10.1109/TFUZZ.2014.2310242>.
- CLEMENTE, J.; RAMÍREZ, J.; ANTONIO, A. de. A proposal for student modeling based on ontologies and diagnosis rules. **Expert Systems with Applications**, v. 38, n. 7, p. 8066–8078, 2011. ISSN 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2010.12.146>.

CONATI, C.; GERTNER, A.; VANLEHN, K. Using bayesian networks to manage uncertainty in student modeling. **User Modeling and User-Adapted Interaction**, v. 12, n. 4, p. 371–417, 2002. ISSN 1573-1391. DOI: <https://doi.org/10.1023/A:1021258506583>.

CORBETT, A. T.; ANDERSON, J. R. Knowledge tracing: Modeling the acquisition of procedural knowledge. **User Modeling and User-Adapted Interaction**, v. 4, n. 4, p. 253–278, 1994. ISSN 1573-1391. DOI: <https://doi.org/10.1007/BF01099821>.

DAVIS, J. A. **Elementary Survey Analysis**. [S.l.]: Prentice-Hall, Englewood Cliffs, NJ, 1971.

DESMARAIS, M. C.; BAKER, R. S. A review of recent advances in learner and skill modeling in intelligent learning environments. **User Modeling and User-Adapted Interaction**, v. 22, n. 1-2, p. 9–38, 2012. ISSN 1573-1391. DOI: <https://doi.org/10.1007/s11257-011-9106-8>.

DICKSON, P. E. et al. Student Reactions to Classroom Lecture Capture. In: **Proceedings of the 17th ACM Annual Conference on Innovation and Technology in Computer Science Education**. New York, NY: ACM, 2012. p. 144–149. DOI: <https://doi.org/10.1145/2325296.2325334>.

DIMITROVA, V.; SELF, J.; BRNA, P. Applying interactive open learner models to learning technical terminology. In: BAUER, M.; GMYTRASIEWICZ, P. J.; VASSILEVA, J. (Ed.). **User Modeling 2001**. Berlin, Heidelberg: Springer, 2001. p. 148–157. ISBN 978-3-540-44566-1. DOI: https://doi.org/10.1007/3-540-44566-8_15.

DOLOG, P.; NEJDL, W. The adaptive web. In: BRUSILOVSKY, P.; KOBSA, A.; NEJDL, W. (Ed.). Berlin, Heidelberg: Springer, 2007. cap. Semantic Web Technologies for the Adaptive Web, p. 697–719. ISBN 978-3-540-72078-2. DOI: https://doi.org/10.1007/978-3-540-72079-9_23.

DURLACH, A. M. L. P. J. **Adaptive Technologies for Training and Education**. New York, NY: Cambridge University Press, 2011. ISBN 9781139049580.

DURRANI, S.; DURRANI, D. Intelligent tutoring systems and cognitive abilities. In: **Proceedings of Graduate Colloquium on Computer Sciences**. Vol. 1. FAST-NU Lahore: Department of Computer Science. 2010.

EHIMWENMA, K. E.; BEER, M.; CROWTHER, P. Student modelling and classification rules learning for educational resource prediction in a multiagent system. In: **Proceedings of the 7th Computer Science and Electronic Engineering Conference (CEEC)**. Colchester, UK: IEEE, 2015. p. 59–64. DOI: <https://doi.org/10.1109/CEEC.2015.7332700>.

- FELDER, R. M.; SILVERMAN, L. K. et al. Learning and teaching styles in engineering education. **Engineering education**, v. 78, n. 7, p. 674–681, 1988.
- FERREIRA, H. N. M. et al. Classroom Experience: A Platform for Multimedia Capture and Access in Instrumented Educational Environments. In: **Proceedings of the 2012 Brazilian Symposium on Collaborative Systems**. São Paulo, SP: IEEE, 2012. p. 59–64. DOI: <https://doi.org/10.1109/SBSC.2012.20>.
- _____. Gamificação em ambientes educacionais ubíquos. In: **Proceedings of the 2015 Brazilian Symposium on Computers in Education**. Maceió, AL: SBC, 2015. p. 509–518. DOI: <http://dx.doi.org/10.5753/cbie.sbie.2015.509>.
- GAUDIOSO, E.; MONTERO, M.; OLMO, F. H. del. Supporting teachers in adaptive educational systems through predictive models: A proof of concept. **Expert Systems with Applications**, v. 39, n. 1, p. 621–625, 2012. ISSN 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2011.07.052>.
- GILMAN, E. et al. Towards user support in ubiquitous learning systems. **IEEE Transactions on Learning Technologies**, v. 8, n. 1, p. 55–68, 2015. ISSN 1939-1382. DOI: <https://doi.org/10.1109/TLT.2014.2381467>.
- GOEL, G.; LALLÉ, S.; LUENGO, V. Fuzzy logic representation for student modelling: Case study on geometry. In: **Proceedings of the 11th International Conference on Intelligent Tutoring Systems (ITS)**. Berlin, Heidelberg: Springer-Verlag, 2012. p. 428–433. DOI: https://doi.org/10.1007/978-3-642-30950-2_55.
- GOGUADZE, G. et al. Evaluating a bayesian student model of decimal misconceptions. In: **Proceedings of the 4th International Conference on Educational Data Mining (EDM)**. Eindhoven: EDM, 2011. p. 301–306.
- GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge acquisition**, Academic Press Ltd., v. 5, n. 2, p. 199–220, 1993. ISSN 1042-8143. DOI: <https://doi.org/10.1006/knac.1993.1008>.
- GRUBIŠIĆ, A.; STANKOV, S.; PERAIĆ, I. Ontology based approach to bayesian student model design. **Expert Systems with Applications**, v. 40, n. 13, p. 5363–5371, 2013. ISSN 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2013.03.041>.
- GRÜNINGER, M.; FOX, M. Methodology for the Design and Evaluation of Ontologies. In: **IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing**. Montreal, CA: AAAI Press, 1995.
- GUERRA, J. et al. An intelligent interface for learning content: Combining an open learner model and social comparison to support self-regulated learning

and engagement. In: **Proceedings of the 21st International Conference on Intelligent User Interfaces**. New York, NY: ACM, 2016. p. 152–163. DOI: <https://doi.org/10.1145/2856767.2856784>.

HAIR, J.; ANDERSON, R.; BABIN, B. **Multivariate Data Analysis**. 7. ed. [S.l.]: Prentice Hall, 2009.

HAWKINS, W. J.; HEFFERNAN, N. T.; BAKER, R. S. Learning bayesian knowledge tracing parameters with a knowledge heuristic and empirical probabilities. In: **Proceedings of the 12th International Conference on Intelligent Tutoring Systems (ITS)**. Honolulu, HA: Springer, 2014. p. 150–155. DOI: https://doi.org/10.1007/978-3-319-07221-0_18.

HSIAO, I.-H. et al. Open social student modeling: Visualizing student models with parallel introspectiveviews. In: **Proceedings of the 19th International Conference on User Modeling, Adaption, and Personalization (UMAP)**. Berlin, Heidelberg: Springer-Verlag, 2011. p. 171–182. DOI: https://doi.org/10.1007/978-3-642-22362-4_15.

_____. Progressor: social navigation support through open social student modeling. **New Review of Hypermedia and Multimedia**, v. 19, n. 2, p. 112–131, 2013. ISSN 1740-7842. DOI: <http://dx.doi.org/10.1080/13614568.2013.806960>.

IGLESIAS, A. et al. Learning teaching strategies in an adaptive and intelligent educational system through reinforcement learning. **Applied Intelligence**, v. 31, n. 1, p. 89–106, 2009. ISSN 1573-7497. DOI: <https://doi.org/10.1007/s10489-008-0115-1>.

ISOTANI, S.; BITTENCOURT, I. I. **Dados Abertos Conectados: Em busca da Web do Conhecimento**. São Paulo, SP: Novatec Editora, 2015. DOI: <https://doi.org/10.13140/RG.2.1.4355.6329>.

JAIN, A. K. Data clustering: 50 years beyond k-means. **Pattern Recognition Letters**, v. 31, n. 8, p. 651 – 666, 2010. ISSN 0167-8655. DOI: <https://doi.org/10.1016/j.patrec.2009.09.011>.

JEREMIĆ, Z.; JOVANOVIĆ, J.; GAVSEVIĆ, D. Student modeling and assessment in intelligent tutoring of software patterns. **Expert Systems with Applications**, v. 39, n. 1, p. 210–222, 2012. ISSN 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2011.07.010>.

KÄSER, T.; HALLINEN, N. R.; SCHWARTZ, D. L. Modeling exploration strategies to predict student performance within a learning environment and beyond. In: **Proceedings of the 7th International Conference on Learning Analytics & Knowledge (LAK)**. New York, NY: ACM, 2017. p. 31–40. DOI: <https://doi.org/10.1145/3027385.3027422>.

- KASER, T. et al. Dynamic bayesian networks for student modeling. **IEEE Transactions on Learning Technologies**, v. 10, n. 4, p. 450–462, 2017. ISSN 1939-1382. DOI: <https://doi.org/10.1109/TLT.2017.2689017>.
- KLAŠNJA-MILIĆEVIĆ, A. et al. Personalization and adaptation in e-learning systems. In: **E-Learning Systems: Intelligent Techniques for Personalization**. Cham: Springer, 2017. p. 21–25. ISBN 978-3-319-41163-7. DOI: https://doi.org/10.1007/978-3-319-41163-7_2.
- KLERK, S. de; EGGEN, T. J.; VELDKAMP, B. P. A methodology for applying students' interactive task performance scores from a multimedia- based performance assessment in a bayesian network. **Computers in Human Behavior**, v. 60, p. 264–279, 2016. ISSN 0747-5632. DOI: <https://doi.org/10.1016/j.chb.2016.02.071>.
- KUSUMAWARDANI, S. S.; PRAKOSO, R. S.; SANTOSA, P. I. Using ontology for providing content recommendation based on learning styles inside e-learning. In: **Proceedings of the 2nd International Conference on Artificial Intelligence, Modelling and Simulation**. Madrid, Spain : IEEE, 2014. p. 276–281. DOI: <https://doi.org/10.1109/AIMS.2014.40>.
- LAMB, R. L. et al. A computational modeling of student cognitive processes in science education. **Computers & Education**, v. 79, p. 116–125, 2014. ISSN 0360-1315. DOI: <https://doi.org/10.1016/j.compedu.2014.07.014>.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, International Biometric Society, v. 33, n. 1, 1977. DOI: <https://doi.org/10.2307/2529310>.
- LI, N. et al. A machine learning approach for automatic student model discovery. In: **Proceedings of the 4th International Conference on Educational Data Mining (EDM)**. Eindhoven: EDM, 2011. p. 31–40.
- LIN, C.; SHEN, S.; CHI, M. Incorporating student response time and tutor instructional interventions into student modeling. In: **Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization (UMAP)**. New York, NY, 2016. p. 157–161.
- LOBODA, T. D. et al. Mastery grids: An open source social educational progress visualization. In: **European Conference on Technology Enhanced Learning**. Graz, Austria: Springer, 2014. p. 235–248. DOI: https://doi.org/10.1007/978-3-319-11200-8_18.
- LONG, Y.; ALEVEN, V. Supporting students' self-regulated learning with an open learner model in a linear equation tutor. In: **International Conference on**

Artificial Intelligence in Education. Memphis, Tennessee: AIED, 2013. p. 219–228. DOI: https://doi.org/10.1007/978-3-642-39112-5_23.

_____. Enhancing learning outcomes through self-regulated learning support with an open learner model. **User Modeling and User-Adapted Interaction**, v. 27, n. 1, p. 55–88, 2017. ISSN 1573-1391. DOI: <https://doi.org/10.1007/s11257-016-9186-6>.

LOPEZ, M. F.; PEREZ, A. G.; JURISTO, N. METHONTOLOGY: from Ontological Art towards Ontological Engineering. In: **Proceedings of the American Asociation for Artificial Intelligence**. Stanford, USA: MIT Press, 1997. p. 33–40.

MAHNANE, L.; LASKRI, M. T. An adaptive hypermedia system integrating thinking style (ahs-ts): Model and experiment. **International Journal of Hybrid Information Technology**, v. 5, n. 1, p. 11–28, 2012. ISSN 1738-9968

MARTINEZ-MALDONADO, R. et al. The latux workflow: Designing and deploying awareness tools in technology-enabled learning settings. In: **Proceedings of the 5th International Conference on Learning Analytics & Knowledge (LAK)**. New York, NY: ACM, 2015. p. 1–10.

MAZZA, R.; DIMITROVA, V. Coursevis: A graphical student monitoring tool for supporting instructors in web-based distance courses. **International Journal of Human-Computer Studies**, v. 65, n. 2, p. 125–139, 2007. ISSN 1071-5819. DOI: <https://doi.org/10.1016/j.ijhcs.2006.08.008>.

MENDONÇA, I. E. et al. Explorando Funcionalidades Sociais e Colaborativas em Ambientes Educacionais Ubíquos. In: **Proceedings of the 2014 Brazilian Symposium on Computers in Education**. Dourados, MS: SBC, 2014. p. 486–495. DOI: <http://dx.doi.org/10.5753/cbie.sbie.2014.486>.

MILLÁN, E.; CRUZ, J. L. Pérez-de-la; GARCÍA, F. Dynamic versus static student models based on bayesian networks: An empirical study. In: **Proceedings of the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems**. Oxford, UK: Springer, 2003. p. 1337–1344. DOI: https://doi.org/10.1007/978-3-540-45226-3_181.

MILLÁN, E. et al. Learning Bayesian Networks for Student Modeling. In **Proceedings of the International Conference on Artificial Intelligence in Education (AIED)**. Madrid, Spain: AIED, 2015. p. 718–721. DOI: https://doi.org/10.1007/978-3-319-19773-9_100.

MILLÁN, E.; LOBODA, T.; CRUZ, J. L. Pérez-de-la. Bayesian networks for student model engineering. **Computers & Education**, v. 55, n. 4, p. 1663–1683, 2010. ISSN 0360-1315. DOI: <https://doi.org/10.1016/j.compedu.2010.07.010>.

