Learning Parameters For the Sequence Constraint From Solutions

Émilie Picard-Cantin Mathieu Bouchard Claude-Guy Quimper Jason Sweeney

Abstract

This paper studies the problem of learning parameters for global constraints such as **Sequence** from a small set of positive examples. The proposed technique computes the probability of observing a given constraint in a random solution. This probability is used to select the more likely constraint in a list of candidates. The learning method can be applied to both soft and hard constraints.

Context

- We work in collaboration with **PetalMD**, an expert in medical scheduling.
- We are given schedules and our goal is to learn the limit constraints that were applied in order to create a new schedule.

Here is the schedule of an employee. Can you guess the constraints that created it ?

Is it :

- One in three days ?
- Two in five days ?
- Three of (,) in four days ?

Sequence Constraint

- **Among**$(\ell, u, [x_j, \dots, x_{j+k-1}], V)$
  Ensures that variables $x_1, \dots, x_k$ are assigned to values in V at least ℓ and at most u times.
- **Sequence**$(\ell, u, k, [x_1, \dots, x_d], V)$
  Sliding of **Among**$(\ell, u, [x_j, \dots, x_{j+k-1}], V)$ over all subsequences of k consecutive variables.

Problem Description

Why ?

- Learning the set of **parameters** (ℓ, u, k, V) of **Sequence**.

Why ?

- **Sequence** is one of the most common constraint.
- Mostly depict team preferences.
- Clients express their constraints informally.

How ?

- Statistical algorithm that, from a small set of positive examples, ranks all satisfied sets of parameters by increasing probability of being observed.

Methodology

Input

- A small set of positive examples given by a client
- The scoped variables $x_1, \dots, x_d$
- The probability of assignment $x_i = v$, noted p_v

Steps

- List all sets of parameters satisfied by the given examples (candidates).
- Rank the sets of parameters according to the statistical analysis.

Output

- Set of parameters (ℓ, u, k, V) describing the chosen **Sequence**.

Individual Probability

The individual probability of observing a set of parameters is the sum of probabilities of all its solutions. In the figure below, each dot represent a solution. The bigger the dot, the higher the probability of observing the solution.

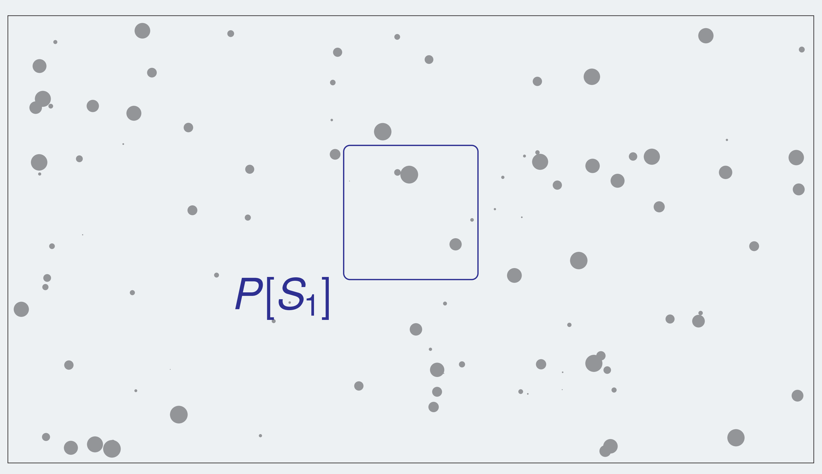
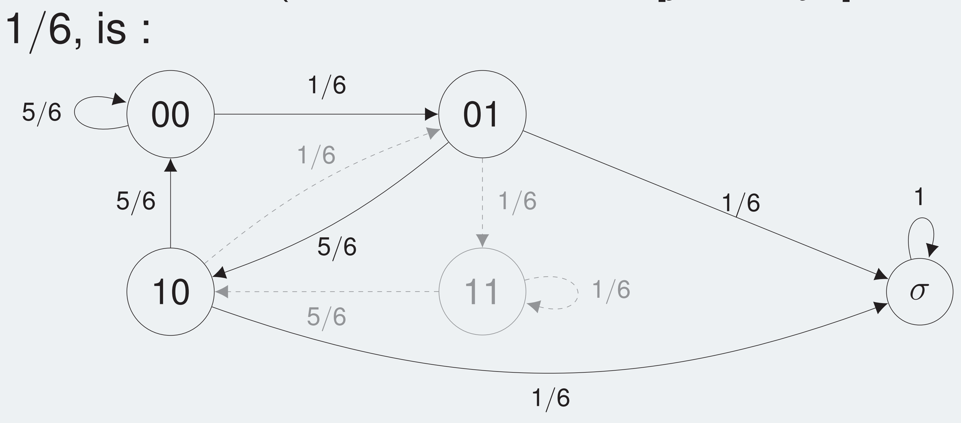


Figure: Probability of observing a set of parameter S_1.

