On Efficacy of Approximating Arbitrary Relations by Partial Orders

Ryszard Janicki[0000−0001−5364−9725] and Ni Yang

Department of Computing and Software, McMaster University,
Hamilton, Ontario, L8S 4K1 Canada
janicki@mcmaster.ca, yangn19@mcmaster.ca

Abstract. The problem of optimal quantitative approximation of an arbitrary binary relation by a partial order is discussed and the results of some experiments are discussed. In general, this problem is NP-hard even for very simple quantitative measures, so some alternative sub-optimal but relatively efficient algorithms are discussed and tested.

Keywords: approximations of relations, partial orders, similarity, efficacy of algorithms

1 Introduction

Consider the following problem (cf. [11]): we have a set of data that have been obtained in an empirical manner. From the nature of the problem we know that the set should be partially ordered, but because the data are empirical it is not. In a general case, this relation may be arbitrary. What is the 'best' partially ordered approximation of an arbitrary relation and how this approximation can be computed?" Some solutions have been proposed in [11,13,22] and their efficacy is discussed in this paper. Areas of immediate applications of any 'best' partial order approximation algorithm include group ranking, social choice, pairwise comparisons based non-numerical ranking, analysis of subjective judgments, etc. ([7,8,10,15]). To illustrate this concept, imagine organizing a diverse collection of books in a library without a clear hierarchy. How do we best arrange them in a manner that respects certain desired orderings, like genre or authorship, yet acknowledges the inherent complexity of literary classification? This analogy simplifies our exploration into the realm of partial orders.

It was shown in [13] that the problem is NP-hard even for very simple similarity measures. On the other hand there are a few $O(n^3)$ partial order approximations, including the classical Schröder's from 1895 [11,13,22], but their efficacy is basically unknown.

In this paper we will analyse experimental efficacy of three partial order approximations, the classical Schröder approximation, denoted $(R^+)^\bullet$ [22], the reverse Schröder [11], denoted $(R^\bullet)^+$, and some randomized approximation from [13], denoted $R^\odot$. For comparisons, we will use four different similarity measures, namely, absolute similarity [13], Kemeny distance [5,16], popular Jaccard index [5,9], and a new asymmetric similarity measure proposed in this paper.

2 Relations, Partial Orders and Partial Order Approximations

In this section we will recall basic concepts and results used in this paper [2,7,11,21].

Let X be a finite set and $R \subseteq X \times X$ be a *relation* on X. If $(a,b) \in R$ we will often write aRb. Every such relation R can be interpreted as a directed graph $G_R = (V,E)$ where $V = X$ and $E = R$. In this paper, we will consider the terms 'relation' and 'graph' as equivalent, and use both of them, dependently on a context.

For two relations R,S on X, their composition, denoted RS, is defined as $xRSz \iff \exists y \in X.\ xRy \wedge ySz$. The identity relation Id_X, is defined as $Id_X = \{(x,x) \mid x \in X\}$. For every relation R on X, and every $i = 0,1,\ldots,\infty$, we define the relation R^i, as $R^0 = Id_X$ and $R^{i+1} = RR^i$, for $i = 0,1,\ldots,\infty$.

The *transitive closure* of R, denoted R^+, is defined as $R^+ = \bigcup_{i=1}^{\infty} R^i$.

A relation $<\in X \times X$ is a *(sharp) partial order* if it is irreflexive and transitive. Formally, $\neg(a < a)$ and $a < b < c \implies a < c$ for all $a,b,c \in X$.

For every relation R on X, its *cyclic closure*, denoted R^{cyc}, is defined as follows $\forall a,b \in X.\ aR^{cyc}b \iff aR^+b \wedge bR^+a$. In graph terminology, if $aR^{cyc}b$ then a and b are *strongly connected* in $G_R = (X,R)$.

For every relation R on X, the *acyclic refinement* of R, denoted $R^\bullet$, is defined as $R^\bullet = R \setminus R^{cyc}$. In graph terminology, the graph $G_{R^\bullet} = (X,R^\bullet)$ has been derived from $G_R = (X,R)$ by deleting all edges from all *strongly connected components* of G_R.