- MINOVIĆ, M. et al. Visualisation of student learning model in serious games. **Computers in Human Behavior**, v. 47, p. 98–107, 2015. ISSN 0747-5632. DOI: <https://doi.org/10.1016/j.chb.2014.09.005>.
- MITROVIC, A.; MARTIN, B. Evaluating the effect of open student models on self-assessment. **International Journal of Artificial Intelligence in Education**, v. 17, n. 2, p. 121–144, 2007. ISSN 1560-4306.
- MITROVIC, A.; THOMSON, D. Towards a negotiable student model for constraint-based itss. **Proceedings of the 17th International Conference on Computers in Education**, 2009. p. 83–90.
- MORAIS, E. A. M.; AMBRÓSIO, A. P. L. **Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens**. [S.l.], 2007.
- MOURI, K. et al. Bayesian network for predicting students' final grade using e-book logs in university education. In: **Proceedings of the 16th International Conference on Advanced Learning Technologies (ICALT)**. Kraków, Poland: IEEE, 2016. p. 85–89. DOI: <https://doi.org/10.1109/ICALT.2016.27>.
- NAFEA, S. et al. Personalized students' profile based on ontology and rule-based reasoning. **EAI Endorsed Transactions on e-Learning**, v. 16, n. 12, p. 1–18. 2016. ISSN 2032-9253. DOI: <https://doi.org/10.4108/eai.2-12-2016.151720>.
- NAVARRO, M. R. **Procesos cognitivos y aprendizaje significativo**. [S.l.]: Comunidad de Madrid. Consejería de Educación. Viceconsejería de Organización Educativa, 2008.
- NGUYEN, C. D. et al. An ontology-based it student model in an educational social network. In: **Proceedings of the 13th International Conference on Information Integration and Web-based Applications and Services**. New York, NY, 2011. p. 379–382. DOI: <https://doi.org/10.1145/2095536.2095609>.
- NOY, N. F.; MCGUINNESS, D. L. **Ontology Development 101: A Guide to Creating Your First Ontology**. 2001.
- OUF, S. et al. A proposed paradigm for smart learning environment based on semantic web. **Computers in Human Behavior**, v. 72, p. 796–818, 2017. ISSN 0747-5632. DOI: <https://doi.org/10.1016/j.chb.2016.08.030>.
- PANAGIOTOPoulos, I. et al. An ontology-based model for student representation in intelligent tutoring systems for distance learning. **Proceedings of the International Conference on Artificial Intelligence Applications and Innovations**, 2012. p. 296–305. DOI: https://doi.org/10.1007/978-3-642-33409-2_31.

PEARL, J. Bayesian networks: A model of self-activated memory for evidential reasoning. In: **Proceedings of the 7th Conference of the Cognitive Science Society**, [S.l.: s.n.], 1985. p. 329–334.

_____. **Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988. ISBN 0-934613-73-7.

PIMENTEL, M. d. G. et al. Supporting educational activities through dynamic web interfaces. **Interacting with computers**, Elsevier, v. 13, n. 3, p. 353–374, 2001. ISSN 1873-7951. DOI: [https://doi.org/10.1016/S0953-5438\(00\)00042-4](https://doi.org/10.1016/S0953-5438(00)00042-4).

PUKKHEM, N. Ontology-based semantic approach for learning object recommendation. **International Journal on Information Technology**, Association of Computer Electronics and Electrical Engineers (ACEEE), v. 3, n. 4, p. 12–21, 2013.

RAMÍREZ-NORIEGA, A.; JUÁREZ-RAMÍREZ, R.; MARTÍNEZ-RAMÍREZ, Y. Evaluation module based on bayesian networks to intelligent tutoring systems. **International Journal of Information Management**, v. 37, n. 1, p. 1488–1498, 2017. ISSN 0268-4012. DOI: <https://doi.org/10.1016/j.ijinfomgt.2016.05.007>.

RAZALI, N. M.; WAH, Y. B. et al. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. **Journal of Statistical Modeling and Analytics**, v. 2, n. 1, p. 21–33, 2011.

REZENDE, P. A. et al. Personna: proposta de ontologia de contexto e perfil de alunos para recomendação de objetos de aprendizagem. **Revista Brasileira de Informática na Educação**, v. 23, n. 01, p. 70–84, 2015. ISSN 1414-5685. DOI: <http://dx.doi.org/10.5753/rbie.2015.23.01.70>.

RICH, E. Users are individuals: individualizing user models. **International Journal of Human-Computer Studies**, v. 51, n. 2, p. 323–338, 1999. ISSN 1071-5819. DOI: [http://dx.doi.org/10.1016/S0020-7373\(83\)80007-8](http://dx.doi.org/10.1016/S0020-7373(83)80007-8).

RIENTIES, B.; CROSS, S.; ZDRAHAL, Z. Implementing a learning analytics intervention and evaluation framework: what works? In: DANIEL, B. K.; BUTSON, R. (Ed.). **Big Data and Learning Analytics in Higher Education: Current Theory and Practice**. Heidelberg: Springer, 2016. DOI: https://doi.org/10.1007/978-3-319-06520-5_10.

RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 3. ed. [S.l.]: Pearson Education, 2013.

- SANI, S.; ARIS, T. N. Proposal for ontology based approach to fuzzy student model design. In: **Proceedings of the 5th International Conference on Intelligent Systems, Modelling and Simulation (ISMS)**. Langkawi, Malaysia: IEEE, 2014. p. 35–37. DOI: <https://doi.org/10.1109/ISMS.2014.14>.
- SEFFRIN, H. et al. Modelling students' algebraic knowledge with dynamic bayesian networks. In: **Proceedings of the 16th International Conference on Advanced Learning Technologies (ICALT)**. Kraków, Poland: IEEE, 2016. p. 44–48. DOI: <https://doi.org/10.1109/ICALT.2016.96>.
- SELF, J. Student models: what use are they. **Artificial Intelligence tools in education**, Elsevier, p. 73–86, 1988.
- SELF, J. A. Bypassing the intractable problem of student modelling. **Intelligent tutoring systems: At the crossroads of artificial intelligence and education**, v. 41, p. 1–26, 1990.
- SETTLE, A.; DETTORI, L.; DAVIDSON, M. J. Does Lecture Capture Make a Difference for Students in Traditional Classrooms. In: **Proceedings of the 16th annual joint conference on Innovation and technology in computer science education**. New York, NY: ACM, 2011. p. 78–82. DOI: <https://doi.org/10.1145/1999747.1999772>.
- SHUM, S. B.; CRICK, R. D. Learning dispositions and transferable competencies: Pedagogy, modelling and learning analytics. In: **Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK)**. New York, NY: ACM, 2012. p. 92–101. DOI: <https://doi.org/10.1145/2330601.2330629>.
- SHUTE, V. J. et al. Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. **Computers & Education**, v. 86, p. 224 – 235, 2015. ISSN 0360-1315. DOI: <https://doi.org/10.1016/j.compedu.2015.08.001>.
- SIEMENS, G.; BAKER, R. S. J. d. Learning analytics and educational data mining: Towards communication and collaboration. In: **Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK)**. New York, NY: ACM, 2012. p. 252–254. DOI: <https://doi.org/10.1145/2330601.2330661>.
- SISON, R.; SHIMURA, M. Student modeling and machine learning. **International Journal of Artificial Intelligence in Education (IJAIED)**, v. 9, p. 128–158, 1998. ISSN 1560-4306.
- SOSNOVSKY, S.; DICHEVA, D. Ontological technologies for user modelling. **International Journal of Metadata, Semantics and Ontologies**, Inderscience Publishers, v. 5, n. 1, p. 32–71, 2010. ISSN 1744-2621. DOI: <https://doi.org/10.1504/IJMSO.2010.032649>.

- STAAB, S. et al. Knowledge processes and ontologies. **IEEE Intelligent Systems**, v. 16, n. 1, p. 26–34, 2001. ISSN 1541-1672. DOI: <https://doi.org/10.1109/5254.912382>.
- SUÁREZ-FIGUEROA, M. C. et al. D5. 4.1. Neon Methodology for Building Contextualized Ontology Networks. **NeOn Project Deliverable D**, v. 5, 2008.
- TARAGHI, B. et al. Bayesian modelling of student misconceptions in the one-digit multiplication with probabilistic programming. In: **Proceedings of the 6th International Conference on Learning Analytics & Knowledge (LAK)**. New York, NY: ACM, 2016. p. 449–453. DOI: <https://doi.org/10.1145/2883851.2883895>.
- TEMPELAAR, D. et al. Student profiling in a dispositional learning analytics application using formative assessment. **Computers in Human Behavior**, v. 78, p. 408–420, 2018. ISSN 0747-5632. DOI: <https://doi.org/10.1016/j.chb.2017.08.010>.
- TING, C.-Y.; PHON-AMNUAISUK, S. Properties of bayesian student model for INQPRO. **Applied Intelligence**, Kluwer Academic Publishers, v. 36, n. 2, p. 391–406, 2012. ISSN 0924-669X. DOI: <https://doi.org/10.1007/s10489-010-0267-7>.
- TORABI, R.; MORADI, P.; KHANTAIMOORI, A. R. Predict student scores using bayesian networks. **Procedia - Social and Behavioral Sciences**, v. 46, p. 4476–4480, 2012. ISSN 1877-0428. 4th World Conference on Educational Sciences (WCES-2012). DOI: <https://doi.org/10.1016/j.sbspro.2012.06.280>.
- TOURTOGLOU, K.; VIRVOU, M. An intelligent recommender system for trainers and trainees in a collaborative learning environment for UML. **Intelligent Decision Technologies**, IOS Press, v. 6, n. 2, p. 79–95, 2012. ISSN 1875-8843. DOI: <https://doi.org/10.3233/IDT-2012-0125>.
- TRUONG, K. N.; HAYES, G. R. Ubiquitous Computing for Capture and Access. **Foundations and Trends in Human-Computer Interaction**, Now Publishers Inc., Hanover, MA, v. 2, n. 2, p. 95–171, 2009. ISSN 1551-3955. DOI: <http://dx.doi.org/10.1561/1100000014>.
- USCHOLD, M.; GRUNINGER, M. Ontologies: Principles, methods and applications. **Knowledge Engineering Review**, v. 11, p. 93–136, 1995. ISSN 1469-8005.
- VERBERT, K. et al. Context-aware recommender systems for learning: A survey and future challenges. **IEEE Transactions on Learning Technologies**, IEEE Computer Society Press, v. 5, n. 4, p. 318–335, 2012. ISSN 1939-1382. DOI: <https://doi.org/10.1109/TLT.2012.11>.
- VIER, J.; GLUZ, J.; JAQUES, P. A. Empregando redes bayesianas para modelar automaticamente o conhecimento dos aprendizes em lógica de programação. **Revista**

- Brasileira de Informática na Educação**, v. 23, n. 2, 2015. p. 45–59. ISSN 1414-5685. DOI: <https://doi.org/10.5753/rbie.2015.23.02.45>.
- WEISER, M. The computer for the 21st century. **Scientific American**, v. 265, n. 3, p. 94–104, 1991. DOI: <https://doi.org/10.1145/329124.329126>.
- WELTY, C.; MCGUINNESS, D. L.; SMITH, M. K. Owl web ontology language guide. **W3C recommendation**, W3C, 2004. DOI: https://doi.org/.1007/978-3-642-15970-1_5.
- WENGER, E. **Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1987. ISBN 0-934613-26-5. DOI: <https://doi.org/10.1016/C2013-0-07697-9>.
- YU, L. **A developer's guide to the semantic Web**. [S.l.]: Springer Science & Business Media, 2011. DOI: <https://doi.org/10.1007/978-3-662-43796-4>.
- YUDELSON, M. V.; KOEDINGER, K. R.; GORDON, G. J. Individualized bayesian knowledge tracing models. In: Springer. **Proceedings of the International Conference on Artificial Intelligence in Education (AIED)**. Memphis, Tennessee: AIED, 2013. p. 171–180. DOI: https://doi.org/10.1007/978-3-642-39112-5_18.
- ZADEH, L. A. Fuzzy sets. **Information and control**, Elsevier, v. 8, n. 3, p. 338–353, 1965. ISSN 0019-9958. DOI: [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X).

