HAL
archives-ouvertes.fr

Semantic(s) of negative sequential patterns

Thomas Guyet

▶ **To cite this version:**

Thomas Guyet. Semantic(s) of negative sequential patterns. JIAF 2019 - 13èmes Journées d'Intelligence Artificielle Fondamentale, Jul 2019, Toulouse, France. pp.1-10. hal-02188501

HAL Id: hal-02188501

https://hal.inria.fr/hal-02188501

Submitted on 18 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Semantic(s) of negative sequential patterns

Thomas Guyet
AGROCAMPUS-OUEST/IRISA – UMR6074
thomas.guyet@irisa.fr

Résumé

Dans le domaine de la fouille de motifs, un motif séquentiel négatif exprime un comportement par une séquence d'événements devant survenir et par des événements qui doivent être absents. Par exemple, le motif $\langle a \neg b\ c\rangle$ décrit l'absence d'un événement b entre les occurrences des événements a et c.

Dans cet article, nous mettons la lumière sur l'ambiguité de cette notation et nous identifions huit sémantiques possibles à la relation d'inclusion d'un motif dans une séquence. Ces sémantiques sont illustrées et nous les étudions formellement. Nous proposons ainsi des relations de dominance et d'équivalence entre ces sémantiques, et nous mettons en évidence de nouvelles propriétés d'anti-monotonie. Ces résultats pourraient être utilisés pour développer de nouveaux algorithmes efficaces pour la fouille de motifs séquentiels négatifs fréquents.

Abstract

In the field of pattern mining, a negative sequential pattern specifies a behavior by a sequence of events that must occur and negative events that must be absent. For instance, the pattern $\langle a \neg b\ c\rangle$ specifies the absence of event b between occurrences of a and c.

In this article, we shed light on the ambiguity of this notation and we identify eight possible semantics for the containment relation between a pattern and a sequence. These semantics are illustrated and formally studied. We propose dominance and equivalence relations between them and we establish new anti-monotonicity properties. These results may be used to develop new algorithms to extract efficiently frequent negative patterns.

1 Introduction

Pattern mining consists in exploring a set of potential patterns to output all but only the most interesting ones. The set of potential patterns can be seen as a search space and the notion of pattern interestingness can be seen as a set of constraints.

The search space in which lies the potential patterns defines a pattern domain. A large number of pattern domains have been proposed in the pattern mining community. The most studied patterns domains [8] are: itemsets, sequential patterns [12] and graph patterns. But a lot of variants from these patterns domains have been proposed. As we focus our attention on temporal data, we can mention: temporal patterns [10], episodes [11] or chronicles [2, 5].

The notion of interestingness studied in pattern mining often refers to the number of its occurrences in a database. A pattern is said to be interesting if it occurs *frequently* in a dataset of examples (frequent sequential pattern mining task). In practice, a pattern is frequent if its number of occurrences is above a user-defined threshold.

Therefore, counting sequences in which the pattern occurs is of paramount importance. It is strongly related to the *containment relation* which decides whether a pattern occurs in a sequence or not. The support measure denotes the counting of sequences in a dataset $\mathcal{D}$ that *contains* the pattern $\boldsymbol{p}$, denoted $supp_{\mathcal{D}}(\boldsymbol{p})$.

The success of pattern mining techniques comes from an anti-monotonicity property of some support measures [1]. Intuitively, if $\boldsymbol{p}$ is not frequent, no "larger" pattern than $\boldsymbol{p}$ is frequent. The pattern mining trick is to prune the search space as soon as an unfrequent pattern as been found. The "is larger than" relation is a topological structure on the set of patterns. As soon as a support measure is anti-monotonic on this topological structure, the frequent pattern mining trick can be used to efficiently prune the search space. Ideally, this structure is a lattice. In this case, the above strategy is complete and correct.

It is worth noticing that the support measure strongly contributes to the semantic of the interestingness of patterns. In fact, output patterns depend on their supports in the dataset(s) of examples. Different support measures have different outputs. For instance, in sequential pattern mining, counting occurrences considering *gap* constraints does not result in the same set of patterns.[1]

In this work, we explore the domain of negative sequential patterns. Sequential patterns describe a sequence of items. The pattern occurs in a sequence while its items appear in the same order in a sequence. For instance, a pattern $\langle a\ b\ c\rangle$ is read as "a occurs and then b occurs and finally c occurs". Negative sequential patterns are sequential patterns with the specification of absent events. Intuitively, the syntax of a negative sequential pattern will be the following $\langle a\ \neg b\ c\rangle$. This pattern is read as "a occurs and then c occurs, but b does not occurs in between".

The apparently intuitive notion of absent event appears to be complex. Few approaches explored what seems to be the same pattern domain [3, 4, 6, 9, 13, 14]. They use the similar notation of negation ($\neg$) but this unique syntax is hiding different semantics.

In this article, we bring the light on eight different semantics of negative sequential patterns and we study them formally: we introduce dominance and equivalence relations between these semantics, and we establish anti-monotony results to thwart the idea that negative sequential patterns are not anti-monotonic.

2 Negative sequential patterns

In the sequel, $[n] = \{1, \dots, n\}$ denotes the set of the first n strictly positive integers. Let $\mathcal{I}$ be the set of items (alphabet). An *itemset* $A = \{a_1\ a_2\ ...\ a_m\} \subseteq \mathcal{I}$ is a set of items. A *sequence* $\boldsymbol{s}$ is a set of sequentially ordered itemsets $\boldsymbol{s} = \langle s_1\ s_2\ ...\ s_n\rangle$: $\forall i, j \in [n]$, $i < j$ means that s_i is located before s_j in sequence $\boldsymbol{s}$ which starts by s_1 and finishes by s_n.

Definition 1 (Negative sequential patterns (NSP)). *A negative pattern* $\boldsymbol{p} = \langle p_1\ \neg q_1\ p_2\ \neg q_2\ \dots\ p_{n-1}\ \neg q_{n-1}\ p_n\rangle$ *is a finite sequence where* $p_i \subseteq \mathcal{I} \setminus \emptyset$ *for all* $i \in [n]$ *and* $q_i \subseteq \mathcal{I}$ *for all* $i \in [n-1]$.

The length *of a NSP, denoted* $|\boldsymbol{p}|$ *is n, its number of itemsets (negative or positive).* $\boldsymbol{p}^+ = \langle p_1\ \dots\ p_n\rangle$ *is so-called the positive part of the NSP.*

We denote by $\mathcal{N}$ the set of negative sequential patterns.

It can be noticed that Definition 1 introduces a syntactic limitation on negative sequential patterns:
- a pattern cannot start neither finish by a negative pattern,
- a pattern cannot have two successive negative itemsets.

Example 1 (Negative sequential pattern). *This example illustrates notations of Definition 1. Let* $\mathcal{I} = \{a, b, c, d\}$ *and* $\boldsymbol{p} = \langle a\ \neg(bc)\ (ad)\ d\ \neg(ab)\ d\rangle$. *We have* $p_1 = \{a\}$, $p_2 = \{ad\}$, $p_3 = \{d\}$, $p_4 = \{d\}$ *and* $q_1 = \{bc\}$, $q_2 = \emptyset$, $q_3 = \{ab\}$. *The length of* $\boldsymbol{p}$ *is* $|\boldsymbol{p}| = 6$ *and* $\boldsymbol{p}^+ = \langle a\ (ad)\ d\ d\rangle$.

3 Semantics of negative sequential patterns

The semantics of negative sequential patterns relies on *negative containment*: a sequence $\boldsymbol{s}$ supports pattern $\boldsymbol{p}$ (or $\boldsymbol{p}$ matches the sequence $\boldsymbol{s}$) iff $\boldsymbol{s}$ contains a sub-sequence $\boldsymbol{s}'$ such that every positive itemset of $\boldsymbol{p}$ is included in some itemset of $\boldsymbol{s}'$ in the same order and for any negative itemset $\neg q_i$ of $\boldsymbol{p}$, q_i is *not included* in any itemset occurring in the sub-sequence of $\boldsymbol{s}'$ located between the occurrence of the positive itemset preceding $\neg q_i$ in $\boldsymbol{p}$ and the occurrence of the positive itemset following $\neg q_i$ in $\boldsymbol{p}$.

Definition 2 (Non inclusion). *We introduce two operators relating two itemsets* $P \subseteq \mathcal{I} \setminus \emptyset$ *and* $I \subseteq \mathcal{I}$:
- *partial non inclusion:* $P \not\preceq I \Leftrightarrow \exists e \in P,\ e \notin I$
- *total non inclusion:* $P \not\sqsubseteq I \Leftrightarrow \forall e \in P, e \notin I$

and, by convention, $\emptyset \not\sqsubseteq I$ *and* $\emptyset \not\preceq I$ *for all* $I \subseteq \mathcal{I}$.

In the sequel we will denote the general form of itemset non inclusion by the symbol $\not\subseteq$, meaning either $\not\preceq$ or $\not\sqsubseteq$.

Intuitively, partial non inclusion considers the itemset as a disjunction of negative constraints, *i.e.* at least one of the item has to be absent, and total non-inclusion consider the itemset as a conjunction of negative constraints: all items have to be absent.

