**Métricas de avaliação de
alinhamentos de ontologias**

Esdras Lins Bispo Junior

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA A
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Ciência da Computação
Orientadora: Profa. Dra. Renata Wassermann

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro do CNPq

São Paulo, setembro de 2011

Métricas de avaliação de alinhamentos de ontologias

Esta dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa realizada por Esdras Lins Bispo Junior em 04/08/2011.

O original encontra-se disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Profa. Dra. Renata Wassermann - IME-USP
- Prof. Dr. Flávio Soares Corrêa da Silva - IME-USP
- Profa. Dra. Anarosa Alves Franco Brandão - POLI-USP

Agradecimentos

Este trabalho é fruto de uma produção coletiva de muitas pessoas. É bem verdade que, de hoje em diante, este trabalho será atribuído como sendo de minha autoria. Porém, considero ser impossível ignorar ou desprezar, durante todo o processo, o quanto ele foi afetado através da influência da vida de várias pessoas.

É também bem triste perceber a minha incapacidade de poder fazer com justiça os meus agradecimentos neste momento. Como serei tentado a citar alguns nomes, ficam aqui já os meus sinceros agradecimentos a todos que contribuíram de alguma maneira para que tudo isto ocorresse. Sem vocês, com certeza, eu não teria condições de dar nenhum passo à frente.

Agradeço primeiramente àquEle que me criou e em quem deposito toda a minha esperança. Agradeço por tudo o que eu sou e o que tenho. Agradeço a D'us que “é bom e a sua misericórdia dura para sempre” (Salmos 100.5).

Agradeço aos meus familiares. Agradeço pelo apoio, carinho e cuidado. Agradeço a cada um por me agüentar por mais de 20 anos (tarefa nada fácil.. rs). Agradeço a minha mãe (Aidê Bispo), meus irmãos (Tarik/Luana e Ingrid Bispo), a minha irmã Simone Bispo (que conseguiu me aturar durante um mês e meio na minha chegada em São Paulo), às minhas tias/tio (Enylce, Ecilda e Emerson Bispo). Muito feliz pela vida de cada um de vocês.

Agradeço à minha orientadora Renata Wassermann. Agradeço pela liberdade e confiança depositadas em mim. Pelas palavras de apoio e de direcionamento. Agradeço muito pela tranquilidade transmitida durante todo o trabalho. Agradeço também pelo exemplo como professora e por ainda também alimentar em mim boas esperanças quanto à carreira docente. Muito obrigado!

Agradeço a todos os professores do Laboratório de Inteligência Artificial e Métodos Formais (LIAMF). Agradeço aos professores Flávio Soares, Marcelo Finger, Leilane Barros e Renata Wassermann pela oportunidade de ter sido aluno dos senhores. Agradeço, pois com vocês aprendi um pouco mais sobre o que é ser excelente no que se faz (apesar de eu ser bastante resmungão e sempre estar reclamando de alguma coisa... rs).

Agradeço a todos os mestrandos e doutorandos do LIAMF. Foram eles que agüentaram durante mais de dois anos as minhas conversas bestas... rs. Agradeço pela companhia, pelas conversas e por todos os cafezinhos. Agradeço a Fábio, Carlos, Bustamante, Raphael, Paulo, Viviane, Erika, Daniel, Rosi, Poliana, Tales, Wesley, Ricardo (os três... rs), Sirley, Alexandre (grande passo quântico), Bóris, Mijail, Herbert. Muito obrigado!

Agradeço a todos do Serviço de Alunos da Pós-Graduação (SVAPG). Agradeço a Pinho, Emerson e Alessandra. Muito bom contar com o trabalho e dedicação de vocês.

Agradeço a todos da Aliança Bíblica Universitária do Brasil (ABUB). Agradeço por ter encontrado um grupo de pessoas que acreditam que não existe necessariamente incompatibilidade entre fé e ciência. Agradeço por ter encontrado pessoas que testemunham no seu dia-a-dia com vidas

dignas de um Reino que já começou aqui, e permanecerá eternamente.

Agradeço a todos os meus amigos no Conjunto Residencial da USP (CRUSP). Agradeço aos meus vizinhos de quarto do apartamento 408 (Iremar e Fellipe, seus burgueses... rs). Agradeço pela experiência de conviver junto com vocês dois. Tenho aprendido muito com a nossa amizade e todas as nossas diferenças. Agradeço a Wesley (*Xylala*) pela companhia. Aos adjacentes do meu apartamento (Lucas, Vinícius, Luís, Maraca, Júlio, Dário, Stênio) e a todos que já passaram para tomar um chá, jogar dominó ou jogar conversa fora mesmo. Aos porteiros Joílson e Rosângela por todo o carinho, pelos vários sorrisos e longas conversas. Desempenham muito além do que a sua função. Muito obrigado!

Agradeço a algumas pessoas que acompanham de perto a minha caminhada e me conhecem mais de perto. Meus agradecimentos singulares a Fábio Franco, Riziely Herrera e Thiago Rodgers. Foi (e ainda é) muito bom ter vocês por perto. Espero continuar tendo sempre a oportunidade de ter grandes amizades na minha vida.

E, finalmente, ficam os meus agradecimentos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio para a realização deste trabalho (136210/2009-0). Agradeço por ter acreditado e investido em mim e no meu projeto de mestrado.

Resumo

Na área de emparelhamento de ontologias, são utilizadas algumas métricas para avaliar os alinhamentos produzidos. As métricas baseadas em alinhamento têm como princípio básico confrontar um alinhamento proposto com um alinhamento de referência. Algumas destas métricas, entretanto, não têm alcançado êxito suficiente porque (i) não conseguem discriminar sempre entre um alinhamento totalmente errado e um quase correto; e (ii) não conseguem estimar o esforço do usuário para refinar o alinhamento resultante. Este trabalho tem como objetivo apresentar uma nova abordagem para avaliar os alinhamentos de ontologias. A nossa abordagem apresenta uma métrica na qual utilizamos as próprias consultas normalmente já realizadas nas ontologias originais para julgar a qualidade do alinhamento proposto. Apresentamos também alguns resultados satisfatórios de nossa abordagem em relação às outras métricas já existentes e largamente utilizadas.

Palavras-chave: emparelhamento de ontologias, mapeamento de ontologias, alinhamento de ontologias, interoperabilidade, heterogeneidade.

Abstract

In the ontology matching field, different metrics are used to evaluate the resulting alignments. Metrics based on alignment adopt the basic principle of verifying a proposed alignment against a reference alignment. Some of these metrics do not achieve good results because (i) they cannot always distinguish between a totally wrong alignment and one which is almost correct; and (ii) they cannot estimate the effort for the user to refine the resulting alignment. This work aims to present a new approach to evaluate ontology alignments. Our approach presents a measure that uses the usual queries in the original ontologies to assess the quality of the proposed alignment. We also present some satisfactory results of our approach with regard to widely used metrics.

Keywords: ontology matching, ontology mapping, ontology alignment, interoperability, heterogeneity.

Sumário

Lista de Abreviaturas	ix
Lista de Símbolos	xi
Lista de Figuras	xiii
Lista de Tabelas	xv
1 Introdução	1
1.1 Motivações	2
1.2 Organização do Texto	3
2 Ontologias	5
2.1 Definição	5
2.2 Classificação de Ontologias	7
2.3 Expressividade e Modelos de Representação	7
2.4 Consultas em Ontologias com SPARQL	8
2.5 Considerações Finais	9
3 Emparelhamento de Ontologias	11
3.1 Aplicações	11
3.1.1 Comunicação de agentes	11
3.1.2 Integração de catálogos	11
3.2 Definições	12
3.2.1 Elemento mapeado, alinhamento e emparelhamento	12
3.2.2 Terminologia	13
3.3 Técnicas de Emparelhamento de Ontologias	14
3.4 Considerações Finais	15
4 Sistemas de Emparelhamento de Ontologias	17
4.1 Anchor-Prompt	17
4.2 Lily	18
4.3 ASMOV	20
4.4 RiMOM	21
4.5 Considerações Finais	21

5	Métricas baseadas em Alinhamento	23
5.1	Precisão e cobertura padrão	23
5.2	Precisão e cobertura relaxada	25
5.3	<i>Ontology Alignment Evaluation Initiative</i> (OAEI)	26
5.4	Considerações Finais	27
6	Proposta de Trabalho	29
6.1	Benefícios de uma métrica baseada em consultas	30
6.2	Proposta de implementação	30
6.3	Dificuldades da abordagem	31
6.4	Considerações Finais	32
7	Experimentos e Resultados	33
7.1	Criação das consultas de referência	33
7.1.1	Algoritmo proposto	34
7.2	Tradução de consultas	35
7.3	População de indivíduos das ontologias	37
7.4	Resultados	38
8	Conclusões	41
A	Ontologia O - Universidade A	43
B	Ontologia O' - Universidade B	49
C	Consultas - Alinhamento de Referência	53
D	Consultas - Alinhamento A_1	57
E	Consultas - Alinhamento A_2	61
F	Consultas - Alinhamento A_3	65
	Referências Bibliográficas	67

Lista de Abreviaturas

ARP	<i>Another RDF (Resource Description Framework) Parser.</i>
ASMOV	<i>Automated Semantic Mapping of Ontologies with Validation.</i>
B2B	<i>Business-to-Business.</i>
DAWG	<i>Data Access Working Group.</i>
DL	<i>Description Logic.</i>
DOLCE	<i>Descriptive Ontology for Linguistic and Cognitive Engineering.</i>
e.g.	<i>exempli gratia</i> (por exemplo).
ENIA	Encontro Nacional de Inteligência Artificial.
FMA	<i>Foundational Model of Anatomy Ontology.</i>
GOM	<i>Generic Ontology Matching.</i>
I3CON	<i>Information Interpretation and Integration Conference.</i>
i.e.	<i>id est</i> (isto é).
ISWC	<i>International Semantic Web Conference.</i>
LOM	<i>Large Scale Ontology Matching.</i>
MBA	Métrica baseada em Alinhamento.
MBC	Métrica baseada em Consultas.
OAEI	<i>Ontology Alignment Evaluation Initiative.</i>
OMD	<i>Ontology Mapping Debugging.</i>
OWL	<i>Web Ontology Language.</i>
RDF	<i>Resource Description Framework.</i>
RiMOM	<i>Risk Minimization based Ontology Mapping.</i>
SAT	<i>Boolean SATisfiability Problem.</i>
SMA	Sistemas Multiagentes.
SOM	<i>Semantic Ontology Matching.</i>
SPARQL	<i>Simple Protocol and RDF Query Language.</i>
SUMO	<i>Suggested Upper Merged Ontology.</i>
UML	<i>Unified Modeling Language.</i>
UMLS	<i>Unified Medical Language System.</i>
vs.	<i>versus.</i>
W3C	<i>World Wide Web Consortium.</i>
XML	<i>eXtensible Markup Language.</i>

Lista de Símbolos

A	Alinhamento entre ontologias.
e	Entidade de uma ontologia.
C	Conjunto de classes em uma ontologia.
C_O	Conjunto de consultas realizadas em uma ontologia O .
I	Conjunto de indivíduos em uma ontologia.
n	Medida de confiança em uma correspondência.
O	Ontologia.
p	Parâmetros de um emparelhamento.
r	Recursos externos de um emparelhamento.
R	Alinhamento de referência entre ontologias.
T	Conjunto de tipos de dados em uma ontologia.
U	Conjunto de relações em uma ontologia.
V	Conjunto de valores de dados em uma ontologia.
$\boxtimes$	Relação genérica entre entidades de duas ontologias.
$\perp$	Relação de exclusão (ou disjunção) em uma ontologia.
$\in$	Relação de instanciação em uma ontologia.
$=$	Relação de atribuição em uma ontologia.
$\equiv$	Relação de equivalência em uma ontologia.
$\supseteq$	Relação do tipo “mais geral” em uma ontologia.
$\sqsubseteq$	Relação do tipo “menos geral” em uma ontologia (especialização).
$\sqcap$	Relação de sobreposição em uma ontologia.

Lista de Figuras

1.1	Abordagem híbrida para interoperabilidade de modelos de reputação (Nardin, 2009).	3
2.1	Duas ontologias (O e O') que representam informações sobre o quadro de funcionários em duas universidades diferentes.	6
2.2	Várias formas de ontologias ordenadas pela sua expressividade (Euzenat e Shvaiko, 2007).	8
3.1	O processo de emparelhamento de ontologias	13
3.2	As ontologias da Figura 2.1 alinhadas através de um elemento mapeado $\langle id, Docente, ProfessorDE, .9, \equiv \rangle$	13
3.3	As ontologias da Figura 2.1 alinhadas através de vários elementos mapeados.	14
3.4	Classificação das técnicas elementares de emparelhamento de ontologias. A divisão superior é baseada na granularidade e na interpretação da entrada; a divisão inferior é baseada no tipo de entrada. A camada intermediária caracteriza classes de técnicas básicas.	16
4.1	A interface do Prompt no Protégé.	18
4.2	A arquitetura do sistema Lily (Wang e Xu, 2009).	19
4.3	A interface do usuário do Lily.	19
4.4	O processo de emparelhamento do ASMOV (Jean-Mary <i>et al.</i> , 2009).	20
5.1	Sobreposição entre as correspondências do alinhamento de referência e as correspondências do alinhamento proposto.	24
6.1	Esquema da abordagem de uma métrica baseada em consultas.	31

Lista de Tabelas

2.1	Consulta SPARQL que lista todos os indivíduos da classe Departamento da ontologia O e o seu respectivo resultado.	9
7.1	Alinhamentos A_1 , A_2 , A_3 e R para as ontologias da Figura 2.1 e suas respectivas correspondências.	33
7.2	Indivíduos das classes Docente da ontologia O e da classe ProfessorDE da ontologia O'	38
7.3	Resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura padrão, e a medida-F correspondente.	38
7.4	Resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura relaxadas, e a medida-F correspondente.	39
7.5	Resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura baseadas em consultas, e a medida-F correspondente.	39

Capítulo 1

Introdução

Vivemos em um momento histórico em que a produção da informação acontece em uma velocidade nunca antes ocorrida. Entre os diversos meios de comunicação hoje existentes, a Web tem ocupado um lugar de destaque. A Web é um dos principais meios utilizados pela sociedade civil, academia e organizações no acesso, transmissão e consumo do montante informacional produzido.

Um dos grandes desafios ainda hoje é a manipulação das informações disponíveis na Web. Já em meados dos anos 90, a heterogeneidade semântica já havia sido identificada como um dos maiores problemas ao lidar com a interoperabilidade e cooperação entre múltiplas fontes de dados (Kashyap e Sheth, 1996). O problema da heterogeneidade semântica está presente em várias áreas como, por exemplo, o gerenciamento de banco de dados (Sheth e Kashyap, 1992).

Com a necessidade de resolver problemas deste gênero, surge um novo conceito de Web. A Web Semântica traz à rede uma característica a mais: a possibilidade de tratar a semântica dos dados disponíveis. A fim de implementar este novo conceito, pesquisas recentes vêm recomendando fortemente o uso de ontologias (Maedche, 2001).

