International Journal of Approximate Reasoning 98 (2018) 177–195

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Approximations of arbitrary relations by partial orders

Ryszard Janicki

Department of Computing and Software, McMaster University, Hamilton L8S 4K1, Canada

ARTICLE INFO

Article history:
Received 7 October 2017
Received in revised form 9 February 2018
Accepted 27 April 2018
Available online 2 May 2018

Keywords:
Approximation
Relation
Partial order
NP-completeness
Rough Sets

ABSTRACT

The problem of optimal quantitative approximation of an arbitrary binary relation by a partial order is discussed and some solutions are provided. It is shown that even for a very simple quantitative measure the problem is NP-hard. Some quantitative metrics are also applied for known property-driven approximations by partial orders. Some relationship to Rough Sets is discussed.

1. Introduction

A motivation for this kind of work has been clearly described in [12]:

> "Consider the following problem: 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?"

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. [8,11,15,17].

Defining what is the 'best' partially ordered approximation is itself a problematic task. It could be approached in at least two ways.

One approach is just to propose some similarity metrics for binary relations and then just choose a partial order that is closest to a given arbitrary relation. This is the main subject of this paper. The first question is how these similarity metrics should be defined. Should we look for some generic similarity measure between arbitrary relations, or should we take into account that one of the relations is always a partial order and include this fact into the definition of similarity? Partial ordering means acyclity and transitivity, should our similarity measures also make this distinction? From the application point of view, the roles of acyclity and transitivity are different. Lack of transitivity may not be an error at all, it could just be a fact that a given data set is of minimum (or optimal, sufficient) size (as Hasse diagrams, directed acyclic graphs or dags [24], or direct causality graphs [20], etc.). Acyclity on the other hand is usually an indication of some errors. Moreover, the relation that we are going to approximate is not absolutely arbitrary. It represents, with perhaps some errors or incompleteness, some real data or phenomena. If its partially ordered approximation is chosen only on the basis of some numerical calculations, some structural properties of the original relation might be lost or wrongly changed.

E-mail address: janicki@mcmaster.ca.

https://doi.org/10.1016/j.ijar.2018.04.012
0888-613X/

The other orthogonal approach is not to use quantitative metrics at all. This approach is called *property-driven* and is based on the assumption that partial order approximations must satisfy certain properties. It stems from the 1895 paper by Schröder [25]. As partial orders, the approximations must be acyclic and transitive, but they also should satisfy some other properties. In [12] property-driven partial order approximations of an arbitrary binary relation were provided and discussed in both the classical algebraic model and the Rough Sets settings [23]. No quantitative metrics were used in [12].

In this paper we propose two simple metrics for measuring similarity and difference between relations, and a definition of *optimal* approximation. We also provide some justification of both metrics and the definition. One metric is just a simple adaptation of a metric used for sets, the other is a special modification designed specially for relations. We show that, at least for approximations by partial orders, both metrics lead to identical results.

In [12] and [13], a special attention is paid to two partially ordered approximations of R, denoted by $(R^\bullet)^+$ and $(R^+)^\bullet$ for a given relation R. Using graph terminology, $R^\bullet$ is derived from R by erasing all arcs from all strongly connected components (or equivalently, removing all arcs from all cycles). The relation R^+ is a transitive closure of R. The relation $(R^+)^\bullet$ is a classical approximation, first proposed by Schröder in 1895 [25], which is often regarded as 'the' partially ordered approximation. We will show that with respect to our metrics, $(R^\bullet)^+$ is a better approximation of R than the well known Schröder's $(R^+)^\bullet$.

We will also show that finding quantitative optimal approximation, with respect to simple metrics proposed in this paper, is NP-hard.

Finally, we will argue that while an arbitrary quantitative optimal approximation (with any reasonable metrics) might somehow be inconsistent with property-driven approximations of [12,13], our model satisfies most properties required from property-driven approximations.

We will also show how the presented model relates to the Rough Sets approach for specialized relational approximations.

This paper is a substantially extended, revised and corrected version of the conference paper [14].

The paper is organized as follows. In Section 2 we recall the basic notions of the theory of relations, directed graphs and partial orders. The basic concepts of similarity and distance for relations that are used in this paper, are presented in Section 3. The problems encountered when trying to define the concept of optimal approximation are discussed in Section 4, while Section 5 is devoted to the property-driven partial orders approximations of [12,13]. Quantitative properties of property-driven partial order approximations $(R^\bullet)^+$ and $(R^+)^\bullet$ are presented in Section 6. Some intuitions and properties that led to our concept of optimal partial order approximation are analyzed in Section 7. The main result of this paper, namely introduction and discussion of partial order approximations based on absolute similarity and distance, are presented in Section 8. Another version of a distance for relations is proposed and its properties are discussed in Section 9. In Section 10 the main results of the paper are presented in Rough Sets setting, and Section 11 contains final comments.

2. Relations, directed graphs and partial orders

In this section we recall some fairly known concepts and results that will be used later in this paper [3,7,24].

Let X be a set. *We assume all sets considered in this paper are finite.* Note that every relation $R \subseteq X \times X$ can be interpreted as a *directed graph* $G_R = (V, E)$ where $V = X$ is the set of vertices and $E = R$ is the set of edges (cf. [3]).

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

We write $a \sim_< b$ if $\neg(a < b) \wedge \neg(b < a)$, that is if a and b are either *distinctly incomparable* (w.r.t. $<$) or *identical* elements. We also write

$$a \equiv_< b \iff (\{x \mid a < x\} = \{x \mid b < x\} \ \wedge \ \{x \mid x < a\} = \{x \mid x < b\}).$$

The relation $\equiv_<$ is an *equivalence relation* (i.e. it is reflexive, symmetric and transitive) and it is called *the equivalence with respect to* $<$, since if $a \equiv_< b$, there is nothing in $<$ that can distinguish between a and b (cf. [7]). We always have $a \equiv_< b \implies a \sim_< b$.

- Let $\mathbb{PO}(X)$ denote the set on all partial orders included in $X \times X$.

For every relation $R \subseteq X \times X$, we define $R^0 = Id_X = \{(a, a) \mid a \in X\}$, the identity relation on X, and $R^{i+1} = R^i R$ for all $i \geq 0$. Furthermore the relation $R^+ = \bigcup_{i=1}^{\infty} R^i$ is called the *transitive closure* of R, the relation $R^{-1} = \{(b, a) \mid (a, b) \in R\}$ is called the *inverse* of R, and a relation R is *acyclic* if and only if $\neg xR^+x$ for all $x \in X$. In graph terminology, if R is acyclic then G_R is DAG (*Directed Acyclic Graph*), while if for all $x \in X$ we have xR^+x then the graph G_R is *strongly connected*. Also for a given relation R and $a \in X$, we define: $aR = \{x \mid aRx\}$ and $Ra = \{x \mid xRa\}$.

For every relation R we can define the relations R^{cyc}, $R^\bullet$ and $\equiv_R$ as follows

- $aR^{cyc}b \iff aR^+b \wedge bR^+a$,
- $aR^\bullet b \iff aRb \wedge \neg(aR^{cyc}b)$, i.e. $R^\bullet = R \setminus R^{cyc}$,
- $a \equiv_R b \iff aR = bR \wedge Ra = Rb$.

In [12,13], the relation $R^\bullet$ is called an *acyclic refinement of* R. In graph terminology, if $aR^{cyc}b$ then a and b are *strongly connected* in G_R, and 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. The relation $\equiv_R$ is an extension of $\equiv_<$ for an arbitrary relation R and it was proposed in [12]. In principle, similarly as for partial orders, if $a \equiv_R b$, then there is nothing in R that can distinguish between a and b (with respect to R).

Lemma 1.

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* $\equiv_R \subseteq \equiv_{R^\bullet}$.

Proof. (1), (2) and (3) From their definitions.
(4) From Lemma 5 of [12]. □

The above corollary presents basic hints for property-driven partial order approximations as proposed in [12,13] and discussed later in Section 5. 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.

Let $|X| = n$. Given R, the complexity of calculating R^+ is $O(n^3)$ (Floyd–Warshall Algorithm). Calculating $R^\bullet$ comprises of finding all strongly connected components of G_R (Tarjan Algorithm can be used) and then deleting all edges from all strongly connected components so the time complexity is $O(|X| + |R|) = O(n^2)$ (cf. [3,26]).

3. Absolute similarity and distance for relations

Let R and S be two relations on X and $G_R = (X, R)$, $G_S = (X, S)$ their appropriate graph representations. Without losing any generality we may assume that both R and S are irreflexive, i.e. $(a, a) \notin R \cup S$ for any $a \in X$. How can we measure a difference or similarity between R and S? One possibility is just to count common edges of the graphs G_R and G_S, which leads to

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

We will call $sim(R, S)$ an *absolute similarity* between relations R and S. We added *absolute* to distinguish it from *similarity* as formally defined for instance in [16,28].

The other possibility is to count the edges that were removed from R and the edges that were added to R to get S. In this case we can define:

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

We will call $dist(R, S)$ an *absolute distance* between relations R and S. Note that $dist(R, S)$, which is just a cardinality of *symmetric difference* [24], is a proper *metric*, i.e. the 'triangle property': $dist(R, S) \leq dist(R, T) + dist(T, S)$ is satisfied for all relations R, S, T (cf. [4]).

We will later argue that if R is an arbitrary relation and S is a partial order, i.e. S is acyclic and transitive; $sim(R, S)$ better measures the relationship between R and S with respect to acyclity, while $dist(R, S)$ measures this relationship better with respect to transitivity, and we cannot replace one by another.

In our case, R and S are relations, but the symmetric difference is defined for general sets. It is often called Fréchet–Nikodym–Aronszyan distance [22] if arguments are general sets, or Kemeny distance [19] if arguments are relations, as in our case (cf. [4]), and the symmetric difference between two sets is often considered a measure of how "far apart" they are [4]. In particular $dist(R, \emptyset) = |R|$.

When we scale both *sim* and *dist* to $[0, 1]$, we get well known and popular *Jaccard similarity* and *Jaccard distance* [10]: $sim_J(R, S) = \frac{|R \cap S|}{|R \cup S|}$ and $dist_J(R, S) = \frac{|R \cup S| - |R \cap S|}{|R \cup S|} = 1 - sim_J(R, S)$. The function $dist_J(R, S)$ is also in general defined for arbitrary sets and it is also a proper metric (cf. [4,22]). However in this paper we will use unscaled measures *sim* and *dist* instead of Jaccard indexes. If $R \cap S = \emptyset$ then $dist_J(R, S) = 1$, for *all* R and S, which would be some oversimplification in our approach. In the Jaccard model, similarity uniquely defines distance and vice versa, but we do not want this relationship in our model. In our approach the relation S is a partial order and we will show that, if S is a partial order, $sim(R, S)$ measures different aspects of approximation than $dist(R, S)$. Moreover, both Jaccard indexes are meaningless when $S = \emptyset$, and $\emptyset$, i.e. the empty set, is a valid and useful partial order.

If S is interpreted as some approximation of R, we may use:

$$dist_{sim}(R, S) = |R| - sim(R, S) = |R| - |R \cap S| = |R \setminus S|$$

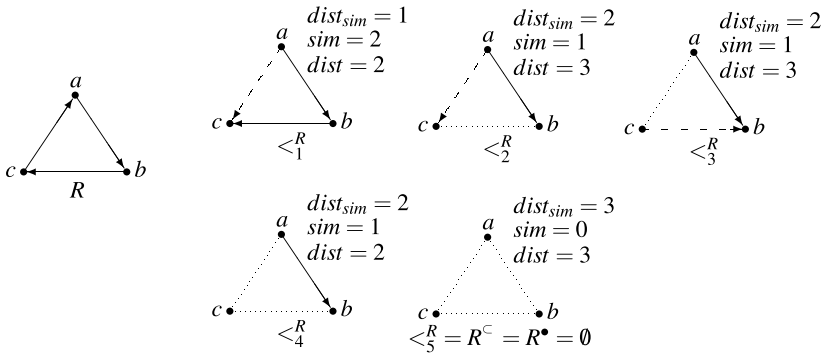

Fig. 1. An example of a relation R and its all potential partial order approximations (up to isomorphism). Dashed edges are added, dotted lines represent incomparability. For example $R \setminus <_1^R = \{(c,a)\}$ and $<_1^R \setminus R = \{(a,c)\}$, so $dist(R, <_1^R) = 2$.

as a *measure of closeness* of S to R. If $R = S$ then $dist_{sim}(R,S) = 0$. As opposed to $sim(R,S)$ and $dist(R,S)$, $dist_{sim}(R,S)$ is *asymmetric*, and it is assumed that S is an approximation of R.

The following result provides simple properties of *sim* and *dist* for some specialized relations.

Lemma 2. *Let R, S, Q be relations on X.*

1. *If $R \subseteq S$ then: $dist(R,S) = |S| - |R| = |S| - sim(R,S) = dist_{sim}(S,R)$.*
2. *If $sim(R,S) = sim(R,Q)$ and $S \subseteq Q$ then:*
   (a) *$dist(R,S) \leq dist(R,Q)$,*
   (b) *$dist(R,S) = dist(R,Q) \iff S = Q$.*
3. *If $R^{cyc} = \emptyset$ then*

$$dist(R, R^+) = \min\{dist(R,<) \mid < \in \mathsf{PO}(R)\} = \max\{sim(R,<) \mid < \in \mathsf{PO}(R)\},$$

where $\mathsf{PO}(R) = \{ < \mid R \subseteq < \wedge < \in \mathbb{PO}(X)\}$.

