

店舗横断検索のための関係代数拡張の一考察

清光 英成

神戸大学大学院国際文化学研究所

1 はじめに

電子商店街の店舗業態はそのほとんどが小売である。小売という業態は、ユーザが求める商品を扱うことはもとより、商品の仕様や他の関連製品との位置づけのような商品の情報ならびに利便性を提供することで対価を得る。電子商店街においてユーザは電子商店街が用意した検索システムにより目的の商品を見つけることができる。また、単一の商品に対してある側面（価格や評価など）から整列することができる。これに加えて、「複数種類の商品を同じ店舗で購入したい」あるいは、「指定した店舗数で購入したい」といった要求を満足するような整列や検索結果の表示方法が提供できれば有用である。そこで、本論文はこのような要求に応えるための検索方法を関係代数の拡張により提案する。

2 データの表現と構造

商品には商品を唯一識別するためのデータとして品番や型番などの表現方法がある。また、商品の種類や性質、特徴などを端的に表現するデータとして商品名がある。本論文で扱う商品、店舗、ユーザが所望する商品 (want list: W)、電子商店街における商品を扱う店舗と価格情報など (M) のリレーション表現を以下のように定義する。

商品=item_ID, item_name,

店舗=store_ID, store_name, sh_fee, free_sh, ...

W =item_ID, priority, ...

M =item_ID, store_ID, price, qty, ...

商品購入支援のための演算法についての本質的な議論をするために、便宜上、 W 内の item_ID と M の item_ID は同一のドメイン (domain) の要素とする。

3 関係代数を用いた演算法

リレーション W 内のすべての商品を単一の店舗から購入するために、購入店舗候補集合を求める。 M の射影 $M[\text{item_ID}, \text{store_ID}]$ と W の射影 $W[\text{item_ID}]$ との商

$$(M[\text{item_ID}, \text{store_ID}]) \div (W[\text{item_ID}])$$

が、購入店舗候補集合を表現するリレーションである。

つぎに、任意の店舗数から W 内の商品を購入する方法を考察する。つまり、商品 k を店舗 l で購入するという組の集合の集合が商品-購入店舗候補集合となる。これは、属性値にリレーションを持つようなリレーション表現となる。いわゆる、関係値属性 (Relation-Valued Attributes) を属性として持つリレーションである [1]。関係値属性を扱う問題はデータベースモデルの研究課題として多くの挑戦がなされてきたが、ここでは商品とその商品を扱う店舗の組の集合を表現できればよく、ドメインの異なる他の属性集合を扱う必要がないため、関係値集合をリレーションの属性集合に展開する。商品をどの店舗で購入するかを表現するための属性に、ドメインを店舗 ID、購入する商品と対応する商品 ID を属性名とするリレーション

$$T = \{\text{item}_1, \dots, \text{item}_m\}$$

をリレーション W とリレーション M から導出する。item_1 ($1 \leq i \leq m$) を扱う店舗を表現するリレーションは、制限と射影

$$(M[\text{item_id} = \text{'item_i'}])[\text{store_ID}]$$

によって導出される。これは、リレーション T の射影 $T[\text{item}_i]$ である。ここで、 $T[\text{item}_i]$ の属性値は商品 i を販売している店舗の店舗 ID である。さらに、リレーション T の属性名を $\text{item}_1, \dots, \text{item}_m$ と記述しているが属性集合であるため順序はない。

3.1 拡張の利用

リレーション T の各タプルは W 内の商品をどの店舗で購入するかのパターンを表現している。リレーション T の全属性集合をキーとして、そのパターンにおける店舗集合を求めるために関係代数の拡張 (extend) を用いる。まず、リレーション T における商品 i を購入する店舗の属性表現 item_i を拡張属性 v_stores として追加する。

$$\text{EXTEND } T \text{ ADD (item_i AS v_stores)}$$

得られたリレーションを E_i とする。 E_i 中の各タプルにおける拡張属性 v_stores の属性値と item_i の属性値は等しい。 T のすべての属性に対して上記の拡張で得られた E_1, \dots, E_m の和 E を

A Study for Store Search over a Virtual Shopping Mall by an Extension of Relational Algebra
Hidenari KIYOMITSU,
Graduate School of Intercultural Studies, Kobe University

$$E = E_1 \cup \dots \cup E_m$$

とする。

3.2 要約の利用

関係代数の応用的な演算として要約 (summarise) がある。合計 (sum), 平均 (average), 最大 (max), 最小 (min), 数え上げ (count) の機能は多くのデータベース管理システムで実装されている。しかしながら、要約の対象は対象となるリレーションのある一属性に対してドメインが数値であれば合計, 平均, 最大, 最小を求めることができ, 数え上げはタプルの数を求めることができる。リレーション E の属性集合 $\{item_1, \dots, item_m\}$ における要約を用いて W 内の商品をどの店舗で購入するかの各パターンにおける店舗数を数え上げによって属性 NoS (Number of Stores) の値として得る。

```
SUMMARISE E PER E{item_1, ..., item_m}
      (ADD COUNT(v_stores) AS NoS)
```