Choosing one non inclusion interpretation or the other has consequences on extracted patterns as well as on pattern search. Let's illustrate this with following sequence dataset:

$$\mathcal{D} = \left\{ \begin{array}{l} \boldsymbol{s}_1 = \langle (bc)\ f\ a\rangle \\ \boldsymbol{s}_2 = \langle (bc)\ (cf)\ a\rangle \\ \boldsymbol{s}_3 = \langle (bc)\ (df)\ a\rangle \\ \boldsymbol{s}_4 = \langle (bc)\ (ef)\ a\rangle \\ \boldsymbol{s}_5 = \langle (bc)\ (cdef)\ a\rangle \end{array} \right\}.$$

Table 1 compares the support of patterns under the two semantics of itemset non-inclusion. Let's consider pattern $\boldsymbol{p}_2$ on sequence $\boldsymbol{s}_2$. Considering that the positive part of $\boldsymbol{p}_2$ is in $\boldsymbol{s}_2$, $\boldsymbol{p}_2$ occurs in the sequence iff $(cd) \not\subseteq (cf)$. In case of total non inclusion, it is false that $(cd) \not\sqsubseteq (cf)$ because of c that occurs in (cf), and thus $\boldsymbol{p}_2$ does not occur in $\boldsymbol{s}_2$. But in case of a partial non inclusion, it is true that $(cd) \not\preceq (cf)$, because of d that does not occur in (cf), and thus $\boldsymbol{p}_2$ occurs in $\boldsymbol{s}_2$.

Lemma 1. [2] *Let* $P, I \subseteq \mathcal{I}$ *be two itemsets:*

$$P \not\sqsubseteq I \implies P \not\preceq I \tag{1}$$

Now, we formulate the notions of sub-sequence, non inclusion and absence by means of the concept of embedding.

1. For instance, a *max*-gap constraint specifies a maximum delay between two successive items of an occurrences.

2. All proofs are provided in Appendix A.

Table 1 – Lists of supported sequences in $\mathcal{D}$ by negative patterns $(\boldsymbol{p}_i)_{i=1..4}$ under the total and partial non inclusion semantics. Each pattern has the shape $\langle a \neg q_i\ b\rangle$ where q_i are itemsets such that $q_i \subset q_{i+1}$.

	partial non inclusion $\not\preceq$	total non inclusion $\not\sqsubseteq$
$\boldsymbol{p}_1 = \langle b\ \neg c\ a\rangle$	$\{\boldsymbol{s}_1, \boldsymbol{s}_3, \boldsymbol{s}_4\}$	$\{\boldsymbol{s}_1, \boldsymbol{s}_3, \boldsymbol{s}_4\}$
$\boldsymbol{p}_2 = \langle b\ \neg(cd)\ a\rangle$	$\{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4\}$	$\{\boldsymbol{s}_1, \boldsymbol{s}_4\}$
$\boldsymbol{p}_3 = \langle b\ \neg(cde)\ a\rangle$	$\{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4\}$	$\{\boldsymbol{s}_1\}$
$\boldsymbol{p}_4 = \langle b\ \neg(cdeg)\ a\rangle$	$\{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4, \boldsymbol{s}_5\}$	$\{\boldsymbol{s}_1\}$

Definition 3 (Positive pattern embedding)**.** *Let* $\boldsymbol{s} = \langle s_1 \ldots s_n\rangle$ *be a sequence and* $\boldsymbol{p} = \langle p_1 \ldots p_m\rangle$ *be a (positive) sequential pattern.* $\boldsymbol{e} = (e_i)_{i\in[m]} \in [n]^m$ *is an* embedding *of pattern* $\boldsymbol{p}$ *in sequence* $\boldsymbol{s}$ *iff* $\forall i \in [m]$, $p_i \subseteq s_{e_i}$ *and* $e_i < e_{i+1}$ *for all* $i \in [m-1]$.

Definition 4 (Strict and soft embeddings of negative patterns)**.** *Let* $\boldsymbol{s} = \langle s_1 \ldots s_n\rangle$ *be a sequence and* $\boldsymbol{p} = \langle p_1\ \neg q_1\ \ldots\ \neg q_{m-1}\ p_m\rangle$ *be a negative sequential pattern.*

$\boldsymbol{e} = (e_i)_{i\in[m]} \in [n]^m$ *is a* ***soft-embedding*** *of pattern* $\boldsymbol{p}$ *in sequence* $\boldsymbol{s}$ *iff:*

— $p_i \subseteq s_{e_i}$, $\forall i \in [m]$

— $q_i \not\subseteq s_j$, $\forall j \in [e_i + 1, e_{i+1} - 1]$ *for all* $i \in [m-1]$

$\boldsymbol{e} = (e_i)_{i\in[m]} \in [n]^m$ *is a* ***strict-embedding*** *of pattern* $\boldsymbol{p}$ *in sequence* $\boldsymbol{s}$ *iff:*

— $p_i \subseteq s_{e_i}$, $\forall i \in [m]$

— $q_i \not\subseteq \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$ *for all* $i \in [m-1]$

Intuitively, the constraint of a negative itemset q_i is checked on the sequence's itemsets at positions in interval $[e_i + 1, e_{i+1} - 1]$, *i.e.* between occurrences of the positive itemset surrounding the negative itemset in the pattern. The soft embedding considers individually each of the sequence's itemsets of $[e_i+1, e_{i+1}-1]$ while strict embedding consider them as a whole.

Notation 1. Soft-*embedding is denoted* $\circ$*-embedding, and* strict-*embedding is denoted* $\bullet$*-embedding.*

Example 2 (Itemset absence semantics)**.** *Let* $\boldsymbol{p} = \langle a\ \neg(bc)\ d\rangle$ *be a pattern and four sequences:*

Sequence	$\not\sqsubseteq$ $\bullet$	$\not\sqsubseteq$ $\circ$	$\not\preceq$ $\bullet$	$\not\preceq$ $\circ$
$\boldsymbol{s_1} = \langle a\ c\ b\ e\ d\rangle$				✓
$\boldsymbol{s_2} = \langle a\ (bc)\ e\ d\rangle$				
$\boldsymbol{s_3} = \langle a\ b\ e\ d\rangle$			✓	✓
$\boldsymbol{s_4} = \langle a\ e\ d\rangle$	✓	✓	✓	✓

One can notice that each sequence contains a unique occurrence of $\langle a\ d\rangle$*, the positive part of pattern* $\boldsymbol{p}$*. Using soft-embedding and partial non inclusion* ($\not\subseteq := \not\preceq$)*,* $\boldsymbol{p}$ *occurs in* $\boldsymbol{s_1}$*,* $\boldsymbol{s_3}$ *and* $\boldsymbol{s_4}$ *but not in* $\boldsymbol{s_2}$*. Using strict-embedding and partial non-inclusion,* $\boldsymbol{p}$ *occurs in sequence* $\boldsymbol{s_3}$ *and* $\boldsymbol{s_4}$*. Indeed, items b and c occur between occurrences of a and d in sequences* 1 *and* 2*. With total non inclusion* ($\not\subseteq := \not\sqsubseteq$) *and either type of embeddings, the absence of an itemset is satisfied if any of its item is absent. As a consequence,* $\boldsymbol{p}$ *occurs only in sequence* $\boldsymbol{s_4}$*.*

Lemma 2. *If* $\boldsymbol{e}$ *is a* $\bullet$*-embedding, then* $\boldsymbol{e}$ *is a* $\circ$*-embedding, whatever is the itemset non-inclusion* ($\not\subseteq$)*.*

Lemma 3. $\boldsymbol{e}$ *is a* $\circ$*-embedding iff* $\boldsymbol{e}$ *is a* $\bullet$*-embedding when* $\not\subseteq := \not\sqsubseteq$*.*

Lemma 4. *Let* $p = \langle p_1\ \neg q_1\ \ldots\ \neg q_{n-1}\ p_n\rangle \in \mathcal{N}$ *s.t.* $|q_i| \leq 1$ *for all* $i \in [n-1]$*, then* $\boldsymbol{e}$ *is a* $\circ$*-embedding iff* $\boldsymbol{e}$ *is a* $\bullet$*-embedding.*

Lemma 4 shows that in the simple case of patterns with negative singleton only, strict and soft-embeddings are equivalent.

Lemma 5. *Let* $p = \langle p_1\ \neg q_1\ \ldots\ \neg q_{n-1}\ p_n\rangle \in \mathcal{N}$*, if* $\boldsymbol{e}$ *is an embedding of pattern* $\boldsymbol{p}$ *in some sequence* $\boldsymbol{s}$*, then* $\boldsymbol{e}$ *is an embedding of the positive sequential pattern* $\boldsymbol{p}^+$ *in* $\boldsymbol{s}$*.*

Example 2 illustrates the impact of itemset non-inclusion operator and of embedding type.

Another point that determines the semantics of negative containment concerns the multiple occurrences of some pattern in a sequence: should all or at least one occurrence(s) of the pattern positive part in the sequence satisfy the non inclusion constraints?

Definition 5 (Negative pattern occurrence)**.** *Let* $\boldsymbol{s}$ *be a sequence,* $\boldsymbol{p}$ *be a negative sequential pattern, and* $\boldsymbol{p}^+$ *the positive part of* $\boldsymbol{p}$*. Let* $\not\subseteq \in \{\not\sqsubseteq, \not\preceq\}$ *be a itemset non-inclusion operator, and* $\mathbin{\text{◐}} \in \{\circ, \bullet\}$ *correspond to the embedding strategy (*$\circ$*: soft-embedding, and* $\bullet$*: strict-embedding).*

— *Pattern* $\boldsymbol{p}$ softly-occurs *in sequence* $\boldsymbol{s}$*, denoted* $\boldsymbol{p} \preceq^{\not\subseteq}_{◐} \boldsymbol{s}$*, iff there exists at least one embedding of* $\boldsymbol{p}$ *in* $\boldsymbol{s}$*.*

— *Pattern* $\boldsymbol{p}$ strictly-occurs *in sequence* $\boldsymbol{s}$*, denoted* $\boldsymbol{p} \sqsubseteq^{\not\subseteq}_{◐} \boldsymbol{s}$*, iff for each embedding* $\boldsymbol{e}$ *of* $\boldsymbol{p}^+$ *in* $\boldsymbol{s}$*,* $\boldsymbol{e}$ *is also an embedding of* $\boldsymbol{p}$ *in* $\boldsymbol{s}$*, and there exists at least one embedding* $\boldsymbol{e}$ *of* $\boldsymbol{p}^+$*.*

Definition 5 allows for capturing two semantics for negative sequential patterns depending on the occurrences of the positive part:

— *strict occurrence*: a negative pattern $\boldsymbol{p}$ occurs in a sequence $\boldsymbol{s}$ iff there exists at least one occurrence of the positive part of pattern $\boldsymbol{p}$ in sequence $\boldsymbol{s}$ and **every** such occurrence satisfies the negative constraints,

— *soft occurrence*: a negative pattern $\boldsymbol{p}$ occurs in a sequence $\boldsymbol{s}$ iff there exists at least one occurrence of the positive part of pattern $\boldsymbol{p}$ in sequence $\boldsymbol{s}$ and **at least one** of these occurrences satisfies the negative constraints.

