

Refatoração de Alinhamentos de Ontologias Inconsistentes

Miriam Oliveira dos santos

Programa de Pós-Graduação em Informática (PPGI) - Universidade Federal do Estado do Rio de Janeiro (UNIRIO) - Rio de Janeiro, RJ - Brasil

miriam.santos@uniriotec.br

Abstract. *The efficient communication between information systems, of different origins, requires a uniformity that can not be achieved without there being a conformity between its concepts; This can be achieved with the use of Ontologies, through a process called Alignment of Ontologies. It happens that this process may fail, causing what we call Alignments inconsistent. In order to solve this problem, there are several approaches that are concerned both to identify problems in the alignments generated and to deal with them, redefining the correspondences between these networks and thus enabling the communication between systems.*

Resumo. *A comunicação eficiente entre sistemas de informação, de origens distintas, requer uma uniformidade que não é possível alcançar sem que haja uma conformidade entre seus conceitos; isto pode ser alcançado com o uso de Ontologias, através de um processo intitulado de Alinhamento de Ontologias. Ocorre que este processo pode apresentar falhas, ocasionando o que chamamos de Alinhamentos Incoerentes. Para resolução deste problema existem diversas abordagens que se preocupam tanto em identificar problemas nos alinhamentos gerados como em tratá-los, redefinindo as correspondências entre estas redes e assim possibilitando a comunicação entre sistemas.*

1. Introdução

Uma ontologia tipicamente fornece um vocabulário descrevendo um domínio de interesse e uma especificação do significado destes termos [Euzenat e Shvaiko 2007]. Uma Ontologia também é descrita como um meio formal de modelar a estrutura de um sistema, ou seja, as entidades relevantes, conceitos e relações que delas emergem que sejam úteis para atingir os objetivos do sistema. A estrutura principal de uma ontologia, portanto, é uma hierarquia de conceitos: uma taxonomia. A essa hierarquia são evidenciadas relações e associações as quais, em um paradigma lógico, são modeladas por meio de predicados [Guarino et al. 2009].

A partir do momento que nos deparamos com o uso de ontologias em dois ou mais sistemas de informação, mesmo que para um mesmo domínio, fatalmente iremos encontrar conceitos especificados de formas distintas, próprias de cada modelo de conhecimento; consequentemente, para que ocorra a comunicação entre estes sistemas faz-se necessário realizar correspondências entre estes conceitos; a este processo chamamos de *Ontology Matching*.

O processo de *Ontology Matching* visa encontrar correspondências entre entidades semanticamente relacionadas de diferentes ontologias, a partir deste processo

são identificadas, além de relações de equivalência, relações como: consequência, subjunção, ou disjunção entre as entidades de uma ontologia [Euzenat e Shvaiko 2007]. Algumas vezes é possível encontrar alinhamentos que se mostram inconsistentes, os alinhamentos assim gerados têm efeitos negativos em diferentes tipos de aplicações [Meilicke 2011]. Esta anomalia é chamada de *Alignment Incoherence*, que ocorre por exemplo, quando há correspondência entre sub-entidades de classes disjuntas, havendo necessidade de executar uma ação para restaurar sua consistência.

O Trabalho a ser desenvolvido tem como base os conceitos acima elencados, tendo como proposta uma abordagem para Refatoração de Alinhamentos inconsistentes.

2. Apresentação do problema

O processo de alinhamento de ontologias não é 100% preciso, necessitando que haja um processo posterior para detectar inconsistências e corrigir os alinhamentos gerados, normalmente são utilizadas nesta etapa técnicas semânticas para detecção de inconsistências; e para restaurar a consistência: técnicas manuais, semi-automáticas ou automáticas, que envolvem desde o uso de especialistas a construção de heurísticas.

Trabalhos relacionados tem como ponto em comum o estudo de soluções para reparação de alinhamentos de ontologias inconsistentes, atuando tanto na detecção do problema como na resolução dos mesmos.

Em [Jean-Mary et al 2009], é proposto antipadrões de correspondências, que indicam categorias de inconsistências que devem ser corrigidas. Um antipadrão é uma descrição de uma situação problemática que é comumente aplicada sobre um cenário e que gera, definitivamente, consequências negativas. Quando devidamente documentado, um antipadrão descreve a situação problemática de forma geral, as causas primárias que levam a essa situação problemática, sintomas que descrevem como reconhecer o antipadrão, consequências da utilização desta situação problemática e uma solução refatorada que pode mudar o antipadrão em uma solução mais adequada [Silva 2014]. Uma outra forma de detectar inconsistências, se dá através do uso de lógica de descrição; Em [Meilicke 2011], é proposto uma solução para automatização do processo de correção de alinhamentos inconsistentes, tendo sido desenvolvido uma ferramenta para esta finalidade, intitulada ALCOMO (*Applying Logical Constraints on Matching Ontologies*).