Proof. (1) Since in this case $R \cup S = S$ and $R \cap S = R$.

(2a) $sim(R,S) = sim(R,Q)$ means $|R \cap S| = |R \cap Q|$. Hence we have $dist(R,S) = |R \cup S| - |R \cap S| = |R \cup S| - |R \cap Q| \subseteq |R \cup Q| - |R \cap Q| = dist(R,Q)$.

(2b) Only $\Longrightarrow$ needs to be shown. $dist(R,S) = dist(R,Q)$ means $|R \cup S| = |R \cup Q|$. Since $S \subseteq Q$ this implies $S = Q$.

(3) Since $R^+ \in \mathbb{PO}(X)$ and for every $< \in \mathbb{PO}(X)$ such that $R \subseteq <$, $R^+ \subseteq <$ (cf. [7,24]), so $dist(R,R^+) = |R^+| - |R| \leq dist(R,<) = | < | - |R|$. Since $R \subseteq R^+$, by (1) of this lemma, $dist(R,R^+) = |R^+| - sim(R,R^+)$. □

Lemma 2(1) simply states that if $R \subseteq S$ then minimal $dist(R,S)$ corresponds to maximal $sim(R,S)$ and vice versa. The case (2) says that when one argument is fixed, the function $dist(R,x)$ is strictly monotone. The third case characterizes transitive closures of acyclic relations. It states that in this case the transitive closure is the optimal approximation with respect to both *dist* and *sim*.

4. Problems with optimal approximation

Let $R \subseteq X \times X$ be an arbitrary relation. It is tempting to say that a partial order $<^R$ on X is the best partial order approximation of R if $sim(R, <^R)$ is *maximal* for all partial orders on X, and/or if $dist(R, <^R)$ is *minimal* for all partial orders on X. However such straightforward approach may lead to unexpected and maybe undesired results.

Consider the relation R from Fig. 1. There are five non-isomorphic partial orders on the three element set $\{a,b,c\}$, and they are named $<_i^R$, $i = 1,\ldots,5$, in Fig. 1. If only values of *sim* and *dist* are taken into account, a partial order $<_1^R$ (or *any order isomorphic to it*) is an optimal or best approximation, as for all partial orders $<$ on $\{a,b,c\}$, $sim(R,<) \leq 2$, $dist(R,<) \geq 2$, and $sim(R,<_1^R) = 2$, $dist(R,<_1^R) = 2$. The order $<_1^R$ had been obtained by 'flipping' edge (c,a) of G_R, two additional isomorphic orders can be obtained by 'flipping' edges (a,b) and (b,c) respectively. In all three cases the values of *sim* and *dist* are the same as for $<_1^R$. But why have we chosen 'flipping' (c,a)? Why not (a,b) or (b,c)? Are we allowed to flip at all without seriously alternating input data, especially if choice of what to flip appears to be random?

In decision and ranking theories, where outcomes are expected to be weak (stratified) partial orders, cycles in input relation R are usually interpreted as *indifference* or *incomparability* [8,11]. With this interpretation, $<_5^R$ would be considered as the *only* acceptable partially ordered approximation of R, but $sim(R,<_5^R) = 0$ and $dist(R,<_5^R) = 3$, so according to the values of *sim* and *dist*, the partial order $<_5^R$ is the *worst* partial order approximation.

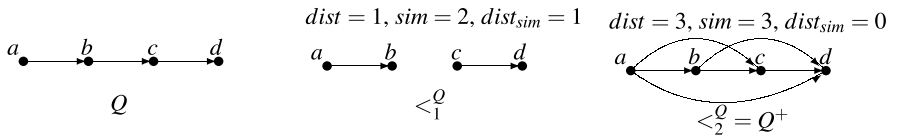

Fig. 2. A relation Q and its two potential partial order approximation. Note that $Q^{\subset} = Q^{cyc} = \emptyset$ and $Q^{\bullet} = Q$. The picture describing Q is *not* a Hasse diagram, it describes the full relation Q, so Q is *not* a partial order!

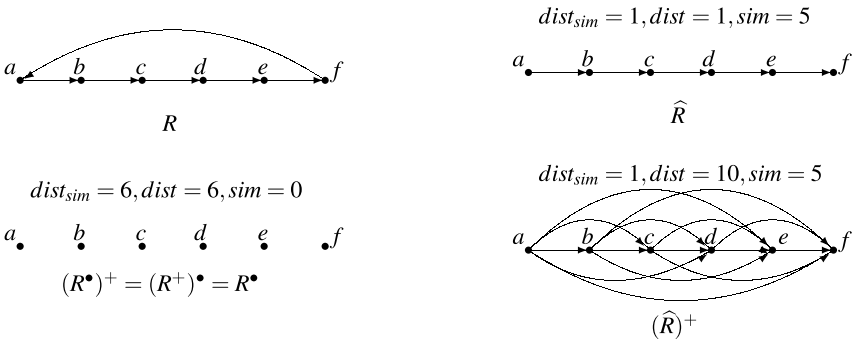

Fig. 3. A relation R, one of its optimal acyclic approximation $\widehat{R}$, and its two potential partial order approximations: $(R^{\bullet})^{+}$ and $(\widehat{R})^{+}$.

Consider now the relation Q from Fig. 2, which is acyclic but not a partial order. It can be shown by inspection that for all partial orders over the set $\{a, b, c, d\}$ the value $dist(Q, <_1^Q) = 1$ is minimal. The partial order $<_1^Q$ resulted from deleting the edge (b, c) from the graph G_Q. But why should we delete (a, b), i.e. make a and b incomparable (recall that Q represents empirical data but the nature of problem demands that Q should be a partial order)? The relation Q is acyclic but it lacks transitivity, which most likely results from the fact that the empirical data represented by Q are incomplete, or because providing explicit transitivity was considered unneeded. In this case the most natural and proper way to transform Q into an appropriate partial order is to compute Q^{+}, the transitive closure of Q. We have $dist(Q, <_2^Q) = 3 > 1 = dist(Q, <_1^Q)$, but on the other hand $sim(Q, <_2^Q) = 3 > 2 = sim(Q, <_1^Q)$ and additionally $dist_{sim}(Q, <_2^Q) = 0$. So, which approximation is better, $<_1^Q$ or $<_2^Q$?

The situation that, for some S_1 and S_2, $sim(R, S_1) > sim(R, S_2)$, so S_1 is a better approximation of R with respect to $sim(\ldots)$, and $dist(R, S_2) < dist(R, S_1)$, so S_2 is a better approximation of R with respect to $dist(\ldots)$, occurs quite often. Consider the relation R from Fig. 3. The graph G_R is strongly connected and $\widehat{R}$ is one of the optimal acyclic approximations of R, while $(R^{\bullet})^{+}$ and $(\widehat{R})^{+}$ are two partial orders that can be regarded as partial order approximations of R. We have $sim(R, (\widehat{R})^{+}) > sim(R, (R^{\bullet})^{+})$ while $dist(R, (R^{\bullet})^{+}) < dist(R, (\widehat{R})^{+})$. The relation $(\widehat{R})^{+}$ has the same problem as the relation $<_1^R$ of Fig. 1, $\widehat{R}$ is one of six different optimal acyclic approximations, it resulted by removing the edge (f, a). But why this edge, not any other?

Consider $Q = \widehat{R}$, where both R and $\widehat{R}$ are these from Fig. 3. In this case $Q^{\bullet} = Q$, $Q^{+} = (Q^{\bullet})^{+} = (Q^{+})^{\bullet}$ and $Q^{+} = (\widehat{R})^{+}$. Moreover, $sim(Q, Q^{+}) = 5$, $dist_{sim}(Q, Q^{+}) = 0$ and $dist(Q, Q^{+}) = 9$. Since $dist_{sim}(Q, Q^{+}) = 0$, Q^{+} is the best partial order approximation of Q with respect to measure $sim(\ldots)$.

These results indicate that using only $sim(R, S)$ and/or $dist(R, S)$ (or any particular numerical measure in fact) may not be sufficient when we are looking for a proper optimal approximation. We also have to preserve properties that all partial orders posses. An approach where we manipulate rather a given relation R to get desirable properties instead of just measuring differences is called *property-driven approximation* [13].

5. Property-driven partial order approximations

In [12] the definitions of property driven partial order approximation and weak partial order approximation were presented and discussed. Ranking process, pairwise comparisons paradigm [11,17] and the results of [27,30] provided the main motivation for interpretation of partial orders in these definitions. The definition below is a slightly modified and rephrased version of definitions proposed in [12] and used in [13].

Definition 1 *([12]).* 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. $a < b \implies \neg a R^{cyc} b$ (or, equivalently $a < b \implies \neg b R^+ a$). □

The conditions of Definition 1 were motivated by the following intuitions [13]. 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 (cf. [13]). 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 from $aR^{cyc}b$ it is interpreted that a and b are indifferent [8]. Similar interpretations take place in concurrency theory [20].

The following result characterizing property-driven partial order approximations has been proven in [12] (they hold for new version of Definition 1 too).

Theorem 1 *([12]).*

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

Unfortunately, the strong but well justified conditions for approximations by partial orders employed in Definition 1 do not always play well with quantitative similarity and distance measures proposed in Section 3.

Only the condition (1) of Definition 1 is satisfied by all examples from Figs. 1–4. The partial orders $<_i^R$, $i = 1, \ldots, 4$ of Fig. 1 *do not* satisfy the condition (4) of Definition 1. For example $b <_1^R a$ and bR^+a. From Fig. 1, only $<_5^R$ satisfies the condition (4) of Definition 1. The condition (2) implies that if R is acyclic, i.e. G_R is DAG, then $R \subseteq <$. The partial order $<_1^Q$ from Fig. 2 *does not satisfy* the condition (2), as $(b, c) \in Q = Q^\bullet$ but $(b, c) \notin <_1^Q$. On the other hand, the partial order $<_2^Q$ from Fig. 2 satisfies all four conditions of Definition 1.

All these examples indicate that strict *property-driven* partial order approximations in the style of [12,13] do not fit well to numerical estimations given by the functions $sim(\ldots)$ and $dist(\ldots)$ proposed in the previous section. To implement quantitative approximations and make them somehow consistent with property-driven approximations we have to weaken the required conditions that dealt with preserving appropriate properties.

However, we may use quantitative estimations for established property-driven approximations as $(R^\bullet)^+$ and $(R^+)^\bullet$.

6. Quantitative properties of $(R^\bullet)^+$ and $(R^+)^\bullet$

In this section we will apply measures $sim(\ldots)$, $dist(\ldots)$ and Jaccard index $sim_J(\ldots)$ to well established property-driven approximations $(R^\bullet)^+$ and $(R^+)^\bullet$. We will start with characterization of their intersections with a given relation R.

Lemma 3. *For every relation* $R \subseteq X \times X$, *we have:* $R \cap (R^\bullet)^+ = R \cap (R^+)^\bullet = R^\bullet$.

Proof. Since $(R^\bullet)^+ \subseteq (R^+)^\bullet$, then $R \cap (R^\bullet)^+ \subseteq R \cap (R^+)^\bullet$. Assume $(a, b) \in R \cap (R^+)^\bullet$. From the definition of acyclic refinement '$\bullet$', we have

$$(a, b) \in (R^+)^\bullet \iff (a, b) \in R^+ \wedge (b, a) \notin R^+.$$

Hence: $(a, b) \in R \cap (R^+)^\bullet \iff (a, b) \in R \wedge (a, b) \in R^+ \wedge (b, a) \notin R^+ \iff (a, b) \in R \wedge (b, a) \notin R^+ \iff (a, b) \in R^\bullet$. Hence $R \cap (R^\bullet)^+ \subseteq R \cap (R^+)^\bullet = R^\bullet$. On the other hand $R^\bullet \subseteq R$ and obviously $R^\bullet \subseteq (R^\bullet)^+$, so $R \cap (R^\bullet)^+ = R \cap (R^+)^\bullet = R^\bullet$. □

We may now formulate the main result of this section.

Proposition 1. *For every relation* R, *we have:*

1. $sim(R, (R^\bullet)^+) = sim(R, (R^+)^\bullet) = |R^\bullet|$,
2. $dist(R, (R^\bullet)^+) \leq dist(R, (R^+)^\bullet)$,
3. $sim_J(R, (R^\bullet)^+) \geq sim_J(R, (R^+)^\bullet)$.

Proof. (1) A consequence of Lemma 3.

(2) Since $(R^\bullet)^+ \subseteq (R^+)^\bullet$, then $|R \cup (R^\bullet)^+| \leq |R \cup (R^+)^\bullet|$. By 3, $|R \cap (R^\bullet)^+| = |R \cap (R^+)^\bullet|$. Hence $dist(R, (R^\bullet)^+) = |R \cup (R^\bullet)^+| - |R^\bullet| \leq |R \cup (R^+)^\bullet| - |R^\bullet| = dist(R, (R^+)^\bullet)$.

