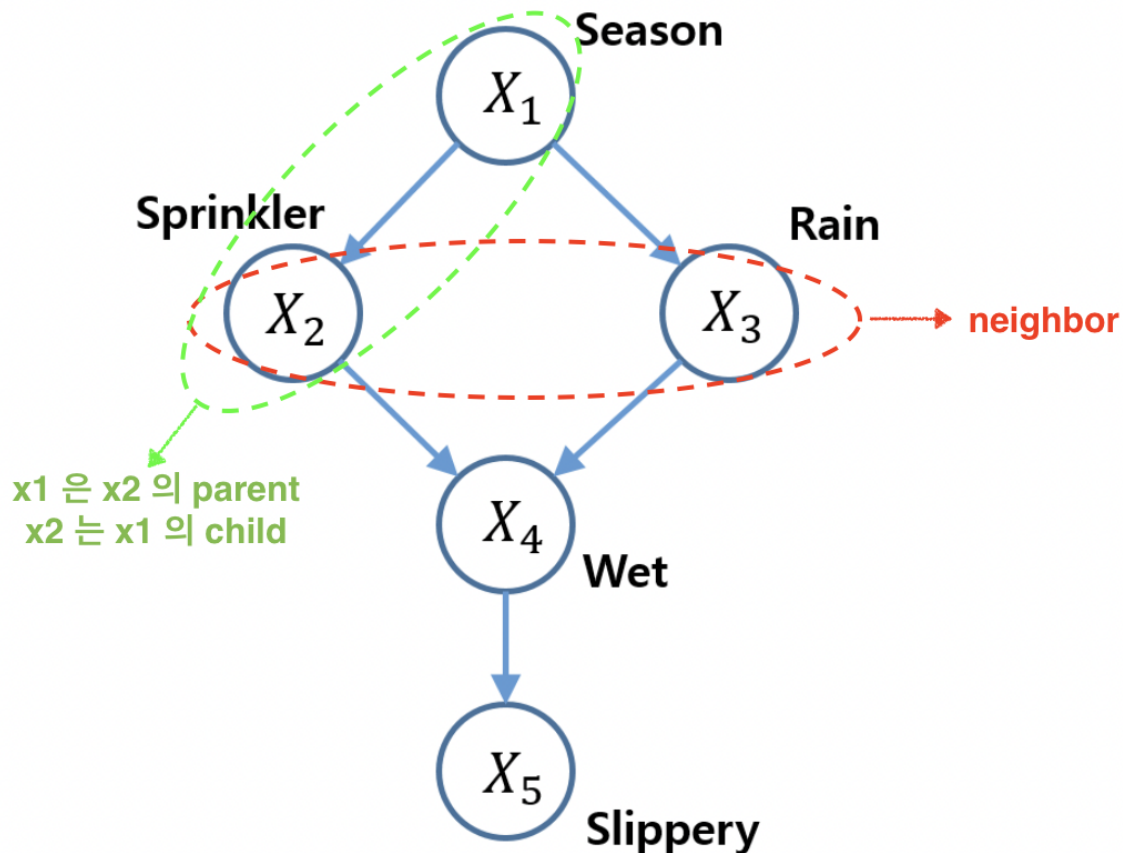


D-Separation

Recap. Bayesian Network



- Bayesian Network : 방향성을 갖는 비순환 그래프와 조건부 확률 테이블을 통하여 집합을 조건부 독립으로 표현하는 확률의 그래픽 모델
- $P(x_1, x_2, x_3, x_4, x_5) = P(x_1)P(x_2|x_1)P(x_3|x_1)P(x_4|x_2, x_3)P(x_5|x_4)$
 - **conditional distribution** 을 방향성 link 로 표현하고 있음
 - $P(x_1)$ 자체가 링크가 아닌 노드로 표현되고 있음. 즉, 하나의 노드는 random variable 이 되는 것
 - $x_1 \rightarrow x_2$ 와 같이 화살표로 표시하는데 이는 x_2 의 부모 노드로 위에서 $P(x_2|x_1)$ 과 같이 표시되고 있음.
- 위와 같은 그래프는 fully connected graph 가 아니기 때문에 일반화 한 식은 아래와 같음

