# Attribute reduction from congruences

Roberto G. Aragón Jesús Medina Eloísa Ramírez-Poussa

Department of Mathematics. University of Cádiz, Spain

Universidad de Cádiz
Departamento de Matemáticas

III HARMONIC 2018
November 1st.

# Outline

Introduction

Preliminary notions

Reducing a context in FCA based on RST

Introducing congruences

Conclusions and future work

# Outline

Introduction

Preliminary notions

Reducing a context in FCA based on RST

Introducing congruences

Conclusions and future work

# FCA and RST

- Two fundamental mathematical tools for modelling and processing incomplete information in databases are Rough Set Theory (RST) and Formal Concept Analysis (FCA).
- Both theories extract information from databases, which contain a set of objects, a set of attributes and a relationship between them.
- In spite of considering different philosophies, rough set theory and formal concept analysis are closely related.

# Attribute reductions

- One of the principal targets in both theories is to reduce the number of attributes, preserving the information that can be obtained from the database.
- To this end, reducts (minimal set of attributes preserving the main information) have been studied in a number of papers, in these two frameworks.
- These theories have been related in different papers but few of them have studied the connections between the attribute reduction mechanisms given in both frameworks.

# Contributions

- In this work, we will present a new mechanism to reduce formal contexts in FCA, based on the philosophy of attribute reduction in RST.
- We reduce a context in FCA considering the reducts of the associated context information system.
- We will show that this kind of reduction satisfies interesting properties.
- We will illustrate them by means of an example.
- Hence, this work introduces a new and different way of reducing a formal context in FCA.

# Outline

Introduction

Preliminary notions

Reducing a context in FCA based on RST

Introducing congruences

Conclusions and future work

# Rough set theory

## Definition

$(U, \mathcal{A})$, where $U$ and $\mathcal{A}$ are finite, non-empty sets of objects and attributes, respectively. Each $a$ in $\mathcal{A}$ corresponds to a mapping $\bar{a}\colon U \to V_a$, where $V_a$ is the value set of $a$ over $U$.

## Example

| | Temperature | Headache |
|---|---|---|
| $x_1$ | Hight | Yes |
| $x_2$ | Normal | Yes |
| $x_3$ | Hight | Yes |
| $x_4$ | Low | No |

# $D$-indiscernibility relation

## $D$-indiscernibility relation

For every subset $D$ of $\mathcal{A}$, the *$D$-indiscernibility relation*, $\text{Ind}(D)$, is defined as the equivalence relation

$$\text{Ind}(D) = \{(x_i, x_j) \in U \times U \mid \text{ for all } a \in D, \bar{a}(x_i) = \bar{a}(x_j)\}$$

where each class given by this relation can be written as $[x]_D = \{x_i \mid (x, x_i) \in \text{Ind}(D)\}$. $\text{Ind}(D)$ produces a partition on $U$.

# Consistent set and reduct in RST

The following notions will be essential in the relationship between RST and FCA considered in this work.

**Definition**

Let $(U, \mathcal{A})$ be an information system and a subset of attributes $D \subseteq \mathcal{A}$. $D$ is a *consistent set* of $(U, \mathcal{A})$ if

$$\text{Ind}(D) = \text{Ind}(\mathcal{A})$$

Moreover, if for each $a \in D$ we have that $\text{Ind}(D \setminus \{a\}) \neq \text{Ind}(\mathcal{A})$, then $D$ is called *reduct* of $(U, \mathcal{A})$.

# Discernibility matrix and function

The discernibility matrix of $(U, \mathcal{A})$ is the $n \times n$ matrix $O$, defined by, for $i$ and $j$ in $\{1, ..., n\}$,

## Definition

Given an information system $(U, \mathcal{A})$, its *discernibility matrix* is a matrix with order $|U| \times |U|$, denoted as $M_{\mathcal{A}}$, in which the element $M_{\mathcal{A}}(i,j)$ for each pair of objects $(i,j)$ is defined by:

$$M_{\mathcal{A}}(i,j) = \{a \in \mathcal{A} \mid \bar{a}(i) \neq \bar{a}(j)\}$$

and the *discernibility function* of $(U, \mathcal{A})$ is defined by:

$$\tau_{\mathcal{A}} = \bigwedge \left\{ \bigvee (M_{\mathcal{A}}(i,j)) \mid i,j \in U \text{ and } M_{\mathcal{A}}(i,j) \neq \varnothing \right\}$$

# Generating all the reducts

The following result relates the discernibility function to the reducts of an information system.