Apêndices

APÊNDICE A

Exemplos de Categorias e Regras para Inferência do Nível de Desempenho

□ Desempenho do estudante a partir das atividades de Socialização/Colaboração.

a) Desempenho do estudante que participa ativamente das atividades de colaboração através de comentários sobre os conteúdos apresentados pelos professores e pelos colegas nas disciplinas matriculadas.

```
integer[>HLow,<=HHigh] (?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
hasComent(?K, ?Com) ∧ hasContent(?Com, ?C) ∧ comentVal(?Com, high)
colabVal(?Co, high) → hasLevel(?S, A)
```

b) Desempenho do estudante que participa ativamente das atividades de colaboração através de classificação de slides (*rating*) sobre os conteúdos apresentados pelos professores e pelos colegas nas disciplinas matriculadas.

```
integer[>LLow,<=LHigh] (?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
hasRating(?K, ?Rat) ∧ hasContent(?Rat, ?C) ∧ comentVal(?Rat, high)
colabVal(?Co, high) → hasLevel(?S, B)
```

c) Desempenho do estudante que possui pouca participação em atividades de colaboração com os demais usuários do ambiente.

```
integer[>LLow,<=LHigh] (?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
colabVal(?Co, low) → hasLevel(?S, E)
```

□ Desempenho do estudante a partir das atividades de Gamificação.

a) Desempenho do estudante que se enquadra como jogador do tipo *Explorers*.

```
integer[>LLow,<=LHigh](?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
hasContent(?K, ?Ga) ∧ hasGamification(?S, ?Ga) ∧ hasContent(?Ga, ?C) ∧
colabVal(?Co, low) ∧ gamificationType(?Ga, explorer) -> hasLevel(?S, B)
```

b) Desempenho do estudante que se enquadra como jogador do tipo *Achievers*.

```
integer[>HLow,<=HHigh](?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
hasContent(?K, ?Ga) ∧ hasGamification(?S, ?Ga) ∧ hasContent(?Ga, ?C) ∧
colabVal(?Co, high) ∧ gamificationType(?Ga, achiever) -> hasLevel(?S, A)
```

c) Desempenho do estudante que se enquadra como jogador do tipo *Socializers*.

```
integer[>HLow,<=HHigh](?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
hasContent(?K, ?Ga) ∧ hasGamification(?S, ?Ga) ∧ hasContent(?Ga, ?C) ∧
colabVal(?Co, high) ∧ gamificationType(?Ga, socializer) -> hasLevel(?S, A)
```

d) Desempenho do estudante que se enquadra como jogador do tipo *Killers*.

```
integer[>MLow,<=MHigh](?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
hasContent(?K, ?Ga) ∧ hasGamification(?S, ?Ga) ∧ hasContent(?Ga, ?C) ∧
colabVal(?Co, low) ∧ gamificationType(?Ga, killers) -> hasLevel(?S, C)
```

e) Desempenho dos estudantes que utilizam mecanismos para burlar as atividades de gamificação (*Gaming the System*)

```
integer[>HLow,<=HHigh](?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) ∧ hasColaboration(?S, ?Co) ∧ hasContent(?Co, ?C) ∧
hasContent(?K, ?Ga) ∧ hasGamification(?S, ?Ga) ∧ hasContent(?Ga, ?C) ∧
colabVal(?Co, high) ∧ gamificationType(?Ga, cheat) -> hasLevel(?S, E)
```

□ Desempenho do estudante a partir de instrumentos avaliativos.

a) Desempenho do estudante que possui um alto nível de acerto a partir dos instrumentos avaliativos disponíveis no ambiente.

```
integer[>HLow,<=HHigh](?gra) ∧ grade(?K, ?gra) ∧ hasKnowledge(?S, ?K) ∧
hasContent(?K, ?C) -> hasLevel(?S, A)
```

- b) Desempenho do estudante que possui um baixo nível de acerto a partir dos instrumentos avaliativos disponíveis no ambiente.

$$\text{integer}[\text{>LLow}, \text{<=LHigh}](\text{?gra}) \wedge \text{grade}(\text{?K}, \text{?gra}) \wedge \text{hasKnowledge}(\text{?S}, \text{?K}) \wedge \text{hasContent}(\text{?K}, \text{?C}) \rightarrow \text{hasLevel}(\text{?S}, \text{D})$$

- c) Desempenho do estudante que possui um alto nível de resposta e um baixo nível de acerto, caracterizando assim, um comportamento de “chute”.

$$\text{integer}[\text{>HLow}, \text{<=HHigh}](\text{?gra}) \wedge \text{grade}(\text{?K}, \text{?gra}) \wedge \text{hasKnowledge}(\text{?S}, \text{?K}) \wedge \text{hasContent}(\text{?K}, \text{?C}) \wedge \text{hasBehaviour}(\text{?S}, \text{?B}) \wedge \text{BehaviourType}(\text{?B}, \text{lie}) \rightarrow \text{hasLevel}(\text{?S}, \text{E})$$

- d) Desempenho do estudante que costuma não responder os instrumentos avaliativos disponíveis no ambiente.

$$\text{integer}[\text{>LLow}, \text{<=LHigh}](\text{?gra}) \wedge \text{grade}(\text{?K}, \text{?gra}) \wedge \text{hasKnowledge}(\text{?S}, \text{?K}) \wedge \text{hasContent}(\text{?K}, \text{?C}) \wedge \text{hasBehaviour}(\text{?S}, \text{?B}) \wedge \text{BehaviourType}(\text{?B}, \text{absent}) \rightarrow \text{hasLevel}(\text{?S}, \text{D})$$

□ Desempenho do estudante a partir das atividades desempenhada durante a utilização do ambiente “rastro”.

- a) Desempenho do estudante que acessa várias vezes ou permanece conectado (logado) durante um longo período de tempo desempenhando muitas atividades.

$$\text{integer}[\text{>HLow}, \text{<=HHigh}](\text{?gra}) \wedge \text{grade}(\text{?K}, \text{?gra}) \wedge \text{hasKnowledge}(\text{?S}, \text{?K}) \wedge \text{hasContent}(\text{?K}, \text{?C}) \wedge \text{hasBehaviour}(\text{?S}, \text{?B}) \wedge \text{BehaviourType}(\text{?B}, \text{high}) \wedge \text{BehaviourActivities}(\text{?B}, \text{high}) \rightarrow \text{hasLevel}(\text{?S}, \text{A})$$

- b) Desempenho do estudante que acessa poucas vezes ou permanece conectado (logado) durante um curto período de tempo desempenhando muitas atividades.

$$\text{integer}[\text{>HLow}, \text{<=HHigh}](\text{?gra}) \wedge \text{grade}(\text{?K}, \text{?gra}) \wedge \text{hasKnowledge}(\text{?S}, \text{?K}) \wedge \text{hasContent}(\text{?K}, \text{?C}) \wedge \text{hasBehaviour}(\text{?S}, \text{?B}) \wedge \text{BehaviourType}(\text{?B}, \text{low}) \wedge \text{BehaviourActivities}(\text{?B}, \text{high}) \rightarrow \text{hasLevel}(\text{?S}, \text{A})$$

- c) Desempenho do estudante que acessa várias vezes ou permanece conectado (logado) durante um longo período de tempo desempenhando poucas atividades.

$$\text{integer}[\text{>HLow}, \text{<=HHigh}](\text{?gra}) \wedge \text{grade}(\text{?K}, \text{?gra}) \wedge \text{hasKnowledge}(\text{?S}, \text{?K}) \wedge \text{hasContent}(\text{?K}, \text{?C}) \wedge \text{hasBehaviour}(\text{?S}, \text{?B}) \wedge \text{BehaviourType}(\text{?B}, \text{low}) \wedge \text{BehaviourActivities}(\text{?B}, \text{low}) \rightarrow \text{hasLevel}(\text{?S}, \text{C})$$

APÊNDICE B

Redes Bayesianas Utilizadas no Experimento 1

A seguir serão apresentadas cinco Rede Bayesiana (RB)s utilizadas no Experimento 1 (5.2). Em *verde* são apresentados os tópicos utilizados como entrada de informação e *vermelho* os tópicos utilizados como saída, aqueles que comparados com as notas reais.

- ❑Figura 35: Esta RB representa o domínio de conhecimento da disciplina Programação Procedimental, no curso Ciência da Computação, no segundo semestre de 2015.
- ❑Figura 36: Esta RB representa o domínio de conhecimento da disciplina Interação Humano-Computador, no curso Sistemas de Informação, no primeiro semestre de 2016.
- ❑Figura 37: Esta RB representa o domínio de conhecimento da disciplina Interação Humano-Computador, no curso de Pós-graduação em Ciência da Computação, no primeiro semestre de 2016.
- ❑Figura 38: Esta RB representa o domínio de conhecimento da disciplina Interação Humano-Computador, no curso Sistemas de Informação, no segundo semestre de 2016.
- ❑Figura 39: Esta RB representa o domínio de conhecimento da disciplina Arquitetura e Organização de Computadores, no curso Sistemas de Informação, no segundo semestre de 2016.

Figura 35 – RB utilizada no Experimento 1 (T1)

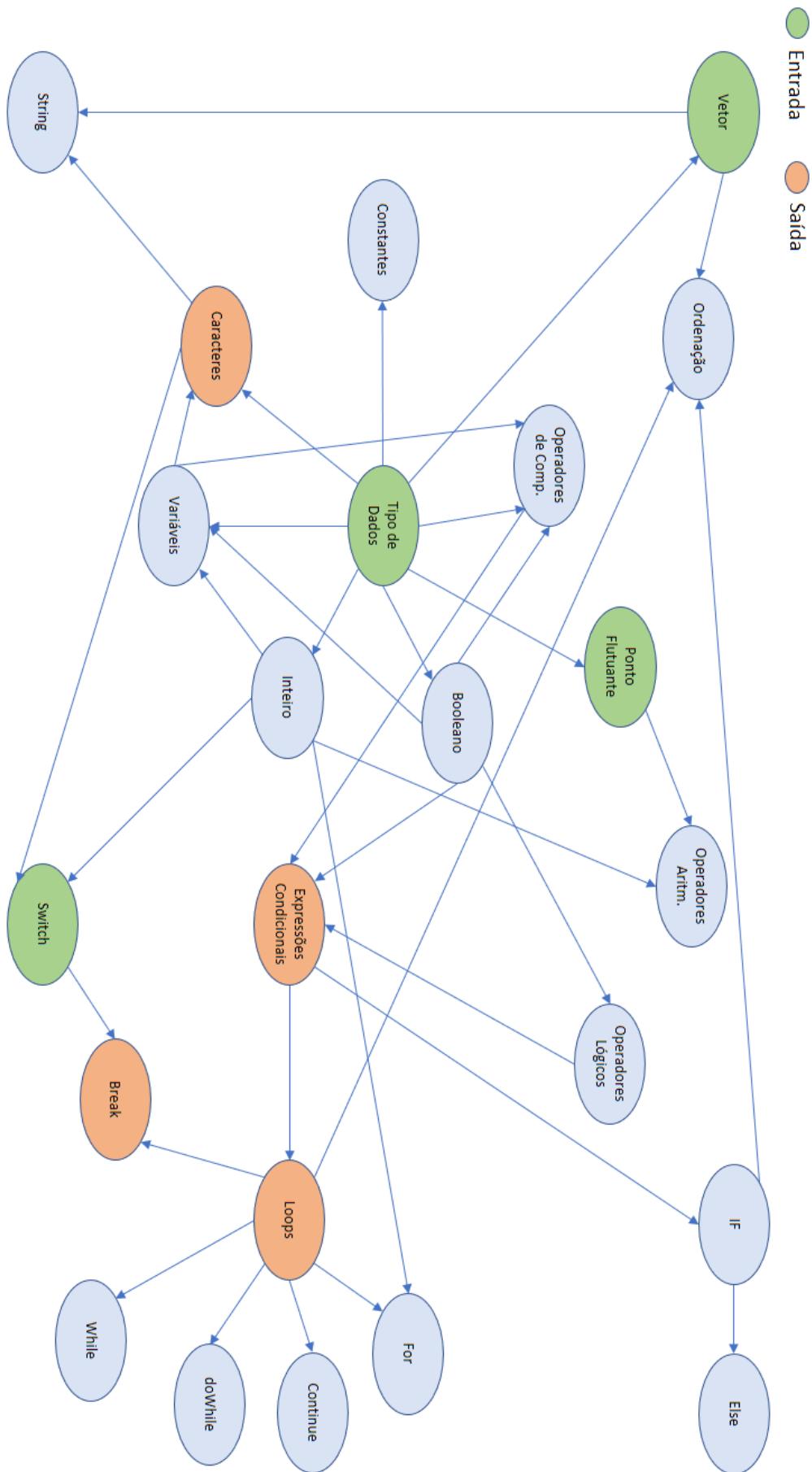


Figura 36 – RB utilizada no Experimento 1 (T2)

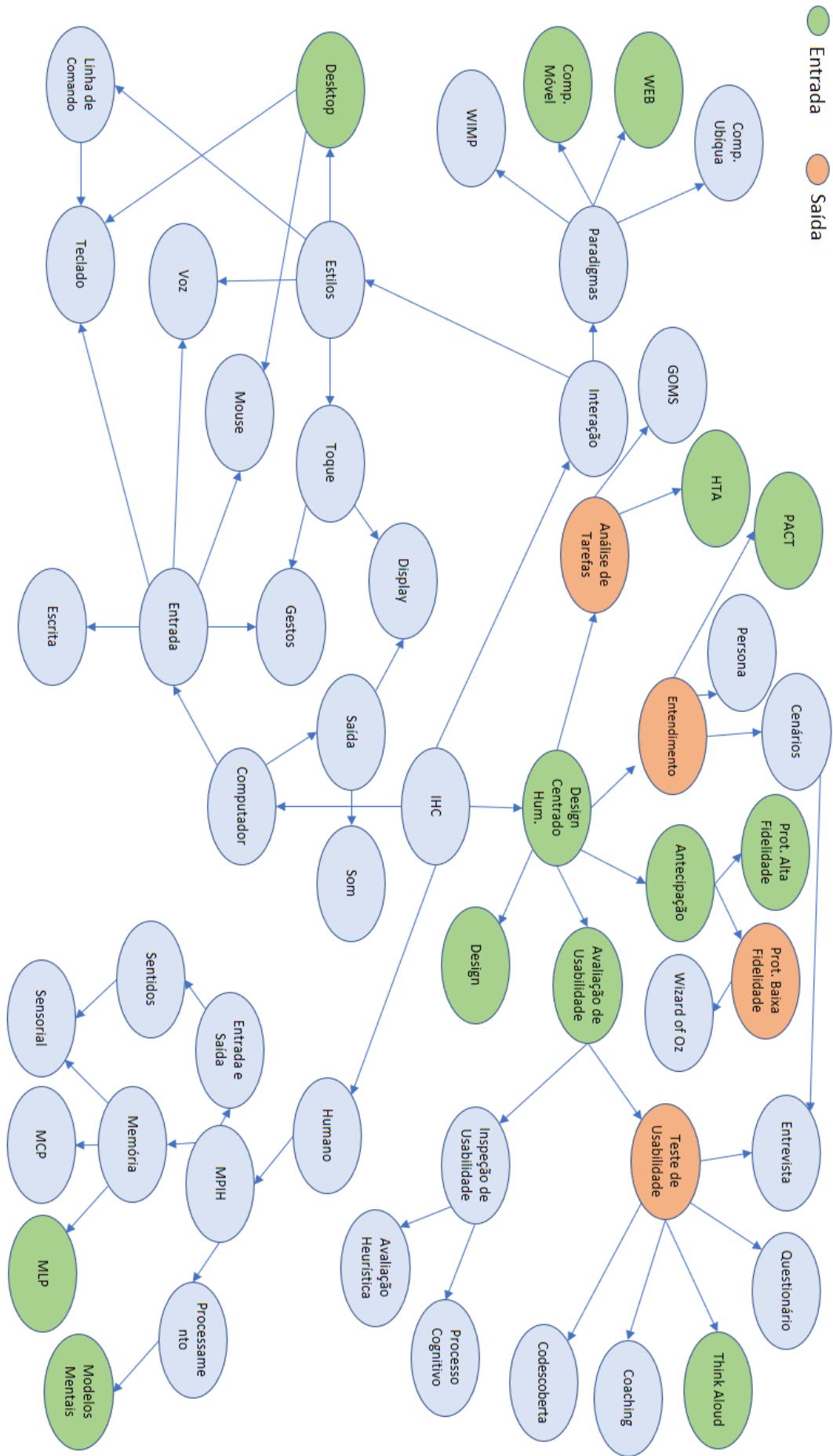


Figura 37 – RB utilizada no Experimento 1 (T3)

