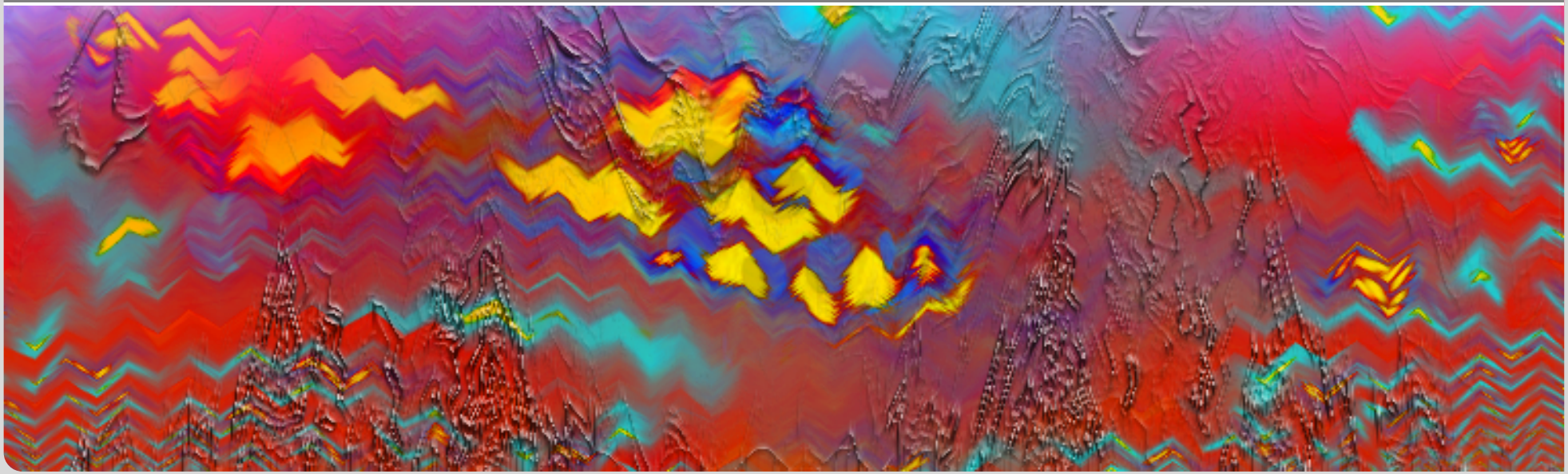
Karlsruhe Institute of Technology

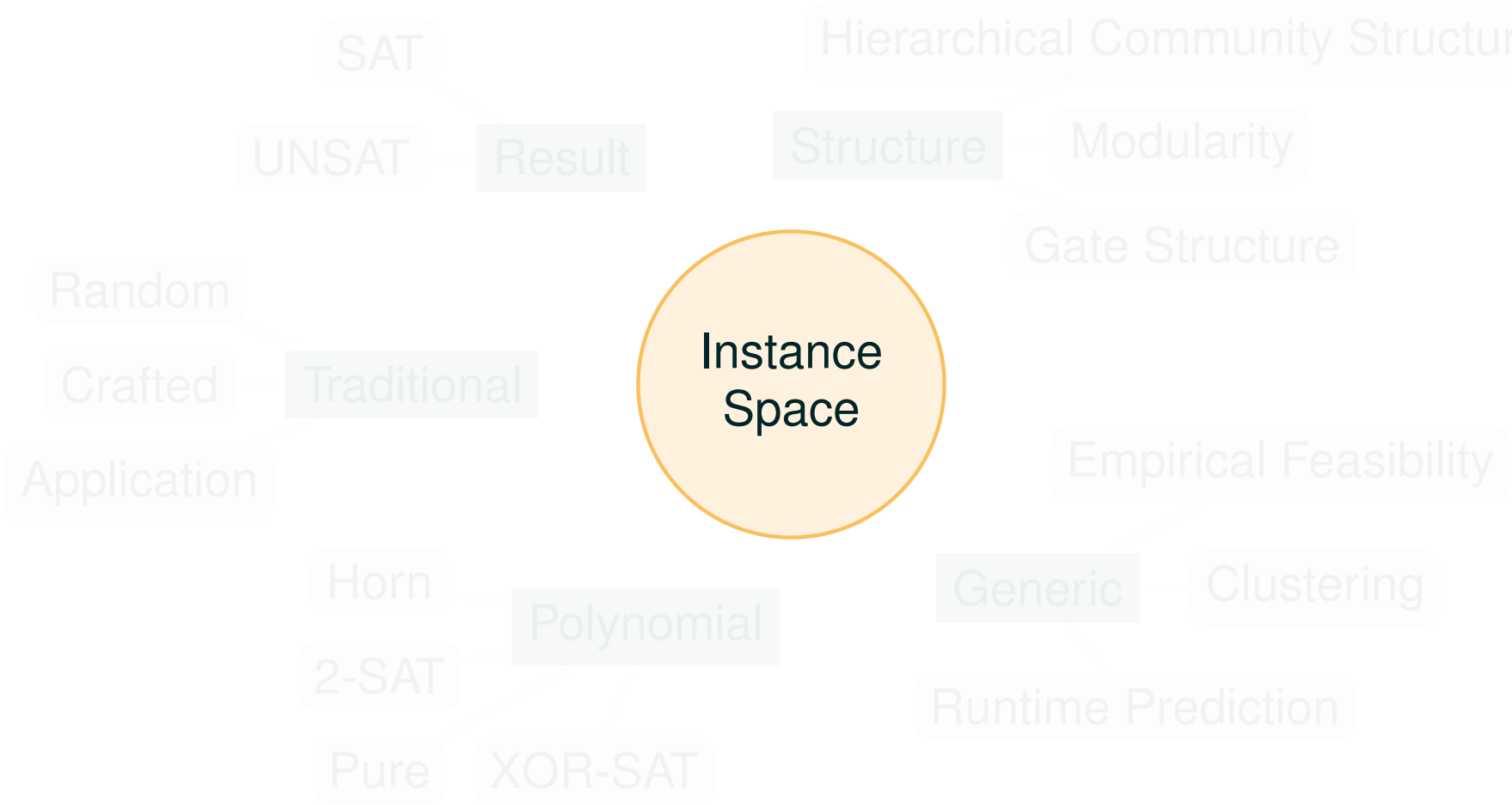
Beyond Portfolios of Stable Solvers

Markus Iser

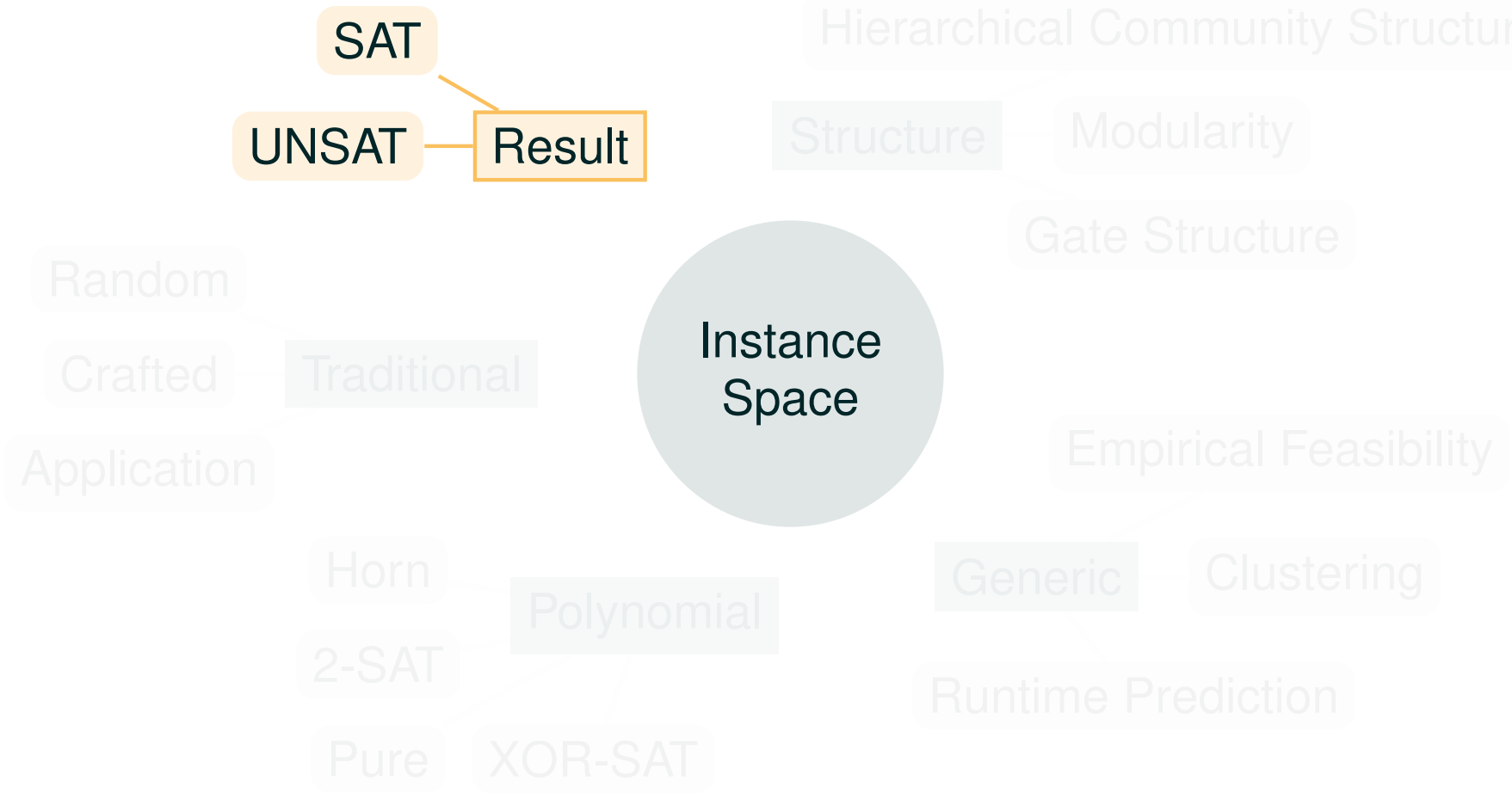
Algorithm Engineering · Institute of Theoretical Informatics



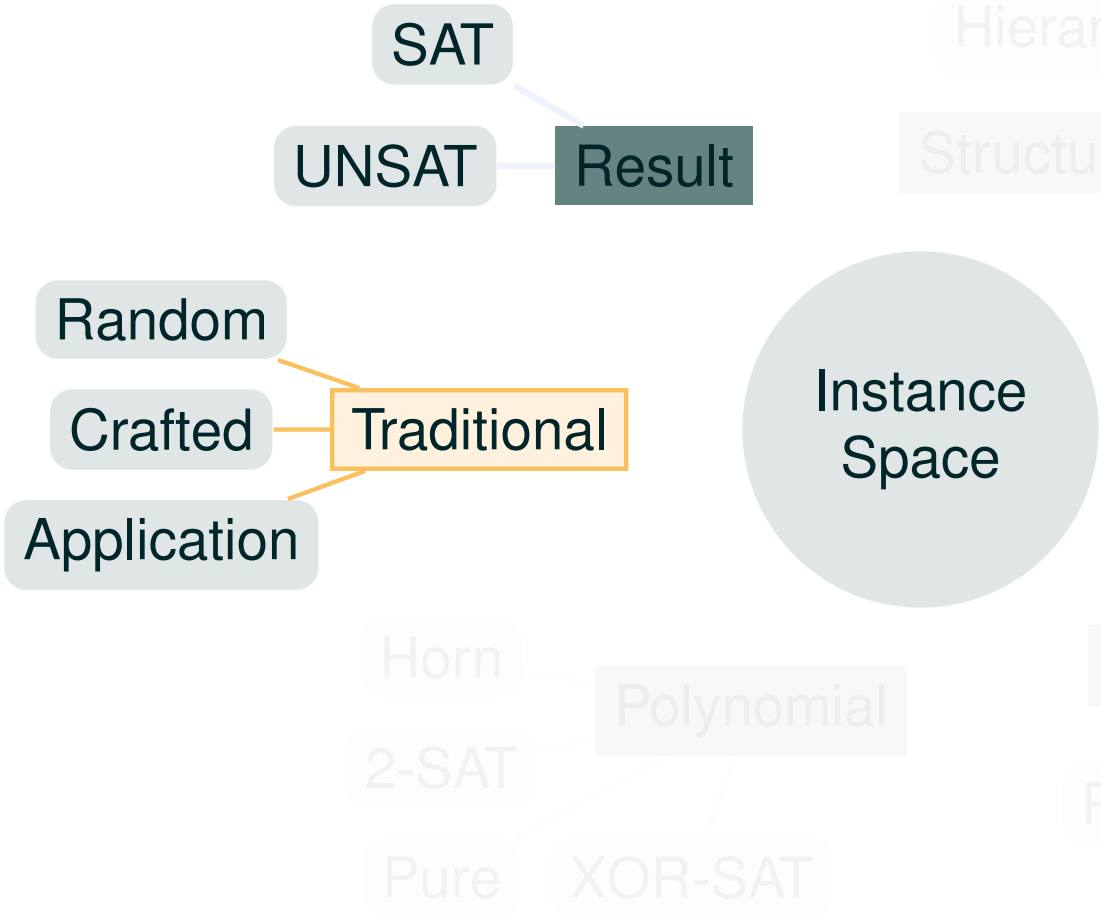
Categories of SAT Instances



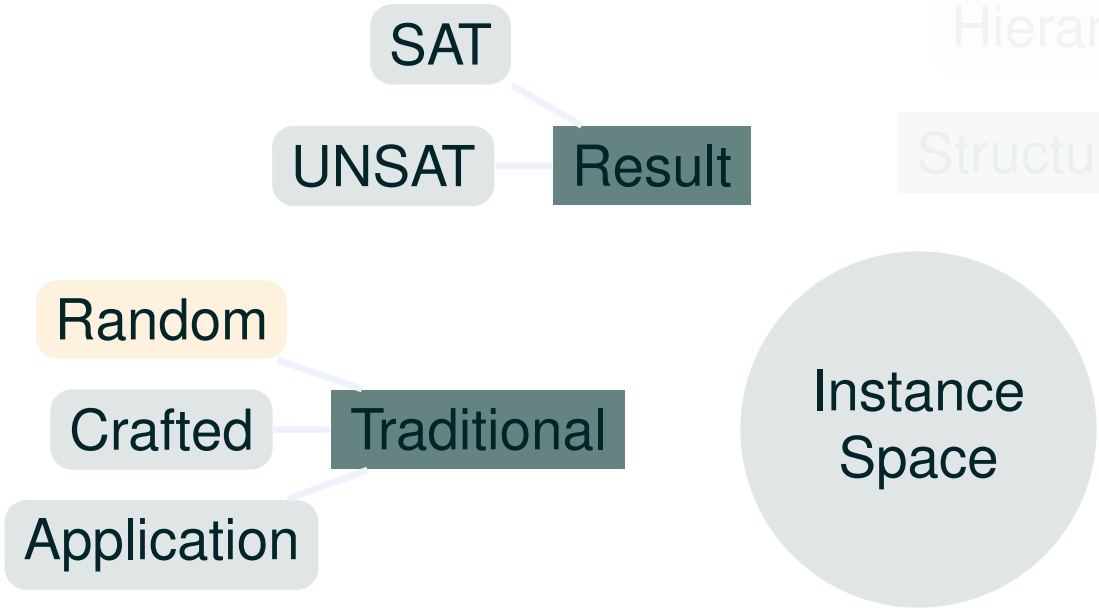
Categories of SAT Instances



Categories of SAT Instances



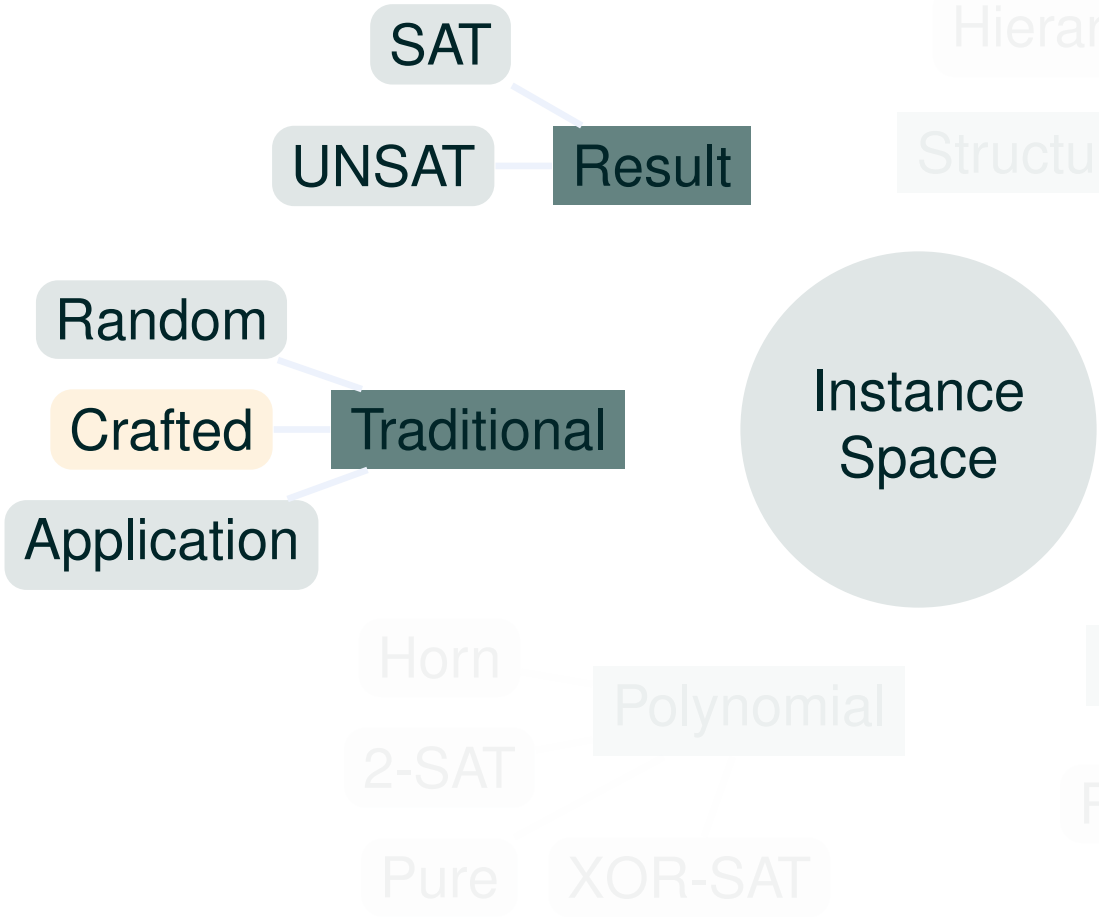
Categories of SAT Instances



Random

- Uniform Random
- Planted Solutions
- Structured Random
- ...

