Using Predictive Clustering and Probabilistic Constraint Solving for Structural Predictions

Kurt Driessens and Hendrik Blockeel

KU LEUVEN

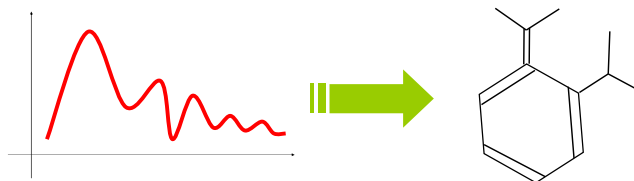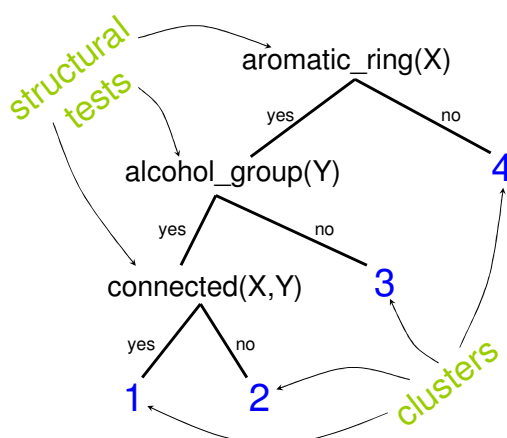**DTAI** Declaratieve Talen & Artificiële Intelligentie

Predicting Structured Output from Unstructured Input:

• Single Ended Line Testing: Discovering topologies of buried telephone lines from signal echo
• Molecular Structure from mass-spectrometry
• Protein Folding from Amino Acid Sequences
• …

Predictive (relational) Clustering Trees:

structural tests

aromatic_ring(X)
yes — no
alcohol_group(Y) — 4
yes — no
connected(X,Y) — 3
yes — no
1 — 2

clusters

Concept:
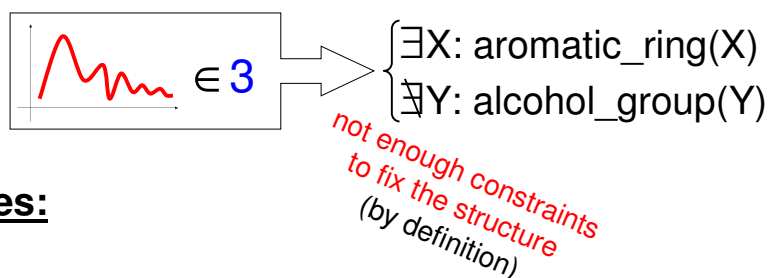
Clustering but using a language bias

Implementation:

TDIDT but using a clustering heuristic

Matching a new example with a cluster:

$\in 3$ → { $\exists X$: aromatic_ring(X), $\exists Y$: alcohol_group(Y)

not enough constraints to fix the structure (by definition)

Using a Forest of Clustering Trees:

1
2   3
$\in 3$ → Set of constraints

3
1   2
$\in 2$ → Set of constraints

1   2 3   4
$\in 3$ → Set of constraints
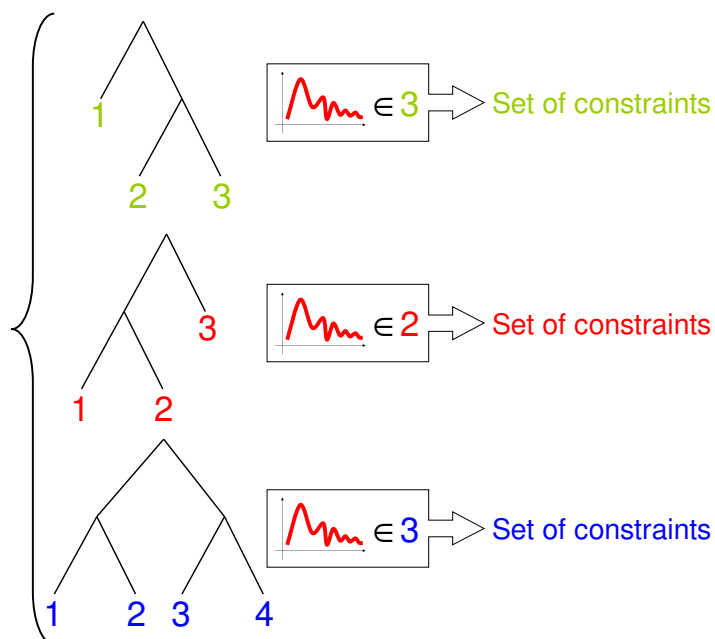
Generate multiple sets of constraints:

• Different sets of constraints can contain contradictions
• Distance to cluster translates to probability for constraints
• Leaves a set of probabilistic, relational constraints to be solved.

To what extend can SRL contribute to this field?

Are there concrete ideas on how to use existing SRL techniques for this?