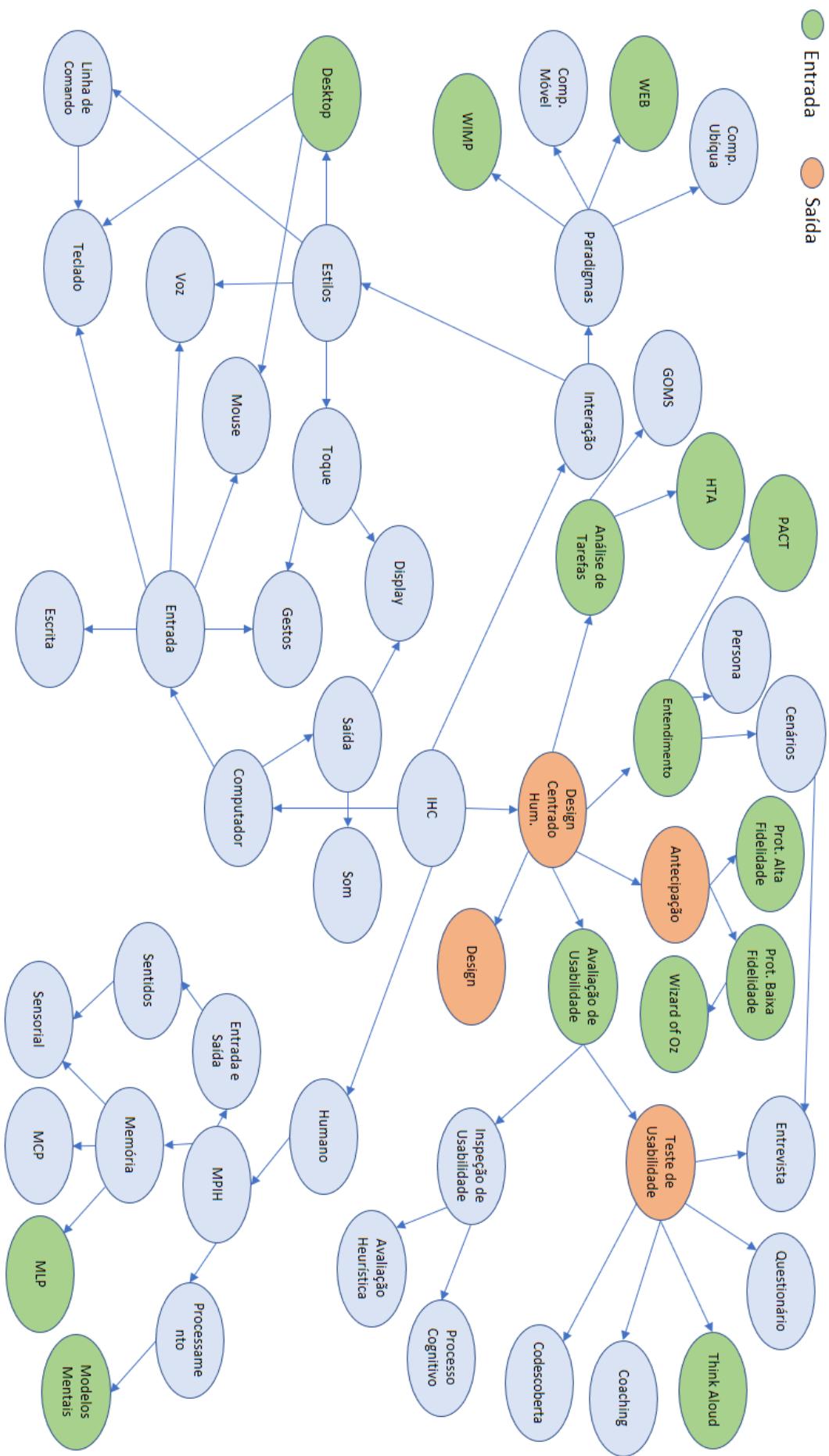


Figura 38 – RB utilizada no Experimento 1 (T4)

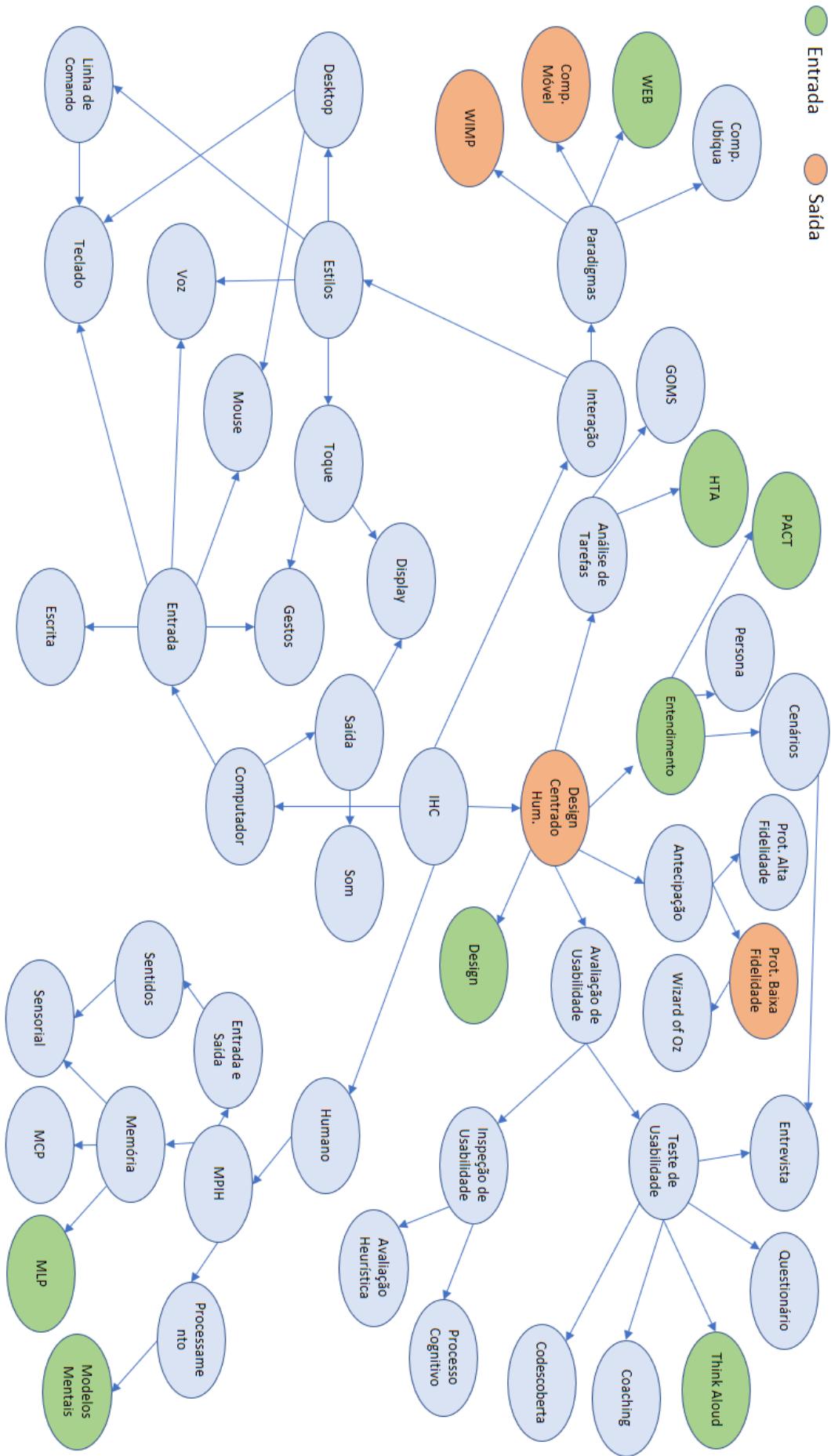
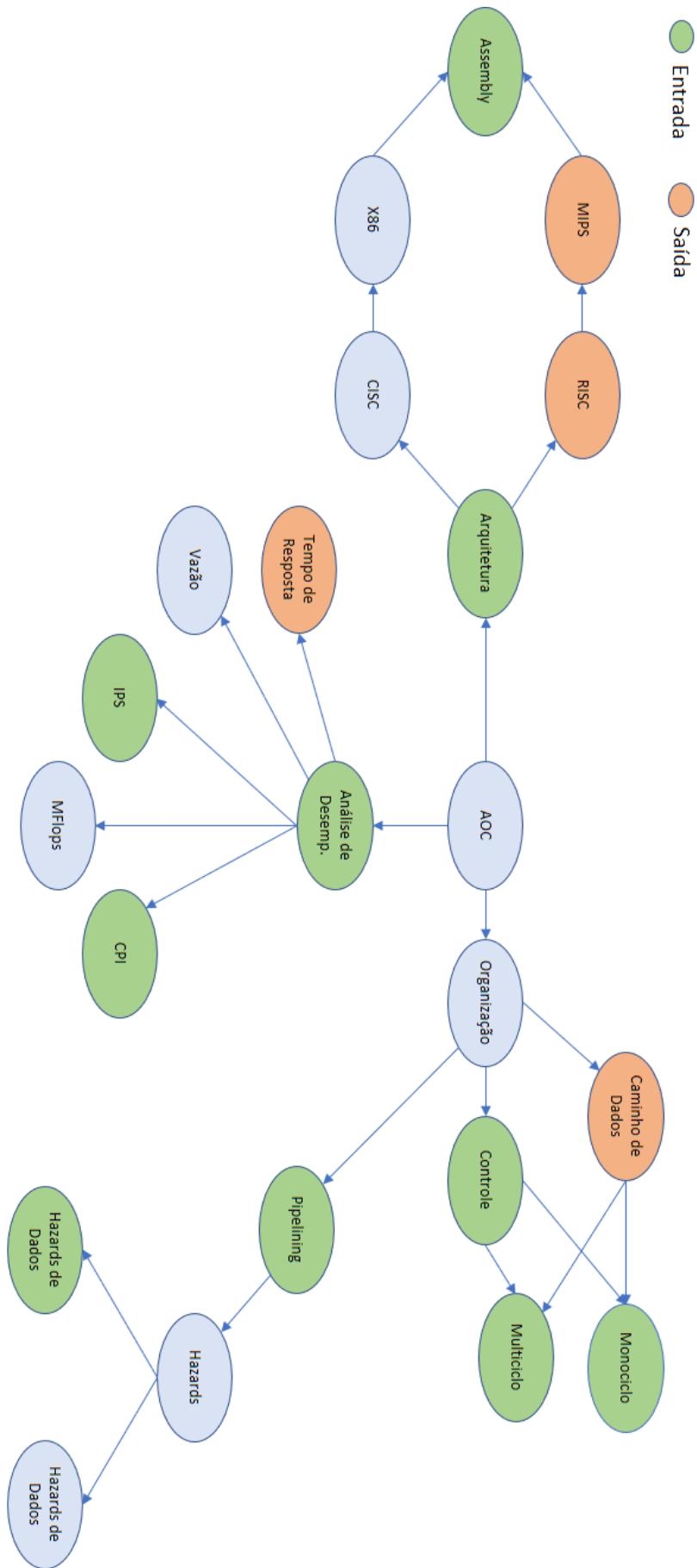


Figura 39 – RB utilizada no Experimento 1 (T5)



APÊNDICE C

Redes Bayesianas Utilizadas no Experimento 2

A seguir serão apresentadas as tabelas com os resultados da análise de correlação entre os valores reais e os valores inferidos sobre o nível de conhecimento dos estudantes. Também são apresentadas figuras que apresentam as Redes Bayesianas (RBs) utilizadas no Experimento 2.

- ❑ Disciplina IHC-2016-1: RBs utilizadas no experimento. Figuras 40 e Figura 41.
- ❑ Disciplina AOC-2016-2: RBs utilizadas no experimento. Figuras 42 e Figura 43.
- ❑ Disciplina IHC-2016-2. RBs utilizadas no experimento. Figuras 44 e Figura 45.
- ❑ Disciplina AOC-2017-1. RBs utilizadas no experimento. Figuras 46 e Figura 47.
- ❑ Disciplina IHC-2017-1. RBs utilizadas no experimento. Figuras 48 e Figura 49.

Tabela 19 – Resultado da análise de correlação entre os valores reais e os valores inferidos sobre o nível de conhecimento dos estudantes (Parte 1).