As ontologias fornecem uma especificação formal de um modelo conceitual (Gruber, 1993). Elas representam o conhecimento de um dado domínio e possibilitam a extração de informações implícitas a partir de informações explicitamente declaradas. Esta extração é chamada de inferência lógica e geralmente é automatizada com o uso de motores de inferências.

Entretanto, o uso de ontologias não é suficiente para resolver o problema da heterogeneidade. Visto que as fontes produtoras de informações são diversas, é muito possível que surjam também divergências entre as especificações dos modelos propostos. Ontologias diferentes podem ser criadas para especificar o mesmo domínio de interesse.

O emparelhamento de ontologias é uma alternativa para a resolução do problema da heterogeneidade semântica. O seu principal objetivo é estabelecer correspondências entre as entidades das ontologias em questão. Como resultado do processo de emparelhamento, tem-se um alinhamento entre as ontologias (Kalfoglou e Schorlemmer, 2003).

Com o intuito de avaliar os alinhamentos produzidos, algumas métricas são utilizadas para este propósito. As métricas baseadas em alinhamento têm como princípio básico confrontar um alinhamento proposto com um alinhamento de referência. Algumas destas métricas, entretanto, não têm alcançado êxito suficiente porque (i) não conseguem discriminar sempre entre um alinhamento totalmente errado e um quase correto; e (ii) não conseguem estimar o esforço do usuário para refinar o alinhamento resultante.

Este trabalho tem como objetivo apresentar uma nova abordagem de avaliar os alinhamentos

de ontologias. A nossa abordagem apresenta uma métrica na qual utilizamos as próprias consultas normalmente já realizadas nas ontologias originais para julgar a qualidade do alinhamento proposto. Ao final do trabalho, apresentamos também alguns resultados satisfatórios obtidos com a nossa abordagem em relação às outras métricas já existentes e largamente utilizadas.

1.1 Motivações

Os Sistemas Multiagentes (SMA) são considerados cada vez mais uma opção tecnológica viável ao desenvolvimento de sistemas complexos e abertos. Em uma sociedade de agentes, o controle social pode ser efetivado através da utilização de sistemas e modelos de reputação (Nardin, 2009).

Como não há como determinar, de maneira não intrusiva, o modelo de reputação utilizado pelos agentes do sistema, existe uma grande possibilidade de que os agentes utilizem modelos de reputação distintos. Porém, para que haja um controle efetivo da sociedade de agentes, todos os agentes têm que ser capazes de interagir uns com os outros sobre reputação.

O problema consiste, conseqüentemente, na co-existência desta diversidade de modelos de reputação em um mesmo sistema. Isto dificulta a interoperabilidade sobre reputação entre agentes, já que eles se baseiam em semânticas distintas para representar os conceitos de reputação.

Desta forma, suponha a instanciação do cenário de avaliação de pinturas com três agentes avaliadores, sendo eles Ana, Beto e Clara. Suponha ainda que a agente Ana receba uma solicitação de um agente cliente para avaliar uma pintura da época impressionista. Como ela identifica que seu conhecimento não é suficiente para avaliar uma pintura desta época, ela decide solicitar a opinião de outro agente avaliador. Devido a interações passadas, Ana sabe que Clara possuía conhecimento para avaliar pinturas da época impressionista, mas como já se passou muito tempo desde a última interação entre elas, Ana não está certa se Clara ainda tem boa reputação em relação à avaliação de pinturas desta época.

Contudo, Ana conhece o agente Beto, e Ana sabe que tal agente teve interações recentes com Clara. Assim, Ana resolve consultar Beto sobre a reputação de Clara como avaliadora de pinturas desta época. Ana consulta Beto, que responde informando que Clara possui uma boa reputação como avaliadora de pinturas da época impressionista. Desta forma, Ana atualiza sua reputação sobre Clara como avaliadora de pinturas desta época e resolve solicitar sua opinião. Clara responde à solicitação e recebe em troca o pagamento por sua opinião a Ana, que por sua vez responde ao agente cliente e recebe o pagamento pelo serviço prestado.

Para que este cenário possa ocorrer, todos os agentes têm que conhecer a semântica dos conceitos utilizados para a troca de informação, dentre os quais encontram-se os conceitos relacionados na noção de reputação. Considerando um ambiente dinâmico e com heterogeneidade de agentes e modelos de reputação, como garantir que estes agentes se comuniquem sobre reputação sem que haja necessidade de conhecimento explícito sobre os modelos usados pelos outros agentes?

Na Figura 1.1, é ilustrado o uso da abordagem híbrida na interoperabilidade entre dois agentes com modelos heterogêneos de reputação, utilizando uma arquitetura orientada a serviços.

Considerando esta abordagem e o cenário de avaliação de pinturas, quando Ana (**Agente A**) solicita a reputação de Clara a Beto (**Agente B**), ela elabora a solicitação utilizando os conceitos de seu modelo interno de reputação (**Modelo de Reputação A**). Porém, antes de encaminhá-la a Beto, solicita a tradução dos conceitos de sua ontologia interna pelos conceitos de uma **Ontologia**

Figura 1.1: Abordagem híbrida para interoperabilidade de modelos de reputação (Nardin, 2009).

Comum no Domínio da Reputação ao Serviço de Emparelhamento de Ontologias (Indicação 1). Uma vez traduzidos, substituem-se os conceitos na mensagem, que é encaminhada para Beto (Indicação 2). Ao receber a mensagem, Beto solicita a tradução dos conceitos da Ontologia Comum no Domínio da Reputação pelos conceitos de seu modelo interno de reputação (Modelo de Reputação B)-(Interação 3), substitui-os na mensagem, que pode então ser processada a fim de responder a Ana.

No trabalho analisado (Nardin, 2009), o emparelhamento é feito de forma manual.

1.2 Organização do Texto

O restante do trabalho está dividido como se segue. O Capítulo 2 faz uma breve apresentação sobre o conceito de ontologias. O Capítulo 3 descreve o emparelhamento de ontologias e suas técnicas básicas. O Capítulo 4 trata de alguns sistemas importantes de emparelhamento de ontologias. O Capítulo 5 apresenta as principais métricas baseadas em alinhamento. O Capítulo 6 propõe uma nova abordagem de avaliar alinhamentos utilizando uma métrica baseada em consultas. O Capítulo 7 descreve o experimento realizado e comenta sobre os seus resultados e, por fim, o Capítulo 8 encerra o trabalho trazendo algumas conclusões obtidas e perspectivas de trabalhos futuros.

Capítulo 2

Ontologias

O termo “ontologia” designa originalmente uma das áreas de estudo da Filosofia que se ocupa do conhecimento dos princípios e fundamentos últimos da realidade e dos seres (Chauí, 1999). Inicialmente, utilizava-se o termo “metafísica” para designar essa área de estudo. O uso do termo ontologia como sinônimo de metafísica foi introduzido no século XVII pelo filósofo alemão Jacobus Thomasius.

Em áreas como Inteligência Artificial, Lingüística Computacional e Teoria de Banco de Dados, o termo “ontologia” adquiriu um outro significado específico. Uma ontologia normalmente provê um vocabulário que descreve um domínio de interesse e uma especificação do significado dos termos utilizados no vocabulário. Dependendo da precisão desta especificação, a noção de ontologia inclui vários dados e modelos conceituais, por exemplo, conjunto de termos, classificações, esquemas de banco de dados (Shvaiko e Euzenat, 2008).

Apresentaremos a seguir, uma definição resumida de ontologia (Seção 2.1), como as ontologias se classificam (Seção 2.2) e como podemos enxergá-las sob o nível de expressividade de seus modelos de representação (Seção 2.3). Por fim, apresentaremos também a linguagem de consultas em ontologias SPARQL (Seção 2.4)

2.1 Definição

Existem na literatura diversas notações e definições para ontologias. Neste trabalho, seguiremos a definição de Euzenat e Shvaiko (2007), que é a usual na área de emparelhamento de ontologias. Segundo eles, uma ontologia é caracterizada da seguinte forma:

Definição 1. Uma ontologia é uma tupla $O = \langle C, I, U, T, V, \sqsubseteq, \perp, \in, = \rangle$, tal que:

C - conjunto de classes ou conceitos (são as principais entidades de uma ontologia. Elas são interpretadas como um conjunto de indivíduos no domínio);

I - conjunto de indivíduos ou objetos ou instâncias (são interpretados como um indivíduo particular de um domínio);

U - conjunto de relações (são interpretadas como um subconjunto do produto do domínio);

T - conjunto de tipos de dados (são partes de um domínio particular o qual se especifica valores a indivíduos);

V - conjunto de valores de dados - são os valores propriamente ditos (sendo C, I, U, T, V disjuntos entre si);

$\sqsubseteq$ - relação sobre $(C \times C) \cup (U \times U) \cup (T \times T)$ chamada *especialização*;

$\perp$ - relação sobre $(C \times C) \cup (U \times U) \cup (T \times T)$ chamada *exclusão*;

$\in$ - relação sobre $(I \times C) \cup (V \times T)$ chamada *instanciação*;

$=$ - relação sobre $I \times U \times (I \cup V)$ chamada *atribuição*.

Podemos apresentar como exemplo a ontologia O de uma universidade (conforme na Figura 2.1). **Pessoas** e **Estudante** são classes da ontologia O . **nome** e **id** são relações da classe **Pessoas**. **string** e **uri** são tipos de dados das relações **nome** e **id** respectivamente. **José Amabis** (em que “**José**” e “**Amabis**” são valores de dados do tipo **string**) é indivíduo da classe **Biologia**. **Docente** é uma especialização da classe **Pessoas** (referente à tupla $(\text{Docente}, \text{Pessoas}) \in (C \times C)$). **José Amabis** é uma instância de **Ciencia** (referente à tupla $(\text{José Amabis}, \text{Ciencia}) \in (I \times C)$), pois **Biologia** é uma especialização de **Ciencia**. “**Amabis**” é uma atribuição de **sobrenome** do indivíduo **José Amabis** (referente à tupla $(\text{José Amabis}, \text{sobrenome}, \text{“Amabis”}) \in (I \times U \times V)$).

Figura 2.1: Duas ontologias (O e O') que representam informações sobre o quadro de funcionários em duas universidades diferentes.

Segundo a formalização apresentada¹, a ontologia O' (ver Figura 2.1) pode ser descrita como se segue:

$$\begin{aligned}
 C &= \{ \text{Funcionario, Professor, ProfessorDE, ProfessorAssistente, ProfessorAssociado,} \\
 &\quad \text{ProfessorVisitante, Assistente, ProfAssociado, ProfVisitante, Doutorando, Sala,} \\
 &\quad \text{Referencia, Aula} \} \\
 I &= \{ \text{José Amabis} \} \\
 U &= \{ \text{nome, dataContr, departamento, publicacoes, leciona} \} \\
 T &= \{ \text{string, date} \} \\
 V &= \{ \text{"José Amabis"} \} \\
 \sqsubseteq &= \{ (\text{Professor, Funcionario}); (\text{Assistente, Funcionario}); (\text{Doutorando, Funcionario}); \\
 &\quad (\text{ProfessorDE, Professor}); (\text{ProfessorAssistente, Professor}); \\
 &\quad (\text{ProfessorAssociado, Professor}); (\text{ProfessorVisitante, Professor}); \\
 &\quad (\text{ProfAssociado, Assistente}); (\text{ProfVisitante, Assistente}) \} \\
 \perp &= \emptyset \\
 \in &= \{ (\text{José Amabis, ProfessorVisitante}) \} \\
 = &= \{ (\text{José Amabis, nome, "José Amabis"}) \}
 \end{aligned}$$

2.2 Classificação de Ontologias

As ontologias podem ser classificadas de várias formas, dependendo de qual enfoque é dado. É possível classificar as ontologias a partir de vários critérios. Maedche (2002) e Guarino (1998) classificam de acordo com o seu nível de generalidade:

- **Ontologias de alto nível:** descrevem conceitos muito gerais como espaço, tempo, evento;
- **Ontologias de domínio:** descrevem o vocabulário relacionado a um domínio genérico, através da especialização de conceitos introduzidos nas ontologias de alto nível;
- **Ontologias de tarefa:** Descrevem um vocabulário relacionado a uma tarefa ou atividade genérica, através da especialização de conceitos introduzidos nas ontologias de alto nível;
- **Ontologias de aplicação:** são as ontologias mais específicas por serem utilizadas dentro das aplicações.

Neste trabalho, ontologias de domínio serão utilizadas como parâmetros de entrada para o processo de emparelhamento.

2.3 Expressividade e Modelos de Representação

Em várias áreas da Ciência da Computação, existem diferentes modelos de dados e modelos conceituais que podem ser concebidos a partir de ontologias. Temos por exemplo folksonomias, esquemas de bancos de dados, modelos UML, hierarquias informais (e.g. diretórios), *thesauri*, esquemas XML e ontologias formais. Estes exemplos podem ser vistos na Figura 2.2 em função do nível de formalidade.

¹Para fins ilustrativos e com o propósito de não fatigar o leitor, foi considerado apenas um único indivíduo na formalização. As ontologias O e O' completamente populadas podem ser consultadas nos Apêndices A e B, respectivamente.

Logo, ao depararmos com uma ontologia que esteja no nível mais alto, espera-se que ela tenha uma semântica explícita e bem definida (considerando que a interpretação de uma estrutura de diretórios em um sistema de arquivos é mais fortemente implícita). A bem da verdade, depende tão somente de que o desenvolvedor tinha em mente no momento da construção do modelo, ou seja, o significado dos rótulos, o conhecimento e o contexto no qual os rótulos aparecem são todos implícitos e, por isso, não fazem parte da especificação do diretório (Euzenat e Shvaiko, 2007).

Figura 2.2: Várias formas de ontologias ordenadas pela sua expressividade (Euzenat e Shvaiko, 2007).

Entretanto, as ontologias com um nível de expressividade maior têm traços distintos. A interpretação destas não é delegada aos usuários que lêem o diagrama, mas especificada explicitamente. Nestas ontologias, a semântica que o desenvolvedor tinha em mente é devidamente documentada e passível de ser processada posteriormente.

Neste trabalho, utilizaremos como objeto de estudo ontologias projetadas no formato RDF (*Resource Description Framework*)².

2.4 Consultas em Ontologias com SPARQL

SPARQL (*Simple Protocol and RDF Query Language*) é uma linguagem para consultas e um protocolo para acessar ontologias no formato RDF. Ela vem sendo discutida pelo grupo de trabalho do *World Wide Web Consortium* (W3C) de Acesso de Dados em RDF (*RDF Data Access Working Group* - DAWG³). O W3C tem recomendado a sua utilização⁴.

Como uma linguagem de consulta, SPARQL é “orientada a dados”, isto é, não existe inferência na própria linguagem em si. Ela apenas extrai da ontologia aquilo que se é solicitado pela consulta.

Podemos apresentar como exemplo, uma consulta SPARQL que lista todos os indivíduos da classe *Departamento* da ontologia *O* apresentada na Figura 2.1. O resultado desta consulta é apresentado na Tabela 2.1. A consulta foi executada na ontologia *O* já populada (ver Apêndice A).

