CP 2019 25 YEARS

Stamford, CT, U.S. | September 30 to October 4, 2019

The 25th International Conference on Principles and Practice of Constraint Programming

Modeling Pattern Set Mining using Boolean Circuits

John Aoga, Siegfried Nijssen, Pierre Schaus

UCLouvain icteam Artificial Intelligence & Algorithms fnrs LA LIBERTÉ DE CHERCHER

CP 2019 25 YEARS

Stamford, CT, U.S. | September 30 to October 4, 2019

The 25th International Conference on Principles and Practice of Constraint Programming

Modeling Pattern Set Mining using Boolean Circuits

John Aoga, Siegfried Nijssen, Pierre Schaus

UCLouvain icteam Artificial Intelligence & Algorithms fnrs LA LIBERTÉ DE CHERCHER

Our Motivation

Why we are doing this research ?

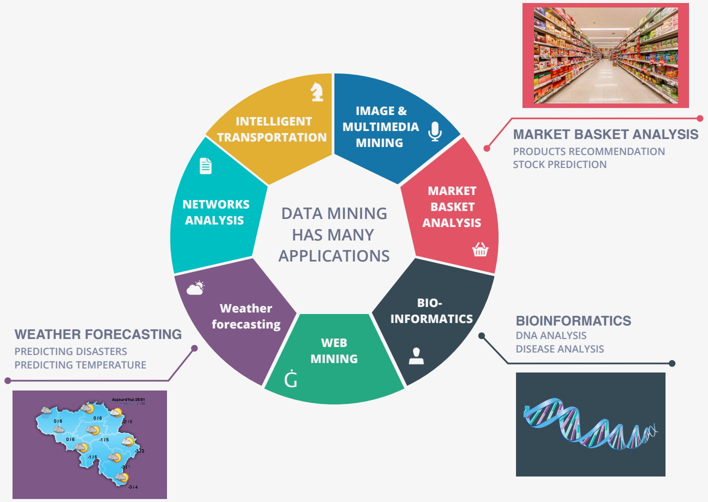
GENERAL OVERVIEW OF DATA MINING

Many applications

GENERAL OVERVIEW OF DATA MINING

Many applications

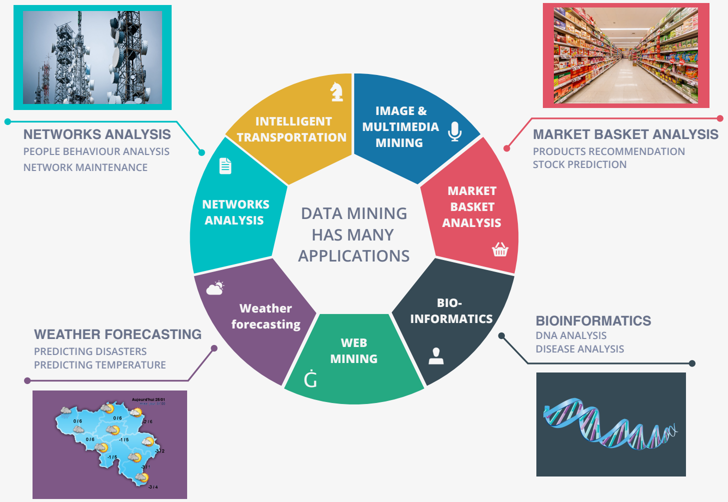
Aoga et al. - CP'19
Oct. 2nd 2019



GENERAL OVERVIEW OF DATA MINING

Many applications

Aoga et al. - CP'19
Oct. 2nd 2019



PATTERN-BASED LEARNING

Pattern mining vs Pattern set Mining

Aoga et al. - CP'19
Oct. 2nd 2019

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

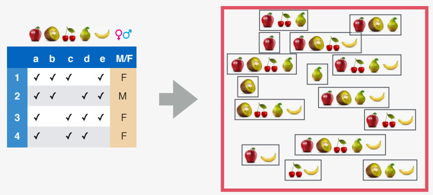
PATTERN-BASED LEARNING

Pattern mining vs Pattern set Mining

Aoga et al. - CP'19
Oct. 2nd 2019

Pattern Mining

- Find (**all**) patterns in data
  - Find novel information.
  - Better understanding data.

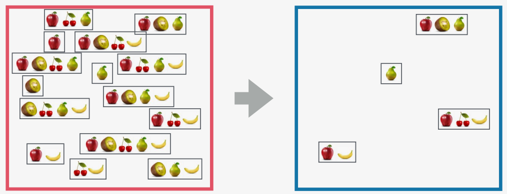


PATTERN-BASED LEARNING
Pattern mining vs Pattern set Mining

Pattern Mining
- Find (**all**) patterns in data
  - Find novel information.
  - Better understanding data.

Pattern Set Mining
- Find **small set** relevant patterns in data
  - Build classifiers/predictors.
  - Describe/Explain data.

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F



Aoga et al. - CP'19
Oct. 2nd 2019

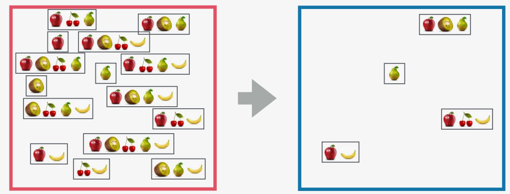
PATTERN-BASED LEARNING
Pattern mining vs Pattern set Mining

PATTERN MINING ≠ PATTERN SET MINING

Pattern Mining
- Find (**all**) patterns in data
  - Find novel information.
  - Better understanding data.

Pattern Set Mining
- Find **small set** relevant patterns in data
  - Build classifiers/predictors.
  - Describe/Explain data.

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

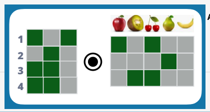


Aoga et al. - CP'19
Oct. 2nd 2019

PATTERN SET MINING
Examples

IF { } THEN ♀
ELSE IF { } THEN ♀
ELSE IF { } THEN ♂
ELSE (default case) THEN ♀

RULE LEARNING
FIND A LIST OF K RULES R SUCH THAT ERROR(COVER(R)) IS MINIMAL



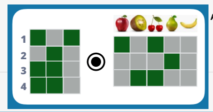
BOOLEAN MATRIX FACTORIZATION
FIND TWO BOOLEAN MATRICES $A_{n \times k}$ & $B_{k \times m}$, SUCH THAT ERROR(A∘B,D) IS MINIMAL

Aoga et al. - CP'19
Oct. 2nd 2019

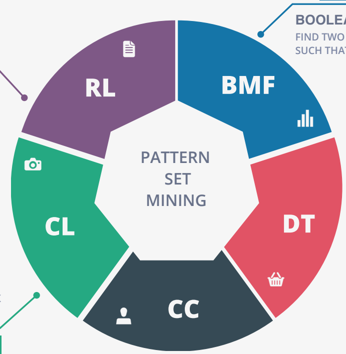
PATTERN SET MINING
Examples

IF { } THEN ♀
ELSE IF { } THEN ♀
ELSE IF { } THEN ♂
ELSE (default case) THEN ♀

RULE LEARNING
FIND A LIST OF K RULES R SUCH THAT ERROR(COVER(R)) IS MINIMAL

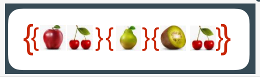


BOOLEAN MATRIX FACTORIZATION
FIND TWO BOOLEAN MATRICES $A_{n \times k}$ & $B_{k \times m}$, SUCH THAT ERROR(A∘B,D) IS MINIMAL



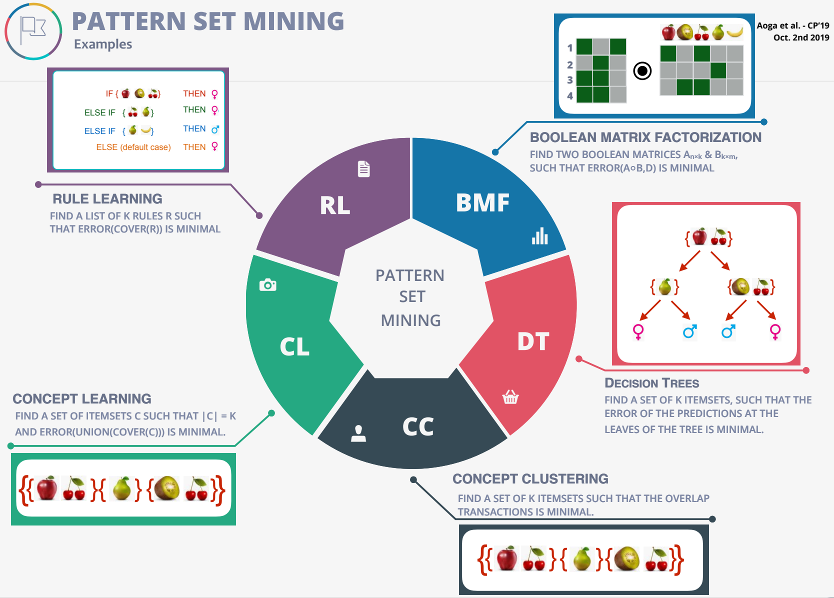
CONCEPT LEARNING
FIND A SET OF ITEMSETS C SUCH THAT |C| = K AND ERROR(UNION(COVER(C))) IS MINIMAL.

CONCEPT CLUSTERING
FIND A SET OF K ITEMSETS SUCH THAT THE OVERLAP TRANSACTIONS IS MINIMAL.



