An explicitly modeled algorithm for mining frequent itemsets in MDE settings

T. Leys , I. Dávid , C.G. Gomes , H. Vangheluwe

Content

1. Motivation
   a. Recommender systems
   b. Frequent itemset mining
   c. Limits of frequent itemset mining
   d. Explicit modeling
2. Project
   a. Transaction formalisms
   b. Explicitly modeled algorithm
3. Motivating example
   a. Visual formalism for floor tiles
   b. Mining models
   c. Formalisms

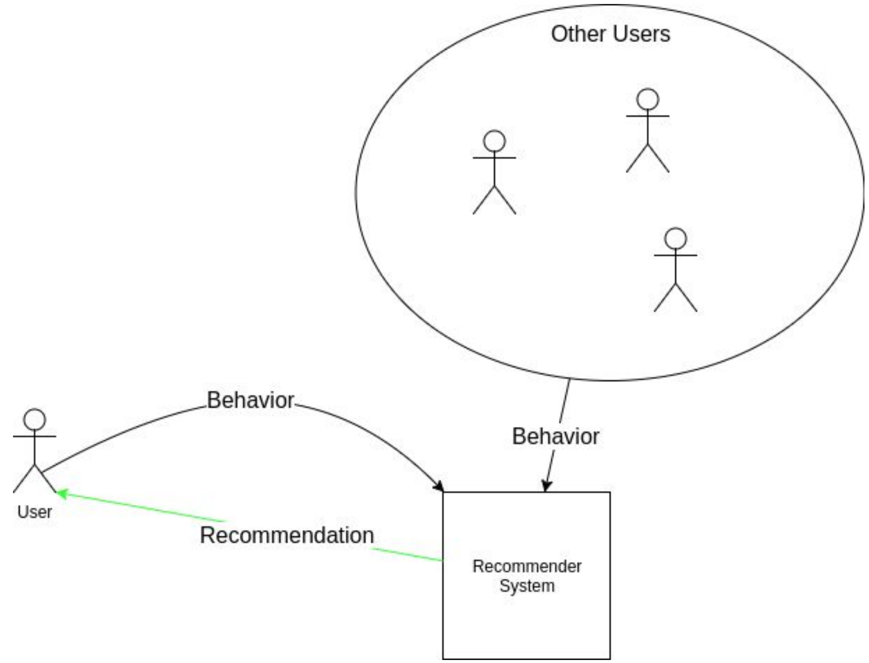
Motivation

Recommender systems

- Recommend a user items, he is likely interested in
- Inspect past behavior of users, give recommendations based on current behavior
- Association rule mining and frequent itemset mining
- Used in applications such as amazon.com, netflix, youtube
- Modeling environments can benefit from it. [1]

[1] Andrej Dyck, Andreas Ganser, and Horst Lichter. Model recommenders for command-enabled editors. MDEBE2013, 2013.

Recommender systems



Frequent Itemset mining

- Apriori [3]
- Eclat [4]
- FP-growth [5]
- CATS-trees [6]

[3] Rakesh Agrawal, Ramakrishnan Srikant, et al. Fast algorithms for mining association rules. In Proc. 20th int. conf. very large data bases, VLDB, volume 1215, pages 487–499, 1994.

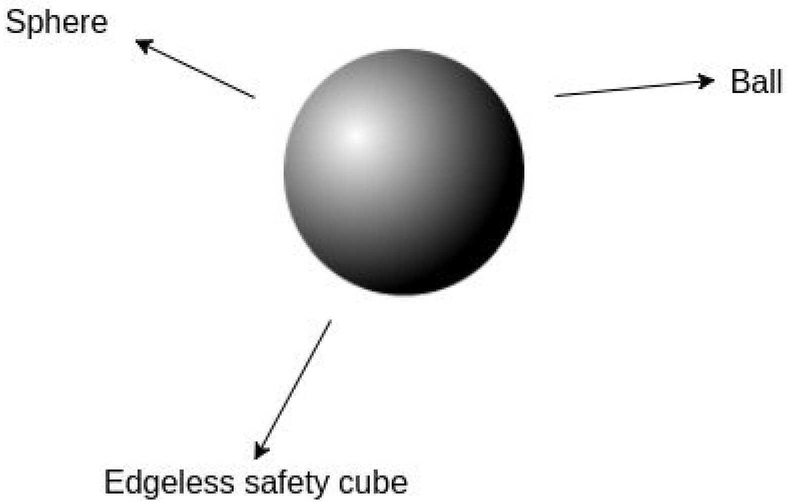
[4] Mohammed Javeed Zaki. Scalable algorithms for association mining. IEEE Transactions on Knowledge and Data Engineering, 12(3):372–390, 2000.

