Membership Constraints in Formal Concept Analysis

Sebastian Rudolph, Christian Săcărea, and Diana Troancă

TU Dresden and Babeş-Bolyai University of Cluj-Napoca
sebastian.rudolph@tu-dresden.de, {csacarea,dianat}@cs.ubbcluj.ro

July 29, 2015

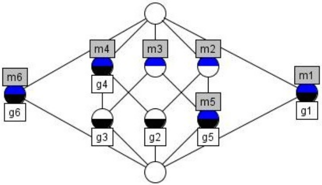
Formal Concept Analysis

Definition

*A **formal context** is a triple $\mathbb{K} = (G, M, I)$ with a set G called **objects**, a set M called **attributes**, and $I \subseteq G \times M$ the binary **incidence relation** where gIm means that object g has attribute m.*

*A **formal concept** of a context $\mathbb{K}$ is a pair (A, B) with **extent** $A \subseteq G$ and **intent** $B \subseteq M$ satisfying $A \times B \subseteq I$ and A, B are maximal w.r.t. this property, i.e., for every $C \supseteq A$ and $D \supseteq B$ with $C \times D \subseteq I$ must hold $C = A$ and $D = B$.*

	m_1	m_2	m_3	m_4	m_5	m_6
g_1	×					
g_2		×		×		
g_3			×	×		
g_4				×		
g_5		×	×		×	
g_6						×



Constraints on Formal Contexts

Definition (inclusion/exclusion constraint)

A ***inclusion/exclusion constraint*** *(MC) on a formal context* $\mathbb{K} = (G, M, I)$ *is a quadruple* $\mathbb{C} = (G^+, G^-, M^+, M^-)$ *with*

- $G^+ \subseteq G$ *called* ***required objects***,
- $G^- \subseteq G$ *called* ***forbidden objects***,
- $M^+ \subseteq M$ *called* ***required attributes***, *and*
- $M^- \subseteq M$ *called* ***forbidden attributes***.

A formal concept (A, B) *of* $\mathbb{K}$ *is said to* ***satisfy*** *a MC if all the following conditions hold:*

$$G^+ \subseteq A, \qquad G^- \cap A = \emptyset, \qquad M^+ \subseteq B, \qquad M^- \cap B = \emptyset.$$

An MC is said to be ***satisfiable*** *with respect to* $\mathbb{K}$*, if it is satisfied by one of its formal concepts.*

Problem (MCSAT)

input: *formal context* $\mathbb{K}$*, membership constraint* $\mathbb{C}$
output: YES *if* $\mathbb{C}$ *satisfiable w.r.t.* $\mathbb{K}$, NO *otherwise.*

Theorem

MCSAT is NP*-complete, even when restricting to membership constraints of the form* $(\emptyset, G^-, \emptyset, M^-)$*.*

Proof.

In NP: guess a pair (A, B) with $A \subseteq G$ and $B \subseteq M$, then check if it is a concept satisfying the membership constraint. The check can be done in polynomial time.
NP-hard: We polynomially reduce the NP-hard 3SAT problem to MCSAT. □

Reduction from 3SAT to MCSAT (by example)

Satisfiability of formula

$$\varphi = (r \vee s \vee \neg q) \wedge (s \vee \neg q \vee \neg r) \wedge (\neg q \vee \neg r \vee \neg s)$$

corresponds to satisfiability of MC

$$(\emptyset, \{(r \vee s \vee \neg q), (s \vee \neg q \vee \neg r), (\neg q \vee \neg r \vee \neg s)\}, \emptyset, \{\tilde{q}, \tilde{r}, \tilde{s}\})$$

in the context

	q	r	s	$\neg q$	$\neg r$	$\neg s$	$\tilde{q}$	$\tilde{r}$	$\tilde{s}$
$(r \vee s \vee \neg q)$	×				×	×	×	×	×
$(s \vee \neg q \vee \neg r)$	×	×				×	×	×	×
$(\neg q \vee \neg r \vee \neg s)$	×	×	×				×	×	×
q		×	×	×	×	×		×	×
r	×		×	×	×	×	×		×
s	×	×		×	×	×	×	×	
$\neg q$	×	×	×		×	×		×	×
$\neg r$	×	×	×	×		×	×		×
$\neg s$	×	×	×	×	×		×	×	

Bijection between valuations making φ true (here: $\{q \mapsto \mathit{true}, r \mapsto \mathit{false}, s \mapsto \mathit{true}\}$)
and concepts satisfying MC (here: $(\{r, \neg q, \neg s\}, \{q, s, \neg r\})$).