**Theorem**

Given a boolean information system $(U, \mathcal{A})$. An arbitrary set $D$, where $D \subseteq \mathcal{A}$, is a reduct of the information system if and only if the cube $\bigwedge_{a \in D} a$ is a cube in the restricted disjunctive normal form(RDNF) of $\tau_{\mathcal{A}}$.

# Formal concept analysis

We consider a set of attributes $A$, a set of objects $B$, both of them non empty, and a crisp relationship between them $R\colon A \times B \to \{0,1\}$. A *context* is the triple $(A, B, R)$ and we can define the mappings:

$$X^{\uparrow} = \{a \in A \mid \text{for all } b \in X, aRb\} \quad (1)$$

$$Y^{\downarrow} = \{b \in B \mid \text{for all } a \in Y, aRb\} \quad (2)$$

## Definition

A *concept* in the context $(A, B, R)$ is a pair $(X, Y)$, where $X \subseteq B$, $Y \subseteq A$, $X^{\uparrow} = Y$ and $Y^{\downarrow} = X$ hold. The subset of objects $X$ is called *extent* and $Y$ the *intent*.

The set of all the concepts is denoted as $\mathcal{B}(A, B, R)$, which has a complete lattice structure, when we consider the inclusion ordering on the left argument.

# Formal concept analysis

Note that the operators defined in previous equations form a Galois connection. Taking into account this fact:

- Given an attribute $a \in A$, *the concept generated by* $a$, that is $(a^{\downarrow}, a^{\downarrow\uparrow})$, will be called *attribute-concept*.
- Given an object $b \in B$, the concept generated by $b$, that is $(b^{\uparrow\downarrow}, b^{\uparrow})$, will be called *object-concept*.

# Consistent set and reduct in FCA

**Definition**

- Let $(A, B, R)$ be a context, if there exists a set of attributes $Y \subseteq A$ such that $\mathcal{B}(A, B, R) \cong \mathcal{B}(Y, B, R_{|Y})$, then $Y$ is called a *consistent* set of $(A, B, R)$.
- If $\mathcal{B}(Y \setminus \{y\}, B, R_{|Y \setminus \{y\}}) \ncong \mathcal{B}(A, B, R)$, for all $y \in Y$, then $Y$ is called *reduct* of $(A, B, R)$.

# Notational conventions

- A reduct of the information system $(U, \mathcal{A})$ will be called *RS-reduct* and a reduct of the context $(A, B, R)$ as *CL-reduct*.
- A consistent set of the information system $(U, \mathcal{A})$ will be written in short as *RS-consistent set* and a consistent set of the context $(A, B, R)$ as *CL-consistent set*.

In addition, from now on, as it is usual in real-life knowledge systems, the sets of attributes and the set of objects will be considered finite.

# Irreducible elements of a lattice

Finally, we will recall the notions of meet-irreducible and join-irreducible elements of a lattice.

**Definition**

Given a lattice $(L, \preceq)$, such that $\wedge, \vee$ are the meet and the join operators, and an element $x \in L$ verifying

1. If $L$ has a top element $\top$, then $x \neq \top$.
2. If $x = y \wedge z$, then $x = y$ or $x = z$, for all $y, z \in L$.

we call $x$ *meet-irreducible ($\wedge$-irreducible) element* of $L$. Condition (2) is equivalent to

2′. If $x < y$ and $x < z$, then $x < y \wedge z$, for all $y, z \in L$.

A *join-irreducible ($\vee$-irreducible) element* of $L$ is defined dually.

# Outline


Introduction

Preliminary notions

Reducing a context in FCA based on RST

Introducing congruences

Conclusions and future work

# Context information system

**Definition**

Let $(A, B, R)$ be a context, a *context information system* is defined as the pair $(B, A)$ where the mappings $\bar{a} : B \to V_a$, with $V_a = \{0, 1\}$, are defined as $\bar{a}(b) = R(a, b)$, for all $a \in A, b \in B$.