Let R be a relation on X and $a \in X$. We define: $aR = \{x \mid aRx\}$ and $Ra = \{x \mid xRa\}$.

For any relation R on X, we define the *equivalence relation with respect to R*, denoted $\equiv_R$, as follows $\forall a,b \in X.\ a \equiv_R b \iff aR = bR \wedge Ra = Rb$.

This relation was first proposed for R being a partial order ([7]). It was later generalized for arbitrary relations in ([11]). Elements a and b are equivalent with respect to R if they have the same set of elements related to them and they relate to the same set of elements. In principle, if $a \equiv_R b$, then there is nothing in R that can distinguish between a and b (with respect to the relation R).

The following proposition characterizes the basic properties of the above concepts.

Proposition 1 ([11]).

1. $R \subseteq R^+$.
2. $R^\bullet \subseteq R$, *$R^\bullet$ is acyclic (i.e. also irreflexive), and* $aR^\bullet b \iff aRb \wedge \neg(bR^+a)$.
3. *If R is a partial order then* $R = R^+ = R^\bullet$.
4. $\equiv_R \subseteq \equiv_{R^+}$ *and* $\equiv_R \subseteq \equiv_{R^\bullet}$. □

The above proposition presents basic hints for partial order approximations. R^+ is a kind of an upper approximation of R, $R^\bullet$ is a kind of lower approximations, and each partial order is acyclic and transitive.

We say that a partial order $<_R$ is an *approximation* of a relation R if R and $<_R$ are 'similar' or 'close'. The definition is very imprecise and relies heavily on intuition and what R is supposed to represent. Unfortunately, a widely accepted formal definition of *relation approximation* does not exist so far [13].

The standard approach is to modify a given relation R so it has the desired properties, for partial ordering this would be *acyclity* and *transitivity*.

This standard approach was first proposed by E. Schröder in 1895 [22]. A slightly more formal definition of partial ordered approximation, motivated by ranking and pairwise comparisons, and called *property-driven partial order approximation* was proposed in [11] and [12].

Schröder's approximation has many equivalent formulations, but in the terminology of this paper, for a given arbitrary relation R, it is just $(R^+)^\bullet$.

Theorem 1 (E. Schröder, 1895, [22]). *For every relation R on X, the relation $(R^+)^\bullet$ is a partial order.* □

Despite its popularity, the efficacy of Schröder's approximation most likely has never been sufficiently studied. This is one of the goals of this paper.

In [11,12] the definition of *property-driven partial order approximation* was presented and its properties, motivation and interpretations were discussed.

Definition 1 ([11]). *A partial order $<\subseteq X \times X$ is a **property-driven partial order approximation** of a relation $R \subseteq X \times X$ if it satisfies the following four conditions:*

1. $a < b \implies aR^+b$,
2. $aR^\bullet b \implies a < b$,
3. $a \equiv_R b \implies a \equiv_< b$.
4. $a < b \implies \neg aR^{cyc}b$ *(or, equivalently* $a < b \implies \neg bR^+a$*).* □

Definition 1 was motivated by the following intuitions [12]. Since R^+ is the smallest transitive relation containing R, and due to informational noise, imprecision, randomness, etc., some parts of R might be missing, it is reasonable to assume that R^+ is the upper bound of $<$, so condition (1). Condition (2) defines the lower bound. The greatest partial order included in R usually does not exist, but when R is interpreted as an estimation of a ranking, $R^\bullet$ appears to be a reasonable lower bound [12]. Condition (3) ensures preservation of the equivalence with respect to R. Condition (4) says that if $aR^{cyc}b$ then usually a and b are incomparable. If R is interpreted as an estimation of a ranking, then in most cases $aR^{cyc}b$ it is interpreted that a and b are indifferent [8]. Similar interpretations take place in concurrency theory [17].