$$p(x) = \prod_{k=1}^K p(x_k|pa_k)$$

- $p(a_k)$ 는 x_k 의 parents node 의 집합

- 이는 joint distribution P 가 Bayesian network G 로 표현이 가능하다는 것이고, 이 경우 P 가 G 에 대해 *Markov* 하다고 말하고, G 와 P 는 서로 *Compatible* 하다고 할 수 있다.
- 방향성 그래프의 중요한 제약은 cycle 이 형성되지 않는 것이고, 이런 그래프를 DAG(directed acyclic graph) 라고 부른다. (이전에 발표했었던 DAG-SVM 또한 Directed Acyclic Graph 임)

Separation

- 어떤 path 가 다른 node 에 의해 경로가 막힌 경우를 의미
 - 그래프의 edge 가 화살표가 아닌 undirected edge 인 경우,



- j 가 i 와 k 사이의 path 를 **separate** 하고 있다고 할 수 있음.
- But! Bayesian Network 는 Directed Graph 이기 때문에 여기서의 separation 개념을 정의해야 하는데, 이를 **D-Separation** 이라고 하는 것.

D-Separation

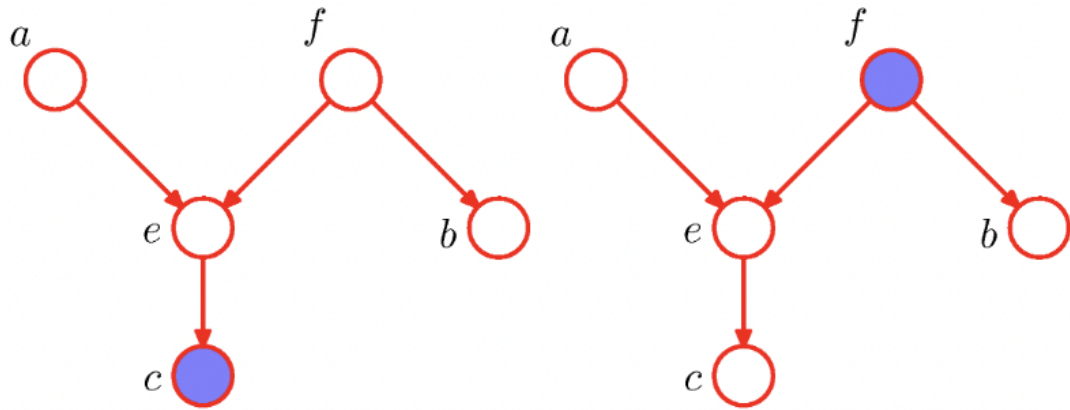
- D-Separation 은 그래프에서 조건부독립인 노드들을 판단할 수 있다.
- Factorization 과 Independent 간의 관계를 생각해보면, 확률분포 P 를 Factorization 한다는 것은 두 변수가 독립적임을 말하는 것이고, 그래프 G 확률 분포에 대한 Factorization 이 있을 때 그래프 내 변수들 사이의 독립성을 어떻게 확인 할 수 있을까?
 - 두 노드 X 와 Y , 그리고 observation Z 가 주어졌을 때, X 와 Y 를 잇는 모든 사슬을 Z 가 막고(block) 있는 경우 이다. (**D-Separation**)
 - $X \perp\!\!\!\perp Y | Z$
 - block 의 정의
 1. X 와 Y 경로의 화살표가 head-to-tail 이나 tail-to-tail 이고, 지나가는 노드 중 일

부가 Z 에 포함되는 경우

2. X 와 Y 경로의 화살표가 head-to-head 이고, 지나가는 노드 중 일부가 Z 노드 집합에 없는 경우

- X 에서 Y 까지 모든 경로가 차단되면, X 와 Y 는 Z 에 의해 D-separated 되었다고 얘기할 수 있다.