Theorem

When restricted to membership constraints of the form $(G^+, \emptyset, M^+, M^-)$ or $(G^+, G^-, M^+, \emptyset)$ MCSAT is in AC_0.

Proof.

$(G^+, \emptyset, M^+, M^-)$ is satisfiable w.r.t. $\mathbb{K}$ if and only if it is satisfied by $(M^{+\prime}, M^{+\prime\prime})$. By definition, this is the case iff

1. $G^+ \subseteq M^{+\prime}$ and
2. $M^{+\prime\prime} \cap M^- = \emptyset$.

These conditions can be expressed by the first-order sentences

1. $\forall x, y.(x{\in}G^+ \wedge y{\in}M^+ \rightarrow xIy)$ and
2. $\forall x.(x{\in}M^- \rightarrow \exists y.(\forall z.(z{\in}M^+ \rightarrow yIz) \wedge \neg yIx))$.

Due to descriptive complexity theory, first-order expressibility of a property ensures that it can be checked in AC_0. □

Triadic FCA

Definition

*A **tricontext** is a quadruple* $\mathbb{K} = (G, M, B, I)$ *with*

- *a set* G *called **objects**,*
- *a set* M *called **attributes**, and*
- *a set* B *called **conditions**, and*
- $Y \subseteq G \times M \times B$ *the ternary **incidence relation** where* $(g, m, b) \in Y$ *means that object* g *has attribute* m *under condition* b*.*

Definition

*A **triconcept** of a tricontext* $\mathbb{K}$ *is a triple* (A_1, A_2, A_3) *with **extent*** $A_1 \subseteq G$*, **intent*** $A_2 \subseteq M$*, and **modus*** $A_3 \subseteq B$ *satisfying* $A_1 \times A_2 \times A_3 \subseteq Y$ *and for every* $C_1 \supseteq A_1$*,* $C_2 \supseteq A_2$*,* $C_3 \supseteq A_3$ *that satisfy* $C_1 \times C_2 \times C_3 \subseteq Y$ *holds* $C_1 = A_1$*,* $C_2 = A_2$*, and* $C_3 = A_3$*.*

Membership constraints in triadic FCA

Definition

A ***triadic inclusion exclusion constraint*** *(3MC) on a tricontext* $\mathbb{K} = (G, M, B, Y)$ *is a sextuple* $\mathbb{C} = (G^+, G^-, M^+, M^-, B^+, B^-)$ *with*

- $G^+ \subseteq G$ *called* ***required objects***, $G^- \subseteq G$ *called* ***forbidden objects***,
- $M^+ \subseteq M$ *called* ***required attributes***, $M^- \subseteq M$ *called* ***forbidden attributes***,
- $B^+ \subseteq B$ *called* ***required conditions***, *and* $B^- \subseteq B$ *called* ***forbidden conditions***.

A triconcept (A_1, A_2, A_3) *of* $\mathbb{K}$ *is said to* ***satisfy*** *such a 3MC if all the following conditions hold:* $G^+ \subseteq A_1$, $G^- \cap A_1 = \emptyset$, $M^+ \subseteq A_2$, $M^- \cap A_2 = \emptyset$, $B^+ \subseteq A_3$, $B^- \cap A_3 = \emptyset$.
A 3MC constraint is said to be satisfiable with respect to $\mathbb{K}$, *if it is satisfied by one of its triconcepts.*