The following result characterizing property-driven partial order approximations has been proven.

Theorem 2 ([11]).

1. *$(R^\bullet)^+$ is a partial order.*
2. *The relations $(R^\bullet)^+$ and $(R^+)^\bullet$ are property-driven partial order approximations of R.*
3. $(R^\bullet)^+ \subseteq (R^+)^\bullet$. □

The relation $(R^\bullet)^+$, introduced in [11], is in this paper called *reverse Schröder approximation*. Some other property-driven partial order approximations we considered in [10] and [12], however, $(R^\bullet)^+$ and $(R^+)^\bullet$ are the two most important ones, the only ones that are also analysed in [13]. Again, how good both $(R^\bullet)^+$ and $(R^+)^\bullet$ really are, have never being tested.

3 Similarities and Optimal Approximation

The definition and analysis of *optimal approximation* problem requires using the concept of *similarity*. As opposed to the orthogonal concept of a *distance*, the concept of *similarity* does not have standard indisputable axiomatization ([5,13,20,23]). Depending on the area of application, some desirable properties may vary ([5,13,14,18,20,23]), sometimes substantially ([13,20]). We will say that a (scaled) similarity *sim* is *metrical,* if the function $1 - sim(A,B)$ is a *distance* for appropriate objects A,B (cf. [5]). Many useful and popular similarities are not metrical (cf. [13,23]).

Recall that for any set of objects O, a *distance* is any function $dist : O \times O \to [0,\infty)$ such that $dist(\alpha,\beta) = 0 \iff \alpha = \beta$, $dist(\alpha,\beta) = dist(\beta,\alpha)$ and $dist(\alpha,\gamma) \leq dist(\alpha,\beta) + dist(\beta,\gamma)$, for all $\alpha,\beta,\gamma \in O$. The latter inequality is called the 'triangle property' ([5]).

In this paper we will use three similarities and one distance that work well when one of the relations is a partial order, namely *absolute similarity* [13], *Kemeny distance* [16], *Jaccard index* [9], and newly defined asymmetric similarity, called *weighted absolute similarity*.

- The *absolute similarity* between two relation R and S is defined as [13]:

$$sim_A(R,S) = |R \cap S|.$$

It is argued in ([13]) that the *absolute similarity* is most likely the simplest, yet adequate, measure of similarity between two binary relations.

- The *Kemeny distance* [16] is defined as as

$$dist_K(R,S) = |(R \setminus S) \cup (S \setminus R)| = |R \setminus S| + |S \setminus R| = |R \cup S| - |R \cap S|.$$

Note that $dist_K(R,S)$, which is just a cardinality of *symmetric difference* ([21]), is a proper *metric*, i.e. the 'triangle property': $dist_K(R,S) \leq dist_K(R,T) + dist_K(T,S)$ is satisfied for all relations R,S,T (cf. [5,16]). The symmetric difference between two relations is often considered a measure of how "far apart" they are ([5]).

It was argued in [13] that $sim_A(R,S)$ better measures the relationship between R and S with respect to acyclity, while $dist_K(R,S)$ measures this relationship better with respect to transitivity, and we cannot replace one by another.

- *Jaccard similarity* [9] is defined as:

$$sim_J(R,S) = \frac{|R \cap S|}{|R \cup S|}.$$

Jaccard index, proposed in 1901, is the oldest quantitative formula to measure similarity. It is a special case of more general Marczewski-Steinhaus similarity [18]; and it is still the most popular one with a lot of various applications. Jaccard similarity has values in the interval $[0,1]$, and we can define *Jaccard distance*, $dist_J$, simple as $dist_J(R,S) = 1 - sim_J(R,S) = \frac{|R \cup S| - |R \cap S|}{|R \cup S|}$. It can easily be shown that $dist_J$ is a proper distance, i.e. it has the 'triangle property'. One can also interpret the Jaccard similarity as a (*symmetric*) *weighted* version of the absolute similarity sim_A.