Example 3 (Strict vs soft occurrence semantics)**.** *Let* $\boldsymbol{p} = \langle a\ b\ \neg c\ d\rangle$ *be a pattern,* $\boldsymbol{s_1} = \langle a\ b\ e\ d\rangle$ *and* $\boldsymbol{s_2} = \langle a\, b\, c\, a\, d\, e\, b\, d\rangle$ *be two sequences.* $\boldsymbol{p}^+ = \langle a\, b\, d\rangle$ *occurs once in* $\boldsymbol{s_1}$ *so there is no difference for occurrences under the two semantics. But, it occurs fourth in* $\boldsymbol{s_2}$ *with embeddings* $(1,2,5)$*,* $(1,2,8)$*,* $(1,7,8)$ *and* $(4,7,8)$*. The two first occurrences do not satisfy the negative constraint* $(\neg c)$ *while the two last occurrences do. Under the soft occurrence semantics, pattern* $\boldsymbol{p}$ *occurs in sequence* $\boldsymbol{s_2}$ *whereas it does not under the strict occurrence semantics.*

Lemma 6. *Let* $\boldsymbol{p}$ *be a NSP and* $\boldsymbol{s}$ *a sequence,*

$$\boldsymbol{p} \sqsubseteq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s} \implies \boldsymbol{p} \preceq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s} \tag{2}$$

where $\LEFTcircle \in \{\circ, \bullet\}$ *and* $\not\subseteq \in \{\not\preceq, \not\sqsubseteq\}$*.*

Lemma 7. *Let* $\boldsymbol{p}$ *be a NSP and* $\boldsymbol{s}$ *a sequence,*

$$\boldsymbol{p} \dashv\!\exists^{\not\sqsubseteq}_{\LEFTcircle} \boldsymbol{s} \implies \boldsymbol{p} \dashv\!\exists^{\not\preceq}_{\LEFTcircle} \boldsymbol{s} \tag{3}$$

where $\dashv\!\exists \in \{\preceq, \sqsubseteq\}$ *and* $\LEFTcircle \in \{\circ, \bullet\}$

In this section, we shown that there are several semantics associated to negative patterns. This leads to eight different types of pattern occurrences. We denote Θ the set of considered pattern occurrence operators:

$$\Theta = \left\{\preceq^{\not\sqsubseteq}_{\circ}, \preceq^{\not\sqsubseteq}_{\bullet}, \preceq^{\not\preceq}_{\circ}, \preceq^{\not\preceq}_{\bullet}, \sqsubseteq^{\not\sqsubseteq}_{\circ}, \sqsubseteq^{\not\sqsubseteq}_{\bullet}, \sqsubseteq^{\not\preceq}_{\circ}, \sqsubseteq^{\not\preceq}_{\bullet}\right\}$$

These operators allows to disambiguate the semantics of negative pattern containment. But, is there no useless distinctions between containment relations? Is there some equivalent containment relations in Θ? The next section answers these questions by introducing the notion of dominance between semantics. Then, we provide some results about the anti-mononicity of these containment relations.

4 Dominance and equivalence between containment relations

Definition 6 (Dominance)**.** *For all* $\theta, \theta' \in \Theta$*,* θ *dominates* θ'*, denoted* $\theta \geqslant \theta'$*, iff* $\boldsymbol{p}\theta\boldsymbol{s} \implies \boldsymbol{p}\theta'\boldsymbol{s}$ *for all* $\boldsymbol{p} \in \mathcal{N}$ *and all sequence* $\boldsymbol{s}$*.*

The idea behind the dominance relation between two containment relations θ and θ' is related to the sequences in which a pattern occurs. By definition, if $\theta \geqslant \theta'$ then for any pattern $\boldsymbol{p} \in \mathcal{N}$, if $\boldsymbol{p}$ occurs in a sequence $\boldsymbol{s}$ according to the θ pattern containment relation, then it also occurs in $\boldsymbol{s}$ according to the θ' pattern containment relation. In the context of pattern mining, such kinds of relation are useful to propose algorithms which could benefit from properties of a dominating containment relation to extract efficiently the patterns according to dominated containment relations.

Notation 2. *We denote by* $\theta \ngeqslant \theta'$ *iff* $\theta \geqslant \theta'$ *is false.*

Lemma 8. *Dominance relation is a pre-order.*

Definition 7 (Equivalent containment relations)**.** *For all* $\theta, \theta' \in \Theta$*,* θ *is equivalent to* θ'*, denoted* $\theta \sim \theta'$ *iff* $\theta \geqslant \theta'$ *and* $\theta' \geqslant \theta$*.*

Lemma 9. $\sim$ *is an equivalence relation on* Θ*.*

Equivalent containment relations have the same semantic. The sets of sequences in which a given pattern occurs are the same and, reciprocally, the sets of negative patterns which occur in a sequence are the same considering these *equivalent* containment relations.

We now study the practical dominance relations we have between the elements of Θ.

Proposition 1. *The following dominances between containment relations hold:*

$$\dashv\!\exists^{\not\subseteq}_{\bullet} \geqslant \dashv\!\exists^{\not\subseteq}_{\circ} \tag{4}$$

$$\dashv\!\exists^{\not\sqsubseteq}_{\circ} \geqslant \dashv\!\exists^{\not\sqsubseteq}_{\bullet} \tag{5}$$

$$\sqsubseteq^{\not\subseteq}_{\LEFTcircle} \geqslant \preceq^{\not\subseteq}_{\LEFTcircle} \tag{6}$$

$$\dashv\!\exists^{\not\sqsubseteq}_{\LEFTcircle} \geqslant \dashv\!\exists^{\not\preceq}_{\LEFTcircle} \tag{7}$$

and the following non-dominance statements hold:

$$\dashv\!\exists^{\not\preceq}_{\LEFTcircle} \ngeqslant \dashv\!\exists^{\not\sqsubseteq}_{\LEFTcircle} \tag{8}$$

$$\sqsubseteq^{\not\subseteq}_{\LEFTcircle} \ngeqslant \preceq^{\not\subseteq}_{\LEFTcircle} \tag{9}$$

$$\dashv\!\exists^{\not\preceq}_{\bullet} \ngeqslant \dashv\!\exists^{\not\preceq}_{\circ} \tag{10}$$

where $\not\subseteq \in \left\{\not\sqsubseteq, \not\preceq\right\}$*,* $\dashv\!\exists \in \{\preceq, \sqsubseteq\}$ *and* $\LEFTcircle \in \{\circ, \bullet\}$*.*

Proposition 1 gathers the results from the previous section. Each line expresses several relationships between pairs of containment relations. Equations 4-7 are dominances deduced from Lemmas 2, 3, 6 and 7. Equations 8-10 states the absence of dominance for which we can exhibit counterexamples. Figure 1 summarizes them. [3]

In addition, many other dominance and non dominance relationships can be deduced from Proposition 1 using the transitivity of dominance (Lemma 8). Nonetheless, in this work, some relationships between pairs of containment relations can not be deduced from Proposition 1. For instance, the relationship between $\sqsubseteq^{\not\preceq}_{\circ}$ and $\preceq^{\not\preceq}_{\bullet}$ is not determined by Proposition 1.

One interesting result of Proposition 1 is that there are two pairs of containment relations, $\left(\sqsubseteq^{\not\sqsubseteq}_{\circ}, \sqsubseteq^{\not\sqsubseteq}_{\bullet}\right)$ and $\left(\sqsubseteq^{\not\preceq}_{\circ}, \sqsubseteq^{\not\preceq}_{\bullet}\right)$, whose elements are equivalent.

3. Assuming that two containment relations are not equals iff they are different, the dominance relation is a pre-order. The property is not central in the following. Figure 1 illustrates clearly that it is not a partial order. Indeed, two pairs of different containment relations are symmetrically dominated.

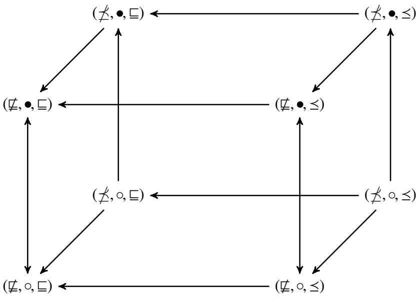

Figure 1 – Summary of the dominance relation between relations of Θ. An arrow shows that the operator at the destination is dominated by the operator at the origin. Unidirectional arrows indicates that the dominance relation holds in one direction but not in the other direction.

It ensues that they are six equivalent classes of containment relations: $\preceq_{\circ}^{\nsubseteq}$, $\preceq_{\bullet}^{\nsubseteq}$, $\preceq_{\circ}^{\not\subseteq}$, $\preceq_{\bullet}^{\not\subseteq}$, $\left\{\sqsubseteq_{\circ}^{\nsubseteq}, \sqsubseteq_{\bullet}^{\nsubseteq}\right\}$ and $\left\{\sqsubseteq_{\circ}^{\not\subseteq}, \sqsubseteq_{\bullet}^{\not\subseteq}\right\}$.