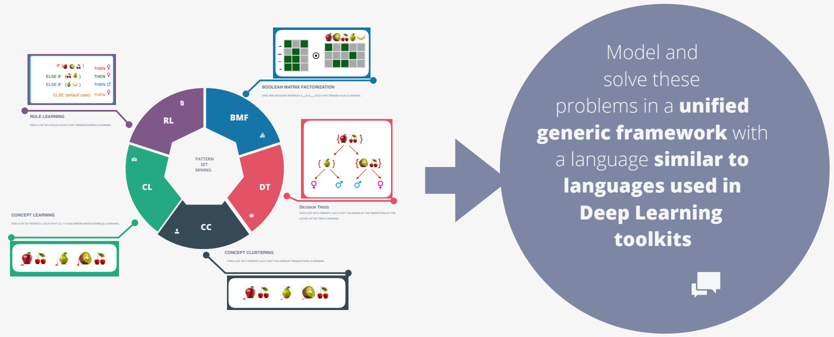
PATTERN SET MINING

Examples



OUR GOAL

Pattern-set problems model and solve in a generic framework



WHY ?

Pattern-set problems model and solve in a generic framework

- Unified Generic Framework for Pattern set learning problems
  - **Increasingly investigated problems** (interpretable solutions, many applications but many challenges as well).
  - Solving each problems **need a lot of effort** and the solution is often solver-dependent.
  - Making Pattern set mining **more accessible to CSP community**.

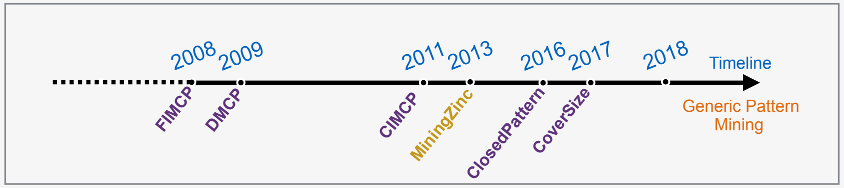
WHY ?

Pattern-set problems model and solve in a generic framework

- Unified Generic Framework for Pattern set learning problems
  - **Increasingly investigated problems** (interpretable solutions, many applications but many challenges as well).
  - Solving each problems **need a lot of effort** and the solution is often solver-dependent.
  - Making Pattern set mining **more accessible to CSP community**.
- Similar to languages used in Machine Learning (deep learning)
  - Researchers in Machine Learning are **used to model problems like this**.
  - We want our framework to be **easy and natural** to use for them.

RELATED WORK

Generic Framework for Pattern Mining vs Generic Framework for Pattern set Mining



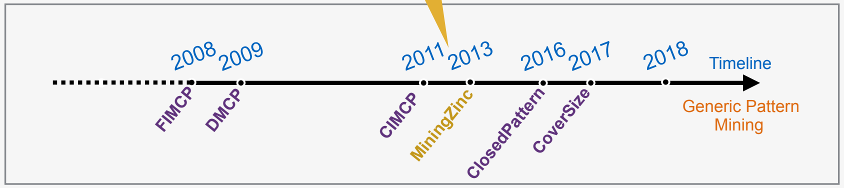
Guns, T., Dries, A., Tack, G., Nijssen, S., De Raedt, L.: MiningZinc: a modeling language for constraint-based mining. In: 23rd IJCAI (2013)

Aoga et al. - CP'19
Oct. 2nd 2019

RELATED WORK

Generic Framework for Pattern Mining vs Generic Framework for Pattern set Mining

Language for constraint-based pattern mining ⇏ Not available for Pattern **set** Mining



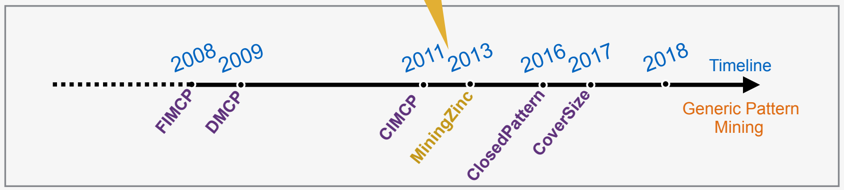
Guns, T., Dries, A., Tack, G., Nijssen, S., De Raedt, L.: MiningZinc: a modeling language for constraint-based mining. In: 23rd IJCAI (2013)

Aoga et al. - CP'19
Oct. 2nd 2019

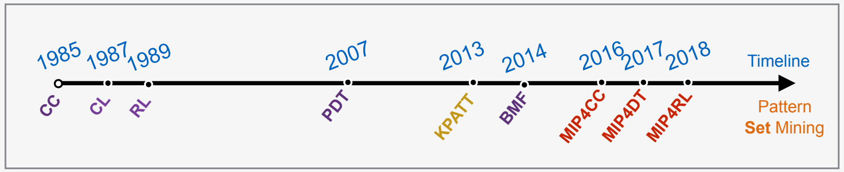
RELATED WORK

Generic Framework for Pattern Mining vs Generic Framework for Pattern set Mining

Language for constraint-based pattern mining ⇏ Not available for Pattern **set** Mining



Guns, T., Dries, A., Tack, G., Nijssen, S., De Raedt, L.: MiningZinc: a modeling language for constraint-based mining. In: 23rd IJCAI (2013)



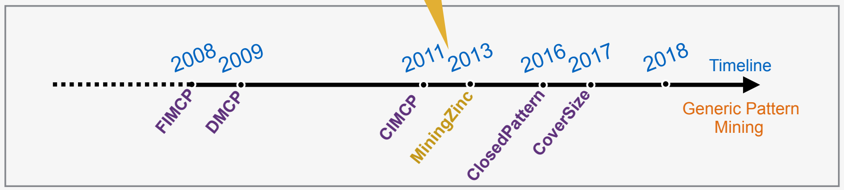
Guns, T., Nijssen, S., De Raedt, L.: k-Pattern set mining under constraints. IEEE Trans. Knowl. Data Eng. 25(2), 402–418 (2013)

Aoga et al. - CP'19
Oct. 2nd 2019

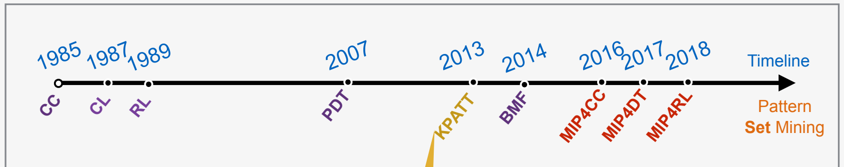
RELATED WORK

Generic Framework for Pattern Mining vs Generic Framework for Pattern set Mining

Language for constraint-based pattern mining ⇏ Not available for Pattern **set** Mining



Guns, T., Dries, A., Tack, G., Nijssen, S., De Raedt, L.: MiningZinc: a modeling language for constraint-based mining. In: 23rd IJCAI (2013)



Guns, T., Nijssen, S., De Raedt, L.: k-Pattern set mining under constraints. IEEE Trans. Knowl. Data Eng. 25(2), 402–418 (2013)

Formalize Pattern set Mining Problems as Constraint Optimizations Problems ⇏ Apply to relatively small datasets (scalability issues)

Pattern set Problems as Parameter Learning Approach

How we do that?

PARAMETER LEARNING

What is parameter learning ?

CONTRIBUTION

C — Constraints

W — Variables

X — Inputs

CSP Model!

$$\textit{Minimize} \quad error(W, y)$$
$$\text{s.t. } C(W, X) \text{ are satisfied}$$

Complete Assignment of W

PARAMETER LEARNING

What is parameter learning ?

CONTRIBUTION

C — Constraints

W — Variables

X — Inputs

CSP Model!

$$\textit{Minimize} \quad error(W, y)$$
$$\text{s.t. } C(W, X) \text{ are satisfied}$$

Complete Assignment of W

HOW CAN WE REFORMULATE PATTERN SET MINING LIKE THIS?

PATTERN SET MINING

Definition

Aoga et al. - CP'19
Oct. 2nd 2019

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

- Pattern: Itemset { }
  - Cover ({ }) = {1,3,4}
- Pattern set: Set of Itemsets
  - {{ } { } { }}
- Task: Minimize Covers Error
  - Error(Cover({ })) = 1

PATTERN SET MINING

Definition

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

- Pattern: Itemset { }
  - Cover ({ }) = {1,3,4}
- Pattern set: Set of Itemsets
  - {{ } { } { }}
- Task: Minimize Covers Error
  - Error(Cover({ })) = 1

Aoga et al. - CP'19
Oct. 2nd 2019

PATTERN SET MINING

Definition

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

- Pattern: Itemset { }
  - Cover ({ }) = {1,3,4}
- Pattern set: Set of Itemsets
  - {{ } { } { }}
- Task: Minimize Covers Error
  - Error(Cover({ })) = 1

Aoga et al. Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4 11

PATTERN SET MINING

Example: learning a list of rules

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	F/M
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

IF { } THEN ♀
ELSE IF { } THEN ♀
ELSE IF { } THEN ♂
ELSE (default case) THEN ♀

Rule Learning
Find a list of k rules R such that error(cover(R)) is minimal

Aoga et al. - CP'19
Oct. 2nd 2019

PATTERN SET MINING

Example: learning a list of rules

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	F/M
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

Error = 0

IF { } THEN ♀
ELSE IF { } THEN ♀
ELSE IF { } THEN ♂
ELSE (default case) THEN ♀

Rule Learning
Find a list of k rules R such that error(cover(R)) is minimal

Aoga et al. Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4 12

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

PATTERN SET MINING
Example: Boolean Factorization

Aoga et al. - CP'19
Oct. 2nd 2019

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

Aoga et al. Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4 13

PATTERN SET MINING
Example: Boolean Factorization

Aoga et al. - CP'19
Oct. 2nd 2019

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

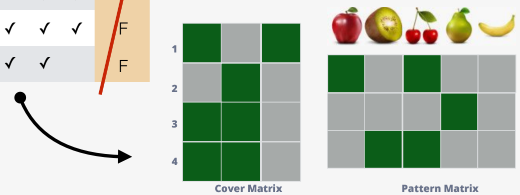
Aoga et al. Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4 13

