Learning constraint models from data

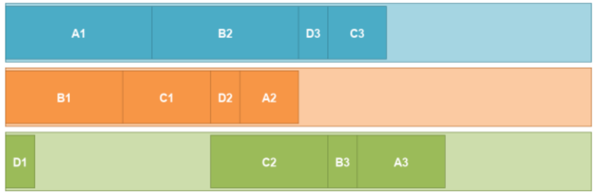
Dimos Tsouros, Kostas Stergiou , Tias Guns

dimos.tsouros@kuleuven.be, kstergiou@uowm.gr , tias.guns@kuleuven.be

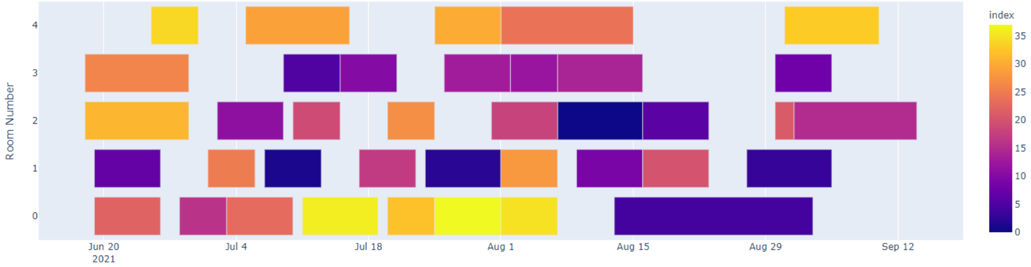
07/02/2023

Introduction

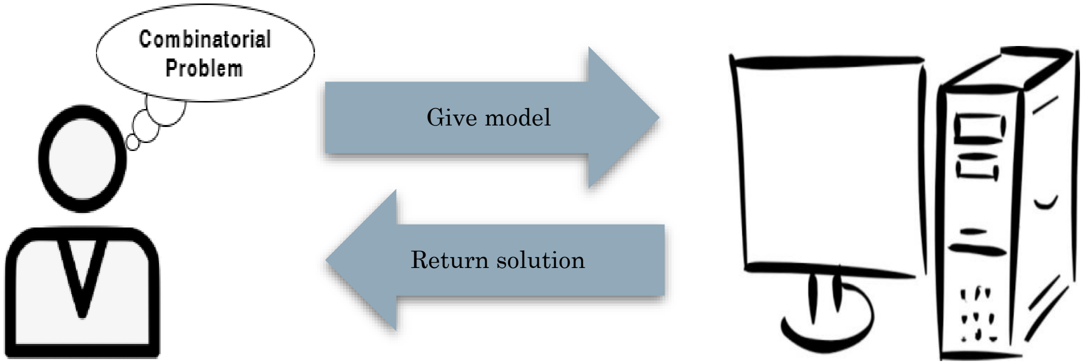
- Constraint programming (CP)
  - ❑ Solving combinatorial problems in AI