²<http://www.w3.org/TR/rdf-syntax/>

³<http://www.w3.org/2001/sw/DataAccess/>

⁴<http://www.w3.org/TR/rdf-sparql-query/>

Tabela 2.1: *Consulta SPARQL que lista todos os indivíduos da classe Departamento da ontologia O e o seu respectivo resultado.*

Consulta	<pre>prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> prefix onto: <http://www.owl-ontologies.com/universidadeA.owl#> select ?nome where { ?nome rdf:type onto:Departamento}</pre>
Resultado	<pre>depBio depComp depFilo depHist depMed</pre>

2.5 Considerações Finais

Este capítulo apresentou uma definição para ontologia, como as ontologias são classificadas e como podemos enxergá-las sob o nível de expressividade de seus modelos de representação. A linguagem de consultas em ontologias SPARQL também foi apresentada. De posse desta notação necessária, apresentaremos no próximo capítulo o processo de emparelhamento de ontologias.

Capítulo 3

Emparelhamento de Ontologias

A natureza distribuída do desenvolvimento de ontologias tem levado a ontologias dissimilares para domínio iguais ou sobrepostos. Como consequência, as várias partes que utilizam ontologias diferentes não se compreendem umas às outras perfeitamente. Para resolver estes problemas, é necessário utilizar um emparelhamento de ontologias para estabelecer a interoperabilidade entre os dois sistemas (Shvaiko e Euzenat, 2005).

Apresentaremos a seguir, duas aplicações nas quais se faz necessário o emparelhamento de ontologias (Seção 3.1), definições importantes relativas ao processo (Seção 3.2), e as técnicas elementares de emparelhamento (Seção 3.3).

3.1 Aplicações

3.1.1 Comunicação de agentes

Agentes são entidades computacionais caracterizadas pela autonomia e capacidade de interação. Eles se comunicam através de linguagens inspiradas na fala humana que determinam o “envelope” das mensagens e habilitam os agentes a posicioná-los dentro de um contexto particular de interação. As mensagens são expressas em linguagens de representação do conhecimento e freqüentemente se referem a alguma ontologia. Como consequência, quando dois agentes autônomos e independentemente projetados se encontram, eles têm a possibilidade de trocar mensagens, mas têm pouca possibilidade de compreender uns aos outros se não compartilharem a mesma linguagem e ontologia. Então, é necessário prover um meio para estes agentes emparelharem suas ontologias com o fim de ou traduzir suas mensagens ou integrar axiomas como pontes em seus próprios modelos (van Eijk *et al.*, 2001).

Uma solução para este problema é ter um protocolo de alinhamento de ontologias que pode ser intercalado com qualquer outro protocolo de interação do agente e que pode ser disparado pelo recebimento de uma mensagem expressa em uma ontologia alheia. Como consequência, os agentes que se encontram pela primeira vez e utilizam ontologias diferentes deveriam estar aptos a negociar o emparelhamento dos termos em suas respectivas ontologias e traduzir o conteúdo da mensagem que eles trocam com o auxílio do alinhamento.

3.1.2 Integração de catálogos

Em aplicações B2B (*Business-to-Business*), os parceiros de negócios armazenam os seus produtos em catálogos eletrônicos. Os catálogos são estruturas semelhantes a árvores, conhecidas como hierarquias de conceitos com atributos. Exemplos típicos de catálogos são os diretórios de produtos do Amazon¹ e do eBay². Para que uma empresa privada possa participar no mercado (e.g. eBay), é necessário determinar as correspondências entre entradas de seus catálogos e entradas do catálogo simples do mercado. Este processo de mapear entradas entre catálogos é conhecido como o problema de emparelhamento de catálogos (Bouquet *et al.*, 2003). Havendo identificado as correspondências entre as entradas dos catálogos, elas são posteriormente analisadas com o intuito de gerar expressões de consulta que automaticamente traduzam as instâncias de dados entre os catálogos (Velegrakis *et al.*, 2005). Com os catálogos alinhados, os usuários têm um acesso unificado aos produtos que estão à venda.

3.2 Definições

3.2.1 Elemento mapeado, alinhamento e emparelhamento

De acordo com Shvaiko e Euzenat (2005), podemos definir um *elemento mapeado* (ou correspondência) como uma 5-upla $\langle id, e, e', n, \boxtimes \rangle$, em que

- id é um identificador único de um elemento mapeado dado;
- e e e' são entidades (e.g. tabelas, elementos XML, propriedades, classes) da primeira e segunda ontologia respectivamente;
- n é uma *medida de confiança* em alguma estrutura matemática (normalmente no intervalo $[0; 1]$) apropriada para a correspondência entre as entidades e e e' ;
- $\boxtimes$ é uma relação (e.g. *equivalência* ($\equiv$); *mais geral* ($\supseteq$); *disjunção* ($\perp$); *sobreposição* ($\cap$)) apropriada entre entidades e e e' .

Um *alinhamento* é um conjunto de elementos mapeados. Uma operação de *emparelhamento* determina o alinhamento (A') para um par de ontologias (O e O'). Existem alguns outros parâmetros que podem estender a definição de um processo de emparelhamento, a saber: (i) o uso de um alinhamento de entrada (A) o qual está para ser completado pelo processo; (ii) os parâmetros do emparelhamento, p (e.g. pesos, limiares); e (iii) recursos externos utilizados pelo processo, r (e.g. *thesauri*); ver Figura 3.1.

Podemos apresentar como exemplo o elemento mapeado na Figura 3.2. O id poderia ser um identificador único qualquer (com o propósito de diferenciar os elementos mapeados entre si) como por exemplo 00006. As entidades e e e' são *Docente* e *ProfessorDE* respectivamente. A medida de confiança tem o valor de .9 (informando que o elemento mapeado admite uma confiabilidade de 90%, a partir dos critérios adotados pela regra de mapeamento). A relação entre as entidades é do tipo “equivalente” ($\equiv$). Logo, a 5-upla correspondente seria $\langle 00006, Docente, ProfessorDE, .9, \equiv \rangle$.

¹<http://www.amazon.com>

²<http://www.ebay.com>

Figura 3.1: O processo de emparelhamento de ontologias

Figura 3.2: As ontologias da Figura 2.1 alinhadas através de um elemento mapeado $\langle id, Docente, ProfessorDE, .9, \equiv \rangle$

Também podemos nos referenciar aos elementos mapeados de uma forma mais concisa, omitindo os seus identificadores. Como por exemplo, $Docente \equiv_{.9} ProfessorDE$ (ao invés da 5-upla $\langle 00006, Docente, ProfessorDE, .9, \equiv \rangle$) e $sala \equiv departamento$ (ao invés da 5-upla $\langle 00003, sala, departamento, 1.0, \equiv \rangle$).

Como exemplo de alinhamento, podemos apresentar o alinhamento R apresentado na Figura 3.3. As correspondências de R são descritas a seguir:

R	$Estudante \sqsubseteq Doutorando$	$Docente \equiv_{.9} ProfessorDE$
	$sala \equiv departamento$	$Disciplina \equiv_{.7} Aula$
	$leciona \equiv leciona$	$Departamento \equiv Sala$

Figura 3.3: As ontologias da Figura 2.1 alinhadas através de vários elementos mapeados.

3.2.2 Terminologia

Na área de emparelhamento de ontologias, diferentes autores utilizam termos diferentes para se referir a conceitos semelhantes (Euzenat, 2001; Kalfoglou e Schorlemmer, 2003; Noy e Klein, 2004). Como não há um consenso explícito, é mister apresentar as definições de alguns termos que serão utilizados no decorrer deste trabalho:

Emparelhamento é o processo de encontrar relacionamentos ou correspondências entre entidades de ontologias diferentes.

Alinhamento é um conjunto de elementos mapeados entre duas ou mais (em caso de emparelhamento múltiplo) ontologias. O alinhamento é o resultado de um processo de emparelhamento.

Correspondência ou **Elemento Mapeado** é uma relação entre entidades de duas ontologias diferentes representada como uma 5-upla $\langle id, e, e', n, \boxtimes \rangle$ (ver Seção 3.2).

Fusão de ontologias é a criação de uma nova ontologia a partir de duas, possivelmente sobrepostas, ontologias-fonte.

Integração de ontologias é a inclusão de uma ontologia em outra ontologia.

Tradução de ontologias é o processo de transformar uma ontologia de uma linguagem para uma outra linguagem.

3.3 Técnicas de Emparelhamento de Ontologias

Para classificação de técnicas elementares, Shvaiko e Euzenat (2005) introduziram duas classificações sintéticas baseadas nas propriedades mais evidentes das dimensões do emparelhamento:

- A classificação de **Granularidade / Interpretação de Entrada** é baseada (i) na granularidade do emparelhador, i.e., nível elementar ou estrutural, e então (ii) em como as técnicas geralmente interpretam a informação de entrada;
- A classificação de **Tipo de Entrada** é baseada no tipo de entrada em que as técnicas de emparelhamento elementares utilizam.

A classificação geral da Figura 3.4 (apresentada por [Euzenat e Shvaiko \(2007\)](#)) pode ser lida tanto de forma descendente (focando em como as técnicas interpretam a informação de entrada) como de forma ascendente (focando nos tipos de objetos manipulados) com o propósito de alcançar a camada de **Técnicas Básicas**.

Os emparelhadores elementares são diferenciados pela camada de **Granularidade / Interpretação de Entrada** de acordo com os seguintes critérios de classificação:

- **Nível elementar vs. Nível estrutural**: As técnicas de emparelhamento de nível elementar calculam correspondências analisando entidades ou instâncias daquelas entidades isoladamente, ignorando suas relações com outras entidades ou suas instâncias. As técnicas de nível estrutural calculam correspondências analisando como as entidades ou suas instâncias aparecem junto à estrutura. Este critério para abordagens baseadas em esquema é o mesmo adotado pela primeira vez por [Rahm e Bernstein \(2001\)](#), enquanto que a separação de nível elementar *vs.* nível estrutural para abordagens baseadas em instâncias vem do trabalho de [Kang e Naughton \(2003\)](#).
- **Sintático vs. externo vs. semântico**: A característica-chave das técnicas sintáticas é que elas interpretam a entrada somente no que diz respeito à sua estrutura, seguindo algum algoritmo claramente definido. As técnicas externas exploram recursos auxiliares (externos) de um domínio e o conhecimento comum com o objetivo de interpretar a entrada. Estes recursos podem ser entradas inseridas por um humano ou acessível em algum *thesaurus*, expressando o relacionamento entre os termos. As técnicas semânticas utilizam alguma semântica formal, e.g., semânticas teóricas de modelo, para interpretar a entrada e justificar seus resultados. No caso de um sistema de emparelhamento baseado em semântica, os algoritmos exatos são completos no que diz respeito à semântica, i.e., eles garantem descobrir todos os alinhamentos possíveis, enquanto que os algoritmos aproximados tendem a ser incompletos.

A classificação da camada **Tipo de Entrada** diz respeito ao tipo de entrada considerada por uma técnica particular:

- A primeira camada é classificada dependendo de qual tipo de dados os algoritmos trabalham: *strings* (terminológico), estrutura (estrutural), modelos (semântica) ou instâncias de dados (extensional). As duas primeiras são encontradas nas descrições das ontologias. A terceira requer alguma interpretação semântica da ontologia e normalmente utiliza algum motor de inferência semanticamente apropriado para deduzir correspondências. A última constitui na atual população de uma ontologia.
- A segunda camada dessa classificação decompõe além destas categorias, se necessário: métodos terminológicos que podem ser baseados em *string* (considerando os termos como seqüências

de caracteres) ou baseado na interpretação destes termos como objetos lingüísticos (lingüístico). A categoria de métodos estruturais é dividida em dois tipos de métodos: aqueles que consideram a estrutura interna das entidades, e.g., atributos e seus tipos (interno); e aquelas que consideram a relação de entidades com outras entidades (relacional).

3.4 Considerações Finais

Este capítulo apresentou duas aplicações nas quais se faz necessário o emparelhamento de ontologias, definições importantes relativas ao processo, e as técnicas elementares de emparelhamento. Apresentaremos no próximo capítulo alguns sistemas importantes de emparelhamento de ontologias.

Figura 3.4: Classificação das técnicas elementares de emparelhamento de ontologias. A divisão superior é baseada na granularidade e na interpretação da entrada; a divisão inferior é baseada no tipo de entrada. A camada intermediária caracteriza classes de técnicas básicas.

Capítulo 4

Sistemas de Emparelhamento de Ontologias

Na operação de emparelhamento de ontologias, encontrar manualmente as correspondências é tedioso, suscetível a erro, e obviamente não é possível em grande escala (a Web, por exemplo). O desenvolvimento de ferramentas para auxiliar no processo de emparelhamento de ontologias é crucial para o sucesso da Web Semântica (Doan *et al.*, 2002).

Descreveremos a seguir o funcionamento de quatro importantes sistemas de emparelhamento de ontologias: Anchor-Prompt, Lily, ASMOV e RiMOM. O Anchor-Prompt foi um dos sistemas pioneiros de emparelhamento de ontologias e até hoje é largamente citado¹ e utilizado por especialistas da área (Seção 4.1). O Lily (Seção 4.2), o ASMOV (Seção 4.3) e o RiMOM (Seção 4.4) destacaram-se nos últimos anos da *Ontology Alignment Evaluation Initiative*² (OAEI) ocupando as melhores posições em relação a outros sistemas na competição (Caracciolo *et al.*, 2008; Euzenat *et al.*, 2007, 2009).

4.1 Anchor-Prompt

O Anchor-Prompt (Noy e Musen, 2001) é uma ferramenta de fusão e alinhamento de ontologia com um sofisticado mecanismo de interação com usuário (*prompt*) para mapeamento de possíveis termos. É uma extensão do Prompt, também conhecido como SMART. O algoritmo de alinhamento Anchor-Prompt recebe duas ontologias como entrada e um conjunto de âncoras-pares de termos relacionados, os quais são identificados com o auxílio de técnicas baseada em *string*, ou definidos por um usuário, ou por um outro algoritmo que calcula a similaridade lingüística entre os nomes dos rótulos nos nós. Então o algoritmo refina os mapeamentos pela análise estrutural da ontologia e através da interação com o usuário.

O algoritmo constrói um grafo rotulado direcionado representando a ontologia por uma hierarquia de conceitos (chamada classes no algoritmo) e por uma hierarquia de relações (chamada *slots* no algoritmo), em que os nós no grafo são conceitos e as arestas são relações que denotam os relacionamentos entre os conceitos (os rótulos são os nomes das relações). Uma lista de âncoras-pares de conceitos relacionados definida pelos usuários ou automaticamente identificada por um emparelhador léxico é a entrada do algoritmo. O Anchor-Prompt analisa então os caminhos no sub-grafo

¹ <http://citeseerx.ist.psu.edu/showciting?doi=10.1.1.18.7666&sort=date>

² <http://oei.ontologymatching.org/>

limitado pelas âncoras e determina quais conceitos freqüentemente aparecem em posições similares sobre caminhos similares. Baseado nestas freqüências, o algoritmo decide se estes conceitos são semanticamente similares.