PATTERN SET MINING
Example: Boolean Factorization

Aoga et al. - CP'19
Oct. 2nd 2019

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

	a	b	c	d	e	M/F
1	✓	✓	✓		✓	F
2	✓	✓		✓	✓	M
3	✓		✓	✓	✓	F
4	✓		✓	✓		F

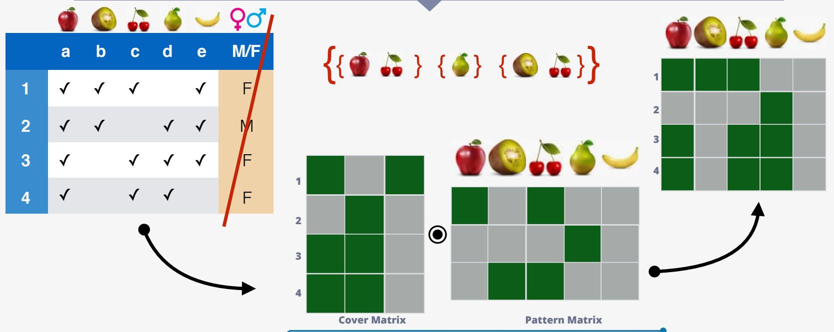


BOOLEAN MATRIX FACTORIZATION

FIND TWO BOOLEAN MATRICES $A_{n \times k}$ & $B_{k \times m}$, SUCH THAT ERROR($A \circ B$,D) IS MINIMAL

PATTERN SET MINING
Example: Boolean Factorization

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL

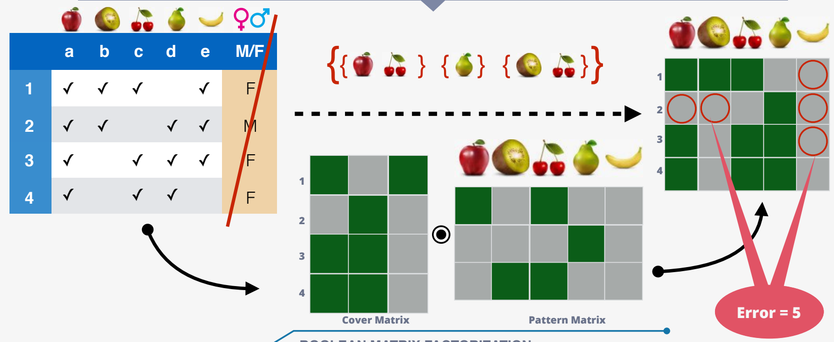


BOOLEAN MATRIX FACTORIZATION
FIND TWO BOOLEAN MATRICES $A_{n\times k}$ & $B_{k\times m}$, SUCH THAT ERROR($A \circ B$,D) IS MINIMAL

Aoga et al. - CP'19
Oct. 2nd 2019

PATTERN SET MINING
Example: Boolean Factorization

PATTERN SET MINING = FINDING A SET OF PATTERNS THAT TOGETHER PERFORM A GIVEN TASK WELL



BOOLEAN MATRIX FACTORIZATION
FIND TWO BOOLEAN MATRICES $A_{n\times k}$ & $B_{k\times m}$, SUCH THAT ERROR($A \circ B$,D) IS MINIMAL

CONTRIBUTION

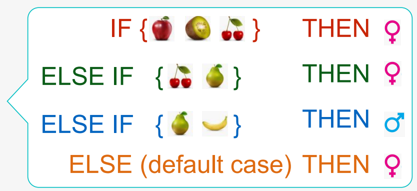
OUR APPROACH - GOAL & CONTRIBUTION
Main contributions

Goal: Build new Framework to Model Pattern Set Mining as Parameter Learning using Boolean Circuits

- ☑ *Generic setting of parameter learning in boolean circuits.*
- ☑ *Modeling language close to that of deep learning toolkits.*
- ☑ *Models for a number of mining problems in this language.*
- ☑ *Two approaches for solving: based on Mixed Integer Programming and greedy.*
- ☑ *Performance that is better than of alternative generic approach.*

CONTRIBUTION

KEY INSIGHTS
We represent pattern sets in boolean circuits

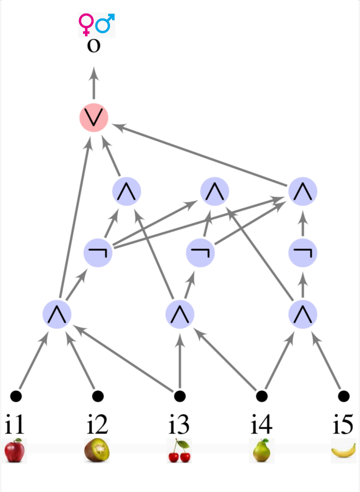


Rule Learning
Find a list of k rules R such that error(cover(R)) is minimal

CONTRIBUTION

KEY INSIGHTS

We represent pattern sets in boolean circuits



IF { 🍎 🥝 🍒 } THEN ♀
ELSE IF { 🍒 🍐 } THEN ♀
ELSE IF { 🍐 🍌 } THEN ♂
ELSE (default case) THEN ♀

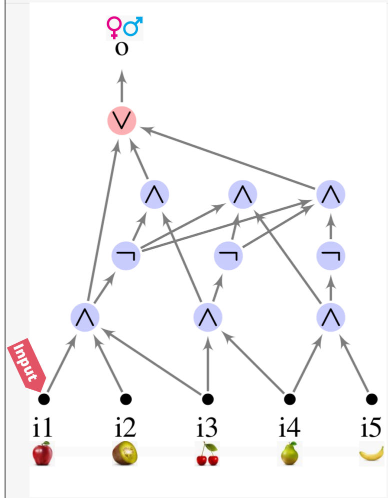
Rule Learning

Find a list of k rules R such that error(cover(R)) is minimal

CONTRIBUTION

KEY INSIGHTS

We represent pattern sets in boolean circuits



IF { 🍎 🥝 🍒 } THEN ♀
ELSE IF { 🍒 🍐 } THEN ♀
ELSE IF { 🍐 🍌 } THEN ♂
ELSE (default case) THEN ♀

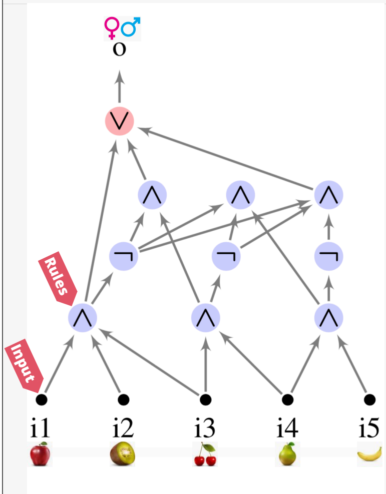
Rule Learning

Find a list of k rules R such that error(cover(R)) is minimal

CONTRIBUTION

KEY INSIGHTS

We represent pattern sets in boolean circuits



IF { 🍎 🥝 🍒 } THEN ♀
ELSE IF { 🍒 🍐 } THEN ♀
ELSE IF { 🍐 🍌 } THEN ♂
ELSE (default case) THEN ♀

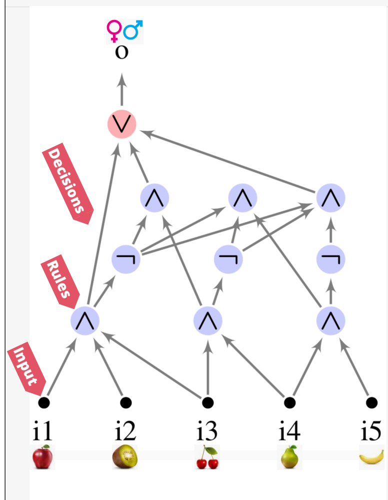
Rule Learning

Find a list of k rules R such that error(cover(R)) is minimal

CONTRIBUTION

KEY INSIGHTS

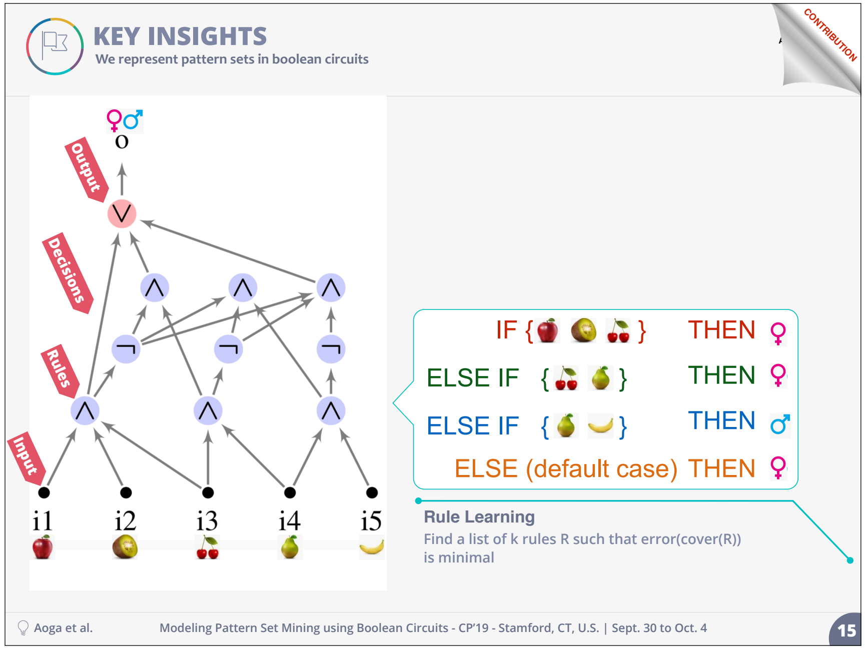
We represent pattern sets in boolean circuits

