WordNet-based Semantic Similarity Measures for Process Model Matching

Khurram Shahzad, Ifrah Pervaz, Rao Muhammad Adeel Nawab[*]

Punjab University College of Information Technology, University of the Punjab, Lahore.
[*]Department of Computer Science, COMSATS Institute of Information Technology, Lahore.
khurram|ifrah@pucit.edu.pk, adeelnawab@ciitlahore.edu.pk

Abstract. Process Model Matching (PMM) refers to the automatic identification of corresponding activities between a pair of process models. Due to the wider applicability of PMM techniques several semantic matching techniques have been proposed. However, these techniques focus on utilizing few word-to-word (word-level) similarity measures, without giving due consideration to activity-level aggregation methods. The inadequate attention to the choice of activity-level methods limit the effectiveness of the matching techniques. Furthermore, there are some WordNet-based semantic similarity measures that have shown promising results for various text matching tasks. However, the effectiveness of these measures has never been evaluated in the context of PMM. To that end, in this paper we have used five word-level semantic similarity measures and three sentence-level aggregation methods to experimentally evaluate the effectiveness of their 15 combinations for PMM. The experiments are performed on the three widely used PMMC'15 datasets. From the results we conclude that, a) Jiang similarity is more suitable than the mostly used Lin similarity, and b) QAP is the most suitable sentence-level aggregation method.

Keywords: Business Process Models, Process Model Matching, Semantic Similarity, WordNet-based similarity measures.

1 Introduction

Business process models are the conceptual models that explicitly represent the business operations of an enterprise. These models are widely accepted as a useful resource for a variety of purposes ranging from representing requirements for software development to configuring ERP systems. Process Model Matching (PMM) refers to identifying the activities between two process models that represent similar or identical functionality [1]. A pair of activities that represent similar or identical functionality is called a corresponding pair and the involved activities are called corresponding activities [2]. Figure 1 shows the example process models of two universities, University A and University B, and correspondences between their activities. In the figure, each correspondence between a pair of activities is marked by a shaded area.

An accurate identification of corresponding activities is of higher significance for the BPM community due to its widespread application areas, such as identifying clones of process models, searching process models and harmonizing process models [3]. To that end, a plethora of automatic techniques have been proposed [4]. Despite the existence of several matching techniques, the need for enhancing the accuracy of matching techniques has been widely pronounced during the recent years [4, 5]. For instance, a comprehensive survey of the state-of-the-art has made imperative revelations about process model matching techniques [5]. The two notable ones are, 1) 21 out of 35 techniques use the most basic syntactic measures, and 2) Lin similarity is the most prominent semantic similarity measure.