Disciplina	Tópico	Teste de Spearman ¹			
		Exec1	Exec2	Exec3	Exec4
IHC-2016-1	Antecipação	0.852	0.836	0.817	-
IHC-2016-1	Análise de Tarefas	0.977	-	-	-
IHC-2016-1	Avaliação de usabilidade	0.151	0.158	0.187	0.406
IHC-2016-1	Computação Móvel	0.742	0.715	0.550	0.653
IHC-2016-1	Design	0.54	0.545	0.512	0.506
IHC-2016-1	Design Centrado no Humano	0.703	0.701	-	0.587
IHC-2016-1	Desktop	0.38	0.489	-	0.512
IHC-2016-1	Entedimento	-	-	0.807	0.587
IHC-2016-1	HTA	-	0.987	0.987	0.987
IHC-2016-1	MLP	0.38	-	0.431	-
IHC-2016-1	Modelos Mentais	0.42	0.364	0.354	-
IHC-2016-1	PACT	0.478	0.498	0.516	0.504
IHC-2016-1	Prototipo Alta Fidelidade	0.842	-	-	0.666
IHC-2016-1	Prototipo Baixa Fidelidade	-	0.841	0.808	0.666
IHC-2016-1	Teste de Usabilidade	0.969	0.968	0.948	0.65
IHC-2016-1	Think Aloud	-	-	-	0.65
IHC-2016-1	Web	0.184	0.17	0.142	0.108
AOC-2016-2	Análise de Desempenho	-	-	0.747	0.654
AOC-2016-2	Arquitetura	-	-	0.807	0.579
AOC-2016-2	Assembly	0.936	0.937	-	0.608
AOC-2016-2	CPI	-	0.512	-	-
AOC-2016-2	Caminho de Dados	0.706	-	0.704	-
AOC-2016-2	Controle	0.711	0.939	0.714	-
AOC-2016-2	Hazards de Dados	0.578	0.719	0.426	0.504
AOC-2016-2	IPS	-	-	0.669	0.58
AOC-2016-2	MIPS	0.795	0.802	0.9	0.675
AOC-2016-2	Monociclo	-	0.72	-	0.758
AOC-2016-2	Multiciclo	0.761	0.624	0.768	-
AOC-2016-2	Pipelining	0.696	-	0.696	0.696
AOC-2016-2	RISC	0.715	0.722	0.878	0.603
AOC-2016-2	Tempo de Resposta	0.777	0.808	-	0.617
IHC-2016-2	Computação Móvel	0.278	-	-	-
IHC-2016-2	Design	0.749	0.737	0.705	0.737
IHC-2016-2	Design Centrado no Humano	-	-	0.575	-
IHC-2016-2	Entedimento	-	-	0.541	-
IHC-2016-2	HTA	-	-	0.436	0.318
IHC-2016-2	MLP	-	0.774	0.618	0.774
IHC-2016-2	Modelos Mentais	-	0.428	-	0.428
IHC-2016-2	PACT	0.575	0.59	0.614	0.589
IHC-2016-2	Prototipo Baixa Fidelidade	0.732	0.734	0.667	0.734
IHC-2016-2	Think Aloud	0.293	-	0.533	-
IHC-2016-2	WIMP	0.9	-	-	-
IHC-2016-2	Web	0.243	-	-	-

¹Valores obtidos para o coeficiente de correlação de Spearman em cada uma das execuções.

Tabela 20 – Resultado da análise de correlação entre os valores reais e os valores inferidos sobre o nível de conhecimento dos estudantes (Parte 2).

Disciplina	Tópico	Teste de Spearman ¹			
		Exec1	Exec2	Exec3	Exec4
AOC-2017-1	Análise de Desempenho	-	-	-	0.913
AOC-2017-1	Assembly	-	0.794	0.793	0.783
AOC-2017-1	Bolhas	-	0.475	-	-
AOC-2017-1	CPI	0.711	0.731	0.717	-
AOC-2017-1	Controle	0.579	0.582	0.567	0.578
AOC-2017-1	Hazards de Dados	0.574	-	0.508	-
AOC-2017-1	IPS	0.711	0.731	0.717	0.796
AOC-2017-1	MIPS	0.855	0.959	0.75	-
AOC-2017-1	Multiciclo	-	-	-	-
AOC-2017-1	Organização	0.728	0.737	0.711	0.731
AOC-2017-1	Pipelining	0.539	0.622	0.54	0.515
AOC-2017-1	Ponto Flutuante	0.763	0.713	-	0.661
AOC-2017-1	Procedimentos	-	0.013	-0.025	-
AOC-2017-1	RISC	0.843	-	0.701	0.938
AOC-2017-1	Tempo de Resposta	0.751	0.738	0.748	-
IHC-2017-1	Antecipação	0.335	0.366	-	-
IHC-2017-1	Análise de Tarefas	-	-	-	0.359
IHC-2017-1	Avaliação de Usabilidade	0.501	0.516	0.59	0.615
IHC-2017-1	Design	0.762	0.718	0.435	-
IHC-2017-1	Design Centrado no Humano	0.072	0.161	0.209	0.121
IHC-2017-1	Estilos	0.957	0.957	-	0.571
IHC-2017-1	IHC	0.369	0.485	-	0.593
IHC-2017-1	Interação	-	-	0.781	0.571
IHC-2017-1	MLP	0.984	0.981	-	0.977
IHC-2017-1	MPIH	0.326	-	0.347	-
IHC-2017-1	Memória	-	-	0.957	-
IHC-2017-1	Modelos Mentais	-	-	-	-
IHC-2017-1	PACT	0.485	-	-	0.515

¹Valores obtidos para o coeficiente de correlação de Spearman em cada uma das execuções.

Figura 40 – RB utilizada no Experimento 2 - IHC 2016-1 (Execuções 1 e 2)

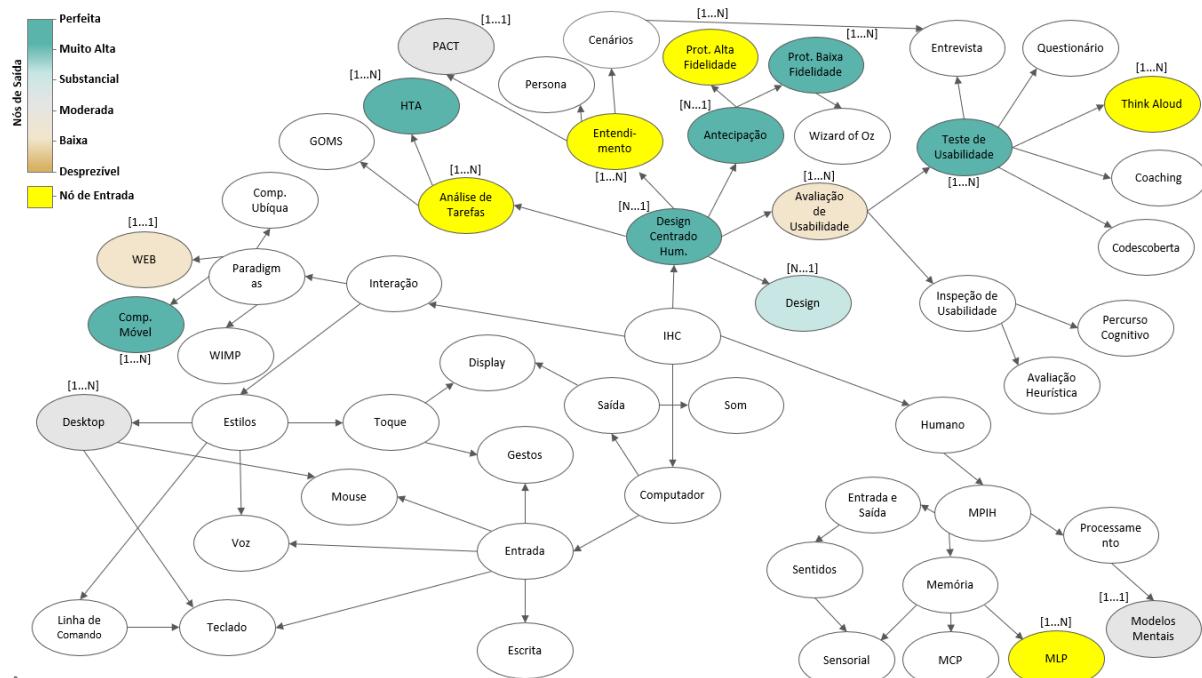
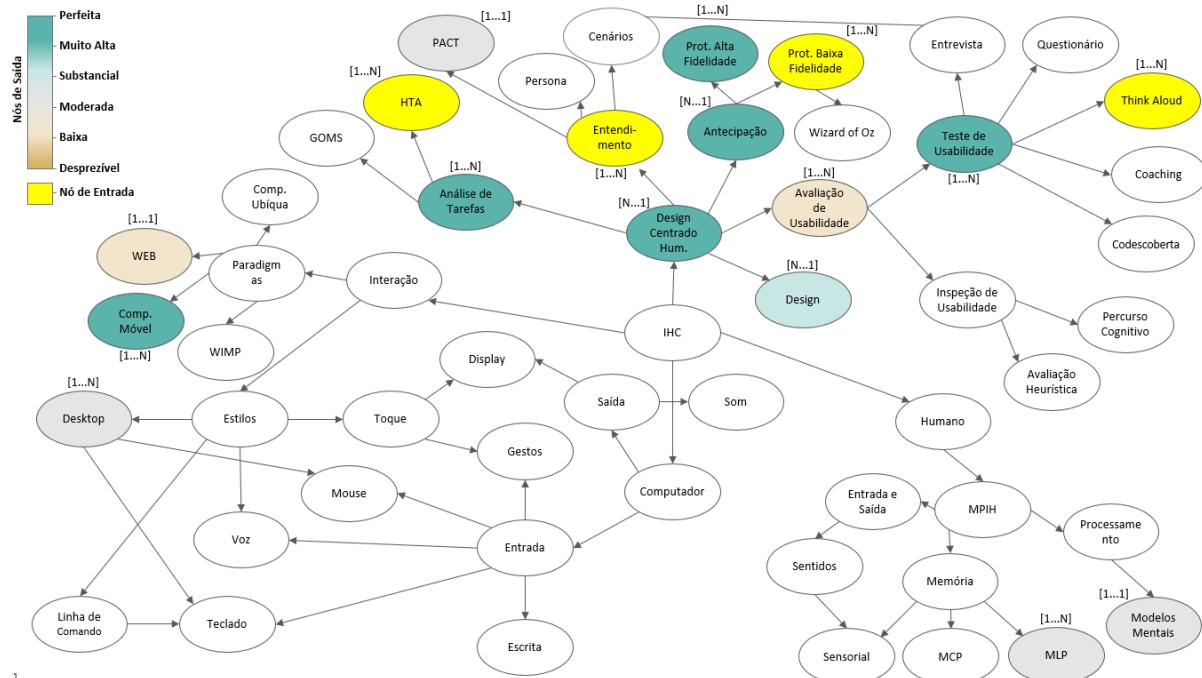


Figura 41 – RB utilizada no Experimento 2 - IHC 2016-1 (Execuções 3 e 4)

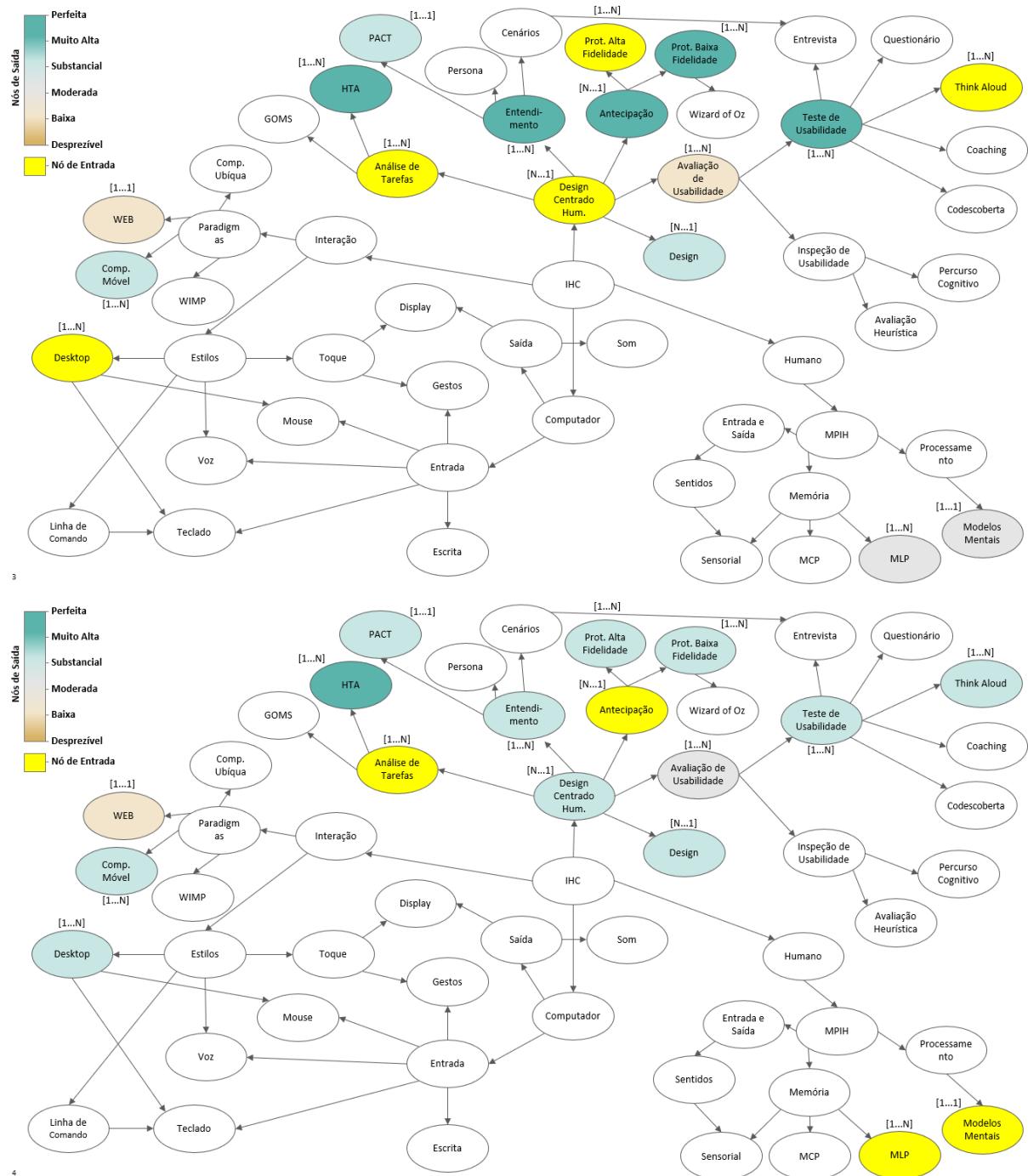


Figura 42 – RB utilizada no Experimento 2 - AOC 2016-2 (Execuções 1 e 2)