We can finally point out that Lemma 4 adds a dominance relation when negative sequential patterns are restricted to have singleton negative itemsets. In this case, the equivalent classes become: $\left\{\preceq_{\circ}^{\nsubseteq}, \preceq_{\bullet}^{\nsubseteq}\right\}$, $\left\{\preceq_{\circ}^{\not\subseteq}, \preceq_{\bullet}^{\not\subseteq}\right\}$, $\left\{\sqsubseteq_{\circ}^{\nsubseteq}, \sqsubseteq_{\bullet}^{\nsubseteq}\right\}$ and $\left\{\sqsubseteq_{\circ}^{\not\subseteq}, \sqsubseteq_{\bullet}^{\not\subseteq}\right\}$.

5 Anti-monotonicity

Our question is now to know whether there are containment relations that have more interesting properties. In our original context of mining frequent negative sequential patterns, we investigate the anti-monotonicity properties.

According to Zheng et *al.* [14], "the APriori principle doesn't apply to negative sequential pattern". The "APriori principle" can be understood as the anti-monotonicity property. We will see that assertion is actually only partially true.

The anti-monotonicity makes sense only with a partial order on the set of NSPs. We first introduce different possible partial orders and then we introduce the anti-monotonicity.

For sake of conciseness, the remaining of the section assumes that $\not\subseteq := \nsubseteq$. Thus, we can count on the anti-monotonicity of the non inclusion of itemsets: $q \subseteq q' \implies q' \not\subseteq q$ for all $q, q' \subseteq \mathcal{I}$. The following results can be extended to the case $\not\subseteq := \not\subseteq$ by reversing the inclusion relations for negatives in the partial orders.

5.1 Partial orders

Definition 8 introduces several relations between negative sequential patterns that are partial orders (see Proposition 2).

Definition 8 (NSP relations). *Let $\boldsymbol{p} = \langle p_1\ \neg q_1\ p_2\ \neg q_2\ \cdots\ p_{k-1}\ \neg q_{k-1}\ p_k\rangle$ and $\boldsymbol{p}' = \langle p'_1\ \neg q'_1\ p'_2\ \neg q'_2\ \cdots\ p'_{k'-1}\ \neg q'_{k'-1}\ p'_{k'}\rangle$ be two NSPs.*

By definition, $\boldsymbol{p} \lhd \boldsymbol{p}'$ iff $k \leq k'$ and $\exists (u_i)_{i\in[k]} \in [k']^k$ s.t.:

1. $\forall i \in [k],\ p_i \subseteq p'_{u_i}$
2. $\forall i \in [k-1],\ q_i \subseteq \bigcup_{j\in[u_i,u_{i+1}-1]} q'_j$
3. $k = k' \implies \exists j \in [k],\ p_j \neq p'_j$ *or* $q_j \neq q'_j$
4. $u_i < u_{i+1}$, *for all* $i \in [k-1]$

by definition, $\boldsymbol{p} \lhd \boldsymbol{p}'$ iff $k \leq k'$ s.t.:

1. $\forall i \in [k],\ p_i \subseteq p'_i$
2. $\forall i \in [k-1],\ q_i \subseteq q'_i$
3. $k = k' \implies p_k \neq p'_k$ *or* $\exists j \in [k-1]$ *s.t.* $q_j \neq q'_j$

and, by definition, $\boldsymbol{p} \lhd \boldsymbol{p}'$ iff $k = k'$ s.t.:

1. $\forall i \in [k],\ p_i = p'_i$
2. $\forall i \in [k-1],\ q_i \subseteq q'_i$
3. $\exists j \in [k-1]$ *s.t.* $q_j \neq q'_j$

The $\lhd$ relation can be seen as the "classical" inclusion relation between sequential patterns [12]. A NSP $\boldsymbol{p}$ is less specific than $\boldsymbol{p}'$ iff $\boldsymbol{p}^+$ is a subsequence of $\boldsymbol{p}'^+$ and negative constraints are satisfied. The principal difference with $\lhd$ is that $\lhd$ permits to insert new positive itemsets in the middle of the sequence while $\lhd$ permits only insertion of new positive itemsets at the end.[4, 5] Nonetheless, it is still possible to insert items to the positive itemsets. The $\lhd$ does not even permit such differences: each pair of positive itemsets must be equals to have comparable NSP.

Lemma 10. *For all $\boldsymbol{p}, \boldsymbol{p}' \in \mathcal{N}$,*

$$\boldsymbol{p} \lhd \boldsymbol{p}' \implies \boldsymbol{p} \lhd \boldsymbol{p}' \implies \boldsymbol{p} \lhd \boldsymbol{p}' \qquad (11)$$

Proposition 2 (Strict partial orders). *$\lhd$, $\lhd$ and $\lhd$ are partial orders on $\mathcal{N}$.*

We can notice that the third conditions in Definition 8 enforce the relations to be irreflexive. Removing these conditions enables to define non-strict partial orders.

5.2 Anti-monotonicity

Let us first define the anti-monotonicity property of a containment relation $\theta \in \Theta$ considering a strict partial order $\ltimes \in \{\lhd, \lhd, \lhd\}$.

Definition 9 (Anti-monotonicity on $(\mathcal{N}, \ltimes)$). *Let $\theta \in \Theta$ be a containment relation, θ is anti-monotonic on $(\mathcal{N}, \ltimes)$ iff for all $\boldsymbol{p}, \boldsymbol{p}' \in \mathcal{N}$ and all sequence $\boldsymbol{s}$:*

$$\boldsymbol{p} \ltimes \boldsymbol{p}' \implies (\boldsymbol{p}'\theta\boldsymbol{s} \implies \boldsymbol{p}\theta\boldsymbol{s})$$

4. In sequential pattern mining, it is called a *backward*-extension of the patterns.

5. We remind that, by Definition 1, $p_i \neq \emptyset$ and that we never have two successive negative itemsets in a NSP.

First of all, we provide an example showing that none of the containment relation is monotonic on $(\mathcal{N}, \unlhd)$. Let $\boldsymbol{p} = \langle b\ \neg c\ a\rangle$, $\boldsymbol{p}' = \langle b\ \neg c\ d\ a\rangle$ and $\boldsymbol{s} = \langle b\ e\ d\ c\ a\rangle$. Then, we have $\boldsymbol{p} \unlhd \boldsymbol{p}'$.[6] Nonetheless, for each $\theta \in \Theta$, $\boldsymbol{p}'\theta\boldsymbol{s}$ but it is false that $\boldsymbol{p}\theta\boldsymbol{s}$. In fact, the presence of the item d in the sequence changes the scope for checking the absence of c.

This example illustrates the case of Zheng et *al.* [14] to argue for the absence of anti-monotonic property for negative patterns. But, using partial orders that prevent from changing the scope for absent items enables to exhibit anti-monotonicity properties.

Proposition 3. $\preceq^{\nsubseteq}_{\LEFTcircle}$ *is anti-monotonic on* $(\mathcal{N}, \lhd)$*, where* $\LEFTcircle \in \{\circ, \bullet\}$.

Proposition 3 shows that using the $\lhd$ order leads to have anti-monotonicity only for containment with soft-occurrence, but not with strict-occurrence. Let us give a counterexample illustrating the problem with strict-occurrence. Let $\boldsymbol{p} = \langle a\ \neg b\ c\rangle$, $\boldsymbol{p}' = \langle a\ \neg b\ c\ d\rangle$ and $\boldsymbol{s} = \langle a\ c\ d\ a\ b\ c\rangle$. Then, we have $\boldsymbol{p} \lhd \boldsymbol{p}'$.[7] Nonetheless, $\boldsymbol{p}' \sqsubseteq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s}$ holds but it is false that $\boldsymbol{p} \sqsubseteq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s}$. In fact, without the presence of the item d in the pattern, there are three possible embeddings of $\boldsymbol{p}$ in $\boldsymbol{s}$. Considering $\sqsubseteq^{\not\subseteq}_{\LEFTcircle}$ each embedding must satisfy the negation of b, which is not the case, while it is sufficient to have only one embedding satisfying negations for $\preceq^{\not\subseteq}_{\LEFTcircle}$.

The previous example illustrates the problem while extending the pattern with additional itemsets. The same issue is encountered with the following example considering same length patterns but with an extended itemset. Let $\boldsymbol{p} = \langle a\ \neg b\ c\rangle$, $\boldsymbol{p}' = \langle a\ \neg b\ (cd)\rangle$ and $\boldsymbol{s} = \langle a\ (cd)\ a\ b\ c\rangle$. Then, we have $\boldsymbol{p} \lhd \boldsymbol{p}'$. Nonetheless, $\boldsymbol{p}' \sqsubseteq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s}$ holds but it is false that $\boldsymbol{p} \sqsubseteq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s}$.

Proposition 4. $\dashv^{\nsubseteq}_{\LEFTcircle}$ *is anti-monotonic on* $(\mathcal{N}, \unlhd)$*, where* $\dashv \in \{\preceq, \sqsubseteq\}$ *and* $\LEFTcircle \in \{\circ, \bullet\}$.

We remind that this section presented the case of total non-inclusion ($\not\subseteq := \nsubseteq$) but similar results can be obtained with partial non-inclusion.

6 Come back to pattern mining

The definitions of pattern support, frequent pattern and pattern mining derive naturally from the notion of occurrence of a negative sequential pattern, no matter the choices for embedding (soft or strict), non inclusion (partial or total) and occurrences (soft or strict). However, these choices concerning the semantics of NSPs impact directly the number of frequent patterns (under the same minimal threshold constraint) and further the computation time. The stronger the negative constraints, the lesser the number of sequences containing a pattern, and the lesser the number of frequent patterns.