Os sistemas Prompt e Anchor-Prompt também têm contribuído para o projeto de outros algoritmos tais como PromptDiff que encontra diferenças entre duas ontologias e fornece a opção de transformar uma ontologia em outra. Na Figura 4.1, temos um *snapshot* do *plugin* do Prompt no Protégé.

Figura 4.1: A interface do Prompt no Protégé.

4.2 Lily

Lily (Wang e Xu, 2009) é um sistema que utiliza estratégias de emparelhamento híbridas para executar a tarefa de emparelhamento de ontologias. Lily pode ser usado para descobrir correspondências tanto para ontologias de pequeno porte quanto para ontologias de grande porte.

A idéia que norteia a estratégia de emparelhamento no Lily é combinar técnicas de emparelhamento atuais e eficientes para encontrar alinhamentos. Lily realiza quatro funções principais:

1. *Generic Ontology Matching (GOM)* é um método utilizado para tarefas de emparelhamento ordinárias com ontologias de pequeno porte;
2. *Large scale Ontology Matching (LOM)* é um método utilizado para tarefas de emparelhamento com ontologias de grande porte;
3. *Semantic Ontology Matching (SOM)* é um método utilizado para descobrir relações semânticas entre as ontologias. Lily utiliza conhecimento da Web para reconhecer estas relações

semânticas por meio de um mecanismo de busca;

4. *Ontology Mapping Debugging (OMD)* é utilizado para melhorar os alinhamentos resultantes.

O processo de emparelhamento consiste basicamente em três passos: a) **pré-processamento**: Lily analisa as ontologias e prepara os dados necessários para os passos subsequentes; b) **processamento**: o sistema utiliza métodos para calcular a similaridade entre os elementos de ontologias diferentes; c) **pós-processamento**: os alinhamentos são extraídos e então refinados pelo OMD. A arquitetura do Lily é apresentada na Figura 4.2.

Figura 4.2: A arquitetura do sistema Lily (Wang e Xu, 2009).

A última versão do Lily (versão 2.0) fornece uma interface gráfica amigável para o usuário. Um *snapshot* do Lily em execução é apresentado na Figura 4.3.

Figura 4.3: A interface do usuário do Lily.

4.3 ASMOV

Automated Semantic Mapping of Ontologies with Validation (ASMOV) (Jean-Mary *et al.*, 2009) é um algoritmo que automatiza o processo de emparelhamento de ontologias utilizando uma média ponderada das medidas de similaridade considerando juntamente alguns aspectos diferentes das ontologias em questão. Obtém-se o pré-alinhamento baseado nestas medidas e, então, semanticamente, verifica-se o alinhamento para garantir que não existem inconsistências semânticas.

O algoritmo ASMOV calcula iterativamente a similaridade entre entidades de um par de ontologias através da análise de quatro aspectos:

- elementos léxicos (id, rótulos e comentários);
- estrutura relacional (hierarquia ancestral-descendente);
- estrutura interna (restrições de propriedade para conceitos; tipos, domínios, e extensão para propriedades; valores de dados para indivíduos);
- extensão (instâncias de classes e valores de propriedades).

As medidas obtidas pela comparação destes quatro aspectos são combinadas em um simples valor utilizando uma soma ponderada de forma semelhante em (Euzenat e Valtchev, 2004).

O processo de emparelhamento do ASMOV é ilustrado na Figura 4.4. Na fase de pré-processamento, as ontologias são carregadas na memória utilizando o *Jena ARP Parser* e o componente de modelagem de ontologia do ASMOV. Um *thesaurus* é utilizado com o objetivo de calcular similaridades léxicas entre cada par de conceitos, propriedades e indivíduos. O ASMOV pode ser configurado para usar ou o *UMLS Metathesaurus* ou o *WordNet* para obter as medidas de similaridade. O usuário pode escolher por não usar um *thesaurus*; neste caso, um algoritmo de emparelhamento de texto é utilizado para calcular a distância léxica.

Figura 4.4: O processo de emparelhamento do ASMOV (Jean-Mary *et al.*, 2009).

Após isto, as similaridades entre pares de entidades juntamente a estrutura relacional, estrutura interna e dimensões de extensão são calculadas e uma medida de similaridade geral (ou valor de confiança) é armazenada em três matrizes bidimensionais, cada uma para conceitos, propriedades e indivíduos. Destas matrizes de similaridade, um pré-alinhamento é obtido pela seleção da entidade

de uma ontologia com o valor de confiança mais alto para uma entidade correspondente em outra ontologia.

Este pré-alinhamento então passa por uma verificação semântica, que detecta semanticamente correspondências inconsistentes e as suas causas. Estas correspondências inconsistentes são removidas do pré-alinhamento e são armazenadas para que o algoritmo não tente mapear as mesmas entidades em uma iteração subsequente; os mapeamentos são removidos a partir de um registro de inconsistências quando a causa subjacente desaparece.

4.4 RiMOM

Risk Minimization based Ontology Mapping (RiMOM) (Wang *et al.*, 2010) é um sistema de emparelhamento de ontologias dinâmico e multiestratégico. Ele implementa várias estratégias de emparelhamento as quais são baseadas em uma informação ontológica diferente. Para cada tarefa de emparelhamento individual, o RiMOM pode combinar automaticamente e dinamicamente múltiplas estratégias para gerar um resultado de emparelhamento composto. Recentemente, algumas novas características foram adicionadas na nova versão do RiMOM que o habilita a lidar com emparelhamento de ontologias desbalanceado, emparelhamento de ontologias com interação do usuário e emparelhamento de instâncias em grande escala. O RiMOM foi projetado com um *framework* de emparelhamento de ontologias em que diferentes tipos de estratégias de alinhamento podem ser incorporadas e configuradas.

4.5 Considerações Finais

Este capítulo descreveu o funcionamento de quatro importantes sistemas de emparelhamento de ontologias: Anchor-Prompt, Lily, ASMOV e RiMOM. Apresentaremos no próximo capítulo as métricas comumente utilizadas para avaliar os alinhamentos de ontologias.

Capítulo 5

Métricas baseadas em Alinhamento

As métricas baseadas em alinhamento (MBAs) têm como princípio básico confrontar um alinhamento proposto com um alinhamento de referência. Estas métricas surgiram da necessidade de avaliar a qualidade de alinhamentos produzidos por um algoritmo de emparelhamento de ontologias (Do *et al.*, 2003).

A primeira tarefa é fazer o emparelhamento entre as ontologias manualmente. O alinhamento resultante obtido manualmente é utilizado como o padrão de referência para avaliar a qualidade do resultado determinado automaticamente pelo algoritmo. Normalmente é um especialista no domínio o qual as ontologias representam que executa esta tarefa. Este alinhamento é chamado de alinhamento de referência (R).

A próxima etapa é comparar as correspondências do alinhamento de referência em relação às correspondências do alinhamento proposto pelo sistema (A). Podemos classificar esta relação entre as correspondências destes dois alinhamentos em quatro conjuntos (ver Figura 5.1):

- **os falsos positivos** ($A \setminus R$): são correspondências desnecessárias em R , mas propostas em A ;
- **os falsos negativos** ($R \setminus A$): são correspondências necessárias em R , mas não propostas em A ;
- **os autênticos negativos** ($(A \cup R)^c$): são correspondências desnecessárias em R que não são propostas em A ;
- **os autênticos positivos** ($A \cap R$): são correspondências necessárias em R que são propostas em A .

Baseadas na cardinalidade destes conjuntos, duas medidas são comumente utilizadas: precisão e cobertura. Estas medidas são originárias da área de recuperação da informação (van Rijsbergen, 1975). Elas vêm sido utilizadas de várias formas nas métricas baseadas em alinhamento. Descreveremos a seguir, a precisão e cobertura padrão (Seção 5.1) e a precisão e cobertura relaxada (Seção 5.2). Por fim, apresentaremos a OAEI e os seus objetivos (Seção 5.3).

5.1 Precisão e cobertura padrão

A precisão e a cobertura padrão são dadas a seguir:

Figura 5.1: Sobreposição entre as correspondências do alinhamento de referência e as correspondências do alinhamento proposto.

Definição 1 (Precisão Padrão). Dado um alinhamento de referência R , a precisão de um alinhamento A é dada por

$$P(A, R) = \frac{|A \cap R|}{|A|}$$

Definição 2 (Cobertura Padrão). Dado um alinhamento de referência R , a cobertura de um alinhamento A é dada por

$$C(A, R) = \frac{|A \cap R|}{|R|}$$

A precisão padrão tem como objetivo indicar a presença dos falsos positivos no alinhamento proposto. Podemos reescrever a Definição 1 da seguinte forma:

$$P(A, R) = \frac{|A \cap R|}{|A \cap R| \cup |A \setminus R|}$$

Assim um alinhamento de precisão máxima (valor 1) é aquele que não apresenta falsos positivos, isto é, $|A \setminus R| = \emptyset$.

Já a cobertura padrão tem como objetivo indicar a presença dos falsos negativos no alinhamento proposto. Podemos reescrever a Definição 2 da seguinte forma:

$$C(A, R) = \frac{|A \cap R|}{|A \cap R| \cup |R \setminus A|}$$

Assim um alinhamento de cobertura máxima (valor 1) é aquele que não apresenta falsos negativos, isto é, $|R \setminus A| = \emptyset$.

Um alinhamento ideal seria aquele que não contivesse falsos positivos, nem falsos negativos. O alinhamento ideal seria composto por apenas autênticos positivos. Logo, é necessário levar em conta tanto a precisão quanto a cobertura na avaliação da qualidade do alinhamento proposto.

Por isso, utiliza-se também em conjunto com a precisão e a cobertura, uma medida combinada de ambas chamada medida-F:

Definição 3 (Medida-F). *Dado um alinhamento de referência R e um número $\alpha \in [0; 1]$, a medida-F de um alinhamento A é dada por*

$$M_\alpha(A, R) = \frac{P(A, R) \times C(A, R)}{(1 - \alpha) \times P(A, R) + \alpha \times C(A, R)}$$

Se $\alpha = 1$, então a medida-F é igual à precisão, e se $\alpha = 0$, a medida-F é igual à cobertura. Quanto maior for o valor de α , maior importância é dada à precisão em relação à cobertura. Frequentemente, utiliza-se o valor de $\alpha = 0,5$, isto é,

$$M_{0,5}(A, R) = \frac{2 \times P(A, R) \times C(A, R)}{P(A, R) + C(A, R)}$$

Como os autênticos negativos não são propostos em R e nem no alinhamento A , tornam-se relevantes para fins de avaliação apenas os outros três conjuntos anteriormente apresentados.

5.2 Precisão e cobertura relaxada

Existem algumas críticas às medidas de precisão e cobertura padrão. A primeira é que elas não conseguem discriminar sempre entre um alinhamento totalmente errado e um quase correto. Isto deve-se principalmente à função de sobreposição utilizada. Ao calcular a intersecção entre A e R , as medidas de precisão e cobertura padrão ignoram a proximidade entre as correspondências e aceitam somente correspondências idênticas.

A função de sobreposição adotada nas medidas de precisão e cobertura padrão trata da mesma forma estes dois pares de correspondências:

- (1) (Estudante $\sqsupseteq$ Doutorando, Estudante $\equiv$ Doutorando)
- (2) (Estudante $\sqsupseteq$ Doutorando, leciona $\equiv$ leciona)

embora seja óbvio que o par (1) tenha uma maior similaridade entre as correspondências do que o par (2).

Uma outra crítica é uma consequência da primeira. Estas métricas não conseguem estimar o esforço do usuário para adaptar o alinhamento. Após proposto, o alinhamento normalmente é refinado pelo usuário, corrigindo as eventuais inadequações que quase sempre aparecem em um alinhamento proposto automaticamente. Ao não informar qualquer indicativo da proximidade das correspondências, estas métricas podem avaliar dois alinhamentos com a mesma valoração, mesmo que um destes seja de maior qualidade do que o outro.

Com o intuito de resolver esta fragilidade, foram propostas medidas de precisão e cobertura relaxadas (Ehrig e Euzenat, 2005). A idéia é substituir a função de sobreposição utilizada na abordagem padrão por uma outra função que calcule a similaridade entre as correspondências. As medidas de precisão e a cobertura relaxadas são dadas a seguir:

Definição 4 (Precisão Relaxada). *Dado um alinhamento de referência R e uma função de sobreposição ω_P , a precisão relaxada de um alinhamento A é dada por*

$$P_\omega(A, R) = \frac{\omega_P(A, R)}{|A|}$$

Definição 5 (Cobertura Relaxada). *Dado um alinhamento de referência R e uma função de sobreposição ω_C , a cobertura relaxada de um alinhamento A é dada por*

$$C_\omega(A, R) = \frac{\omega_C(A, R)}{|R|}$$

A função de sobreposição ω atende às seguintes propriedades:

- (1) $\forall A, B \quad \omega(A, B) \geq 0$ (positividade)
- (2) $\forall A, B \quad \omega(A, B) \leq \min(|A|, |B|)$ (maximalidade)
- (3) $\forall A, B \quad \omega(A, B) \geq |A \cap B|$ (minimalidade)

O propósito destas propriedades é que as medidas de precisão e a cobertura relaxadas sejam uma extensão da abordagem padrão. As propriedades (1) e (2) garantem que elas variam no intervalo $[0; 1]$. A propriedade (3) garante que elas terão a mesma valoração da abordagem padrão (na pior das hipóteses) ou contribuirão informando uma valoração maior.

5.3 *Ontology Alignment Evaluation Initiative* (OAEI)

Nos últimos anos, uma campanha vem sendo feita anualmente com o intuito de avaliar os métodos disponíveis para emparelhamento de ontologias: *Ontology Alignment Evaluation Initiative* (OAEI)¹. A OAEI vem acontecendo desde de 2004 com a organização da *Information Interpretation and Integration Conference* (I3CON)² e da *EON Ontology Alignment Contest*³. De 2006 em diante, a campanha vem sendo realizada em conjunto com o *International Semantic Web Conference* (ISWC), associada ao *ISWC Ontology Matching*.

Os objetivos da OAEI são:

- avaliar as potencialidades e fragilidades dos sistemas de emparelhamentos de ontologias;
- comparar o desempenho de técnicas;
- aumentar a comunicação entre os desenvolvedores de algoritmos;
- melhorar as técnicas de avaliação;
- ajudar o aperfeiçoamento do trabalho em emparelhamento de ontologias;

através de avaliação experimental e controlada do desempenho das técnicas.

Esta avaliação experimental é realizada através de uma biblioteca de testes que é submetida a cada sistema participante. A OAEI utiliza como métrica de avaliação as medidas de precisão e cobertura relaxadas. [Euzenat et al. \(2010\)](#) apresentam os resultados da competição em sua edição em 2010.

¹<http://oaei.ontologymatching.org/>

²<http://www.atl.external.lmco.com/projects/ontology/i3con.html>

³<http://oaei.ontologymatching.org/2004/Contest/>

5.4 Considerações Finais

Este capítulo apresentou duas métricas baseadas em alinhamento: as medidas de precisão e cobertura padrão e as medidas de precisão e cobertura relaxada. Também foi apresentado o OAEI e os seus objetivos. Apresentaremos no próximo capítulo a proposta de trabalho que se utiliza de uma métrica baseada em consultas.