- The *asymmetric weighted absolute similarity* is defined as follows:

$$sim_{AWA}(R,S) = \frac{|R \cap S|}{|R|}.$$

The above similarity measure is asymmetric, it measures 'how S is similar to R', but not necessarily vice versa. Asymmetric similarity measures do exist and have right interpretations for many applications. The classical example is to compare a variant to a prototype [23]. In our case, we are interesting in *approximating R by S*, and not necessarily vice versa. For this reason we think that the asymmetric weighted absolute similarity might be a better measure than the standard Jaccard similarity. The similarity sim_{AWA} is not metrical. It can also be considered as a special case of a weighted version of the Jaccard index [1,19].

Choosing a proper similarity measure is a complex issue. It has been discussed in some detail in [3,4], and for some special cases in [13,23].

Proposition 2 ([13]).

1. $sim_A(R,(R^+)^\bullet) = sim_A(R,(R^\bullet)^+) = |R^\bullet|$.
2. $dist_K(R,(R^\bullet)^+) \leq dist_K(R,(R^+)^\bullet)$.
3. $sim_J(R,(R^\bullet)^+) \geq sim_J(R,(R^+)^\bullet)$. □

Corollary 1. $sim_{AWA}(R,(R^+)^\bullet) = sim_{AWA}(R,(R^\bullet)^+) = \frac{|R^\bullet|}{|R|}$. □

Proposition 2 indicates that for Kemeny distance and Jaccard similarity, the partial order $(R^\bullet)^+$ performs better than $(R^+)^\bullet$, but how better is beyond the theoretical analysis.

Let $sim(\ldots)$ be some similarity for binary relations over some set X, $dist(\ldots)$ be some distance for binary relations over X, and let $\mathbb{PO}(X)$ be the set of all partial orders over the set X.

A relation $S \in$ is an *optimal partial order approximation* of the relation R with respect to *sim*, if and only if $S \in \mathbb{PO}(X)$, and

$$\forall S' \in \mathbb{PO}(X).\ sim(R,S) \geq sim(R,S').$$

Similarly, S is an *optimal partial order approximation* of the relation R with respect to *dist*, if and only if $S \in \mathbb{PO}(X)$, and

$$\forall S' \in \mathbb{PO}(X).\ dist(R,S) \leq dist(R,S').$$

In many cases, we have more than one optimal approximation, for both definitions.

The main technical contribution of the paper [13] is that finding optimal approximation is NP-hard even for such simple similarities as *absolute similarity* $sim_A(R,S) = |R \cap S|$. On the other hand, the property-driven partial order approximations $(R^\bullet)^+$ and $(R^+)^\bullet$ are $O(n^3)$, but we really do not know how accurate they are.

The following simple randomized algorithm for partial order approximation, denoted $R^\odot$, again with $O(n^3)$ time complexity, has been proposed in [13].

Algorithm 1 (Calculation of $R^{\odot}$ [13]) *Let $R \subseteq X \times X$ and $|X| = n$.*

1. *Calculate $R^{\bullet}$. Set $\bar{R} = R^{\bullet}$.*
2. *Set $R^{\odot} := \bar{R}$.*
3. *Pick **randomly** $(a,b) \in R \setminus \bar{R}$ and add (a,b) to $\bar{R}$, i.e., $\bar{R} := \bar{R} \cup \{(a,b)\}$.*
4. *If $\bar{R}$ is acyclic, i.e., $(\bar{R})^{cyc} = \emptyset$, go to (2).*
5. *Return $R^{\odot} := (R^{\odot})^{+}$.*

How the partial order $R^{\odot}$ performs when compared with $(R^{\bullet})^{+}$ and $(R^{+})^{\bullet}$ was a open problem, and the main motivation of this paper.

4 Experiments and Their Results

Since finding an optimal partial order approximation is not feasible for even medium size relations, when it comes to applications, we have to decide which feasible approximation, $(R^{\bullet})^{+}$, $(R^{+})^{\bullet}$ and $R^{\odot}$ should be chosen. Since theoretical analysis is inconclusive ([13]), the only choice is to conduct some empirical experiments using various similarity measures.