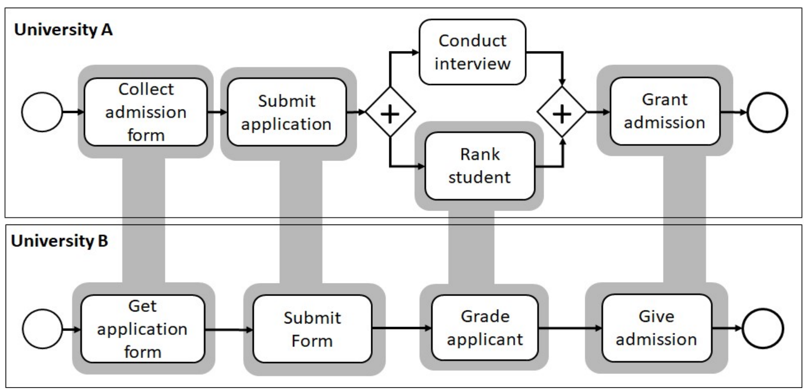

Fig. 1. Illustration of process model matching

In this study we contend that there are several word-level semantic similarity measures that have shown promising results for various text processing tasks [6, 7]. However, an empirical assessment of these competing measures has never been conducted in the context of process model matching. Consequently, a well-grounded recommendation about the choice of a semantic similarity measure is non-existent. Furthermore, the presently used similarity measures merely focus on the word-to-word semantic similarity, without paying adequate attention to the aggregation of word-level similarity scores to an activity-level similarity score. This arbitrary selection of sentence-level aggregation method, such as average score, may impede the effectiveness of the matching techniques. To that end, in this study we evaluate the effectiveness of five WordNet-based word-to-word semantic similarity measures and three sentence-level methods, which extend word-level semantic similarity scores to an activity-level similarity score. The effectiveness of all (fifteen) combinations of five word-level semantic similarity measures and three sentence-level methods are evaluated using the three PMMC'15 datasets.

The rest of the paper is organized as follows: Section 2 provides an overview of the word-level and sentence-level semantic similarity measures. Section 3 and 4 presents the experimental setup and results of the experiments, respectively. Section 5 provides an overview of the related work. Finally, Section 6 concludes the paper.

2 WordNet-based Semantic Similarity Measures

The semantic similarity methods that we have applied for identifying corresponding activities between process models are based on WordNet. WordNet is widely acknowledged as a valuable source to find semantic similarity between two words as it organizes words based on lexical relations and then defines semantic relations between those lexically related synsets. The lexical relationships are categorized into two subcategories: synsets, and antonyms, whereas semantic relations are categorized into five sub categories: hyponyms, meronyms, co-ordinate terms, entailment of a verb, and troponym of a verb. Synsets are related with other synsets to form a hierarchical structure of conceptual relations. In the WordNet version 2.0, there are nine noun hierarchies that include 80,000 concepts and 554 verb hierarchies that are made up of 13,500 concepts. All the concepts are linked to a unique root, called *entity*.

2.1 Word-level Semantic Similarity Measures

We have selected five well established and widely used word-level semantic similarity measures to compute the degree of semantic similarity between activity pairs. These are: Resnik similarity [7], Jiang similarity [8], Leacock similarity [6], Lin similarity [9], and Wu similarity [10]. The methods have been previously used for lexical and textual semantic relatedness [11, 13], word sense disambiguation [12], gene and sequence matching [14], generating sentences from pictures [15], paraphrasing [16], sentiment analysis [17] and topic modeling [18, 19]. A brief overview of these measures is as follows:

Resnik Similarity. Resnik similarity relies on *is-a* relationship in the WordNet taxonomy, where each node represents a unique WordNet synset or concept. According to this measure two nodes are considered more similar if they share more information. This shared information is specified by Information Content (IC) of the nodes that subsumes these nodes in a taxonomy. Formally, IC is calculated as follows:

$$IC = -\log P(C)$$

Let C1 and C2 be the two concept nodes in WordNet taxonomy and concept node C is the lowest common subsumer node of nodes C1 and C2. Furthermore, let P(C) is the probability of occurrence of longest common subsumer node C and probability of node C is simply found by normalizing occurrences of concepts with total number of nouns in the taxonomy.

$$P(C) = \frac{f(c)}{N} \; and \; f(c) = \sum_{n \in W(c)} count(n)$$

Where, W(C) is the set of concepts in which word w occurs and each occurrence of a word is considered as occurrence of all concepts containing that word. The Resnik similarity is referred as maximal IC over all concepts to which both words belong. Formally, it is defined as follows:

$$sim_{res}(C1, C2) = IC\,(LCS\,(C1, C2))$$

Where, LCS is the lowest common subsumer of concept nodes C1 and C2 defined as the common parent of these with minimum node distance.

Jiang Similarity. This method uses corpus statistical information i.e. Information Content (IC) and nodes path in *is-a* taxonomy for computing similarity, where, IC is measure of occurrence of the concept in the corpus. Given a word pair C1 and C2, this measure computes similarity between the words by using following equation:

$$sim_{jnc}(C1, C2) = \frac{1}{IC(C1)+IC(C2)-2*IC(LCS)}$$

Where, IC stands for information content and LCS is the Lowest Common Subsumer of concepts C1 and C2 defined as the common parent of these with minimum node distance.