(3) Since $|R \cap (R^\bullet)^+| = |R \cap (R^+)^\bullet|$ and $(R^\bullet)^+ \subseteq (R^+)^\bullet$, then $sim_J(R, (R^\bullet)^+) = \frac{|R \cap (R^\bullet)^+|}{|R \cup (R^\bullet)^+|} \geq \frac{|R \cap (R^+)^\bullet|}{|R \cup (R^+)^\bullet|} = sim_J(R, (R^+)^\bullet)$. □

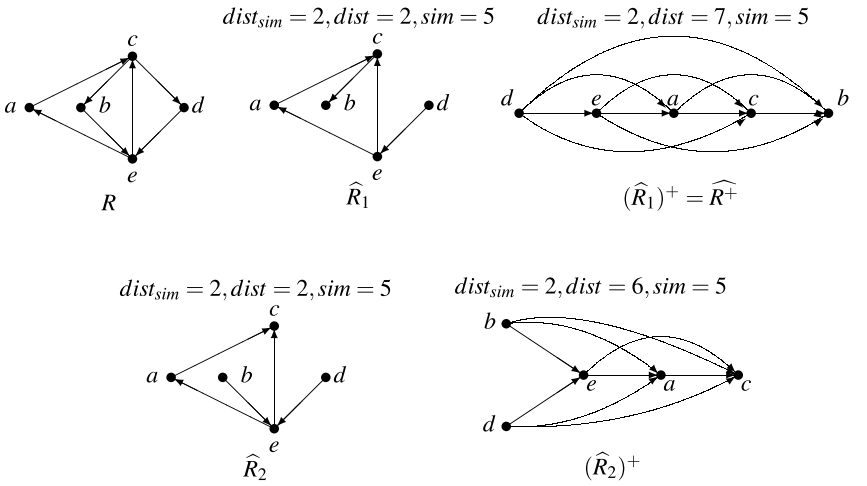

Fig. 4. The relations $\widehat{R}_1$ and $\widehat{R}_2$ are examples of two different *optimal* acyclic approximations of the relation R, and $\widehat{R^+}$ is an optimal acyclic approximation of R^+. The relation $(\widehat{R}_2)^+$ can be regarded as an optimal partial order approximation of R, but $(\widehat{R}_1)^+$ can not.

It appears that with respect to numerical similarity and distance measures, including the Jaccard index (and all indexes consistent with Jaccard index, as defined in [16]), the relation $(R^\bullet)^+$ seems to be a better approximation than the classical Schröder's approximation $(R^+)^\bullet$.

While the time complexity of calculating $(R^\bullet)^+$ and $(R^+)^\bullet$ is $O(n^3)$ for both relations (as calculating transitive closure is a dominating factor in both cases), practical time complexity is always smaller for $(R^\bullet)^+$ as $|R^\bullet| \leq |R|$.

7. Towards optimal approximations

Property-driven approximations of [12,13] are based on an observation that complex properties, as partial orders, are seldom enforced by 'natural' well defined, but structurally complex, operators (cf. [13,30]). On the other hand, many 'simple' or 'elementary' properties such as acyclity, symmetry, transitivity, etc., can easily be enforced by using kernels and closures (cf. [1]), or specialized lower and upper approximations [11,13,30]. In particular, $(R^\bullet)^+$ is the result of applying acyclic refinement (which is a lower approximation or a kernel) first, and transitive closure (i.e. upper approximation) later; while for $(R^+)^\bullet$ we proceed in the opposite order.

The above reasoning, supported by Theorem 1 and Lemma 2(3), suggests the following approach for finding optimal partial order approximation. For a given arbitrary relation R first find its optimal acyclic approximation $\widehat{R}$ and then construct its transitive closure[1] $\widehat{R}^+$. This would be mimicking the property-driven partial order approximation $(R^\bullet)^+$. Orthogonally, we can calculate the transitive closure R^+ first then find an optimal acyclic approximation of R^+, i.e. $\widehat{R^+}$. This would be mimicking the pattern for the property-driven partial order approximation $(R^+)^\bullet$.

An optimal acyclic approximation can be defined as follows. A relation $\widehat{R}$ is an *optimal acyclic approximation* of R if and only if $\widehat{R} \subseteq R$, $\widehat{R}^{cyc} = \emptyset$ and $sim(R, \widehat{R})$ is maximal, or equivalently, since $\widehat{R} \subseteq R$, if $dist(R, \widehat{R})$ is minimal. An optimal acyclic approximation is usually not unique. For the relation R from Fig. 1 deleting an arbitrary arc result in an optimal acyclic approximation, so in this case we have three optimal acyclic approximations of R.

Fig. 4 illustrates both procedures described above. Both $\widehat{R}_1$ and $\widehat{R}_2$ are two different optimal acyclic approximations of R. There are actually six different optimal cyclic approximations of R, $\widehat{R}_3$: obtained from R by deleting arcs (c, b) and (d, e), $\widehat{R_4}$: obtained by deleting arcs (c, b) and (c, d), $\widehat{R}_5$: by deleting (e, c) and (e, a), and $\widehat{R}_6$: by deleting (e, c) and (a, c). It can be shown that $(\widehat{R}_2)^+$, $(\widehat{R_4})^+$ and $(\widehat{R}_5)^+$ are better partial order approximations with respect to *dist* than $(\widehat{R}_1)^+$, $(\widehat{R}_3)^+$ and $(\widehat{R}_6)^+$. We have $dist(R, (\widehat{R}_i)^+) = 5$ for $i = 2, 4, 5$, $dist(R, (\widehat{R}_3)^+) = 6$ and $dist(R, (\widehat{R}_i)^+) = 7$ for $i = 1, 3$.

For the orthogonal procedure, i.e., calculating the transitive closure first, for the relation R from Fig. 4, we have $R^+ = X \times X$, where $X = \{a, b, c, d, e\}$. It can be shown that any total order on X is an optimal acyclic approximation of $X \times X$, so for example $\widehat{R^+}$ from Fig. 4 is such an approximation. In this case $dist(R, (\widehat{R}_2)^+) < dist(R, \widehat{R^+})$. We will show later that this relationship, with '$<$' replaced with '$\leq$' is a general rule, i.e. starting with finding optimal acyclic approximation and applying transitive closure next leads to better results than starting with transitive closure and finding optimal acyclic approximation of this closure next.

[1] There is only one transitive closure so it is already 'optimal'.

8. Approximations based on absolute similarity and distance

In Section 4 we have discussed problems related to quantitative optimal approximation. In this section a solution, based on analysis from previous sections, is proposed. This section also contains the main results of this paper. We start with a definition of an *optimal simple partial order approximation* of a given relation R. A relation is a partial order when it is acyclic and transitive and in our approach we will use different quantitative penalties for not having these properties.

Definition 2. For every relation R on X, a partial order $R^{\oplus}$ on X is an *optimal simple partial order approximation* of R if and only if the following conditions are satisfied:

1. $R^{\bullet} \subseteq R^{\oplus}$,
2. $sim(R, R^{\oplus}) = \max\{sim(R, <) \mid < \in \mathbb{PO}(X)\}$,
3. $dist(R, R^{\oplus}) = \min\{dist(R, <) \mid sim(R, <) = sim(R, R^{\oplus})\}$.

Let $\mathsf{POA}(R)$ denote the set of all optimal simple partial order approximations of R. □

The condition (1) above defines a lower bound. The relation $R^{\bullet}$ is R with all cycles removed and is considered as an absolutely necessary part of $R^{\oplus}$.

Lack of acyclicity is considered a bigger problem than lack of transitivity. The latter could be intensional (cf. Hasse diagrams, dependency graphs, etc. [5,8,20]), the former is usually a serious error [8,11]. The conditions (2) and (3) of Definition 2 capture this asymmetry by making absolute similarity the dominant measure and absolute distance the secondary measure.

Definition 2 does not explicitly provide any requirements for `the` upper bound of $R^{\oplus}$. However we will show later that the conditions (1) and (2) imply the following intuitive upper bound, namely: $R^{\oplus} \subseteq R^{+}$.

We call this approximation '*simple*' as most of the properties from Definition 1 are no longer required. They are just too restrictive for quantitative optimization.

We will consider two distinct cases:

Case 1. R is acyclic, i.e. $R^{cyc} = \emptyset$.
Case 2. R contains a cycle, i.e. $R^{cyc} \neq \emptyset$.

Deciding between these two cases can be done in $O(n)$ time, where $n = |X|$, by using *depth first traversal* on the graph $G_R = (X, R)$ [3].

The case 1 is simple, one just has to use transitive closure.

Proposition 2. *If R is acyclic, then $R^{\oplus} = R^{+}$.*

Proof. For any partial order $<$ containing R the condition (2) of Definition 2 is satisfied as then $dist_{sim}(R, <) = 0$ and $sim(R, <) = |R|$. Since R^{+} is the smallest partial order containing R (cf. [7]), then, by Lemma 2(3), the condition (3) of Definition 2 is satisfied too. □

The second case involves removing cycles and it is much more complex.

For every relation $R \subseteq X \times X$, let

$$\mathbb{MDAG}(R) = \{\widehat{R} \mid sim(R, \widehat{R}) = \max\{sim(R, S) \mid S \subseteq R \wedge S^{cyc} = \emptyset\}\}.$$

The elements of $\mathbb{MDAG}(R)$ are maximal (with respect to number of arcs) directed acyclic graphs included in R.

We will now recall the concept of *feedback edge set* for directed graphs. Let $G = (V, E)$ be a directed graph.

- A *feedback edge set* is a subset $F \subseteq E$ such that every cycle of G contains a vertex in F [3].

Theorem 2 *(Karp 1972 [18]). Minimum feedback edge set problem is NP-complete.* □

It turns out the minimum feedback edge problem is involved in effectively finding the elements of $\mathbb{MDAG}(R)$.

Lemma 4.

1. *For every $S \in \mathbb{MDAG}(R)$, the relation $R \setminus S$, interpreted as a set edges, is a minimum feedback edge set of the graph $G_R = (X, R)$.*
2. *Constructing an element of $\mathbb{MDAG}(R)$ is NP-complete.*

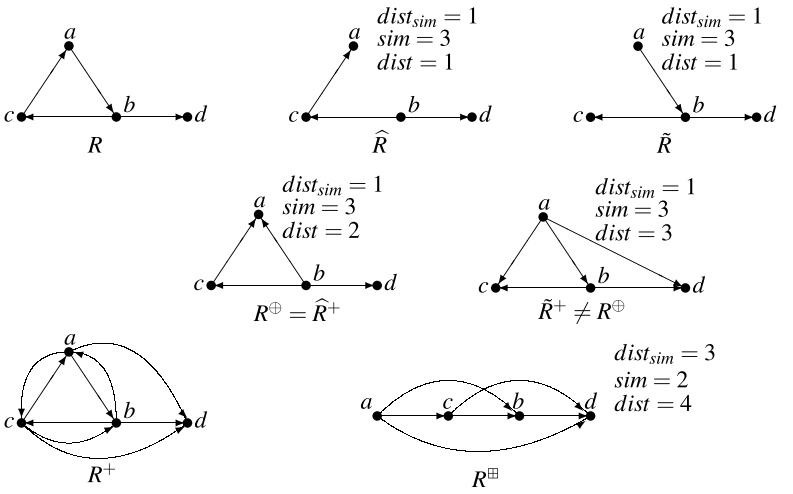

Fig. 5. An example of R, where $\widehat{R}, \tilde{R} \in \mathbb{MDAG}(R)$, and $\widehat{R}^{+} = R^{\oplus}$ but $\tilde{R}^{+} \neq R^{\oplus}$. Also an example that calculating R^{+} first and then taking any $R^{\boxplus} \in \mathbb{MDAG}(R^{+})$ is worse approximation than $R^{\oplus}$.

Proof. (1) Since for every $S \in \mathbb{MDAG}(R)$, $G_S = (X, S)$ is a maximal acyclic subgraph included in G_R.
(2) From (1) and Theorem 2. □

The next theorem is the main result of this section and provides a solution to the case when R contains a cycle.

Theorem 3. *If R contains a cycle, i.e. $R^{cyc} \neq \emptyset$ then:*

1. $R^{\oplus} = \widehat{R}^{+}$, *where* $\widehat{R}$ *is* some *relation from* $\mathbb{MDAG}(R)$.
2. *Finding an optimal simple partial order approximation* $R^{\oplus}$ *is* NP-*complete.*
3. *There are R such that for some* $\tilde{R} \in \mathbb{MDAG}(R)$, $\tilde{R}^{+} \neq R^{\oplus}$.