We generated randomly relations of various sizes, and for each relation R we calculated:

- $sim_A(R,(R^{\bullet})^{+}) = sim_A(R,(R^{+})^{\bullet}) = |R^{\bullet}|$, $sim_A(R,R^{\odot})$,
- $dist_K(R,(R^{\bullet})^{+})$, $dist_K(R,(R^{+})^{\bullet})$, $dist_K(R,R^{\odot})$,
- $sim_J(R,(R^{\bullet})^{+})$, $sim_J(R,(R^{+})^{\bullet})$, $sim_J(R,R^{\odot})$,
- $sim_{AWA}(R,(R^{\bullet})^{+}) = sim_{AWA}(R,(R^{+})^{\bullet}) = \frac{|R^{\bullet}|}{|R|}$, $sim_{AWA}(R,R^{\odot})$.

Then, we aggregated and analysed the obtained results. The algorithms used to compute all the similarity measures are standard, their descriptions can be found for example in [2]. All programming was done in Python and all codes can be found in [24]. The only non standard algorithm was the one used for random graphs generations.

For random graphs generation we used Erdös-Rényi algorithm [6]. This algorithm has two parameters, edge probability ep, and cycle probability cp. The version we used assumed uniform distribution for both edges and cycles. We conducted nine series of experiments, for all pairs (ep,cp) where $ep,cp \in \{0.25,0.5,0.75\}$. The value $ep = 0.25$ indicates that the graphs generated are rather sparse, while $ep = 0.75$ indicates that the graphs are rather dense. The value $cp = 0.25$ indicates that most graphs are DAGs (Directed Acyclic Graphs) and $cp = 0.75$ indicates a lot of cycles.

In each series we generated graphs with numbers of nodes from 1 to 49, 100 graphs for each number of nodes, and then perform appropriate calculations. All the detailed results, including forty eight graphs, illustrating all different cases and aspects, can be found in [24].

It was reasonable to expect different results for different types of graphs, however, the results turned out to be surprisingly homogeneous. The number of cycles (parameter $cp \in \{0.25,0.5,0.75\}$) practically does not matter. For sparse graphs, i.e. $ep = 0.25$, the results were slightly different than for $ep = 0.5$ and $ep = 0.75$, but not significantly.

Below, in Figure 1, we present the results for $ep = 0.5$ and $cp = 0.5$, i.e. the most average case. This is a representative example for all cases except the Kemeny distance for $ep = 25$, i.e. sparse graphs. For the asymmetric similarity sim_{AWA} we use the percentage scale from 1 to 100.