Leacock Similarity. This similarity measure is based on a node based approach using *is-a* taxonomy in the WordNet. When considering the WordNet taxonomy, each node represents a unique concept (or synset) in the taxonomy. Subsequently, the degree of similarity between a word pair is computed by calculating the shortest path between two concepts (represented as nodes), and dividing it by twice the maximum depth of the graph. Formally, it is represented as follows:

$$sim_{lch}(C1, C2) = -log\frac{shortest\ path\ lenghth}{2*Dept}$$

In the equation, C1 and C2 are the two concepts represented by nodes, shortest path length is the minimum path length from node C1 to node C2 by using node counting, and depth is the number of maximum nodes from root node to a leaf node.

Lin Similarity. According to this measure, similarity between two concepts is expressed as the similarity between generic terms belonging to these concept classes, rather than measuring similarity between all terms. For instance, word 'design' belongs to the concept class 'blueprint', and the word 'construct' belongs to the concept class named 'concept'. According to Lin, the similarity between these words should be the same as the similarity between two synsets 'blueprint' and 'concept' to which these words belong. Formally, if there is a word $x_1 \in C1$ and a word $x_2 \in C2$ the information shared by two words can be expressed by C, the most specific class that subsumes both. The similarity is then computed as:

$$sim_{lin}(x_1, x_2) = \frac{2*\log P(C)}{\log P(C1)+\log P(C2)}$$

Where C is the most specific class that subsumes concepts C1 and C2 is the common parent of the two concepts with a minimum node distance. log P (C), log P (C1) and log P (C2) are log likelihood of the occurrence of concepts C, C1 and C2.

Wu Similarity. This similarity measure relies on the depth of both concept nodes, and depth of lowest common subsumer. The similarity between two concepts is then computed by using the following Equation.

$$sim_{wup}(C1, C2) = \frac{2*depth(LCS)}{depth(C1)+depth(C2)}$$

Where LCS is the lowest common subsumer of concepts C1 and C2 defined as the common parent of C1 and C2 with minimum node distance. Depth (C1) represents the number of nodes from C1 to LCS node C, depth (C2) represents the number of nodes from C2 to LCS node C and depth (LCS) represents the number of nodes from LCS node C to root node.

2.2 Sentence-Level Similarity Methods

The preceding section presented various WordNet-based semantic similarity measures for computing word-level similarity. These measures compute similarity between a pair of words, however, PMM refers to computing similarity between activity pairs. Therefore, there is a need to combine the word-level measures with sentence-level methods, where each label is considered as a sentence. For this purpose, we applied three methods which extend word-level measures to sentence-level methods. These methods are Greedy Pairing [21], Optimal Matching [22], and Quadratic Assignment Problem (QAP) [23]. A brief overview of each method is described below.

Greedy Pairing. Using this method, at first both sentences are tokenized. After that, word-level semantic similarity method is used to search the maximum semantic similarity of each token in the first sentence with all the tokens in the second sentence. These maximum similarities are weighted using Inverse Document Frequency (IDF) scores. Maximum similarity scores of all the tokens in the first sentence are computed, summed up and resulting score is normalized with maximum sentence length. The same steps are repeated to find the maximum mappings for each token of the second sentence with the first sentence. The final similarity score between sentence pair is obtained by computing the average scores obtained using sentence one and two.

Optimal Matching. This method is based on combinatorial matching problem, where for a given weighted bipartite graph the problem is to find maximum matching of graph. Bipartite graph is the graph whose nodes can be divided into two disjoint sets. Using this approach, the two sentences S1 and S2 are considered part of a weighted bipartite graph $G = S1 \cup S2$, where words in sentences are represented as nodes of graph and the weight of edges between these nodes corresponds to similarity score between the respected nodes. The task is to select those node pairs in matching M such that overall sum of all selected node pair is maximum.