3 x 3
1 x 1
5 x 5
6 x 6
4 x 4
2 x 2
7 x 7
8 x 8



- Model + Solve paradigm

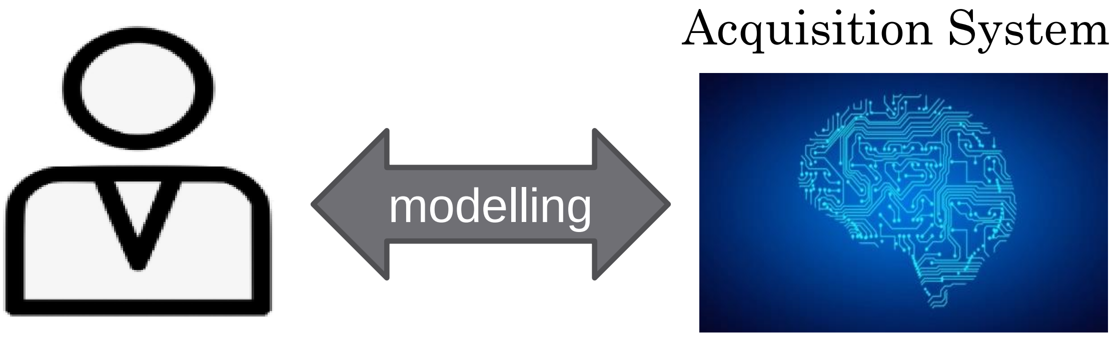


Combinatorial Problem

Give model

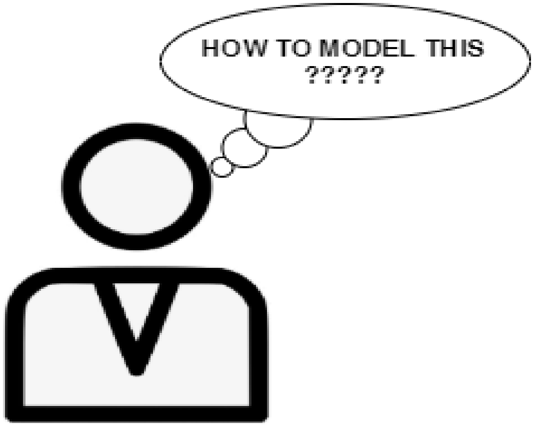
Return solution

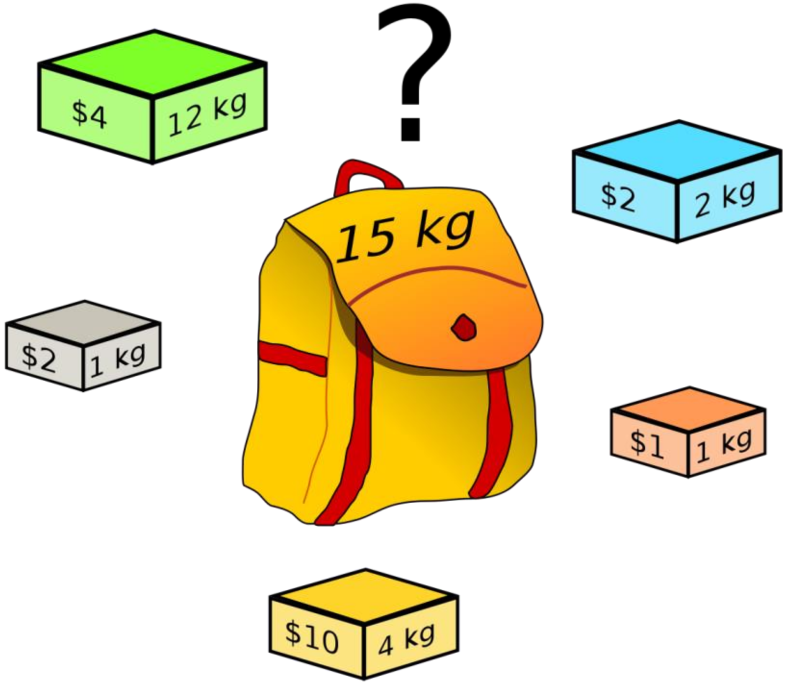
Introduction



Modelling is not always trivial

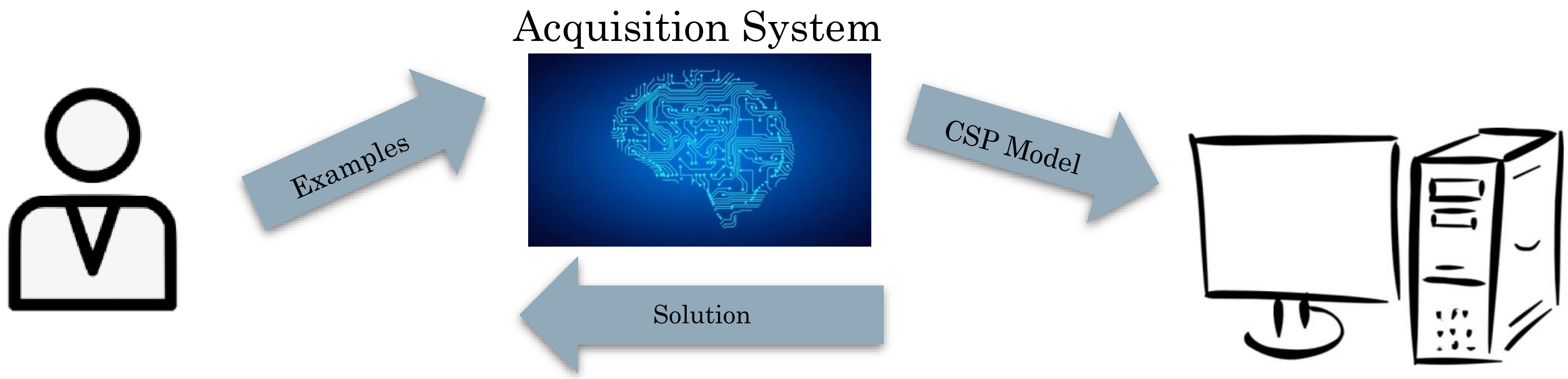
- Requires expertise
- Bottleneck for the wider use of CP





Introduction (4/4)