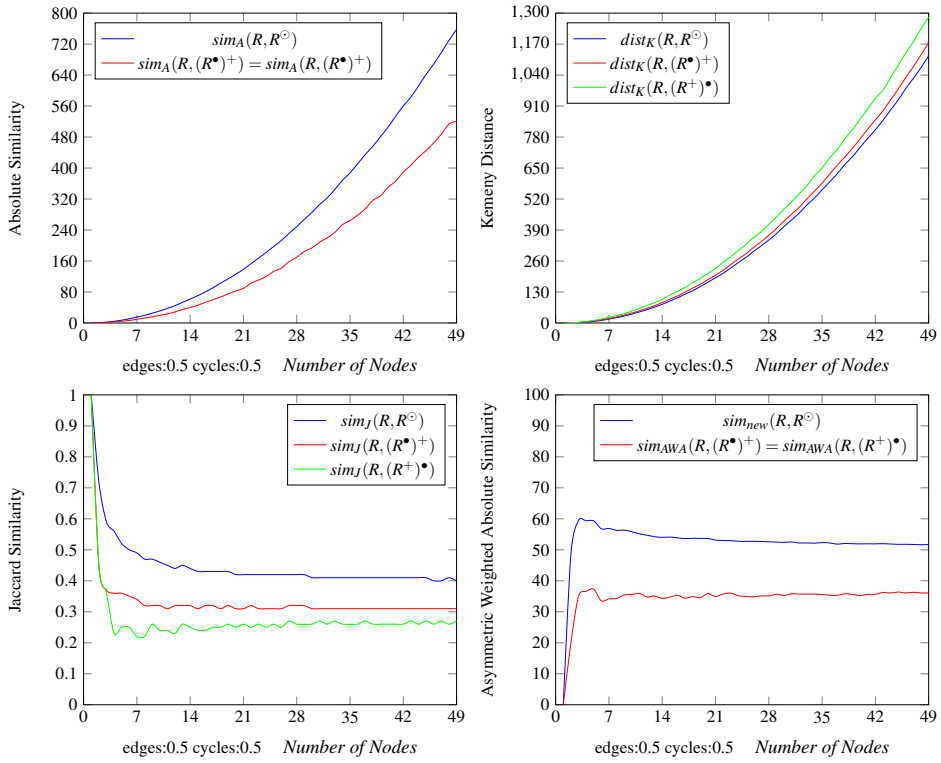

Fig. 1: The graphs of similarity measures for $ep = 0.5$ and $cp = 0.5$.

In this case, the approximation $R^{\odot}$ always outperforms both $(R^{\bullet})^{+}$ and $(R^{\bullet})^{+}$. For Absolute similarity sim_A and Asymmetric weighted absolute similarity sim_{AWA}, the results are the same (Proposition 2 and Corollary 1) for $(R^{\bullet})^{+}$ and $(R^{+})^{\bullet}$, but for Kemeny distance and Jaccard similarity, $(R^{\bullet})^{+}$ - the reverse Schröder, performs better than $(R^{+})^{\bullet}$ - the classical Schröder. For Kemeny distance the differences are the smallest.

Both Jaccard similarity and Asymmetric weighted absolute similarity stabilize at about 25 nodes, so we may safely assume that the results for all $n \geq 30$ would be almost identical.

While $R^{\odot}$ outperforms both $(R^{\bullet})^{+}$ and $(R^{\bullet})^{+}$ rather substantially, the differences between $(R^{\bullet})^{+}$ and $(R^{+})^{\bullet}$ are not that big.

The results for other values of ep and cp, are practically the same. Of course, the denser graphs have for edges for the same amount of vertices, so numerical values of functions sim_A, $dist_K$, sim_J and sim_{AWA} vary, but the shapes of functions and proportional differences between $R^{\odot}, (R^{\bullet})^{+}$ and $(R^{+})^{\bullet}$ are almost identical.

In all cases, both Jaccard similarity and asymmetric weighted absolute similarity stabilize at less than 25 nodes. The dense functions/graphs stabilize slightly earlier, at 20 nodes for $ep = 0.75$. The values of functions for $n \geq 30$ are called *saturated values*, and they are presented in Table 1 below (bold fonts). They are all the same for all $n \geq 30$.

Values for *number of nodes* ≥ 30 (**saturated values**, bold fonts)								
Jaccard similarity					Asymmetric weighted absolute similarity			
ep	cp	$R^{\odot}$	$(R^{\bullet})^{+}$	$(R^{+})^{\bullet}$	ep	cp	$R^{\odot}$	$(R^{\bullet})^{+}$ and $(R^{+})^{\bullet}$
0.25	0.25	**0.28**	**0.24**	**0.19**	0.25	0.25	**0.58**	**0.33**
0.25	0.5	**0.29**	**0.2**	**0.18**	0.25	0.5	**0.55**	**0.32**
0.25	0.75	**0.30**	**0.24**	**0.21**	0.25	0.75	**0.52**	**0.31**
0.5	0.25	**0.39**	**0.30**	**0.25**	0.5	0.25	**0.53**	**0.37**
0.5	0.5	**0.41**	**0.31**	**0.28**	0.5	0.5	**0.52**	**0.37**
0.5	0.75	**0.41**	**0.30**	**0.25**	0.5	0.75	**0.51**	**0.37**
0.75	0.25	**0.47**	**0.35**	**0.26**	0.75	0.25	**0.52**	**0.38**
0.75	0.5	**0.48**	**0.36**	**0.30**	0.75	0.5	**0.51**	**0.37**
0.75	0.75	**0.49**	**0.36**	**0.30**	0.75	0.75	**0.51**	**0.38**

Table 1: Saturated values for Jaccard similarity and Asymmetric weighted absolute similarity.