Proof. (1) Since X is finite clearly $R^{\oplus}$ exists. Define $\widehat{R} = R \cap R^{\oplus}$. Clearly $\widehat{R} \subseteq R$, $\widehat{R}^{cyc} = \emptyset$ and $\widehat{R}^{+}$ is a partial order. Moreover $\widehat{R} \subseteq R \cap \widehat{R}^{+}$. Assume $(a,b) \in R \cap \widehat{R}^{+}$ and $(a,b) \notin \widehat{R} = R \cap R^{\oplus}$. But this means that $sim(R, \widehat{R}^{+}) > sim(R, R^{\oplus})$, a contradiction as $sim(R, R^{\oplus}) = \max\{sim(R, <) \mid < \in \mathbb{PO}(X)\}$. Hence $R \cap R^{\oplus} = R \cap \widehat{R}^{+}$, i.e. $sim(R, R^{\oplus}) = sim(R, \widehat{R}^{+})$. We also have $\widehat{R} = R \cap R^{\oplus} \subseteq R^{\oplus}$, so $\widehat{R}^{+} \subseteq (R^{\oplus})^{+} = R^{\oplus}$. By Lemma 2(2a) this means that $dist(R, \widehat{R}^{+}) \leq sim(R, R^{\oplus})$. However by the definition $dist(R, R^{\oplus}) = \min\{dist(R, <) \mid sim(R, <) = sim(R, R^{\oplus})\}$, so $dist(R, \widehat{R}^{+}) = sim(R, R^{\oplus})$. This, by Lemma 2(2b), means that $R^{\oplus} = \widehat{R}^{+}$.

We will now show that $\widehat{R} \in \mathbb{MDAG}(R)$. Suppose that $\widehat{R} \notin \mathbb{MDAG}(R)$. This means there is $(a,b) \in R \setminus \widehat{R}$ such that $\widehat{R} \cup \{(a,b)\} \subseteq R$ and $(\widehat{R} \cup \{(a,b)\})^{cyc} = \emptyset$. Clearly $(\widehat{R} \cup \{(a,b)\})^{+}$ is a partial order and, since $(a,b) \notin \widehat{R}$, we additionally have $sim(R, (\widehat{R} \cup \{(a,b)\})^{+}) = |R \cap (\widehat{R} \cup \{(a,b)\})^{+}| > |\widehat{R}| = sim(R, R^{\oplus})$. But this means that $sim(R, R^{\oplus}) < \max\{sim(R, <) \mid < \in \mathbb{PO}(X)\}$, a contradiction, i.e. $\widehat{R} \in \mathbb{MDAG}(R)$.

In general $R^{\oplus}$ and corresponding $\widehat{R}$ might not be unique. Note that we only have proven that there is some $\widehat{R} \in \mathbb{MDAG}(R)$ such that $R^{\oplus} = \widehat{R}^{+}$. It does not have to be true for all members of $\mathbb{MDAG}(R)$.

(2) Suppose $R^{\oplus}$ is known. We can derive $\widehat{R} = R \cap R^{\oplus}$ from R and $R^{\oplus}$ in $O(n^2)$ where $n = |X|$, which means that constructing $\widehat{R}$ is polynomially reduced to constructing $R^{\oplus}$. By (1) above, $\widehat{R} \in \mathbb{MDAG}(R)$, and by Lemma 4(2) constructing the elements of $\mathbb{MDAG}(R)$ is NP-complete. Hence constructing $R^{\oplus}$ is NP-complete.

(3) Consider R, $\widehat{R}$ and $\tilde{R}$ from Fig. 5. We have $\widehat{R}, \tilde{R} \in \mathbb{MDAG}(R)$, $\widehat{R}^{+} = R^{\oplus}$ but $\tilde{R}^{+} \neq R^{\oplus}$. □

If $R^{cyc} \neq \emptyset$, feasible construction of $R^{\oplus}$ is problematic. Not only finding any element of $\mathbb{MDAG}(R)$ is NP-complete, but only for *some* $\widehat{R} \in \mathbb{MDAG}(R)$ we have $R^{\oplus} = \widehat{R}^{+}$.

However, if some suboptimal solution is acceptable, there are many efficient approximation, heuristic algorithms, or exact, but feasible ones, that can be used for $\tilde{R} \in \mathbb{MDAG}(R)$ [2,6,9,26], and then we can calculate $\tilde{R}^{+}$ in $O(|X|^3)$ time. We may then consider $\tilde{R}^{+}$ as an approximation of $R^{\oplus}$.

For our case studies, we have: in Fig. 1: $R^{\oplus} = <_1^R$; in Fig. 2: $Q^{\oplus} = <_2^Q$, in Fig. 3: $R^{\oplus} = \widehat{R}^{+}$, and in Fig. 5 $R^{\oplus}$ is $\widehat{R}^{+}$.

Most known approximation algorithms have some substantial overhead time. The following simple randomized algorithm provides an approximation or exact value of $R^{\oplus}$ in polynomial time and it has very little overhead time.

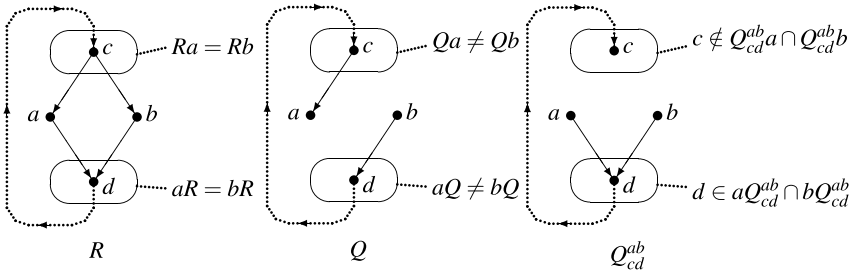

Fig. 6. Transforming Q into Q^{ab}_{cd} in the proof of Proposition 4(4).

Algorithm 1 *(Constructs $R^{\odot}$ – some partial order approximation of R).*
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 $R^{\odot}$. □

A detailed analysis of complexity and accuracy of the above algorithm is an open problem, however some limited random tests for not so big examples indicated that it usually works quite well. In most times the value of $R^{\odot}$ differs from any $R^{\oplus}$ but it is very close.

Two propositions below show the relationship between $R^{\oplus}$ and property-driven partial order approximations. Because our constriction of $R^{\oplus}$ was guided by the fact that partial ordering is a conjunction of acyclicity and transitivity, three out of four conditions required from property-driven partial order approximations, are alway satisfied; and one is never satisfied. The properties of $R^{\oplus}$ corresponding to the conditions (1), (2) and (4) of Definition 1 are rather expected and easy to prove.

Proposition 3.

1. *$R^{\oplus} \subseteq R^{+}$, so the condition (1) of Definition 1 is always satisfied.*
2. *$(R^{\bullet})^{+} \subseteq R^{\oplus}$ and $R^{\bullet} \subseteq R^{\oplus}$, so the condition (2) of Definition 1 is always satisfied.*
3. *If $R^{cyc} \neq \emptyset$ and $(R^{\bullet})^{+} \neq R^{\oplus}$ then: $R^{\oplus} \cap R^{cyc} \neq \emptyset$, so the condition (4) of Definition 1 is never satisfied.*

Proof. (1) From Proposition 2 and Theorem 3(1).
(2) Since $R^{\bullet} \subseteq \widehat{R}$ for each $\widehat{R} \in \mathbb{MDAG}(R)$.
(3) $\widehat{R} \cap R^{cyc} \neq \emptyset$ for each $\widehat{R} \in \mathbb{MDAG}(R)$ different from $R^{\bullet}$, so $\widehat{R}^{+} \cap R^{cyc} \neq \emptyset$, and $R^{\oplus} = \widehat{R}^{+}$. □

On the other hand the result about $R^{\oplus}$ and the condition (3) of Definition 1 is not necessarily expected and its proof is rather complicated. As Fig. 4 shows, the elements of $\mathbb{MDAG}(R)$ may or may not satisfy the equivalence of condition (3), the relation $\widehat{R}_2$ satisfies it but $\widehat{R}_1$ does not. Similarly a partial order approximation of R, the relation $(\widehat{R}_2)^{+}$ satisfies the condition (3) of Definition 1, but a partial order approximation $(\widehat{R}_1)^{+}$ does not.

Proposition 4. *For all $a, b \in X$, $a \equiv_R b \implies a \equiv_{R^{\oplus}} b$, so the condition (3) of Definition 1 is always satisfied.*

Proof. As $\widehat{R}_1$ in Fig. 4 indicates, it might happen that $a \equiv_R b$ but $\neg(a \equiv_Q b)$ for some $Q \in \mathbb{MDAG}(R)$. We will show that there is always some $\widehat{R} \in \mathbb{MDAG}$ such that for all $a, b \in X$ $a \equiv_R b \implies a \equiv_{\widehat{R}} b$. Recall that $a \equiv_R b \iff aR = bR \wedge Ra = Rb$. The case $c \in Ra = Rb$ and $d \in aR = bR$ is illustrated in Fig. 6.

Suppose $a \equiv_R b$ and $\neg(a \equiv_Q b)$ for some $Q \in \mathbb{MDAG}(R)$. This means $Ra = Rb$ and $aR = bR$, but $Qa \neq Qb$ or $aQ \neq bQ$. Suppose $c \in Ra \setminus Rb$. This means that the $(c, b) \in R$ and $(c, b) \notin Q$. Since $Q \in \mathbb{MDAG}(R)$, the only reason for deleting (c, b) is that $dR^{+}c$. But erasing only (c, b) does not break the cycle $cR^{+}dR^{+}c$. We have to additionally delete either (c, a) or (a, d). But deleting (c, a) implies $c \notin Ra$, contradicting our assumption $c \in R$. So we have to delete (a, d) instead. Analysis of remaining cases $c \in Rb \setminus Ra$, $d \in aR \setminus bR$, and $d \in bR \setminus aR$ is the same. The removal of these two arcs from R was needed to break the cycle $cR^{+}dR^{+}c$ in R. Suppose the arcs (c, b) and (a, d) do not belong to Q, as for Q in Fig. 6. We can also break the cycle $cR^{+}dR^{+}c$ in R by deleting arcs (c, a) and (c, b) instead. Let denote this new relations as Q^{ab}_{cd}. Clearly $|Q| = |Q^{ab}_{cd}|$ and $Q \in \mathbb{MDAG}(R)$ implies $Q^{ab}_{cd} \in \mathbb{MDAG}(R)$.

We may repeat the same process for all remaining elements of $aR = bR$ and $Ra = Rb$ and name the outcome relation as Q^{ab}. Now we have $a \equiv_{Q^{ab}} b$. Moreover, if $Q \in \mathbb{MDAG}(R)$ then $Q^{ab} \in \mathbb{MDAG}(R)$ too. Now we may repeat the same process

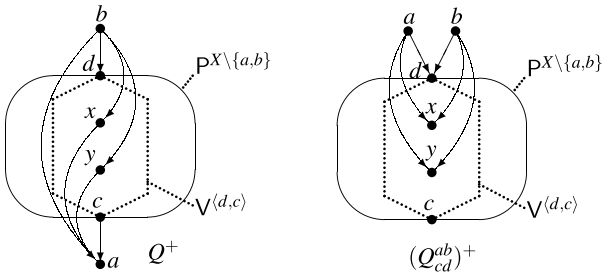

Fig. 7. An idea of the proof that $|Q^+| = |\mathsf{P}^{X\setminus\{a,b\}}| + 2\kappa + 1$ and $|(Q^{ab}_{cd})^+| = |\mathsf{P}^{X\setminus\{a,b\}}| + 2\kappa$, where $\kappa = |\mathsf{V}^{\langle d,c\rangle}|$ (from the proof of Proposition 4(4)).

for all x, y such that $x \equiv_R y$ and $\neg(x \equiv_Q y)$. Let us give the name Q' to the outcome of all these transformations. Clearly if $Q \in \mathbb{MDAG}(R)$ then $Q' \in \mathbb{MDAG}(R)$ as well.

We will show that $dist(R, Q^+) > dist(R, (Q')^+)$. It suffices to show that $dist(R, Q^+) > dist(R, (Q^{ab}_{cd})^+)$ and we will actually show that $dist(R, Q^+) = dist(R, (Q^{ab}_{cd})^+) + 1$. This part of the proof is illustrated in Fig. 7.

First note that $Q^+ \cap (X \setminus \{a,b\}) \times (X \setminus \{a,b\}) = (Q^{ab}_{cd})^+ \cap (X \setminus \{a,b\}) \times (X \setminus \{a,b\})$, and define $\mathsf{P}^{X\setminus\{a,b\}} = Q^+ \cap (X \setminus \{a,b\}) \times (X \setminus \{a,b\})$. Moreover we have $\{x \mid cQ^+xQ^+d\} = \{x \mid c(Q^{ab}_{cd})^+x(Q^{ab}_{cd})^+d\}$. Define $\mathsf{V}^{\langle d,c\rangle} = \{x \mid cQ^+xQ^+d\}$, and let $\kappa = |\mathsf{V}^{\langle d,c\rangle}|$. We now have $Q^+ = \mathsf{P}^{X\setminus\{a,b\}} \cup \{(b,x) \mid x \in \mathsf{V}^{\langle d,c\rangle}\} \cup \{(x,a) \mid x \in \mathsf{V}^{\langle d,c\rangle}\} \cup \{(b,a)\}$. Hence $|Q^+| = |\mathsf{P}^{X\setminus\{a,b\}}| + 2\kappa + 1$. On the other hand $(Q^{ab}_{cd})^+ = \mathsf{P}^{X\setminus\{a,b\}} \cup \{(a,x) \mid x \in \mathsf{V}^{\langle d,c\rangle}\} \cup \{(b,x) \mid x \in \mathsf{V}^{\langle d,c\rangle}\}$, i.e. $|(Q^{ab}_{cd})^+| = |\mathsf{P}^{X\setminus\{a,b\}}| + 2\kappa$. By the definition $dist(R, Q^+) = |R \cup Q^+| - |R \cap Q^+|$ and $dist(R, (Q^{ab}_{cd})^+) = |R \cup (Q^{ab}_{cd})^+| - |R \cap (Q^{ab}_{cd})^+|$. Since $R \cap Q^+ = Q$, $R \cap (Q^{ab}_{cd})^+ = Q^{ab}_{cd}$ and $|Q| = |Q^{ab}_{cd}|$, then $dist(R, Q^+) - dist(R, (Q^{ab}_{cd})^+) = |R \cup Q^+| - |R \cup (Q^{ab}_{cd})^+| = |Q^+| - |Q^{ab}_{cd})^+| = 1$. Hence, $dist(R, Q^+) = dist(R, (Q^{ab}_{cd})^+) + 1$.