Passive acquisition



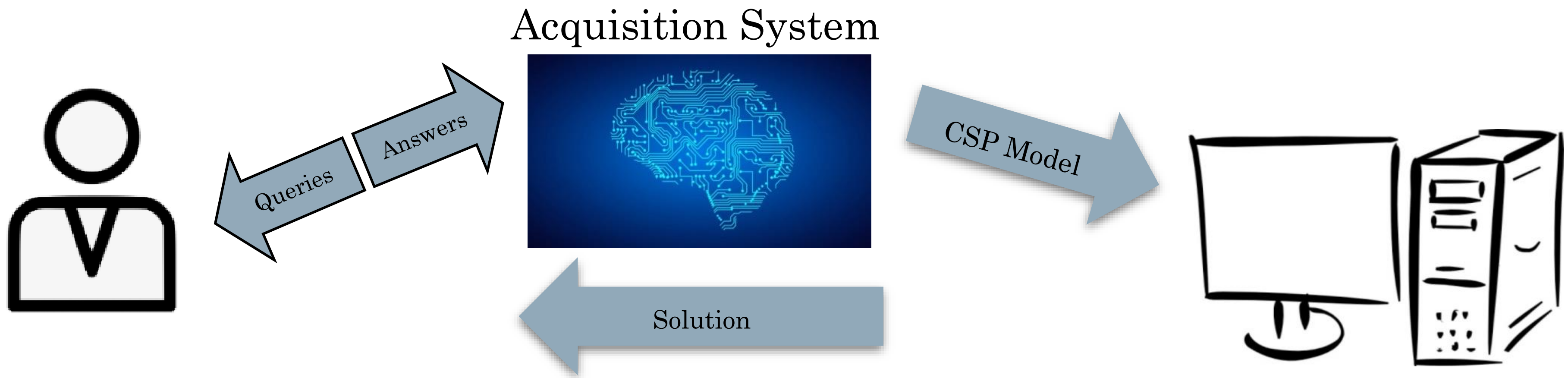
Acquisition System

Examples

CSP Model

Solution

Interactive acquisition



Acquisition System

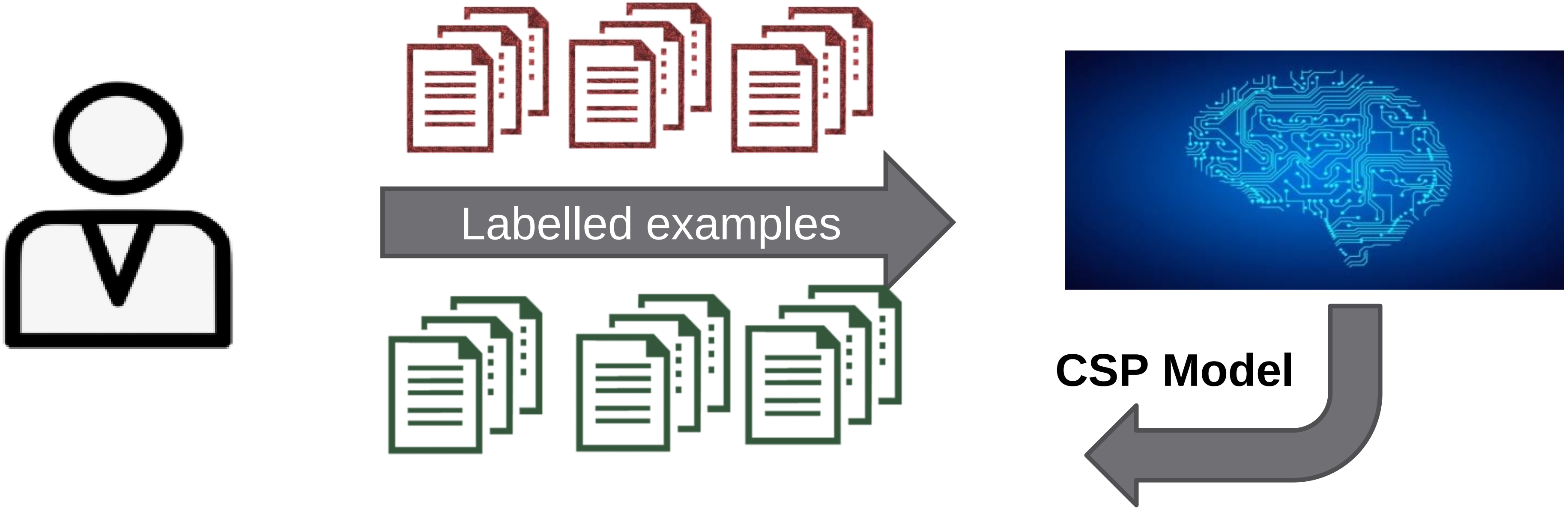
Queries

Answers

CSP Model

Solution

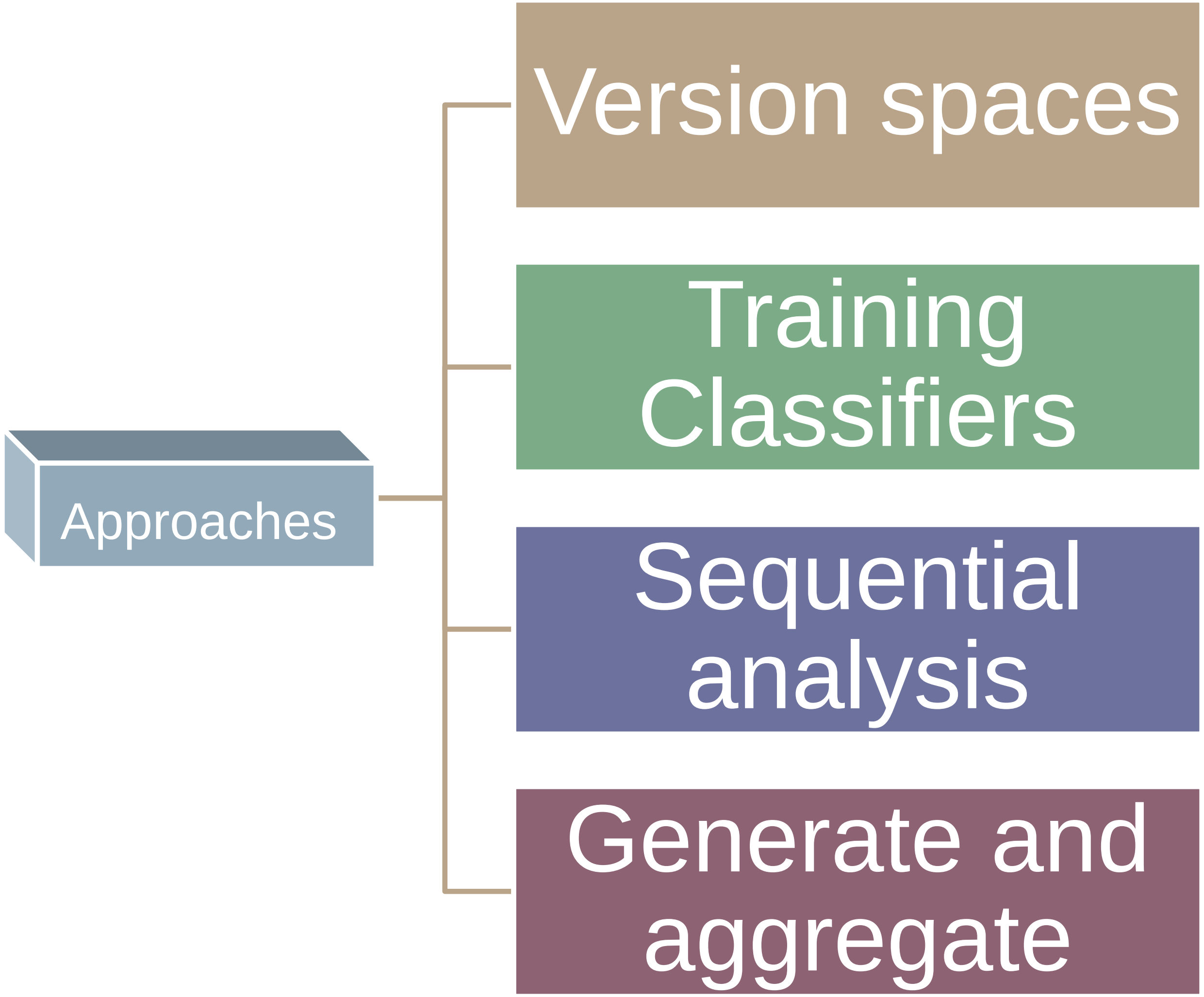
Passive Acquisition



Assignment to all variables of the problem, labelled as:

- a solution
- or a non solution

Passive Acquisition



Approaches

Version spaces

- ConAcq
- Modelseeker
- CABSC
- MineAsk

Training Classifiers

- ClassAcq
- BAYESACQ
- EML

Sequential analysis

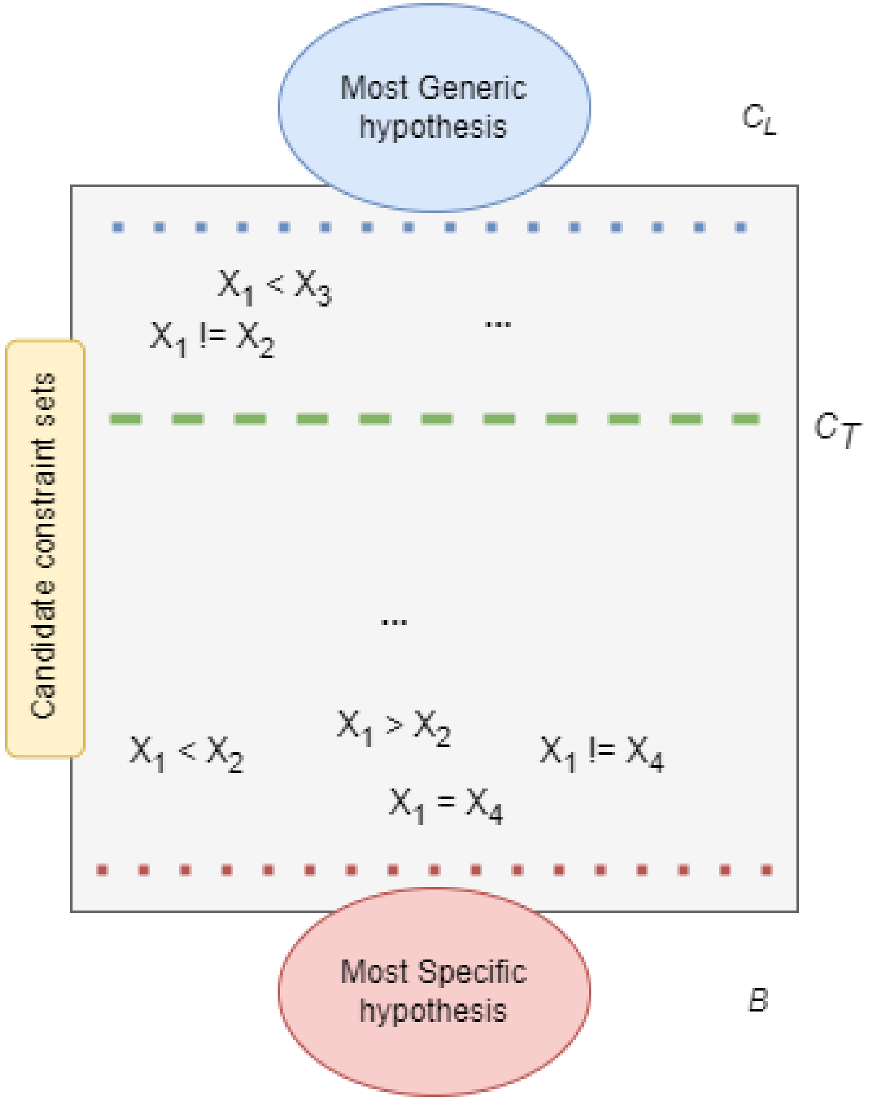
- SEQACQ

Generate and aggregate

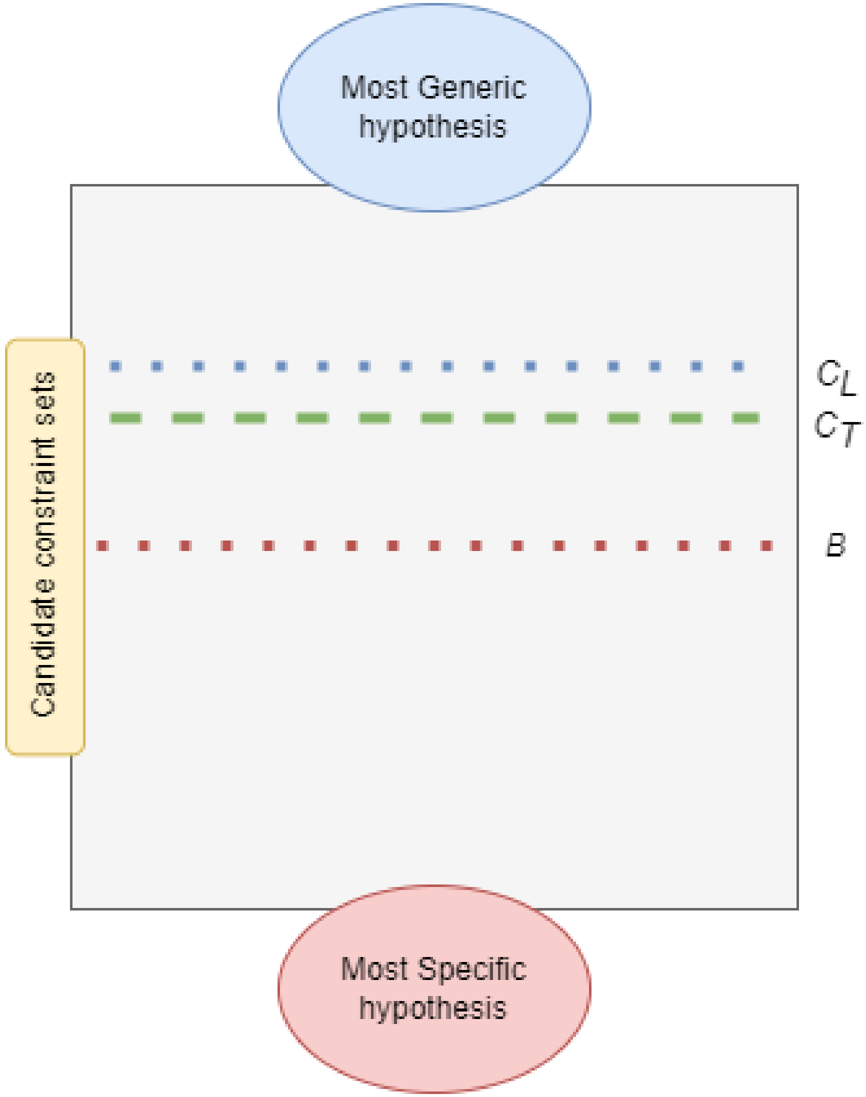
- COUNT-CP

Adapting Candidate Elimination

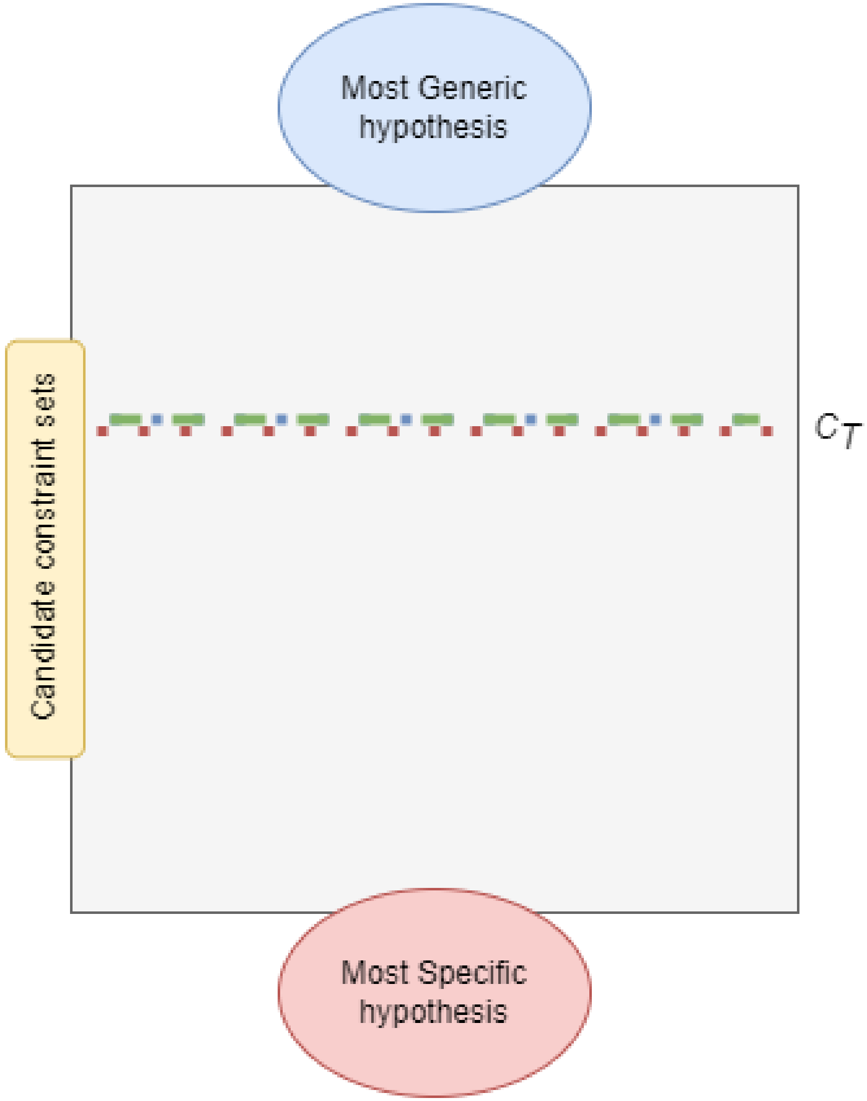
- **C_L**: learned set of constraints
- **B**: set of all candidate constraints
- **C_T**: target set of constraints