With respect to relation/graph density, the Jaccard similarity, sim_J, is better for dense graphs and worse for sparse graphs, for all three approximations $R^{\odot}, (R^{\bullet})^{+}$ and $(R^{+})^{\bullet}$. This could be explained by the fact that all three approximations use transitive closure at some point, and for sparse graphs, on average, the transitive closure proportionally adds more edges than in the case of dense graphs.

For the asymmetric absolute similarity, sim_{AWA}, we have a little bit different story. Since $sim_{AWA}(R, (R^{\bullet})^{+}) = sim_{AWA}(R, (R^{+})^{\bullet}) = |R^{\bullet}|/|R|$, we have only two columns of numbers, and in this case, the results are much more uniform. For the approximation $R^{\odot}$, the results are almost identical, around 0.52, except the values in two top rows of the above table, which are slightly bigger, 0.55 and 0.58. As opposed to Jaccard similarity, the asymmetric absolute similarity performs slightly better for sparse graphs. This would indicate that for sparse graphs, the relation $R^{\odot} \setminus R$ is proportionally bigger than for dense graphs. The saturated values for $(R^{\bullet})^{+}$ and $(R^{+})^{\bullet}$ are practically constant at about 0.37, with the exception of sparse graphs, where they are around 0.32. In this case, the trend is the same as for Jaccard similarity, the worse performance for sparse graphs, but the difference is smaller.

5 Summary and Final Comments

This paper deals with the efficacy of approximating arbitrary binary relations by partial orders. The problem is NP-hard in general [13], however there are three, intuitively promising, $O(n^3)$ approximations, namely the classical Schröder approximation from 1895, $(R^+)^\bullet$, [22], the reverse Schröder approximation $(R^\bullet)^+$ proposed in [11], and randomized approximation $R^\odot$ proposed in [13]. In this paper we tested the efficacy of these three approximations. Four similarity measures were used, absolute similarity [13], Kemeny distance [16], Jaccard similarity [9], and a new similarity, sim_{AWA}, introduced in Section 3 of this paper. The tests were performed on randomly generated relations of various sizes and densities. Erdös-Rényi algorithm [6] was used to generate random graphs.

The randomized approximation $R^\odot$ substantially outperformed both $(R^\bullet)^+$ and $(R^+)^\bullet$ in all cases. The approximation $(R^\bullet)^+$, as expected due to Proposition 2, performed better than $(R^+)^\bullet$, but the differences were much smaller.

We would like to point out that this study still does not say much about the relationship of $R^\odot$ to any optimal approximation (as defined in Section 3 and [13]). The approximation $R^\odot$ is a randomized algorithm, but its theoretical relationship to optimal approximations is an open problem not discussed in this paper. Assume that R^{opt} is an optimal approximation of R, and $sim_J(R,R^{opt}) = \alpha, sim_{AWA}(R,R^{opt}) = \beta$. From the results of this paper. all we can derive is that on average, for sparse graphs: $\alpha \geq 0.3, \beta \geq 0.58$, for medium density graphs: $\alpha \geq 0.41, \beta \geq 0.53$, and for dense graphs $\alpha \geq 0.49, \beta \geq 0.52$. We also can say that, on average, both $(R^\bullet)^+$ and $(R^+)^\bullet$ are rather far away from optimal approximations. Most likely because, in both cases, we are removing much more cycles than might be needed.

Acknowledgment

This research has been partially supported by a Discovery NSERC grant of Canada. The authors gratefully acknowledge three anonymous referees for their helpful comments.