- 이 경우 $X \perp\!\!\!\perp Y | Z$ 가 성립한다고 말할 수 있는 것이다.

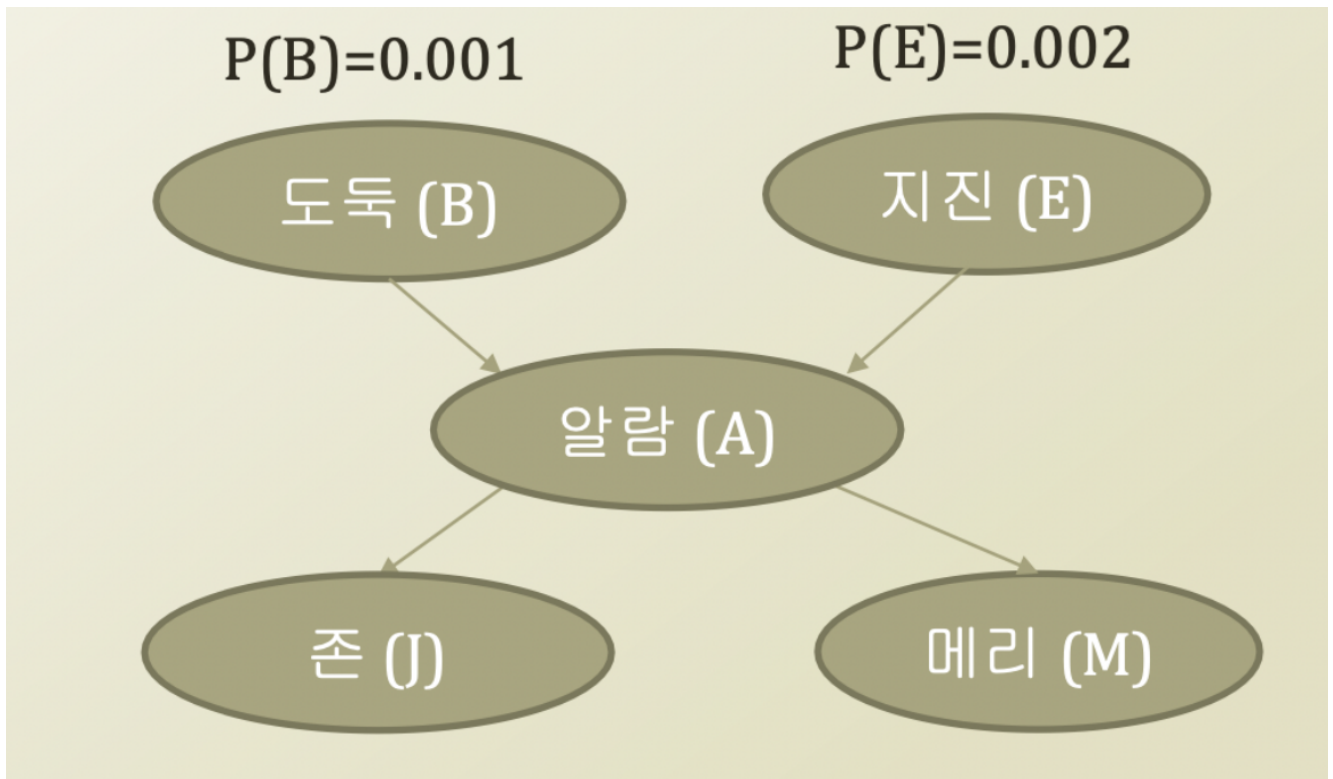


- 예시1. (왼쪽 그림)