Quadratic Assignment Problem (QAP). This approach finds an optimal assignment of word in first sentence to second sentence using word-level similarity measure and at the same time maximizes the similarity between syntactic dependencies of words pair. The Koopmans-Beckmann formulation of the QAP problem is used. The goal is to maximize the objective function QAP (F, D, B) where F and D captures syntactic dependencies between words in two sentences respectively and B captures the word-to-word similarity across the two sentences.

3 Experimental Setup

For the experiments we have used three well established datasets, developed by experts and used in Process Model Matching Contest 2015 (PMMC'15). Since the competition, the datasets are widely used for the evaluation of process model matching techniques [24]. The datasets are named as, University Admissions (UA), Birth Registration (BR), and Asset Management (AM) datasets. Below, we present a brief overview of the three datasets. The UA dataset is composed of 9 process models about admission to nine German universities and 36 pairs of process models. In addition to that, the dataset includes gold standard correspondences between equivalent activities. The specifications of the three datasets are given below in Table 1.

The BR dataset includes 9 process models, 36 pairs of these nine models and gold standard correspondences. The models represent birth registration process of different countries: Germany, Russia, South Africa, and the Netherlands. The collection includes both 1:1 and 1: n correspondences. The AM dataset consists of 36 process model pairs selected from 72 process models of the SAP reference model collection [24]. The selected process models cover different aspects from the area of asset management.

Table 1. Specifications of the collected PMMC'15 datasets.

	UA dataset	BR dataset	AM dataset
No of Activities (Min)	12	9	1
No of Activities (Max)	45	25	43
No of Activities (Avg)	24.2	17.9	18.6
No of 1:1 Correspondences	268	156	140
No of 1:n Correspondences	360	427	82

4 Results and Analysis

This section presents the analysis of the results, which are obtained by applying five combinations of five word-level semantic similarity measures and 3 sentence-level methods.

The main goal of experiments is to classify each activity pair as 'equivalent' or 'non-equivalent'. Since, the semantic similarity methods used in this study return a numeric score between 0 and 1, we have converted these numeric scores into binary 0 (non-equivalent) and 1 (equivalent) at nine different thresholds from 0.1 to 0.9 with a gap of 0.1. However, due to space limitations we have used a cut-off threshold 0.7 because multiple matching systems participating in the latest episode of the Process Model Matching Contest 2015 achieved promising results at this threshold. This threshold value represents that each activity pair for which the similarity score is greater than or equal to 0.7 is marked as equivalent (or 1) and non-equivalent (or 0) otherwise. The F_1 scores at 0.7 threshold are presented in Table 2 and the remaining results are made available for download. For each dataset, the word-level measure that obtained the highest F_1 score for a sentence-level method is highlighted in *bold*. Therefore, each sentence-level method has at least one bold value. Furthermore, we have *underlined* the word-level measure that obtained the highest F_1 score for a dataset, independent of any sentence-level aggregation method.

Table 2. Results of all the techniques for PMMC'15 datasets.

Sentence Level	Word-level	UA Dataset	BR Dataset	AM Dataset
Greedy Pair	Resnik	**0.516**	0.532	**<u>0.464</u>**
	Jiang	**0.516**	**<u>0.534</u>**	**<u>0.464</u>**
	Leacock	0.487	0.509	0.453
	Lin	0.513	0.533	0.455
	Wu	0.495	0.516	0.456
Optimal Pairing	Resnik	0.519	0.532	**<u>0.464</u>**
	Jiang	0.516	**<u>0.534</u>**	**<u>0.464</u>**
	Leacock	0.489	0.516	0.454
	Lin	**0.520**	0.533	0.456
	Wu	0.495	0.525	0.457
QAP	Resnik	0.520	0.532	**<u>0.464</u>**
	Jiang	0.520	**<u>0.534</u>**	**<u>0.464</u>**
	Leacock	0.498	0.522	0.455
	Lin	**<u>0.525</u>**	**<u>0.534</u>**	0.457
	Wu	0.508	0.525	0.459

A brief analysis of the results is as follows.