Markov Chains

The Markov chain for **Sequence**$(\ell = 0, u = 1, k = 3, [y_1, \dots, y_d], V = \{1\})$, when $p_0 = 5/6$ and $p_1 = 1/6$, is :



Gray transitions and node are absorbed in the rejection state because they violate the given constraint.

Probability Algorithms

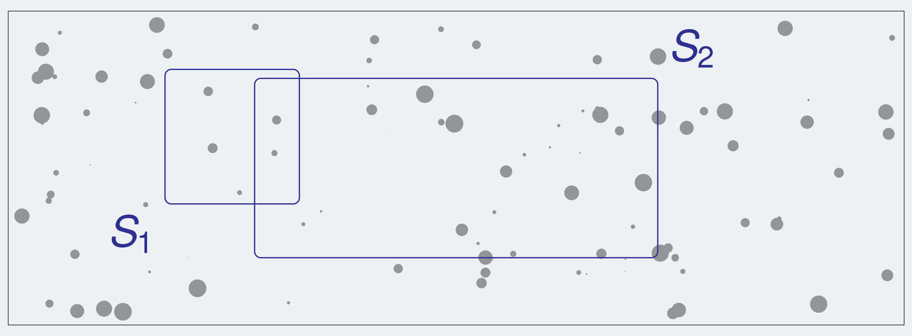
We propose three different algorithms to compute the probability of a given set of parameters for **Sequence**$(\ell, u, k, [x_1, \dots, x_d], V)$.
Let α be the vector of initial probabilities and P be the matrix of one step probabilities associated with the Markov chain.

Method	αP^n	Complexity
Adaptation of Zanarini & Pesant	$((((\alpha P)P)\dots)P)$	$O(d2^k)$
Spectral Decomposition	$\alpha(V^{-1}D^nV)$	$O(8^k)$
Decrease & Conquer	$\alpha(P^{n/2})^{1/2}$	$O(2^{\omega k}\log(d-k))$

Note : $O(n^\omega)$ is the complexity of multiplying two matrices.

Constraint Ranking

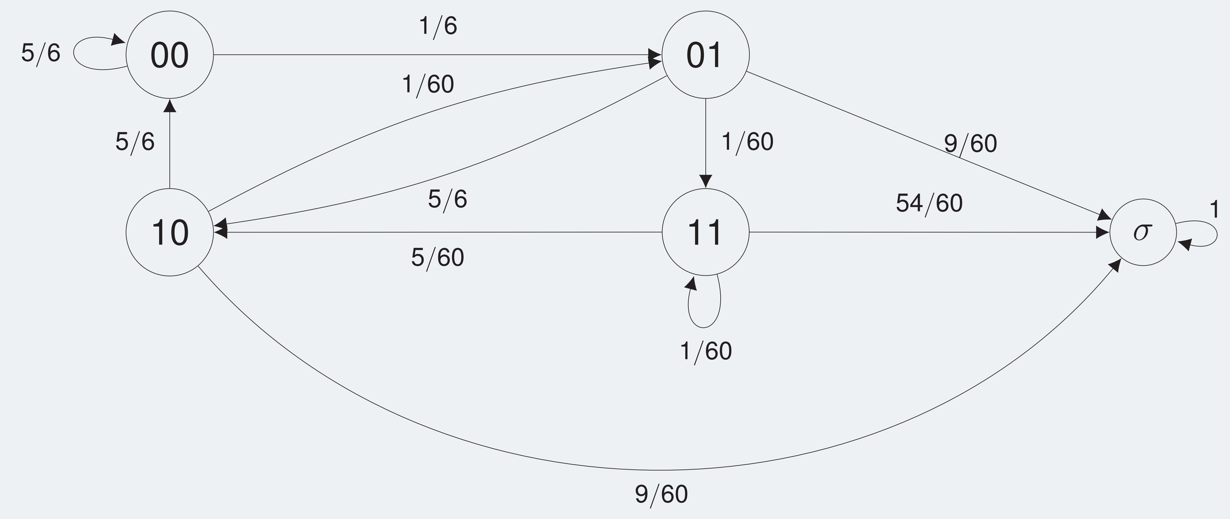
- $P[S_1 \wedge S_2 \mid S_1]$ is the probability of observing both sets of parameters knowing we observed the first set.
- The best choice is the constraint that has the lowest individual probability of being observed in a random solution.



$$P[S_1 \wedge S_2 \mid S_1] > P[S_1 \wedge S_2 \mid S_2]$$
$$\frac{P[S_1 \wedge S_2]}{P[S_1]} > \frac{P[S_1 \wedge S_2]}{P[S_2]}$$
$$P[S_1] < P[S_2]$$

Soft Constraints

The Markov chain for the soft constraint **Sequence**$(\ell = 0, u = 1, k = 3, [y_1, \dots, y_d], V = \{1\})$, when $p_0 = 5/6$ and $p_1 = 1/6$ and when we accept $1/10$ of the violations, is :