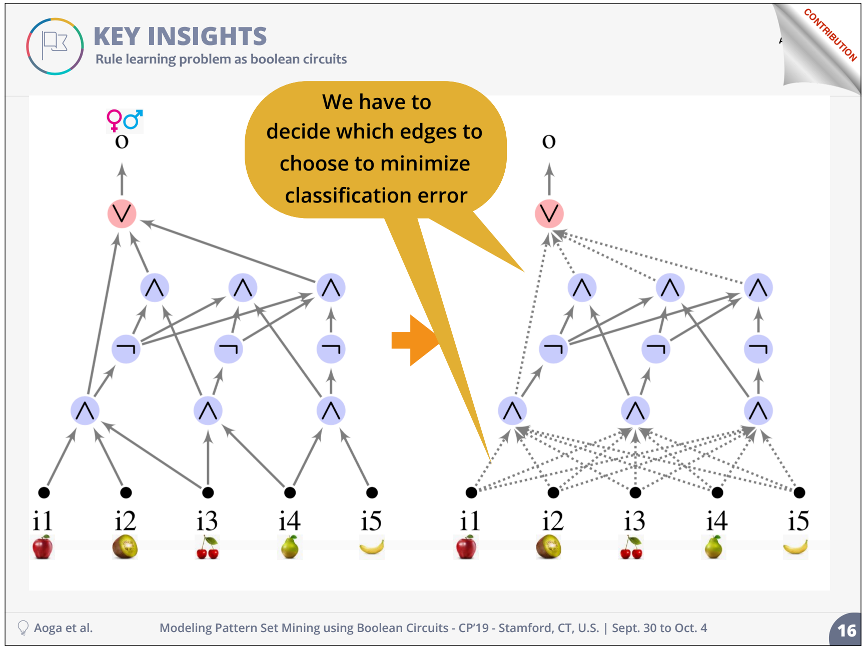


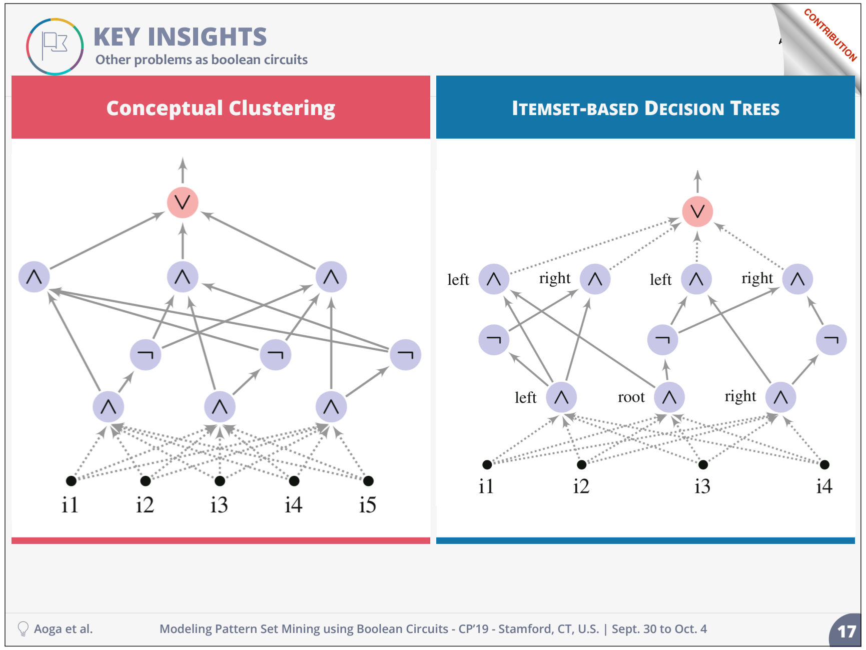
IF { 🍎 🥝 🍒 } THEN ♀
ELSE IF { 🍒 🍐 } THEN ♀
ELSE IF { 🍐 🍌 } THEN ♂
ELSE (default case) THEN ♀

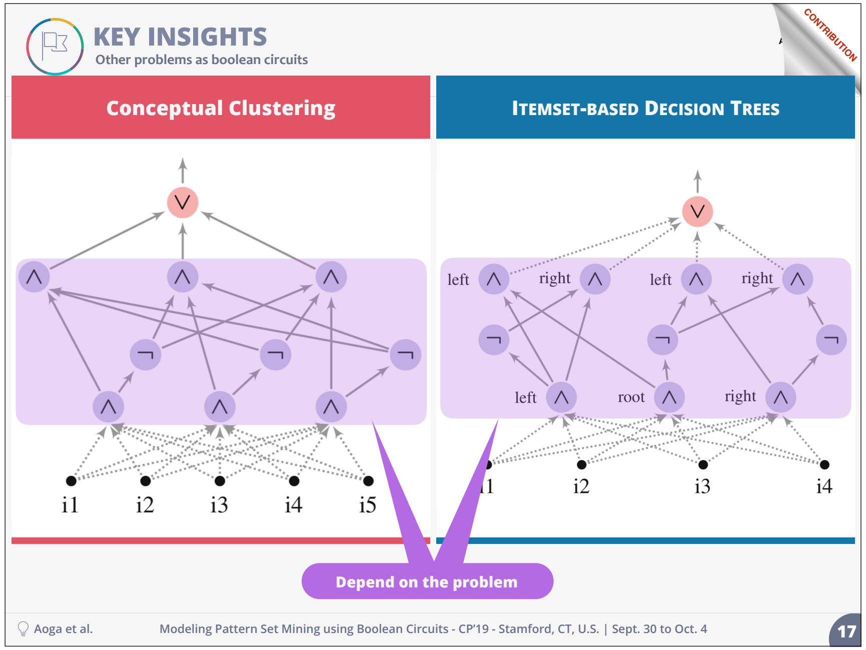
Rule Learning

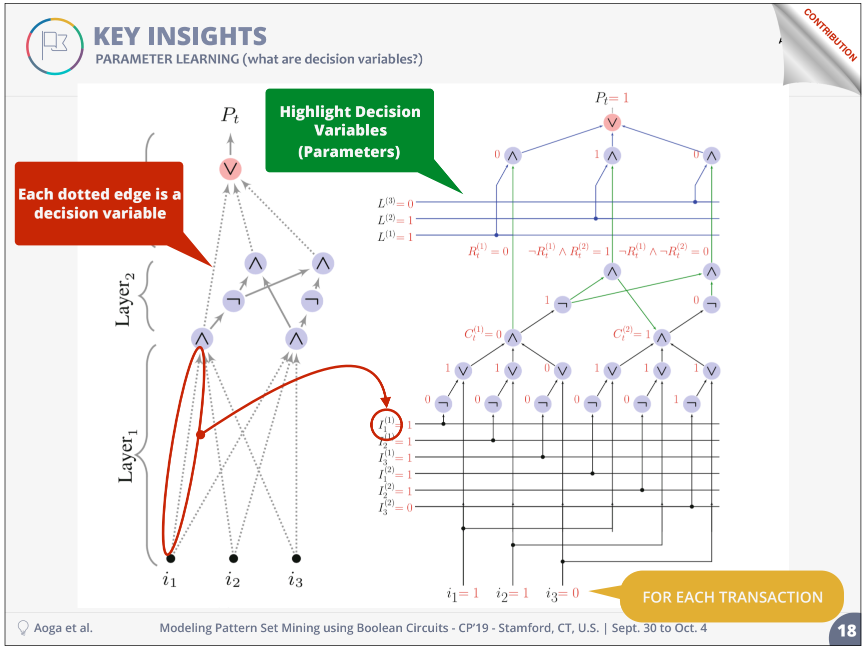
Find a list of k rules R such that error(cover(R)) is minimal


KEY INSIGHTS
We represent pattern sets in boolean circuits
CONTRIBUTION
Output
Decisions
Rules
Input
o
i1
i2
i3
i4
i5
IF { } THEN ♀
ELSE IF { } THEN ♀
ELSE IF { } THEN ♂
ELSE (default case) THEN ♀
Rule Learning
Find a list of k rules R such that error(cover(R)) is minimal
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
15