Definition 10 (Pattern supports)**.** *Let* $\mathcal{D} = \{\boldsymbol{s}_i\}_{i\in[n]}$ *be a set of n sequences and* $\boldsymbol{p}$ *be a NSP. The support of* $\boldsymbol{p}$ *in* $\mathcal{D}$*, denoted* $supp^{\mathcal{D}}_{\theta}(p)$ *is the number of sequences of* $\mathcal{D}$ *in which* $\boldsymbol{p}$ *occurs according to the* $\theta \in \Theta$ *containment relation.*

Notation 3. *When there is no ambiguity on the dataset of sequences,* $supp^{\mathcal{D}}_{\theta}(p)$ *is denoted* $supp_{\theta}(p)$*. And, for sake of readability, the* θ *operators are represented as triplets.*

It is clear that if a containment relation θ is dominated by another containment relation θ', then the Proposition 1 implies that the support of the pattern evaluated with θ is lower than the support of the pattern evaluated with θ'. Thus, we have the following proposition.

Proposition 5. *For all pattern* $\boldsymbol{p} \in \mathcal{N}$*:*

$$supp_{\not\subseteq,\bullet,\dashv}(\boldsymbol{p}) \leq supp_{\not\subseteq,\circ,\dashv}(\boldsymbol{p}) \tag{12}$$

$$supp_{\nsubseteq,\circ,\dashv}(\boldsymbol{p}) \leq supp_{\nsubseteq,\bullet,\dashv}(\boldsymbol{p}) \tag{13}$$

$$supp_{\not\subseteq,\LEFTcircle,\sqsubseteq}(\boldsymbol{p}) \leq supp_{\not\subseteq,\LEFTcircle,\preceq}(\boldsymbol{p}) \tag{14}$$

$$supp_{\nsubseteq,\LEFTcircle,\dashv}(\boldsymbol{p}) \leq supp_{\not\subset,\LEFTcircle,\dashv}(\boldsymbol{p}) \tag{15}$$

In addition, we can also deduce anti-monotonicity properties for support measures from the Propositions 3 and 4.

Proposition 6. *For all pairs of NSPs* $\boldsymbol{p}, \boldsymbol{p}' \in \mathcal{N}$*:*

$$\boldsymbol{p} \lhd \boldsymbol{p}' \implies supp_{\nsubseteq,\LEFTcircle,\preceq}(\boldsymbol{p}') \leq supp_{\nsubseteq,\LEFTcircle,\preceq}(\boldsymbol{p}) \tag{16}$$

$$\boldsymbol{p} \unlhd \boldsymbol{p}' \implies supp_{\nsubseteq,\LEFTcircle,\dashv}(\boldsymbol{p}') \leq supp_{\nsubseteq,\LEFTcircle,\dashv}(\boldsymbol{p}) \tag{17}$$

Then, there is two ways to use these results to implement efficient frequent NSP mining algorithms. On the one hand, the results from Proposition 6 can be directly used to implement algorithms with efficient and correct strategies to prune the search space.[8] For $\preceq^{\nsubseteq}_{\LEFTcircle}$ containment relation, Equation 16 fully exploits the $\lhd$ partial order to early prune a priori unfrequent patterns. For $\sqsubseteq^{\nsubseteq}_{\LEFTcircle}$ containment relation, the $\unlhd$ partial order must be used to ensure the correctness of the algorithm (Equation 17). Unfortunalety, this partial order is less interesting than $\lhd$ because it gives a priori information on less patterns than $\lhd$ does. On the other hand, the support evaluated with $\preceq^{\nsubseteq}_{\LEFTcircle}$ is an upper bound for the support of $\sqsubseteq^{\nsubseteq}_{\LEFTcircle}$ (Equation 14). Thus, it is possible also to prune patterns accessible with $\lhd$ partial order without losing the correctness of the pruning strategy.

7 Conclusion and perspectives

In this article, we explored the semantics of negation in sequential patterns. We gave eight possible semantics

6. In this case, we do not have $\boldsymbol{p} \lhd \boldsymbol{p}'$ nor $\boldsymbol{p} \unlhd \boldsymbol{p}'$

7. In this case, we also have $\boldsymbol{p} \lhd \boldsymbol{p}'$ but not $\boldsymbol{p} \unlhd \boldsymbol{p}'$

8. The completeness of the algorithms requires to study how to traverse the search space. It is out of the scope of this article.

where the state of the art in sequential pattern mining did not notice these differences. We investigated the formal properties of these semantics: their respective relations (dominance and equivalence) and anti-monotonicity properties. These results may be used to develop new efficient algorithms to extract negative sequential patterns.

It is worth noticing that no semantics is "more" correct or relevant than another one. It depends on the information to be captured. Our objective is to give the opportunity to make an informed choice. Even if, in the context of pattern mining, the choice is constrained by computational considerations. With the three proposed partial orders, we have seen that interesting anti-monotonicity holds only for some semantics. Thus, dominance relations may be used to propose alternative search heuristics.

Hence, the first perspective of this work is to develop negative sequential pattern mining algorithms. This has been done for containment relations $\prec_{◐}^{\nsubseteq}$ [7].[9] In this algorithm, we also introduced *maxgap* and *maxspan* constrains negative sequential patterns. The anti-mononicity properties still hold with these constraints. A step further, the framework presented in this article would make possible to propose a complete and correct algorithm to mine closed NSP.

Our second perspective is to provide a complete results on the dominance between containment relations. There are still few cases that are undetermined to achieve this objective. Further studies aim at having results on the quotient set $\Theta/\sim$.

Our third perspective is to propose new intuitive syntax(es) of containment relations. The main problem was the semantic overload of the negation symbol ($\neg$) in the literature of negative sequential patterns. We are currently working on some proposals and plan to evaluate them. To be intuitive, these proposals are based on the equivalent classes we highlight in this work.

References

[1] Agrawal, Rakesh et Ramakrishnan Srikant: *Fast Algorithms for Mining Association Rules in Large Databases*. Dans *Proc. of the Int. Conf. on Very Large Data Bases (VLDB)*, pages 487–499, 1994.

[2] Besnard, Philippe et Thomas Guyet: *Chronicles*. à paraître, 2019.

[3] Cao, Longbing, Xiangjun Dong et Zhigang Zheng: *e-NSP: Efficient negative sequential pattern mining*. Artificial Intelligence, 235:156–182, 2016.

[4] Chen, Yen-Liang, Mei-Ching Chiang et Ming-Tat Ko: *Discovering time-interval sequential patterns in sequence databases*. Expert System with Applications, 25(3):343–354, 2003.

[5] Dauxais, Yann, Thomas Guyet, David Gross-Amblard et André Happe: *Discriminant chronicles mining - application to care pathways analytics*. Dans *Proc. of Conf. on Artificial Intelligence in Medicine (AIME)*, pages 234–244, 2017.

[6] Gong, Yongshun, Tiantian Xu, Xiangjun Dong et Guohua Lv: *e-NSPFI: Efficient mining negative sequential pattern from both frequent and infrequent positive sequential patterns*. Int. Journal of Pattern Recognition and Artificial Intelligence, 31(02):1750002, 2017.

[7] Guyet, Thomas et René Quiniou: *NegPSpan: efficient extraction of negative sequential patterns with embedding constraints*. CoRR, abs/1804.01256, 2018.

[8] Han, Jiawei, Micheline Kamber et Jian Pei: *Advanced pattern mining*. Dans *Data Mining (Third Edition)*, pages 279–325. Morgan Kaufmann, 2012.

[9] Hsueh, Sue Chen, Ming Yen Lin et Chien Liang Chen: *Mining negative sequential patterns for e-commerce recommendations*. Dans *Proc. of Asia-Pacific Services Computing Conference*, pages 1213–1218, 2008.

[10] Laxman, Srivatsan et P. S. Sastry: *A survey of temporal data mining*. Sadhana, 31(2):173–198, 2006.

[11] Mannila, Heikki, Hannu Toivonen et A. Inkeri Verkamo: *Discovery of frequent episodes in event sequences*. Data Mining and Knowledge Discovery, 1(3):259–289, 1997.

[12] Mooney, Carl H. et John F. Roddick: *Sequential pattern mining – approaches and algorithms*. ACM Computing Survey, 45(2):1–39, 2013.

[13] Xu, Tiantian, Xiangjun Dong, Jianliang Xu et Yongshun Gong: *E-msNSP: Efficient negative sequential patterns mining based on multiple minimum supports*. Int. Journal of Pattern Recognition and Artificial Intelligence, 31(02):1750003, 2017.

[14] Zheng, Zhigang, Yanchang Zhao, Ziye Zuo et Longbing Cao: *Negative-GSP: An efficient method for mining negative sequential patterns*. Dans *Proc. of the Australasian Data Mining Conference*, pages 63–67, 2009.

A Proofs

Proof of Lemma 1. Let $P, I \subseteq \mathcal{I}$ s.t. $P \nsubseteq I$. If $P = \emptyset$, by definition, $P \npreceq I$. Otherwise, because P is not empty, then there exists $e \in P$ s.t. $e \notin I$, *i.e.* $P \npreceq I$. □

9. A demo of such an algorithm is accessible online: `http://people.irisa.fr/Thomas.Guyet/negativepatterns/evalnegpat.php`.