This means that for every $Q \in \mathbb{MDAG}(R)$ such that $\neg(\equiv_R \subseteq \equiv_Q)$, there exist $Q' \in \mathbb{MDAG}(R)$ such that $\equiv_R \subseteq \equiv_{Q'}$ and additionally $dist(R, Q^+) > dist(R, (Q')^+)$.

By Lemma 1(4), $\equiv_{Q'} \subseteq \equiv_{(Q')^+}$, so $\equiv_R \subseteq \equiv_{(Q')^+}$ as well. Since $R^\oplus = \widehat{R}^+$ where $dist(R, R^\oplus) = \min\{dist(R, Q^+) \mid Q \in \mathbb{MDAG}(R)\}$, then for all $a, b \in X$, $a \equiv_R b \implies a \equiv_{R^\oplus} b$. □

The statement 'if $Q \in \mathbb{MDAG}(R)$ and for all a, b, $a \equiv_R b \implies a \equiv_{Q^+} b$ then $Q^+ = R^\oplus$' is false. For example $Q = \tilde{R}$, where $\tilde{R}$ is the relation from Fig. 5, satisfies it, but $\tilde{R}^+ \neq R^\oplus$.

The optimal simple partial order approximation, described in Definition 2, is based on the schema used for defining the relation $(R^\bullet)^+$. In the similar way we may follow the pattern $(R^+)^\bullet$. We may define a relation $R^\boxplus$ as follows.

- For a given $R \subseteq X \times X$, the relation $R^\boxplus$ is *any* element of $\mathbb{MDAG}(R^+)$.

Any $R^\boxplus$ can be interpreted as some partial order approximation of R, however it is never a better approximation than $R^\oplus$. We will now prove a slightly generalized equivalence of Proposition 1.

Proposition 5. *For every relation $R \subseteq X \times X$ and every $\widehat{R} \in \mathbb{MDAG}(R)$:*

1. *$sim(R, \widehat{R}^+) \geq sim(R, R^\boxplus)$, i.e. $sim(R, R^\oplus) \geq sim(R, R^\boxplus)$.*
2. *$dist(R, \widehat{R}^+) \leq dist(R, R^\boxplus)$, i.e. $dist(R, R^\oplus) \leq dist(R, R^\boxplus)$.*

Proof. (1) Consider $Q_R = R \cap R^\boxplus$. Clearly $Q_R \subseteq R$, and since $R^\boxplus$ is a partial order, $Q_R^{cyc} = \emptyset$. Since $\mathbb{MDAG}(R)$ is the set of all maximal acyclic relations included in R, then $|Q_R| \leq |\widehat{R}|$. Hence $sim(R, \widehat{R}^+) \geq sim(R, R^\boxplus)$.

An example when $sim(R, \widehat{R}^+) \neq sim(R, R^\boxplus)$ is provided in Fig. 5, where $sim(R, R^\oplus) = 3$ and $sim(R, R^\boxplus) = 2$, and by definition, $R^\oplus = \widehat{R}^+$ for some $\widehat{R} \in \mathbb{MDAG}(R)$.

(2) Since $\widehat{R} \subseteq R$, then $\widehat{R}^+ \subseteq R^+$. Moreover $(\widehat{R}^+)^{cyc} = \emptyset$ so there is $S \in \mathbb{MDAG}(R)$ such that $\widehat{R}^+ \subseteq S$. This means $|\widehat{R}^+| \leq |R^\boxplus|$. Clearly $R \cap \widehat{R}^+ = \widehat{R}$ and from the proof of (1) above we have $|\widehat{R}| \geq |R \cap R^\boxplus|$, i.e. $|R \cap \widehat{R}^+| \geq |R \cap R^\boxplus|$. Hence:

$$dist(R, \widehat{R}^+) = |(R \setminus \widehat{R}^+) \cup (\widehat{R}^+ \setminus R)| = |R| + |\widehat{R}^+| - 2|R \cap \widehat{R}^+|,$$

$$dist(R, R^\boxplus) = |(R \setminus R^\boxplus) \cup (R^\boxplus \setminus R)| = |R| + |R^\boxplus| - 2|R \cap R^\boxplus|.$$

Since $|\widehat{R}^+| \leq |R^\boxplus|$ and $|R \cap \widehat{R}^+| \geq |R \cap R^\boxplus|$, then $dist(R, \widehat{R}^+) \leq dist(R, R^\boxplus)$.

An example when $dist(R, \widehat{R}^+) \neq dist(R, R^\boxplus)$ is provided in Fig. 5, where $dist(R, R^\oplus) = 2$ and $dist(R, R^\boxplus) = 4$. □

The computational complexity of finding $R^\boxplus$ is an open research problem. Calculating R^+ is $O(n^3)$, but a polynomial algorithm for constructing $\mathbb{MDAG}(R^+)$ is unknown. The latter problem can be reduced to minimum feedback problem but

for *transitive* relations. Unfortunately the minimum feedback problem but for *transitive* relations is an open problem too (at least it was at the time this paper was written).

9. Another similarity and distance

In the approach taken so far in this paper, for every relation R and its approximation S, we consider three distinct cases:

1. $(a,b) \in R$ but $(a,b) \notin S \wedge (b,a) \notin S$. In this case, to transform R into S, we just remove (a,b), and the cost of this operation is assumed to be *one*.
2. $(a,b) \notin R \wedge (b,a) \notin R$ but $(a,b) \in S$. In this case, to transform R into S, we just add (a,b), and the cost of this operation is again *one*.
3. $(a,b) \notin R \wedge (b,a) \in R$ but $(a,b) \in S \wedge (b,a) \notin R$. Now, to transform R into S, we add (a,b) *and* remove (b,a), so the cost of this operation is *two*.

Such an approach makes formulas for $sim(\ldots)$ and $dist(\ldots)$ very simple and easy to handle. However, some may argue that for the case (3), which is often called 'flipping', the cost also should be *one* not *two*.

How would the results change if we assume that the cost of 'flipping' is *one*, instead of *two*? This will be the subject of this section.

Assume that the cost of flipping is also one, so we can treat *flippings, additions* and *removals* of edges as 'atomic changes' with the same cost.

Let R and S again be two irreflexive relations on X and let $G_R = (X,R)$, $G_S = (X,S)$ be their appropriate graph representations. In this case we may define $sim_{f1}(R,S)$ and $dist_{f1}(R,S)$, where the subscript '$_{f1}$' indicates that the cost of flipping is one, not two, as follows. Since similarity just means counting common elements, the cost of flipping does not matter, so

$$sim_{f1}(R,S) = sim(R,S) = |R \cap S|.$$

The flipping, i.e. the case 3 above, involves the case when $(a,b) \in S \setminus R$ and $(b,a) \in R \setminus S$, so in the formula $dist(R,S) = |R \setminus R| + |S \setminus R|$, the cost of replacing (b,a) with (a,b) equals 2. If $(a,b) \in S \setminus R$ and $(b,a) \in R \setminus S$ then $(b,a) \in (S \setminus R) \cap (R \setminus S)^{-1}$ and $(a,b) \in (R \setminus S) \cap (S \setminus R)^{-1}$, and clearly

$$|(S \setminus R) \cap (R \setminus S)^{-1}| = |(R \setminus S) \cap (S \setminus R)^{-1}|.$$

Hence, we may define $dist_{f1}(R,S)$ as follows:

$$dist_{f1}(R,S) = |(R \setminus S)| + |(S \setminus R)| - |(S \setminus R) \cap (R \setminus S)^{-1}|.$$

Clearly $dist_{f1}(R,S) \leq dist(R,S)$ for all R, S. We will show that $dist_{f1}$ is also a proper metric.

While in $dist(R,S)$, the argument R and S might be just arbitrary sets, in $dist_{f1}(R,S)$, R and S *must be binary relations.*

Lemma 5. *For all relations $R, S \subseteq X \times X$, the function $dist_{f1}(R,S)$ is a proper metric.*

Proof. Clearly, by the definition:

- $dist_{f1}(R,S) \geq 0$,
- $dist_{f1}(R,S) = dist_{f1}(S,R)$.

We need to show that

1. $dist_{f1}(R,S) = 0 \iff R = S$,
2. $dist_{f1}(R,S) \leq dist_{f1}(R,T) + dist_{f1}(T,S)$.

(1) By the definition again we have $R = S \implies dist_{f1}(R,S) = 0$. Assume that $|(S \setminus R) \cap (R \setminus S)^{-1}| = k > 0$, i.e. $(S \setminus R) \cap (R \setminus S)^{-1} = \{(a_1,b_1), \ldots, (a_k,b_k)\}$. Hence $\{(a_1,b_1), \ldots, (a_k,b_k)\} \subseteq S \setminus R$, so $|S \setminus R| \geq k$, and $\{(b_1,a_1), \ldots, (b_k,a_k)\} \subseteq R \setminus S$, so $|R \setminus S| \geq k$. This means that $dist_{f1}(R,S) = 0$ implies $|(S \setminus R) \cap (R \setminus S)^{-1}| = 0$, i.e. $dist_{f1}(R,S) = |R \setminus S| + |S \setminus R| = 0$, which occurs only if $R = S$. Thus $dist_{f1}(R,S) = 0 \implies R = S$, i.e. $dist_{f1}(R,S) = 0 \iff R = S$.

(2) We set $Left = 0, Right = 0$, and $\mathbb{P} = R \cup S \cup T$. Then we pick up an arbitrary $(a,b) \in \mathbb{P}$, analyze it with respect to its relationship with R and S, and either add zero or one to *Left*. Next we analyze (a,b) with respect to it relationship with R, T and T, S and either add zero, or one or two to *Right*. Then we subtract $\{(a,b)\}$ from $\mathbb{P}$, i.e. now $\mathbb{P} := \mathbb{P} \setminus \{(a,b)\}$, pick up another $(a',b') \in \mathbb{P}$ and repeat the whole process again. We stop when $\mathbb{P} = \emptyset$. We will show that at the end $Left = dist_{f1}(R,S)$, $Right = dist_{f1}(R,T) + dist_{f1}(T,S)$ and in the entire process, we never add one to *Left* and zero to *Right* for the same (a,b). Hence $Left \leq Right$.

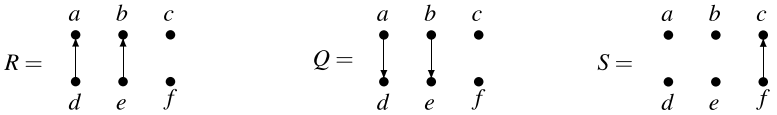

Fig. 8. An example of inconsistency of $dist$ and $dist_{f1}$. We have $dist(R, Q) = 4 > 3 = dist(R, S)$ and $dist_{f1}(R, Q) = 2 < 3 = dist_{f1}(R, S)$.

Let $(a, b) \in \mathbb{P}$. Its relationship to R and S can be divided into the following cases:

1. $(a, b) \in R \cap S$ or $(a, b) \notin R \cup S$, we add zero to *Left*,
2. $(a, b) \in R$, $(a, b) \notin S$ and $(b, a) \notin S$, we add one to *Left*,
3. $(a, b) \in S$, $(a, b) \notin R$ and $(b, a) \notin R$, we add one to *Left*,
4. $(a, b) \in R$, $(a, b) \notin S$ and $(b, a) \in S$, we add one to *Left* (we would add two if computing $dist(R, S)$ instead of $dist_{f1}(R, S)$),
5. $(a, b) \in S$, $(a, b) \notin R$ and $(b, a) \in R$, we add one to *Left* (we would add two if computing $dist(R, S)$ instead of $dist_{f1}(R, S)$).

Repeating this process until $\mathbb{P} = \emptyset$, we clearly get $Left = dist_{f1}(R, S)$.

For the same $(a, b) \in \mathbb{P}$, its relationship with R, T is given by points (1)–(5) above when replacing S with T, and its relationship with T, S is also given by points (1)–(5) above, but when replacing R with T. Hence $Right = dist_{f1}(R, T) + dist_{f1}(T, S)$.

Initially $Left = Right = 0$, so to show that always we have $Left \leq Right$, it suffices to show that there is no such $(a, b) \in X \times X$ such that its processing results in adding one to *Left* and zero to *Right*. Adding one to *Left* means that either $(a, b) \in R \setminus S$ or $(a, b) \in S \setminus R$.

Consider the case $(a, b) \in R \setminus S$. If $(a, b) \in R \setminus T$ then $(a, b) \notin T \cup S$, so we add one to *Right* as well. Suppose now that $(a, b) \in R \setminus S$ and $(a, b) \in T$. Hence $(a, b) \in R \cap S$ and $(a, b) \in T \setminus S$, so we add one to *Right* again. For the case $(a, b) \in S \setminus R$, reasoning is the same. Hence we never add one to *Left* and zero to *Right*.