Capítulo 6

Proposta de Trabalho

Apresentamos até agora as métricas mais comumente utilizadas para fornecer a avaliação de alinhamentos de ontologias. Estas métricas são dependentes diretamente de um alinhamento de referência, isto é, são métricas baseadas em um alinhamento. Os cálculos de precisão e cobertura realizados por estas métricas são todos obtidos em relação às correspondências pertencentes a este alinhamento.

Estas métricas têm sido bastante utilizadas em competições como a OAEI. Os alinhamentos propostos pelos sistemas de emparelhamento participantes da competição são comparados aos respectivos alinhamentos de referência. Estes alinhamentos de referências são produzidos por especialistas no domínio de conhecimento no qual a ontologia representa.

Porém, na maioria dos casos, a realidade não funciona como nas competições. Na prática, não se julga um alinhamento proposto a partir de um alinhamento de referência. Pois se há um alinhamento de referência para as ontologias dadas, não haveria necessidade real de propor um outro alinhamento para elas. Para a lógica e estrutura de uma competição, é bem pertinente a idéia de um alinhamento de referência. Contudo isto não parece plausível para a solução de problemas de heterogeneidade em aplicações do mundo real.

É bem verdade que o espírito de uma competição como a OAEI é, na realidade, utilizar os alinhamentos de referência com o objetivo de avaliar não só os alinhamentos propostos, mas principalmente a robustez dos sistemas de emparelhamento participantes. Aquele sistema que propõe os melhores alinhamentos dentro da competição, terá maiores possibilidades de propor bons alinhamentos em aplicações do mundo real. A biblioteca de testes (*benchmark*¹) a ser executada pelos sistemas no OAEI cobre uma gama bem variada de dificuldades críticas do processo de emparelhamento de ontologias.

Nosso propósito com este trabalho é apresentar uma métrica baseada em consultas (MBC) como uma abordagem alternativa ao uso de métricas baseadas em alinhamento. Na Seção 6.1, apontaremos os benefícios de uma métrica baseada em consultas. Na Seção 6.2, proporemos uma implementação para tal métrica. E, por fim, na Seção 6.3, evidenciaremos as dificuldades da abordagem.

¹<http://oei.ontologymatching.org/2010/>

6.1 Benefícios de uma métrica baseada em consultas

Existem vários benefícios do uso de uma métrica baseada em consultas. O primeiro deles é a possibilidade real do reuso da informação disponível acompanhada às ontologias existentes. Uma quantidade considerável de consultas normalmente já são realizadas para cada ontologia em seus respectivos contextos antes de serem alinhadas. Utilizar-se destas consultas com o propósito de avaliar a qualidade do alinhamento proposto é uma estratégia de abordagem coerente com procedimentos comuns, por exemplo, na área de construção de ontologias (Noy e McGuinness, 2001).

Um outro benefício é que cada consulta realizada em uma das ontologias alinhadas está diretamente associada a um fragmento da ontologia. Os fragmentos associados a uma consulta revelam uma parte da ontologia bastante requisitada pela aplicação que a utiliza. Com isto, uma consulta identifica através de seu fragmento um conjunto de entidades como fortes candidatas a participarem das correspondências pertencentes ao alinhamento sugerido.

Temos por fim um problema que consiste em obter o alinhamento para ontologias de grande porte. Conforme apontado por Gal e Shvaiko (2009), este é um dos grandes desafios para a área de emparelhamento de ontologias hoje. Para ontologias deste porte, até construir um alinhamento de referência torna-se uma tarefa bastante difícil (visto que o processo é manual, tedioso, bastante suscetível a erros e consome um tempo considerável (Noy e Musen, 2000)). Além disso, nem sempre temos à disposição um especialista da área para se dedicar a uma tarefa como essa.

É bem razoável que um engenheiro do conhecimento escreva um conjunto de consultas em parceria com o especialista da área para serem realizadas sobre o conjunto alinhamento-ontologias. Parece ser bem menos custoso e coerente com os propósitos da aplicação.

6.2 Proposta de implementação

Sejam O e O' as duas ontologias alinhadas pelo alinhamento A . Sejam C_O e $C_{O'}$ os respectivos conjuntos de consultas usualmente realizadas para as ontologias O e O' . O Algoritmo 1 apresenta a solução para aplicações que exigem apenas um alinhamento que seja unidirecional e o Algoritmo 2 apresenta a solução para aplicações que exigem um alinhamento bidirecional. Os Algoritmos 4 e 5 apresentam as funções que auxiliam no cálculo da precisão e cobertura aplicado a instâncias. A Figura 6.1 ilustra com um esquema a proposta apresentada.

Algoritmo 1 - Métrica Baseada em Consultas (Unidirecional)

METRICACONSULTASUNIDIRECIONAL(O, O', C_O, A)

1. $precisao \leftarrow 0$
 2. $cobertura \leftarrow 0$
 3. $t_p \leftarrow 0$
 4. $t_c \leftarrow 0$
 5. $precisao, cobertura, t_p, t_c \leftarrow \text{METRICACONSULTASAUZ}(O, O', C_O, A, precisao, cobertura, t_p, t_c)$
 6. $precisao \leftarrow precisao/t_p$
 7. $cobertura \leftarrow cobertura/t_c$
 8. **retorne** $precisao, cobertura$
-

Existe uma diferença fundamental entre uma métrica baseada em alinhamento e uma métrica baseada em consultas. Para a primeira, o cálculo de precisão e cobertura é realizado a partir das correspondências dos dois alinhamentos (A e R). Já para a segunda, o cálculo é realizado a partir das instâncias obtidas como respostas das consultas.

Algoritmo 2 - Métrica Baseada em Consultas (Bidirecional)

 METRICACONSULTASBIDIRECIONAL($O, O', C_O, C_{O'}, A$)

1. $precisao \leftarrow 0$
 2. $cobertura \leftarrow 0$
 3. $t_p \leftarrow 0$
 4. $t_c \leftarrow 0$
 5. $precisao, cobertura, t_p, t_c \leftarrow METRICACONSULTASAux(O, O', C_O, A, precisao, cobertura, t_p, t_c)$
 6. $precisao, cobertura, t_p, t_c \leftarrow$
 7. $METRICACONSULTASAux(O', O, C_{O'}, INVERTERALINHAMENTO(A), precisao, cobertura, t_p, t_c)$
 8. $precisao \leftarrow precisao/t_p$
 9. $cobertura \leftarrow cobertura/t_c$
 10. **retorne** $precisao, cobertura$
-

Algoritmo 3 - Métrica Baseada em Consultas: Algoritmo Auxiliar

 METRICACONSULTASAux($O, O', C_O, A, precisao, cobertura, t_p, t_c$)

1. **para** toda consulta c de C_O **faça**
 2. $c_{trans} \leftarrow TRANSFORMARCONSULTA(c, A)$
 3. $resp_{trans} \leftarrow CONSULTARONTOLOGIA(c_{trans}, O')$
 4. $resp_{prop} \leftarrow TRADUZIRINSTANCIA(resp_{trans}, INVERTERALINHAMENTO(A))$
 5. $resp_{ref} \leftarrow CONSULTARONTOLOGIA(c, O)$
 6. $interseccao, n_p \leftarrow PRECISAOINSTANCIA(resp_{prop}, resp_{ref})$
 7. $precisao \leftarrow precisao + interseccao$
 8. $t_p \leftarrow t_p + n_p$
 9. $interseccao, n_c \leftarrow COBERTURAINSTANCIA(resp_{prop}, resp_{ref})$
 10. $cobertura \leftarrow cobertura + interseccao$
 11. $t_c \leftarrow t_c + n_c$
 12. **fim para**
 13. **retorne** $precisao, cobertura, t_p, t_c$
-

Para uma dada consulta de referência c_R , já são esperadas um conjunto de instâncias como respostas. Ao consultarmos c_R em uma ontologia, podemos sobrepor os resultados esperados com os resultados obtidos e construir um diagrama semelhante ao da Figura 5.1.

Figura 6.1: Esquema da abordagem de uma métrica baseada em consultas.

6.3 Dificuldades da abordagem

Para as funções utilizadas no Algoritmo 3, não há dificuldades na implementação de CONSULTARONTOLOGIA e INVERTERALINHAMENTO. A linguagem SPARQL², por exemplo, foi criada com o

²<http://www.w3.org/TR/rdf-sparql-query/>

Algoritmo 4 - Cálculo da Precisão com InstânciasPRECISAOINSTANCIA(A, R)

```

1. interseccao  $\leftarrow 0$ 
2.  $n \leftarrow 0$ 
3. para toda instância  $a$  de  $A$  faça
4.   para toda instância  $r$  de  $R$  faça
5.     se  $a = r$  então
6.       interseccao  $\leftarrow$  interseccao + 1
7.     break
8.   fim se
9.   fim para
10.   $n \leftarrow n + 1$ 
11. fim para
12. retorne interseccao,  $n$ 

```

Algoritmo 5 - Cálculo da Cobertura com InstânciasCOBERTURAINSTANCIA(A, R)

```

1. retorne PRECISAOINSTANCIA( $R, A$ )

```

objetivo de realizar consultas em ontologias. E o processo de inversão de alinhamento é bem trivial. Para isto, invertem-se todas as correspondências do alinhamento (como $\text{Docente} \equiv_g \text{ProfessorDE}$ para $\text{ProfessorDE} \equiv_g \text{Docente}$).

Porém, existem dificuldades reais para a implementação das funções TRANSFORMARCONSULTA e TRADUZIRINSTANCIA. Euzenat *et al.* (2008) apresentou uma proposta para a utilização de consultas em linguagem SPARQL com o intuito de resolver o problema de tradução de dados adicionando novas triplas com a declaração CONSTRUCT. Entretanto o problema de como criar dinamicamente tais consultas, explorando o alinhamento entre as ontologias ainda é uma problema em aberto e um campo de pesquisa relativamente pouco estudado dentro da comunidade da Web Semântica (Correndo *et al.*, 2010).

Também é interessante salientar que a abordagem baseada em consultas é altamente dependente dos indivíduos da ontologia. Sem a presença deles, é impossível obter as medidas de precisão e cobertura. Torna-se, portanto, um pré-requisito que as ontologias alinhadas estejam populadas.

6.4 Considerações Finais

Neste capítulo, apresentamos a proposta de uma métrica baseada em consultas, apontamos os seus benefícios, propomos uma implementação para tal métrica, e evidenciamos as dificuldades da abordagem. Apresentaremos no próximo capítulo os experimentos e resultados realizados neste trabalho.

Capítulo 7

Experimentos e Resultados

Este capítulo tem como propósito apresentar os experimentos e resultados realizados no decorrer das atividades do mestrado. Nos experimentos, foram utilizadas as duas ontologias apresentadas na Figura 2.1 e os quatros alinhamentos da Tabela 7.1. Foi utilizado o Algoritmo 1 para a abordagem baseada em consultas e foram calculadas as medidas de precisão e cobertura padrão e as medidas de precisão e cobertura relaxadas.

Tabela 7.1: Alinhamentos A_1 , A_2 , A_3 e R para as ontologias da Figura 2.1 e suas respectivas correspondências.

A_1	Departamento $\equiv$ departamento leciona $\equiv$ leciona sala $\equiv$ Sala	nome $\sqsubseteq$ nome Estudante $\sqsupseteq$ Doutorando
A_2	Docente $\sqsubseteq$ Professor Disciplina $\equiv$ Aula Pessoas $\equiv$ Funcionario	Departamento $\equiv$ Sala sala $\equiv$ departamento leciona $\equiv$ leciona
A_3	Docente $\sqsupseteq_{.9}$ ProfessorDE nome $\equiv$ nome leciona $\equiv$ leciona	Estudante $\equiv$ Doutorando Funcionario $\equiv$ Assistente
R	Estudante $\sqsupseteq$ Doutorando sala $\equiv$ departamento leciona $\equiv$ leciona	Docente $\equiv_{.9}$ ProfessorDE Disciplina $\equiv_{.7}$ Aula Departamento $\equiv$ Sala

Os experimentos realizados neste trabalho têm como foco atender a dois propósitos: (i) executar a abordagem proposta pelo autor de uma MBC; e (ii) comparar os resultados desta com os resultados gerados pelas MBAs. Para cumprir estes objetivos, foi necessária a criação de consultas de referência para a ontologia O (Seção 7.1). Logo depois, as consultas criadas foram traduzidas a partir de cada alinhamento proposto para a ontologia O' (Seção 7.2). Em seguida, as ontologias foram efetivamente construídas e populadas (Seção 7.3). E, por fim, as abordagens deveriam ser executadas e seus resultados comparados (Seção 7.4).

7.1 Criação das consultas de referência

Uma MBC parte do princípio de que a ontologia a ser alinhada já está sendo utilizada por uma aplicação ou para uma finalidade específica. Naturalmente várias consultas são comumente

realizadas para a ontologia. Estas consultas revelam o propósito do uso da ontologia e o que se espera em seus resultados. Visto que para as ontologias apresentadas na Figura 2.1 não existia um conjunto de consultas usualmente realizadas, houve a necessidade de criá-lo.

A criação do conjunto de consultas de referência foi feita a partir da ontologia O e o alinhamento de referência R . A idéia é ter o alinhamento R também como referência para os experimentos utilizando a MBC. Com isto existe a possibilidade de se estabelecer um critério de julgamento objetivo entre as duas abordagens, de modo que se permita realizar comparações entre elas em condições de igualdade.

Em termos estatísticos, poderíamos dizer que é possível desta forma extrair das abordagens variáveis quantitativas contínuas passíveis de ser comparáveis. Estas variáveis quantitativas contínuas são as medidas de precisão e cobertura fornecidas pelas abordagens.

A possibilidade de comparação entre as abordagens consiste em que as MBAs fornecem os seus resultados baseados em seu alinhamento de referência. E, por sua vez, a MBC, para os resultados apresentados a seguir, tem as suas consultas originadas do mesmo alinhamento de referência utilizado pela outra abordagem. Isto indica que as duas métricas, nestas condições, apresentarão uma correlação entre si.

7.1.1 Algoritmo proposto

Seja O a ontologia de referência e R o alinhamento de referência dado. O Algoritmo 6 apresenta o processo de criação de consultas.

Algoritmo 6 - Criação de Consultas a partir do Alinhamento de Referência

```

CRIARCONSULTAS( $O, O', R$ )
1.  $consultas \leftarrow \emptyset$  // consultas é uma coleção de variáveis
2. para toda instância  $r$  de  $R$  faça
3.   se ENTRECLASSES( $r$ ) então
4.      $consultas \leftarrow consultas \cup \text{LISTARINDIVIDUOS}(\text{CLASSE}(r, O))$ 
5.   fim se
6.   se ENTRERELACOES( $r$ ) então
7.      $consultas \leftarrow consultas \cup \text{LISTARINDIVIDUOS}(\text{CLASSEDOMINIO}(r, O))$ 
8.      $consultas \leftarrow consultas \cup \text{LISTARINDIVIDUOS}(\text{CLASSEIMAGEM}(r, O))$ 
9.   fim se
10. fim para
11. retorne  $consultas$ 

```

O princípio utilizado no Algoritmo 6 é que as entidades envolvidas nas correspondências em R trazem consigo informações relevantes sobre o alinhamento de referência. Para as correspondências entre classes (ENTRECLASSES(r)), a classe da ontologia de referência é identificada e cria-se uma consulta que lista os seus indivíduos. Esta torna-se uma consulta relativa à correspondência entre classes utilizada em R .