References

1. Baioletti M., Coletti G. and Petturiti D.,Weighted Attribute Combinations Based Similarity Measures, *Communications in Computer and Information Science*, 299 (2012), 211–220.
2. Cormen T. H., Leiserson C. E., Rivest D. L., Stein C., *Introduction to Algorithms*, MIT Press 2001.
3. Coletti G., Bouchon-Meunier B., A study of similarity measures through the paradigm of measurement theory: the classic case, *Soft Computing*, 23 ,16 (2019), 6827–6845.
4. Coletti G., Petturiti D., Bouchon-Meunier B., Weighted and Choquet L^p distance representation of comparative dissimilarity relations on fuzzy description profiles, *Annals of Mathematics and Artificial Intelligence*, 2024, doi: 10.1007/s10472-024-09924-y.
5. Deza M. M., Deza E., *Encyclopedia of Distances*, Springer, Berlin 2012.
6. Erdös P., Rényi A., On Random Graphs, *Publicationes Mathematicae* 6 ,3-4 (1959) 290–297.
7. Fishburn, P. C., *Interval Orders and Interval Graphs*, J. Wiley, New York 1985.
8. French S., *Decision Theory*, Ellis Horwood, New York 1986.

9. Jaccard P., Étude comparative de la distribution florale dans une portion des Alpes et des Jura, *Bulletin de la Société Vaudoise des Sciences Naturalles* 37 (1901) 547-549.
10. Janicki R., Pairwise Comparisons Based Non-Numerical Ranking, *Fundamenta Informaticae* 94, 2 (2009), 197-217.
11. Janicki R., Approximations of Arbitrary Binary Relations by Partial Orders. Classical and Rough Set Models, *Transactions on Rough Sets*, 13 (2011), 17-38.
12. Janicki R., Property-Driven Rough Sets Approximations of Relations, In A. Skowron, Z. Suraj (eds.), Rough Sets and Intelligent Systems - Professor Zdzisław Pawlak in Memoriam (Series: *Intelligent Systems Reference Library Vol. 42*), Springer 2013, pp. 333-357.
13. Janicki R., Approximations of Arbitrary Relations by Partial Orders, *International Journal of Approximate Reasoning*, 98 (2018) 177–195.
14. Janicki R., Lenarčič A., Optimal Approximations with Rough Sets and Similarities in Measure Spaces, *International Journal of Approximate Reasoning*, 71 (2016) 1-14.
15. Janicki R., Zhai Y., On a Pairwise Comparison Based Consistent Non-Numerical Ranking, *Logic Journal of IGPL*, 20, 4, (2012) 667-676.
16. Kemeny, J. G., Mathematics without numbers, *Daedalus,* 88, 571—591 (1959)
17. Kleijn, J., Koutny, M., Formal Languages and Concurrent Behaviour. *Studies in Computational Intelligence* 113, 125–182 (2008)
18. Marczewski E., Steinhaus H., On a certain distance of sets and corresponding distance of functions, *Colloquium Mathematicum*, 4 (1958) 319-327.
19. Marsala C., Petturiti D., Vantaggi B., Adding Semantics to Fuzzy Similarity Measures Through the d-Choquet Integral, *Lecture Notes in Computer Science*, 14294 (2023), 386–399.
20. Rezai H, Emoto M, Mukaidono M., New Similarity Measure Between Two Fuzzy Sets, *Journal of Advanced Computational Intelligence and Intelligent Informatics* 10, 6 (2006) 946-953.
21. Rosen K. H., *Discrete Mathematics and Its Applications*, McGraw-Hill, New York 1999.
22. Schröder E., *Algebra der Logik*, Teuber, Leipzig 1895.
23. Tversky A., Features of Similarity, *Psychological Reviews* 84, 4 (1977) 327–352.
24. Yang N., Evaluating the Efficacy of Approximating Arbitrary Relations by Partial Orders, Master Thesis, McMaster University, Hamilton, Ontario, Canada, December 2023.