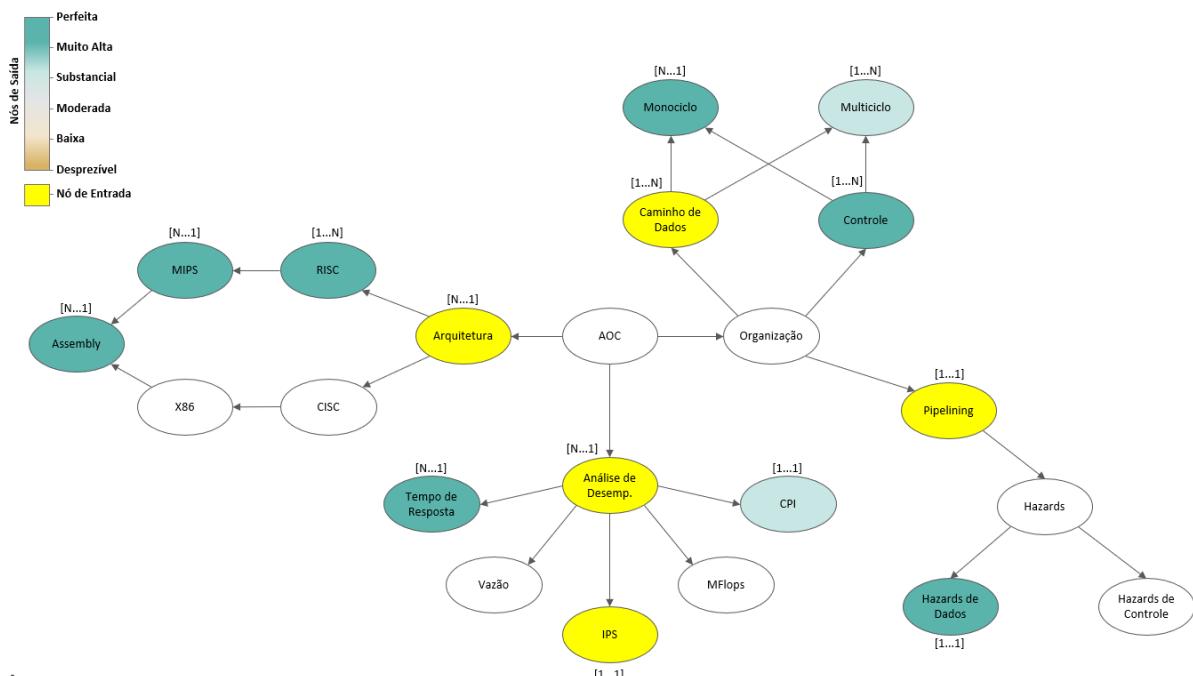
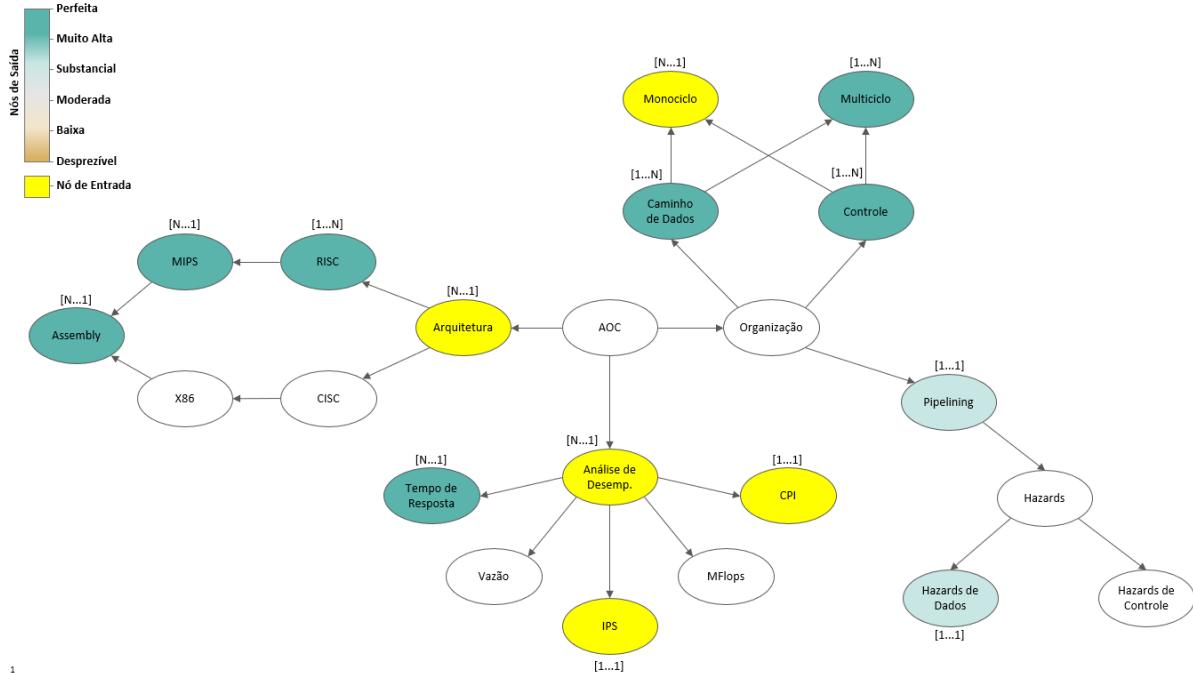


Figura 43 – RB utilizada no Experimento 2 - AOC 2016-2 (Execuções 3 e 4)

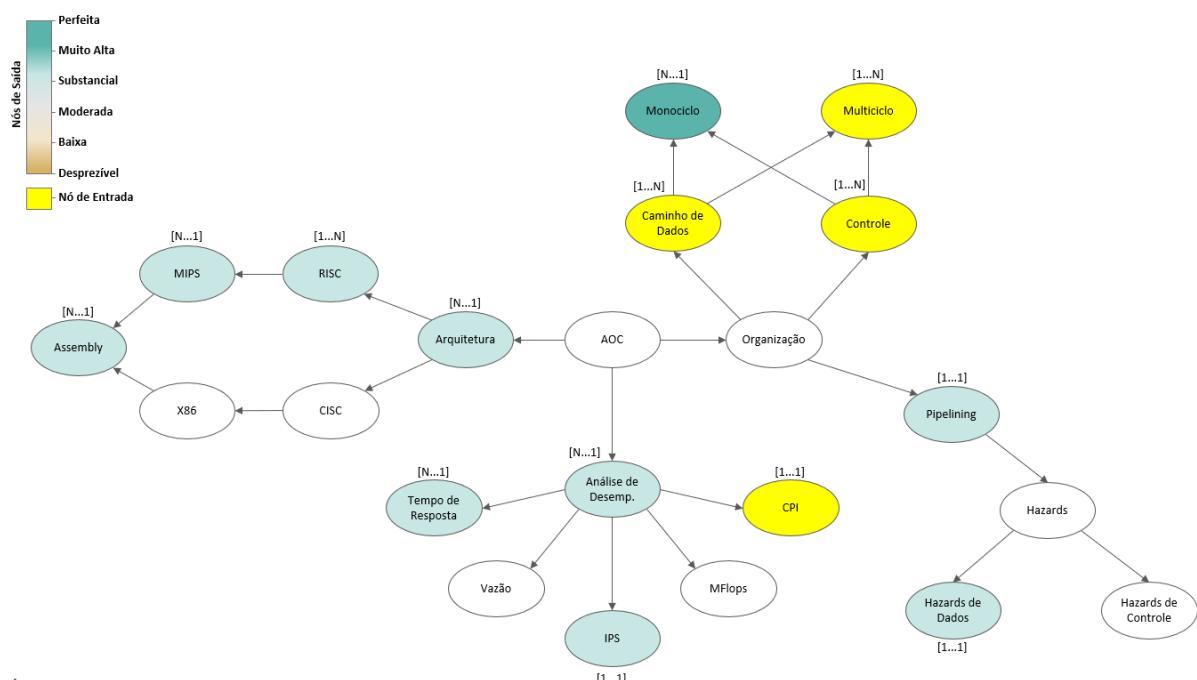
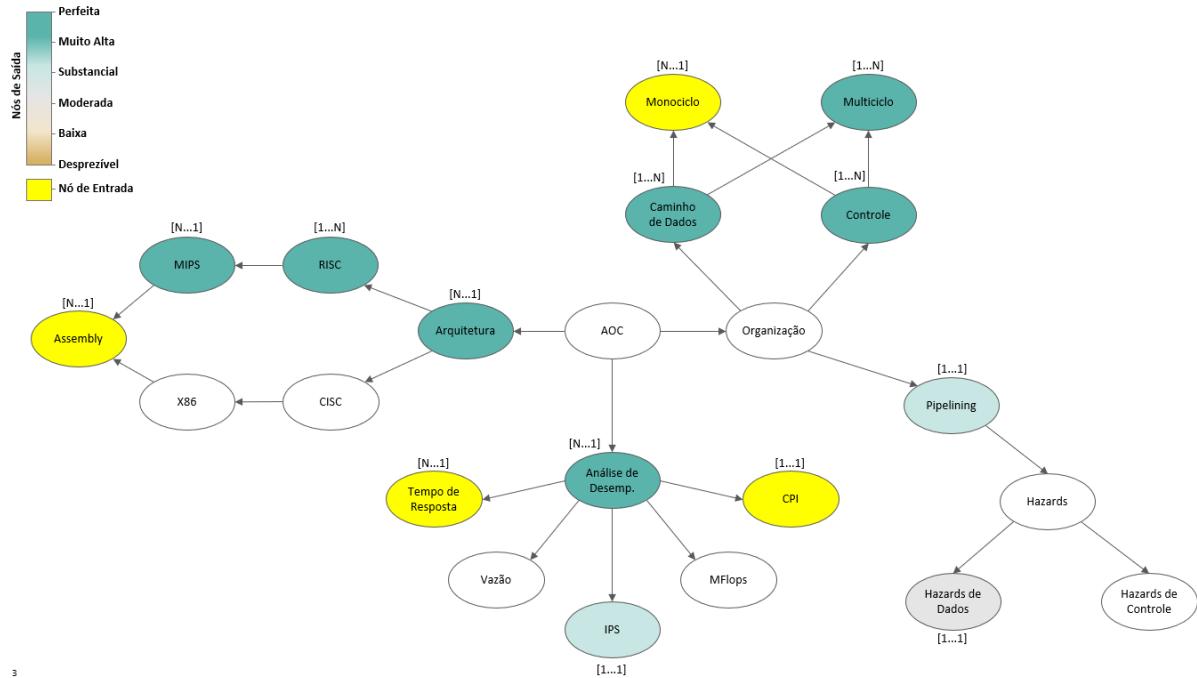


Figura 44 – RB utilizada no Experimento 2 - IHC 2016-2 (Execuções 1 e 2)

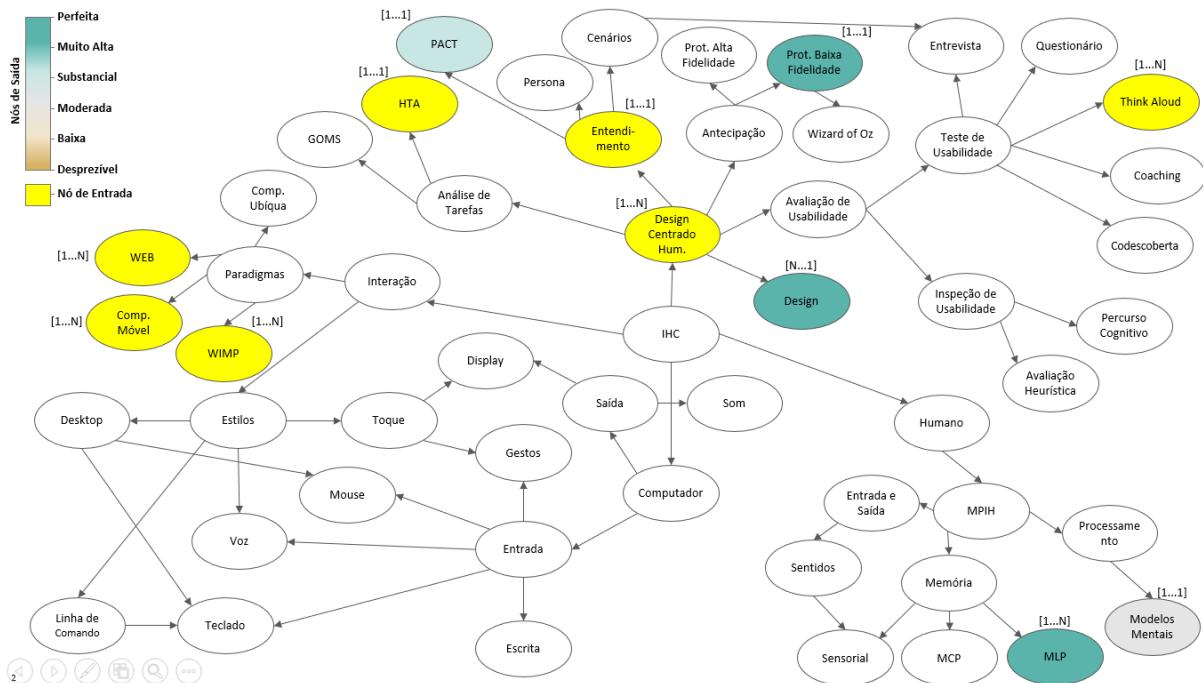
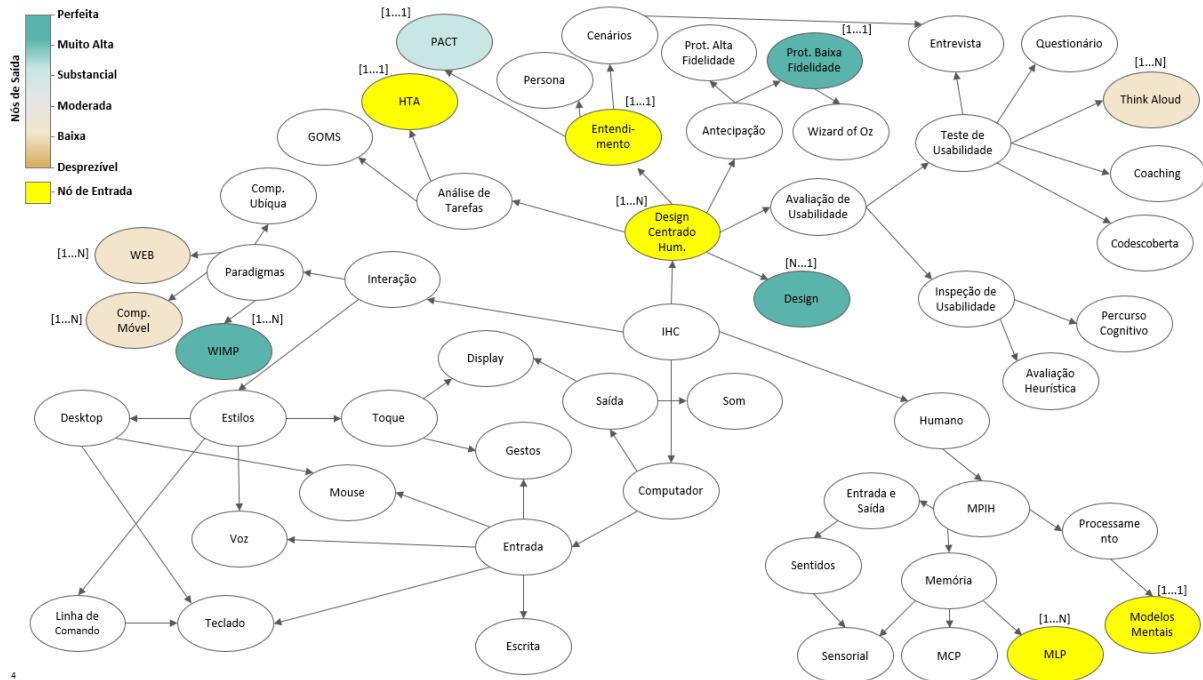