A partir do momento que a inconsistência é detectada, devemos restabelecer a consistência, procedendo a análise dos alinhamentos em conflito. Existem dois tipos de estratégias que norteiam as abordagens existentes na resolução do problema, são elas: Estratégia global de reparação e Estratégia local de reparação. Na estratégia global, é realizado uma análise completa de todos os mapeamentos que encontram-se em conflito, para só então proceder a resolução dos mesmos [Ferrara et al. 2008], [Jiménez-Ruiz et al. 2013], [Santos et al. 2015]. Nas estratégias locais, a reparação ocorre a partir da análise de cada par de correspondências que resultem em inconsistência, sem analisar como um todo os problemas resultantes de todo o processo de mapeamento. Diferentes abordagens foram utilizadas para uso de estratégias locais, como as citadas nos trabalhos de [Jiménez-Ruiz et al. 2013], [Euzenat 2014],[Santos et al 2015].

Nos trabalhos que consideram a estratégia global de reparação, busca-se estabelecer uma métrica para estabelecer a ordem com que os conflitos são analisados; este é um fator de extrema importância. Na maior parte dos trabalhos, esta métrica envolve valor de confiança, Valores de confiança são fornecidos como resultado do processo de alinhamento de ontologias e determinam a força entre os alinhamentos gerados. Algumas abordagens de reparação de alinhamento de ontologias, utilizam-se do valor de confiança como base para decisão de qual alinhamento deve ser removido ou utiliza-se deste valor como base de sua estratégia [Ferrara et al. 2008], [Jean-Mary et al 2009], [Meilicke e Stuckenschmidt 2009], [Atencia et al. 2012], [Jiménez-Ruiz et al. 2013] .

3. Proposta de Solução

Através da análise de trabalhos relacionados, pode-se perceber que muitos problemas ainda encontram-se em aberto, seja na possível exclusão de mapeamentos corretos; seja nas abordagens baseadas em especialistas em que se deseja diminuir a quantidade de questões a serem apresentadas ou na própria forma de apresentação apoiada na demonstração de implicações; seja no tratamento de ontologias extensas que também ocasionam problemas de performance; seja no tratamento de ontologias que possuem poucos axiomas de disjunção; Seja no uso de raciocínios incompletos para tratar de ontologias mais extensas e manter a escalabilidade; seja no uso de padrões e antipadrões que também são considerados incompletos para detecção de problemas, ou seja há uma variedade de questões que ainda podem ser aperfeiçoadas em problemas apresentados tanto em abordagens locais ou globais.

Existem três tipos de ações que podem ser utilizadas para o restabelecimento da coerência a serem aplicadas nas correspondências em conflito: remoção da correspondência em conflito, enfraquecimento ou adição de relacionamentos [Euzenat, 2014].

Em nosso trabalho, pretende-se utilizar uma estratégia global de resolução de conflitos, em que será analisado todos os conflitos existentes, e será estabelecido uma métrica que influenciará tanto na ordem com que os conflitos serão analisados, quanto na ação a ser executada. Nosso trabalho prevê uma abordagem de resolução de conflitos executada de forma automática, devemos resolver todos os conflitos possíveis, mas ao mesmo tempo preservar o maior número de correspondências possíveis. Deve ser observado que dependendo da ordem com que estes conflitos são resolvidos, automaticamente outros conflitos podem deixar de existir, portanto é fundamental estabelecer uma métrica para ordem com que estes conflitos sejam resolvidos, ou constatar qual(is) os melhores caminhos a serem analisados e compará-los para estabelecer em qual ordem deve ser resolvido.

Objetivos de pesquisa: Resolver todos os conflitos possíveis, tratando mapeamentos incorretos; Tornar o processo de reparação de alinhamento de ontologias mais eficiente, não permitindo remoção de mapeamentos corretos; Aumentar as métricas de precisão, cobertura e medida-F.

As métricas utilizadas observarão tanto valor de confiança, quanto análise de outros valores que verificarão a quantidade de conflitos em que uma correspondência está envolvida ou ainda uma nova análise da força deste alinhamento como por exemplo o grau de conectividade entre as entidades envolvidas no mapeamento.

Inicialmente pensa-se em utilizar um processo de inferência fuzzy, para estabelecer uma nova métrica que permita aferir uma ordem com que as correspondências sejam analisadas ou qual a solução a ser adotada para resolução do conflito. O processo de inferência fuzzy é subdividido em quatro etapas (1) fuzzyficação, (2) avaliação das regras fuzzy, (3) agregação das regras fuzzy e (4) defuzzyficação. Estas etapas são executadas durante o processamento, porém as configurações para que estas etapas possam ocorrer, são executadas no pré-processamento desta solução. Passo1: Definir tipos de conflitos que serão tratados, por exemplo, contradição de disjunção-subjunção; Passo2: Parametrização de regras por domínio, por tamanho ou Regras por tamanho, por domínio ou; Regras Gerais independentes do domínio e tamanho. Passo3: Configuração para o Processo de Inferência Fuzzy. O processo de inferência fuzzy, será determinada por parametrizações realizadas pelo especialista, as decisões tomadas pelo especialista na parametrização, será responsável pela resolução de conflitos, que em outras abordagens eram resolvidas com o apoio do usuário apenas durante o processo.

