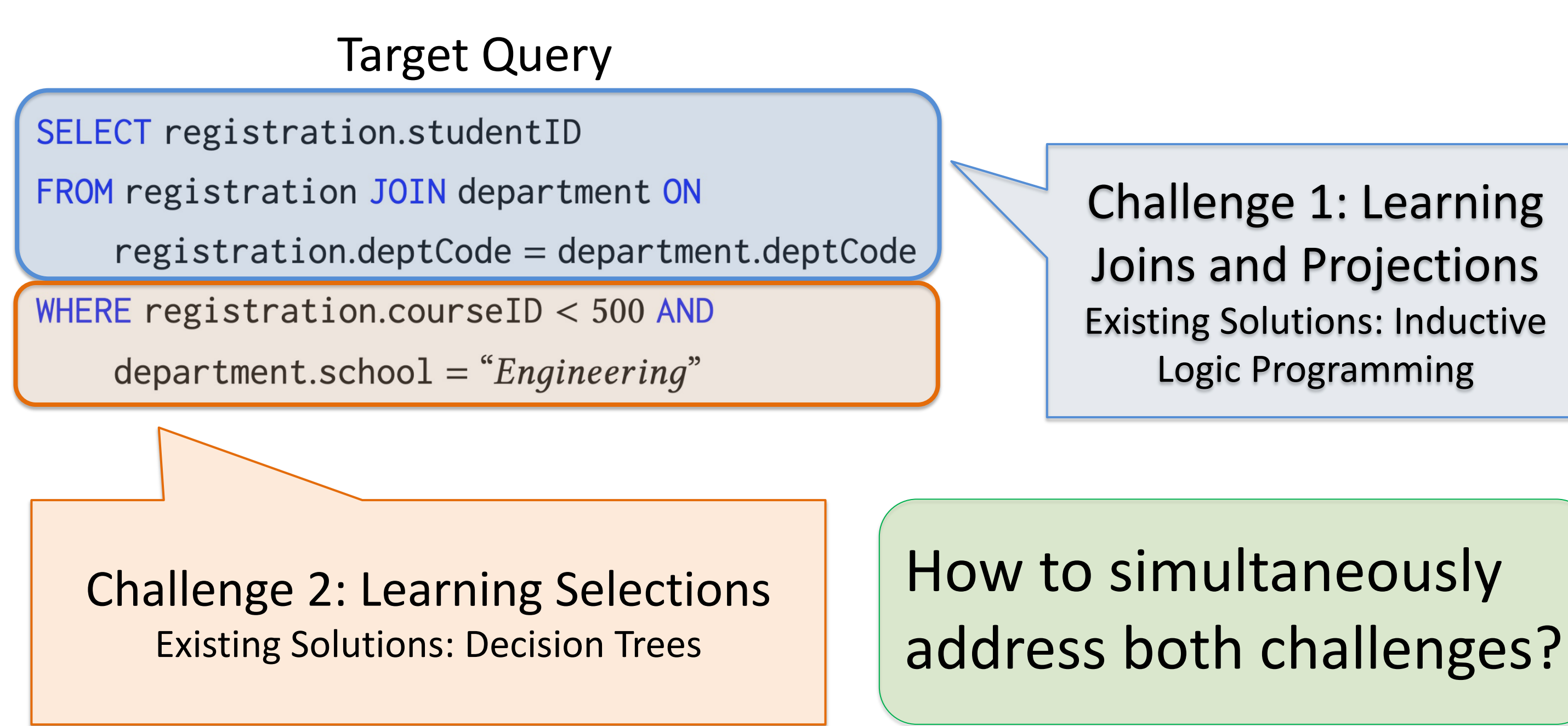
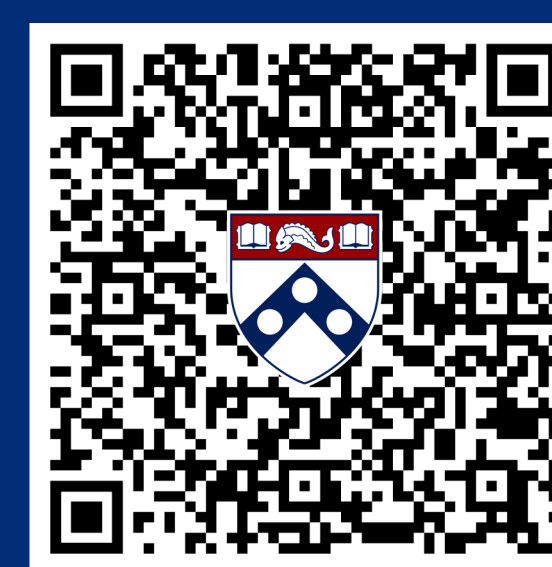