Figura 45 – RB utilizada no Experimento 2 - IHC 2016-2 (Execuções 3 e 4)

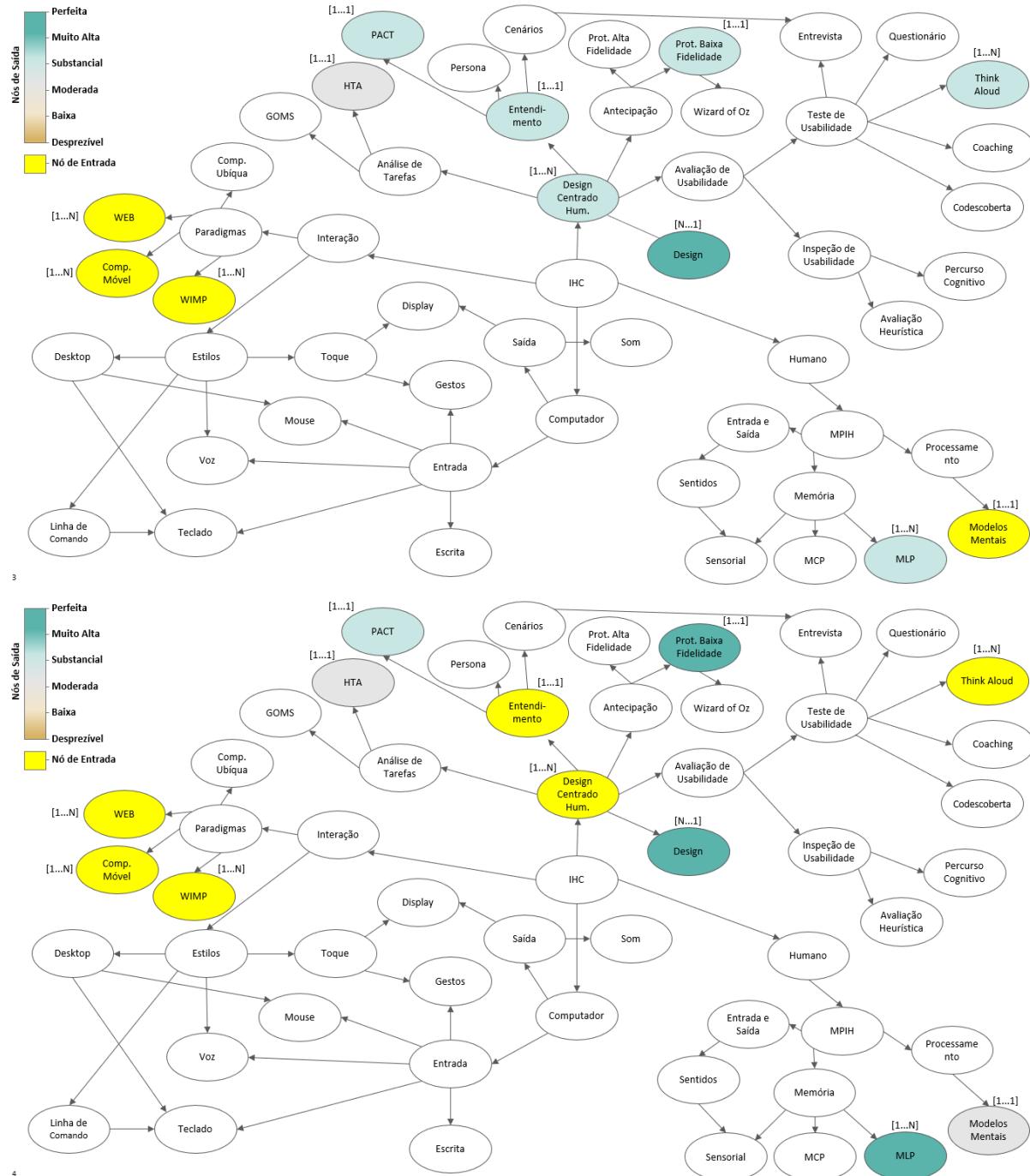
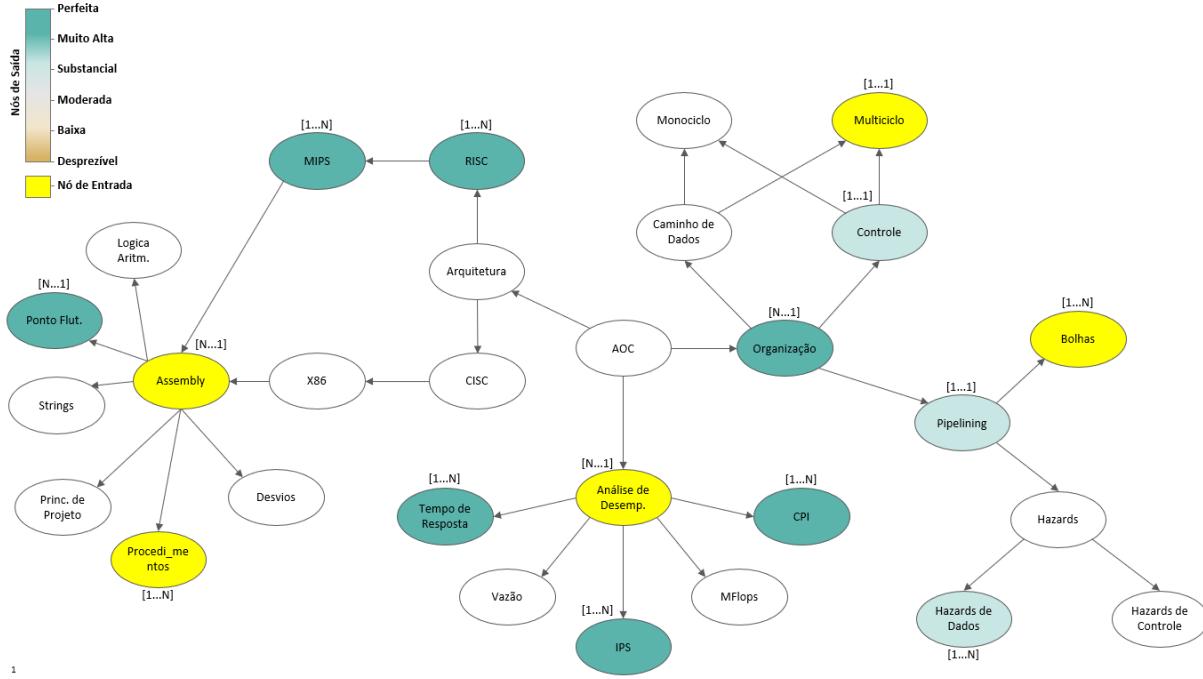
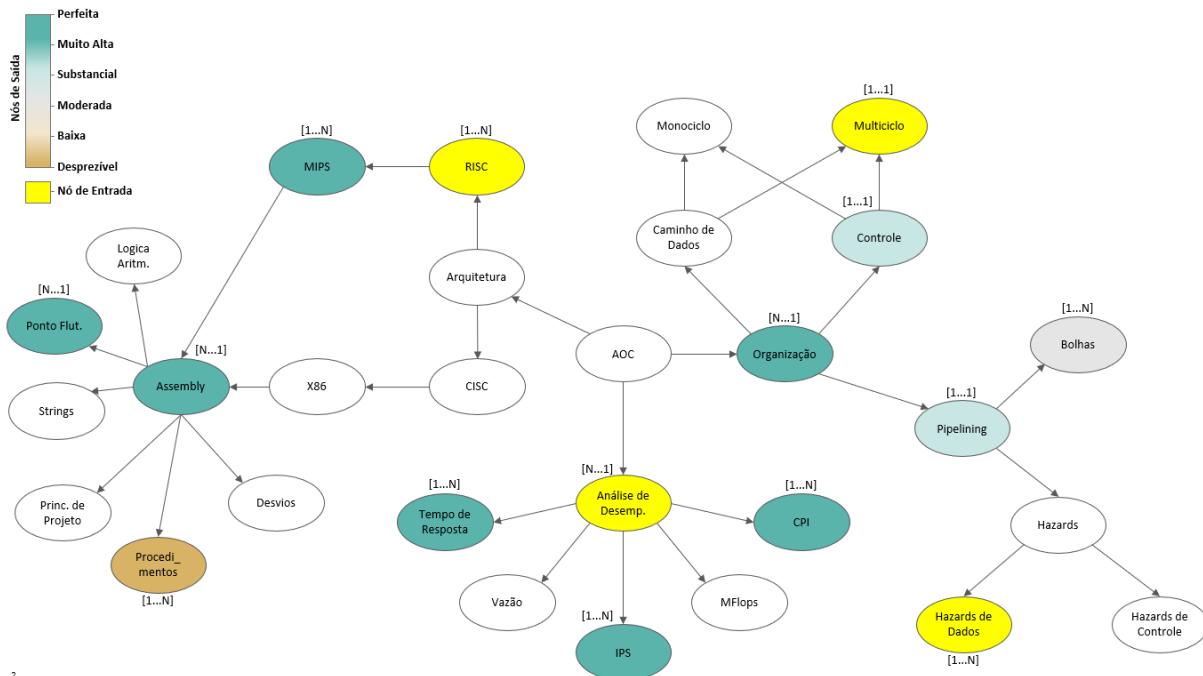


Figura 46 – RB utilizada no Experimento 2 - AOC 2017-1 (Execuções 1 e 2)



1



2

Figura 47 – RB utilizada no Experimento 2 - AOC 2017-1 (Execuções 3 e 4)

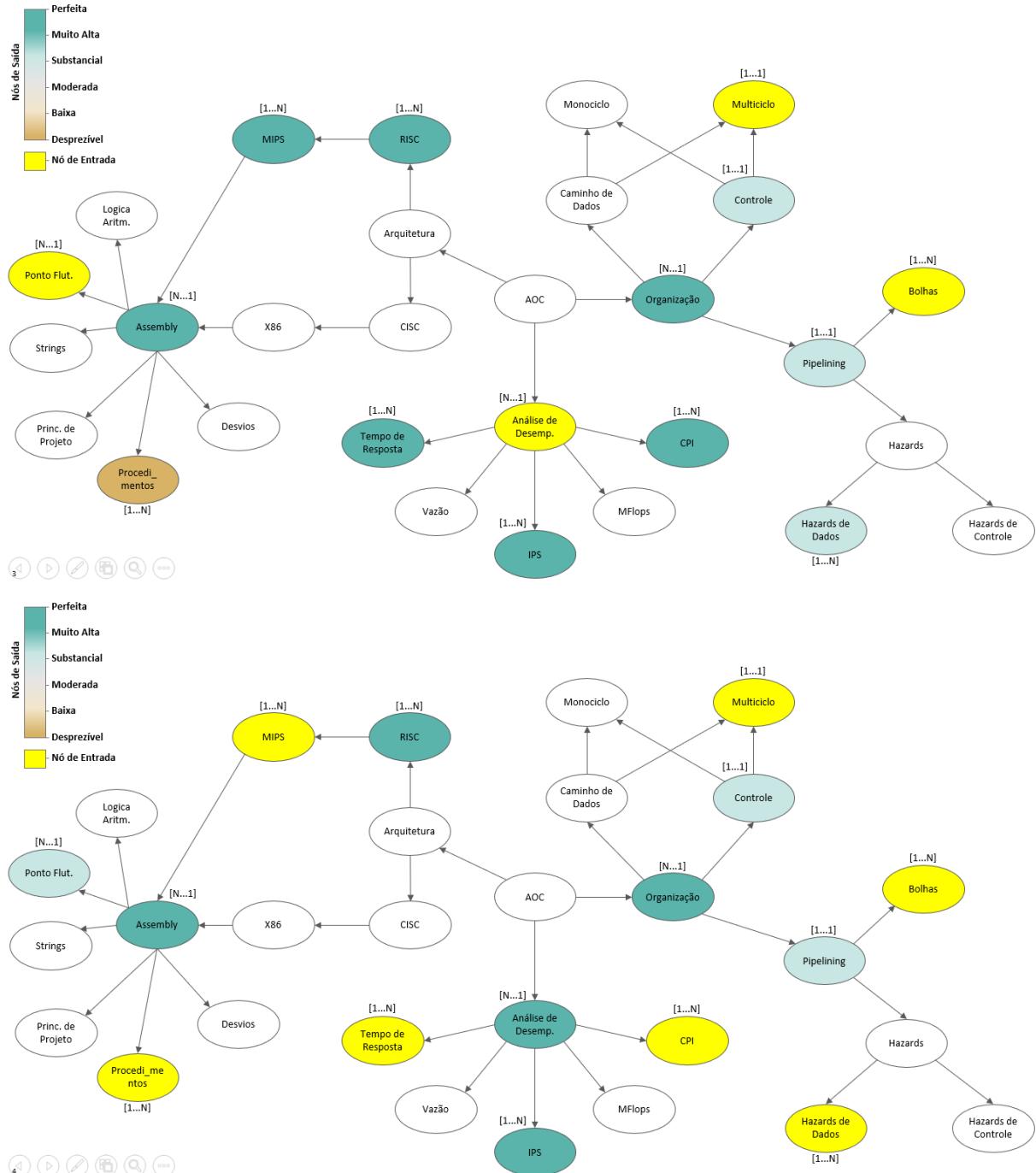


Figura 48 – RB utilizada no Experimento 2 - IHC 2017-1 (Execuções 1 e 2)