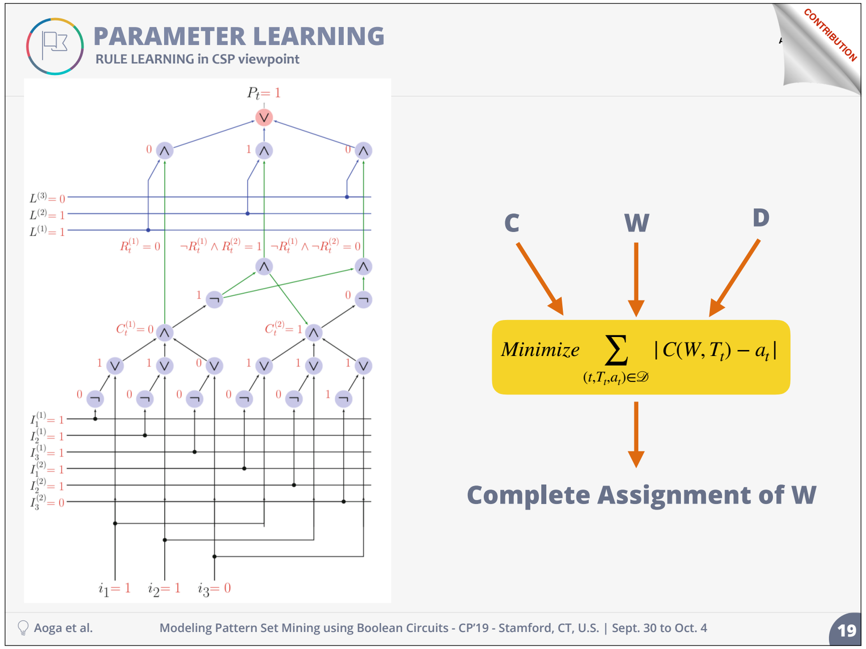

KEY INSIGHTS
Rule learning problem as boolean circuits
CONTRIBUTION
We have to decide which edges to choose to minimize classification error
o
i1
i2
i3
i4
i5
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
16


KEY INSIGHTS
Other problems as boolean circuits
CONTRIBUTION
Conceptual Clustering
Itemset-based Decision Trees
left
right
left
right
left
root
right
i1
i2
i3
i4
i5
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
17

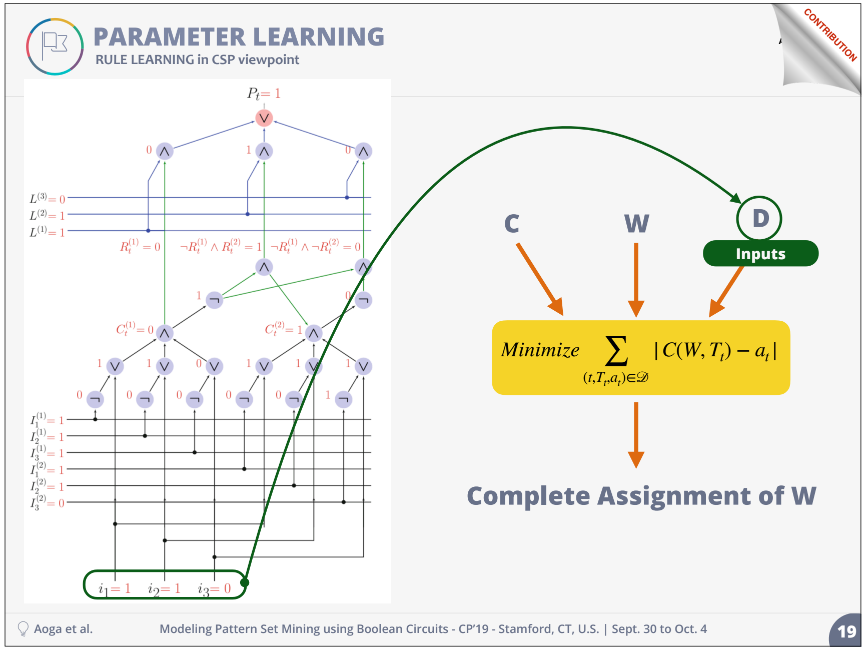

KEY INSIGHTS
Other problems as boolean circuits
CONTRIBUTION
Conceptual Clustering
Itemset-based Decision Trees
left
right
left
right
left
root
right
i1
i2
i3
i4
i5
Depend on the problem
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
17



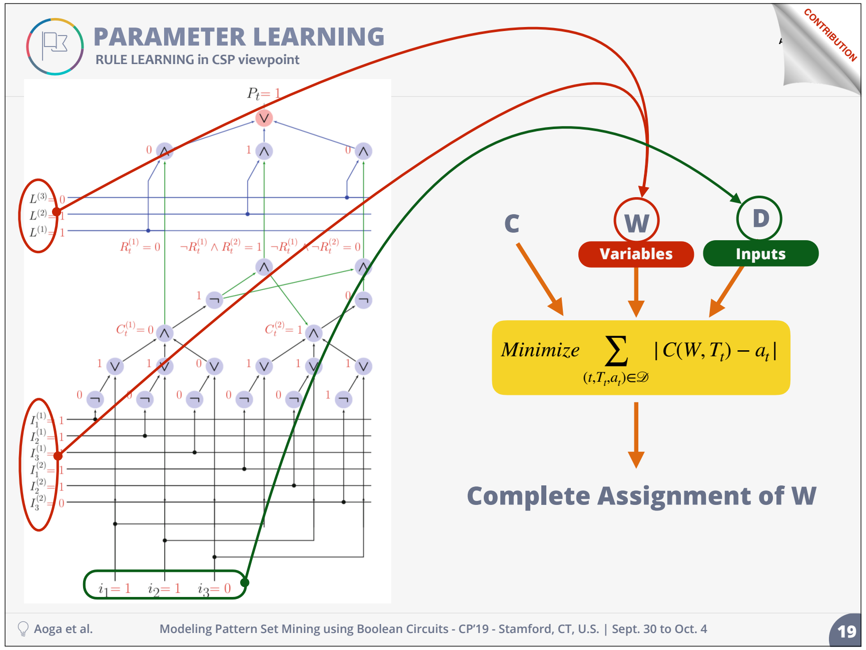
KEY INSIGHTS
PARAMETER LEARNING (what are decision variables?)
CONTRIBUTION
Highlight Decision Variables (Parameters)
Each dotted edge is a decision variable
P_t
Layer$_2$
Layer$_1$
i_1 i_2 i_3
$P_t = 1$
$i_1 = 1$ $i_2 = 1$ $i_3 = 0$
FOR EACH TRANSACTION
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
18



PARAMETER LEARNING
RULE LEARNING in CSP viewpoint
CONTRIBUTION
C
W
D
Minimize $\sum_{(t,T_t,a_t)\in\mathcal{D}} |C(W,T_t) - a_t|$
Complete Assignment of W
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
19



PARAMETER LEARNING
RULE LEARNING in CSP viewpoint
CONTRIBUTION
C
W
D
Inputs
Minimize $\sum_{(t,T_t,a_t)\in\mathcal{D}} |C(W,T_t) - a_t|$
Complete Assignment of W
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
19

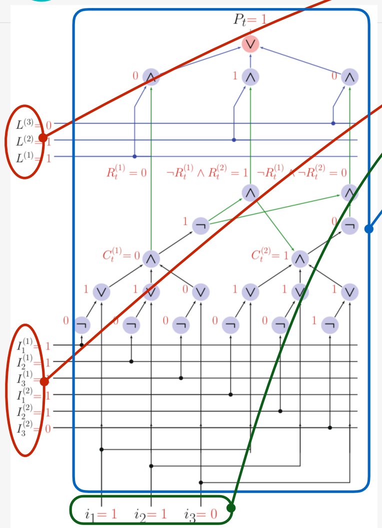


PARAMETER LEARNING
RULE LEARNING in CSP viewpoint
CONTRIBUTION
C
W
Variables
D
Inputs
Minimize $\sum_{(t,T_t,a_t)\in\mathcal{D}} |C(W,T_t) - a_t|$
Complete Assignment of W
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
19

PARAMETER LEARNING

RULE LEARNING in CSP viewpoint

CONTRIBUTION



C Constraints — **W** Variables — **D** Inputs

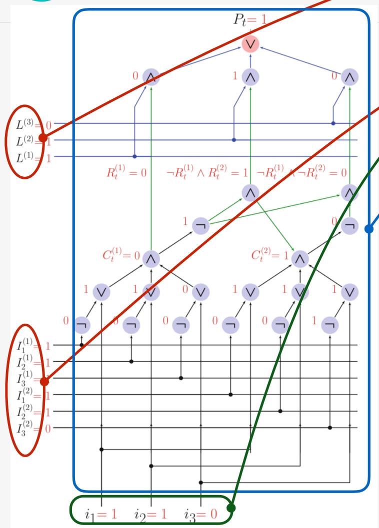
$$Minimize \sum_{(t,T_t,a_t)\in\mathcal{D}} |C(W,T_t) - a_t|$$

Complete Assignment of W

PARAMETER LEARNING

RULE LEARNING in CSP viewpoint

CONTRIBUTION



C Constraints — **W** Variables — **D** Inputs

$$Minimize \sum_{(t,T_t,a_t)\in\mathcal{D}} |C(W,T_t) - a_t|$$

Complete Assignment of W

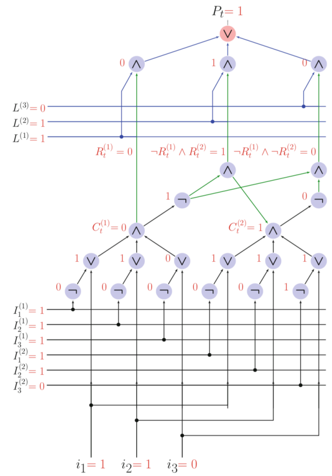
Can be Solved with CP, MIP, Greedily

PARAMETER LEARNING

RULE LEARNING in MIP/CP

CONTRIBUTION

Constraints: Solving Rule Learning problem using **MIP/CP**



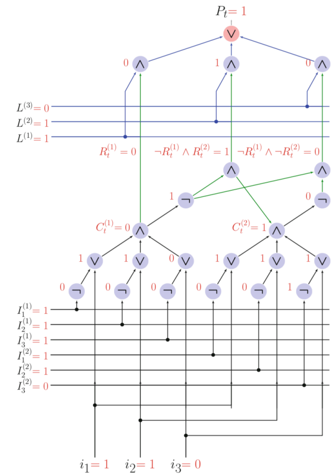
CP Use Logical constraints for each gate