上記によるリレーションを E' とする。リレーション E' は

$$E' = \{item_1, \dots, item_m, NoS\}$$

という属性集合からなり, 各タプルは各属性値で当該商品をどの店舗で購入するかのパターンとそのパターンでの購入店舗数を表す。

3.3 提案演算

上述のような二つのリレーション M' , W'

$$M' = M[item_ID, store_ID]$$

$$W' = W[item_ID]$$

を入力としてリレーションを T を出力する演算を,

$$T = M' \bowtie_{[store_ID]} \widehat{W'}$$

と書くことにする。まず単一の属性からなる W' を定値リレーションを属性として持つ入れ子リレーションととらえて W' を分解する。 W' 中の各定値リレーションを $w'_i \in W'$ ($1 \leq i \leq m$) とする。リレーション M' と定値リレーション w'_i との自然結合を

$$M'_i = M' \bowtie w'_i$$

とすると, T は各 M'_i の射影 $M'_i[store_ID]$ の直積

$$T = M'_1 \times \dots \times M'_m$$

で求めることができる。このとき, リレーション T は

$$T = \{ M'_1.store_ID, \dots, M'_m.store_ID \}$$

という属性集合からなる。

W 内の商品をどの店舗で購入するかのパターンを表現するために, リレーション T を求めたが, $W' = W[item_ID]$ を定値リレーションを属性として持つ入れ子リレーションととらえて W' を分解した。直観的には W' の行と列とを入れ替えたようなものであった。リレーション T は商品とその商品を扱う店舗の組の集合を表現できればよく, ドメインの異なる他の属性集合を扱う必要がないため, すべての属性のドメインが等しいリレーションがである。このため, T の全属性は θ 比較可能である。

リレーション T は, リレーション M' と W' から合成された時点で厳密な意味ではデータベースから切り離されているが, 用途は wantlist 中の商品をどの店舗で購入すればユーザの要求を満たすことになるのかを大まかに推薦する二次データの生成である。これは, 精密なリアルタイム性が不要という意味ではなく, 若干の許容度がある程度の意味である。その時点でのスナップショットではこのような購入店舗候補集合を求めることができ, 店舗数の数え上げができるように拡張しようとしているのである。逆説的に, T が表現する購入店舗候補集合ビューをもとにしてデータベースを更新することを想定しないので, 実装においてはタプルの複製を外部で計算して計算結果の整数値を拡張演算を利用して要約演算を利用してよい。また, Ruby などの集合のサイズを計算する関数を持つ言語をベースとしたデータ操作言語を持つデータベース管理システムや KVS(Key Value Store) であれば, 拡張と要約が同時に記述できる。

4 まとめ

ユーザが購入したい商品を want list に登録し, ユーザ所望の条件で購入支援を行うための演算法を考察した。wantlist 内の商品をどこで買うべきかというユーザの意思 (行動) 決定を支援することを目的として複数種類の商品を同じ店舗で購入したい, あるいは指定した店舗数で購入したいといった要求を満たすような商品と店舗の組の集合を検索する方法について議論した。本稿では特に, m 個の商品を n 店舗で購入するという条件を満たすための演算法を関係値属性を持つリレーションの属性集合への展開をへて関係代数の拡張演算並びに要約演算を用いて示した。

本論文で求めた購入店舗候補集合は, 各店舗候補の買い物カゴに変換されてユーザに評価される。ユーザが実際に購入するのは, 買い物カゴに商品が登録された後のしばらく後であることが想定されるため, スナップショットとの厳密な整合性は保証できないし, 本方式では購入店舗候補集合が合成された時点で厳密な意味ではデータベースから切り離されている。しかしながら, 繰り返しにはなるが, 用途は wantlist 中の商品をどの店舗で購入すればユーザの要求を満たすことになるのかを大まかに推薦する二次データの生成である。実データとの整合性よりもむしろ, 本研究は情報推薦サービスを目的としており, 購入店舗候補集合は数的に十分な絞り込みの方が要求される。限られた購入店舗候補集合から買い物カゴの中身のパターンをユーザの要求にマッチさせるための購入店舗候補集合の評価と順位付けは今後の課題である。

参考文献

- [1] HJ. Schek, MH. Scholl :The relational model with relation-valued attributes, Information Systems, Vol.11, No. 2, pp. 137-147, Pergamon-Elsevier, 1986
- [2] C.J. Date, H. Darwen : Database, Types, and the Relational Model: the third Manifesto, third edition, Adison-Wesley, 2006
- [3] 楊斐, 清光 英成, 大月 一弘, 森下 淳也, Want list を用いた購買支援と情報推薦, 情報処理学会研究報告データベースシステム (DBS), Vol. 2011-DBS-153, No. 15. pp. 1-8, 2011.
- [4] 清光 英成, Wantlist を用いた商品検索のための演算法考察-店舗集合候補提示のための演算を例として-, 情報処理学会研究報告データベースシステム (DBS), Vol. 2013-DBS-158, No. 18, pp. 1-4, 2013.