Difficulty level of datasets. From the table it can be observed that there is a clear difference between the performance of all techniques for the three datasets. That is, all combinations of techniques obtained high F_1 scores for UA dataset, moderate F_1 scores for BR dataset, and low F_1 scores for AM dataset. These results indicate that the corresponding activity pairs of the AM dataset are harder-to-detect than that of UA and BR datasets. Furthermore, the corresponding activities in the BR datasets are harder-to-detect than that of the UA dataset.

Performance variation across word-level measures. From Table 2 it can be observed that in the case of Greedy pairing sentence-level method, Jiang similarity obtained the highest F_1 scores for all the three datasets (0.516, 0.534 and 0.464). Furthermore, for Optimal pairing sentence-level method, Jiang similarity obtained the highest F_1 scores for two datasets, BR and AM datasets (0.534 and 0.464) whereas, Lin similarity obtained the highest F_1 score for one dataset, UA dataset. Similarly, for QAP pairing, Jiang similarity obtained the highest F_1 scores for two datasets, BR and AM datasets, whereas Lin similarity obtained the highest F_1 score for one dataset, UA dataset. Based on these observations and the previous observation about the hardness of the three datasets (AM > BR > UA) we conclude, Jiang similarity is the most suitable word-level semantic similarity measure.

Performance variation across sentence-level methods. From Table 2 it can be observed that for the UA dataset, among the five word-level measures, Lin similarity obtained the highest F_1 score with Optimal and QAP pairing (i.e. 0.520 for Optimal and 0.525 for QAP pairing). However, in the case of Greedy pairing sentence-level method, both Resnik and Jiang measures obtained a higher F_1 score than Lin similarity. Similarly, for the BR dataset, both Lin similarity and Jiang similarity obtained the highest F1 score with QAP pairing. However, the F_1 scores obtained by Jiang similarity with Optimal and Greedy pairing is higher than that of the Lin similarity. These changes in the best performing similarity measures due to the change in sentence-level methods, highlights the significance of sentence-level methods. Hence, we conclude that adequate attention should be given to the choice of the sentence-level methods. Another key observation regarding the sentence-level methods is that, for each dataset, the highest F_1 score obtained by a word-level measure involved QAP pairing. This indicates that QAP pairing is the most suitable sentence-level aggregation method than Optimal and Greedy pairing methods.

5 Related Work

A plethora of process model matching techniques have been developed which can be broadly divided into two types, syntactic and semantic [5]. Syntactic techniques merely rely on the similarity or distance between the labels without taking into consideration the meaning of the words. In contrast, semantic techniques rely on the semantics of words for computing similarity.

A recent survey of PMM has identified a set of semantic matching techniques that are used in literature [5]. A summary of these techniques is presented in Table 3. In the table, Wu, Leacock, Jiang represents Wu & Palmer, Leacock & Chodorow and Jiang & Conrath word-level semantic similarity measures. The '+' sign in the table indicates that the technique uses the respective technique, whereas the '–' sign indicates that the technique is not used in the paper. There are occasions in which the use of synonyms is implicit, these are marked as '+/'.