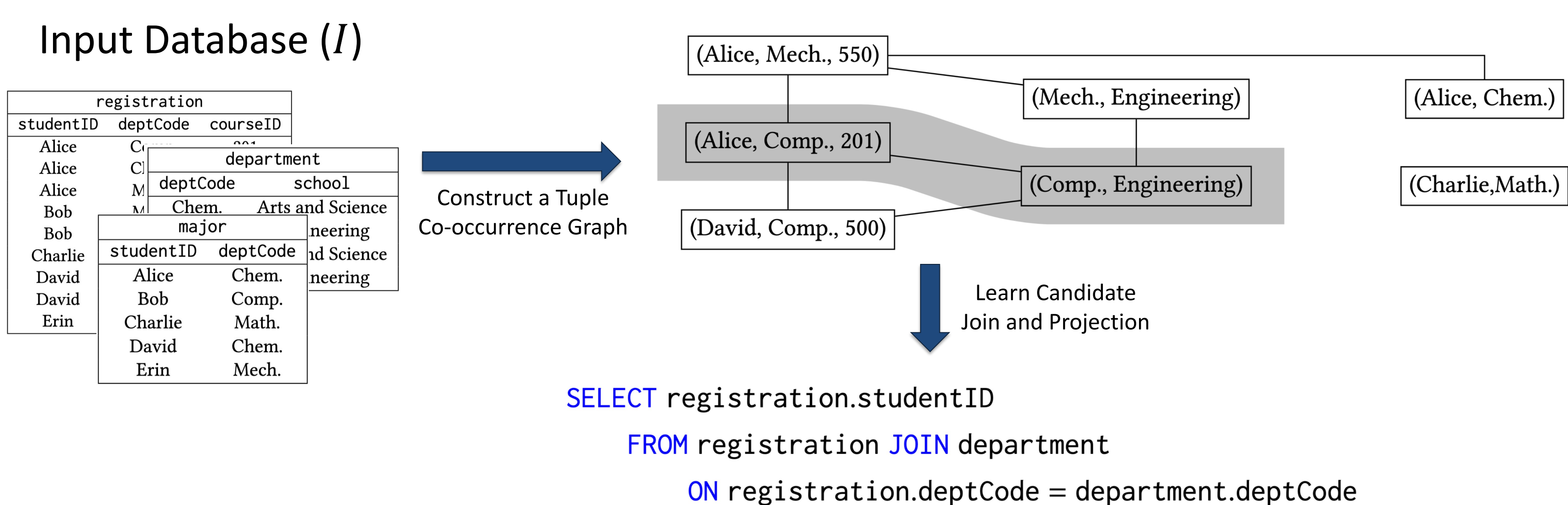
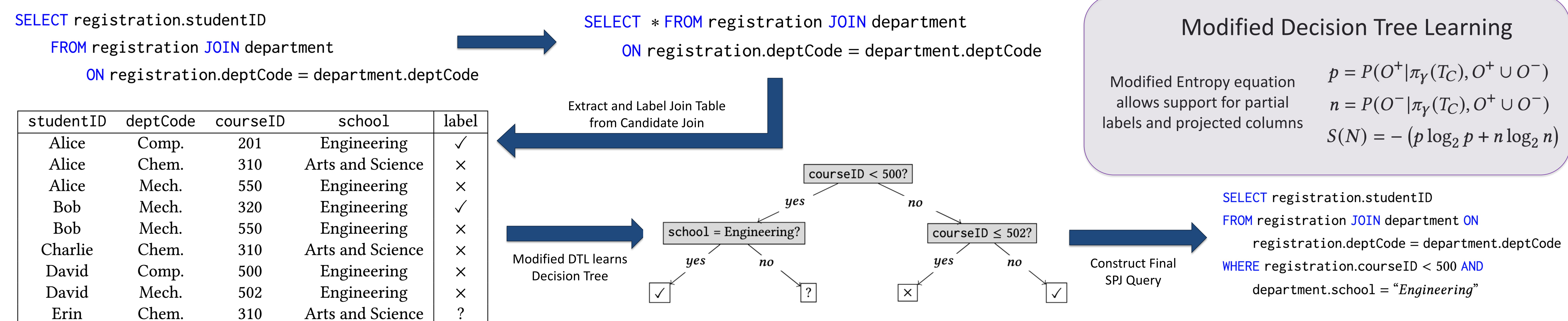