Proof of Lemma 2. Let $\boldsymbol{e} = (e_i)_{i\in[m]} \in [n]^m$ be a $\bullet$-embedding of a NSP $\boldsymbol{p} = \langle p_1\ \neg q_1\ \dots\ \neg q_{m-1}\ p_m\rangle$ in a sequence $\boldsymbol{s} = \langle s_1\ \dots\ s_n\rangle$. For all positive itemsets p_i, the definition of $\bullet$-embedding matches the one for $\circ$-embedding. For a negative itemset q_i, let us start with $\not\subset := \npreceq$. By definition 4, $q_i \npreceq \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$, and by Definition 2, $\exists\alpha \in q_i,\ \alpha \notin \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$. And then, $\exists\alpha \in q_i,\ \forall j \in [e_i+1, e_{i+1}-1],\ \alpha \notin s_j$. That is $\forall j \in [e_i+1, e_{i+1}-1],\ \exists\alpha \in q_i,\ \alpha \notin s_j$. This shows $\forall j \in [e_i+1, e_{i+1}-1],\ q_i \npreceq s_j$ ($\circ$-embedding definition). It remains $\not\subset := \nsubseteq$. By definition 4, $q_i \nsubseteq \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$, and by Definition 2, $\forall\alpha \in p_i,\ \alpha \notin \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$. And then, $\forall\alpha \in q_i,\ \forall j \in [e_i+1, e_{i+1}-1],\ \alpha \notin s_j$. That is $\forall j \in [e_i+1, e_{i+1}-1],\ \forall\alpha \in q_i,\ \alpha \notin s_j$. This shows $\forall j \in [e_i+1, e_{i+1}-1],\ q_i \not\subset s_j$. □

Proof of Lemma 3. Let $\boldsymbol{s} = \langle s_1\ \dots\ s_n\rangle$ be a sequence and $\boldsymbol{p} = \langle p_1\ \neg q_1\ \dots\ \neg q_{m-1}\ p_m\rangle$ be a negative sequential pattern. Lemma 2 shows that $\bullet$-embedding implies $\circ$-embedding. It remains the implication to the left. Let $\boldsymbol{e} = (e_i)_{i\in[m]} \in [n]^m$ be a $\circ$-embedding of pattern $\boldsymbol{p}$ in sequence $\boldsymbol{s}$. Then, the definition matches the one for $\bullet$-embedding for positives, p_i. For negatives, q_i, then $\forall j \in [e_i+1, e_{i+1}-1],\ q_i \nsubseteq s_j$, *i.e.* $\forall j \in [e_i+1, e_{i+1}-1],\ \forall\alpha \in q_i,\ \alpha \notin s_j$ and then $\forall\alpha \in q_i,\ \forall j \in [e_i+1, e_{i+1}-1],\ \alpha \notin s_j$. It thus implies that $\forall\alpha \in p_i,\ \alpha \notin \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$, *i.e.* by definition, $q_i \nsubseteq \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$. □

Proof of Lemma 4. Let $\boldsymbol{s} = \langle s_1\ \dots\ s_n\rangle$ be a sequence and $\boldsymbol{p} = \langle p_1\ \neg q_1\ \dots\ \neg q_{m-1}\ p_m\rangle$ be a NSP s.t. $\forall i,\ |q_i| \leq\leq 1$. Let $\boldsymbol{e} = (e_i)_{i\in[m]} \in [n]^m$ be a $\circ$-embedding of $\boldsymbol{p}$ in $\boldsymbol{s}$ then, by definition, 1) $p_i \subseteq s_{e_i}$ for all $i \in [m]$ and 2) $q_i \not\subset s_j$ for all $j \in [e_i+1, e_{i+1}-1]$. In case $|q_i| = 0$, there is no constraint. In case $|q_i| = 1$, and 2) becomes $q_i \notin s_j$ for all $j \in [e_i+1, e_{i+1}-1]$ whatever $\not\subset \in \{\npreceq, \nsubseteq\}$. Hence, $q_i \notin \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$ *i.e.* $q_i \not\subset \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$ (no matter $\npreceq$ or $\nsubseteq$). As a consequence $\boldsymbol{e}$ is a $\bullet$-embedding of $\boldsymbol{p}$. □

Proof of Lemma 5. Let $\boldsymbol{s} = \langle s_1\ \dots\ s_n\rangle$ be a sequence and $\boldsymbol{p} = \langle p_1\ \neg q_1\ \dots\ \neg q_{m-1}\ p_m\rangle \in \mathcal{N}$ be a pattern. By definition 4, if $\boldsymbol{e} = (e_i)_{i\in[m]} \in [n]^m$ is an embedding of pattern $\boldsymbol{p}$ in sequence $\boldsymbol{s}$ then $\forall i \in [l]$: $p_i \subseteq s_{e_i}$ because p_i is positive. According to definition 3, $\boldsymbol{e}$ is an embedding of the positive pattern $\boldsymbol{p}^+$. □

Proof of Lemma 6. Let $\boldsymbol{s} = \langle s_1\ \dots\ s_n\rangle$ be a sequence and $\boldsymbol{p} = \langle p_1\ \neg q_1\ \dots\ \neg q_{m-1}\ p_m\rangle \in \mathcal{N}$ be a pattern s.t. $p \sqsubseteq^{\not\subset}_{◐} s$. Then, there exists $\boldsymbol{e}$ an embedding of $\boldsymbol{p}^+$ in $\boldsymbol{s}$ and, by definition, it is also an embedding of $\boldsymbol{p}$ in $\boldsymbol{s}$. This means that $p \preceq^{\not\subset}_{◐} s$. □

Proof of Lemma 7. Let $\boldsymbol{s} = \langle s_1\ \dots\ s_n\rangle$ be a sequence and $\boldsymbol{p} = \langle p_1\ \neg q_1\ \dots\ \neg q_{m-1}\ p_m\rangle \in \mathcal{N}$ be a pattern.

We start by considering relations between semantics at the embedding level, and then we will conclude at the pattern level.

Let's first assume that $◐ := \bullet$. Then, for all $(\bullet, \nsubseteq)$-embedding $e = (e_i)_{i\in[m]}$ of pattern $\boldsymbol{p}$ in sequence $\boldsymbol{s}$. Hence, $\forall i \in [m],\ p_i \subseteq s_{e_i}$ and $\forall i \in [m-1], \forall j \in [e_i+1, e_{i+1}-1],\ q_i \nsubseteq s_j$. According to eq. 1, we have $q_i \npreceq s_j$. It comes that e is a $(\bullet, \preceq)$-embedding.

Let's now assume that $◐ := \circ$. Then, let $e = (e_i)_{i\in[m]}$ be a $(\circ, \nsubseteq)$-embedding of pattern $\boldsymbol{p}$ in sequence $\boldsymbol{s}$. Hence, $\forall i \in [m],\ p_i \subseteq s_{e_i}$ and $\forall i \in [m-1],\ q_i \nsubseteq \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$. In addition $\bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j \subseteq \mathcal{I}$, then, $\forall i \in [m-1],\ q_i \npreceq \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$ according to eq. 1. It comes that e is an $(\circ, \preceq)$-embedding.

Let's come back at the pattern level. if $p \dashv^{\nsubseteq}_{◐} s$, in the two cases ($\dashv \in \{\preceq, \sqsubseteq\}$). In the first case the existing $(◐, \nsubseteq)$-embedding is a $(◐, \npreceq)$-embedding, and in the second case, all $(◐, \nsubseteq)$-embeddings are $(◐, \npreceq)$-embeddings. Therefore, we have that $p \dashv^{\npreceq}_{◐} s$. □

Proof of Lemma 8. A pre-order is a reflexive, transitive binary relation. The reflexivity of the relation comes with Definition 6. Let $\theta, \theta', \theta'' \in \Theta$ be three dominance relations s.t. $\theta \geqslant \theta'$ and $\theta' \geqslant \theta''$. Then, for all $\boldsymbol{p} \in \mathcal{N}$ and sequence $\boldsymbol{s}$: $\boldsymbol{p}\theta\boldsymbol{s} \implies \boldsymbol{p}\theta'\boldsymbol{s}$ and $\boldsymbol{p}\theta'\boldsymbol{s} \implies \boldsymbol{p}\theta''\boldsymbol{s}$. Hence, we have, $\boldsymbol{p}\theta\boldsymbol{s} \implies \boldsymbol{p}\theta''\boldsymbol{s}$, *i.e.* $\theta \geqslant \theta''$. □

Proof of Lemma 9. Let $\theta, \theta' \in \Theta$, by reflexivity of $\geqslant$ we have that $\sim$ is reflexive. By definition $(\theta \geqslant \theta' \wedge \theta' \geqslant \theta)$, $\sim$ is symmetric. And , $\sim$ is also transitive. Let $\theta, \theta', \theta'' \in \Theta$ be three dominance relation s.t. $\theta \sim \theta'$ and $\theta' \sim \theta''$ then, $\theta \geqslant \theta'$, $\theta' \geqslant \theta''$, $\theta' \geqslant \theta$ and $\theta'' \geqslant \theta'$. Hence, by transitivity of $\geqslant$, $\theta \geqslant \theta''$ and $\theta'' \geqslant \theta$, $\theta \sim \theta''$. □

Proof of Proposition 1. Let $\boldsymbol{p} \in \mathcal{N}$ and $\boldsymbol{s}$ a sequence.

According to Lemma 6, $\boldsymbol{p} \sqsubseteq^{\not\subset}_{◐} \boldsymbol{s} \implies \boldsymbol{p} \preceq^{\not\subset}_{◐} \boldsymbol{s}$. Thus we obtain Equality 6 by Definition 6. is immediately deduced from .

According to Lemma 7, $\boldsymbol{p} \dashv^{\nsubseteq}_{◐} \boldsymbol{s} \implies \boldsymbol{p} \dashv^{\npreceq}_{◐} \boldsymbol{s}$. Thus we obtain Equality 7 by Definition 6.

According to Lemma 2, a $\bullet$-embedding is a $\circ$-embedding whatever the itemset non-inclusion operator. Then, we can conclude that $\boldsymbol{p} \dashv^{\not\subset}_{\bullet} \boldsymbol{s} \implies \boldsymbol{p} \dashv^{\not\subset}_{\circ} \boldsymbol{s}$ (Equality 4).

