Thesis Report

March 1, 2016

1   Introduction and Background

We would use existing solver as a black-box to build upon and call it as subroutine. Our new solver's search strategy would be guided by the structure and call IDP to solve subproblems.

2   Design Frame

Eclat is a program to find frequent item sets (also closed and maximal as well as generators) with the Eclat algorithm [2], which carries out a depth first search on the subset lattice and determines the support of item sets by intersecting transaction lists.

The design frame is showed in Figure 1. Given the data which could be a set of graphs, sequences, and transactions, after a user giving inputs which include the type of problem to solve, the matching pattern, the constraints of the problem, and the dominance.

- Use gSpan [1] or eclat [2] to get frequent patterns (frequent graphs or itemsets).

- Use Parser method to get python objects from frequent patterns in text.

- Given these frequent graphs or itemsets and user inputs (especially constraints), use IDP [3] to find results (such as, closed itemsets) from python objects.

References

[1] X. Yan and J. Han, "gspan: Graph-based substructure pattern mining," in *Data Mining, 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on.* IEEE, 2002, pp. 721–724.

[2] M. J. Zaki, S. Parthasarathy, M. Ogihara, W. Li *et al.*, "New algorithms for fast discovery of association rules." in *KDD*, vol. 97, 1997, pp. 283–286.
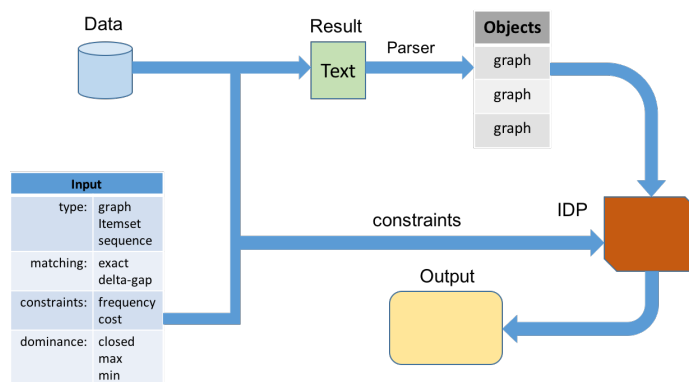
Figure 1: Design frame.

[3] B. De Cat, B. Bogaerts, M. Bruynooghe, and M. Denecker, "Predicate logic as a modelling language: The idp system," 2014. [Online]. Available: dtai.cs.kuleuven.be/krr/software/idp