Já para as correspondências entre relações (ENTRERELACOES(r)), a relação da ontologia de referência é identificada e criam-se duas consultas: (i) a primeira que lista todos os indivíduos do domínio da relação; e (ii) a segunda que lista todos os indivíduos da imagem da relação. Tornam-se estas as duas consultas relativas à correspondência entre relações utilizada em R .

Podemos apresentar como exemplo a correspondência $\text{Estudante} \sqsubseteq \text{Doutorando}$ que pertence a R . A consulta SPARQL criada lista todos os indivíduos da classe **Estudante**:

```

prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeA.owl#>

```

```
select ?nome
where { ?nome rdf:type onto:Estudante}
```

E para a correspondência $sala \equiv departamento$, são criadas duas consultas SPARQL a partir da relação *sala*. A primeira lista todos os indivíduos que fazem parte da imagem de *sala*:

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeA.owl#>
prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#>
```

```
select ?nome
where { onto:sala rdfs:range ?obj.
       ?nome rdf:type ?obj}
```

E a segunda, lista todos os indivíduos que fazem parte do domínio de *sala*:

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeA.owl#>
prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#>
```

```
select ?nome
where { onto:sala rdfs:domain ?obj.
       ?nome rdf:type ?obj}
```

O Algoritmo 6 indica apenas dois casos a serem tratados para a criação das consultas de referência porque já se era conhecido de antemão os tipos de correspondências. O algoritmo foi proposto para criar as consultas para este experimento específico. A proposição de um algoritmo mais geral requer um nível de tratamento mais refinado dependendo do nível de expressividade do alinhamento. Todavia, para o escopo deste trabalho, o Algoritmo 6 foi suficiente e atendeu satisfatoriamente às nossas necessidades.

No Apêndice C, são apresentadas todas as consultas de referência que foram criadas para o experimento.

7.2 Tradução de consultas

Após a criação das consultas de referência, para cada alinhamento proposto, é necessário o processo de tradução de consultas. Cada consulta de referência é traduzida a partir de alguma correspondência existente no alinhamento proposto. Na ausência de uma correspondência que sirva durante o processo de tradução, a consulta de referência deixa de ser traduzida.

O objetivo do processo de tradução de consultas é verificar o resultado das consultas traduzidas na ontologia-alvo (O'). De posse destes resultados, a MBC tem condições de calcular as medidas de precisão e a cobertura, utilizando os resultados obtidos com a ontologia de referência (O).

Seja C_O o conjunto de consultas de referência, O' a ontologia-alvo e A o alinhamento proposto. O Algoritmo 7 descreve o processo de tradução de consultas.

Algoritmo 7 - Processo de Tradução de Consultas

```

TRADUZIRCONSULTAS( $C_O, O', A$ )
1.  $consultas \leftarrow \emptyset$  // consultas é uma coleção de variáveis
2. para toda instância  $c$  de  $C_O$  faça
3.   se PROVEMDECLASSE( $c$ ) então
4.      $consultas \leftarrow consultas \cup \text{ENCONTRARCLASSERELATIVA}(c, A)$ 
5.   fim se
6.   se PROVEMDERELACAO( $c$ ) então
7.      $consultas \leftarrow consultas \cup \text{ENCONTRARRELACAORELATIVA}(c, A)$ 
8.   fim se
9. fim para
10. retorne  $consultas$ 

```

São dois os casos tratados no Algoritmo 7. Para as consultas que provêm de classes (PROVEMDECLASSES(c)), a classe da ontologia alvo é identificada (ENCONTRARCLASSERELATIVA(c, A)) e cria-se uma consulta que lista os seus indivíduos. Esta torna-se a consulta traduzida baseada na correspondência do alinhamento proposto.

Já para as consultas que provêm de relações (PROVEMDERELAÇÕES(c)), a relação da ontologia alvo é identificada (ENCONTRARRELACAORELATIVA(c, A)) e cria-se uma consulta que (i) lista todos os indivíduos do domínio da relação (se a consulta original for referente ao domínio); ou (ii) lista todos os indivíduos da imagem da relação (se a consulta original for referente à imagem). Torna-se esta consulta traduzida baseada na correspondência entre relações do alinhamento proposto.

Podemos apresentar como exemplo a consulta SPARQL de referência criada

```

prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:    <http://www.owl-ontologies.com/universidadeA.owl#>

select ?nome
where { ?nome rdf:type onto:Estudante}

```

e sua respectiva consulta traduzida a partir da correspondência $\text{Estudante} \sqsubseteq \text{Doutorando}$ (pertencente ao alinhamento A_1) é

```

prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:    <http://www.owl-ontologies.com/universidadeB.owl#>

select ?nome
where { ?nome rdf:type onto:Doutorando}

```

Para a mesma consulta de referência, não é possível encontrar nenhuma correspondência que sirva ao processo de tradução para o alinhamento A_2 . Logo, neste caso, a tradução não é feita para esta consulta.

E para a consulta SPARQL de referência criada

```

prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:    <http://www.owl-ontologies.com/universidadeA.owl#>
prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#>

select ?nome
where { onto:sala rdfs:range ?obj.
       ?nome rdf:type ?obj}

```


sua respectiva consulta traduzida a partir da correspondência $sala \equiv departamento$ (pertencente ao alinhamento A_2) é

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:    <http://www.owl-ontologies.com/universidadeB.owl#>
prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#>

select ?nome
where { onto:departamento rdfs:range ?obj.
        ?nome rdf:type ?obj }
```

Para a mesma consulta de referência, não é possível encontrar nenhuma correspondência que sirva ao processo de tradução para o alinhamento A_1 . Logo, neste caso, a tradução não é feita para esta consulta.

Semelhante como justificado na Seção 7.1, o Algoritmo 7 foi proposto para traduzir as consultas para este experimento específico. A proposição de um algoritmo mais geral requer um nível de tratamento mais refinado dependendo do nível de expressividade do alinhamento. Todavia, para o escopo deste trabalho, ele foi suficiente e atendeu satisfatoriamente às nossas necessidades.

Nos Apêndices D, E e F são apresentadas todas as consultas traduzidas através dos alinhamentos A_1 , A_2 e A_3 , respectivamente.

7.3 População de indivíduos das ontologias

Nas etapas de criação e tradução de consultas, não são utilizadas todas as informações de cada correspondência. Foram utilizadas apenas as entidades e a “natureza” da correspondência (i.e. $ENTRECLASSES(r)$). A relação entre as entidades e a medida de confiança são duas informações muito importantes que não são foram utilizadas nestas etapas.

A etapa de população de indivíduos das ontologias é a responsável por utilizar estas informações. A relação e a medida de confiança de cada correspondência do alinhamento de referência R são utilizadas. Utilizamos a relação da correspondência para indicar a relação do conjunto de indivíduos das duas ontologias: relação de continência ($\sqsubseteq$ e $\sqsupseteq$), de igualdade ($\equiv$), de intercessão ($\sqcap$) e de disjunção ($\sqcup$).

Já a medida de confiança é utilizada como uma fonte de incertezas na relação entre os conjuntos de indivíduos. Por exemplo, para a correspondência $Disciplina \equiv_{0,7} Aula$, a classe *Disciplina* terá os mesmos indivíduos da classe *Aula* com uma margem de erro de 30% (na ontologia de referência ou na ontologia alvo). Uma medida de confiança 0,7 indica um erro de 30% na atribuição da correspondência (o que poderíamos dizer também 15% para mais ou para menos).

Para ilustrar o procedimento, tomemos a correspondência $Docente \equiv_{0,9} ProfessorDE$. As duas classes foram populadas como apresentadas na Tabela 7.2. A classe *ProfessorDE* tem 11 indivíduos da classe *Docente* (que tem 13 indivíduos). Foi utilizada a relação de equivalência e a medida de confiança 0,9 (temos (i) $10\% \times 13 = 1,3$; (ii) $[1, 3] = 2, 0$; logo (iii) $13 - 2 = 11$).

Nos Apêndices A e B, são apresentadas as duas ontologias construídas e populadas, a ontologia fonte e a ontologia alvo, respectivamente. As ontologias foram construídas e populadas utilizando o ambiente Protégé 3.4.6.

Tabela 7.2: *Indivíduos das classes Docente da ontologia O e da classe ProfessorDE da ontologia O'.*

	Docente	ProfessorDE
Indivíduos	docAlmeidaJr	docAlmeidaJr
	docCezar	docCezar
	docJoseAmabis	--
	docFlavioSoares	--
	docLeilaneBarros	docLeilaneBarros
	docMarceloFinger	docMarceloFinger
	docRenataWassermann	docRenataWassermann
	docMarilenaChaui	docMarilenaChaui
	docNewtonCosta	docNewtonCosta
	docRenatoJanine	docRenatoJanine
	docNelsonWolosker	docNelsonWolosker
	docRonaldoGryschek	docRonaldoGryschek
	docThelmaOkay	docThelmaOkay

7.4 Resultados

Na Tabela 7.3, apresentamos os resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura padrão, e a medida-F correspondente. Percebe-se pelos resultados que A_2 é o melhor alinhamento dentre os três propostos. Porém não é possível diferenciar pelos resultados se o alinhamento A_1 é melhor do que o alinhamento A_3 . Os valores de precisão e cobertura dos dois são idênticos.

Tabela 7.3: *Resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura padrão, e a medida-F correspondente.*

Métrica	(A_1, R)	(A_2, R)	(A_3, R)
$P(A, R)$	0,40	0,50	0,40
$C(A, R)$	0,33	0,50	0,33
$M_{0.5}(A, R)$	0,37	0,50	0,37

Na Tabela 7.4, apresentamos os resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura relaxadas, e a medida-F correspondente. O alinhamento A_2 é apontado ainda como o melhor alinhamento. Porém é possível agora diferenciarmos o alinhamento A_1 do A_3 . A precisão e cobertura relaxada consegue apontar a diferença entre eles, indicando o A_3 como melhor do que A_1 .

Na Tabela 7.5, apresentamos os resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura baseadas em consultas, e a medida-F correspondente utilizando o Algoritmo 1. Embora os valores da precisão e cobertura sejam diferentes dos apresentados na Tabela 7.4, é possível diferenciar os alinhamentos de forma a concluir que $A_1 < A_3 < A_2$ em termos de qualidade.

Os resultados aqui apresentados foram sintetizados no artigo (Bispo Jr. e Wassermann, 2011).

Tabela 7.4: Resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura relaxadas, e a medida- F correspondente.

Métrica	(A_1, R)	(A_2, R)	(A_3, R)
$P_\omega(A, R)$	0,40	0,69	0,58
$C_\omega(A, R)$	0,33	0,69	0,48
$M_{0.5}(A, R)$	0,37	0,69	0,53

Tabela 7.5: Resultados da avaliação dos alinhamentos A_1 , A_2 e A_3 utilizando as medidas de precisão e cobertura baseadas em consultas, e a medida- F correspondente.

Métrica	(A_1, C_R)	(A_2, C_R)	(A_3, C_R)
$P_C(A, C_R)$	0,79	0,60	0,83
$C_C(A, C_R)$	0,33	0,67	0,42
$M_{0.5}(A, C_R)$	0,46	0,64	0,56

Capítulo 8

Conclusões

Inicialmente a avaliação de alinhamentos de ontologias era realizada utilizando as medidas de precisão e cobertura padrão. Porém, ultimamente, têm sido propostas as medidas de precisão e cobertura relaxadas com o intuito de discriminar de uma forma mais acurada os alinhamentos entre si. Esta diferenciação é importante para estimar o esforço do usuário para refinar o alinhamento após o processo.

Apresentamos uma métrica baseada em consultas como uma abordagem alternativa para a avaliação de alinhamentos. Esta métrica conseguiu diferenciar os alinhamentos propostos no exemplo de motivação da mesma forma que as medidas de precisão e cobertura relaxadas. Dentre os vários benefícios desta métrica, podemos destacar o reuso das consultas que já são usualmente realizadas nas ontologias emparelhadas e o fato de não haver necessidade da criação de um alinhamento de referência.

Uma dificuldade significativa da métrica baseada em consultas consiste no problema de tradução de dados que está diretamente relacionado à expressividade das lógicas utilizadas. Uma possibilidade de abordar o problema em trabalhos futuros seria resolvê-lo em lógicas menos expressivas e ir gradualmente incrementando para linguagens mais complexas.

Como um próximo passo bastante importante, seria relevante propor uma forma mais eficiente de comparação de métricas de avaliação de alinhamentos de ontologias. As comparações realizadas na área de emparelhamento de ontologias (como a utilizada neste trabalho) até então são carentes de uma validação estatística mais séria.

Também, como um trabalho futuro, seria importante a validação da proposta com ontologias maiores já existentes (e.g. catálogos, currículos, ontologias biomédicas). Um outro trabalho interessante seria a validação da proposta a partir das ontologias que fazem parte do conjunto de testes do OAEI. Estes resultados podem proporcionar bons indicadores a respeito da qualidade da abordagem.

Apêndice A

Ontologia *O* - Universidade A

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns="http://www.owl-ontologies.com/universidadeA.owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://www.owl-ontologies.com/universidadeA.owl">
  <owl:Ontology rdf:about=""/>
  <rdfs:Class rdf:ID="Medicina">
    <rdfs:subClassOf>
      <rdfs:Class rdf:ID="Docente"/>
    </rdfs:subClassOf>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Pessoas"/>
  <rdfs:Class rdf:ID="Funcionario">
    <rdfs:subClassOf rdf:resource="#Pessoas"/>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Estudante">
    <rdfs:subClassOf rdf:resource="#Pessoas"/>
  </rdfs:Class>
  <rdfs:Class rdf:ID="CienciaComp">
    <rdfs:subClassOf>
      <rdfs:Class rdf:ID="Ciencia"/>
    </rdfs:subClassOf>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Filosofia">
```

```

    <rdfs:subClassOf>
      <rdfs:Class rdf:about="#Docente"/>
    </rdfs:subClassOf>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Biologia">
    <rdfs:subClassOf>
      <rdfs:Class rdf:about="#Ciencia"/>
    </rdfs:subClassOf>
  </rdfs:Class>
  <rdfs:Class rdf:about="#Docente">
    <rdfs:subClassOf rdf:resource="#Pessoas"/>
  </rdfs:Class>
  <rdfs:Class rdf:ID="Departamento"/>
  <rdfs:Class rdf:ID="Disciplina"/>
  <rdfs:Class rdf:about="#Ciencia">
    <rdfs:subClassOf rdf:resource="#Docente"/>
  </rdfs:Class>
  <owl:ObjectProperty rdf:ID="sala">
    <rdfs:range rdf:resource="#Departamento"/>
    <rdfs:domain rdf:resource="#Docente"/>
  </owl:ObjectProperty>
  <owl:ObjectProperty rdf:ID="curso">
    <rdfs:domain rdf:resource="#Estudante"/>
    <rdfs:range rdf:resource="#Disciplina"/>
  </owl:ObjectProperty>
  <owl:ObjectProperty rdf:ID="leciona">
    <rdfs:domain rdf:resource="#Docente"/>
    <rdfs:range rdf:resource="#Disciplina"/>
  </owl:ObjectProperty>
  <owl:DatatypeProperty rdf:ID="sobrenome">
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
    <rdfs:domain rdf:resource="#Pessoas"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="id">
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int"/>
    <rdfs:domain rdf:resource="#Pessoas"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="nome">
    <rdfs:domain rdf:resource="#Pessoas"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="dataNasc">
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#date"/>