**Lemma**

Given a context $(A, B, R)$ and the corresponding context information system $(B, A)$, the following equality holds, for each $a \in A$:

$$a^{\downarrow} = \bar{a}$$

# Relating CL-consistent sets to RS-consistent sets

The following result was given in **[Wei and Qi. 2010]** shows that, in some sense, the attribute reduction in FCA implies an attribute reduction in RST.

**Theorem**

Given a context $(A, B, R)$ and the corresponding context information system $(B, A)$. If $D \subseteq A$ is a CL-consistent set then $D$ is an RS-consistent set.

The counterpart of this Theorem is not true, even though considering a small information system.

# Attribute reduction in FCA based on RST

The proposed attribute reduction mechanism is carried out in the following way:

- Given a context $(A, B, R)$, we consider the corresponding context information system.
- We compute the RS-reducts of this information system.
- We reduce the original context according to the obtained RS-reducts.
- We analyze the properties satisfied by such reduction.

# Preserving the number of object-concepts

The first one proves that different object-concepts in the original context provides different object-concepts in the reduced contexts.

**Proposition**

Let $(A, B, R)$ be a context and $(B, A)$ the corresponding context information system. Considering $D \subseteq A$ a RS-consistent set of $(B, A)$ and the objects $k, j \in B$, if $k^{\uparrow} \neq j^{\uparrow}$, then $k^{\uparrow_D} \neq j^{\uparrow_D}$.

This result implies that the reduction given by an RS-consistent set preserves the number of object-concepts.

# Preserves the inequality between object-contexts

Now, we show that the reduction given by a RS-consistent set also preserves the (strict) inequality between object-concepts.

## Proposition

Given a context $(A, B, R)$ and its corresponding context information system $(B, A)$. If $D \subseteq A$ is a RS-consistent set of $(B, A)$ and we consider two objects $k, j \in B$ satisfying that $k^{\uparrow} < j^{\uparrow}$, then the inequality $k^{\uparrow_D} < j^{\uparrow_D}$ holds.

# No new join-irreducible elements in the lattice

The following theorem proves that the join-irreducible elements in the reduced concept lattice by an RS-consistent set are also join-irreducible elements of the original concept lattice.

## Proposition

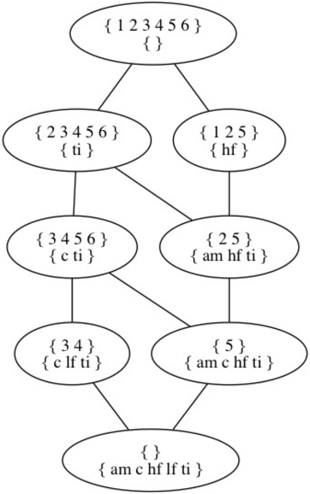
Given a context $(A, B, R)$, the corresponding context information system $(B, A)$ and $D \subseteq A$ a RS-consistent set. If an object $j \in B$ generates a join-irreducible concept in the concept lattice associated with the context $(D, B, R)$, then it also generates a join-irreducible concept of the concept lattice associated with $(A, B, R)$.

# Example

Let us consider the formal context $(A, B, R)$, where $B$ represents a group of six patients and $A$ is the set of symptoms (attributes).

| $R$ | l.fever(lf) | h.fever(hf) | cough(c) | tonsil infla.(ti) | a.muscle(am) |
|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 1 | 1 |
| 3 | 1 | 0 | 1 | 1 | 0 |
| 4 | 1 | 0 | 1 | 1 | 0 |
| 5 | 0 | 1 | 1 | 1 | 1 |
| 6 | 0 | 0 | 1 | 1 | 0 |

# Example


{ 1 2 3 4 5 6 }
{ }
{ 2 3 4 5 6 }
{ ti }
{ 1 2 5 }
{ hf }
{ 3 4 5 6 }
{ c ti }
{ 2 5 }
{ am hf ti }
{ 3 4 }
{ c lf ti }
{ 5 }
{ am c hf ti }
{ }
{ am c hf lf ti }