PARAMETER LEARNING

RULE LEARNING in MIP/CP

CONTRIBUTION

Constraints: Solving Rule Learning problem using **MIP/CP**



CP Use Logical constraints for each gate

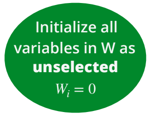
MIP Use the linearization formulations of each gate

$$\wedge \quad \begin{cases} y \geq x_i & \forall i \in [1,m] \\ y \leq \Sigma_i x_i \\ y \leq 1 \end{cases}$$

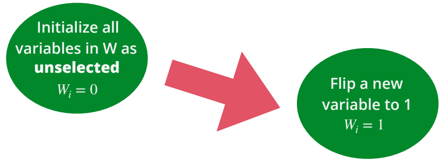
$$\vee \quad \begin{cases} y \leq x_i & \forall i \in [1,m] \\ y \geq \Sigma_i x_i - (m-1) \\ y \geq 0 \end{cases}$$

$$\neg \quad y = \neg x \equiv y = 1 - x$$

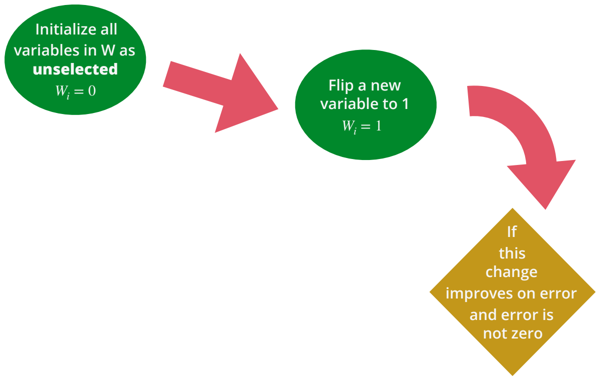
Constraints: Solving Rule Learning problem using **Greedy algorithm**



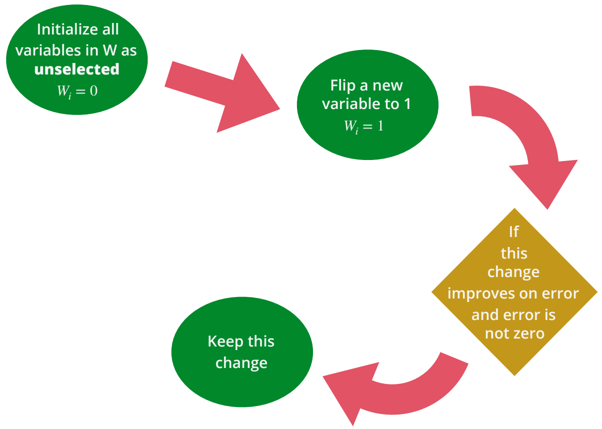
Constraints: Solving Rule Learning problem using **Greedy algorithm**

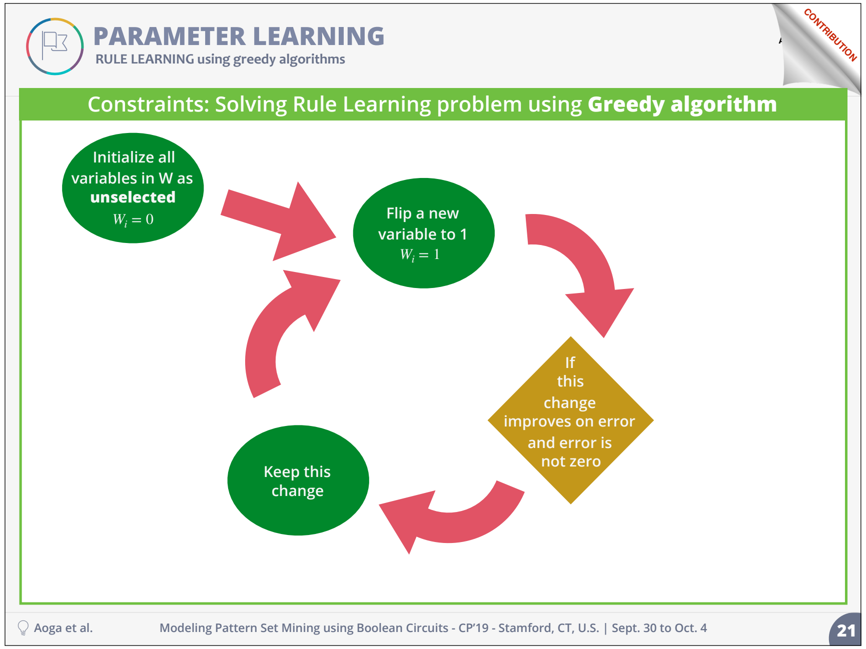


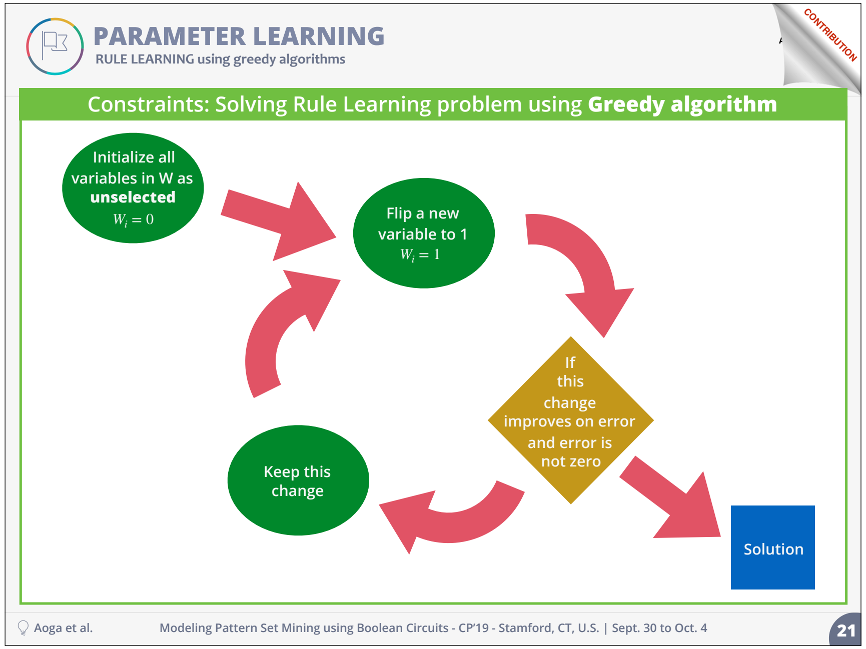
Constraints: Solving Rule Learning problem using **Greedy algorithm**

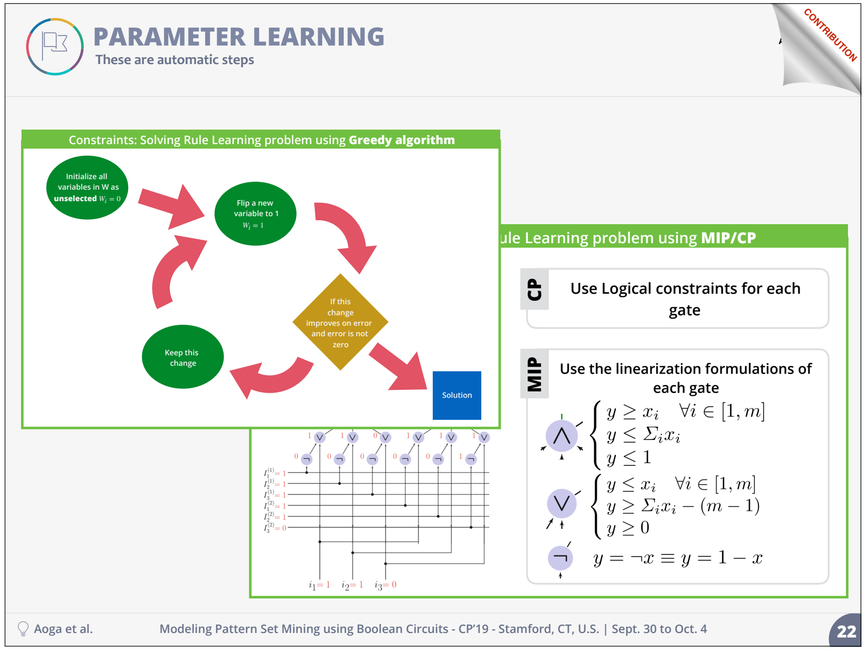


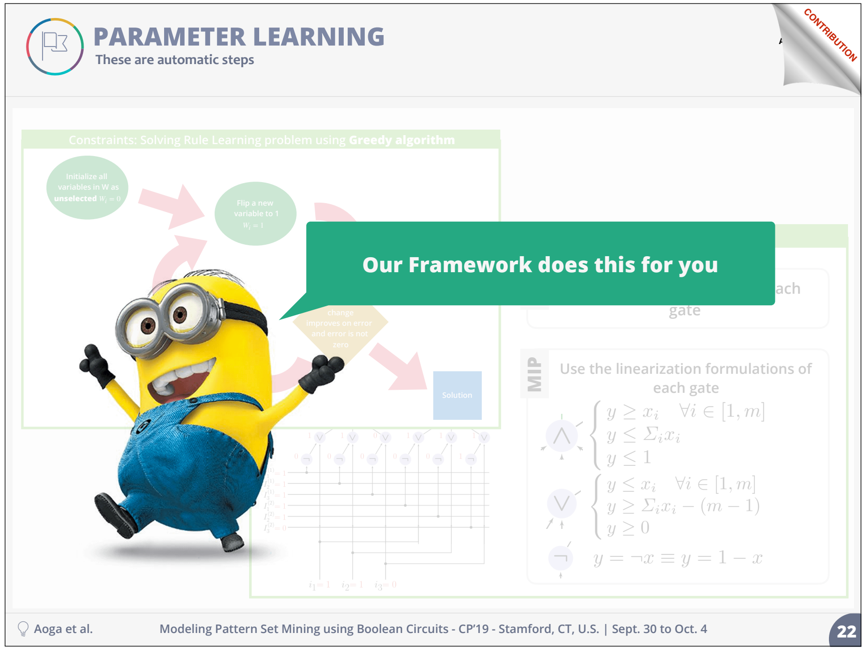
Constraints: Solving Rule Learning problem using **Greedy algorithm**