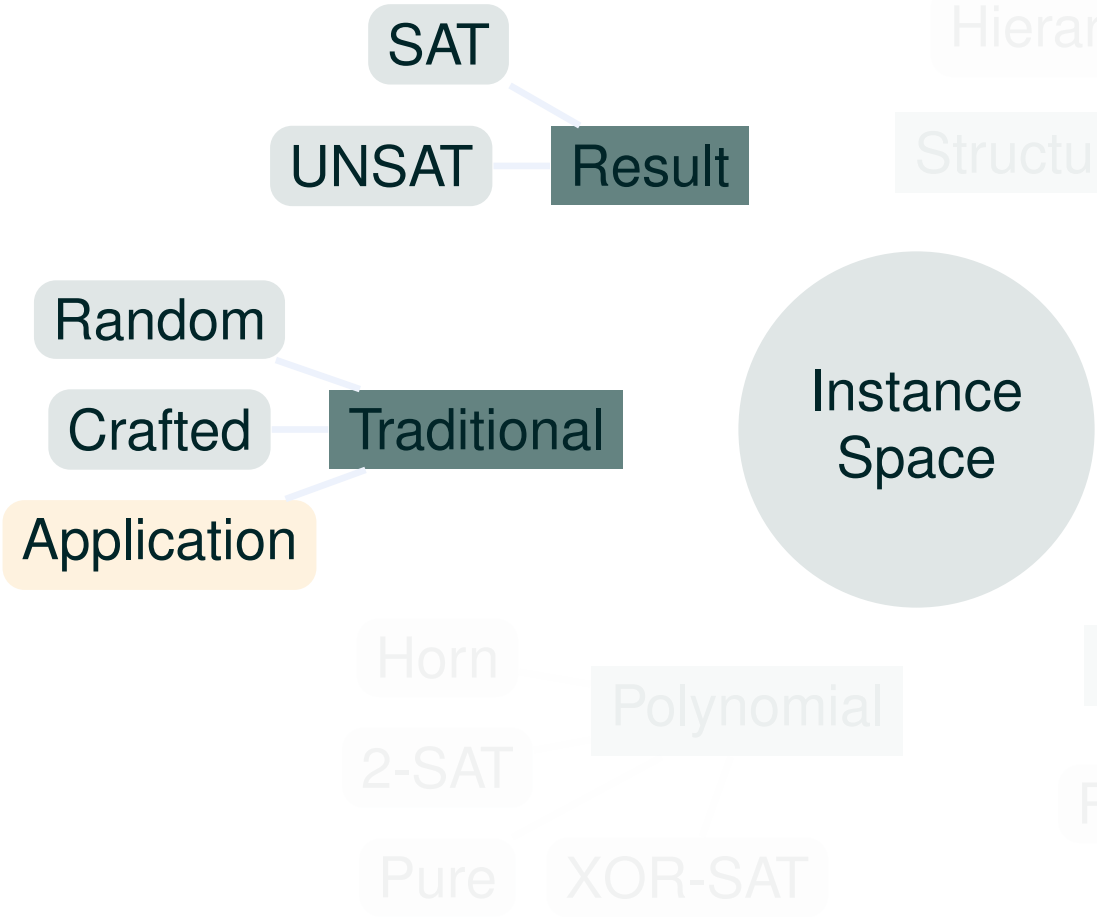
Categories of SAT Instances



Crafted

- Pigeon Hole
- Mutilated Chessboard
- Tseitin Formulas
- ...

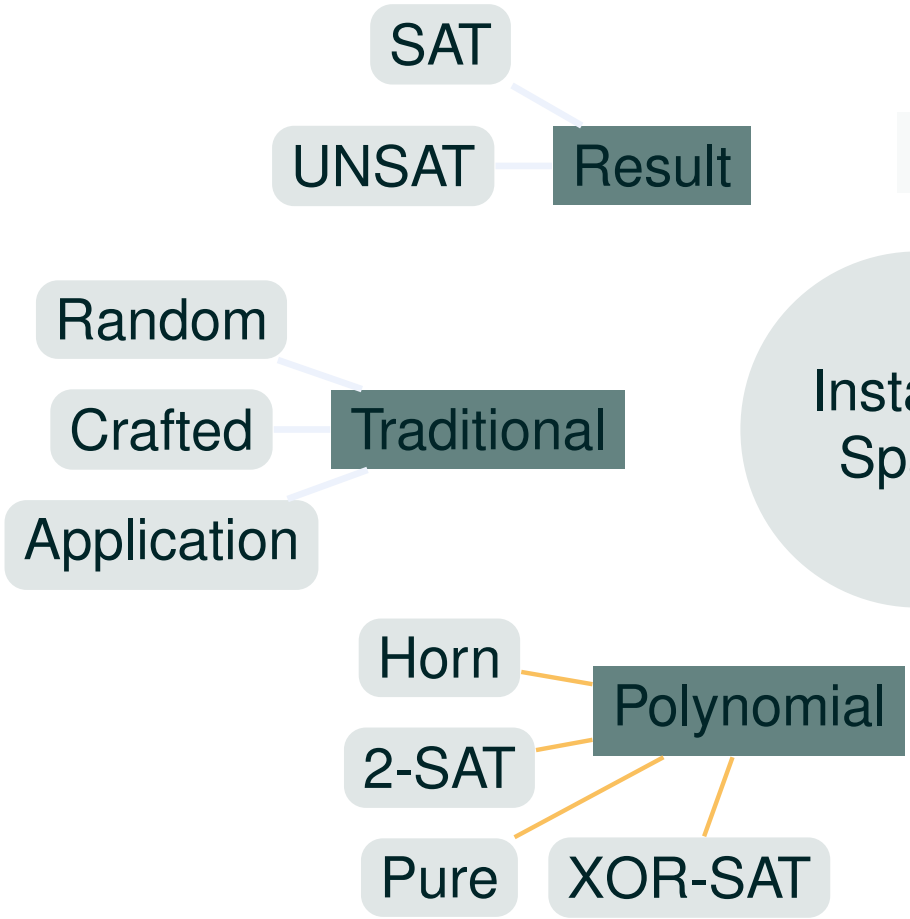
Categories of SAT Instances



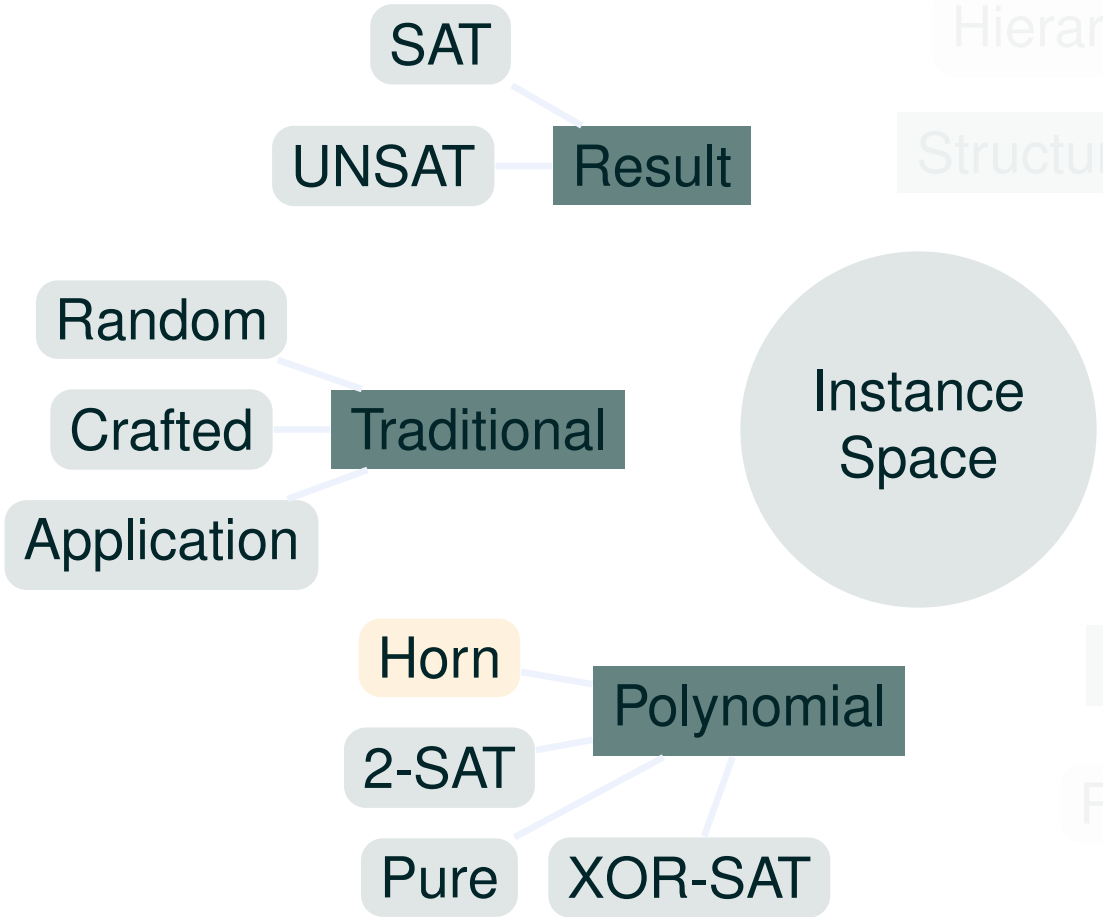
Application

- Planning
- Hardware Verification
- Product Configuration
- Cryptography
- …

Categories of SAT Instances



Categories of SAT Instances



Horn

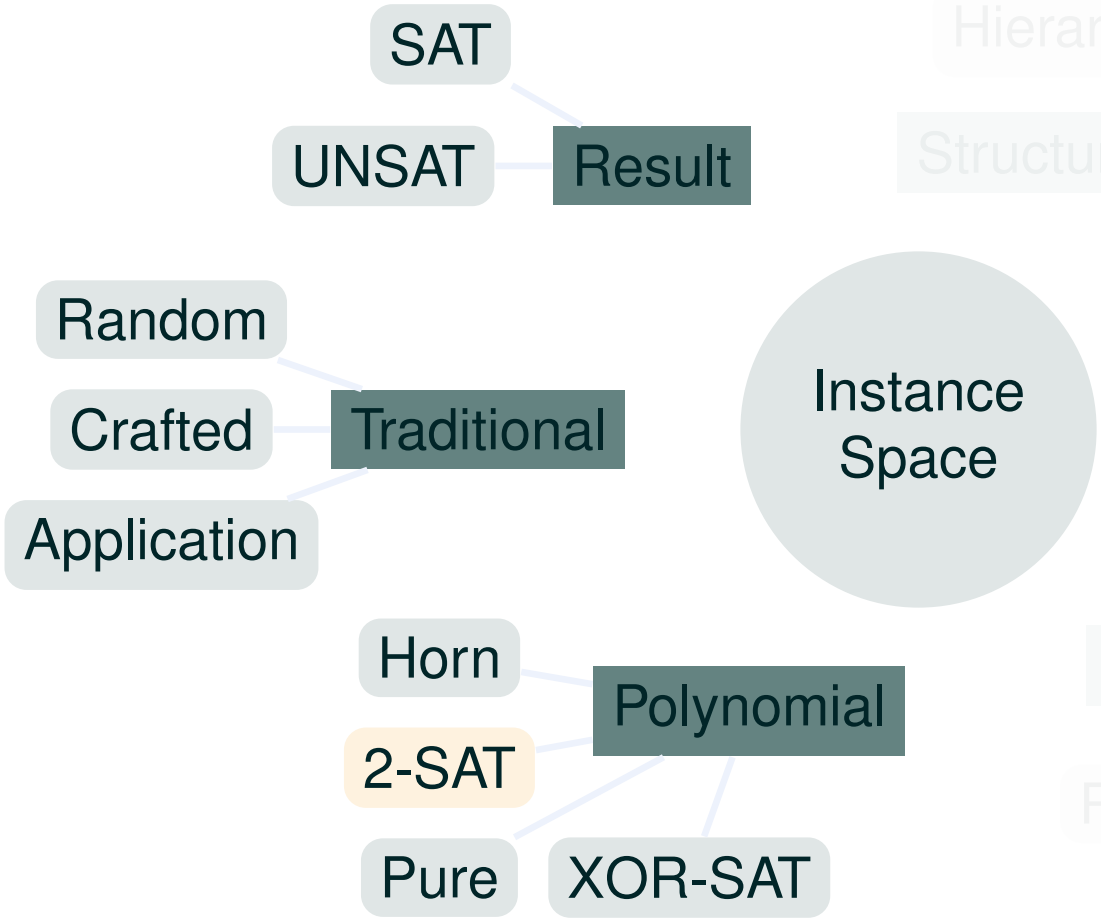
- At most one *negative* literal per clause

Inverted Horn

- At most one *positive* literal per clause

$\Rightarrow$ Unit-propagation $\in \mathcal{O}(n)$

Categories of SAT Instances

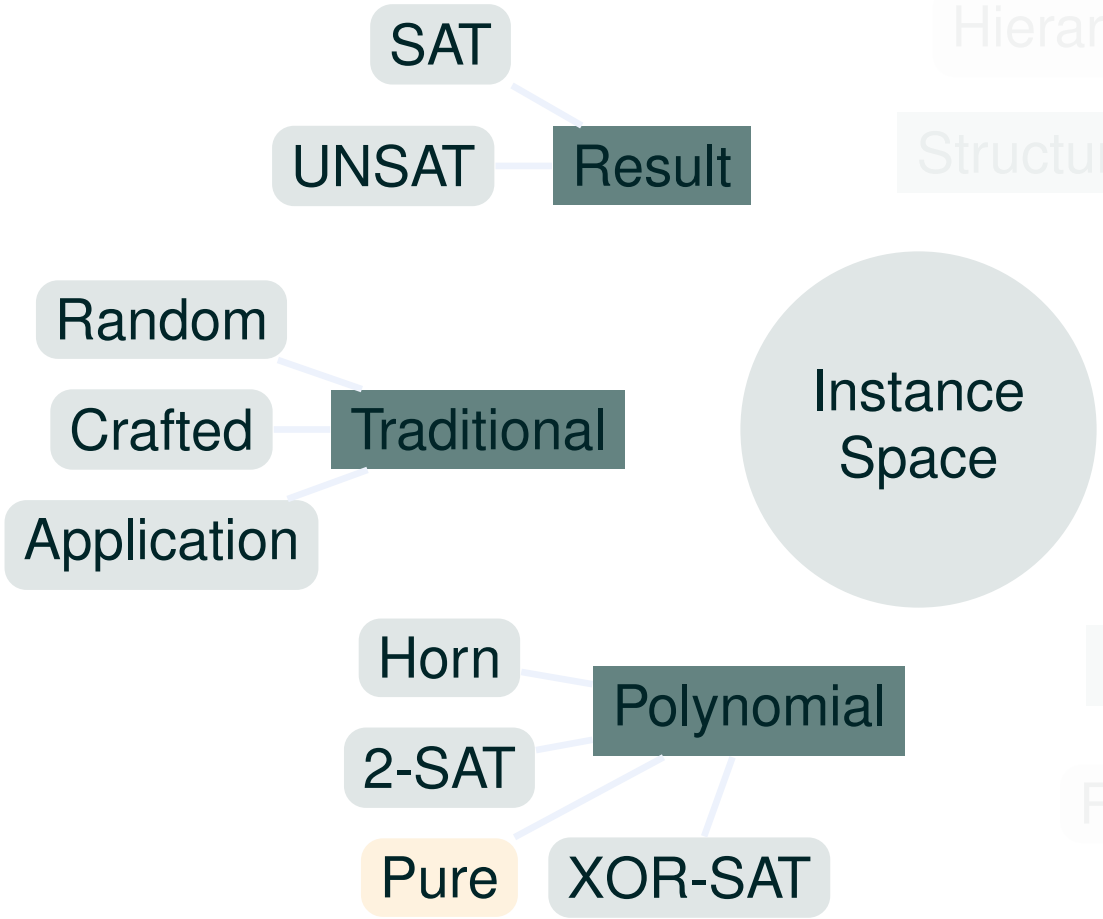


2-SAT

- All clauses have length two

$\Rightarrow$ determine connected components in implication graph $\in \mathcal{O}(n+m)$