# Example

Now, we will reduce the context taking into account RS-reducts.
In this case, we obtain the following discernibility matrix:

$$\begin{pmatrix} \varnothing & & & & & \\ \{\mathrm{ti}, \mathrm{am}\} & \varnothing & & & & \\ \{\mathrm{lf}, \mathrm{hf}, \mathrm{c}, \mathrm{ti}\} & \{\mathrm{lf}, \mathrm{hf}, \mathrm{c}, \mathrm{am}\} & \varnothing & & & \\ \{\mathrm{lf}, \mathrm{hf}, \mathrm{c}, \mathrm{ti}\} & \{\mathrm{lf}, \mathrm{hf}, \mathrm{c}, \mathrm{am}\} & \varnothing & \varnothing & & \\ \{\mathrm{c}, \mathrm{ti}, \mathrm{am}\} & \{\mathrm{c}\} & \{\mathrm{lf}, \mathrm{hf}, \mathrm{am}\} & \{\mathrm{lf}, \mathrm{hf}, \mathrm{am}\} & \varnothing & \\ \{\mathrm{hf}, \mathrm{c}, \mathrm{ti}\} & \{\mathrm{hf}, \mathrm{c}, \mathrm{am}\} & \{\mathrm{lf}\} & \{\mathrm{lf}\} & \{\mathrm{hf}, \mathrm{am}\} & \varnothing \end{pmatrix}$$

From this matrix, we obtain the discernibility function,

$$\tau_A = \{\mathrm{lf} \wedge \mathrm{c} \wedge \mathrm{am}\} \vee \{\mathrm{lf} \wedge \mathrm{hf} \wedge \mathrm{c} \wedge \mathrm{ti}\}$$

# Example

Consequently, we have two RS-reducts:

$$\begin{aligned} D_1 &= \{\text{low fever},\ \text{cough}, \text{ache muscle}\} \\ D_2 &= \{\text{low fever}, \text{high fever}, \text{cough}, \text{tonsil inflam.}\} \end{aligned}$$

- We will use these RS-reducts to reduce the original context.
- Once we have the reduced contexts, we will build the concept lattices associated with these two RS-reducts.
- We will see that the structure of the original concept lattice is not necessarily preserved when we consider RS-reducts.

# Example

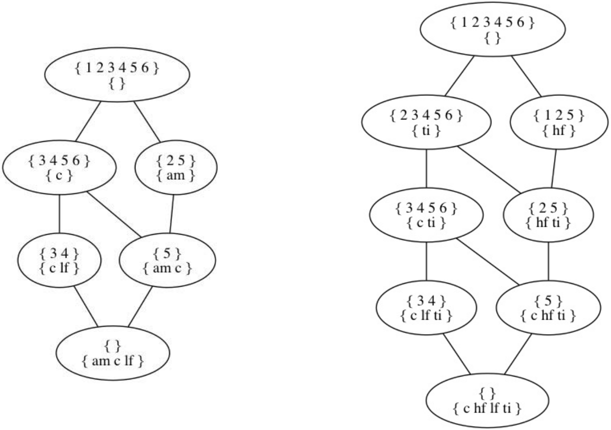



**Figure:** Concept lattices built from the RS-reducts $D_1$ (left) and $D_2$ (right).

# Example

- Considering $D_2$ we do not alter the original structure of the concept lattice since this RS-reduct is also a CL-reduct. Hence, the previous results trivially hold.
- When we consider the concept lattice obtained from the RS-reduct $D_1$:
  - We also reduce the size of the concept lattice.
  - The objects $2, 3, 4, 5$ generate join-irreducible concepts of the concept lattice $\mathcal{B}(D_1, B, R_{|D_1})$ and they also generate join-irreducible concepts of $\mathcal{B}(A, B, R)$.
  - No new join-irreducible element is created after the reduction.

# Example

The inequalities among object-concepts are preserved:

- For example, we have that $2^{\uparrow} < 5^{\uparrow}$ in the original context, and the inequality $2^{\uparrow_1} < 5^{\uparrow_1}$ holds after the reduction.

This is interesting because it shows that two objects that were differentiated, continue being different after the reduction.