PARAMETER LEARNING
RULE LEARNING using greedy algorithms
CONTRIBUTION
Constraints: Solving Rule Learning problem using Greedy algorithm
Initialize all variables in W as unselected $W_i = 0$
Flip a new variable to 1 $W_i = 1$
If this change improves on error and error is not zero
Keep this change
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
21


PARAMETER LEARNING
RULE LEARNING using greedy algorithms
CONTRIBUTION
Constraints: Solving Rule Learning problem using Greedy algorithm
Initialize all variables in W as unselected $W_i = 0$
Flip a new variable to 1 $W_i = 1$
If this change improves on error and error is not zero
Keep this change
Solution
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
21


PARAMETER LEARNING
These are automatic steps
CONTRIBUTION
Constraints: Solving Rule Learning problem using Greedy algorithm
Initialize all variables in W as unselected $W_i = 0$
Flip a new variable to 1 $W_i = 1$
If this change improves on error and error is not zero
Keep this change
Solution
ule Learning problem using MIP/CP
CP
Use Logical constraints for each gate
MIP
Use the linearization formulations of each gate
$\wedge$ $y \geq x_i \quad \forall i \in [1, m]$, $y \leq \Sigma_i x_i$, $y \leq 1$
$\vee$ $y \leq x_i \quad \forall i \in [1, m]$, $y \geq \Sigma_i x_i - (m-1)$, $y \geq 0$
$\neg$ $y = \neg x \equiv y = 1 - x$
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
22


PARAMETER LEARNING
These are automatic steps
CONTRIBUTION
Our Framework does this for you
Aoga et al.
Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4
22

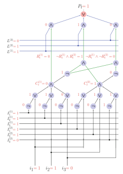
MODELING LANGUAGE
RULE LEARNING (Similar to deep learning language)

CONTRIBUTION

Algorithm 2: DSL to solve a rule learning problem with the architecture

$\triangleright$ *Building the network*
$L_1 \leftarrow \text{InputLayer}(m = 3)$
$L_2 \leftarrow AndSelectionLayer(L_1, k = 3)$
$L_3 \leftarrow NotLayer(L_2)$
$L_4 \leftarrow Layer(L_2[1],\ And(L_3[1], L_2[2]),\ And(L_3[1], L_3[2]))$
$L_5 \leftarrow OrSelectionLayer(L_4)$
$N \leftarrow L_5.network()$

Building Boolean Circuit



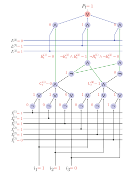
MODELING LANGUAGE
RULE LEARNING (Similar to deep learning language)

CONTRIBUTION

Algorithm 2: DSL to solve a rule learning problem with the architecture

$\triangleright$ *Building the network*
$L_1 \leftarrow \text{InputLayer}(m = 3)$
$L_2 \leftarrow AndSelectionLayer(L_1, k = 3)$
$L_3 \leftarrow NotLayer(L_2)$
$L_4 \leftarrow Layer(L_2[1],\ And(L_3[1], L_2[2]),\ And(L_3[1], L_3[2]))$
$L_5 \leftarrow OrSelectionLayer(L_4)$
$N \leftarrow L_5.network()$
$\triangleright$ *Load inputs and parameters*
$X \leftarrow getDB(\mathcal{D}) \quad y \leftarrow getAttr(\mathcal{D}) \quad \hat{y} \leftarrow L_5[0]$
$obj \leftarrow 1 - y.\hat{y} - (1 - y)(1 - \hat{y})$ $\triangleright$ *Objective function*

Building Boolean Circuit



Providing inputs and defining objective function

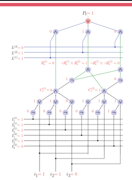
MODELING LANGUAGE
RULE LEARNING (Similar to deep learning language)

CONTRIBUTION

Algorithm 2: DSL to solve a rule learning problem with the architecture

$\triangleright$ *Building the network*
$L_1 \leftarrow \text{InputLayer}(m = 3)$
$L_2 \leftarrow AndSelectionLayer(L_1, k = 3)$
$L_3 \leftarrow NotLayer(L_2)$
$L_4 \leftarrow Layer(L_2[1],\ And(L_3[1], L_2[2]),\ And(L_3[1], L_3[2]))$
$L_5 \leftarrow OrSelectionLayer(L_4)$
$N \leftarrow L_5.network()$
$\triangleright$ *Load inputs and parameters*
$X \leftarrow getDB(\mathcal{D}) \quad y \leftarrow getAttr(\mathcal{D}) \quad \hat{y} \leftarrow L_5[0]$
$obj \leftarrow 1 - y.\hat{y} - (1 - y)(1 - \hat{y})$ $\triangleright$ *Objective function*
$\triangleright$ *Defining the procedures*
$greedy \leftarrow X$ into N using $Greedy.solver$ minimizing obj
$stats \leftarrow greedy.run()$
$mip \leftarrow X$ into N using $MIP.solver$ minimizing obj
$stats \leftarrow mip.run()$

Building Boolean Circuit



Providing inputs and defining objective function

Running the model in several solvers

MODELING LANGUAGE
RULE LEARNING (Similar to deep learning language)

CONTRIBUTION

Algorithm 2: DSL to solve a rule learning problem with the architecture

This all you have to do!

EXPERIMENTS

Our Framework vs existing approaches

THIS WORK IS NOT TO SHOW THAT IT CAN LEAD TO MUCH BETTER CLASSIFICATION OR CLUSTERING RESULTS.

THIS WORK IS NOT TO SHOW THAT IT CAN LEAD TO MUCH BETTER CLASSIFICATION OR CLUSTERING RESULTS.

- Running experiments over CL, RL, PDT, CC, BMF,...
  - **Performance criteria:** Accuracy, Time, Proof of optimality
  - **Setting time limit:** 10min, 1hour
- We mainly show experiments about CL
  - **Compared to KPATT approach**

THIS WORK IS NOT TO SHOW THAT IT CAN LEAD TO MUCH BETTER CLASSIFICATION OR CLUSTERING RESULTS.

- Running experiments over CL, RL, PDT, CC, BMF,...
  - **Performance criteria:** Accuracy, Time, Proof of optimality
  - **Setting time limit:** 10min, 1hour
- We mainly show experiments about CL
  - **Compared to KPATT approach**

methods	Audi.	Aust.	HeCl.	Hepa.	KrKp.	Lymp.	Mush.	PrTu.	Soyb.	Spli.	TTT.	Vote	Zoo
							a) Dataset Features						
$\lvert\mathcal{T}\rvert$	216	653	296	137	3196	148	8124	336	630	3190	958	435	101
$\frac{\lvert\{t\in\mathcal{T} \mid a_t=1\}\rvert}{\mathcal{T}}$	0.26	0.55	0.54	0.81	0.52	0.55	0.52	0.24	0.15	0.52	0.65	0.61	0.41
$\lvert\mathcal{I}\rvert$	148	125	95	68	74	68	119	31	50	287	27	48	36

EXPERIMENTS

Protocols and datasets

THIS WORK IS NOT TO SHOW THAT IT CAN LEAD TO MUCH BETTER CLASSIFICATION OR CLUSTERING RESULTS.

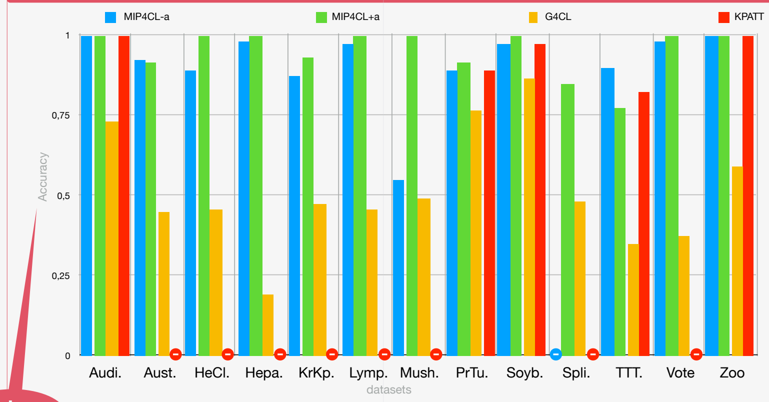
- Running experiments over CL, RL, PDT, CC, BMF,...
  - **Performance criteria:** Accuracy, Time, Proof of optimality
  - **Setting time limit:** 10min, 1hour
- We mainly show experiments about CL
  - **Compared to KPATT approach**

methods	Audi.	Aust.	HeCl.	Hepa.	KrKp.	Lymp.	Mush.	PrTu.	Soyb.	Spli.	TTT.	Vote	Zoo
a) Dataset Features													
$\|\mathcal{T}\|$	216	653	296	137	3196	148	8124	336	630	3190	958	435	101
$\frac{\|\{t \in \mathcal{T} \| a_t = 1\}\|}{\mathcal{T}}$	0.26	0.55	0.54	0.81	0.52	0.55	0.52	0.24	0.15	0.52	0.65	0.61	0.41
$\|\mathcal{I}\|$	148	125	95	68	74	68	119	31	50	287	27	48	36

EXPERIMENTS WITH CONCEPT LEARNING

How do we perform over training sets?