This means that we always have $Left \leq Right$, i.e. $dist_{f1}(R, S) \leq dist_{f1}(R, T) + dist_{f1}(T, S)$.

It does not mean that always $Left = Right$. When for example $(a, b) \notin R \cup S$ but $(a, b) \in T$, we add zero to *Left* and two to *Right*. □

In [16] the concepts of *consistent* similarity and distance measures for sets has been proposed. Since relations are sets we can easily adapt these concepts to our purposes. Because $sim_{f1}(R, S) = sim(R, S)$ we need the concept of consistency for distances only.

- We will say that two distances $dist_1$ and $dist_2$ are *consistent* if for each three relations $R, Q, S \subseteq X \times X$:

$$dist_1(R, Q) \leq dist_1(R, S) \iff dist_2(R, Q) \leq dist_2(R, S).$$

Consistent distances share many important properties, as proofs for one distance measure can often be extended to all consistent distances. It was shown in [16] that many popular distances are consistent with a rather general Marczewski–Steinhaus index [4,16,22]. Unfortunately the distances $dist$ and $dist_{f1}$ are *not consistent*. A simple counterexample is presented in Fig. 8.

However if one relation is included in another then $dist$ and $dist_{f1}$ are identical.

Proposition 6. *If $R \subseteq S$ then $dist(R, S) = dist_{f1}(R, S)$.*

Proof. $R \subseteq S \implies R \setminus S = \emptyset \implies |(S \setminus R) \cap (R \setminus S)^{-1}| = 0$. □

Despite the fact that $dist$ and $dist_{f1}$ are inconsistent, all the results of Section 6 and Section 8 hold for sim_{f1} and $dist_{f1}$ as well. First we show that an equivalent of Lemma 2 always holds.

Lemma 6. *Let R, S, Q be relations on X.*

1. *If $R \subseteq S$ then: $dist_{f1}(R, S) = |S| - |R| = |S| - sim_{f1}(R, S)$.*
2. *If $sim_{f1}(R, S) = sim_{f1}(R, Q)$ and $S \subseteq Q$ then:*
   (a) $dist_{f1}(R, S) \leq dist_{f1}(R, Q)$,
   (b) $dist_{f1}(R, S) = dist_{f1}(R, Q) \iff S = Q$.

3. *If* $R^{cyc} = \emptyset$ *then*

$$dist_{f1}(R, R^{+}) = \min\{dist_{f1}(R, <) \mid < \in \mathsf{PO}(R)\} = \max\{sim_{f1}(R, <) \mid < \in \mathsf{PO}(R)\},$$

where $\mathsf{PO}(R) = \{ < \mid R \subseteq < \wedge < \in \mathbb{PO}(X)\}$.

Proof. (1) From Proposition 6 and Lemma 2(1).

(2) We have:

$$dist_{f1}(R, S) = |R \cup S| - |R \cap S| - |(S \setminus R) \cap (R \setminus S)^{-1}| \text{ and}$$
$$dist_{f1}(R, Q) = |R \cup Q| - |R \cap Q| - |(Q \setminus R) \cap (R \setminus Q)^{-1}|.$$

If $sim_{f1}(R, S) = sim_{f1}(R, Q)$ then $|R \cap S| = |R \cap Q|$. If $S \subseteq Q$ then $R \cap S \subseteq R \cap Q$, so $|R \cap S| = |R \cap Q|$ means $R \cap S = R \cap Q$. Now we have $R \setminus S = R \setminus (R \cap S) = R \setminus (R \cap Q) = R \setminus Q$. But this means that $(S \setminus R) \cap (R \setminus S)^{-1} = (Q \setminus R) \cap (R \setminus Q)^{-1}$. Hence

$$dist_{f1}(R, Q) - dist_{f1}(R, S) = |R \cup Q| - |R \cup Q| \geq 0.$$

(3) From Proposition 6 and Lemma 2(1). □

Moreover, the quantitative relationship between $(R^{\bullet})^{+}$ and $(R^{+})^{\bullet}$ also remains the same.

Proposition 7. *For every relation R, we have:*

1. $sim(R, (R^{\bullet})^{+}) = sim_{f1}(R, (R^{+})^{\bullet}) = |R^{\bullet}|$,
2. $dist_{f1}(R, (R^{\bullet})^{+}) \leq dist_{f1}(R, (R^{+})^{\bullet})$.

Proof. (1) A consequence of Lemma 3.

(2) By Theorem 1(3) we have $(R^{\bullet})^{+} \subseteq (R^{+})^{\bullet}$. We will show a little bit more general result. Let Q be any relation such that $(R^{\bullet})^{+} \subseteq Q \subseteq (R^{+})^{\bullet}$.

We will prove that:

$$dist_{f1}(R, (R^{\bullet})^{+}) \leq dist_{f1}(R, Q).$$

First recall that by the definition:

$$dist_{f1}(R, Q) = |R \cup Q| - |R \cap Q| - |(Q \setminus R) \cap (R \setminus Q)^{-1}|.$$

We will prove our assertion by induction on the size of $|Q|$. This clearly holds for $Q = (R^{\bullet})^{+}$. Assume for some Q such that $(R^{\bullet})^{+} \subseteq Q \subset (R^{+})^{\bullet}$, we have: $dist_{f1}(R, (R^{\bullet})^{+}) \leq dist_{f1}(R, Q)$. Let $(a, b) \in (R^{+})^{\bullet} \setminus Q$, and let $Q' = Q \cup \{(a, b)\}$. We have to consider three cases: $(a, b) \in R$, $(a, b) \notin R \wedge (b, a) \in R \setminus Q'$ and $(a, b) \notin R \wedge (b, a) \notin R \setminus Q'$.

Case 1. $(a, b) \in R$. In this case we have: $R \cup Q' = R \cup Q$, $R \cap Q' = R \cap Q$, $Q' \setminus R = Q \setminus R$ and $R \setminus Q' = R \setminus Q$. Hence $dist_{f1}(R, Q') = dist_{f1}(R, Q)$, i.e. $dist_{f1}(R, (R^{\bullet})^{+}) \leq dist_{f1}(R, Q')$.

Case 2. $(a, b) \notin R \wedge (b, a) \in R \setminus Q'$. Now we have: $|R \cup Q'| = |R \cup Q| + 1$, $R \cap Q' = R \cap Q$ and $|(Q' \setminus R) \cap (R \setminus Q')^{-1}| = |(Q \setminus R) \cap (R \setminus Q)^{-1}| + 1$, hence again $dist_{f1}(R, Q') = dist_{f1}(R, Q)$, i.e. $dist_{f1}(R, (R^{\bullet})^{+}) \leq dist_{f1}(R, Q')$. Note that in this case $dist(R, Q') = dist(R, Q) + 1$.

Case 3. $(a, b) \notin R \wedge (b, a) \notin R \setminus Q'$. In this case we have $|R \cup Q'| = |R \cup Q| + 1$, $R \cap Q' = R \cap Q$ and $|(Q' \setminus R) \cap (R \setminus Q')^{-1}| = |(Q \setminus R) \cap (R \setminus Q)^{-1}|$, hence $dist_{f1}(R, Q') = dist_{f1}(R, Q) + 1$, i.e. $dist_{f1}(R, (R^{\bullet})^{+}) \leq dist_{f1}(R, Q')$. □

We will now replace *sim* and *dist* with sim_{f1} and $dist_{f1}$ in the definition of optimal simple partial order approximation of a given relation R.

Definition 3. For every relation R on X, a partial order $R^{\oplus f1}$ on X is a *flip1-optimal simple partial order approximation* of R if the following conditions are satisfied:

1. $R^{\bullet} \subseteq R^{\oplus f1}$,
2. $sim_{f1}(R, R^{\oplus f1}) = \max\{sim_{f1}(R, <) \mid < \in \mathbb{PO}(X)\}$,
3. $dist_{f1}(R, R^{\oplus f1}) = \min\{dist_{f1}(R, <) \mid sim_{f1}(R, <) = sim_{f1}(R, R^{\oplus f1})\}$.

Let $\mathsf{POA}_{f1}(R)$ denote the set of all flip1-optimal simple partial order approximations of R. □

Proposition 8. *For every relation R on X:* $\mathsf{POA}(R) = \mathsf{POA}_{f1}(R)$.

Proof. Because $sim = sim_{f1}$ then $\mathbb{MDAG}(R) = \mathbb{MDAG}_{f1}(R)$. We can now literally copy proofs of Proposition 2 and Theorem 3 with replacing Lemma 2 by Lemma 6. □

Despite the fact that *dist* and $dist_{f1}$ are not consistent, in our approach to optimal partially ordered approximation they are equivalent, one can be replaced by another and the results remain the same. This is mainly because in our model *sim* plays a stronger role than *dist*, and counting 'flippings' is not included in *sim*, i.e. $sim = sim_{f1}$. Nevertheless, $dist_{f1}$ is a new metric, it is not included in the recent encyclopedia of metrics [4], and may lead to different results than *dist* in other applications, especially those that do not use *sim*. However this topic is beyond the subject of this paper.

10. Relationship to Rough Sets model

Our approach is intuitively close to the ideas of Rough Sets [23]. The transitive closure can be interpreted as some upper approximation, and the cyclic refinement and maximal directed subgraphs can be interpreted as a kind of lower approximations. However, the formal relationship between our model and Rough sets is quite complex and it is based on the concepts and the results of [12,13] and [16].

The principles of Rough Sets [23] can be formulated as follows. Let U be a finite and nonempty universe of elements, and let $E \subseteq U \times U$ be an *equivalence relation*. The elements of U/E are called *elementary* sets and they are interpreted as basic observable or measurable sets. The pair (U, E) is referred to as a Pawlak *approximation space*. A set $X \subseteq U$ is approximated by two subsets of U, $\underline{\mathbf{A}}(X)$ – called the lower approximation of X, and $\overline{\mathbf{A}}(X)$ – called the upper approximation of X, where:

$$\underline{\mathbf{A}}(X) = \bigcup\{[x]_E \mid x \in U \wedge [x]_E \subseteq X\}, \qquad \overline{\mathbf{A}}(X) = \bigcup\{[x]_E \mid x \in U \wedge [x]_E \cap X \neq \emptyset\}.$$

In [16] the concept of *optimal approximation* has been introduced. A set $A \subseteq U$ is *definable* (or *exact*) [23] if it is a union of some equivalence classes of the equivalence relation E. Let $\mathbb{D}$ denote the family of all definable sets defined by the space (U, E), i.e. $\mathbb{D} = \{A \mid A = \bigcup_{x \in A}[x]_E\}$. For every set $X \subseteq U$, a definable set $\mathsf{O} \in \mathbb{D}$ is an *optimal approximation* of X (w.r.t. a given *similarity measure sim*) if and only if:

$$sim(X, \mathsf{O}) = \max_{\mathsf{A} \in \mathbb{D}}(sim(X, \mathsf{A})),$$

and $\mathsf{Opt}_{sim}(X)$ denotes the set of all optimal approximations of X. An efficient polynomial algorithm that find an optimal approximation for any similarity consistent with fairly general Marczewski–Steinhaus index[2] [22] (and that includes Jaccard similarity [10]) was also provided in [16]. Moreover the following result has been proven.

Proposition 9 *([16]). For every similarity measure that satisfies five simple axioms[3] from [16], for every set $X \subseteq U$, and every $\mathsf{O} \in \mathsf{Opt}_{sim}(X)$:*

$$\underline{\mathbf{A}}(X) \subseteq \mathsf{O} \subseteq \overline{\mathbf{A}}(X). \quad \square$$

Since every relation is a set of pairs, this approach can be used for relations as well [27]. Unfortunately, in such cases as ours we want approximations to have some specific properties like irreflexivity, transitivity etc., and most of those properties are not closed under the set union operator. As was pointed out in [30], in general one cannot expect approximations to have the desired properties (see [30] for details). It is also unclear how to define the relation E for cases such as ours.

However the Rough Sets can also be defined in an orthogonal (sometimes called 'topological') manner [23,27,29]. We may start with defining a space as $(U, \mathcal{D})$ where $\mathcal{D}$ is a family of sets that contains $\emptyset$ and for each $x \in U$ there is $X \in \mathcal{D}$ such that $x \in X$ (i.e. $\mathcal{D}$ is some *covering* of U [24]). We may now define $E_{\mathcal{D}}$ as the equivalence relation generated by the set of all *components* (in the sense of [21]) defined by the covering $\mathcal{D}$. If we set $\mathcal{D} = \mathbb{D}$, where $\mathbb{D}$ is the set of all definable sets of a classic approximation space (U, E), then $E_{\mathcal{D}} = E$. Hence both approaches can be considered as equivalent [23,27,30], however now for each $X \subseteq U$ we have:

$$\underline{\mathbf{A}}(X) = \bigcup\{Y \mid Y \subseteq X \wedge Y \in \mathcal{D}\}, \qquad \overline{\mathbf{A}}(X) = \bigcap\{Y \mid X \subseteq Y \wedge Y \in \mathcal{D}\}.$$

[2] Marczewski–Steinhaus index is defined as $sim(X, Y) = \frac{\mu(X \cap Y)}{\mu(X \cup Y)}$, where μ is a finite measure on U and $X, Y \subseteq U$ [22].