Problem (3MCSAT)

input: *formal context* $\mathbb{K}$*, triadic inclusion/exclusion constraint* $\mathbb{C}$

output: YES *if* $\mathbb{C}$ *satisfiable w.r.t.* $\mathbb{K}$, NO *otherwise.*

Theorem

3MCSAT is NP*-complete, even when restricting to 3MCs of the following forms:*

- $(\emptyset, G^-, \emptyset, M^-, \emptyset, \emptyset)$, $(\emptyset, G^-, \emptyset, \emptyset, \emptyset, B^-)$, $(\emptyset, \emptyset, \emptyset, M^-, \emptyset, B^-)$,
- $(G^+, G^-, \emptyset, \emptyset, \emptyset, \emptyset)$, $(\emptyset, \emptyset, M^+, M^-, \emptyset, \emptyset)$, $(\emptyset, \emptyset, \emptyset, \emptyset, B^+, B^-)$.

Proof.

In NP: guess a triple (A_1, A_2, A_3) with $A_1 \subseteq G$ and $A_2 \subseteq M$ and $A_3 \subseteq M$, then check if it is a triconcept satisfying the 3MC. The check can be done in polynomial time.
NP-hard: for the first type, use the same reduction as in the previous proof. For the second type, we polynomially reduce the NP-hard 3SAT problem to 3MCSAT in another way. □

Reduction from 3SAT to 3MCSAT (by example)

Satisfiability of formula

$$\varphi = (r \vee s \vee \neg q) \wedge (s \vee \neg q \vee \neg r) \wedge (\neg q \vee \neg r \vee \neg s)$$

corresponds to satisfiability of 3MC

$$(\{*\}, \{(r \vee s \vee \neg q), (s \vee \neg q \vee \neg r), (\neg q \vee \neg r \vee \neg s)\}, \emptyset, \emptyset, \emptyset, \emptyset)$$

in the tricontext

$*$	$*$	q	r	s
$*$	×	×	×	×
$\neg q$	×		×	×
$\neg r$	×	×		×
$\neg s$	×	×	×	

$(r\vee s\vee\neg q)$	$*$	q	r	s
$*$	×	×		
$\neg q$		×	×	×
$\neg r$	×	×	×	×
$\neg s$	×	×	×	×

$(s\vee\neg q\vee\neg r)$	$*$	q	r	s
$*$	×	×	×	
$\neg q$		×	×	×
$\neg r$		×	×	×
$\neg s$	×	×	×	×

$(\neg q\vee\neg r\vee\neg s)$	$*$	q	r	s
$*$	×	×	×	×
$\neg q$		×	×	×
$\neg r$		×	×	×
$\neg s$		×	×	×

Bijection between valuations making φ true (here: $\{q \mapsto \mathit{true}, r \mapsto \mathit{false}, s \mapsto \mathit{true}\}$)
and triconcepts satisfying 3MC (here: $(\{*\}, \{*, q, s\}, \{*, \neg r\})$).

Theorem

3MCSAT is in AC_0 *when restricting to MCs of the forms* $(\emptyset, G^-, M^+, \emptyset, B^+, \emptyset)$, $(G^+, \emptyset, \emptyset, M^-, B^+, \emptyset)$, *and* $(G^+, \emptyset, M^+, \emptyset, \emptyset, B^-)$.

Proof.

$\mathbb{C} = (\emptyset, G^-, M^+, \emptyset, B^+, \emptyset)$ is satisfiable w.r.t. $\mathbb{K}$ if and only if the triconcept (G_U, M, B) satisfies it (where $G_U = \{g \mid \{g\} \times M \times B \subseteq Y\}$), that is, if $G_U \cap G^- = \emptyset$. This can be expressed by the first-order formula

$$\forall x.x \in G^- \to \exists y, z.(y \in M \wedge z \in B \wedge \neg(x, y, z) \in Y).$$