- Thus, the new mechanism satisfies useful properties and preserves the necessary information to distinguish the objects.
- Specifically, we have removed attributes ensuring that patients with different symptoms will continue being different.

# Open question

**What are the properties a good reduction mechanism should have?**

- Subset??
- Sublattice??
- Fuzzy transformation: hedges, thresholds, etc.??
- Preserving indiscernibility objects??
- etc.

# Outline

Introduction

Preliminary notions

Reducing a context in FCA based on RST

**Introducing congruences**

Conclusions and future work

# Congruences

- We write $a \equiv b \pmod{\theta}$ or $(a, b) \in \theta$ to indicate that $a$ and $b$ are related under the equivalence relation $\theta$.
- An equivalence relation $\theta$ on a set $A$ gives rise to a partition of $A$ into non-empty disjoint subset. These subsets are the blocks of $\theta$, which are of the form $[a]_\theta = \{x \in A \mid x \equiv a \pmod{\theta}\}$.
- We say that an equivalence relation $\theta$ on a lattice $L$ is *compatible with join and meet* if, for all $a, b, c, d \in L$,

$$a \equiv b \pmod{\theta} \quad \text{and} \quad c \equiv d \pmod{\theta}$$

imply

$$a \vee c \equiv b \vee d \pmod{\theta} \quad \text{and} \quad a \wedge c \equiv b \wedge d \pmod{\theta}.$$

# Congruence

## Congruence

An equivalence relation on a lattice $L$, which is compatible with both join and meet is called a *congruence* on $L$.

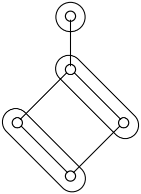


**Figure:** A congruence indicated by placing loops on a lattice.

# Quotient lattices

## Quotient lattices

Given an equivalence relation $\theta$ on a lattice $L$ there is a natural way to try to define operations $\vee$ and $\wedge$ on the set of blocks

$$L/\theta = \{[a]_\theta | a \in L\}.$$

Namely, for all $a, b \in L$, we define

$$[a]_\theta \vee [b]_\theta := [a \vee b]_\theta \text{ y } [a]_\theta \wedge [b]_\theta := [a \wedge b]_\theta.$$

$\vee$ and $\wedge$ are well defined on $L/\theta$ if and only if $\theta$ is a congruence. When $\theta$ is a congruence on $L$, we call $\langle L/\theta, \vee, \wedge\rangle$ *the quotient lattice of L modulo* $\theta$.

# Example of quotient lattices

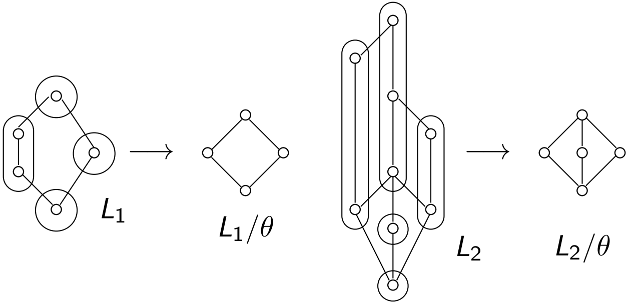



**Figure:** Some examples of congruences and the resulting quotient lattice.

# Congruences

The following lemmas are handy when calculating with congruences.

**Lemma 1**

**i.** An equivalence relation $\theta$ on a lattice $L$ is a congruence if and only if, for all $a, b, c \in L$,

$$a \equiv b \pmod{\theta} \quad \Rightarrow \quad \begin{cases} a \vee c \equiv b \vee c \pmod{\theta} \text{ and} \\ a \wedge c \equiv b \wedge c \pmod{\theta} \end{cases}$$

**ii.** Let $\theta$ be a congruence on $L$ and let $a, b, c \in L$.

**a.** If $a \equiv b \pmod{\theta}$ and $a \leq c \leq b$, then $a \equiv c \pmod{\theta}$.

**b.** $a \equiv b \pmod{\theta}$ if and only if $a \wedge b \equiv a \vee b \pmod{\theta}$.

# Structure properties

**Lemma 2**

Let $\theta$ be a congruence on a lattice $L$ and let $X$ and $Y$ be blocks of $\theta$.