Table 3. Semantic matching techniques using for computing similarities between activity pairs

Author et al.	Synonyms	Lesk	Wu	Leacock	Resnik	Jiang	Lin
Sebu et al. [25]	+	+	-	-	-	-	+
Sonntag et al. [26]	+	-	-	-	-	-	+
Sebu et al. [28]	+	+	-	-	-	-	+
Makni et al. [29]	+	-	-	-	-	-	-
Fengel [30]	-	-	-	-	-	-	-

Pittke et al. [31]	+	-	-	-	-	-	+
Klinkmuller et al. [32]	-	-	-	-	-	-	+
Klinkmuller et al. [33]	+	-	+	-	-	-	+
Klinkmuller et al. [34]	+	-	-	-	-	-	+
Jin et al.[35]	+	-	-	-	-	-	-
Caygolu et al. [36]	-	-	+	-	-	-	-
Belhoul et al. [37]	-	-	+	-	-	-	-
Niemann et al.[38]	+	-	-	-	-	-	-
Leopold et al. [39]	+	-	-	-	-	-	+
Humm et al. [40]	+/-	-	-	-	-	-	-
Dijkman et al. [41]	+/-	-	-	-	-	-	-
Dumas et al. [42]	+	-	-	-	-	-	-
Dongen et al. [43]	+/-	-	-	-	-	-	-
Agnes et al. [44]	+	-	+	-	-	-	-
Ehrig et al. [45]	+	-	+	-	-	-	-
Corrales et al. [46]	+	-	-	-	-	-	-
Schoknecht et al. [8]	-	-	-	-	-	+	-

From the table it can be observed that most of the studies propose to use synonyms for semantic similarity. However, these studies do not explicitly present the measures used for computing similarity. Also, it can be seen from the table that, Lesk, Wu and Lin are the other similarity measures used in literature. Furthermore, it can be observed that Leacock, Resnik and Jiang measures have never been used for identifying corresponding activities between a pair of process models. Additionally, only word-level semantic similarity measures are considered for computing similarities between activity pairs and these word-level similarity measures have not been extended to compute similarity at activity-level.

6 Conclusion

Several semantic Process Model Matching (PMM) techniques have been proposed, however these techniques merely focus on the word-to-word semantic similarity, without due consideration to aggregation of word-level similarity to sentence-level (or activity-level) similarity. Furthermore, the existing studies have only used three semantic similarity measures and ignored the other semantic similarity techniques that have shown promising results for various text processing tasks. To that end, in this paper, we have used five word-level sematic similarity measures and three sentence-level aggregation techniques to experimentally evaluate the effectiveness of all the 15 combinations in the context of PMM. For the experiments we have used well established datasets from PMMC'15. The results reveal the following: a) the hardness of the three datasets are different, with AM dataset being the hardest, BR dataset being the moderate, and UA dataset being the easiest, b) Jiang similarity, is the most suitable matching technique, and c) QAP Pairing is the most effective sentence-level measure. In the future, we plan to compare the performance of these semantic measures with all the existing matching techniques.