Therefore, checking satisfiability of this type of 3MCs is in AC_0. The other cases follow by symmetry. □

n-adic FCA

Definition

*An n-**context** is an $(n+1)$-tuple $\mathbb{K} = (K_1, \ldots, K_n, R)$ with $K_1, \ldots, K_n$ being sets, and $R \subseteq K_1 \times \ldots \times K_n$ the n-ary **incidence relation**.*
*An n-**concept** of an n-context $\mathbb{K}$ is an n-tuple $(A_1, \ldots, A_n)$ satisfying $A_1 \times \ldots \times A_n \subseteq R$ and for every n-tuple $(C_1, \ldots, C_n)$ with $A_i \supseteq C_i$ for all $i \in \{1, \ldots, n\}$, satisfying $C_1 \times \ldots \times C_n \subseteq R$ holds $C_i = A_i$ for all $i \in \{1, \ldots, n\}$.*

Definition

*A n-**adic inclusion/exclusion constraint** (nMC) on a n-context $\mathbb{K} = (K_1, \ldots, K_n, R)$ is a $2n$-tuple $\mathbb{C} = (K_1^+, K_1^-, \ldots, K_n^+, K_n^-)$ with $K_i^+ \subseteq K_i$ called **required sets** and $K_i^- \subseteq K_i$ called **forbidden sets**.*
*An n-concept $(A_1, \ldots, A_n)$ of $\mathbb{K}$ is said to **satisfy** such a membership constraint if $K_i^+ \subseteq A_i$ and $K_i^- \cap A_i = \emptyset$ hold for all $i \in \{1, \ldots, n\}$.*
An n-adic membership constraint is said to be satisfiable with respect to $\mathbb{K}$, if it is satisfied by one of its n-concepts.

Theorem

For a fixed $n > 2$, the nMCSAT problem is

- NP-*complete for any class of constraints that allows for*
  - *the arbitrary choice of at least two forbidden sets or*
  - *the arbitrary choice of at least one forbidden set and the corresponding required set,*
- *in* AC_0 *for the class of constraints with at most one forbidden set and the corresponding required set empty,*
- *trivially true for the class of constraints with all forbidden sets and at least one required set empty.*

Econding in answer set programming

Given an n-context $\mathbb{K} = (K_1, \ldots, K_n, R)$ and nMC $\mathbb{C} = (K_1^+, K_1^-, \ldots, K_n^+, K_n^-)$, let the corresponding problem be given by the following set of ground facts $F_{\mathbb{K},\mathbb{C}}$:

- $\mathtt{set}_i(a)$ for all $a \in K_i$,
- $\mathtt{rel}(a_1, \ldots, a_n)$ for all $(a_1, \ldots, a_n) \in R$,
- $\mathtt{required}_i(a)$ for all $a \in K_i^+$, and
- $\mathtt{forbidden}_i(a)$ for all $a \in K_i^-$.

Let P denote the following fixed answer set program (with rules for every $i \in \{1, \ldots, n\}$):

Program

$$\begin{aligned}
\mathtt{in}_i(x) &\leftarrow \mathtt{set}_i(x) \wedge {\sim}\mathtt{out}_i(x)\\
\mathtt{out}_i(x) &\leftarrow \mathtt{set}_i(x) \wedge {\sim}\mathtt{in}_i(x)\\
&\leftarrow \textstyle\bigwedge_{j\in\{1,\ldots,n\}} \mathtt{in}_j(x_j) \wedge {\sim}\mathtt{rel}(x_1, \ldots, x_n)\\
\mathtt{exc}_i(x_i) &\leftarrow \textstyle\bigwedge_{j\in\{1,\ldots,n\}\setminus\{i\}} \mathtt{in}_j(x_j) \wedge {\sim}\mathtt{rel}(x_1, \ldots, x_n)\\
&\leftarrow \mathtt{out}_i(x) \wedge {\sim}\mathtt{exc}_i(x)\\
&\leftarrow \mathtt{out}_i(x) \wedge \mathtt{required}_i(x)\\
&\leftarrow \mathtt{in}_i(x) \wedge \mathtt{forbidden}_i(x)
\end{aligned}$$

Then the answer sets of P correspond to the n-concepts of $\mathbb{K}$ satisfying $\mathbb{C}$.

Applications

- "concept retrieval"
- guided navigation by interactively narrowing down the search space (“faceted browsing”)
- context debugging

Thank You!