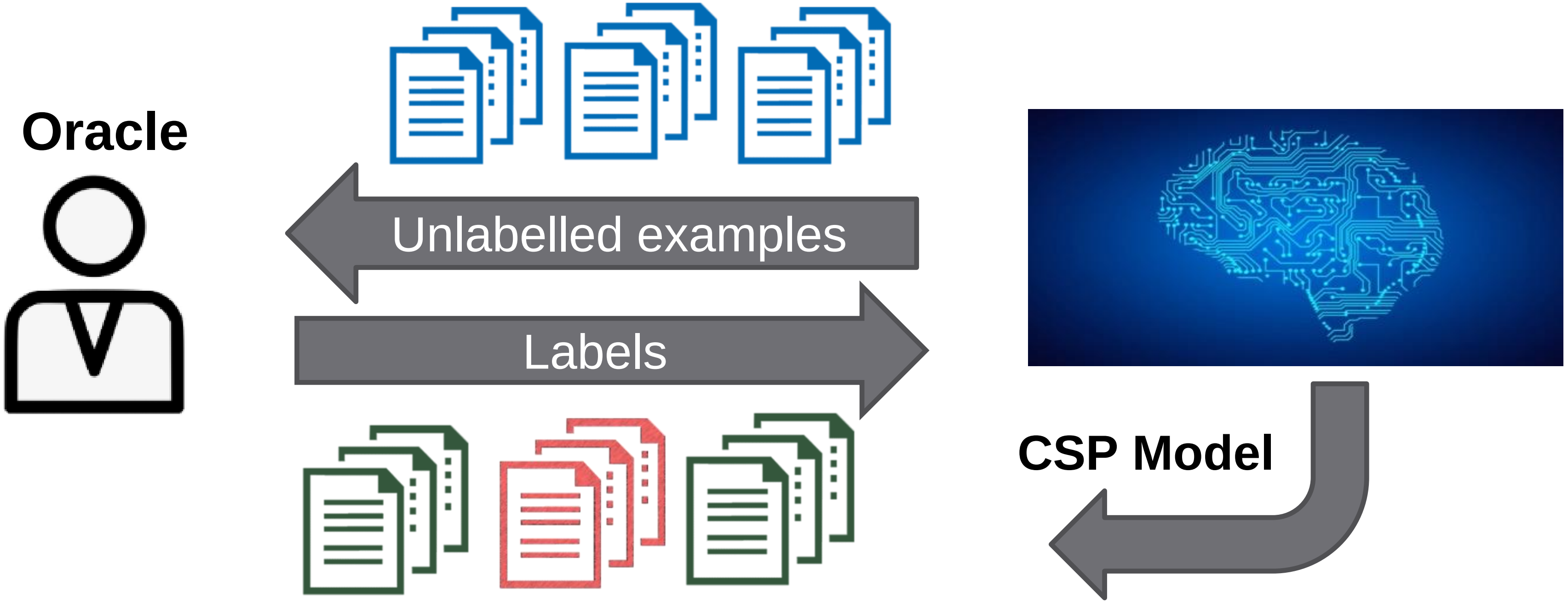
Adapting Candidate Elimination



Adapting Candidate Elimination



Interactive Constraint Acquisition



Oracle

Unlabelled examples

Labels

CSP Model

Membership query

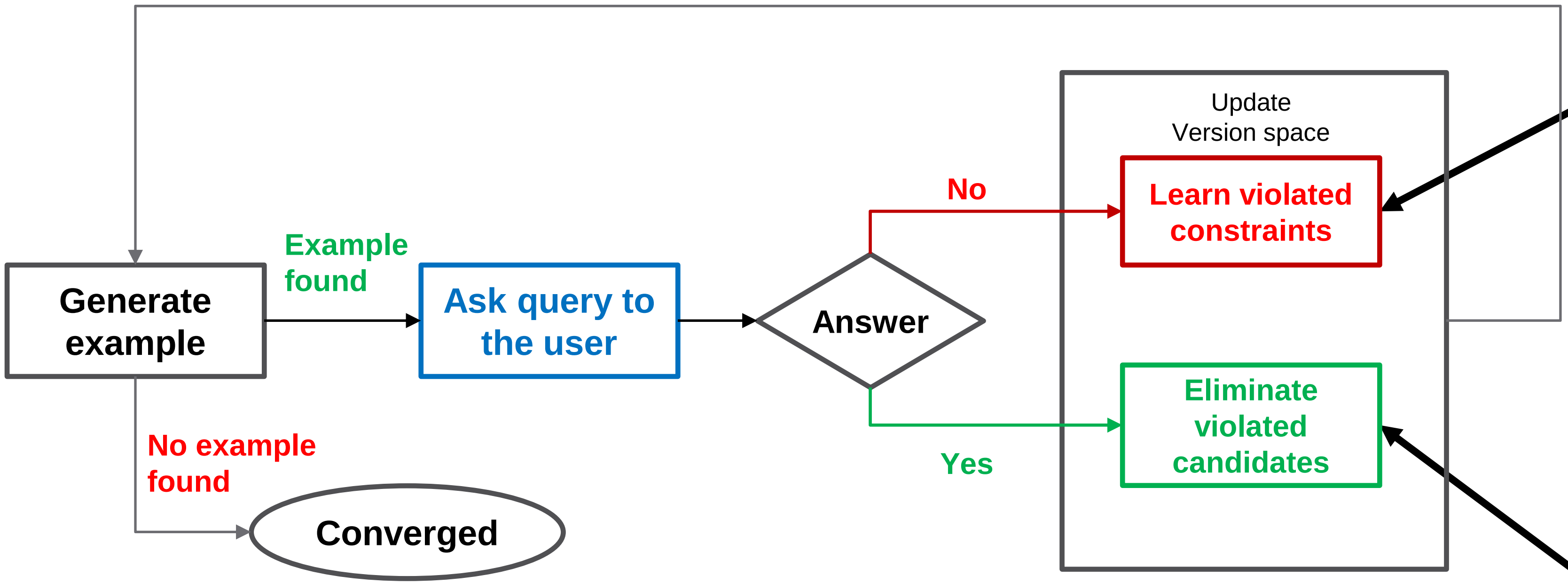
1	1	3	4
3	2	1	1
2	2	3	1
2	3	4	3

Answer: Negative in both of them (a constraint Is violated)

Partial query

1	1	-	4
3	-	1	-
-	-	-	-
2	-	-	-

Interactive Constraint Acquisition



Learn constraints: zoom in violated constraints' scopes using partial queries and add them to C_L