```



```

    <rdfs:domain rdf:resource="#Pessoas"/>
</owl:DatatypeProperty>
<Disciplina rdf:ID="discAbiogenese"/>
<Funcionario rdf:ID="funcRejane"/>
<Estudante rdf:ID="estFernanda"/>
<Estudante rdf:ID="estRoberta"/>
<Estudante rdf:ID="estCarlos"/>
<Estudante rdf:ID="estPaula"/>
<Disciplina rdf:ID="discEpidemiologia"/>
<Disciplina rdf:ID="discPreHistoria"/>
<CienciaComp rdf:ID="docFlavioSoares"/>
<Funcionario rdf:ID="funcJose"/>
<Estudante rdf:ID="estNoberto"/>
<Funcionario rdf:ID="funcRenato"/>
<Departamento rdf:ID="depComp"/>
<Estudante rdf:ID="estFlavio"/>
<Estudante rdf:ID="estLuana"/>
<Medicina rdf:ID="docNelsonWolosker"/>
<Departamento rdf:ID="depMed"/>
<Funcionario rdf:ID="funcClaudio"/>
<Filosofia rdf:ID="docMarilenaChaui">
    <sobrenome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Chauí</sobrenome>
    <nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Marilena</nome>
</Filosofia>
<Estudante rdf:ID="estOlivia"/>
<Disciplina rdf:ID="discHistSocialBrasil"/>
<Funcionario rdf:ID="funcLais"/>
<Funcionario rdf:ID="funcJosefa"/>
<CienciaComp rdf:ID="docLelianeBarros"/>
<Disciplina rdf:ID="discIA"/>
<Departamento rdf:ID="depBio"/>
<Estudante rdf:ID="estMercia"/>
<Estudante rdf:ID="estGustavo"/>
<Funcionario rdf:ID="funcRomulo"/>
<Estudante rdf:ID="estOlavo"/>
<Funcionario rdf:ID="funcAnaEliza"/>
<Disciplina rdf:ID="discFilMedieval"/>
<Estudante rdf:ID="estPatricio"/>
<Funcionario rdf:ID="funcAntonio"/>
<Funcionario rdf:ID="funcNivea"/>
<Estudante rdf:ID="estEduardo"/>

```

```

<Estudante rdf:ID="estJuliana"/>
<Biologia rdf:ID="docCezar"/>
<Estudante rdf:ID="estAmanda">
  <nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Amanda</nome>
  <sobrenome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Vasconcelos</sobrenome>
</Estudante>
<Estudante rdf:ID="estAna"/>
<Disciplina rdf:ID="discProgramacao"/>
<Funcionario rdf:ID="funcMario"/>
<Disciplina rdf:ID="discTeoriaConhecimento"/>
<Medicina rdf:ID="docRonaldoGryschek"/>
<Disciplina rdf:ID="discBacteriologia"/>
<Estudante rdf:ID="estBruno"/>
<CienciaComp rdf:ID="docMarceloFinger"/>
<Estudante rdf:ID="estMarieta"/>
<Estudante rdf:ID="estGabriela"/>
<Biologia rdf:ID="docJoseAmabis"/>
<Funcionario rdf:ID="funcJoao"/>
<Estudante rdf:ID="estJulio"/>
<Estudante rdf:ID="estCarolina"/>
<Filosofia rdf:ID="docNewtonCosta">
  <nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Newton</nome>
  <sobrenome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Costa</sobrenome>
</Filosofia>
<Disciplina rdf:ID="discIntrHistoria"/>
<Filosofia rdf:ID="docRenatoJanine">
  <nome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Renato</nome>
  <sobrenome rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >Janine</sobrenome>
</Filosofia>
<Disciplina rdf:ID="discAntiguidade"/>
<Estudante rdf:ID="estRenato"/>
<Disciplina rdf:ID="discEvolucao"/>
<Disciplina rdf:ID="discEmbriologia"/>
<Estudante rdf:ID="estBarbara"/>
<Disciplina rdf:ID="discFisiologia"/>
<Estudante rdf:ID="estLeandro"/>
<Disciplina rdf:ID="discEstDados"/>

```

```
<Estudante rdf:ID="estBernardo"/>
<Estudante rdf:ID="estDiana"/>
<Estudante rdf:ID="estAndre"/>
<Departamento rdf:ID="depHist"/>
<Estudante rdf:ID="estOnofre"/>
<Departamento rdf:ID="depFilo"/>
<CienciaComp rdf:ID="docRenataWassermann"/>
<Disciplina rdf:ID="discEstetica"/>
<Funcionario rdf:ID="funcPedro"/>
<Disciplina rdf:ID="discAnatomia"/>
<Biologia rdf:ID="docAlmeidaJr"/>
<Funcionario rdf:ID="funcNivaldo"/>
<Estudante rdf:ID="estDiego"/>
<Estudante rdf:ID="estMarcio"/>
<Funcionario rdf:ID="funcRonaldo"/>
<Medicina rdf:ID="docThelmaOkay"/>
</rdf:RDF>

<!-- Created with Protege (with OWL Plugin 3.4.5, Build 608) -->
```


Apêndice B

Ontologia O' - Universidade B

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns="http://www.owl-ontologies.com/universidadeB.owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://www.owl-ontologies.com/universidadeB.owl">
  <owl:Ontology rdf:about=""/>
  <rdfs:Class rdf:ID="Funcionario"/>
  <rdfs:Class rdf:ID="Aula"/>
  <rdfs:Class rdf:ID="Assistente">
    <rdfs:subClassOf rdf:resource="#Funcionario"/>
  </rdfs:Class>
  <rdfs:Class rdf:ID="ProfessorDE">
    <rdfs:subClassOf>
      <rdfs:Class rdf:ID="Professor"/>
    </rdfs:subClassOf>
  </rdfs:Class>
  <rdfs:Class rdf:ID="ProfessorAssociado">
    <rdfs:subClassOf>
      <rdfs:Class rdf:about="#Professor"/>
    </rdfs:subClassOf>
  </rdfs:Class>
  <rdfs:Class rdf:about="#Professor">
    <rdfs:subClassOf rdf:resource="#Funcionario"/>
  </rdfs:Class>
```

```

<rdfs:Class rdf:ID="AssistenteEnsino">
  <rdfs:subClassOf rdf:resource="#Assistente"/>
</rdfs:Class>
<rdfs:Class rdf:ID="Doutorando">
  <rdfs:subClassOf rdf:resource="#Funcionario"/>
</rdfs:Class>
<rdfs:Class rdf:ID="Sala"/>
<rdfs:Class rdf:ID="AssistentePesquisa">
  <rdfs:subClassOf rdf:resource="#Assistente"/>
</rdfs:Class>
<rdfs:Class rdf:ID="ProfessorAssistente">
  <rdfs:subClassOf rdf:resource="#Professor"/>
</rdfs:Class>
<rdfs:Class rdf:ID="ProfessorVisitante">
  <rdfs:subClassOf rdf:resource="#Professor"/>
</rdfs:Class>
<rdfs:Class rdf:ID="Referencia"/>
<owl:ObjectProperty rdf:ID="leciona">
  <rdfs:range rdf:resource="#Aula"/>
  <rdfs:domain rdf:resource="#Professor"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="publicacoes">
  <rdfs:range rdf:resource="#Referencia"/>
  <rdfs:domain rdf:resource="#Funcionario"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="departamento">
  <rdfs:domain rdf:resource="#Funcionario"/>
  <rdfs:range rdf:resource="#Sala"/>
</owl:ObjectProperty>
<owl:DatatypeProperty rdf:ID="nome">
  <rdfs:domain rdf:resource="#Funcionario"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="dataContr">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#date"/>
  <rdfs:domain rdf:resource="#Funcionario"/>
</owl:DatatypeProperty>
<Funcionario rdf:ID="funcRonaldo"/>
<Aula rdf:ID="discBacteriologia"/>
<Aula rdf:ID="discFisiologia"/>
<Funcionario rdf:ID="funcClaudio"/>
<ProfessorVisitante rdf:ID="docJoseAmabis"/>
<Doutorando rdf:ID="estJuliana"/>

```

```
<Sala rdf:ID="depBio"/>
<Funcionario rdf:ID="funcRejane"/>
<ProfessorDE rdf:ID="docNelsonWolosker"/>
<ProfessorDE rdf:ID="docLelianeBarros"/>
<Aula rdf:ID="discAbiogenese"/>
<Doutorando rdf:ID="estLuana"/>
<Aula rdf:ID="curLattes"/>
<Sala rdf:ID="depMed"/>
<Funcionario rdf:ID="funcAnaEliza"/>
<ProfessorAssociado rdf:ID="docReginaldoCoelho"/>
<ProfessorAssociado rdf:ID="docFabioAndrade"/>
<ProfessorDE rdf:ID="docRenatoJanine"/>
<Aula rdf:ID="discAntiguidade"/>
<Funcionario rdf:ID="funcJoao"/>
<Aula rdf:ID="discEpidemiologia"/>
<Funcionario rdf:ID="funcJose"/>
<Doutorando rdf:ID="estCarlos"/>
<Aula rdf:ID="curMetodoSocratico"/>
<Funcionario rdf:ID="funcJosefa"/>
<Sala rdf:ID="depHist"/>
<Aula rdf:ID="curCompNuvens"/>
<Funcionario rdf:ID="funcRomulo"/>
<Sala rdf:ID="depComp"/>
<Doutorando rdf:ID="estFlavio"/>
<Aula rdf:ID="curHistoriaEvolucao"/>
<Funcionario rdf:ID="funcRenato"/>
<ProfessorDE rdf:ID="docMarilenaChaui"/>
<Doutorando rdf:ID="estBarbara"/>
<ProfessorAssociado rdf:ID="docTeodoroAlmeida"/>
<Funcionario rdf:ID="funcNivea"/>
<ProfessorAssociado rdf:ID="docAntonioSilva"/>
<Doutorando rdf:ID="estOlivia"/>
<Aula rdf:ID="discFilMedieval"/>
<Funcionario rdf:ID="funcLais"/>
<Aula rdf:ID="discAnatomia"/>
<ProfessorDE rdf:ID="docNewtonCosta"/>
<Doutorando rdf:ID="estPaula"/>
<Aula rdf:ID="discHisSocialBrasil"/>
<ProfessorDE rdf:ID="docAlmeidaJr"/>
<Aula rdf:ID="discIA"/>
<Aula rdf:ID="discEstDados"/>
<ProfessorDE rdf:ID="docRonaldoGryschek"/>
<Aula rdf:ID="discProgramacao"/>
```

```
<Funcionario rdf:ID="funcMario"/>
<Doutorando rdf:ID="estDiego"/>
<Aula rdf:ID="discEstetica"/>
<Aula rdf:ID="discIntrHistoria"/>
<Funcionario rdf:ID="funcPedro"/>
<Sala rdf:ID="depFilo"/>
<ProfessorDE rdf:ID="docMarceloFinger"/>
<Funcionario rdf:ID="funcNivaldo"/>
<Funcionario rdf:ID="funcAntonio"/>
<Doutorando rdf:ID="estAmanda"/>
<Aula rdf:ID="discEvolucao"/>
<ProfessorAssociado rdf:ID="docEmanuelCarrilho"/>
<Doutorando rdf:ID="estMercia"/>
<Aula rdf:ID="discTeoriaConhecimento"/>
<ProfessorDE rdf:ID="docThelmaOkay"/>
<Aula rdf:ID="curDitaduraMilitar"/>
<ProfessorDE rdf:ID="docCezar"/>
<ProfessorDE rdf:ID="docRenataWassermann"/>
</rdf:RDF>

<!-- Created with Protege (with OWL Plugin 3.4.5, Build 608) -->
```


Apêndice C

Consultas - Alinhamento de Referência

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:    <http://www.owl-ontologies.com/universidadeA.owl#>
```

```
#Lista os indivíduos da classe Estudante
#Estudante >= Doutorando
```

```
select ?nome
where { ?nome rdf:type onto:Estudante}
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:    <http://www.owl-ontologies.com/universidadeA.owl#>
```

```
#Lista os indivíduos da classe Docente
#Docente =.9 ProfessorDE
```

```
select ?nome
where { ?nome rdf:type onto:Docente}
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:    <http://www.owl-ontologies.com/universidadeA.owl#>
```

```
#Lista os indivíduos da classe Departamento
#Departamento = Sala
```

```
select ?nome
where { ?nome rdf:type onto:Departamento}
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeA.owl#>
```

```
#Lista os indivíduos da classe Disciplina
#Disciplina = .7 Aula
```

```
select ?nome
where { ?nome rdf:type onto:Disciplina}
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeA.owl#>
prefix rdfs:     <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que são lecionados por alguém
#leciona = leciona (Parte 1)
```

```
select ?nome
where { onto:leciona rdfs:range ?obj.
       ?nome rdf:type ?obj}
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeA.owl#>
prefix rdfs:     <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que lecionam alguma coisa
#leciona = leciona (Parte 2)
```

```
select ?nome
where { onto:leciona rdfs:domain ?obj.
       ?nome rdf:type ?obj}
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeA.owl#>
prefix rdfs:     <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que são salas de alguém
#sala = departamento (Parte 1)
```

```
select ?nome
where { onto:sala rdfs:range ?obj.
       ?nome rdf:type ?obj}
```

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeA.owl#>
prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que têm salas
#sala = departamento (Parte 2)
```

```
select ?nome
where { onto:sala rdfs:domain ?obj.
        ?nome rdf:type ?obj}
```


Apêndice D

Consultas - Alinhamento A_1

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
#Lista os indivíduos da classe Estudante
#Estudante >= Doutorando
```

```
select ?nome
where { ?nome rdf:type onto:Doutorando}
```

```
#Proposto em A1: Estudante >= Doutorando
```

```
#Lista os indivíduos da classe Docente
#Docente =.9 ProfessorDE
```

```
#Não há correspondente para a classe Docente
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
#Lista os indivíduos da classe Departamento
#Departamento = Sala
```

```
select ?nome
where { ?nome rdf:type onto:departamento}
```

```
#Proposto em A1: Departamento = departamento
```

```
#Lista os indivíduos da classe Disciplina
#Disciplina = .7 Aula
```

```
#Não há correspondente para a classe Disciplina
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que são lecionados por alguém
#leciona = leciona (Parte 1)
```

```
select ?nome
where { onto:leciona rdfs:range ?obj.
        ?nome rdf:type ?obj}
```

```
#Proposto em A1: leciona = leciona
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que lecionam alguma coisa
#leciona = leciona (Parte 2)
```

```
select ?nome
where { onto:leciona rdfs:domain ?obj.
        ?nome rdf:type ?obj}
```

```
#Proposto em A1: leciona = leciona
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que são salas de alguém
#sala = departamento (Parte 1)
```

```
select ?nome
```

```
where { onto:Sala rdfs:range ?obj.  
        ?nome rdf:type ?obj}
```

```
#Proposto em A1: sala = Sala
```