[5] Jiawei Han, Jian Pei, and Yiwen Yin. Mining frequent patterns without candidate generation. In ACM sigmod record, volume 29, pages 1–12. ACM, 2000.

[6] William Cheung and Osmar R Zaiane. Incremental mining of frequent patterns without candidate generation or support constraint. In Database Engineering and Applications Symposium, 2003. Proceedings. Seventh International, pages 111–116. IEEE, 2003.

Limits of frequent dataset mining

- Algorithms only work on sets of strings.
- Code generations
    - Can be hard to understand
    - Not that easy to reason about
    - Prone to inconsistencies
- Explicit modeling and model transformations



Explicit modeling

- Meta-Modeling: explicit specification of a language’s well-formedness constraints [2]
- Advantages:
    - the specification is not hidden in the code of a tool, making it easier to understand and correct
    - the specification can be altered by users of the tool instead of requiring a new tool release
    - one can reason about the specifi- cation and the models it describes
- Explicit modeling of algorithms

[2] Thomas Kühne, Gergely Mezei, Eugene Syriani, Hans Vangheluwe, and Manuel Wimmer. Explicit transformation modeling. In International Conference on Model Driven Engineering Languages and Systems, pages 240–255. Springer, 2009.

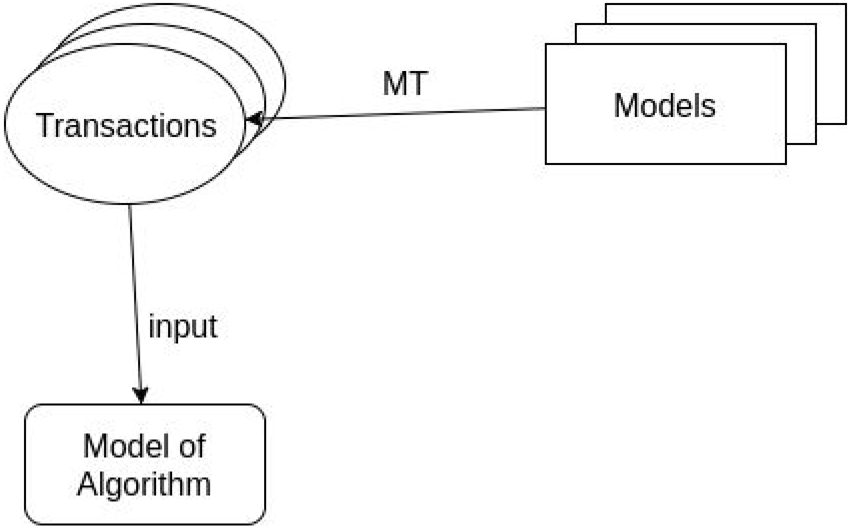
Project

Transaction formalisms

- Modeling a formalism in which we can express transactions
- Formalism can handle visual and textual items
- Model transformation from model to transaction

Explicitly modeled algorithm

- Defining a formalism and model transformation to model an algorithm
- We will model the frequent itemset algorithm



Motivating Example

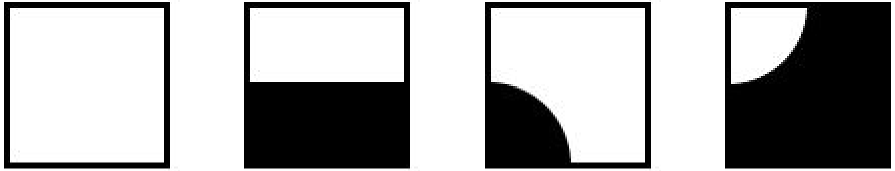
Visual formalism for floor tiles

- As example we will consider a tile manufacturer/retailer
- Create a formalism that:
  - Can model the floorplan of a client
  - Can model the tile placement
- Advantages:
  - A user can make a more precise estimation of the amount of tiles
  - A user can view the result