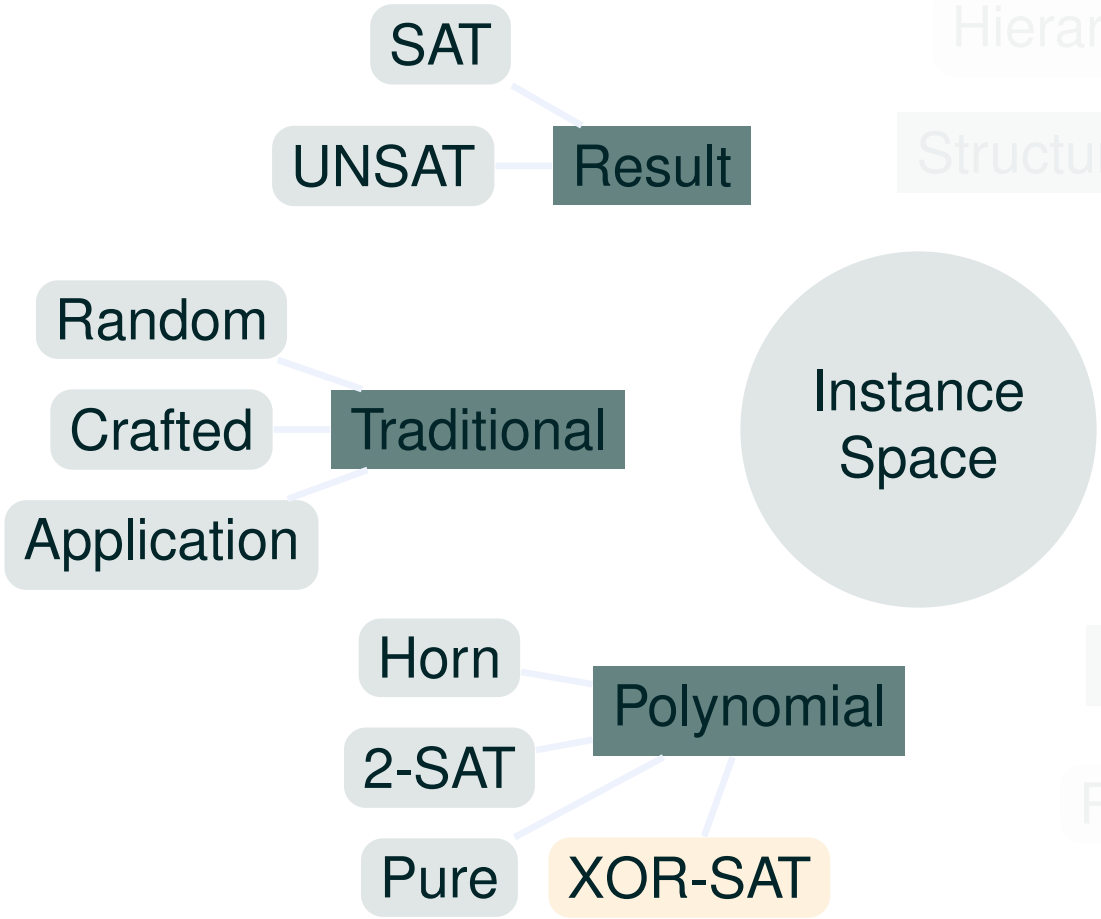
Categories of SAT Instances



Pure

- all literals are positive
- or all literals are negative

Categories of SAT Instances



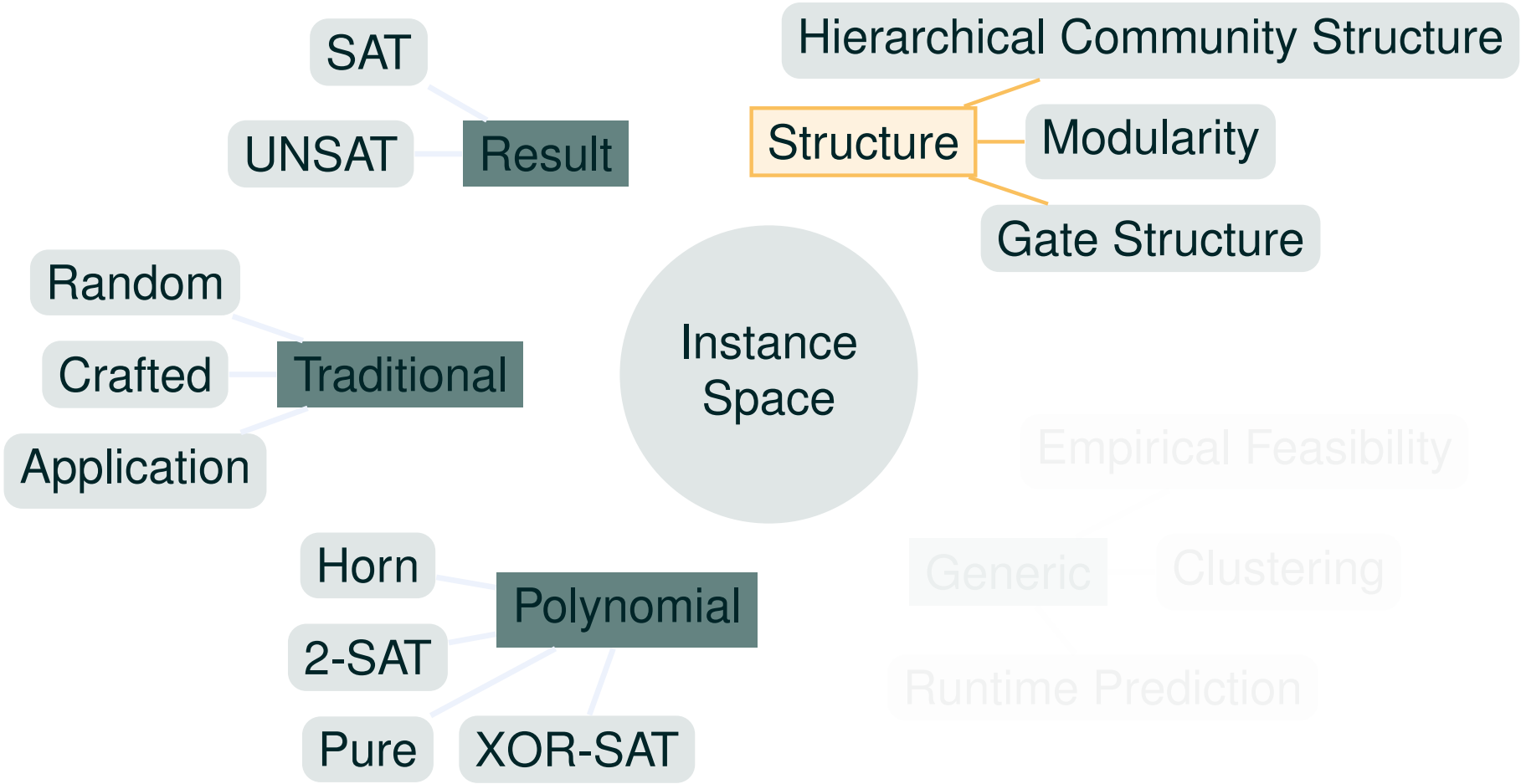
XOR-SAT

- only xors
- aka. affine formulas

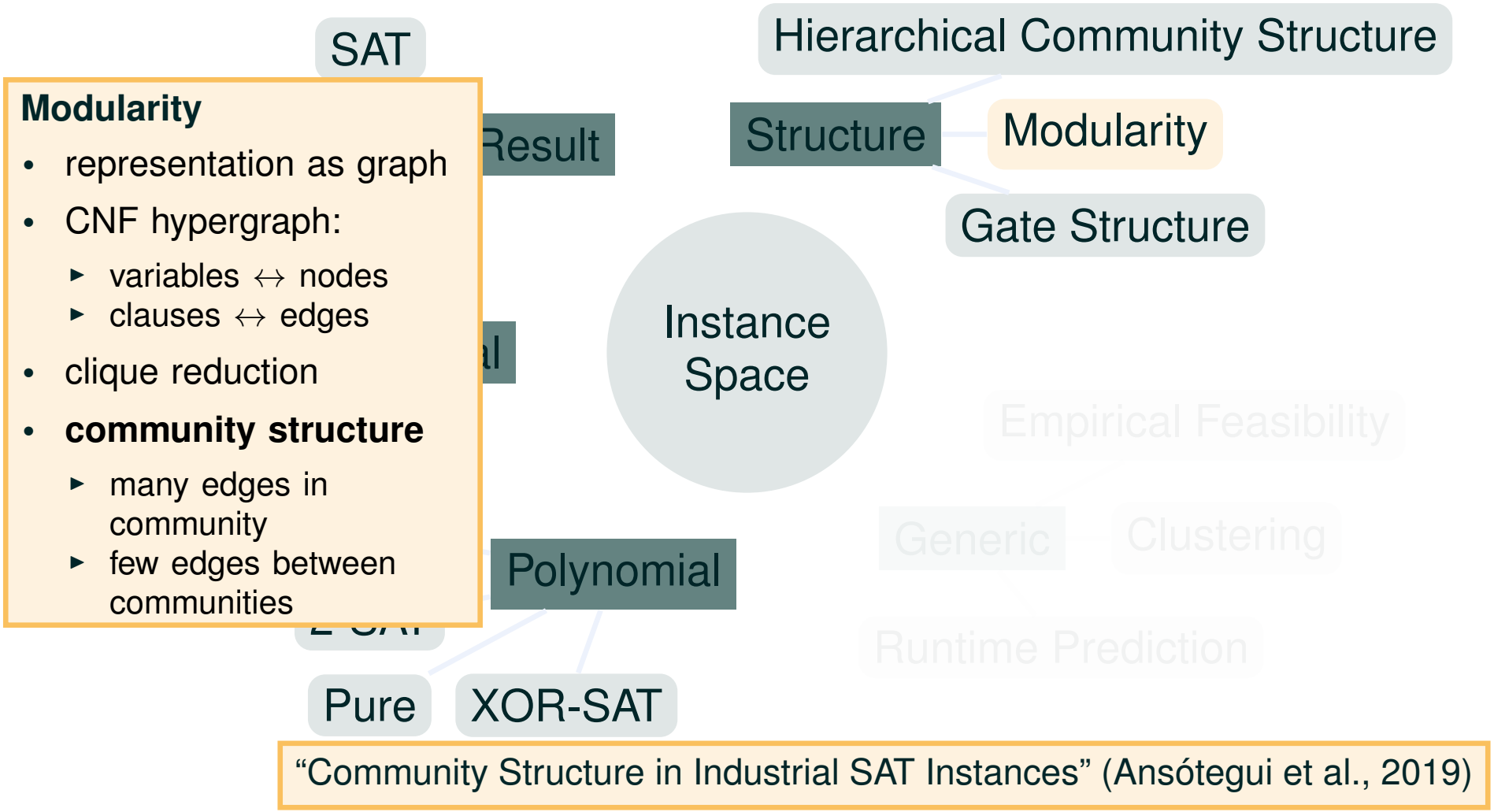
$\Rightarrow$ Gaussian elimination $\in \mathcal{O}(n^3)$

"The Complexity of Satisfiability Problems" (Schaefer, 1978)