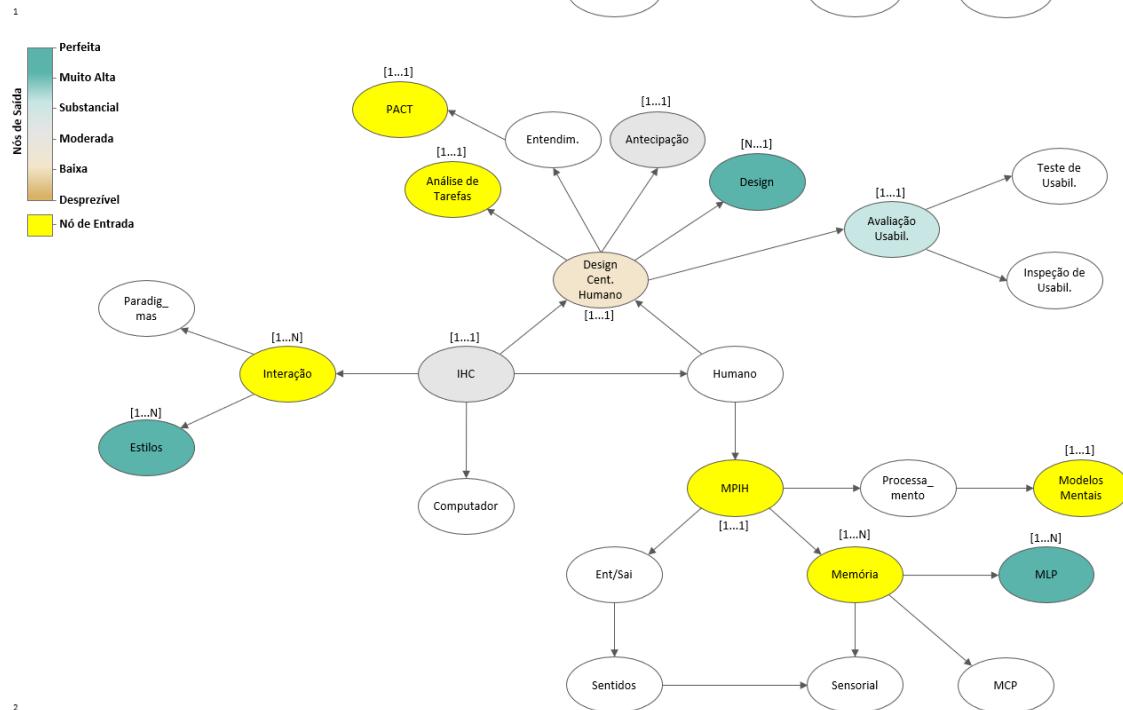
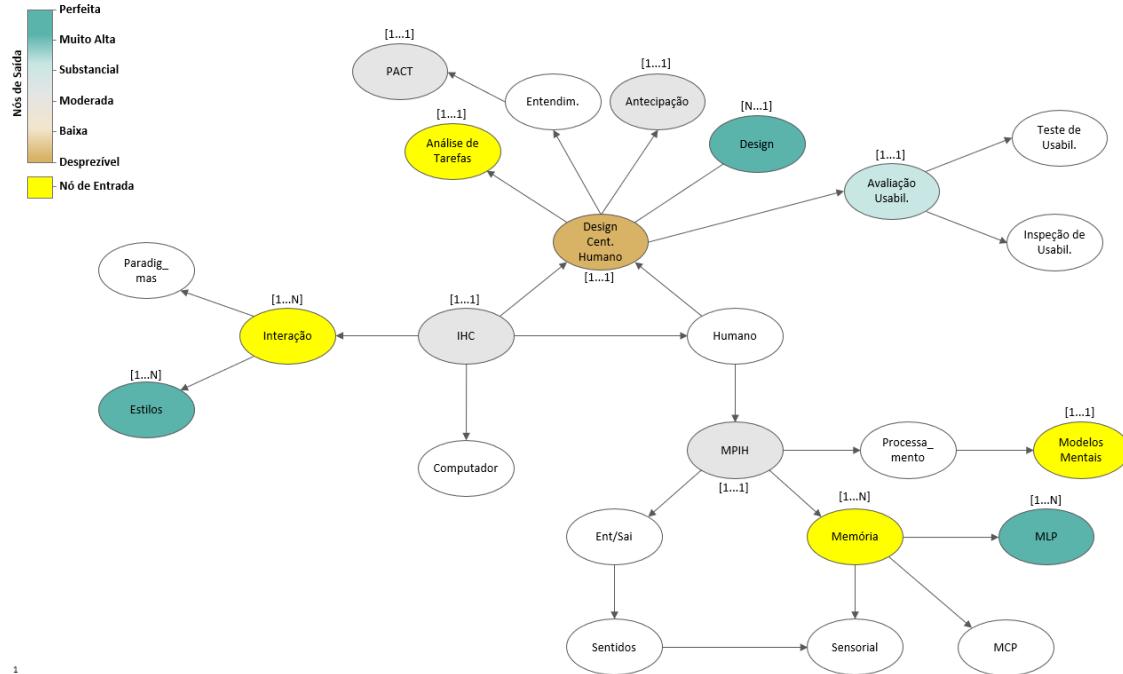
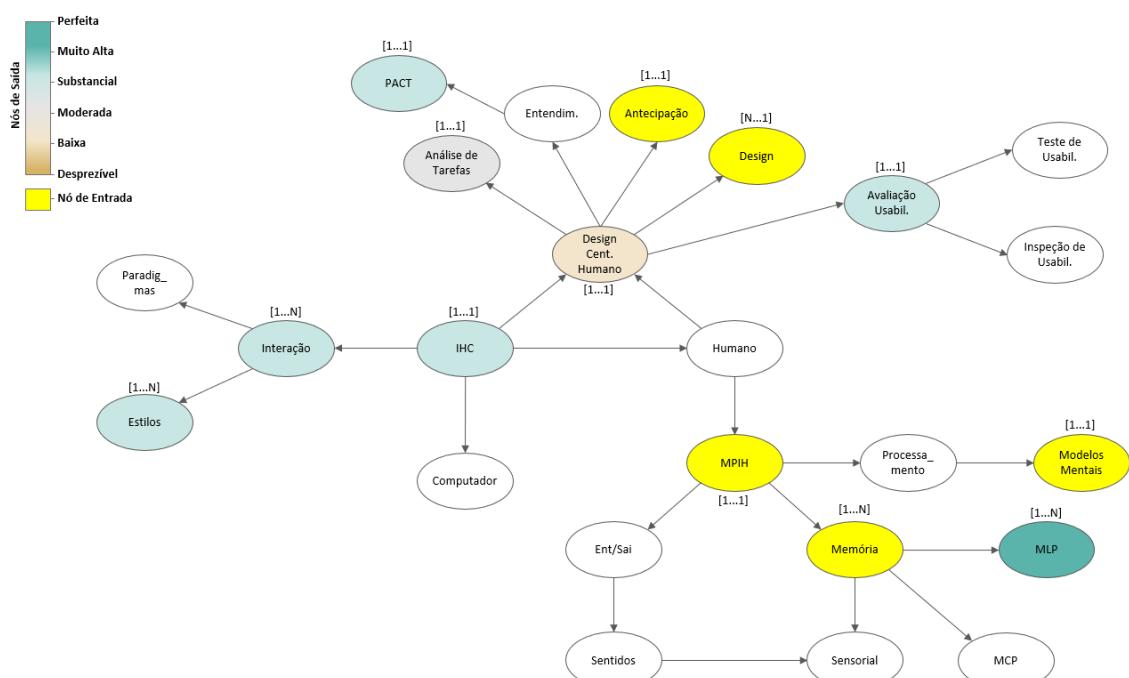
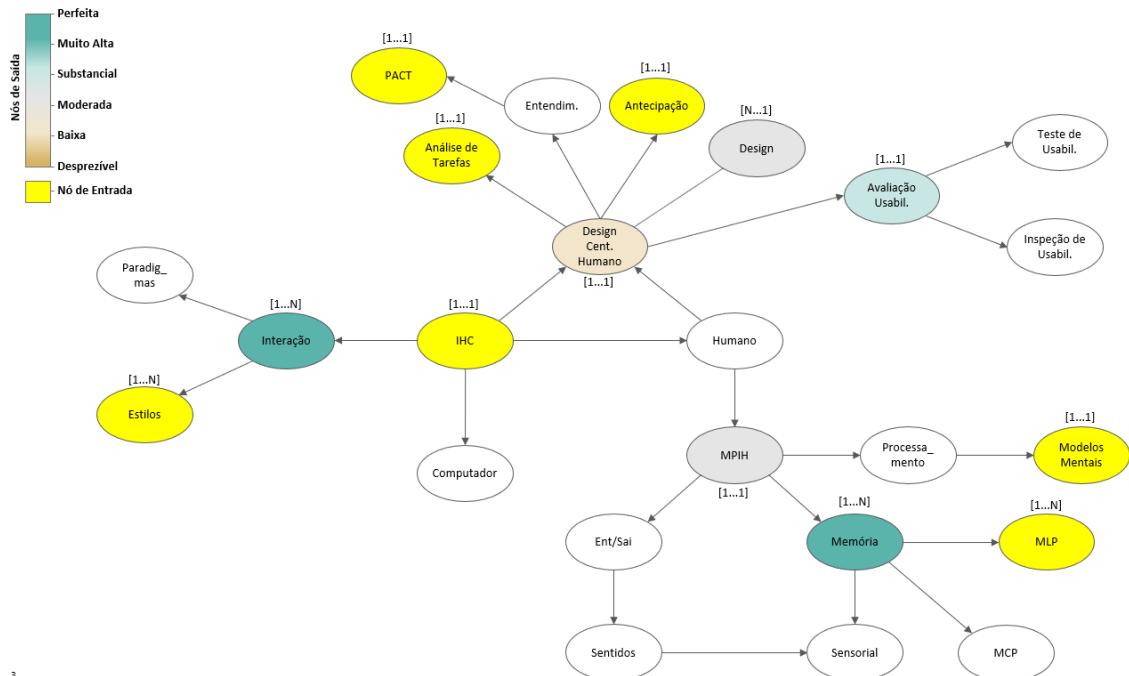


Figura 49 – RB utilizada no Experimento 2 - IHC 2017-1 (Execuções 3 e 4)



APÊNDICE D

Questionários utilizados no Experimento 4

Questionário de avaliação sobre a percepção dos usuários na utilização dos recursos de visualização

1) Com Relação à Percepção de Utilidade.

1.1 A funcionalidade de **visualização de desempenho** por meio de gráficos influenciou positivamente a minha aprendizagem na disciplina

1.2 A funcionalidade de **quizzes** influenciou positivamente a minha aprendizagem na disciplina

1.3 A funcionalidade de **gamificação** (ranking de pontuação e quadro de medalhas) influenciou positivamente a minha aprendizagem na disciplina

1.4 Eu gostaria que o CX fosse utilizado em outras disciplinas

	Discordo	Concordo
	Totalmente	Totalmente
1.1 A funcionalidade de visualização de desempenho por meio de gráficos influenciou positivamente a minha aprendizagem na disciplina	<input type="radio"/>	<input type="radio"/>
1.2 A funcionalidade de quizzes influenciou positivamente a minha aprendizagem na disciplina	<input type="radio"/>	<input type="radio"/>
1.3 A funcionalidade de gamificação (ranking de pontuação e quadro de medalhas) influenciou positivamente a minha aprendizagem na disciplina	<input type="radio"/>	<input type="radio"/>
1.4 Eu gostaria que o CX fosse utilizado em outras disciplinas	<input type="radio"/>	<input type="radio"/>

2) Como os gráficos de desempenho mudaram a sua forma de você estudar?

2.1 Motivação para estudo

2.2 Necessidade de melhora no desempenho

2.3 Competitividade

	Discordo	Concordo
	Totalmente	Totalmente
2.1 Motivação para estudo	<input type="radio"/>	<input type="radio"/>
2.2 Necessidade de melhora no desempenho	<input type="radio"/>	<input type="radio"/>
2.3 Competitividade	<input type="radio"/>	<input type="radio"/>

3) Com relação à sua satisfação, você gostou da visualização de desempenho por meio de gráficos apresentados pelo sistema?

- Sim
- Não
- Não Percebi

4) Qual tipo de gráfico você prefere para visualizar seu rendimento?

- Barra
- Linha
- Radar

5) Qual tipo de gráfico você prefere para visualizar seu rendimento com relação à média da turma?

- Barra
- Linha
- Radar

6) A comparação do meu desempenho com a média da turma influenciou positivamente a minha aprendizagem na disciplina

Discordo	Concordo
Totalmente	Totalmente
<input type="radio"/>	<input type="radio"/>