**i.** $X \leq Y$ in $L/\theta$ if and only if there exist $a \in X$ and $b \in Y$ such that $a \leq b$.

**ii.** $X \prec Y$ in $L/\theta$ if and only if $X < Y$ in $L/\theta$ and $a \leq c \leq b$ implies $c \in X$ or $c \in Y$, for all $a \in X$, all $b \in Y$ and all $c \in L$.

**iii.** If $a \in X$ and $b \in Y$, then $a \vee b \in X \vee Y$ and $a \wedge b \in X \wedge Y$.

# Blocks of a congruence

## Block properties

- The blocks of a congruence are certainly sublattices and are convex.
- Furthermore, let $L$ be a lattice and suppose that $\{a, b, c, d\}$ is a 4-element subset of $L$. Then $a, b$ and $c, d$ are said to be *opposite sides of the quadrilateral* $\langle a, b; c, d\rangle$ if $a < b,\ c < d$ and either

  $$(a \vee d = b \text{ and } a \wedge d = c) \text{ or } (b \vee c = d \text{ and } b \wedge c = a)$$

  Then, we say that the blocks of a partition of $L$ are *quadrilateral-closed* if whenever $a, b$ and $c, d$ are opposite sides of a quadrilateral and $a, b \in A$ for some block $A$ then $c, d \in B$ for some block $B$.

# Example


**Figure:** Opposite sides of a quadrilateral.

# Congruences

## Theorem

Let $L$ be a lattice and let $\theta$ be an equivalence relation on $L$. Then $\theta$ is a congruence if and only if

i. each block of $\theta$ is a sublattice of $L$,
ii. each block of $\theta$ is convex,
iii. the blocks of $\theta$ are quadrilateral-closed.

## The lattice of congruences of a lattice

- We could define congruence to be those subsets of $L^2$ which are both equivalence relations and sublattices of $L^2$.
- The set of congruences on a lattice $L$, denoted by Con $L$, is easily seen to be a topped $\bigcap$-structure on $L^2$. Hence Con $L$, when ordered by inclusion, is a complete lattice.
- The least element, **0**, and the greatest element, **1**, are given by $\mathbf{0} = \{(a, a) | a \in L\}$ y $1 = L^2$.

# Congruences

## Principal congruence

The smallest congruence collapsing a given pair of elements $a$ and $b$ is denoted by $\theta(a, b)$; it is called the *principal congruence generated by* $(a, b)$.

$$\theta(a, b) = \bigwedge\{\theta \in \text{Con } L \mid (a, b) \in \theta\}.$$

The next lemma indicates why principal congruences are important:

## Lemma 3

Let $L$ be a lattice and let $\theta \in \text{Con } L$. Then

$$\theta = \bigvee\{\theta(a, b) \mid (a, b) \in \theta\}.$$

Consequently the set of principal congruences is join-dense in Con $L$.

# Congruences

## The smallest congruence

Let $L$ be a lattice and let $H \subseteq L^2$. We denote by $\Theta(H)$ the *smallest congruence relation* such that $a \equiv b \pmod{\Theta}$ for all $(a, b) \in H$.

We have two important results:

## Lemma 4

For any $H \subseteq L^2$, $\Theta(H)$ exists.

## Lemma 5

$\Theta(H) = \bigvee\{\Theta(a, b) | (a, b) \in H\}$.

One of the most important results is the following one:

## Theorem

The lattice Con $L$ is distributive for any lattice $L$.

# Example

Let us consider a context whose associated concept lattice is given in the Figure below.