In addition, Lemma 3 shows that a $\circ$-embedding is $\bullet$-embedding in case of total itemset non-inclusion. Then, we can conclude that $\boldsymbol{p} \dashv^{\nsubseteq}_{\circ} \boldsymbol{s} \implies \boldsymbol{p} \dashv^{\nsubseteq}_{\bullet} \boldsymbol{s}$ (Equality 5).

Let now gives some counterexamples for known non-dominance relationships. A counterexample for Equation 8 is $p = \langle a\ \neg(bc)\ d\rangle$ and $s = \langle a\ b\ d\rangle$. We have that $p \dashv^{\npreceq}_{◐} s$ but $p \dashv^{\nsubseteq}_{◐} s$ is false.
Similarly, a counterexample for Equation 9 is $p = \langle a\ \neg b\ c\rangle$ and $s = \langle a\ c\ b\ c\rangle$, and a counterexample for Equation 10 is $p = \langle a\ \neg(bc)\ d\rangle$ and $s = \langle a\ b\ c\ d\rangle$. □

Proof of Lemma 10. We start with the implication $\boldsymbol{p} \trianglelefteq \boldsymbol{p}' \implies \boldsymbol{p} \lhd \boldsymbol{p}'$. Let $p, p' \in \mathcal{N}$ s.t. $\boldsymbol{p} \trianglelefteq \boldsymbol{p}'$. By definition, $k = k'$ and 1. $\forall i \in [k],\ p_i = p'_i$, 2. $\forall i \in [k-1],\ q_i \subseteq q'_i$

and 3. $\exists j \in [k-1]$ s.t. $q_i \neq q'_i$. A particular case of 1. is that $\forall i \in [k],\ p_i \subseteq p'_i$. In addition, third condition of $\lhd$ is obtained easily from 3. adding a disjunctive condition. Hence, $\boldsymbol{p} \lhd \boldsymbol{p}'$.

We now prove the second implication: $\boldsymbol{p} \lhd \boldsymbol{p}' \implies \boldsymbol{p} \lessdot \boldsymbol{p}'$. Let $p, p' \in \mathcal{N}$ s.t. $\boldsymbol{p} \lhd \boldsymbol{p}'$. Let's now define the sequence u_i such that $u_i = i$ for all $i \in [k]$. By construction, we have that $u_i < u_{i+1}$, for all $i \in [k-1]$ (4.). In addition, by definition of $\lhd$, we have that $\forall i \in [k],\ p_i \subseteq p'_i = p'_{u_i}$, and $\forall i \in [k-1], q_i \subseteq q'_i = \bigcup_{j\in[i,(i+1)-1]} q'_j = \bigcup_{j\in[u_i,u_{i+1}-1]} q'_j$. Assuming $k = k'$, then $p_k \neq p'_k$ or $\exists j \in [k-1]$ s.t. $q_j \neq q'_j$. If $p_k \neq p'_k$ the third condition of $\lessdot$ is satisfied (with $j = k$). Otherwise, it is also satisfied with the j of the definition of $\lhd$. □

Proof of Proposition 2. We begin with the $\lessdot$ relation. We first remind that $\lessdot$ is a strict partial order iff the three following conditions hold:

1. $\forall p \in \mathcal{N}$, not $p \lessdot p$ (irreflexive),
2. $\forall p, p', p'' \in \mathcal{N}$, $p \lessdot p'$ and $p' \lessdot p'' \implies p \lessdot p''$ (transitivity),
3. $\forall p, p' \in \mathcal{N}$, $p \lessdot p' \implies$ not $p' \lessdot p$ (asymmetry)

Irreflexive. Let's assume that $\exists \in \mathcal{N}$ s.t. $p \lessdot p$. Then, because $k = k'$ the third condition implies that $\exists j \in [k-1]$ s.t. $q_j \neq q_j$, which is absurd. Then $\lessdot$ is irreflexive.
Transitivity. Let $p, p', p'' \in \mathcal{N}$ s.t. $p \lessdot p'$ and $p' \lessdot p''$. Then, for all $i \in [k]$, $p_i = p'_i = p''_i$ and for all $i \in [k-1]$, $q_i \subseteq q'_i \subseteq q''_i$ ($k = k' = k''$). Finally, it is not possible to have $q_i = q''_i$ for all $i \in [k-1]$. In fact, if we have these equalities, then we would have $q_i = q'_i$ and $q'_i = q''_i$ for all $i \in [k-1]$ because $q_i \subseteq q'_i \subseteq q''_i$. But, it is not possible according to 3. Therefore, we have that $p \lessdot p''$.
Asymmetry. Let $p, p' \in \mathcal{N}$ s.t. $p \lessdot p'$. Then, according to 2. and 3., there exists $j \in [k-1]$ s.t. $q_j \subsetneq q'_j$. And then, it is not possible to have $q'_j \subseteq q_j$. As a consequence, we can not have $p' \lessdot p$.
We now prove that $\lhd$ is a strict partial order.
Irreflexive. Let's assume that $\exists \in \mathcal{N}$ s.t. $p \lhd p$. Then, because $k = k'$ and that it is not possible that $p_k \neq p_k$, then the third condition implies that $\exists j \in [k-1]$ s.t. $q_j \neq q_j$, which is also absurd. Then $\lhd$ is irreflexive.
Transitivity. Let $p, p', p'' \in \mathcal{N}$ s.t. $p \lhd p'$ and $p' \lhd p''$. Then, for all $i \in [k]$, $p_i \subseteq p'_i \subseteq p''_i$ and for all $i \in [k-1]$, $q_i \subseteq q'_i \subseteq q''_i$ ($k \leq k' \leq k''$). Finally, if $k = k''$, then $k = k'' = k'$. Assuming that $p_k = p'_k$ and $p'_k = p''_k$ then $p_k = p''_{k''}$. Assuming that $p_k \neq p'_k$ or $p'_k \neq p''_k$, then $\exists j \in [k-1]$ s.t. $q_j \neq q'_j$ or $q'_j \neq q''_j$, and hence $q_j \neq q''_j$. Then, we have that $p \lhd p''$.
Asymmetry. Let $p, p' \in \mathcal{N}$ s.t. $p \lhd p'$. Then, if $k < k'$ we can not have $p' \lhd p$. Assuming that $k = k'$ and that $p_k = p'_k$ for the same reason as the asymmetry of $\lessdot$, it is not possible to have $q_j \neq q'_j$. If $p_k \neq p'_k$ the, $p_k \subsetneq p'_k$ (according to 1.) and then it is not possible to have $p'_j \subseteq p_j$. As a consequence, we can not have $p' \lhd p$.
We now prove that $\lessdot$ is a strict partial order.
Irreflexive. Let's assume that $\exists \in \mathcal{N}$ s.t. $p \lessdot p$. Then, the unique strictly incremental (auto-)mapping is $u_i = i$, for all $i \in [k]$. Then, the third condition implies that $\exists j \in [k-1]$ s.t. $q_j \neq q_j$ or $p_j \neq p_j$, which is absurd. Then $\lessdot$ is irreflexive.
Transitivity. Let $p, p', p'' \in \mathcal{N}$ s.t. $p \lessdot p'$ and $p' \lessdot p''$. We denotes by $(u_i)_i \in [k']^k$ and $(v_i)_i \in [k'']^{k'}$ the respective mapping, and we define $(w_i)_i \in [k'']^[k]$ such that $w_i = v_{u_i}$ for all $i \in [k]$. Then, for all $i \in [k]$, $p_i \subseteq p'_{u_i} \subseteq p''_{v_{u_i}} = p''_{w_i}$; $q_i \subseteq \bigcup_{j\in[u_i,u_{i+1}-1]} q'_j = \bigcup_{j\in[u_i,u_{i+1}-1]} \bigcup_{l\in[v_j,v_{j+1}-1]} q''_l$. The union of the q''_l in the intervals $[v_j, v_{j+1}-1]$ for $j \in [u_i, u_{i+1}-1]$ can be sum up as an union on the interval $[v_{u_i}, v_{(u_{i+1}-1)+1} - 1] = [v_{u_i}, v_{u_{i+1}} - 1] = [w_i, w_{i+1} - 1]$ because intervals are contiguous. Then, $q_i \subseteq \bigcup_{j\in[w_i,w_{i+1}-1]} q''_j$. Finally, if $k = k''$, then $k = k'' = k'$ and then it exists $j \in [k]$, s.t. $p_j \neq p'_j \subseteq p''_j$ or $q_j \neq q'_j \subseteq q''_j$. Thus, $p_j \neq p''_j$ or $q_j \neq q''_j$. As a consequence, we have $p \lessdot p''$.
Asymmetry. Let $p, p' \in \mathcal{N}$ s.t. $p \lessdot p'$. Then, if $k < k'$ we can not have $p' \lessdot p$. Assuming that $k = k'$ (and thus $u_i = i$ for all $i \in [k]$), we have that it exists $j \in [k]$ s.t. $p_j \neq p'_j$ or $q_j \neq q'_j$. If $p_j \neq p'_j$ then, according to 1. $p_j \subsetneq p'_j$, it is not possible to have $p'_j \subseteq p_j$. If $q_j \neq q'_j$, then, according to 2. $q_j \subsetneq \bigcup_{j\in[u_i,u_{i+1}-1]} q'_j = q'_j$. Thus, it is not possible to have $q'_j \subseteq q_j = \bigcup_{j\in[u_i,u_{i+1}-1]} q_j$. As a consequence, we can not have $p' \lessdot p$. □

Proof of Proposition 3. We start this proof by a small result about the anti-monotonicity of $\not\subseteq$. Let $P, Q \in \mathcal{I}$ be two itemsets s.t. $P \subseteq Q$, and $I \in \mathcal{I}$ another itemset. Then, $Q \not\subseteq I \implies P \not\subseteq I$. In fact, $Q \not\subseteq I$ implies that for all $e \in Q$, $e \notin I$, and because $P \subseteq Q$, we also have that $e \in P$, $e \notin I$.

