

**안티모노톤의 속성**, 또는 법칙이라고 불리는 규칙이다.

이 용어는 데이터 마이닝에서 패턴을 찾아낼 때 패턴을 더 조합해봐야 더 이상 빈발하지 않을 것을 알 수 있을 때 더 이상 패턴을 조합하는데 자원을 낭비하지 않고 끊어버림으로써 실행시간에 도움이 되는 상황에 사용한다.

“

어떤 패턴  $P$ 가 빈발하지 않으면 그를 포함하는 모든 수퍼패턴  $P'$ 도 빈발하지 않다

”

기본적인 정의는 위와 같다. 빈발한 패턴이라는 것은 어떤 패턴, 예를 들어 { 사과, 바나나 } 같은 패턴이 트랜잭션 데이터베이스에서 많이 나오는 것을 의미한다. 그런데 { 사과, 바나나 } 라는 패턴이 DB에서 적게 나왔다고 생각해 보자. 그러면 이 패턴을 포함하는 수퍼패턴은 더 긴 패턴을 의미하는데, 이런 패턴이 나오는 수는 이전 패턴의 수와 같거나 많아봐야 같은 수일 것이다. { 사과, 바나나, 수박 }이라는 패턴은 덜 나오거나 적게 나올게 분명한 소리다.

즉 어떤 짧은 서브패턴이 자주 나오지 않는다고 알게 되면 이 서브패턴에서 아이템이 더 붙어서 나오는 모든 수퍼패턴은 모두 자주 나오지 않을거라는 사실을 알 수 있다. 그래서 이런 상황에서 더 깊게 마이닝을 하며 패턴을 찾지 않고 모두 잘라내버릴 수 있다. 이러한 속성을 **안티모노톤의 속성**이라 한다. 다른말로 **downward closure property**라고도 한다.

하지만 모든 패턴마이닝에서 이 속성이 적용되는 아니다. 이 속성은 모든아이템에 가중치 없이 동일한 패턴마이닝에서만 적용된다. 즉 어떤 아이템이 패턴에 붙음으로써 패턴의 빈발도가 갑자기 확 높아질 수 있는 상황이 발생한다면 이 방식을 적용할 때 유의미한 패턴의 손실을 불러일으킨다. 하지만 이 속성은 패턴마이닝의 성능향상에 큰 도움이 된다.

## 2. Generating all possible rules from the frequent itemsets

Once the frequent itemsets are generated, identifying rules out of them is comparatively less taxing. Rules are formed by binary partition of each itemset. If {Bread,Egg,Milk,Butter} is the frequent itemset, candidate rules will look like:

(Egg, Milk, Butter  $\rightarrow$  Bread), (Bread, Milk, Butter  $\rightarrow$  Egg), (Bread, Egg  $\rightarrow$  Milk, Butter), (Egg, Milk  $\rightarrow$  Bread, Butter), (Butter  $\rightarrow$  Bread, Egg, Milk)

From a list of all possible candidate rules, we aim to identify rules that fall above a minimum confidence level (*minconf*). (Just like the anti-monotone property of support, ~~confidence of rules~~ generated from the same itemset also follows an anti-monotone property. It is anti-monotone with respect to the number of elements (in consequent).)

결과부의 개수가 늘어남, 신뢰도가 낮아짐

This means that confidence of (A,B,C  $\rightarrow$  D)  $\geq$  (B,C  $\rightarrow$  A,D)  $\geq$  (C  $\rightarrow$  A,B,D). To remind, confidence for {X  $\rightarrow$  Y} = support of {X,Y}/support of {X}

(As we know that support of all the rules generated from same itemset remains the same) and difference occurs only (in the denominator calculation of confidence). (As number of items in X decrease,) support{X} increases (as follows from the anti-monotone property of support) and hence the confidence value decreases.