Concept Learning: Accuracies over training sets (k=5, timelimit=1hour, ⊖ no solution in allocated time)



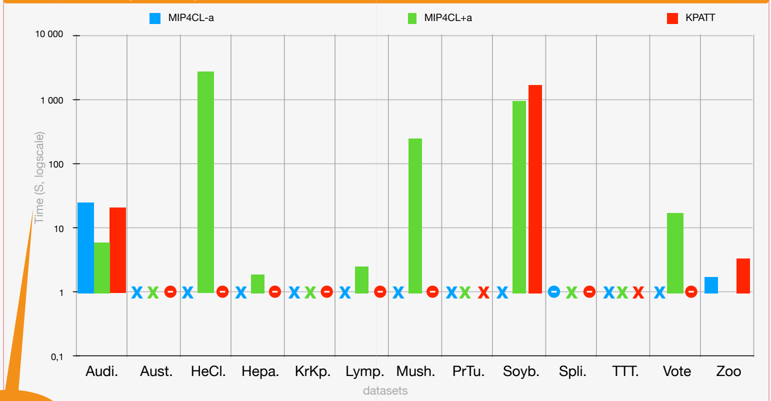
Higher - Better

Guns, T., Nijssen, S., De Raedt, L.: k-Pattern set mining under constraints. IEEE Trans. Knowl. Data Eng. 25(2), 402–418 (2013)

EXPERIMENTS WITH CONCEPT LEARNING

How do we perform over training sets?

Concept Learning: Running time (in second, k=5, **X** = timeout, time limit=1hour, ⊖ no solution in allocated time)



Lower - Better

Guns, T., Nijssen, S., De Raedt, L.: k-Pattern set mining under constraints. IEEE Trans. Knowl. Data Eng. 25(2), 402–418 (2013)

SUMMARY

Takeaways

TAKEAWAYS
Summary
Aoga et al. - CP'19
Oct. 2nd 2019

MOTIVATION AND INSPIRATION FROM DEEP LEARNING TOOLKITS

Aoga et al. Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4 30

TAKEAWAYS
Summary
Aoga et al. - CP'19
Oct. 2nd 2019

MOTIVATION AND INSPIRATION FROM DEEP LEARNING TOOLKITS

UNIFIED MODELING FRAMEWORK FOR VARIOUS K-PATTERN SET MINING PROBLEMS

REFORMULATING PATTERN SET MINING AS PARAMETER LEARNING IN LOGICAL CIRCUITS

UNIFIED MODELING LANGUAGE SIMILAR TO THAT OF DEEP LEARNING TOOLKITS

Aoga et al. Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4 30

TAKEAWAYS
Summary
Aoga et al. - CP'19
Oct. 2nd 2019

MOTIVATION AND INSPIRATION FROM DEEP LEARNING TOOLKITS

UNIFIED MODELING FRAMEWORK FOR VARIOUS K-PATTERN SET MINING PROBLEMS

REFORMULATING PATTERN SET MINING AS PARAMETER LEARNING IN LOGICAL CIRCUITS

UNIFIED MODELING LANGUAGE SIMILAR TO THAT OF DEEP LEARNING TOOLKITS

COMPETITIVE PERFORMANCE COMPARED TO THAT OF STATE OF THE ART

Aoga et al. Modeling Pattern Set Mining using Boolean Circuits - CP'19 - Stamford, CT, U.S. | Sept. 30 to Oct. 4 30

TAKEAWAYS
Summary

- MOTIVATION AND INSPIRATION FROM DEEP LEARNING TOOLKITS
- UNIFIED MODELING FRAMEWORK FOR VARIOUS K-PATTERN SET MINING PROBLEMS
- REFORMULATING PATTERN SET MINING AS PARAMETER LEARNING IN LOGICAL CIRCUITS
- UNIFIED MODELING LANGUAGE SIMILAR TO THAT OF DEEP LEARNING TOOLKITS
- COMPETITIVE PERFORMANCE COMPARED TO THAT OF STATE OF THE ART

RESEARCH QUESTION: CAN WE USE OUR APPROACH TO MAKE PATTERN MINING MORE ACCURATE OR DEEP LEARNING INTERPRETABLE/EXPLAINABLE ?

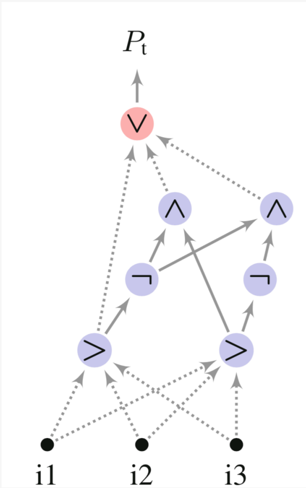
CP 2019 25 YEARS
Stamford, CT, U.S. | September 30 to October 4, 2019
The 25th International Conference on Principles and Practice of Constraint Programming

Modeling Pattern Set Mining using Boolean Circuits

John Aoga, Siegfried Nijssen, Pierre Schaus

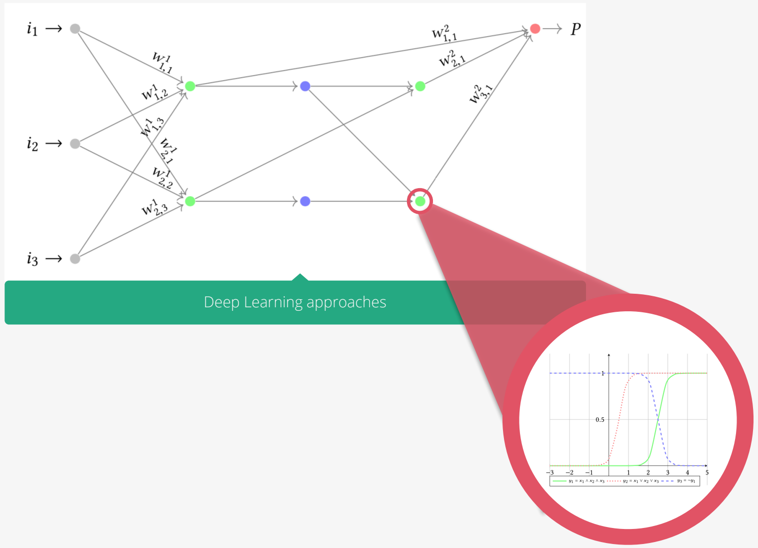
UCLouvain icteam AIA Artificial Intelligence & Algorithms fnrs

Flexibility of the model (building new set of problem)



Soft Gates

Flexibility of the model (building new set of problem)

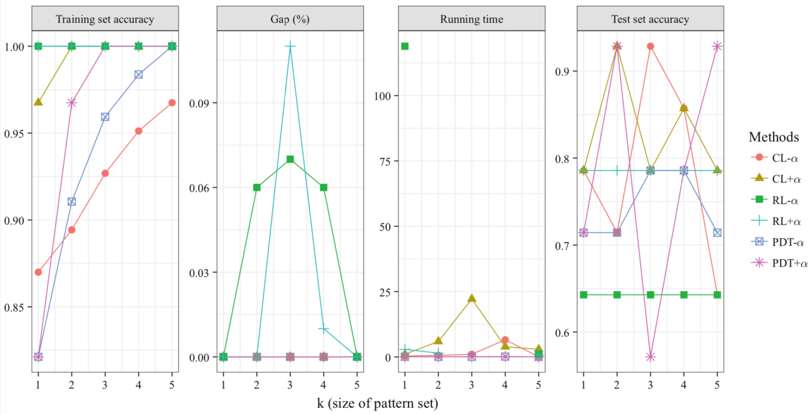


Deep Learning approaches

Sensitivity to the parameter k of our approaches

Time limit = 600s (10min)

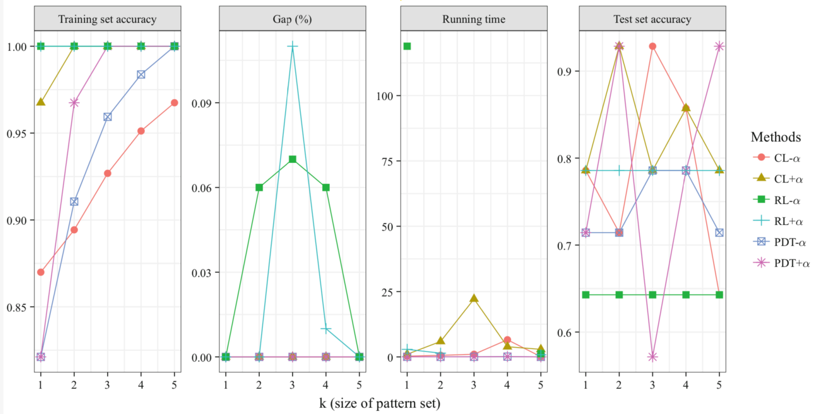
Hepatitis dataset (137 transactions and 68 items)



Sensitivity to the parameter k of our approaches

Time limit = 600s (10min)

Hepatitis dataset (137 transactions and 68 items)



Aoga John
Global constraints for mining sets and sequences - Public Defence (09-07-2019)
34