Let $\boldsymbol{p} = \langle p_1\ \neg q_1\ \ldots\ \neg q_{m-1}\ p_m \rangle \in \mathcal{N}$ and $\boldsymbol{p}' = \langle p'_1\ \neg q'_1\ \ldots\ \neg q'_{m'-1}\ p'_{m'} \rangle \in \mathcal{N}$ be two NSP s.t. $\boldsymbol{p} \lhd \boldsymbol{p}'$.

We first show that an $(\circ, \not\subseteq)$-embedding of $\boldsymbol{p}'$ in a sequence $\boldsymbol{s}$, denoted $\boldsymbol{e} = (e_i)_{i\in[m']}$, induces an $(\circ, \not\subseteq)$-embedding of $\boldsymbol{p}$. By Definition 4, we have $p'_i \subseteq s_{e_i}$, $\forall i \in [m']$ and $q'_i \not\subseteq s_j$, for all $j \in [e_i + 1, e_{i+1} - 1]$ and for all $i \in [m'-1]$.
On the other side, $\boldsymbol{p} \lhd \boldsymbol{p}'$ implies that $p_i \subseteq p'_i$ for all $i \in [m]$. Then, because $m \leq m'$ ($\boldsymbol{p} \lhd \boldsymbol{p}'$), we have that $p_i \subseteq s_{e_i}$ for all $i \in [m]$. In addition, $\boldsymbol{p} \lhd \boldsymbol{p}'$ also implies that $q_i \subseteq q'_i$ for all $i \in [m-1]$ and thus, by anti-monotonicity of $\not\subseteq$ (and $q'_i \not\subseteq s_j$), we have $q_i \not\subseteq s_j$ for all $j \in [e_i + 1, e_{i+1} - 1]$ and for all $i \in [m-1]$. In conclusion, we have that $\boldsymbol{e} = (e_i)_{i\in[m]}$ is an $(\circ, \not\subseteq)$-embedding of $\boldsymbol{p}$.

We now show that an $(\bullet, \not\subseteq)$-embedding of $\boldsymbol{p}'$ in a sequence $\boldsymbol{s}$, denoted $\boldsymbol{e} = (e_i)_{i\in[m']}$, induces an $(\bullet, \not\subseteq)$-embedding of $\boldsymbol{p}$. By Definition 4, we have $p_i \subseteq s_{e_i}$, $\forall i \in [m']$ and $q_i \not\subseteq \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$, for all $i \in [m'-1]$.
On the other side, $\boldsymbol{p} \lhd \boldsymbol{p}'$ implies that $p_i \subseteq p'_i$ for all $i \in [m]$. Then, because $m \leq m'$ ($\boldsymbol{p} \lhd \boldsymbol{p}'$), we have that $p_i \subseteq s_{e_i}$ for all $i \in [m]$. In addition, $\boldsymbol{p} \lhd \boldsymbol{p}'$ also implies that $q_i \subseteq q'_i$ for all $i \in [m-1]$, by anti-monotonicity of $\not\subseteq$, we have

$q_i \not\subseteq \bigcup_{j\in[e_i+1,e_{i+1}-1]} s_j$, for all $i \in [m-1]$. In conclusion, we have that $\boldsymbol{e} = (e_i)_{i\in[m]}$ is an $(\bullet, \not\subseteq)$-embedding of $\boldsymbol{p}$. $\square$

Proof of Proposition 4. Let $\boldsymbol{p} = \langle p_1\ \neg q_1\ \dots\ \neg q_{k-1}\ p_k\rangle \in \mathcal{N}$ and $\boldsymbol{p}' = \langle p'_1\ \neg q'_1\ \dots\ \neg q'_{k'-1}\ p'_{k'}\rangle \in \mathcal{N}$ be two NSP s.t. $\boldsymbol{p} \trianglelefteq \boldsymbol{p}'$. Thus, we have that $k = k'$.

Similarly to the proof of Proposition 3, we can show that any $(\LEFTcircle, \not\subseteq)$-embedding of $\boldsymbol{p}'$ in $\boldsymbol{s}$ induces an $(\LEFTcircle, \not\subseteq)$-embedding of $\boldsymbol{p}$ in $\boldsymbol{s}$. This enables to conclude that $\preceq^{\not\subseteq}_{\LEFTcircle}$ is anti-monotonic on $(\mathcal{N}, \trianglelefteq)$.

The anti-monotonicity of $\sqsubseteq^{\not\subseteq}_{\circ}$ requires that each embedding of $\boldsymbol{p}^+$ in $\boldsymbol{s}$ satisfies the negations. Let us assume that $p' \sqsubseteq^{\not\subseteq}_{\circ} s$, then there exists an embedding $(e_i)_{i\in[k]}$ of $\boldsymbol{p}'$. $(e_i)_{i\in[k]}$ is also an embedding of $\boldsymbol{p}'^+$ (Lemma 5). According to 1. and because $k = k'$, $\boldsymbol{p}'^+ = \boldsymbol{p}+$, and then $(e_i)_{i\in[k]}$ is an embedding of $\boldsymbol{p}$ in $\boldsymbol{s}$. Thus, we shown that there is at least one embedding of $\boldsymbol{p}^+$ in $\boldsymbol{s}$. If $\sqsubseteq^{\not\subseteq}_{\circ}$ is not anti-monotonic, then there exists an embedding $(e_i)_{i\in[k]}$ of $\boldsymbol{p}^+$ such that $\exists j \in [k]$ and $l \in [e_j + 1, e_{j+1} - 1]$, s.t. it is false that $q_j \not\subseteq s_l$ $(\exists\alpha \in q_j,\ \alpha \in s_l)$. According to 2. $q_j \subseteq q'_j$, and thus it is false $q'_j \not\subseteq s_l$. Nonetheless, $(e_i)_{i\in[k]}$ is also an embedding of $\boldsymbol{p}'^+$. And $p' \sqsubseteq^{\not\subseteq}_{\circ} s$, it implies that $q'_j \not\subseteq s_l$. There is a contradiction, thus $\sqsubseteq^{\not\subseteq}_{\circ}$ is anti-monotonic.

The anti-monotonicity of $\sqsubseteq^{\not\subseteq}_{\bullet}$ requires that each embedding of $\boldsymbol{p}^+$ in $\boldsymbol{s}$ satisfies the negations. Let us assume that $p' \sqsubseteq^{\not\subseteq}_{\bullet} s$, then there exists an embedding $(e_i)_{i\in[k]}$ of $\boldsymbol{p}'$. $(e_i)_{i\in[k]}$ is also an embedding of $\boldsymbol{p}'^+$ (Lemma 5). According to 1. and because $k = k'$, $\boldsymbol{p}'^+ = \boldsymbol{p}+$, and then $(e_i)_{i\in[k]}$ is an embedding of $\boldsymbol{p}$ in $\boldsymbol{s}$. Thus, we shown that there is at least one embedding of $\boldsymbol{p}^+$ in $\boldsymbol{s}$. If $\sqsubseteq^{\not\subseteq}_{\bullet}$ is not anti-monotonic, then there exists an embedding $(e_i)_{i\in[k]}$ of $\boldsymbol{p}^+$ such that $\exists j \in [k]$, s.t. it is false that $q_j \not\subseteq \bigcup_{l\in[e_j+1,e_{j+1}-1]} s_l$. According to 2. $q_j \subseteq q'_j$, and thus it is false $q'_j \not\subseteq \bigcup_{l\in[e_j+1,e_{j+1}-1]} s_l$. Nonetheless, $(e_i)_{i\in[k]}$ is also an embedding of $\boldsymbol{p}'^+$. And $p' \sqsubseteq^{\not\subseteq}_{\bullet} s$, it implies that $q'_j \not\subseteq \bigcup_{l\in[e_j+1,e_{j+1}-1]} s_l$. There is a contradiction, thus $\sqsubseteq^{\not\subseteq}_{\bullet}$ is anti-monotonic. $\square$

Proof of Proposition 5. Let $\theta, theta' \in \Theta$, then $\theta \geqslant \theta' \implies supp_{\theta}(\boldsymbol{p}) \leq supp_{\theta'}(\boldsymbol{p})$ for all $\boldsymbol{p} \in \mathcal{N}$ (by Definition 6 of the dominance relation). Thus, Proposition 5 comes immediately with Proposition 1. $\square$

Proof of Proposition 6. Let $\boldsymbol{p}, \boldsymbol{p}' \in \mathcal{N}$ be two negative sequential patterns such that $\boldsymbol{p} \vartriangleleft \boldsymbol{p}'$. According to Proposition 3, $\boldsymbol{p}' \preceq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s} \implies \boldsymbol{p} \preceq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s}$ for all $\boldsymbol{s}$. Thus, $supp_{\not\subseteq,\LEFTcircle,\preceq}(\boldsymbol{p}) \leq supp_{\not\subseteq,\LEFTcircle,\preceq}(\boldsymbol{p})$.

If $\boldsymbol{p} \trianglelefteq \boldsymbol{p}'$. According to Proposition 4, $\boldsymbol{p}' \sqsubseteq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s} \implies \boldsymbol{p} \sqsubseteq^{\not\subseteq}_{\LEFTcircle} \boldsymbol{s}$ for all $\boldsymbol{s}$. Thus, $supp_{\not\subseteq,\LEFTcircle,\sqsubseteq}(\boldsymbol{p}) \leq supp_{\not\subseteq,\LEFTcircle,\sqsubseteq}(\boldsymbol{p})$. $\square$