References

1. Kuss, E., Leopold, H., et al: Probabilistic Evaluation of Process Model Matching Techniques. In: Proceedings of the 35th International Conference on Conceptual Modeling, pp. 279-292, Gifu, Japan, (2016)
2. Cayoglu, U., Dijkman, R., Dumas, M. et al.: Report: The Process Model Matching Contest 2013. In: Proceedings of the Business Process Management Workshops, Springer LNBIP, pp. 442-463, Beijing, China, (2013).
3. Meilicke, C., Leopold, H., et al: Overcoming Individual Process Model Matcher Weaknesses Using Ensemble Matching. DSS, 100 (1), pp. 15-26, (2017).
4. Kuss, E., Leopold, H., et al: Ranking-based evaluation of process model matching. In: Proceedings of the International Conference on Cooperative Information Systems, pp. 298-305, Rhodes, Greece, (2017).
5. Jabeen, F., Leopold, H., Reijers, H. A.: How to make process model matching work better? An analysis of current similarity measures. In: 20th International Conference on BIS, pp. 181-193. Springer LNBIP, Poznan, Poland, (2017).
6. Leacock, C., Chodorow, M.: Combining local context and WordNet similarity for word sense identification. In: WordNet: An electronic lexical database, MIT Press, (1998).
7. Resnik, P.: Using information content to evaluate semantic similarity in a taxonomy. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence, pp. 448-453. Quebec, Canada, (1995).
8. Schoknecht, A., Thaler, T., Fettke, P., Obserweis, A., Laue, R.: Similarity of Business Process models – A State-of-the-Art Analysis. ACM Computing Survey, 50(4), pp. 1-33, (2017).
9. Lin, D.: An information-theoretic definition of similarity. In: Proceedings of the 15th International Conference on Machine Learning, pp. 296-304, (1998).
10. Wu, Z., Palmer, M.: Verbs semantics and lexical selection. In: Proceedings of the 32nd Annual Meeting on ACL, pp. 133-138, Las Cruces, New Maxico, (1994).
11. Budanitsky, A., Hirst, G.: Evaluating WordNet-based measures of lexical semantic relatedness. Computational Linguistics, 32(1), pp. 13-47, (2006).
12. Navigli, R.: Word sense disambiguation: A survey. ACM Computing Surveys, 41(2), pp. 1-12, (2009).
13. Androutsopoulos, I., Malakasiotis, P.: A survey of paraphrasing and textual entailment methods. Journal of AI Research, 38 (1), pp. 135-187, (2010).
14. Gabrilovich, E., Markovitch, S.: Computing semantic relatedness using Wikipedia-based explicit semantic analysis. In: Proceedings of the 20th International Joint Conference on Artificial Intelligence, pp. 1606-1611, Hyderabad, (2007).

15. Lord, W.P., Stevens, R.D, Brass, A., Goble, C.A.: Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation. Bioinformatics, 19(10), pp.1275-1283, (2003).
16. Farhadi, A., Hejrati, M., et al.: Every picture tells a story: Generating sentences from images. In: Proceedings of the European Conference on Computer Vision, Springer LNCS, pp. 15-29, Crete, Berlin, (2010).
17. Androutsopoulos, I., Malakasiotis, P.: A survey of paraphrasing and textual entailment methods. Journal of AI Research, 38 (1), pp. 135-187, (2010).
18. Liu, B.: Sentiment analysis and opinion mining. Morgan & Claypool, (2012).
19. McCallum, A., Xuerui, W., Corrada-Emmanuel, A.: Topic and role discovery in social networks with experiments on enron and academic email. Journal of AI Research, 30 (1), pp. 249-272, (2007).
20. Chang, J., Garber, J.B., Gerrish, S., Wang, C., Blei, D.M.: Reading tea leaves: How humans interpret topic models. In: Proceedings of the 22nd International Conference on NIPS, pp. 288-296, BC, Canada, (2009).
21. Corley, C., Mihalcea, R.: Measuring the semantic similarity of texts. In: Proceedings of the ACL Workshop on empirical modeling of semantic equivalence and entailment, pp. 13-18, Michigan, USA, (2005).
22. Rus, V., Lintean, M.: A comparison of greedy and optimal assessment of natural language student input using word-to-word similarity metrics. In: Proceedings of the Seventh Workshop on Building Educational Applications Using NLP, pp. 157-162, Montreal, Canada, (2012).
23. Lintean, M., Rus, V.: An Optimal Quadratic Approach to Monolingual Paraphrase Alignment. In: Proceedings of the 20th Nordic Conference of Computational Linguistics, pp. 127-134, Vilnius, Lithuania, (2015).
24. Antunes, G., Bakhshandeh, M., Borbinha, J., Cardoso, J., Dadashnia, S., Di Francescomarino, C. & Hake, P. (2015). The process model matching contest 2015. GI-Edition/Proceedings: Lecture notes in informatics, 248, 127-155.
25. Sebu, M. L., Ciocârlie, H.: Similarity of business process models in a modular design. In: Proceedings of the IEEE 11th Symposium on Applied Computational Intelligence and Informatics, pp. 31-36, Timisoara, Romania, (2016).
26. Sonntag, A., Hake, P., Fettke, P., Loos, P.: An approach for semantic business process model matching using supervised learning. In: Proceedings of the European Conference on Information System, pp. 1-10, Istanbul Turkey, (2016).
27. Sebu, M. L., Ciocarlie, H.: Merging business processes for a common workflow in an organizational collaborative scenario. In Proceedings of the 19th International Conference on STCC, pp. 134-139, Cheile Gradistei, Romania, (2015).
28. Makni, L., Haddar, N.Z., Ben-Abdallah, H.: Business process model matching: An approach based on semantics and structure. In Proceedings of the 12th International Conference on e-Business and Telecommunications, pp. 64-71, (2015).
29. Fengel, J.: Semantic technologies for aligning heterogeneous business process models. Business Process Management Journal, 20(4), pp. 549-570, (2014).
30. Pittke, F., Leopold, H., Mendling, J., Tamm, G.: Enabling reuse of process models through the detection of similar process parts. In: Proceedings of the BPM Workshops, Springer LNBIP, pp. 586-597, Tallinn, Estonia, (2012).