[3] Similarity axioms of [16] are the following, for all sets A, B, we have:

$S1$ (Maximum):	$sim(A, B) = 1 \iff A = B$,
$S2$ (Symmetry):	$sim(A, B) = sim(B, A)$,
$S3$ (Minimum):	$sim(A, B) = 0 \iff A \cap B = \emptyset$,
$S4$ (Inclusion):	if $a \in B \setminus A$ then $sim(A, B) < sim(A \cup \{a\}, B)$,
$S5$ (Exclusion):	if $a \notin A \cup B$ and $A \cap B \neq \emptyset$ then $sim(A, B) > sim(A \cup \{a\}, B)$.

As explained in Sections 3 and 8, in this paper we use different, asymmetric, approach to similarity measures, so the axioms $S1$, $S4$ and $S5$ are not satisfied.

We can now define $\mathcal{D}_\alpha$ as a set of relations having the desired property α and then calculate $\underline{\mathbf{A}}(R)$ and/or $\overline{\mathbf{A}}(R)$ with respect to a given $\mathcal{D}_\alpha$. Such an approach was proposed and analyzed in [30], however it seems to have only limited applications. It assumes that the set $\mathcal{D}_\alpha$ is closed under both union and intersection, and few properties of relations do this. For instance, transitivity is not closed under union and having a cycle is not closed under intersection. This and other problems have been discussed in details in [12,13], where a different approach has been proposed.

Let α be any predicate that describe some *property of binary relations*. As an example we can take

$$\alpha = [\forall a, b, c \in X.\ \neg(aRa) \wedge (aRbRc \Rightarrow aRc)]$$

i.e. a definition of (sharp) partial order.

Let X be a finite set and let

$$Rel_\alpha = \{R \mid R \subseteq X \times X \text{ and } R \text{ satisfies the property } \alpha\ \}.$$

If α is a predicate defined above, then Rel_α is the set of all (sharp) partial orders on X.

Let Prop denote the set of predicates such that $\alpha \in \mathsf{Prop} \Rightarrow Rel_\alpha \neq \emptyset \wedge Rel_\alpha \neq \{\emptyset\}$.

Note that we allow the case $\alpha \in \mathsf{Prop}$ and $\emptyset \in Rel_\alpha$. The restrictions of Prop are merely for technical reasons, to avoid considering pathological cases in each result and proof. For more details the reader is referred to [13].

Let $R \subseteq X \times X$ be a non-empty relation and $\alpha \in \mathsf{Prop}$. We say that:

- R has α-lower bound $\iff \exists Q \in Rel_\alpha.\ Q \subseteq R$,
- R has α-upper bound $\iff \exists Q \in Rel_\alpha.\ R \subseteq Q$.

We also define

- $lb_\alpha(R) = \{Q \mid Q \in Rel_\alpha \wedge Q \subseteq R\}$, the set of all α-lower bounds of R, and
- $ub_\alpha(R) = \{Q \mid Q \in Rel_\alpha \wedge R \subseteq Q\}$, the set of all α-upper bounds of R.

Both $lb_\alpha(R)$ and $ub_\alpha(R)$ always exists but they may be empty sets.

For every family of sets $\mathcal{F}$ (relations are sets), let $MIN(\mathcal{F})$ and $MAX(\mathcal{F})$ be defined as follows:

- $MIN(\mathcal{F}) = \{R \mid \forall Q \in \mathcal{F}.\ Q \subseteq R \Rightarrow R = Q\}$, the set of all minimal elements of $\mathcal{F}$,
- $MAX(\mathcal{F}) = \{R \mid \forall Q \in \mathcal{F}.\ R \subseteq Q \Rightarrow R = Q\}$, the set of all maximal elements of $\mathcal{F}$.

Assume that Rel_α is closed under intersection *or* union (could be both, but does not have to). We can now define the lower and upper approximations $\underline{\mathbf{A}}_\alpha(R)$ and $\overline{\mathbf{A}}_\alpha(R)$ as follows:

- If R has α-lower bound then we define its *α-lower approximation* as:

$$\underline{\mathbf{A}}_\alpha(R) = \bigcap\{Q \mid Q \in MAX(lb_\alpha(R))\}.$$

- If R has α-upper bound then we define its *α-upper approximation* as:

$$\overline{\mathbf{A}}_\alpha(R) = \bigcup\{Q \mid Q \in MIN(ub_\alpha(R))\}.$$

If R *does not have* α-lower bound (α-upper bound), then its α-lower approximation (α-upper approximation) *does not exist*.

Directly from the definitions it follows that $\underline{\mathbf{A}}_\alpha(R)$ is well defined if Rel_α is closed under intersection and $\overline{\mathbf{A}}_\alpha(R)$ is well defined if Rel_α is closed under union. The result below shows that both concepts are well defined if Rel_α is closed either under intersection *or* union (or both).

Proposition 10 *([12]).*

1. *If Rel_α is closed under union and R has α-lower bound, then*

$$\underline{\mathbf{A}}_\alpha(R) = \bigcup\{Q \mid Q \subseteq R \wedge Q \in P_\alpha\}.$$

2. *If Rel_α is closed under intersection and R has α-upper bound, then*

$$\overline{\mathbf{A}}_\alpha(R) = \bigcap\{Q \mid R \subseteq Q \wedge Q \in P_\alpha\}. \quad \square$$

In Proposition 10, the formulas for $\underline{\mathbf{A}}_\alpha(R)$ and $\overline{\mathbf{A}}_\alpha(R)$ are practically the same as for standard lower and upper approximations for 'topological case' discussed above, with just P_α instead of $\mathcal{D}$.

In general the properties of $\underline{\mathbf{A}}_\alpha(R)$ and $\overline{\mathbf{A}}_\alpha(R)$ mimic those of $\underline{\mathbf{A}}(R)$ and $\overline{\mathbf{A}}(R)$, in particular, if $\underline{\mathbf{A}}_\alpha(R)$ exists, then $\underline{\mathbf{A}}_\alpha(R) \subseteq R$; and if $\overline{\mathbf{A}}_\alpha(R)$ exists, then $R \subseteq \overline{\mathbf{A}}_\alpha(R)$ (cf. [12,13]).

Let α_{tr}, α_{ac} and α_{po} be the following properties:

- $\alpha_{tr} \stackrel{df}{=} [\forall a, b, c \in X.\ aRb \wedge bRa \Rightarrow aRc]$, i.e. $\alpha_{tr} = \textit{transitivity}$,
- $\alpha_{ac} \stackrel{df}{=} [\forall a, b \in X.\ \neg(aR^{cyc}b)]$, i.e. $\alpha_{ac} = \textit{acyclicity}$,
- $\alpha_{po} \stackrel{df}{=} [\forall a, b, c \in X. \neg(aRa) \wedge (aRbRc \Rightarrow aRc)]$ i.e. $\alpha_{po} = \textit{partial ordering}$,

Clearly $\alpha_{ac} \wedge \alpha_{tr} \iff \alpha_{po}$. While the property α_{po} is not very useful as $\overline{\mathbf{A}}_{\alpha_{po}}(R)$ usually does not exist and $\underline{\mathbf{A}}_{\alpha_{po}}(R)$ is often equal to empty set, since $\alpha_{po} \iff \alpha_{ac} \wedge \alpha_{tr}$, we can use *compositions* of α_{ac}-approximation and α_{tr}-approximation to represent approximations by partial orders It turns out the relations $R^+, R^\bullet, (R^\bullet)^+$ and $(R^+)^\bullet$ can be described as appropriate α-approximations or *compositions* of α-approximations, where $\alpha \in \{\alpha_{ac}, \alpha_{tr}\}$.

Proposition 11 *([12]).*

1. $R^+ = \overline{\mathbf{A}}_{\alpha_{tr}}(R)$,
2. $R^\bullet = \underline{\mathbf{A}}_{\alpha_{ac}}(R)$,
3. $(R^\bullet)^+ = \overline{\mathbf{A}}_{\alpha_{tr}}(\underline{\mathbf{A}}_{\alpha_{ac}}(R))$,
4. $(R^+)^\bullet = \underline{\mathbf{A}}_{\alpha_{ac}}(\overline{\mathbf{A}}_{\alpha_{tr}}(R))$. □

For more details on similar subjects the reader is referred to [12,13]. The compositions of α-approximations can also be used for approximations by the relations different than partial orders, more details in [13]. Optimal approximations in the sense of [16] are not considered in [12,13], but they can be defined in a relatively straightforward manner. Since similarity measures used in this paper are not scaled to $[0, 1]$, we cannot literally copy appropriate definitions from [16].

For every relation $R \subseteq X \times X$ and every property $\alpha \in \mathsf{Prop}$, a relation $\mathsf{O} \in Rel_\alpha$ is an *optimal α-approximation* of R,

(a) w.r.t. *similarity*, iff $\mathsf{O} \in MIN(\mathbb{O})$, where $\mathbb{O}$ is the family of relations such that $\mathsf{O}' \in \mathbb{O} \iff sim(R, \mathsf{O}') = \max\limits_{\mathsf{Q} \in Rel_\alpha} (|R \cap \mathsf{Q}|)$,

(b) w.r.t. *distance*, iff $dist(R, \mathsf{O}) = \min\limits_{\mathsf{Q} \in Rel_\alpha} (|R \cup \mathsf{Q}| - |R \cap \mathsf{Q}|)$.

We need '$\mathsf{O} \in MIN(\mathbb{O})$' in (a) since it may happen that $R \cap \mathsf{O} = R \cap \mathsf{O}'$ and $\mathsf{O} \subsetneq \mathsf{O}'$. The similarity measure $\max\limits_{\mathsf{Q} \in Rel_\alpha} (|R \cap \mathsf{Q}|)$ is different that all considered in [16]. It does not satisfy axioms $S4$ and $S5$ from Proposition 9. Hence we need '$\mathsf{O} \in MIN(\mathbb{O})$' in (a). On the other hand, the distance $\min\limits_{\mathsf{Q} \in Rel_\alpha} (|R \cup \mathsf{Q}| - |R \cap \mathsf{Q}|)$ is a proper measure, so it does not need any specialized modifications.

- The set of all optimal α-approximations of R w.r.t. *similarity* will be denoted by $\mathsf{Opt}^\alpha_{sim}(R)$, and
- the set of all optimal α-approximations of R w.r.t. *distance* will be denoted by $\mathsf{Opt}^\alpha_{dist}(R)$.

The sets $\mathsf{Opt}^\alpha_{sim}(R)$ and $\mathsf{Opt}^\alpha_{dist}(R)$ always exist, even if $\underline{\mathbf{A}}_\alpha(R)$ or $\overline{\mathbf{A}}_\alpha(R)$ do not. For example $\mathsf{Opt}^{\alpha_{po}}_{sim}(R)$ exists while $\overline{\mathbf{A}}_{\alpha_{po}}(R)$ does not if R contains a cycle (cf. [12,13]).

We can now formulate an equivalent of Proposition 9, but for α-approximations.

Proposition 12. *For every $\alpha \in \mathsf{Prop}$, every relation $R \subseteq X \times X$, and every $\mathsf{O} \in \mathsf{Opt}^\alpha_{sim}(R)$, $\mathsf{Q} \in \mathsf{Opt}^\alpha_{dist}(R)$, we have:*

1. *if $\underline{\mathbf{A}}_\alpha(R)$ exists then $\underline{\mathbf{A}}_\alpha(R) \subseteq \mathsf{O}$ and $\underline{\mathbf{A}}_\alpha(R) \subseteq \mathsf{Q}$,*
2. *if $\overline{\mathbf{A}}_\alpha(R)$ exists then $\mathsf{O} \subseteq \overline{\mathbf{A}}_\alpha(R)$ and $\mathsf{Q} \subseteq \overline{\mathbf{A}}_\alpha(X)$.*

Proof. (1) Suppose $(a, b) \in \underline{\mathbf{A}}_\alpha(R) \setminus \mathsf{O}$. Since $\underline{\mathbf{A}}_\alpha(R) \subseteq R$ then $|(\mathsf{O} \cup \{(a, b)\}) \cap R| = |\mathsf{O} \cap R| + 1$, so $\mathsf{O} \notin \mathsf{Opt}^\alpha_{sim}(R)$. Suppose now $(a, b) \in \underline{\mathbf{A}}_\alpha(R) \setminus \mathsf{Q}$. Since $\underline{\mathbf{A}}_\alpha(R) \subseteq R$ then $|R \cup (\mathsf{Q} \cup \{(a, b)\})| - |R \cap (\mathsf{Q} \cup \{(a, b)\})| = |R \cup \mathsf{Q}| - |R \cap (\mathsf{Q} \cup \{(a, b)\})| = |R \cup \mathsf{Q}| - |R \cap \mathsf{Q}| - 1$, so $\mathsf{Q} \notin \mathsf{Opt}^\alpha_{dist}(R)$.