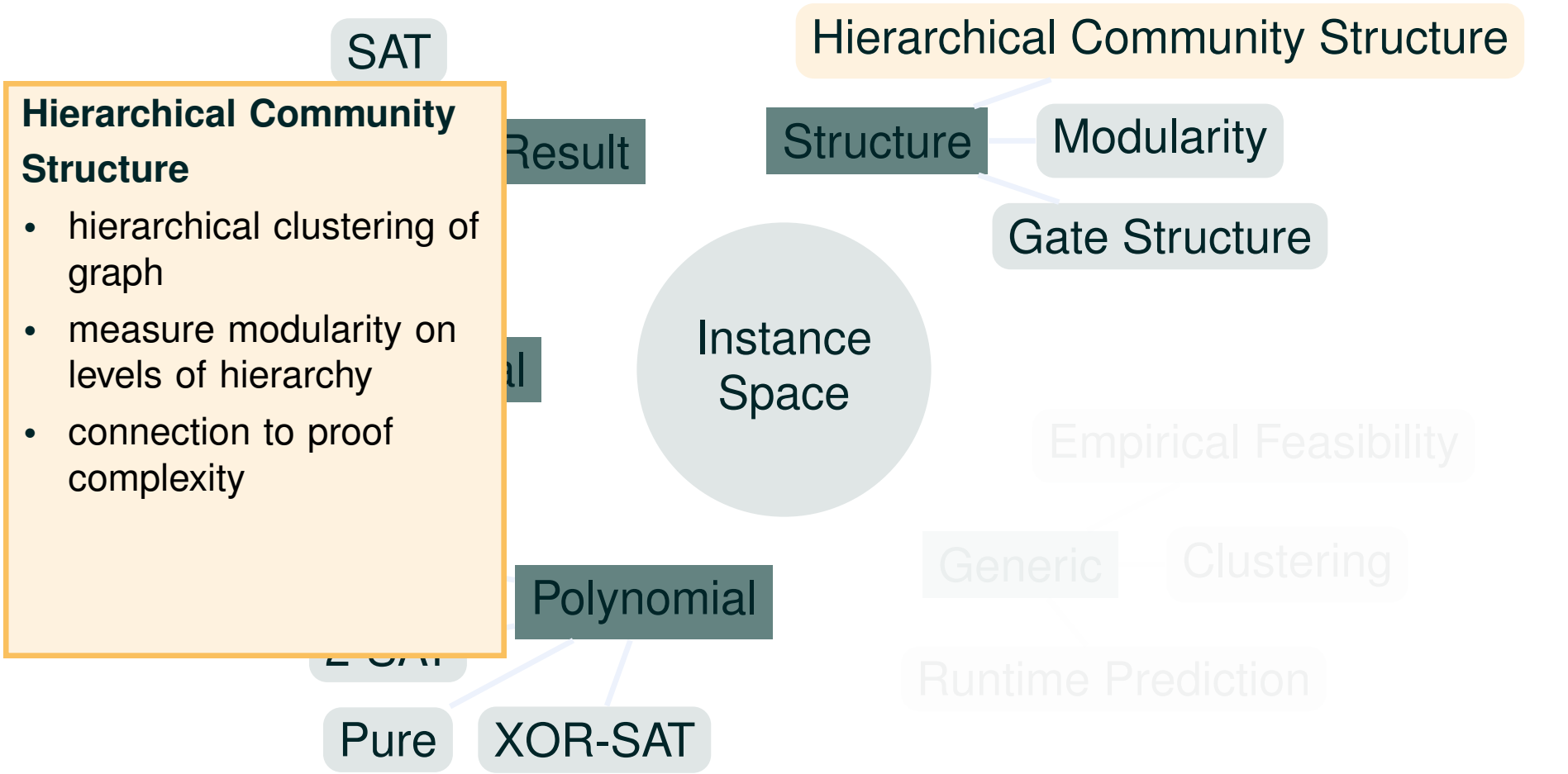
Categories of SAT Instances



Categories of SAT Instances



Categories of SAT Instances

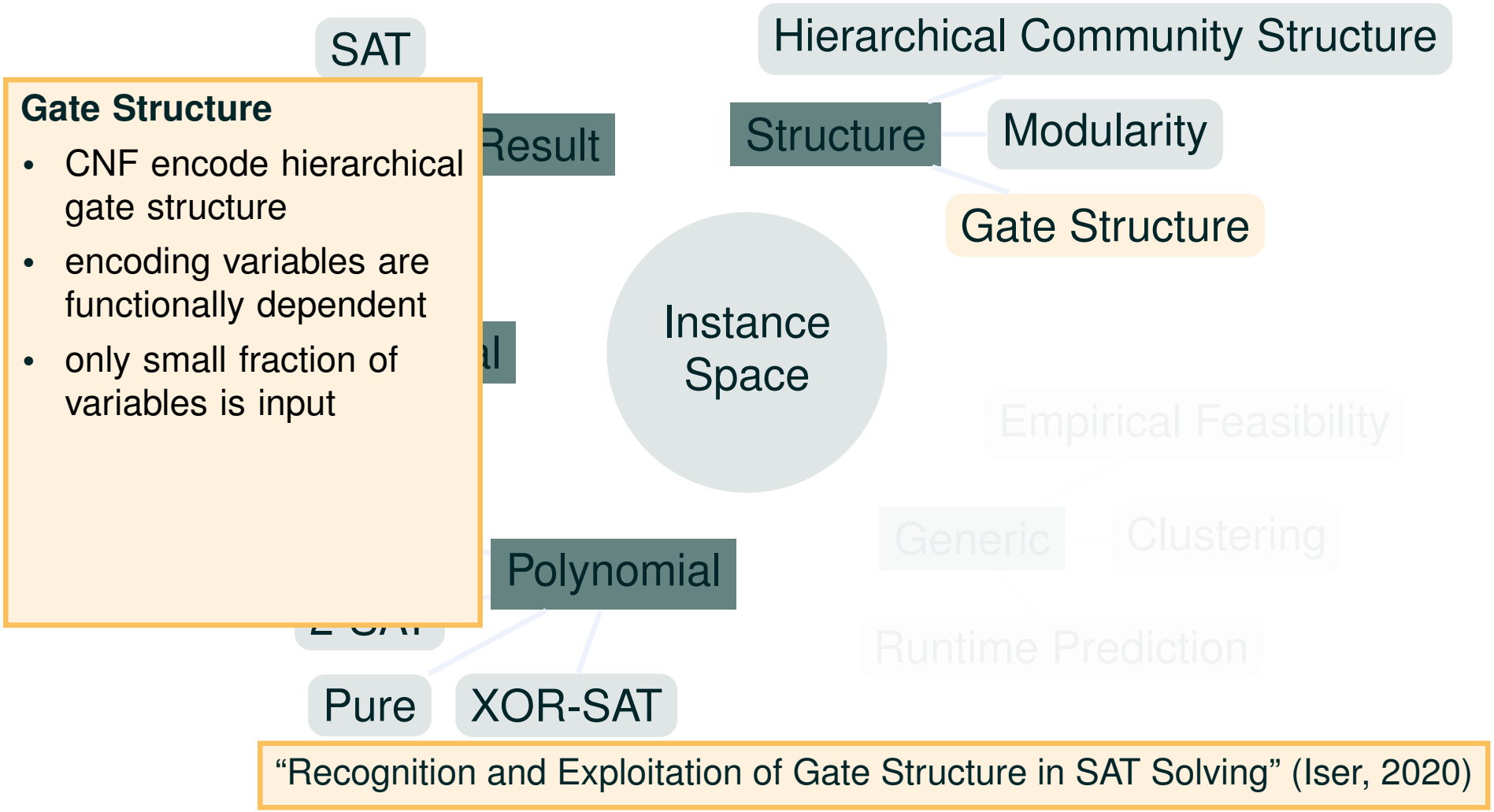


Hierarchical Community Structure

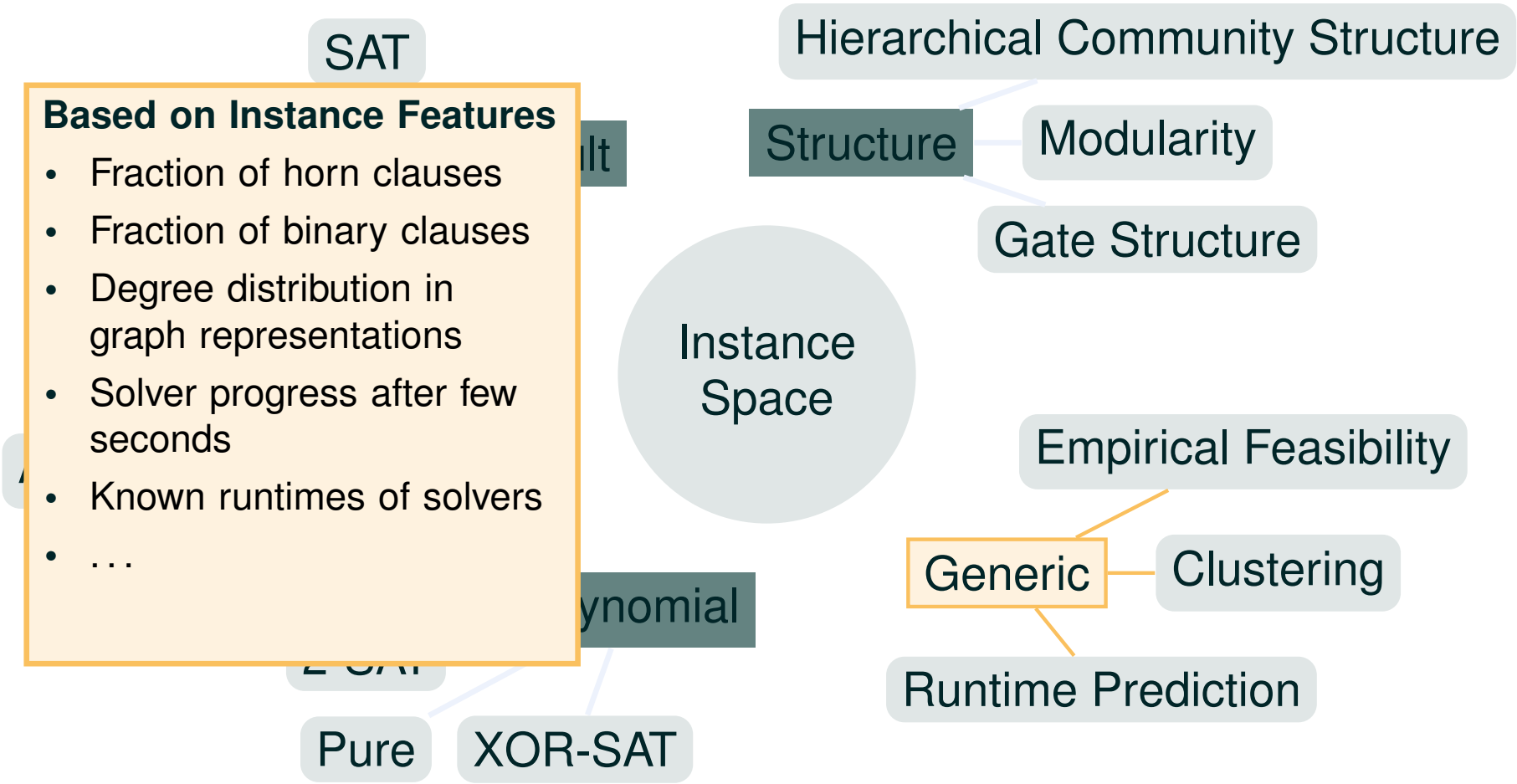
- hierarchical clustering of graph
- measure modularity on levels of hierarchy
- connection to proof complexity

“On the Hierarchical Community Structure of Practical Boolean Formulas” (Li et al., 2021)