Experiments

Table: Task oriented

	Days	
	1	2	3
T_1	A, B	A	
T_2		B	A

GCC

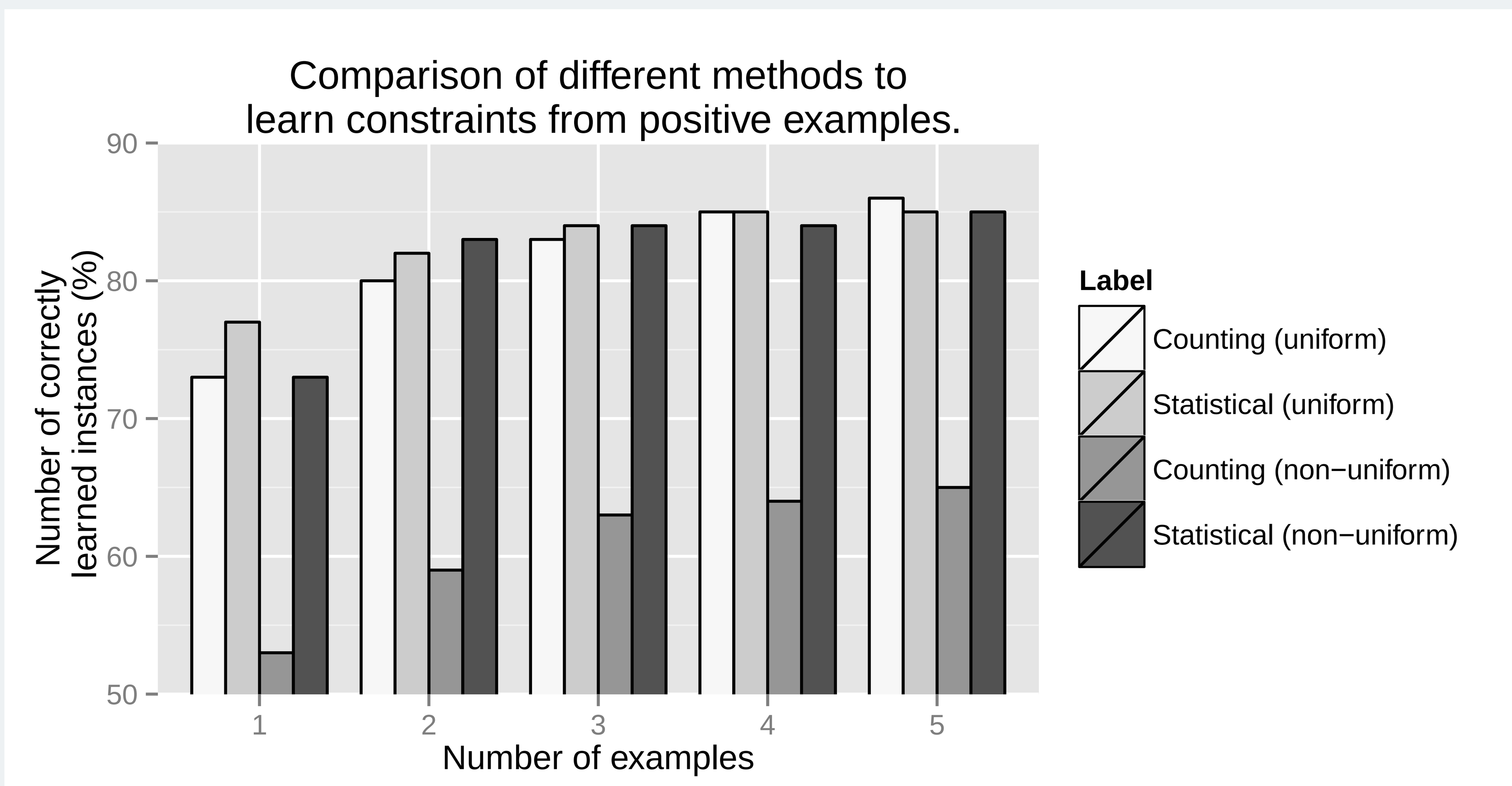
Table: Employee oriented

	Days		
	1	2	3
A	T_1	T_1	T_2
B	T_1	T_2	

One task per day, per employee **Sequence**

Table: Sets of instances

Uniformly distributed tasks	Non-Uniformly distributed tasks
Basic Employee subset Task subset	Basic Employee subset Task subset



Contribution

- Three algorithms to compute the probability of observing a given set of parameters for **Sequence**.
- Improvement on solution counting for the **Regular** constraint using a simplified automaton and a matrix representation.
- Machine learning tool that can be applied to both **soft and hard** global constraints that can be formulated as an automaton, such as **Sequence**, **Among Knapsack**, **Stretch**, etc.
- Requires **less positive examples** to achieve the same results as other methods for instances where values are uniformly distributed.
- Largely better than Counting for instances with **non-uniformly distributed values**.

Emilie Picard-Cantin, PhD student emilie.picard-cantin.1@ulaval.ca