(2) Suppose $(a, b) \in \mathsf{O} \setminus \overline{\mathbf{A}}_\alpha(R)$. Since $R \subseteq \overline{\mathbf{A}}_\alpha(R)$ then $(a, b) \notin R$. Now we have $|(\mathsf{O} \setminus \{(a, b)\}) \cap R| = |\mathsf{O} \cap R|$, and $\mathsf{O} \setminus \{(a, b)\} \subsetneq \mathsf{O}$, so O does not belong to $MIN(\mathbb{O})$, which implies $\mathsf{O} \notin \mathsf{Opt}^\alpha_{sim}(R)$. Suppose now $(a, b) \in \mathsf{Q} \setminus \overline{\mathbf{A}}_\alpha(R)$. Again since $R \subseteq \overline{\mathbf{A}}_\alpha(R)$ then $(a, b) \notin R$. Hence $|R \cup (\mathsf{Q} \setminus \{(a, b)\})| - |R \cap (\mathsf{Q} \setminus \{(a, b)\}) = |R \cup (\mathsf{Q} \setminus \{(a, b)\})| - |R \cap \mathsf{Q}| = |R \cup \mathsf{Q}| - |R \cap \mathsf{Q}| - 1$, so $\mathsf{Q} \notin \mathsf{Opt}^\alpha_{dist}(R)$. □

The above result can be interpreted as some validation of the choice of definitions for $sim(...)$ and $dist(..)$ we have made, and suggest that they may be useful for other type of approximations, not only by partial orders. It may also happen that optimal α-approximations are just lower or upper α-approximations.

Corollary 1. *For each relation R, we have:* $\mathsf{Opt}^{\alpha_{tr}}_{dist}(R) = \overline{\mathbf{A}}_{\alpha_{tr}}(R)$.

Proof. By Proposition 11(1), we have $\overline{\mathbf{A}}_{\alpha_{tr}}(R) = R^+$ and R^+ is the smallest transitive relation that includes R. □

Now we may formulate some relevant results of this paper in terms of Rough Sets.

Proposition 13.

1. $\mathbb{MDGAR}(R) = \mathsf{Opt}^{\alpha_{ac}}_{sim}(R)$.
2. $\mathsf{POA}(R) = \mathsf{Opt}^{\alpha_{tr}}_{dist}(\mathsf{Opt}^{\alpha_{po}}_{sim}(R)) = \overline{\mathbf{A}}_{\alpha_{tr}}(\mathsf{Opt}^{\alpha_{po}}_{sim}(R))$.
3. *For each* $R^{\oplus} \in \mathsf{POA}(R)$, *we have* $\underline{\mathbf{A}}_{\alpha_{ac}}(R) \subseteq R^{\oplus}$.
4. $\mathsf{POA}(R) \subseteq \mathsf{Opt}^{\alpha_{tr}}_{dist}(\mathbb{MDGAR}(R)) = \mathsf{Opt}^{\alpha_{tr}}_{dist}(\mathsf{Opt}^{\alpha_{ac}}_{sim}(R)) = \overline{\mathbf{A}}_{\alpha_{tr}}(\mathsf{Opt}^{\alpha_{ac}}_{sim}(R))$.
5. *There is* R *such that* $\mathsf{POA}(R) \neq \overline{\mathbf{A}}_{\alpha_{tr}}(\mathsf{Opt}^{\alpha_{ac}}_{sim}(R))$.

Proof. (1) From appropriate definitions.
(2) From appropriate definitions and Corollary 1.
(3) From Proposition 11(2) and Definition 2(1).
(4) From appropriate definitions, Theorem 3(1) and Corollary 1.
(5) From Theorem 3(1) or Fig. 5. □

Proposition 13 does not provide any new algorithmic ideas on how to find $R^{\oplus}$ efficiently, however its Rough Sets based formulation provides additional structure to the relationship between $\mathbb{MDGAR}(R)$, $\mathsf{POA}(R)$ and $R^{\oplus}$, the main features of our model. For example, this proposition gives a very natural explanation of why not every transitive closure of an element of $\mathbb{MDGAR}(R)$ is an optimal simple partial order approximation of R. Since $\alpha_{po} \Longrightarrow \alpha_{ac}$, then $\mathsf{Opt}^{\alpha_{po}}_{sim}(R)) \subseteq \mathsf{Opt}^{\alpha_{ac}}_{sim}(R))$ and for some R, we have $\mathsf{Opt}^{\alpha_{po}}_{sim}(R)) \neq \mathsf{Opt}^{\alpha_{ac}}_{sim}(R))$.

Propositions 12 and 13 also do hold if sim is replaced with sim_{f1} and $dist$ with $dist_{f1}$. The proofs are practically identical.

11. Final comment

In many applications of partial orders, full transitivity is never or seldom explicitly used. Quite often use of acyclic relations that uniquely represent partial orders, as Hasse diagrams, dependency graphs, etc., is sufficient and more efficient (cf. [5,8,20]). If $(a,b) \in R, (b,c) \in R$ but $(a,c) \notin R, (c,a) \notin R$, and R is interpreted as partial ordering, then either the relationship between a and c was analyzed and declared that a and c are incomparable, and then we have some inconsistency; or the relationship between a and c was just not analyzed as transitivity of R was implicitly assumed. Cycles, on the other hand, are always a result of errors or data inconsistency. Even if the case $(a,b) \in R, (b,c) \in R$ but $(a,c) \notin R, (c,a) \notin R$ is the result of errors or inconsistencies, this case appears to be less serious problem than the case $(a,b) \in R, (b,c) \in R$ and $(c,a) \in R^+$.

As one of the reviewers rightly pointed out, the lack of transitivity when R is interpreted as a partial order may be due at least to two different circumstances:

1. A pair of elements a and c are declared to be incomparable, while both (a,b) and (b,c) belong to R, so there is some inconsistency.
2. The relation between a and c is not analyzed.

The second case could correspond to a situation where transitivity is implicitly assumed, and therefore the pair (a,c) can be assumed to belong to R. In this paper both case are treated in the same way. First we made the relation R acyclic, and then we assume that for acyclic relations the transitive closure produces the closest partial order.

However, the above two situations may be treated differently. In the second case, a reasonable procedure seems to include the pair (a,c) in our approximation of R, while in the first case we do not know whether the conflict should be solved either by means of adding the pair (a,c) or contrarily by means of removing some of the pairs (a,b) or (b,c). This case is illustrated in Fig. 2 in Section 4 where the problems with defining optimal approximation are discussed. However this would require defining optimal partial order approximation of an acyclic relation R that is included in R, i.e. a lower approximation, and so far we do not have any feasible definition. On the other hand, from Lemma 2(3), it follows that the transitive closure is *the* optimal upper partial order approximation of any acyclic relation, with respect to both closeness measures used in this paper, namely: the similarity $sim(R,S) = |R \cap S|$ and distance $dist(R,S) = |R \cap S| - |R \cap S|$.

In our model the measure *sim* punishes more for having cycles, while *dist* for not being transitive. Since we believe that cycles are more problematic than lack of transitivity (that might be intentional, just the result of specific procedure of an experiment), we emphasize *sim* in Definition 2, and do not use measures like Jaccard index [10] (or others used in [16]), where a differentiation between *sim* and *dist* is also impossible.

We have shown that the quantitative approach to partial order approximations of arbitrary relations is mostly consistent with the property-driven approach presented in [12,13], moreover when quantitative measures are applied to property-driven approximations $(R^\bullet)^+$ and $(R^+)^\bullet$, the relation $(R^\bullet)^+$ is a better approximation of R than $(R^+)^\bullet$.

The main technical result of this paper is somewhat pessimistic as finding our optimal partially ordered approximation is NP-complete, so we have to look for some randomized and approximation algorithms.

When the classical symmetric difference is used as a measure of a distance between relations, the cost of arc flipping is two, while the costs of arc removal and addition are one. An alternative measure of distance, where all three costs are equal to one has been introduced and proved to be a metric. This new measure is inconsistent (in a sense of [16]) with symmetric difference, however as far as approximations by partial orders are concerned, it provides the same results as standard symmetric difference.

We also show how our model can be expressed in Rough Sets settings using α-approximations introduced in [12] and studied in [13]. Rough Sets setting enables additional validation of the similarity measures that are used in the paper, and also provides more intuition for our main result.

In [11,15,13] and many other papers, the relation $(R^+)^\bullet$ (classical Shröder's approximation) was used as a partial order approximation for group ranking and pairwise comparisons non-numerical ranking. In the view of the results of this paper, it should be replaced by either $(R^\bullet)^+$, or any $R^\odot$ produced by Algorithm 1, or any $R^\oplus \in \mathsf{POA}(R)$, as we have $sim((R^+)^\bullet, R) \leq sim((R^\bullet)^+, R) \leq sim(R^\odot, R) \leq sim(R^\oplus, R)$.

Acknowledgements

The author gratefully acknowledges two anonymous referees, whose comments significantly contributed to the final version of this paper.

George Karakostas is thanked for a hint that helped to prove Theorem 3 and Ian Munro for influential comments on the nature of 'flipping'. This research has partially been supported by a Discovery NSERC grant No. RGPIN6466-15.

References

[1] A.D. Bogobowicz, R. Janicki, On approximation of relations by generalized closures and generalized kernels, in: Lecture Notes in Artificial Intelligence, vol. 9920, Springer, 2016, pp. 120–130.
[2] J. Chen, Y. Liu, S. Lu, B. O'sullivan, I. Razgon, A fixed parameter algorithm for the directed feedback vertex set problem, J. ACM 55 (5) (2008) 1–19.
[3] T.H. Cormen, C.E. Leiserson, D.L. Rivest, C. Stein, Introduction to Algorithms, MIT Press, 2001.
[4] M.M. Deza, E. Deza, Encyclopedia of Distances, Springer, Berlin, 2012.
[5] V. Diekert, G. Rozenberg (Eds.), The Book of Traces, World Scientific, Singapore, 1995.
[6] P. Eades, X. Lin, W.F. Smyth, A fast and effective heuristic for the feedback arc set problem, Inf. Process. Lett. 47 (1993) 319–323.
[7] P.C. Fishburn, Interval Orders and Interval Graphs, J. Wiley, New York, 1985.
[8] S. French, Decision Theory, Ellis Horwood, New York, 1986.
[9] R. Hassin, S. Rubinstein, Approximations for the maximum acyclic subgraph problem, Inf. Process. Lett. 51 (3) (1994) 133–140.
[10] P. Jaccard, Étude comparative de la distribution florale dans une portion des Alpes et des Jura, Bull. Soc. Vaud. Sci. Nat. 37 (1901) 547–549.
[11] R. Janicki, Pairwise comparisons based non-numerical ranking, Fundam. Inform. 94 (2) (2009) 197–217.
[12] R. Janicki, Approximations of arbitrary binary relations by partial orders. Classical and rough set models, in: Transactions on Rough Sets 13, 2011, pp. 17–38.
[13] R. Janicki, Property-driven Rough Sets approximations of relations, in: A. Skowron, Z. Suraj (Eds.), Rough Sets and Intelligent Systems – Professor Zdzisław Pawlak in Memoriam, in: Intelligent Systems Reference Library, vol. 42, Springer, 2013, pp. 333–357.
[14] R. Janicki, On optimal approximations of arbitrary relations by partial orders, in: Lecture Notes in Artificial Intelligence, vol. 9920, Springer, 2016, pp. 107–119.
[15] R. Janicki, W.W. Koczkodaj, Weak order approach to group ranking, Comput. Math. Appl. 32 (2) (1996) 51–59.
[16] R. Janicki, A. Lenarčič, Optimal approximations with Rough Sets and similarities in measure spaces, Int. J. Approx. Reason. 71 (2016) 1–14.
[17] R. Janicki, Y. Zhai, On a pairwise comparison based consistent non-numerical ranking, Log. J. IGPL 20 (4) (2012) 667–676.
[18] M. Karp, Reducibility among combinatorial problems, in: R.E. Miller, J.W. Thatcher (Eds.), Complexity of Computer Computations, Plenum, New York, 1972, pp. 85–103.
[19] J.G. Kemeny, Mathematics without numbers, Daedalus 88 (1959) 571–591.
[20] J. Kleijn, M. Koutny, Formal languages and concurrent behaviour, Stud. Comput. Intell. 113 (2008) 125–182.
[21] K. Kuratowski, A. Mostowski, Set Theory, 2nd edition, North Holland, 1976.
[22] E. Marczewski, H. Steinhaus, On a certain distance of sets and corresponding distance of functions, Colloq. Math. 4 (1958) 319–327.
[23] Z. Pawlak, Rough Sets, Kluwer, Dordrecht, 1991.
[24] K.H. Rosen, Discrete Mathematics and Its Applications, McGraw-Hill, New York, 1999.
[25] E. Schröder, Algebra der Logik, Teuber, Leipzig, 1895.
[26] S. Skiena, The Algorithm Design Manual, Springer, 2010.
[27] A. Skowron, J. Stepaniuk, Tolerance approximation spaces, Fundam. Inform. 27 (1996) 245–253.
[28] A. Tversky, Features of similarity, Psychol. Rev. 84 (4) (1977) 327–352.
[29] Y.Y. Yao, Two views of the theory of rough sets in finite universes, Int. J. Approx. Reason. 15 (1996) 291–317.
[30] Y.Y. Yao, T. Wang, On rough relations: an alternative formulations, in: Lecture Notes in Artificial Intelligence, vol. 1711, Springer, 1999, pp. 82–91.