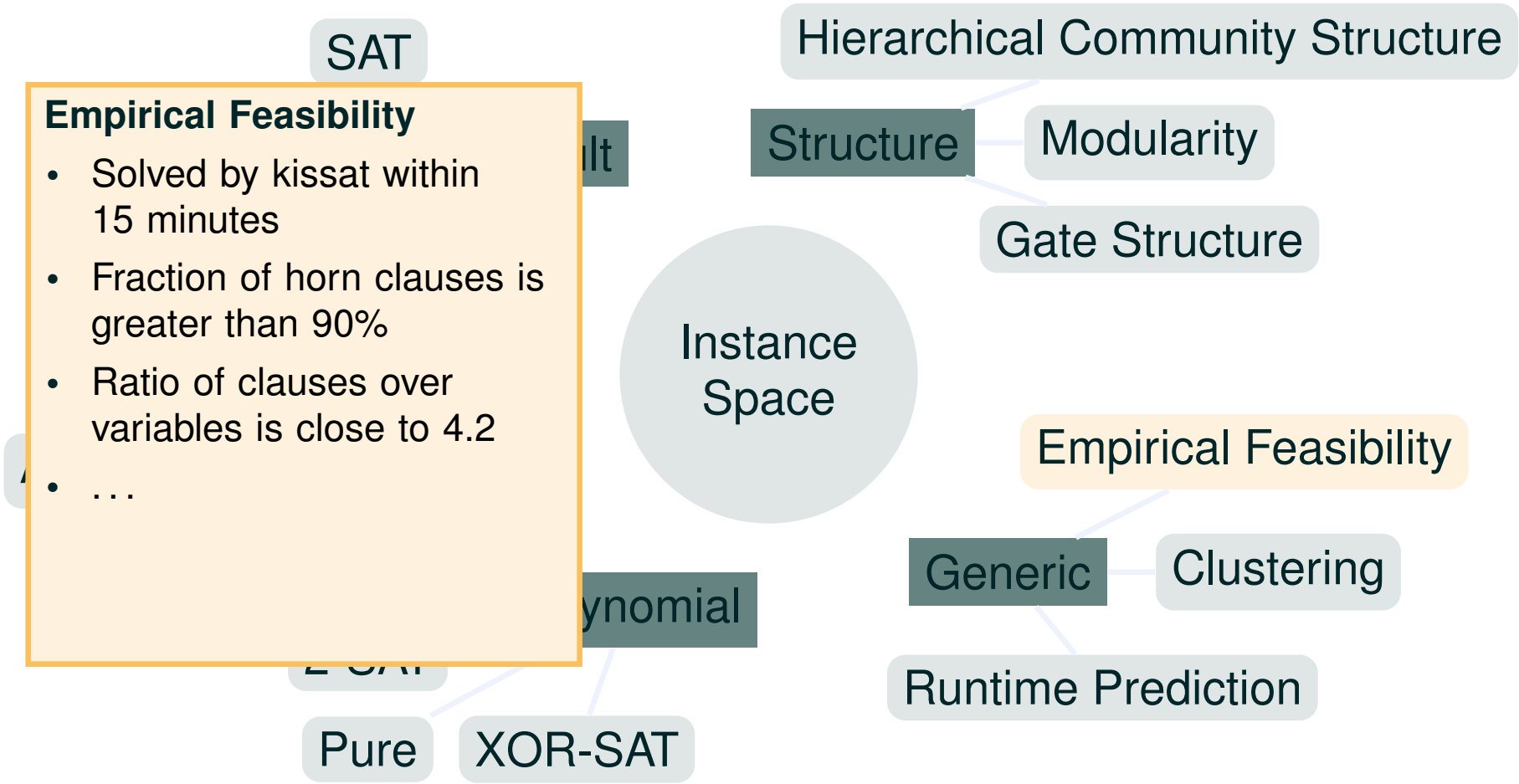
Categories of SAT Instances



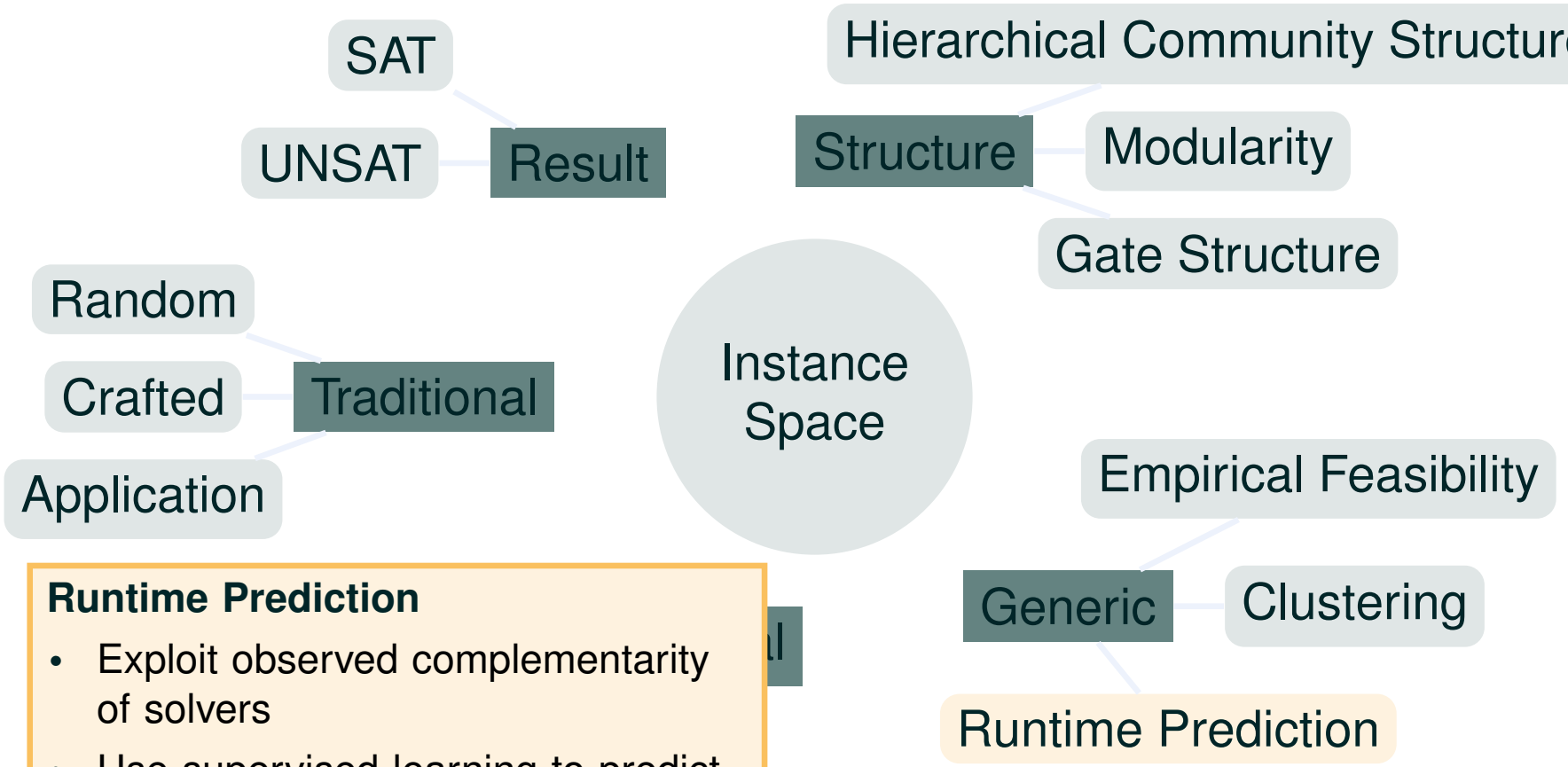
Categories of SAT Instances



Categories of SAT Instances



Categories of SAT Instances

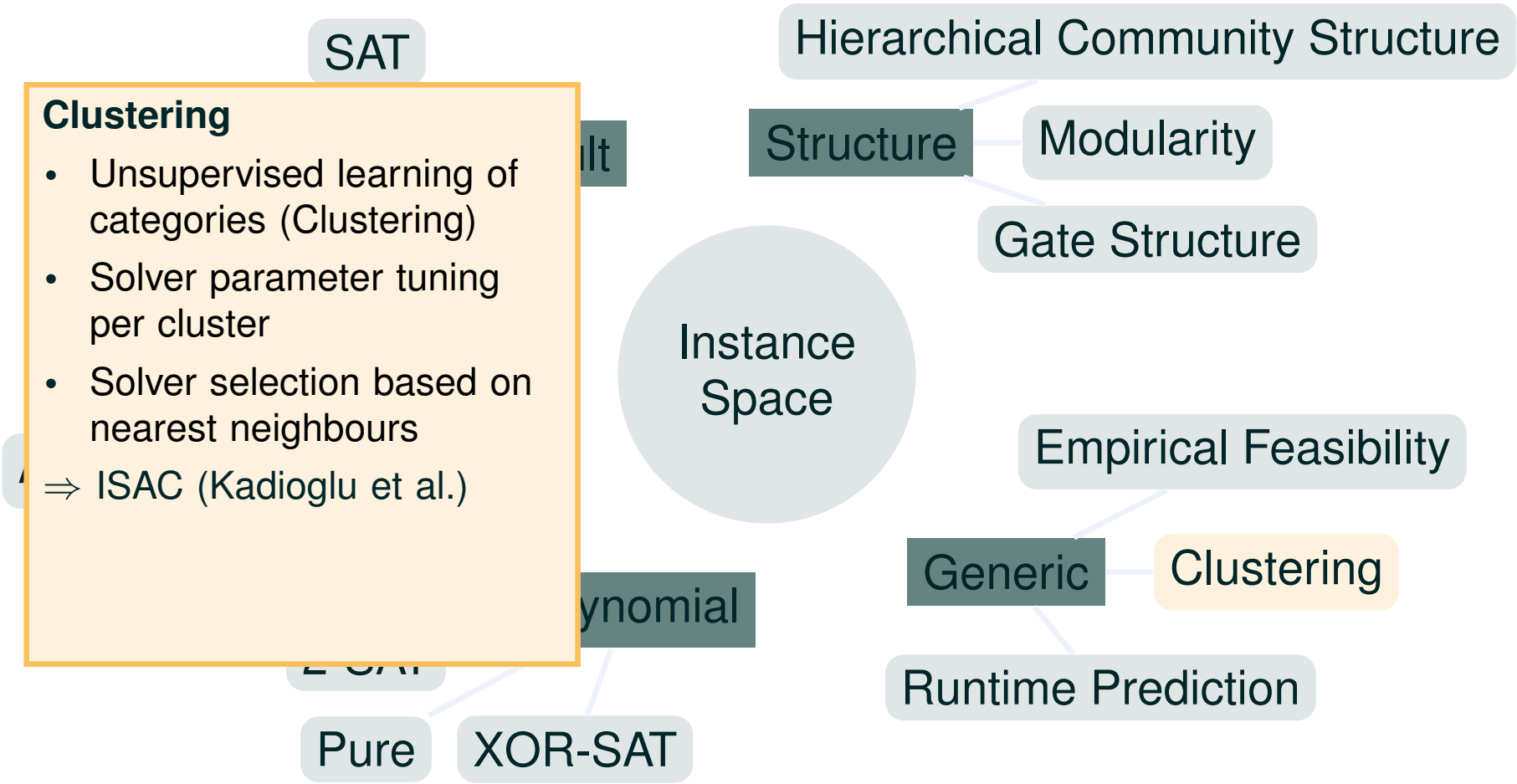


Runtime Prediction

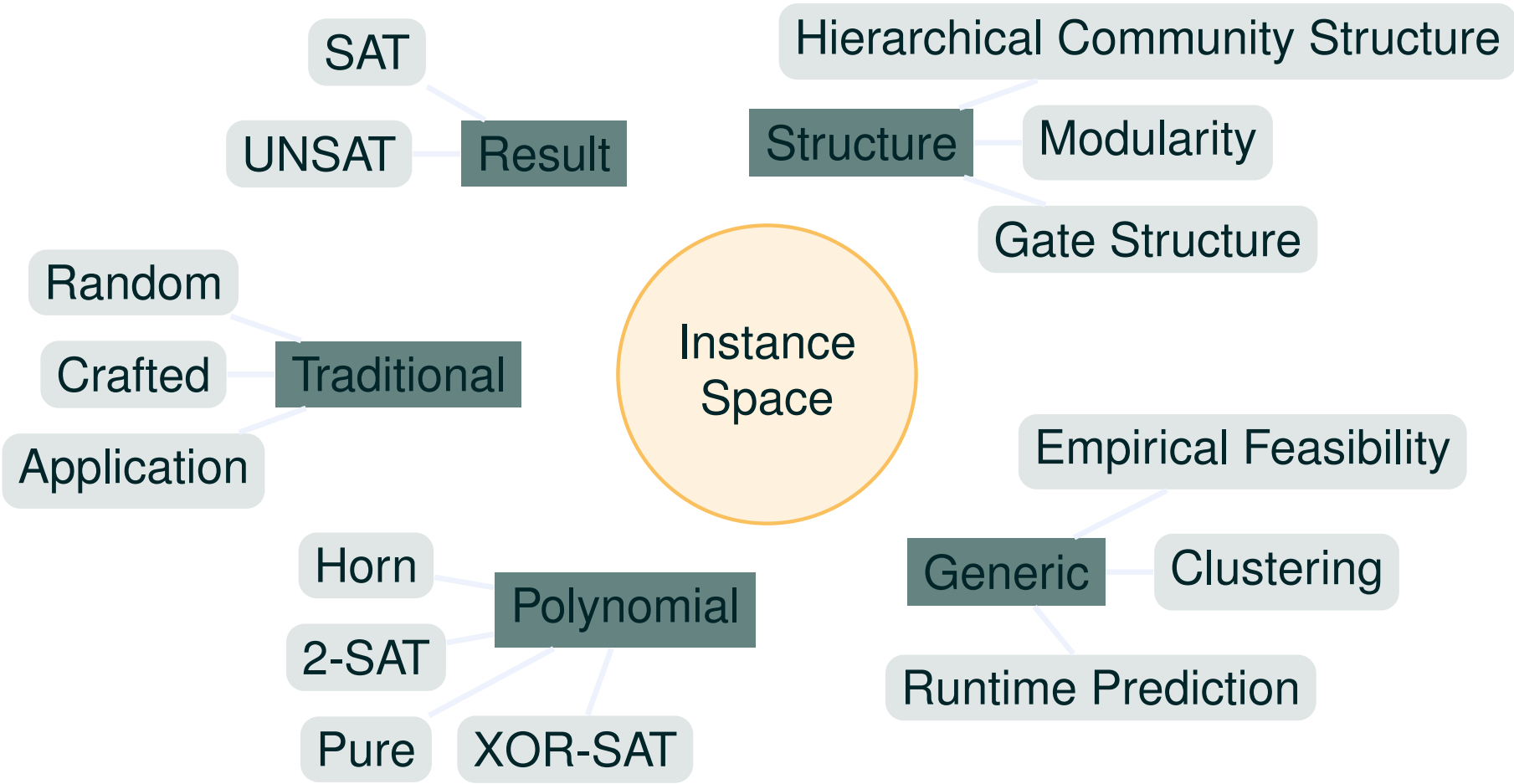
- Exploit observed complementarity of solvers
- Use supervised learning to predict fastest solver

$\Rightarrow$ SATzilla (Xu et al.)

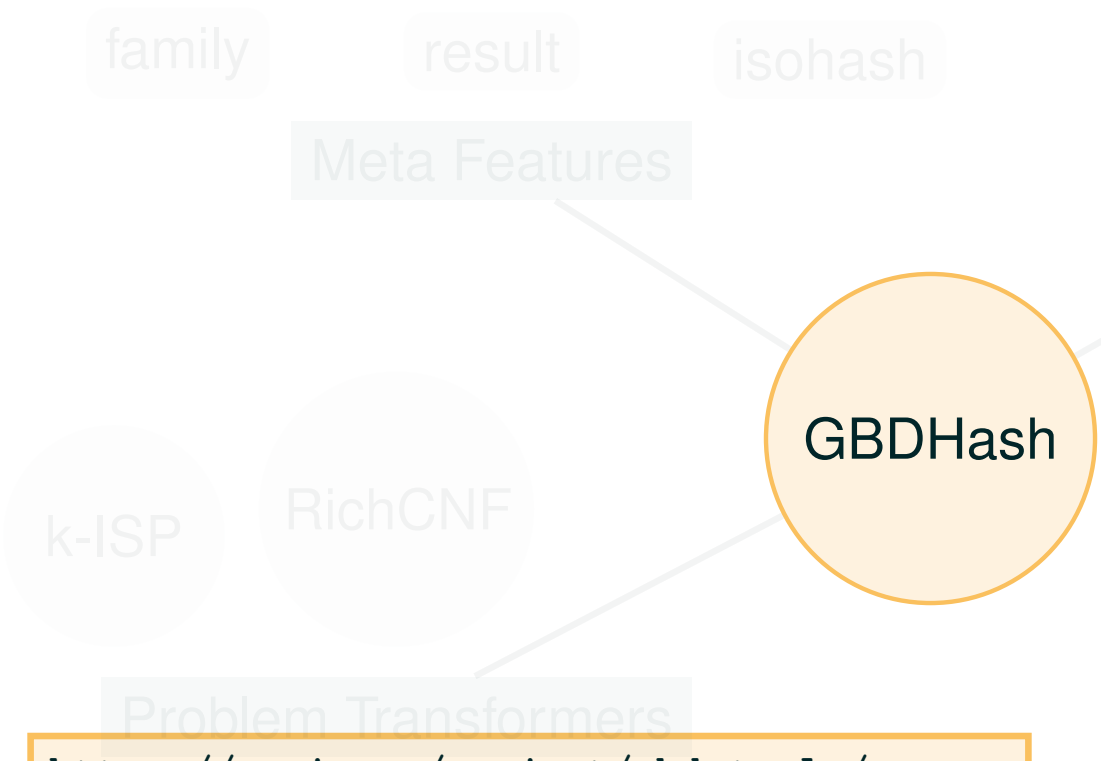
Categories of SAT Instances



Categories of SAT Instances



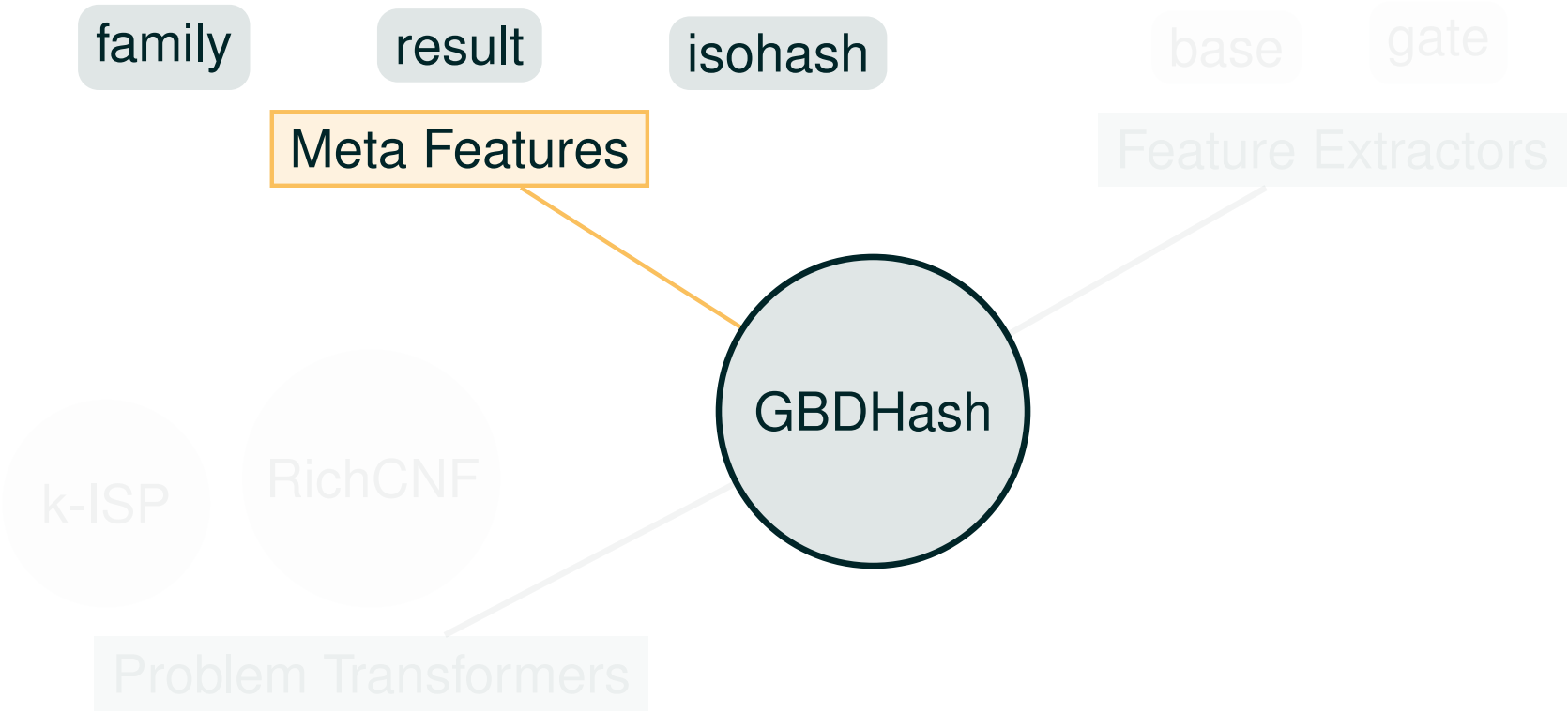
GBD Benchmark Database




```
https://pypi.org/project/gbd-tools/
https://github.com/Udopia/gbd/
https://github.com/sat-clique/cnftools
https://gbd.iti.kit.edu/
```

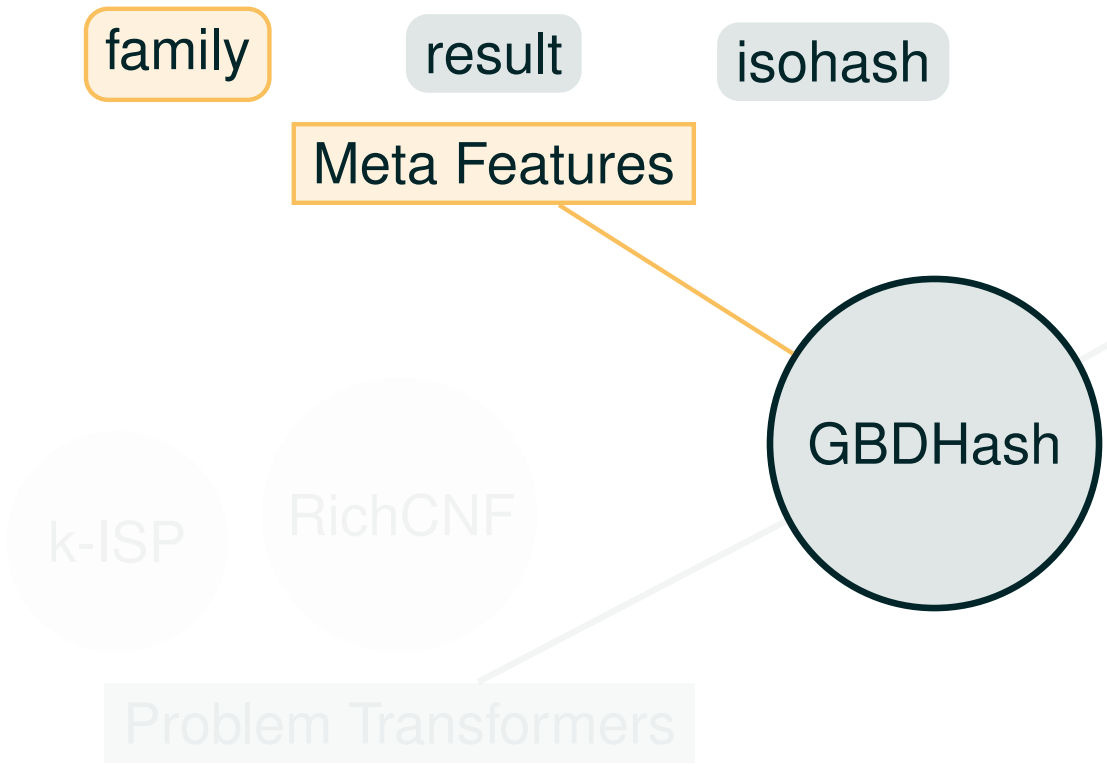
- 27647 CNF Instances
  - SAT Competition benchmarks since 2002
  - further submissions
- **Instance Identification:** GBDHash for DIMACS CNF
  - whitespaces normalized
  - comments removed
  - homogenize line-break
  - etc.