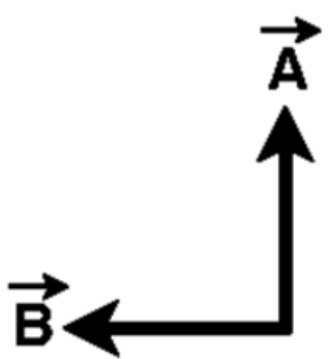
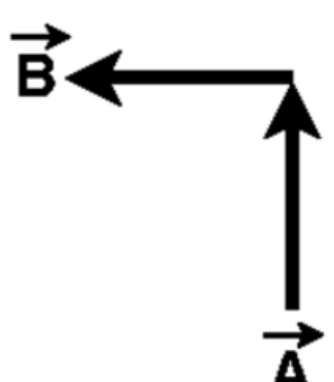
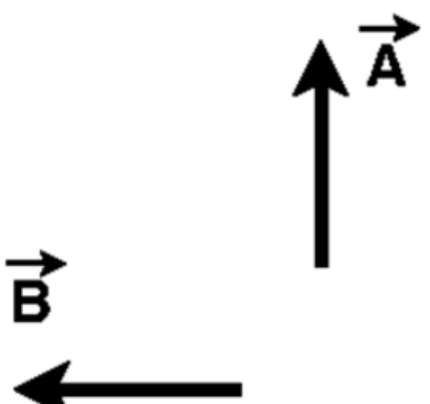
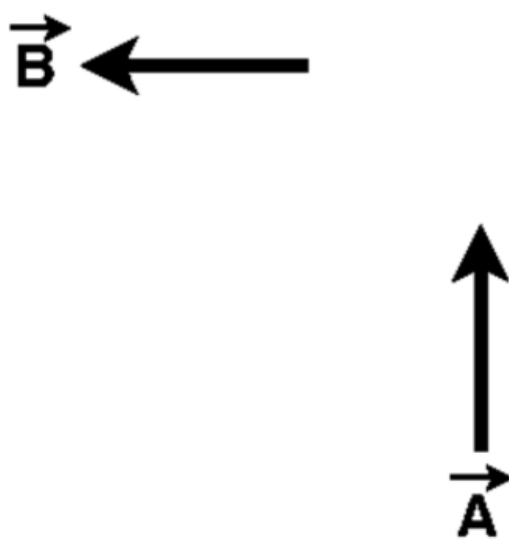
- a → b 경로는 노드 f 에 의해 block 되지 않는다.
 - 이 경로는 tail-to-tail 이고, 관찰되지 않았기 때문(common parents 정의를 생각해보면 된다.)
- e 또한 경로가 block되지 않는다.
 - e 는 head to head 지만, c 라는 자식노드가 고립되어 있는데, 정보를 알고있기 때문에 시작노드를 추정할 수 있다. 이렇게 되면 e 가 간접적으로 관찰 되었다고 할 수 있기 때문에, v-structure 의 정의에 따라 block 되지 않는다.

- 예시 2. (오른쪽 그림)

- a → b 경로는 노드 f 에 의해 block 된다.
- 마찬가지로 e 또한 경로가 block 된다. v-structure 의 정의에 따라 block 되게 된다.



(참고)

	tail to tail	head to tail
together		
separate		

https://www.researchgate.net/figure/The-same-orientation-y-x-z-with-the-vectors-Head-to-tail-Tail-to-tail_fig2_279968301

I-map (Independency map)

- 모든 variable 에 대한 확률 분포 P가 indepent 하는 경우 그래프 G 를 P의 I-map 이라고 정의한다.
 - 그래프 G 의 확률분포 P가 Factorization 된다고 하면 G는 P의 I-map 이다. 왜냐하면, 확률분포 P 를 Factorization 한다는 것은 두 변수가 독립적임을 말하는 것이기 때문이다.

Reference

- PRML 8.2.2
- <https://blueth.tistory.com/7>
 - <https://fabj.tistory.com/18>