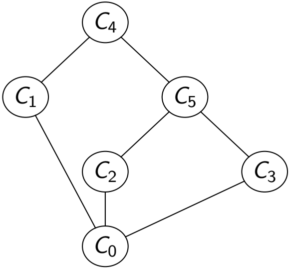



**Figure:** Concept lattice of the context.

# Example

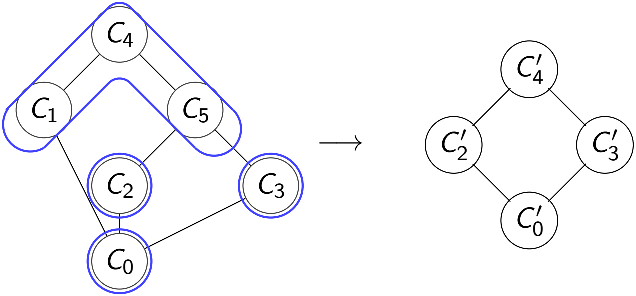



**Figure:** Equivalence relation (blue) obtained from RS-reduct and its concept lattice.

# Example

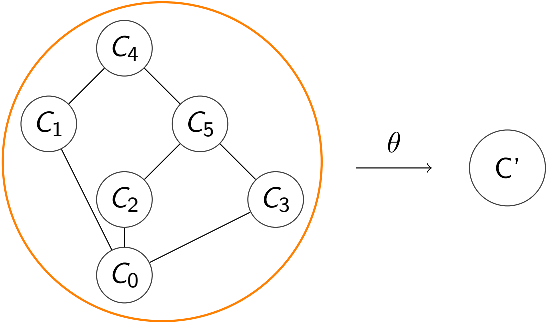



**Figure:** The smallest congruence (orange) containing the equivalence relation obtained from RS-reduct.

# Example

Let us consider a context whose associated concept lattice is given in the Figure below.

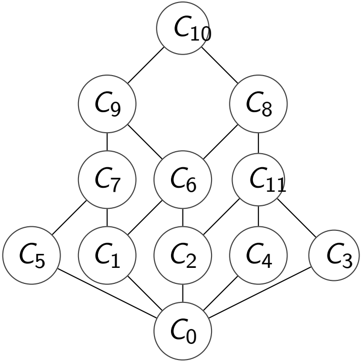



**Figure:** Concept lattice of the context.

# Example

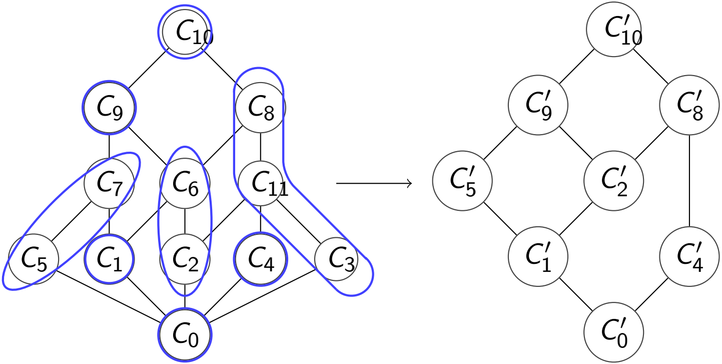



**Figure:** Equivalence relation (blue) obtained from RS-reduct and its concept lattice.

# Example

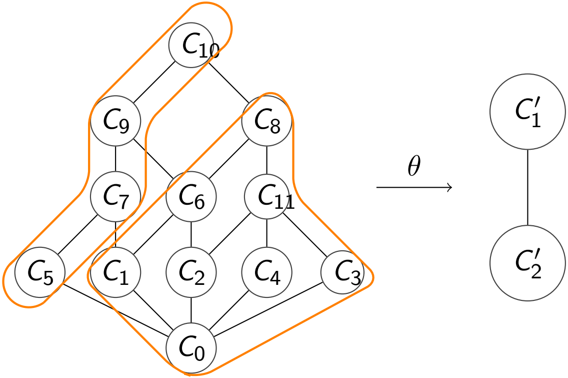



**Figure:** The smallest congruence (orange) containing the equivalence relation obtained from RS-reduct.

# Outline

Introduction

Preliminary notions

Reducing a context in FCA based on RST

Introducing congruences

Conclusions and future work

# Conclusions and future work

- We have shown in this work that the attribute selection procedure given in RST is not equivalent to the attribute reduction in FCA.
- We have proven that the attribute selection mechanism given in RST has different interesting properties when it is applied in the FCA framework.
- These interesting properties provides the possibility of applying the philosophy of RST in order to obtain a reduction in the number of attributes of a context in FCA.
- In the future, we will apply the philosophy of bireducts within the FCA framework and we will also analyze the possible interpretation of this kind of reduction.

# Thank you for your attention