GBD Benchmark Database



`https://gbd.iti.kit.edu/getdatabase/meta_db`

GBD Benchmark Database

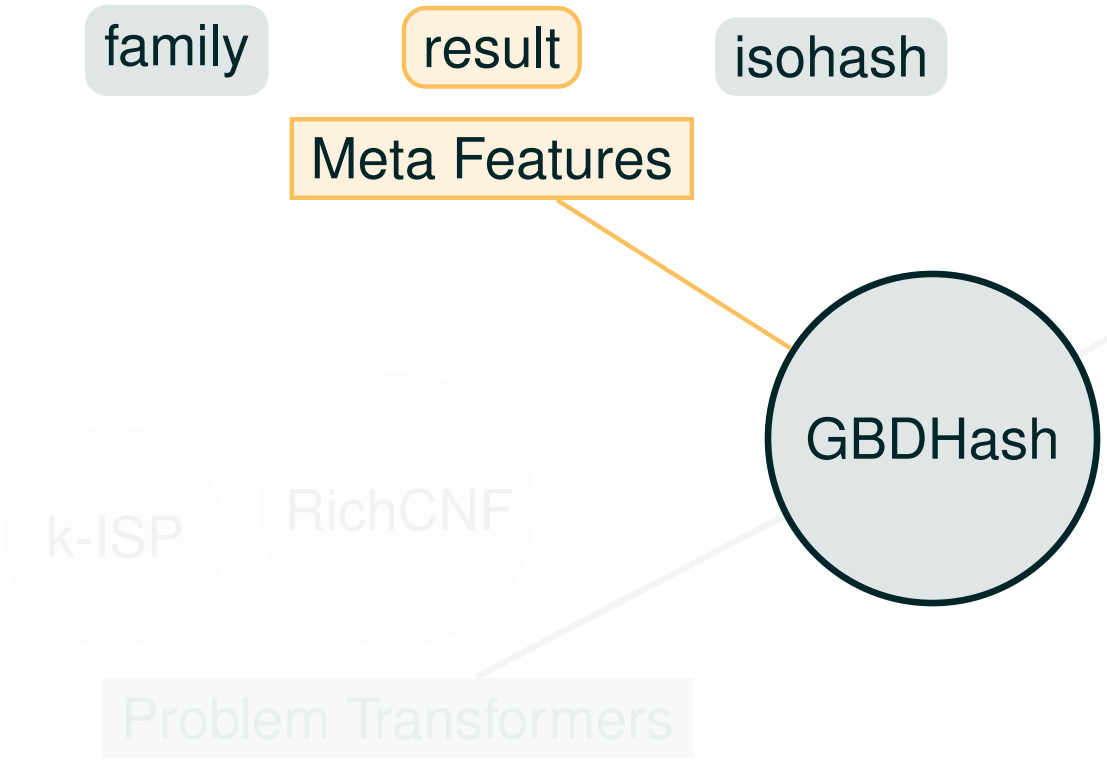


Meta-Feature "Family"

- 26794 entries, 853 unknown
- 135 unique entries

...	...
122	scheduling
210	software-verification
366	graph-coloring
483	planning
570	hardware-verification
841	cryptography
...	...

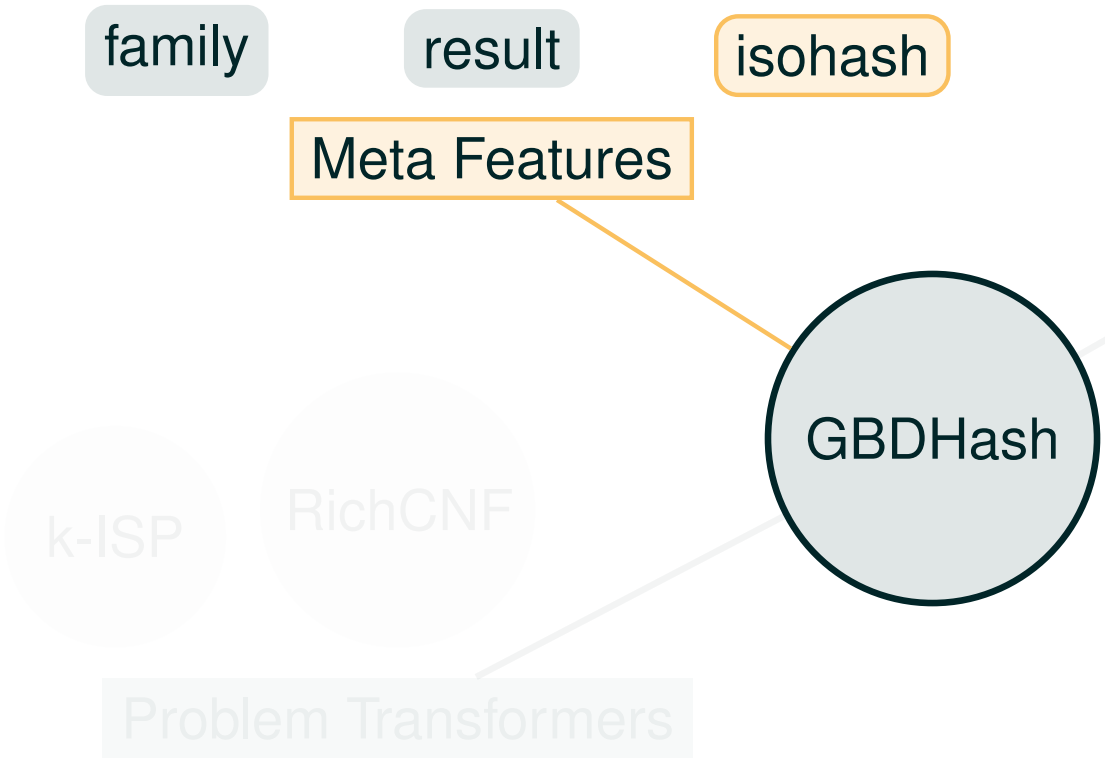
GBD Benchmark Database



Meta-Feature “Result”:

- 11159 entries
  - 5536 unsat
  - 5623 sat
- 16488 unknown

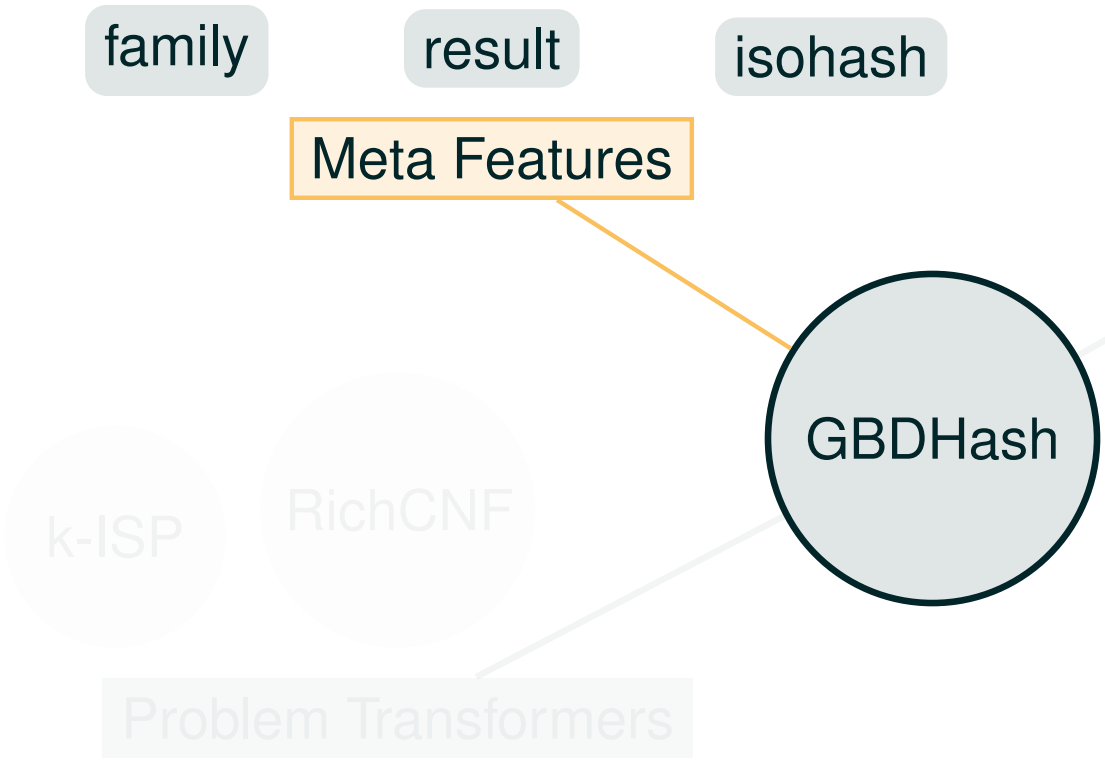
GBD Benchmark Database



Meta-Feature “IsoHash”:

- over-approximates isomorphic instances
  - transform CNF to graph by clique expansion
  - hash sorted degree sequence
- 25713 instances modulo isomorphism

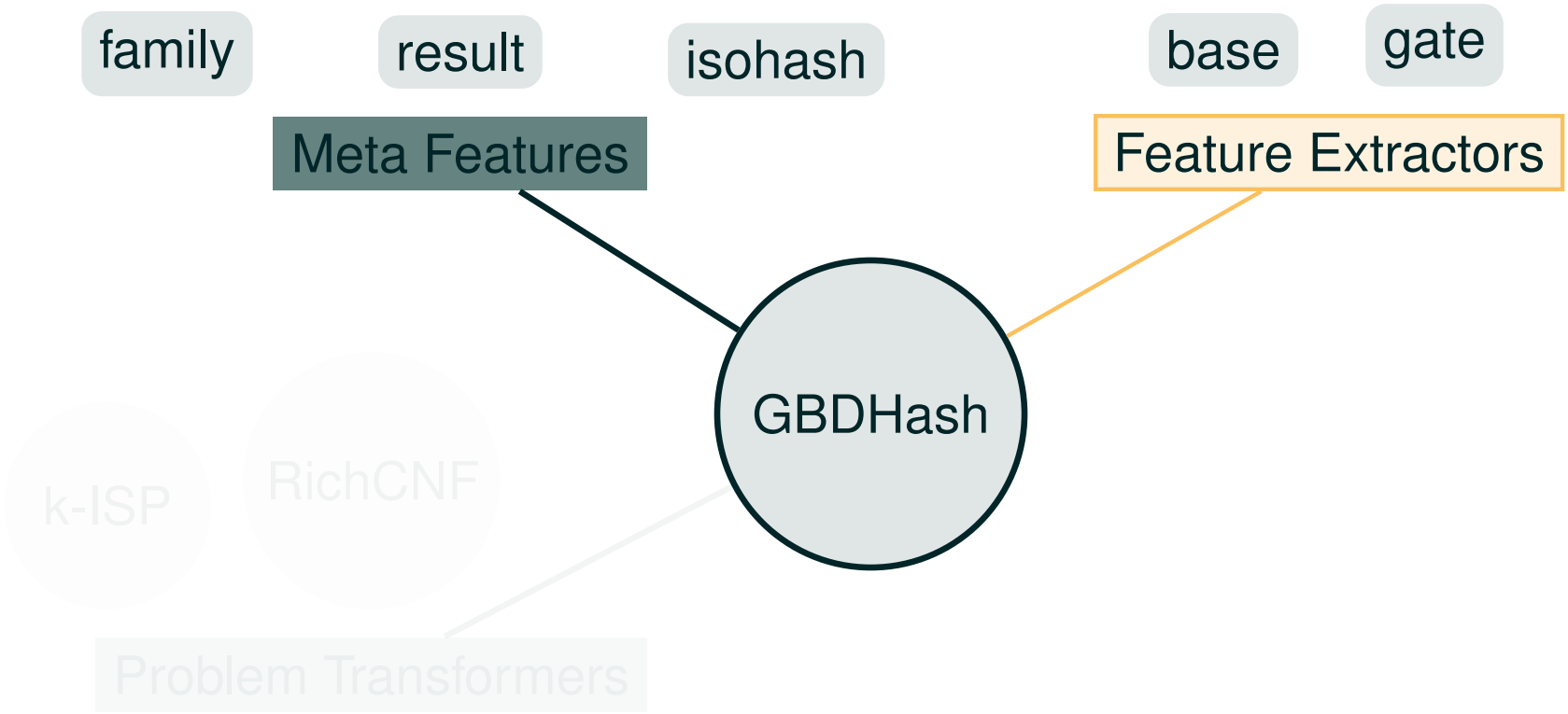
GBD Benchmark Database



More Meta-Features:

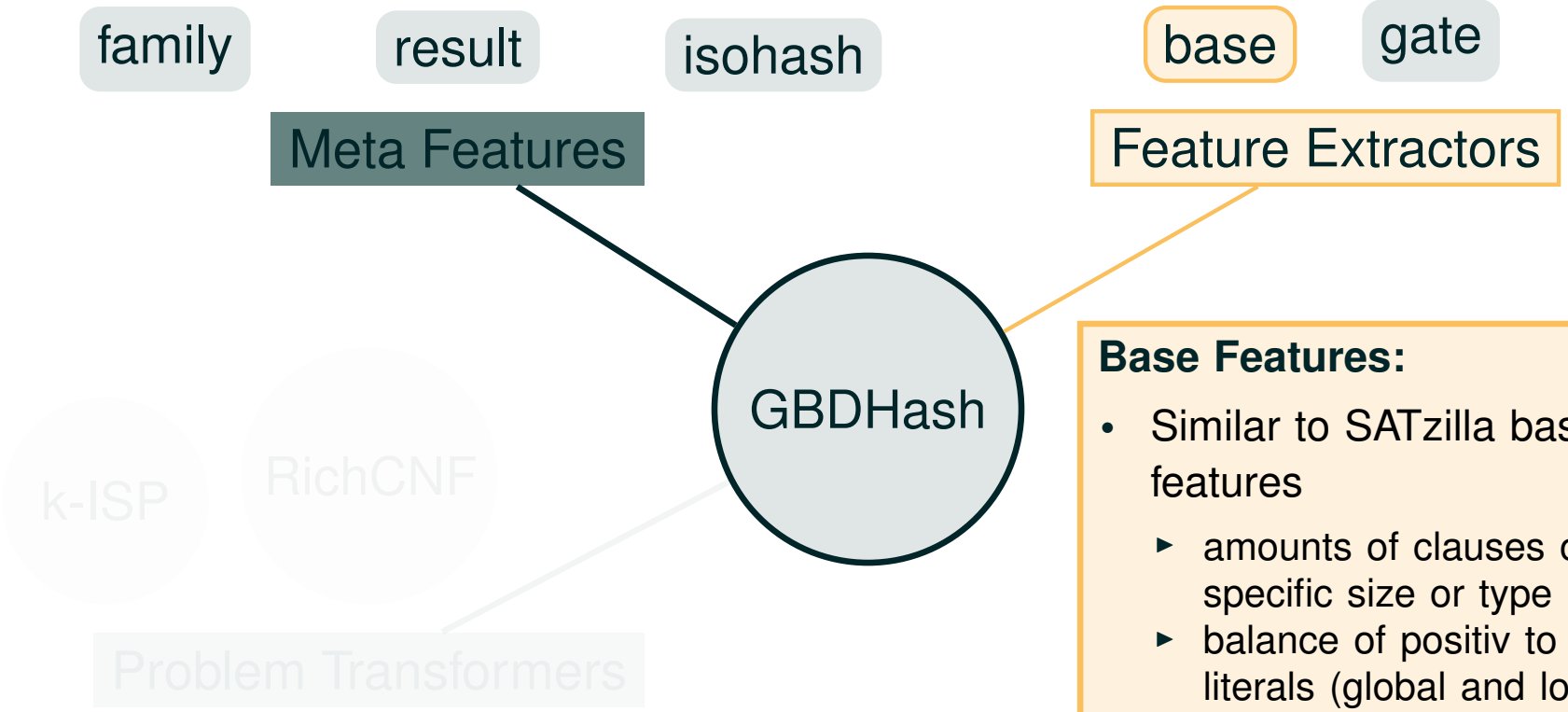
- **author**: who submitted the instance
- **track**: competition benchmark set membership
- **description**: link to benchmark description page in competition proceedings
- ...

