05 Relational data model with infinite flexibility

Finding groups through connection relationships

Abstract

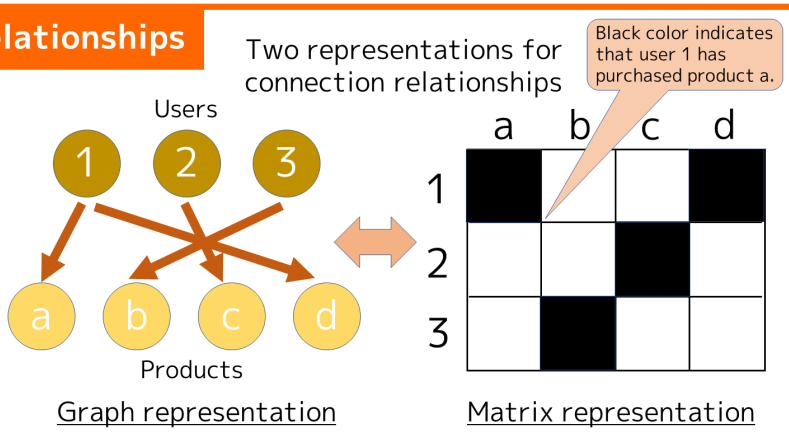
Relational data, including network graphs, such as the connection relationships of users in social networking services and the purchase histories of users for products, appear all around us. In this research, we aim to find the hidden groups in the relational data. In general, when finding groups in relational data, it is difficult to set the number and size of groups appropriately by hand. Therefore, we propose a relational data analysis method that has the ability to automatically adjust the number and size of groups in a data-driven manner according to the size and nature of the input data. By finding appropriate groups in larger relational data, we will contribute to the development of technology for more efficient information storage, search, and retrieval.

Finding groups through connection relationships

- What is connection relationships:
  - Networks that connect people to people,
  - Purchase history of users for products.
- What is groups in connection relationships:
  - Groups of people who are strongly connected each other,
  - Specific group of products that certain people prefer to buy.

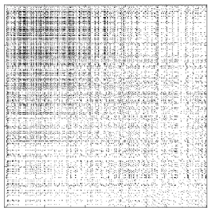
Applications :
- The discovery of potential clusters in a community to prevent the spread of infectious diseases.
- Product recommendation system to present products with high probability of purchase.

Two representations for connection relationships

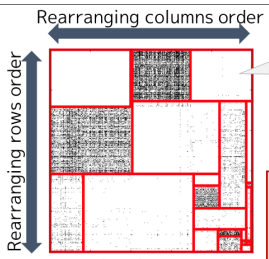


Rectangular cluster extraction from connection matrix:

The size of the input matrix is generally unknown, and it may have an **infinite number of rows and columns**.



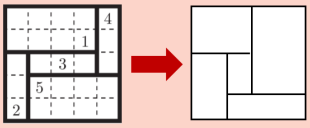
Input data



Analysis result

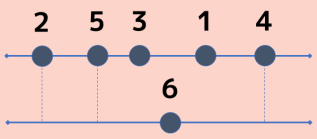
For all possible rectangular partitions **(with infinite number of cases)**, we wish to infer the high degree of fitness to the input data.

Contribution: Modeling the stochastic transformation from floorplan partitions to rectangular partitions



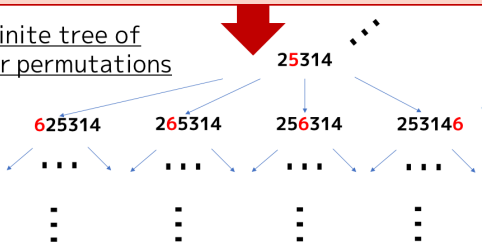
Floorplan partitions (that ignore the size of rooms) Rectangular partitions (that considers the size of rooms)

Contribution: Construction of a generative probabilistic model for the infinite tree of Baxter permutations [Baxter, 1964].



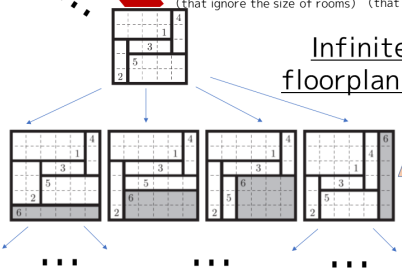
Surjection [Reading,2012]

Infinite tree of Baxter permutations



Bijection [Hong+,2000]

Infinite tree of floorplan partitions



Floorplan partitions have a one-to-one correspondence with Baxter permutations on the diagonal.

References

[1] M. Nakano, A. Kimura, T. Yamada, N. Ueda, "Baxter Permutation Process," in *Proc. Advances in Neural Information Processing Systems 33 (NeurIPS)*, 2020.

Contact

Masahiro Nakano / Recognition Research Group, Media Information Laboratory
Email: cs-openhouse-ml@hco.ntt.co.jp