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
prefix onto:  <http://www.owl-ontologies.com/universidadeB.owl#>  
prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que têm salas  
#sala = departamento (Parte 2)
```

```
select ?nome  
where { onto:Sala rdfs:domain ?obj.  
        ?nome rdf:type ?obj}
```

```
#Proposto em A1: sala = Sala
```


Apêndice E

Consultas - Alinhamento A_2

```
#Lista os indivíduos da classe Estudante
#Estudante >= Doutorando
```

```
#Não há correspondente para a classe Estudante
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
#Lista os indivíduos da classe Docente
#Docente =.9 ProfessorDE
```

```
select ?nome
where { ?nome rdf:type onto:Professor}
```

```
#Proposto em A2: Docente <= Professor
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
#Lista os indivíduos da classe Departamento
#Departamento = Sala
```

```
select ?nome
where { ?nome rdf:type onto:Sala}
```

```
#Proposto em A2: Departamento = Sala
```

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
#Lista os indivíduos da classe Disciplina
#Disciplina = .7 Aula
```

```
select ?nome
where { ?nome rdf:type onto:Aula}
```

```
#Proposto em A2: Disciplina = Aula
```

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeB.owl#>
prefix rdfs:   <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que são lecionados por alguém
#leciona = leciona (Parte 1)
```

```
select ?nome
where { onto:leciona rdfs:range ?obj.
        ?nome rdf:type ?obj}
```

```
#Proposto em A2: leciona = leciona
```

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeB.owl#>
prefix rdfs:   <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que lecionam alguma coisa
#leciona = leciona (Parte 2)
```

```
select ?nome
where { onto:leciona rdfs:domain ?obj.
        ?nome rdf:type ?obj}
```

```
#Proposto em A2: leciona = leciona
```

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que são salas de alguém
```

```
#sala = departamento (Parte 1)
```

```
select ?nome
```

```
where { onto:departamento rdfs:range ?obj.
```

```
        ?nome rdf:type ?obj}
```

```
#Proposto em A2: sala = departamento
```

```
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que têm salas
```

```
#sala = departamento (Parte 2)
```

```
select ?nome
```

```
where { onto:departamento rdfs:domain ?obj.
```

```
        ?nome rdf:type ?obj}
```

```
#Proposto em A2: sala = departamento
```


Apêndice F

Consultas - Alinhamento A_3

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
#Lista os indivíduos da classe Estudante
#Estudante >= Doutorando
```

```
select ?nome
where { ?nome rdf:type onto:Doutorando}
```

```
#Proposto em A3: Estudante = Doutorando
```

```
prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto: <http://www.owl-ontologies.com/universidadeB.owl#>
```

```
#Lista os indivíduos da classe Docente
#Docente =.9 ProfessorDE
```

```
select ?nome
where { ?nome rdf:type onto:ProfessorDE}
```

```
#Proposto em A3: Docente =.9 ProfessorDE
```

```
#Lista os indivíduos da classe Departamento
#Departamento = Sala
```

```
#Não correspondente para a classe Departamento
```

```
#Lista os indivíduos da classe Disciplina
#Disciplina = .7 Aula
```

```
#Não correspondente para a classe Disciplina
```

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeB.owl#>
prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que são lecionados por alguém
#leciona = leciona (Parte 1)
```

```
select ?nome
where { onto:leciona rdfs:range ?obj.
        ?nome rdf:type ?obj }
```

```
#Proposto em A3: leciona = leciona
```

```
prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix onto:  <http://www.owl-ontologies.com/universidadeB.owl#>
prefix rdfs:  <http://www.w3.org/2000/01/rdf-schema#>
```

```
#Lista os indivíduos que lecionam alguma coisa
#leciona = leciona (Parte 2)
```

```
select ?nome
where { onto:leciona rdfs:domain ?obj.
        ?nome rdf:type ?obj }
```

```
#Proposto em A3: leciona = leciona
```

```
#Lista os indivíduos que têm salas
#sala = departamento (Parte 1)
```

```
#Não há correspondente para a propriedade sala
```

```
#Lista os indivíduos que têm salas
#sala = departamento (Parte 2)
```

```
#Não há correspondente para a propriedade sala
```

Referências Bibliográficas

- Bispo Jr. e Wassermann(2011)** Esdras Bispo Jr. e Renata Wassermann. Uma nova abordagem de avaliação de alinhamentos de ontologias baseada em consultas. Em *Anais do XXXI Congresso da Sociedade Brasileira de Computação, VIII Encontro Nacional de Inteligência Artificial (ENIA 2011)*, páginas 926–937, Natal, Brasil. Citado na pág.
- Bouquet et al.(2003)** Paolo Bouquet, Luciano Serafini, e Stefano Zanobini. Semantic coordination: a new approach and an application. Em *Proceedings of the International Semantic Web Conference (ISWC)*, páginas 130–145. Citado na pág.
- Caracciolo et al.(2008)** Caterina Caracciolo, Jérôme Euzenat, Laura Hollink, Ryutaro Ichise, Antoine Isaac, Véronique Malaisé, Christian Meilicke, Juan Pane, Pavel Shvaiko, Heiner Stuckenschmidt, Ondrej Sváb-Zamazal, e Vojtech Svátek. Results of the ontology alignment evaluation initiative 2008. Em *Proceedings of the ISWC 2008 Workshop on Ontology Matching*, páginas 73–119. Citado na pág.
- Chauí(1999)** Marilena Chauí. *Convite à Filosofia*. Editora Atlas, 13^a edição. Citado na pág.
- Correndo et al.(2010)** Gianluca Correndo, Manuel Salvadores, Ian Millard, Hugh Glaser, e Nigel Shadbolt. SPARQL query rewriting for implementing data integration over linked data. Em *Proceedings of the 2010 EDBT/ICDT Workshops*, páginas 4:1–4:11. Citado na pág.
- Do et al.(2003)** Hong-Hai Do, Sergey Melnik, e Erhard Rahm. Comparison of schema matching evaluations. Em *Web, Web-Services, and Database Systems*, Lecture Notes in Computer Science, páginas 221–237. Springer Berlin / Heidelberg. Citado na pág.
- Doan et al.(2002)** AnHai Doan, Jayant Madhavan, Pedro Domingos, e Alon Halevy. Learning to map between ontologies on the semantic web. Em *WWW '02: Proceedings of the 11th International Conference on World Wide Web*, páginas 662–673, New York, USA. Citado na pág.
- Ehrig e Euzenat(2005)** Marc Ehrig e Jérôme Euzenat. Relaxed precision and recall for ontology matching. Em *Proceedings of the K-Cap Workshop on Integrating Ontologies*, páginas 25–32, Banff, CA. Citado na pág.
- Euzenat(2001)** Jérôme Euzenat. Towards a principled approach to semantic interoperability. Em *Proceedings of the IJCAI Workshop on Ontologies and Information Sharing*, páginas 19–25. Citado na pág.
- Euzenat e Shvaiko(2007)** Jérôme Euzenat e Pavel Shvaiko. *Ontology Matching*. Springer. Citado na pág.
- Euzenat e Valtchev(2004)** Jérôme Euzenat e Petko Valtchev. Similarity-based ontology alignment in owl-lite. Em *Proceedings of the 15th ECAI*, páginas 333–337, Valencia, ES. Citado na pág.
- Euzenat et al.(2007)** Jérôme Euzenat, Antoine Isaac, Christian Meilicke, Pavel Shvaiko, Heiner Stuckenschmidt, Ondrej Sváb, Vojtech Svátek, Willem Robert van Hage, e Mikalai Yatskevich. Results of the ontology alignment evaluation initiative 2007. Em *Proceedings of 20th International Joint Conference on Artificial Intelligence (IJCAI)*, páginas 96–132. Citado na pág.

- Euzenat et al.(2008)** Jérôme Euzenat, Axel Polleres, e François Scharffe. Processing ontology alignments with SPARQL. Em *Proceedings of the International Conference on Complex, Intelligent and Software Intensive Systems*, páginas 913–917. Citado na pág.
- Euzenat et al.(2009)** Jérôme Euzenat, Alfio Ferrara, Laura Hollink, Antoine Isaac, Cliff Joslyn, Véronique Malaisé, Christian Meilicke Andriy Nikolov Juan Pane, Marta Sabou, François Scharffe, Pavel Shvaiko, Vassilis Spiliopoulos, Heiner Stuckenschmidt, Ondrej Sváb-Zamazal, Vojtech Svátek, Cássia Trojahn, George Vouros, e Shenghui Wang. Results of the ontology alignment evaluation initiative 2009. Em *Proceedings of the ISWC 2009 Workshop on Ontology Matching*, páginas 73–125. Citado na pág.
- Euzenat et al.(2010)** Jérôme Euzenat, Alfio Ferrara, Christian Meilicke, Andriy Nikolov, Juan Pane, François Scharffe, Pavel Shvaiko, Heiner Stuckenschmidt, Ondrej Sváb-Zamazal, Vojtech Svátek, e Cássia Trojahn. Results of the ontology alignment evaluation initiative 2010. Em *Proceedings of the ISWC 2010 Workshop on Ontology Matching*, páginas 85–117. Citado na pág.
- Gal e Shvaiko(2009)** Avigdor Gal e Pavel Shvaiko. Advances in ontology matching. Em *Advances in Web Semantics I*, volume 4891 of *Lecture Notes in Computer Science*, páginas 176–198. Springer Berlin, Heidelberg. Citado na pág.
- Gruber(1993)** Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220. Citado na pág.
- Guarino(1998)** Nicola Guarino. Formal ontology and information systems. Em *Proceedings of the First International Conference on Formal Ontologies in Information Systems*, páginas 3–15, Amsterdam, Netherlands. IOS Press. Citado na pág.
- Jean-Mary et al.(2009)** Yves R. Jean-Mary, E. Patrick Shironoshita, e Mansur R. Kabuka. AS-MOV: Results for OAEI 2009. Em *Proceedings of the ISWC 2009 Workshop on Ontology Matching*, páginas 152–159. Citado na pág.
- Kalfoglou e Schorlemmer(2003)** Yannis Kalfoglou e Marco Schorlemmer. Ontology mapping: The state of the art. *The Knowledge Engineering Review*, 18(1):1–31. Citado na pág.
- Kang e Naughton(2003)** Jaewoo Kang e Jeffrey F. Naughton. On schema matching with opaque column names and data values. Em *SIGMOD '03: Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, páginas 205–216, New York, NY, USA. ACM. Citado na pág.
- Kashyap e Sheth(1996)** Vipul Kashyap e Amit Sheth. Semantic heterogeneity in global information systems: The role of metadata, context and ontologies. Em *Cooperative Information Systems: current trends and directions*, páginas 139–178. Academic Press Ltd. Citado na pág.
- Maedche(2001)** Alexander Maedche. Ontology learning for the semantic web. *IEEE Intelligent Systems*, 16(2):72–79. Citado na pág.
- Maedche(2002)** Alexander Maedche. *Ontology learning for the Semantic Web*. Kluwer Academic Publishers, Massachusetts. Citado na pág.
- Nardin(2009)** Luís Gustavo Nardin. *Uma Arquitetura de Apoio à Interoperabilidade de Modelos de Reputação de Agentes*. Dissertação de Mestrado, Escola Politécnica, Universidade de São Paulo, Brasil. Citado na pág.
- Noy e Klein(2004)** Natalya Noy e Michel Klein. Ontology evolution: Not the same as schema evolution. Em *Knowledge and Information Systems*, volume 6, páginas 428–440. Springer Berlin, Heidelberg. Citado na pág.

- Noy e McGuinness(2001)** Natalya F. Noy e Deborah L. McGuinness. Ontology development 101: A guide to creating your first ontology. Relatório Técnico Technical Report SMI-2001-0880, Stanford Medical Informatics. Citado na pág.
- Noy e Musen(2000)** Natalya F. Noy e Mark A. Musen. Prompt: Algorithm and tool for automated ontology merging and alignment. Em *Proceedings of the National Conference on Artificial Intelligence (AAAI)*. Citado na pág.
- Noy e Musen(2001)** Natalya F. Noy e Mark A. Musen. Anchor-prompt: Using non-local context for semantic matching. Em *Proceedings of IJCAI Workshop on Ontologies and Information Sharing*, páginas 63–70. Citado na pág.
- Rahm e Bernstein(2001)** Erhard Rahm e Philip A. Bernstein. A survey of approaches to automatic schema matching. *The VLDB Journal*, 10:334–350. Citado na pág.
- Sheth e Kashyap(1992)** Amit P. Sheth e Vipul Kashyap. So far (schematically), yet so near (semantically). *Proceedings of the IFIP WG 2.6 Database Semantics Conference on Interoperable Database Systems (DS-5)*, A-25:283–312. Citado na pág.
- Shvaiko e Euzenat(2005)** Pavel Shvaiko e Jérôme Euzenat. A survey of schema-based matching approaches. Em *Journal on Data Semantics IV*, volume 3730 of *Lecture Notes in Computer Science*, páginas 146–171. Springer Berlin / Heidelberg. Citado na pág.
- Shvaiko e Euzenat(2008)** Pavel Shvaiko e Jérôme Euzenat. Ten challenges for ontology matching. Em *Proceedings of the 7th International Conference on Ontologies, Databases, and Applications of Semantics (ODBASE)*. Citado na pág.
- van Eijk et al.(2001)** Rogier M. van Eijk, Frank S. de Boer, Wiebe van der Hoek, e John-Jules Ch. Meyer. On dynamically generated ontology translators in agent communication. *International Journal of Intelligent System*, 16:587–607. Citado na pág.
- van Rijsbergen(1975)** C. J. van Rijsbergen. *Information retrieval*. Butterworths, London, UK. Acesso em 26 de agosto de 2010. Disponível em <http://www.dcs.gla.ac.uk/Keith/Preface.html>. Citado na pág.
- Velegrakis et al.(2005)** Yannis Velegrakis, Renée J. Miller, e John Mylopoulos. Representing and querying data transformations. Em *Proceedings of the International Conference on Data Engineering (ICDE)*, páginas 81–92. Citado na pág.
- Wang e Xu(2009)** Peng Wang e Baowen Xu. Lily: Ontology alignment results for OAEI 2009. Em *Proceedings of the ISWC 2009 Workshop on Ontology Matching*, páginas 186–192. Citado na pág.
- Wang et al.(2010)** Zhichun Wang, Xiao Zhang, Lei Hou, Yue Zhao, Juanzi Li, Yu Qi, e Jie Tang. RiMOM results for OAEI 2010. Em *Proceedings of the ISWC 2010 Workshop on Ontology Matching*, páginas 194–201. Citado na pág.