Shrink version space: remove violated constraints from B

Query generation

Irredundant query

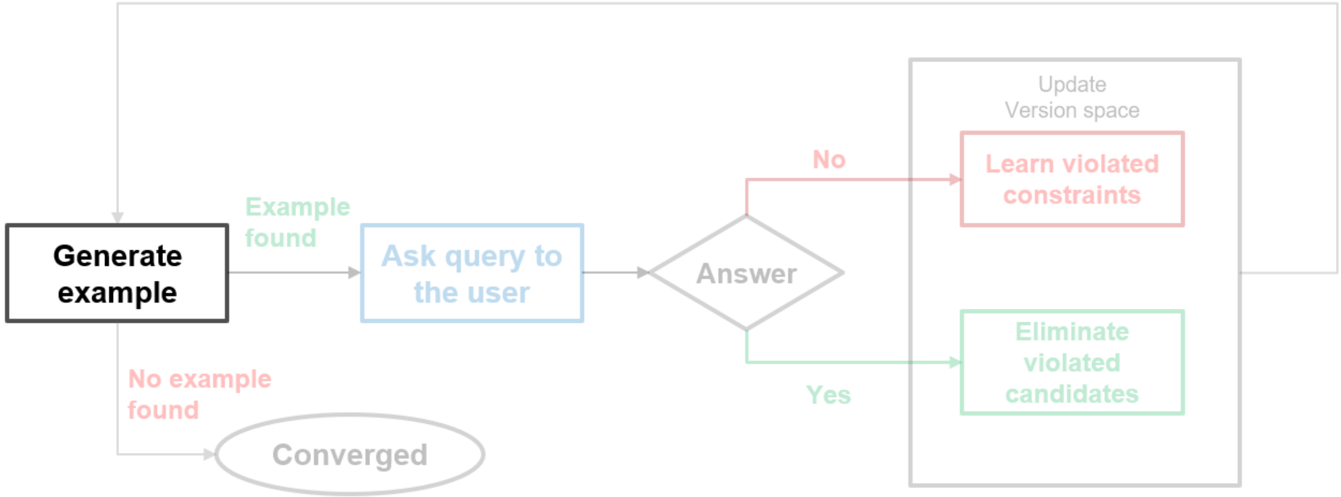
- Not violating any learned constraint
- Violating at least one constraint from *B*

Quality of query

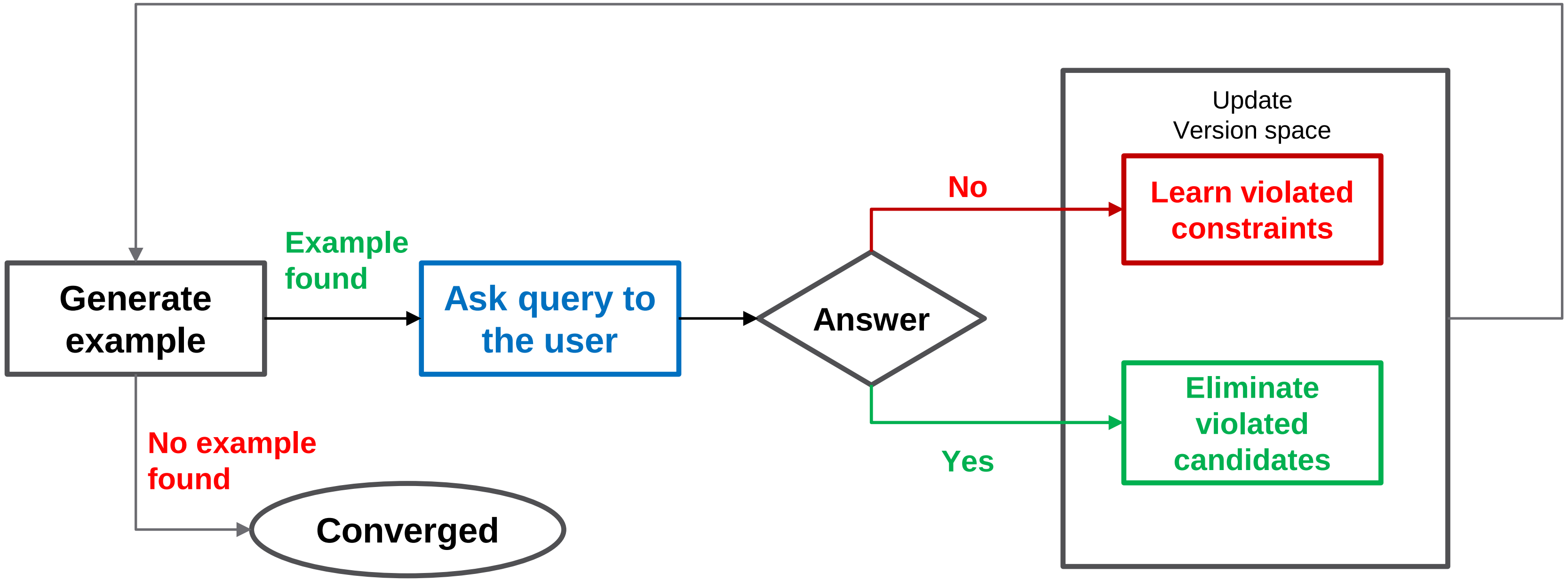
- *Better* generated examples lead to faster convergence
- *Maximizing* violations from B

Convergence

- If no example found



Learning a constraint



Generate example

Example found

Ask query to the user

Answer

No

Yes

Update Version space

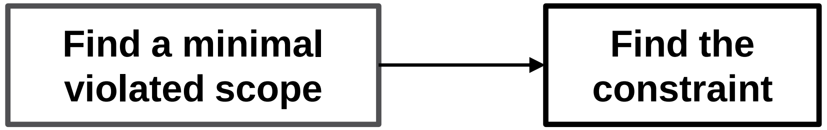
Learn violated constraints

Eliminate violated candidates

No example found

Converged

- 2-step process:



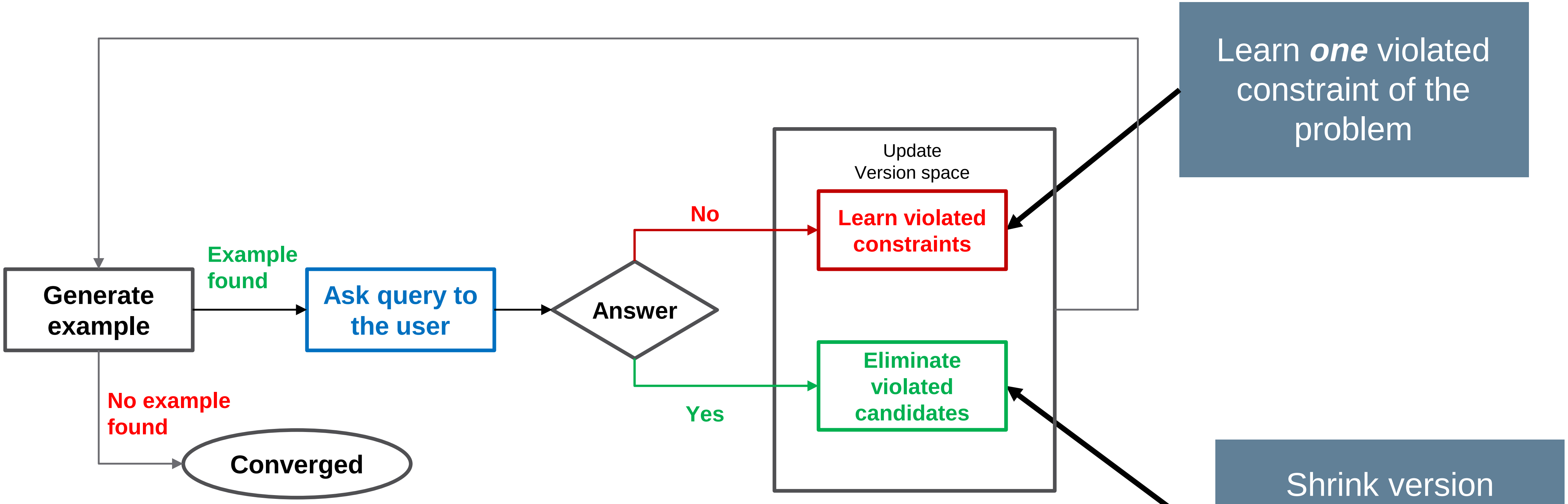
Find a minimal violated scope

Find the constraint

1. *FindScope*: exploit partial (sub)queries to find the problematic part of the assignment
   - logarithmic number of queries
   - splitting variables approximately in half
2. *FindC*: Try different assignments to find the specific constraint in the scope

Interactive Constraint Acquisition

QuAcq



Learn ***one*** violated constraint of the problem

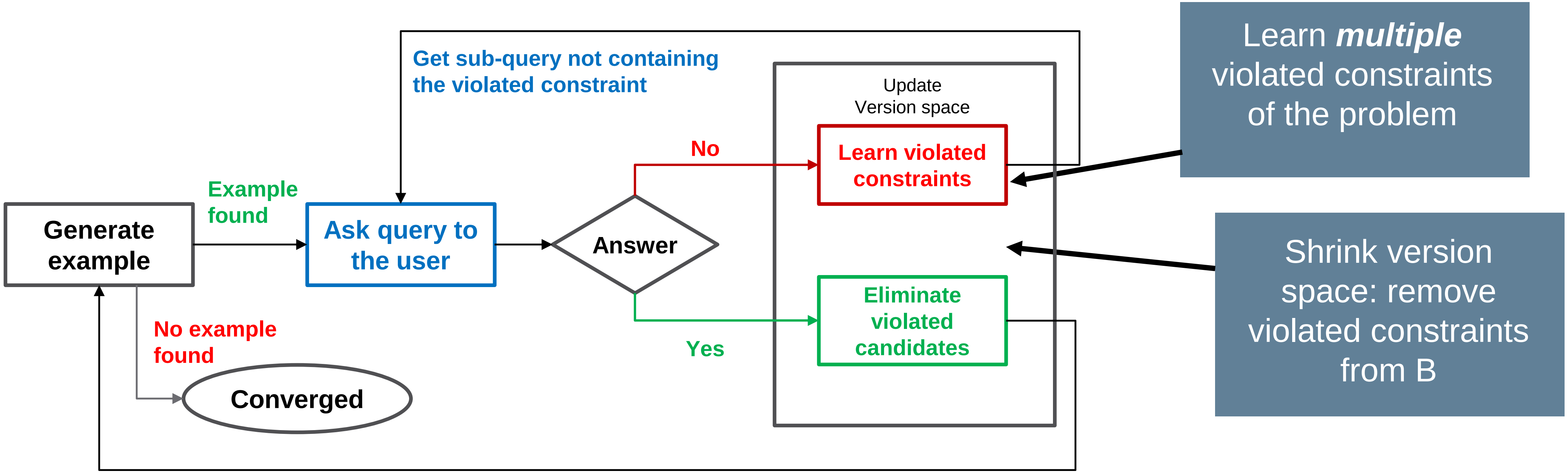
Shrink version space: remove violated constraints from B

QuAcq:

- Learning one violated constraint per generated example
- Logarithmic number of queries for each constraint

Interactive Constraint Acquisition

Multiple Acquisition

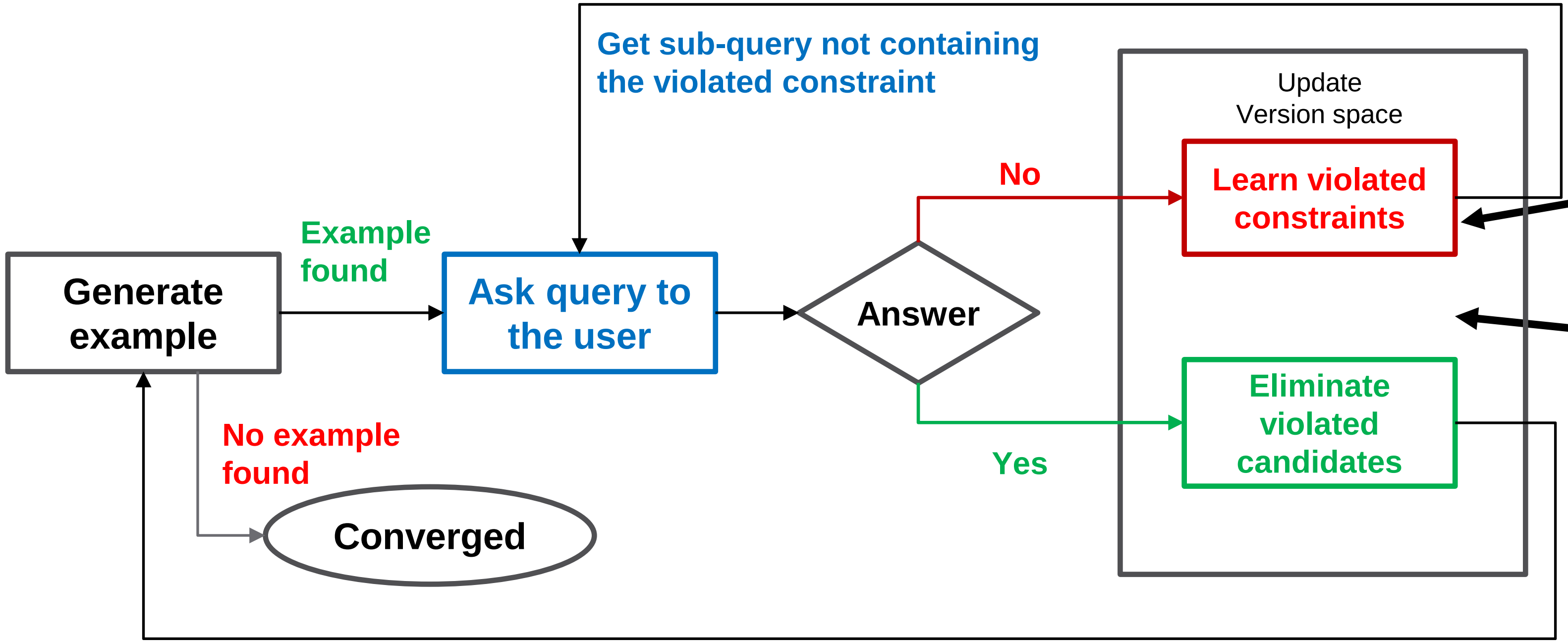

Get sub-query not containing the violated constraint
Update Version space
No
Learn violated constraints
Example found
Generate example
Ask query to the user
Answer
Eliminate violated candidates
Yes
No example found
Converged
Learn *multiple* violated constraints of the problem
Shrink version space: remove violated constraints from B