GBD Benchmark Database




```
https://github.com/sat-clique/cnftools
https://gbd.iti.kit.edu/getdatabase/base_db
https://gbd.iti.kit.edu/getdatabase/gate_db
```

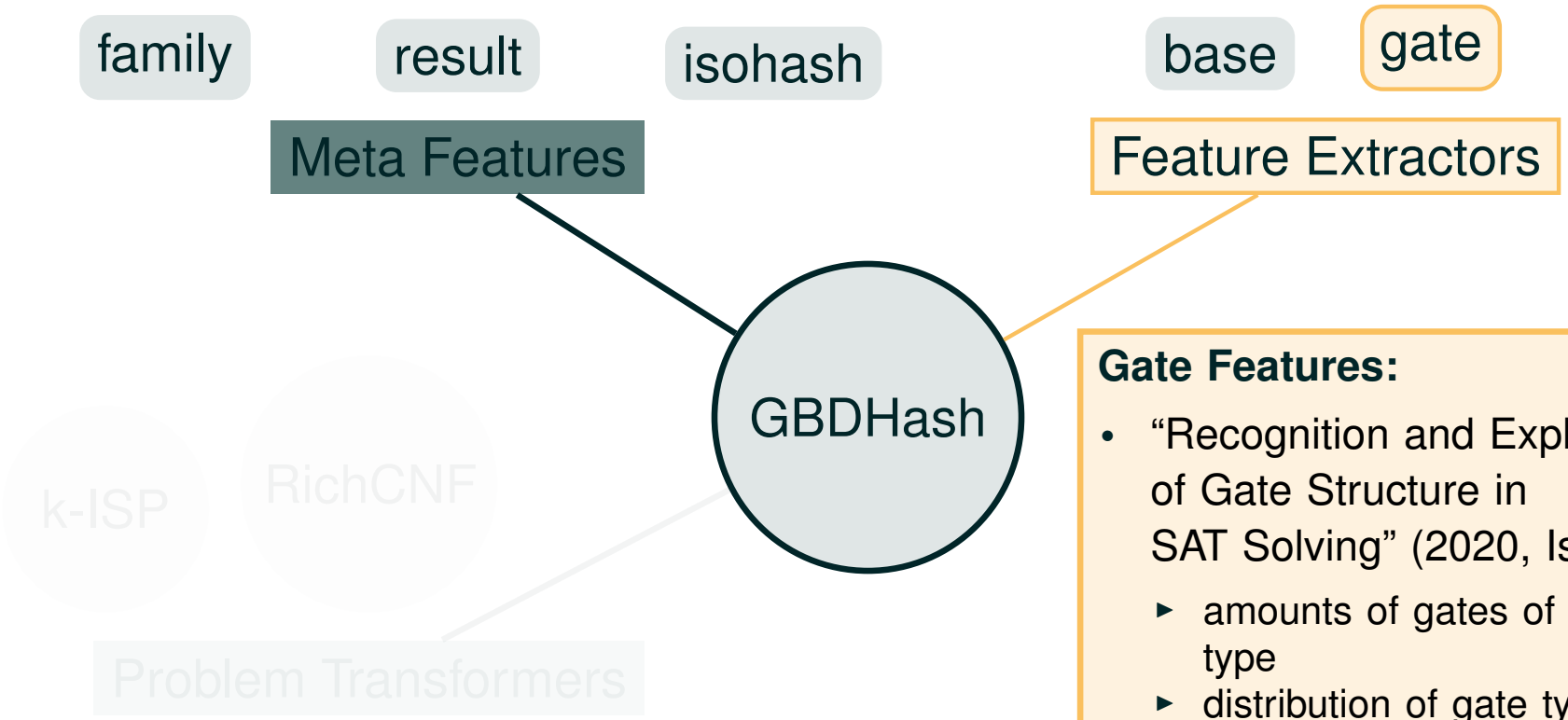
GBD Benchmark Database



Base Features:

- Similar to SATzilla base features
  - amounts of clauses of specific size or type
  - balance of positiv to negative literals (global and local)
  - node degree distributions of various graph representations
  - ...

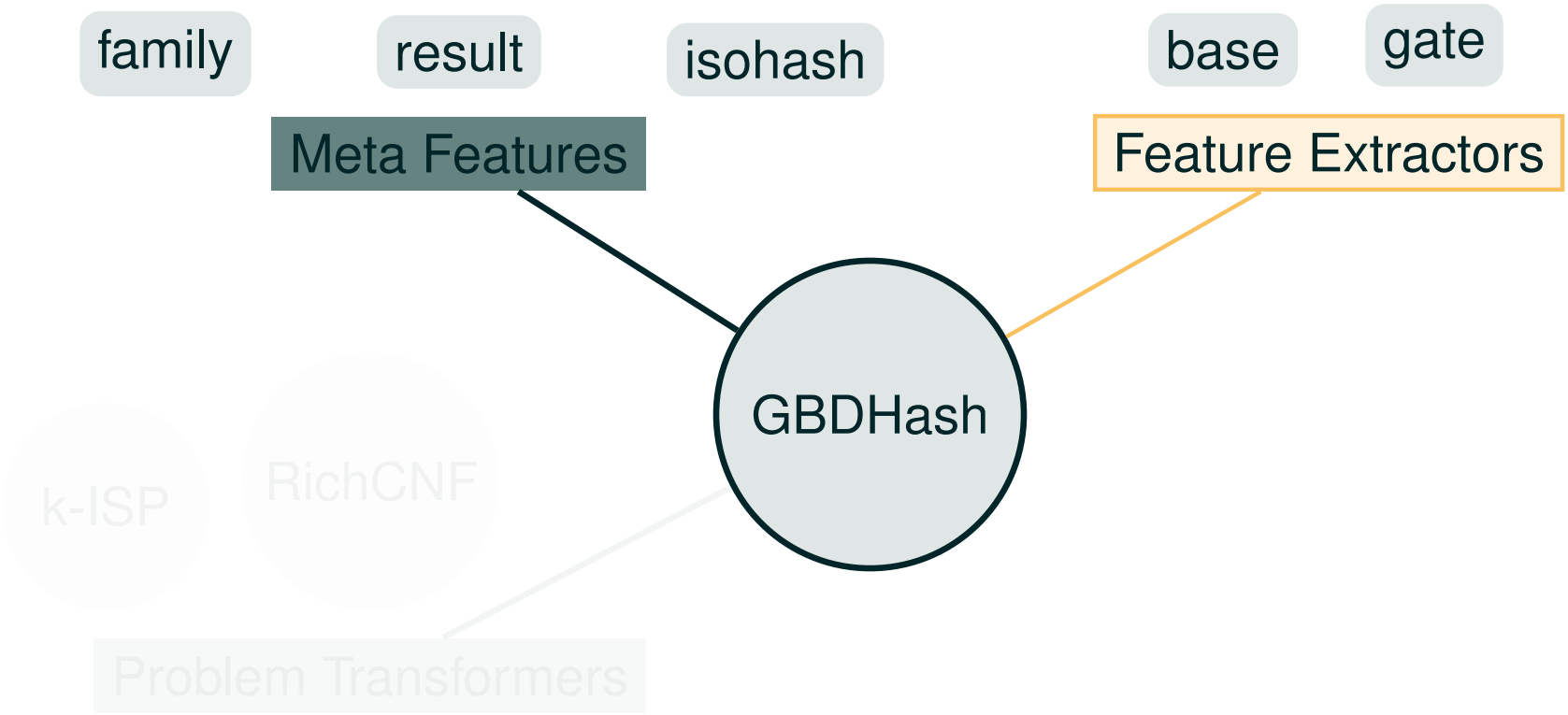
GBD Benchmark Database



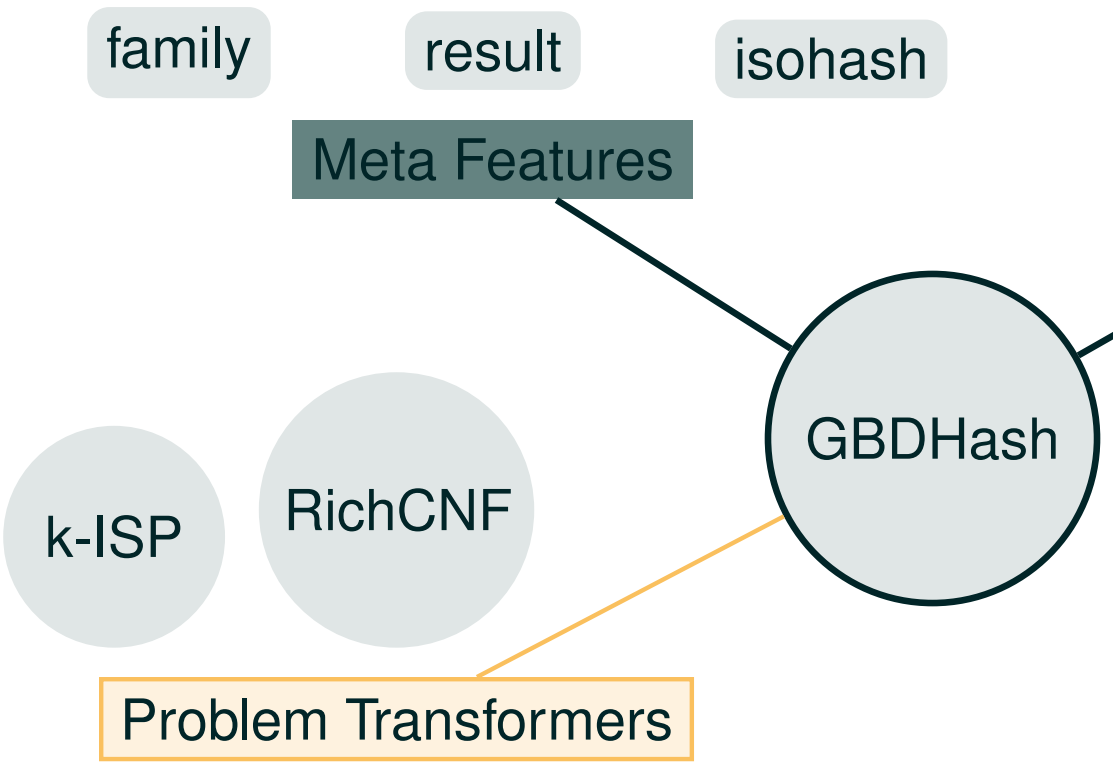
Gate Features:

- "Recognition and Exploitation of Gate Structure in SAT Solving" (2020, Iser)
  - amounts of gates of specific type
  - distribution of gate type over levels in hierarchical gate structure

GBD Benchmark Database



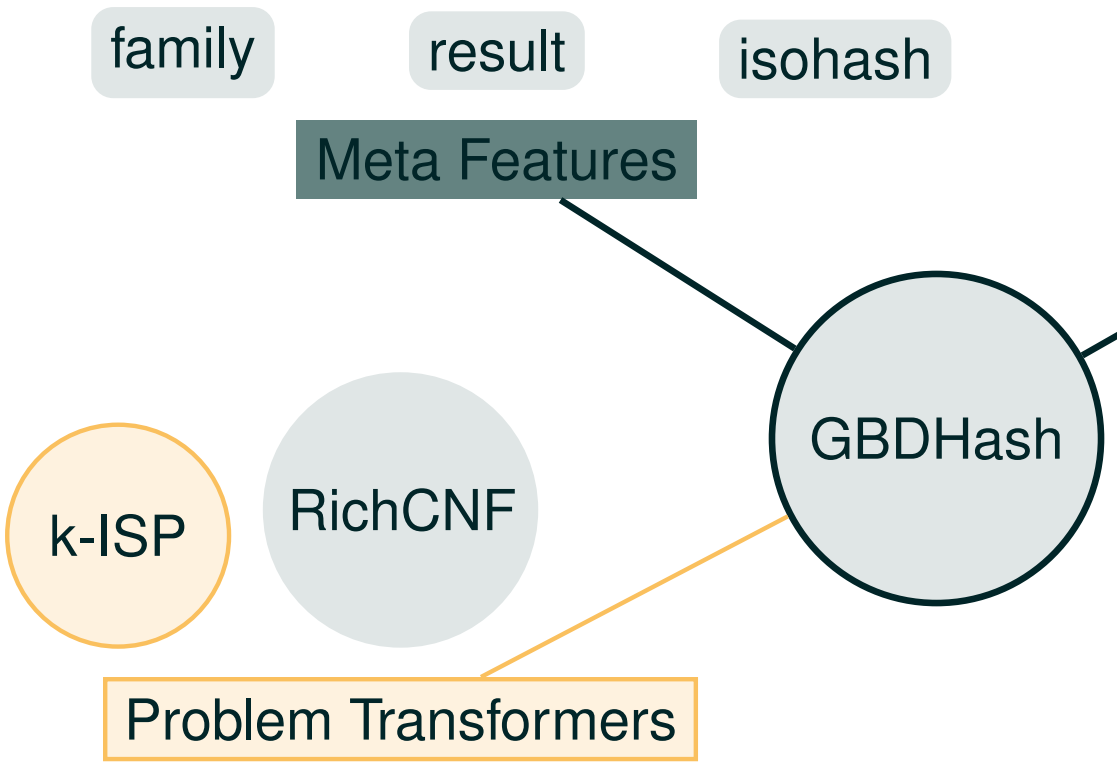
GBD Benchmark Database



Problem Transformers:

- handle instances of other domains
- handle relations between instance of several domains
- scale tool-chain to other domains

GBD Benchmark Database



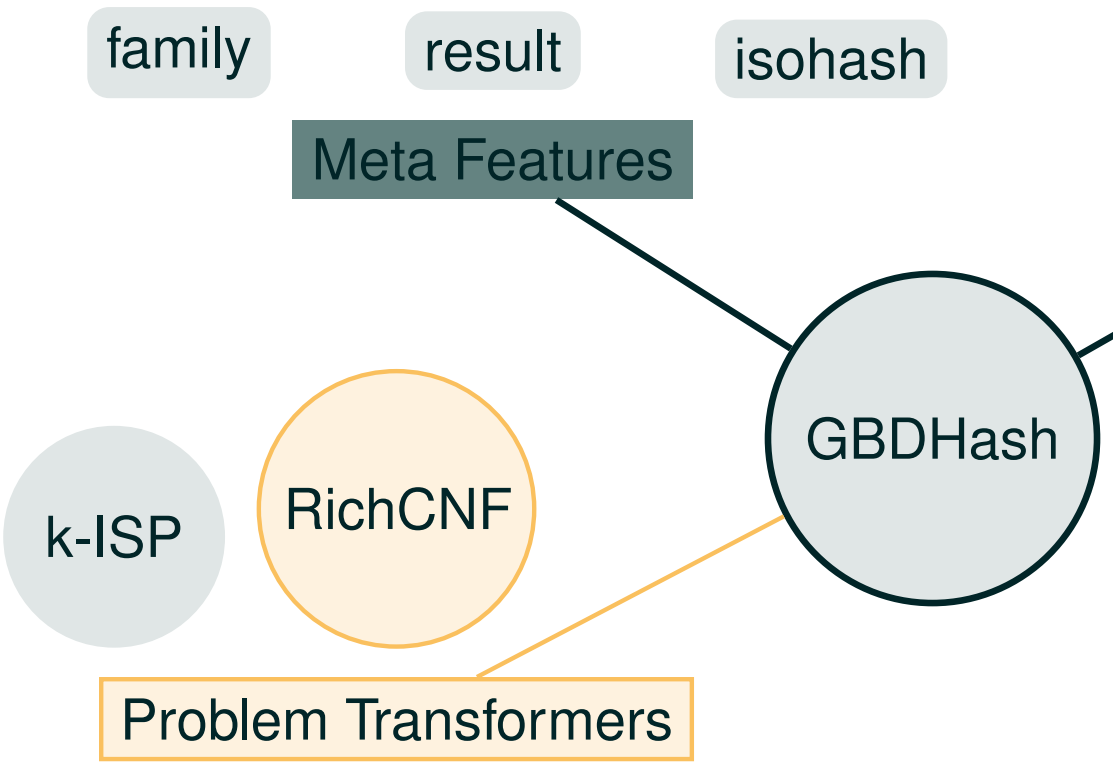
Independent Set Problem:

Given graph (V, E) and $k \in \mathbb{N}$

Decision Problem: Existence of independent set $I \subseteq V$ of size k

Indepentent Set is set of mutually non-adjacent nodes, i.e, $\nexists (v, w) \in E, \{v, w\} \subseteq I$

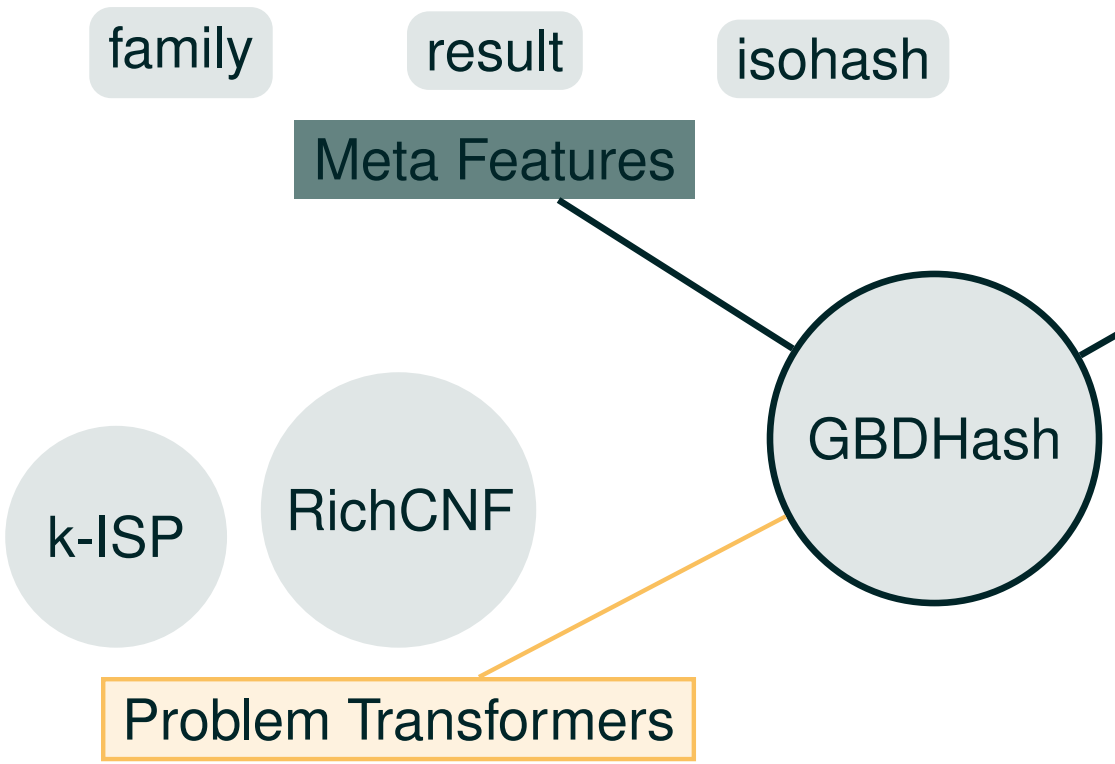
GBD Benchmark Database



RichCNF:

- conjunction of general constraints
- constraints can be:
  - cardinality constraints
  - DNF
  - BDD
  - ...