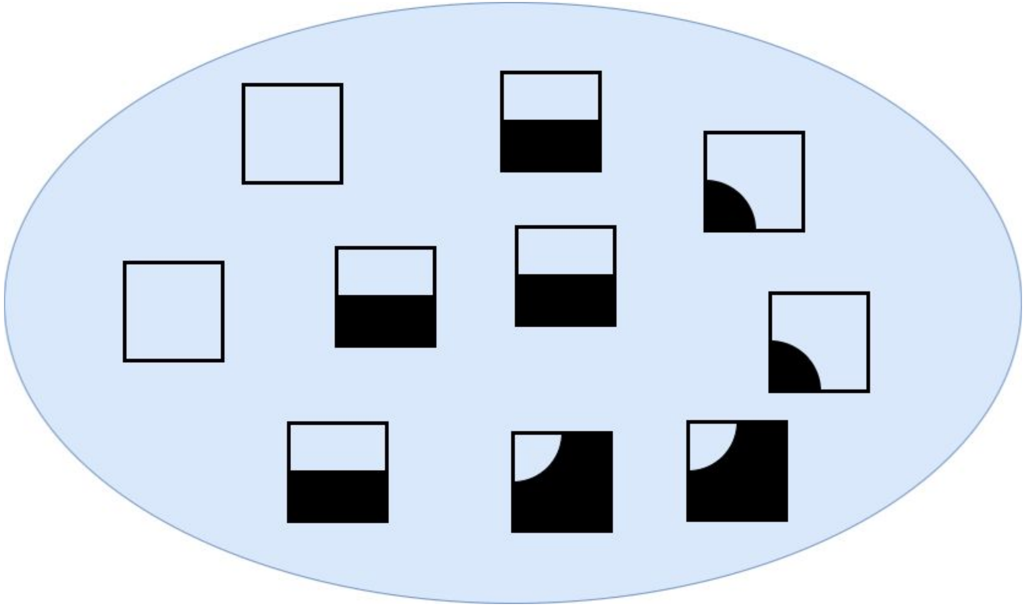
Mining models

- We can mine the database of models
- Mine which tiles are used:
    - Determine association rules
    - Give recommendations to a user about which tiles he can use
- Mine components:
    - Which sets of components appear frequently in orders
    - Proactively and automatically adjust manufacturing process
        - Industry 4.0

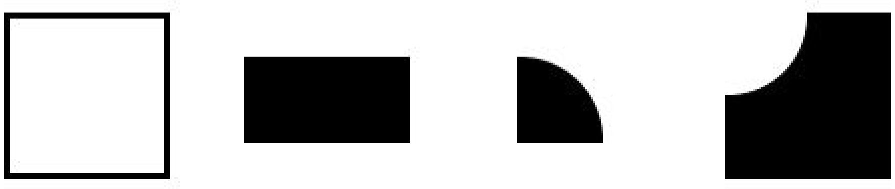
Concrete syntax of the tile formalism



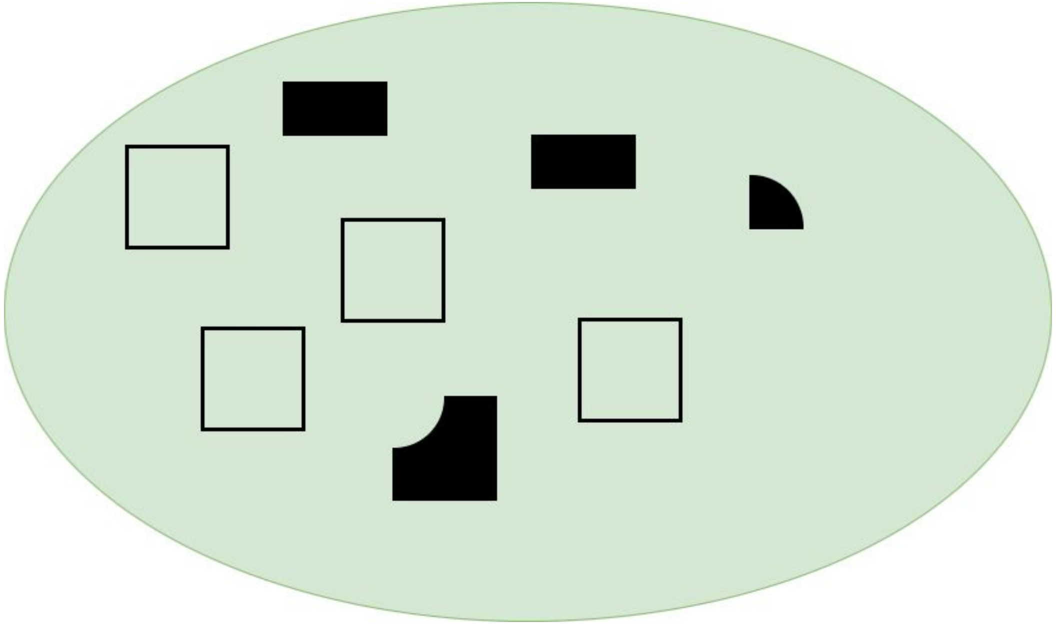
Transaction model of tiles



Concrete syntax for the component formalism



Transaction model of components



Any Questions?