Learning Joins and Projections



Learning Selections



Evaluation Setup

Baselines

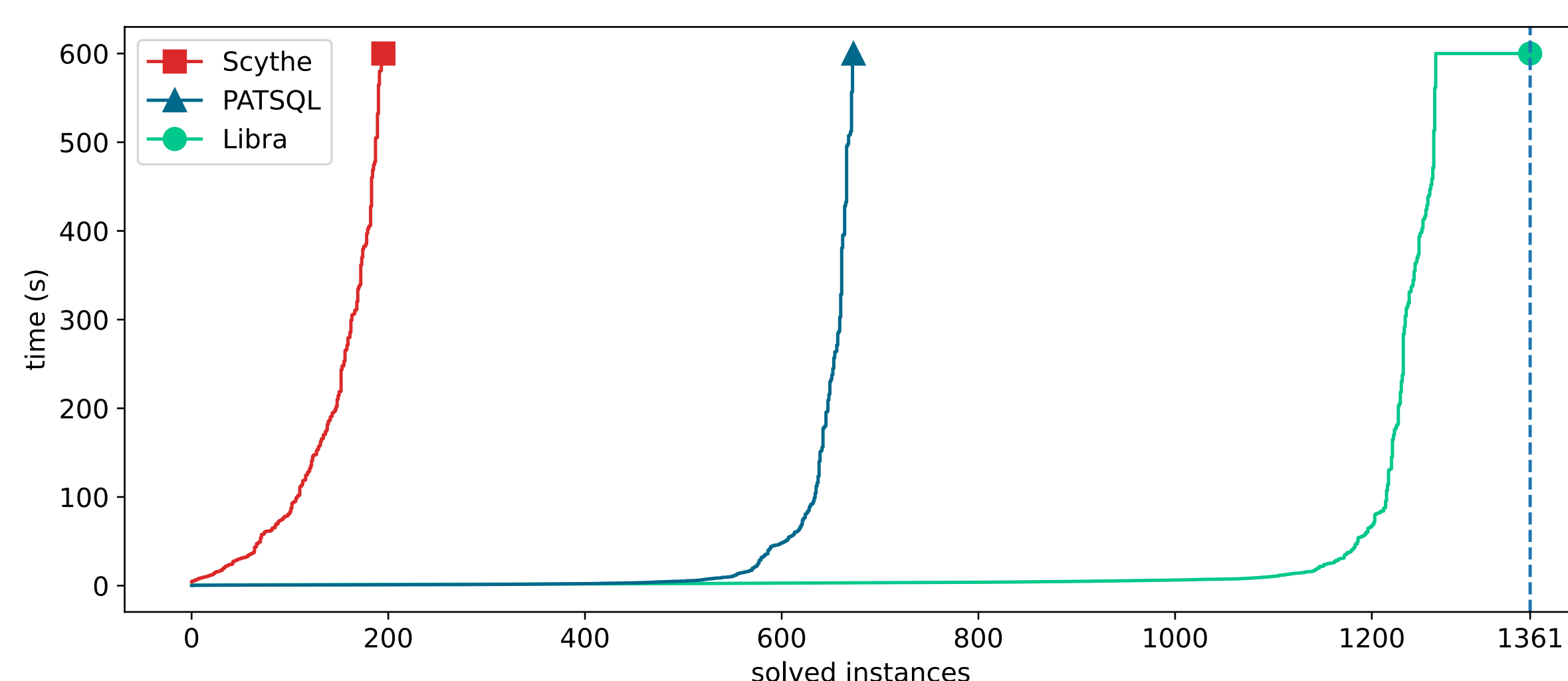
1. Scythe
2. PATSQL

Benchmarks

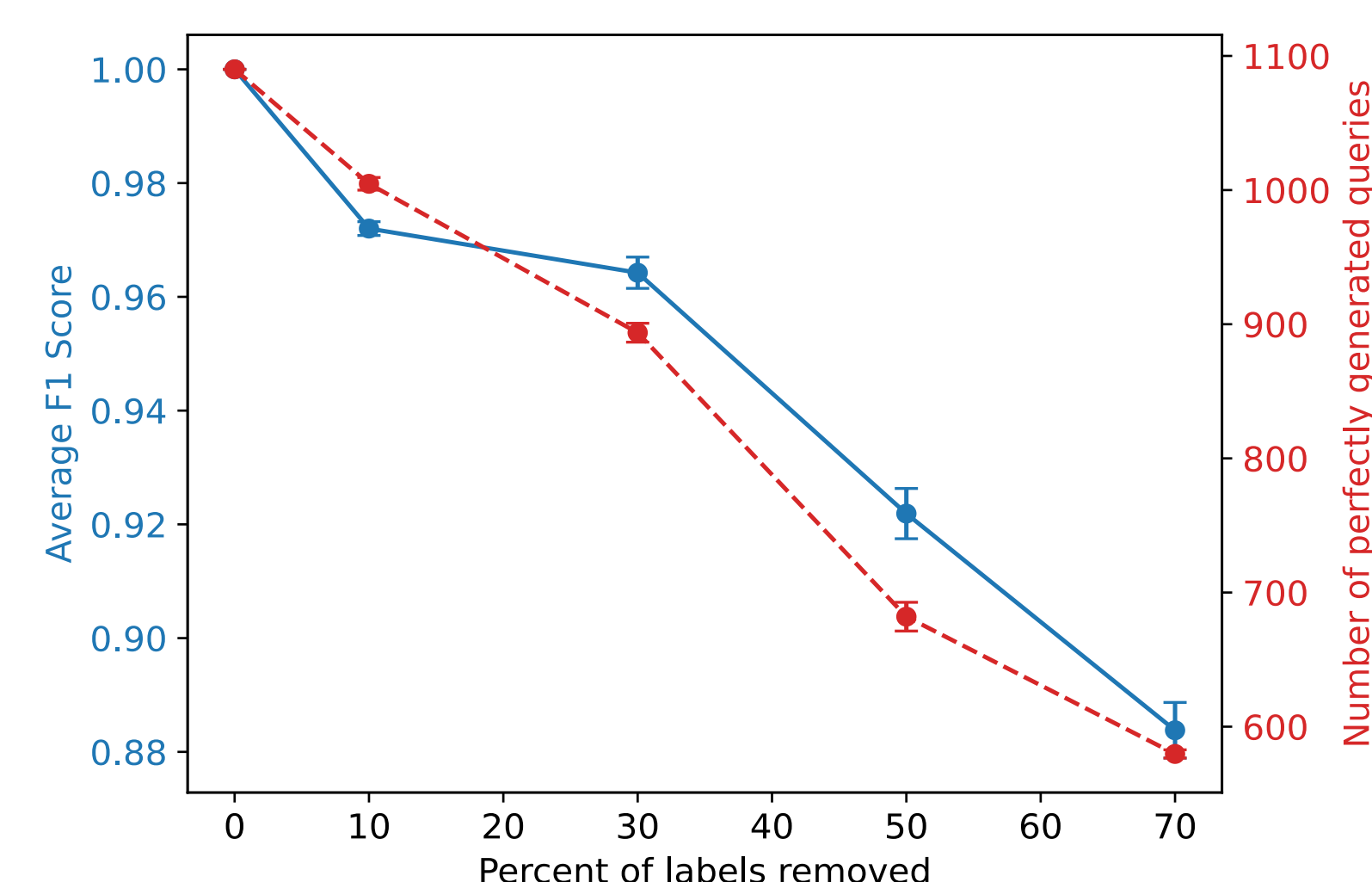
1496 SPJ queries over 160
databases from the Spider and
Geography benchmarks

For each query we provide the full database along with positive and negative labels.