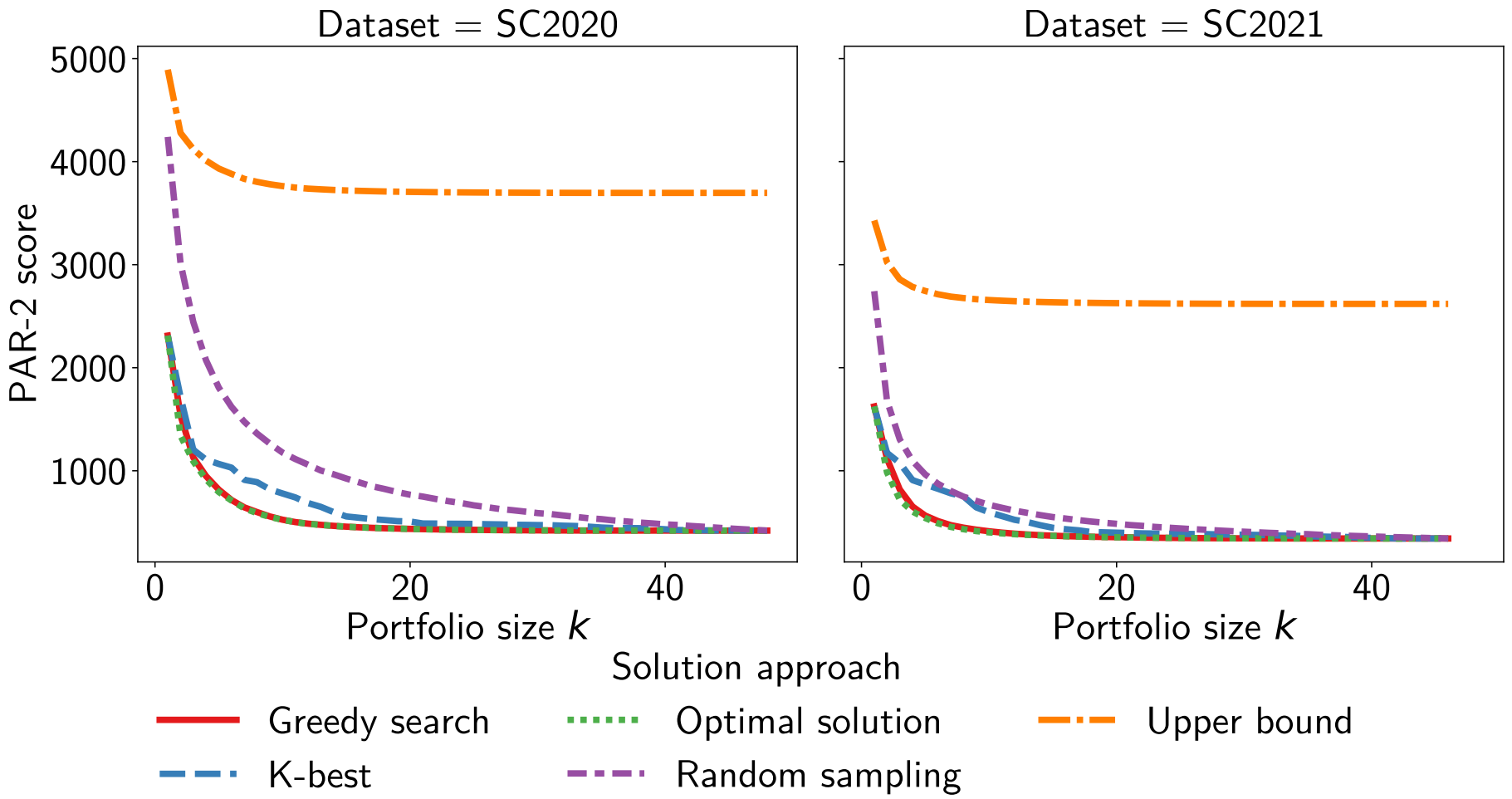
GBD Benchmark Database



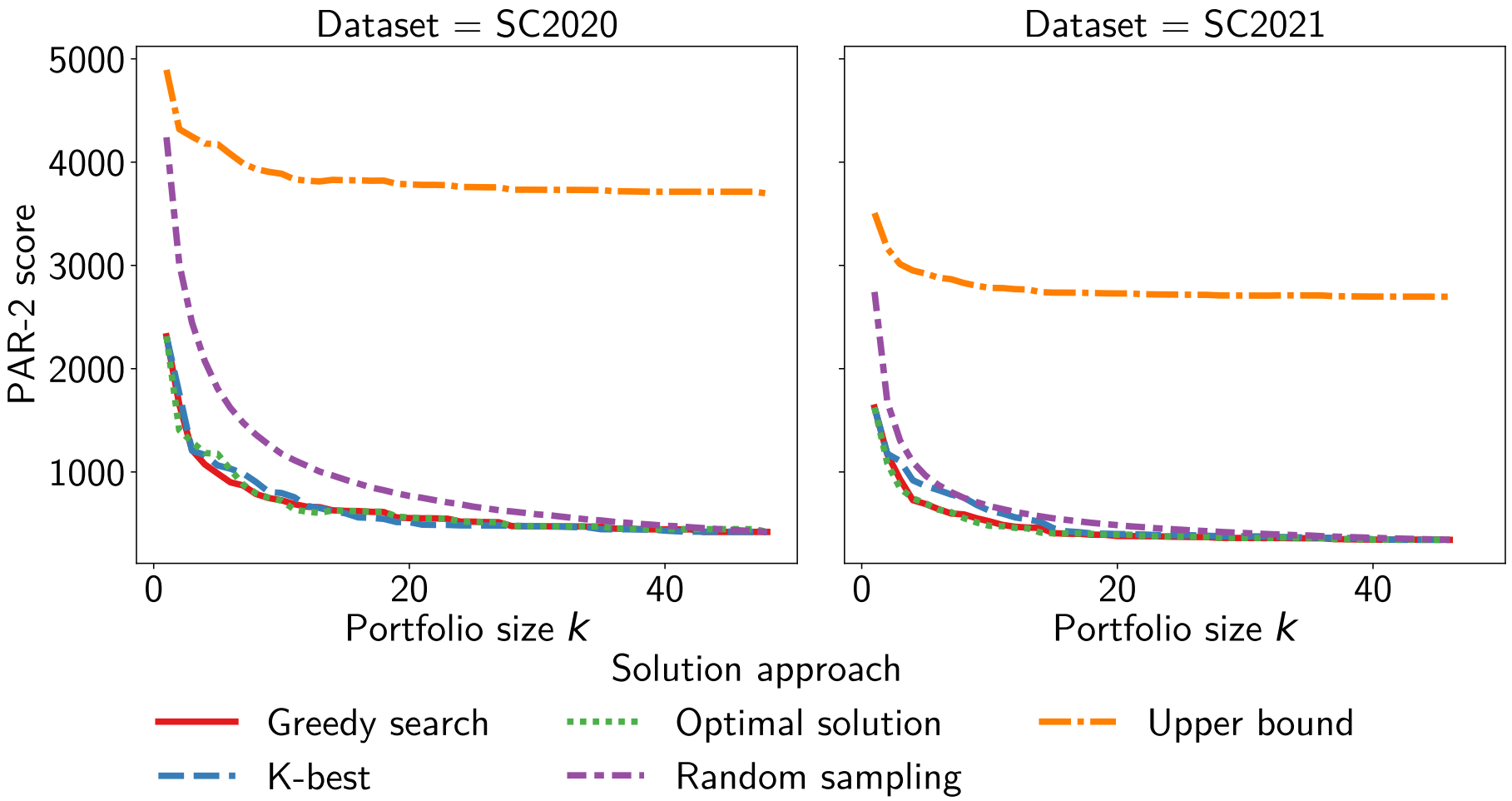
Future Domains:

- MaxSAT
- QBF
- PB
- Graphs
- DRAT Proofs
- ...

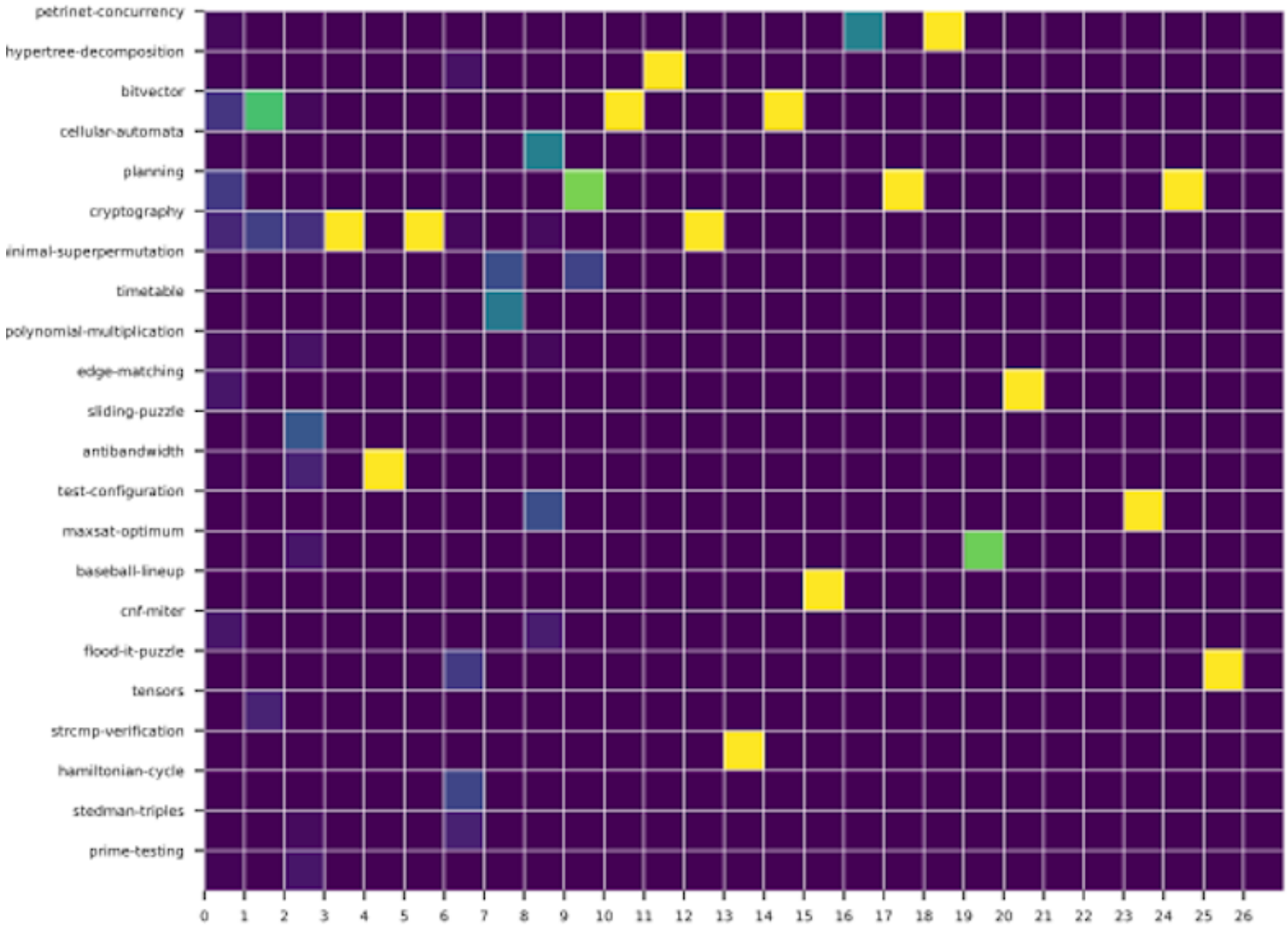
k-Portfolios



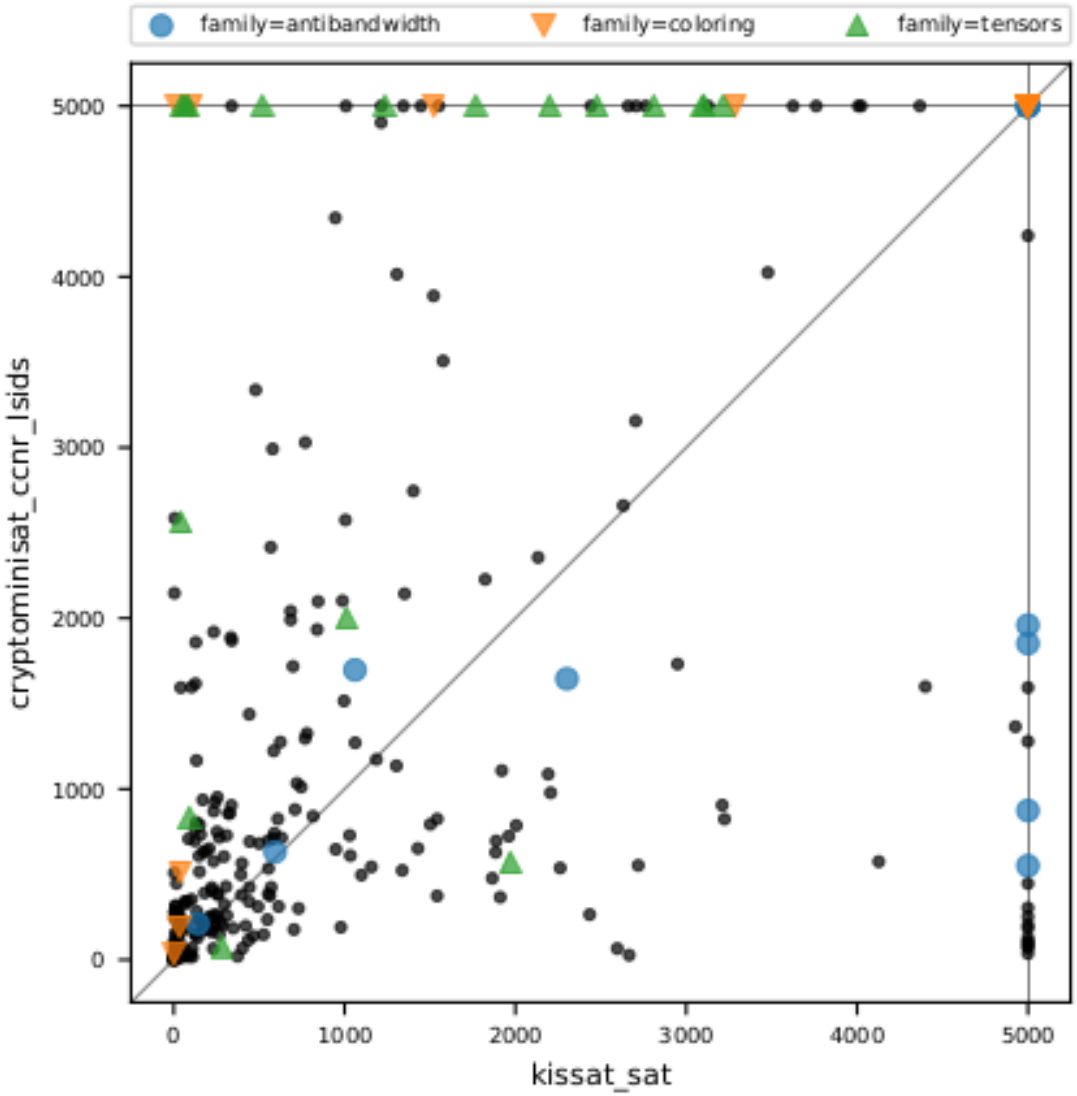
k-Portfolios (test)



Cluster Composition (dbscan)



Families


family=antibandwidth
family=coloring
family=tensors
cryptominisat_ccnr_lsids
kissat_sat
0
1000
2000
3000
4000
5000