Após análise dos trabalhos correlatos, chegou-se a conclusão que poderíamos ter inicialmente as seguintes variáveis: Grau de Confiança, Grau de Conectividade ou localidade, Grau de Conflitos envolvidos ou Grau de Impacto em regras; a idéia é realizar a combinação dos resultados destas variáveis para indicação de ações a serem realizadas ou para obtenção de um novo grau de confiança que não considera apenas o grau de confiança obtido pelos sistemas de mapeamento, mas outros valores igualmente importantes, que reforcem a força do relacionamento em caso de conflitos.

A Tabela 1, exemplifica regras que podem ser configuradas, a partir das variáveis e termos linguísticos a serem estabelecidos.

Tabela 1 – Regras Fuzzy

Regras
(Se Grau de Confiança é alto) AND (Conectividade é alta) AND (Grau de Conflitos é baixo) ENTÃO (Grau de Recomendação é muito baixo)
(Se Grau de Confiança é baixo) AND (Conectividade é baixa) AND (Grau de Conflitos é alto) ENTÃO (Grau de Recomendação é alto)

4. Projeto de avaliação da solução

O Processo de avaliação da solução, contará com uma base de dados amplamente utilizada pela comunidade acadêmica internacional, que estuda o tema Alinhamento de Ontologias e conceitos relacionados. Nesta base são disponibilizadas ontologias e alinhamentos de referências, no qual dada duas ontologias descreve quais correspondências encontram-se corretas. Esta base de dados é disponibilizada pela OAEI, *Ontology Alignment Evaluation Initiative*.

Os resultados são comparados através de três medidas: Precisão, Cobertura e Medida-F. que medem respectivamente, Dentre as correspondências encontradas, a proporção de correspondências corretas; A proporção de correspondências corretas encontradas dentre todas as possíveis; A harmonização entre precisão e cobertura [Meilicke 2011].

5. Conclusão

Neste trabalho foram apresentados os conceitos relacionados a reparação de alinhamento de ontologias e os desafios que permeiam a área, através da apresentação de trabalhos relacionados. Esta Pesquisa tem como objetivo eliminar ou minimizar a existência de alinhamentos inconsistentes e apresenta um estudo inicial cuja proposta de solução se dá através da refatoração do alinhamento através do uso de técnicas de inferência, que atuará como sistema de apoio de decisão, automatizando o processo e eliminando a participação do usuário durante o processo.

Referências

- Atencia, M., Borgida, A., Euzenat, J., Ghidini, C., & Serafini, L. (2012). "A formal semantics for weighted ontology mappings". In: International Semantic Web Conference, p. 17-33, Springer Berlin Heidelberg.
- Meilicke, C. (2011). "Alignment Incoherence in Ontology Matching". Ph.D. thesis, Universitat Mannheim.
- Meilicke, C., Stuckenschmidt, H. (2009) "An Efficient Method for Computing Alignment Diagnoses", In: Web Reasoning and Rule Systems, vol. 5837, p. 182-196.
- Jiménez-Ruiz, E., Meilicke, C., Grau, B. C., Horrocks, I. (2013) "Evaluating Mapping Repair Systems with Large Biomedical Ontologies". Description Logics, vol.13, p.246-257.
- Euzenat, J. (2014). "First experiments in cultural alignment repair (extended version)". In: European Semantic Web Conference Springer International Publishing, p. 115-130.
- Silva, Anselmo (2014). "Identificação e Representação de antipadrões de correspondências em ontologias", Dissertação de Mestrado, UNI-Rio.
- Santos, E., Faria, D., Pesquita, C., Couto, F. M. (2015). "Ontology alignment repair through modularization and confidence-based heuristics". PloS one, v.10, n.12.
- Euzenat, J., & Shvaiko, P. (2007). "Ontology matching". v.18, Heidelberg: Springer.
- Jean-Mary, Y. R., Shironoshita, E. P., Kabuka, M. R. (2009). "Ontology matching with semantic verification". Web Semantics: Science, Services and Agents on the World Wide Web, vol.7, n.3, p.235-251.
- Ferrara, A., Lorusso, D., Stamou, G., Stoilos, G., Tzouvaras, V., & Venetis, T. (2008). "Resolution of conflicts among ontology mappings: a fuzzy approach". In: Proceedings of the 3rd International Conference on Ontology Matching, v.431, p. 13-24.
- Guarino, N.; Oberle, D.; Staab, S. 2009. What is an ontology? In: Staab, S.; Studer, R. Handbook on Ontologies. 2. ed. Berlin: Springer.