31. Klinkmüller, C., Leopold, H., et al.: Listen to me: Improving process model matching through user feedback. In: Proceedings of the International Conference on BPM, Springer LNCS, pp. 84-100, Haifa, Israel, (2014).
32. Klinkmüller, C., Weber, I., Mendling, J. et al.: Increasing recall of process model matching by improved activity label matching. In: Proceedings of the International Conference on BPM, Springer LNCS, pp. 211-218, Beijing, China (2013).
33. Jin, T., Wang, J., La Rosa, M., Ter Hofstede, A., Wen, L. Efficient querying of large process model repositories. Computers in Industry, 64(1), pp.41-49, (2013).
34. Cayoglu, U., Oberweis, A., Schoknecht, A., Ullrich, M.: Triple-S: a matching approach for Petri nets on syntactic, semantic and structural level. In: Proceedings of PMMC'13, co-located with BPM, Beijing, China (2013).
35. Belhoul, Y., Haddad, M., et al.: String comparators based algorithms for process model matchmaking. In: Proceedings of the International Conference on Services Computing (SCC), pp. 649-656, Honolulu, USA, (2012).
36. Niemann, M., Siebenhaar, M., et al.: Comparison and retrieval of process models using related cluster pairs. Computers in Industry, 63(2), pp. 168-180, (2012).
37. Leopold, H., Niepert, M., Weidlich, et al: Probabilistic optimization of semantic process model matching. In: Proceedings of the International Conference on BPM, pp. 319-334, Tallinn, Estonia, (2012).
38. Humm, B. G., Fengel, J.: Semantics-based business process model similarity. In: Proceedings of the International Conference on Business Information Systems, pp. 36-47, Vilnius, Lithuania, (2012).
39. Dijkman, R., Dumas, M., et al.: Similarity of business process models: Metrics and evaluation. Information Systems, 36(2), pp. 498-516. (2011).
40. Dumas, M., Garcia-Banuelos, L., Dijkman, R.M. Similarity search of business process models. IEEE Data Eng. Bull., 32(3), pp. 23-28, (2009).
41. Van Dongen, B., Dijkman, R. et al.: Measuring similarity between business process models. In: Proceedings of the CAiSE, pp. 405-419, Valencia, Spain, (2013).
42. Koschmider, A., Oberweis, A.: How to detect semantic business process model variants? In: Proceedings of the ACM symposium on Applied computing, pp. 1263-1264, Seoul, Korea, (2007).
43. Ehrig, M., Koschmider, A. et al.: Measuring similarity between semantic business process models. In: Proceedings of the 4th Asia-Pacific conference on Conceptual modelling, Springer LNCS, pp. 71-80, Ballarat, Australia, (2007).
44. Corrales, J., Grigori, D. et al.: BPEL processes matchmaking for service discovery. In: Proceedings of the CoopIS, pp. 237-254, Montpellier, France, (2006).