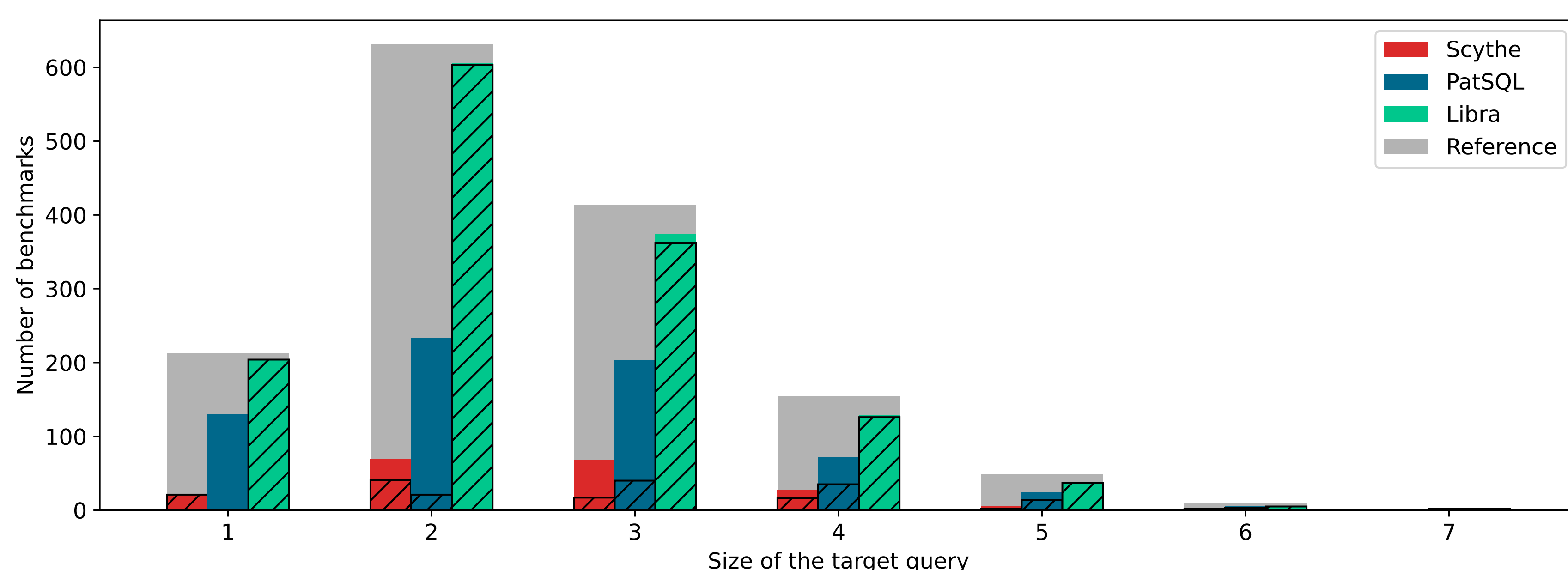
Performance



Sensitivity to Partial Labels



Succinctness



Conclusion

1. We present a novel technique to synthesize SPJ queries by interleaving the synthesis of joins and selections
2. We synthesize joins using an example-guided enumerator and modify the classical decision tree learning algorithm to synthesize selections
3. We implement the algorithm in a tool named Libra and evaluate it over around 1500 queries from 160 databases
4. Libra solves **1,361 benchmarks**, outperforming PATSQL (673) and Scythe (195).
5. **99%** of Libra's solutions are minimal.
6. Libra is robust to partial labels, producing solutions with an F1 score of **0.88** given only **30%** of the labels.