Multiple Acquisition:

- Learn multiple constraints in each loop instance
- Don’t generate a new example when a constraint is learnt

  - Instead, get an example in a subset of variables not violating the constraint found

Interactive Constraint Acquisition
Multiple Acquisition


Get sub-query not containing the violated constraint
Update Version space
No
Learn violated constraints
Example found
Generate example
Ask query to the user
Answer
Eliminate violated candidates
No example found
Yes
Converged

Learn ***multiple*** violated constraints of the problem

Shrink version space: remove violated constraints from B

MultiAcq:

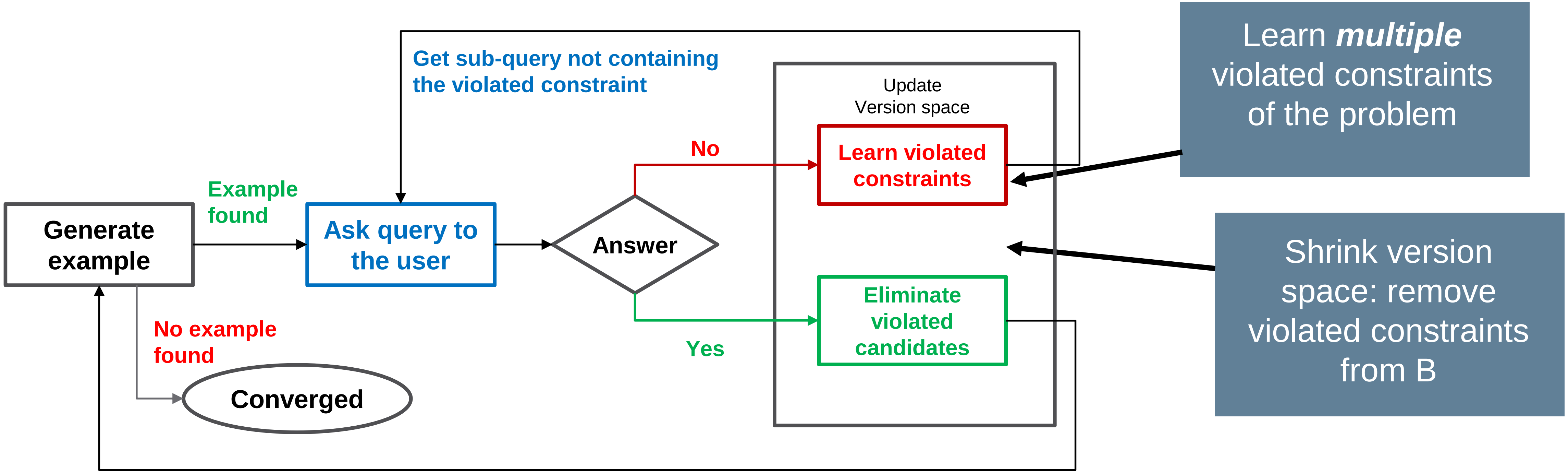
- Learning *all* violated constraints per generated example
- linear number of queries for each constraint

MQuAcq:

- Learning *all* violated constraints per generated example
- Logarithmic number of queries for each constraint
- Using FindScope/FindC repeatedly until no more constraints are found

Interactive Constraint Acquisition

Multiple Acquisition



Generate example
Example found
Ask query to the user
Get sub-query not containing the violated constraint
Answer
No
Yes
Update Version space
Learn violated constraints
Eliminate violated candidates
No example found
Converged

Learn ***multiple*** violated constraints of the problem

Shrink version space: remove violated constraints from B

MQuAcq-2:

- Learning *multiple (not all)* violated constraints per generated example
- Logarithmic number of queries for each constraint
- Avoids the extensive branching needed by MQuAcq and MultiAcq
- Focus the queries on promising parts

  - detecting quasi cliques

Open challenges

Challenges

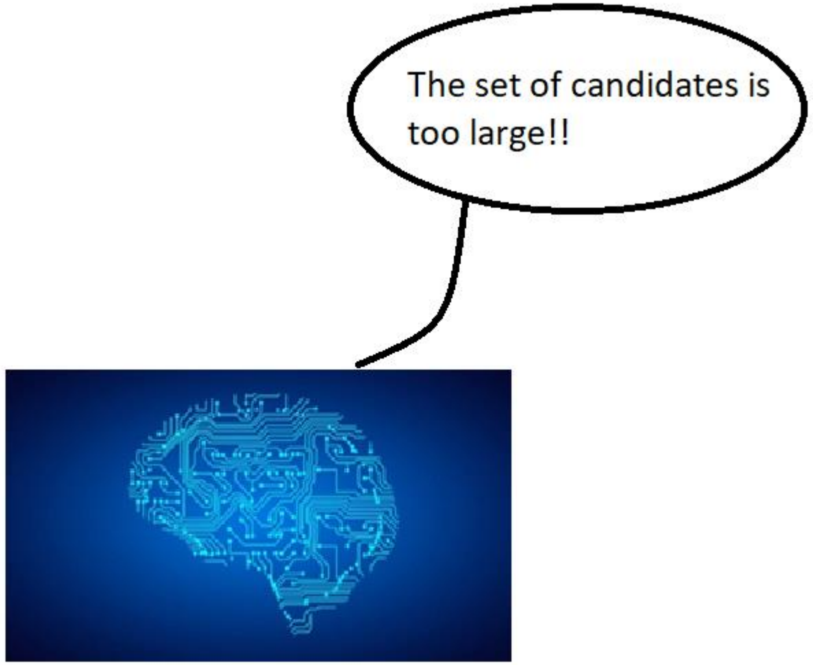
Number of queries

- Number of queries needed to converge is still large.
- Query generation and the acquisition process are highly unguided
- Information from what we have learnt can be used

Challenges

Application level constraint modelling

- Handling of big sets of candidate constraints
- Alleviating the requirement for a more specific knowledge of the constraints that can be present in the problem



Challenges

Specific classes of constraints

- **Global constraints:** Exploding the set of candidate constraints
- **Linear inequalities** with constants: Need to consider all possible constants -> Exploding the set of candidate constraints

COUNT

ALLDIFFERENT

CUMULATIVE

SUM

CIRCUIT

$x_1 + 5 < x_2$

$|x_1 + 12| > x_4$

$x_1 - x_2 != 238$

Challenges

Noisy data

- unlike in machine learning, most constraint acquisition techniques still assume the user always (knows how to) answer correctly
- Tighter integration with modern machine learning techniques

KU LEUVEN

Thank you for your attention

Open for discussion and brainstorming

Learning constraint models from data: http://osullivan.ucc.ie/CPML2023/submissions/01.pdf
Efficient Multiple Constraint Acquisition: https://arxiv.org/abs/2109.05920