# PROBABILISTIC REASONING

Acting under uncertainty

# Agents

- In artificial intelligence, an agent is a computer program or system that is designed to perceive its environment, make decisions and take actions

# Uncertainty

AUTONOMOUS CARS – NEW ROUTE

VACUUM CLEANER – NEW ENVIRONMENT

# Belief states

States + Action [Eg : puncture → A1, break down →A2]

Draw back :

- More belief state
- Contingency plan grow exponentially [same plan not applicable every where]
- No plan yield result → but action must be taken

# Uncertainty example : Einstein to Tirunelveli [catch Nellai express]

01

A30 → start 30 mins before departure of train

02

A45 → start 40 mins before departure of train

03

A120 → start 2 hours before departure of train

04

A1440 → start 24 hours before departure of train

# Uncertainity

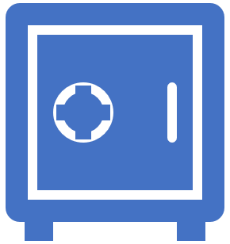
- Traffic jam
- Bad road condition
- Break down
- Train cancel

# Rational decision

- Selecting an option from the actions rationally
- Agent choosing action depends on the situation
- Like going to : excursion, interview
- No guarantee for success

# Agents : Maximize performance measure

On time



Safe

comfortable

# Uncertainty example

- Toothache → cavity
- Toothache → cavity or gum problem or broken tooth etc..

Not Exhaustive

# Change the order

- Cavity → toothache
- Cause → effect [Causal rule]

Not 100% true

# Fails

Laziness → impossible to list all antecedents or consequents

Theoretical ignorance → no complete solution in books

Practical ignorance → even if all rules are available but we may not have results for all test

# Degree if belief (Possibility)

- Represented by probability
- Like A45 → 0.97%
- A1440 → 0.99%

# Agent

1

Logical : True or false

2

Probabilistic : it knows the probabilistic value for the event [0-1]

3

Probabilistic values based on knowledge state

# Utility theory

- Used to represent preferences [tour, interview etc]

# Decision theory

Utility theory + probability theory

If we say the agent is Rational → choose an action to yield maximum utility

# A decision-theoretic agent that selects rational actions.

```
function DT-AGENT(percept) returns an action
   persistent: belief_state, probabilistic beliefs about the current state of the world
               action, the agent's action

   update belief_state based on action and percept
   calculate outcome probabilities for actions,
       given action descriptions and current belief_state
   select action with highest expected utility
       given probabilities of outcomes and utility information
   return action
```

# Probability

- All decisions taken by agents are based on probability

# Throwing a dice

- Total possible outcomes – {1,2,3,4,5,6}
- Event → {5} specific happening or favorable outcome
- Probability of getting 3 → 1 time three is present out of 6 sample space so … (1/6)
- No of elements in favorable/ total number of elements
- $P(n>3)$ → 3/6 = 1/2

# Throwing a dice

- Total possible outcomes – {1,2,3,4,5,6}
- Event → {5} specific happening or favorable outcome
- Probability of getting 3 → 1 time three is present out of 6 sample space so … (1/6)
- No of elements in favorable/ total number of elements
- P(n>3) → 3/6 = ½
- P(3) OR P(4) → 1/6+1/6 = 2/6 =1/3

# Throwing Two dice

- Total possible outcomes
- D1 – {1,2,3,4,5,6}
- D2 – {1,2,3,4,5,6}
- P(sum = 12) → 1/36
- P(sum>10) → 3/36 = 1/12

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | (1, 1) | (1, 2) | (1, 3) | (1, 4) | (1, 5) | (1, 6) |
| 2 | (2, 1) | (2, 2) | (2, 3) | (2, 4) | (2, 5) | (2, 6) |
| 3 | (3, 1) | (3, 2) | (3, 3) | (3, 4) | (3, 5) | (3, 6) |
| 4 | (4, 1) | (4, 2) | (4, 3) | (4, 4) | (4, 5) | (4, 6) |
| 5 | (5, 1) | (5, 2) | (5, 3) | (5, 4) | (5, 5) | (5, 6) |
| 6 | (6, 1) | (6, 2) | (6, 3) | (6, 4) | (6, 5) | (6, 6) |

# Theorems of probability

- Probability values : [0,1]
- Sum of all probabilities

head = ½ tail = ½

Events

$$\sum_{i=0}^{n} p(A_i) = 1$$

Sample Space

# Types of probability

- Unconditional probability – without caring about other events (no information about other events)
- Conditional probability – Dice1 → 6  Dice → ?   [getting sum 11 ]
- E1 i know  I don't have E2 ➔ P(E2/E1)  →**P(E2|E1) = N(E2∩E1)/N(E1)**

**P(A|B) = N(A∩B)/N(B)**

**Or**

**P(B|A) = N(A∩B)/N(A)**

# Types of probability $P(a \mid b) = P(a \wedge b)/P(b)$

Conditional probability – Dice1 → 6  Dice → ?   [getting sum 11 ]

**$P(E2|E1) = P(E2 \cap E1)/P(E1)$**

- P(E1) → dice 1 must be 6 → (1/6)
- P(E2 n E1)→  1/36
- = (1/36) / (1/6)
- = 1/6

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | (1, 1) | (1, 2) | (1, 3) | (1, 4) | (1, 5) | (1, 6) |
| 2 | (2, 1) | (2, 2) | (2, 3) | (2, 4) | (2, 5) | (2, 6) |
| 3 | (3, 1) | (3, 2) | (3, 3) | (3, 4) | (3, 5) | (3, 6) |
| 4 | (4, 1) | (4, 2) | (4, 3) | (4, 4) | (4, 5) | (4, 6) |
| 5 | (5, 1) | (5, 2) | (5, 3) | (5, 4) | (5, 5) | (5, 6) |
| 6 | (6, 1) | (6, 2) | (6, 3) | (6, 4) | (6, 5) | (6, 6) |

# Product rule

- $P(a \land b) = P(a \mid b)\, P(b)$

# Random Variables

- P(Weather = sunny ) = 0.6

Variable

Event

$$P(Weather = sunny) = 0.6$$
$$P(Weather = rain) = 0.1$$
$$P(Weather = cloudy) = 0.29$$
$$P(Weather = snow) = 0.01\ ,$$

but as an abbreviation we will allow

$$\mathbf{P}(Weather) = \langle 0.6, 0.1, 0.29, 0.01 \rangle\ ,$$

probabilities of *all* the possible values of a random variable.

# Random variables domain

- Domain → range of values
- Eg : p(Temp) – 28 to 40 degree Celsius

# Continuous variables domain

- Values keeps on changing. Eg: temperature at a place
- temperature at noon is distributed uniformly between 18 & 26
- This is called probability density function

$$P(NoonTemp = x) = Uniform_{[18C,26C]}(x)$$

$$P(NoonTemp = x) = Uniform_{[18C,26C]}(x) = \begin{cases} \frac{1}{8C} \text{ if } 18C \leq x \leq 26C \\ 0 \text{ otherwise} \end{cases}$$

# Random variables domain

- The random variable domain refers to the set of all possible outcomes of a random experiment or process, regardless of whether the variable is continuous or discrete.
- Domain → range of values
- Eg : p(Temp) – 28 to 40 degree Celsius

# Joint Probability distribution

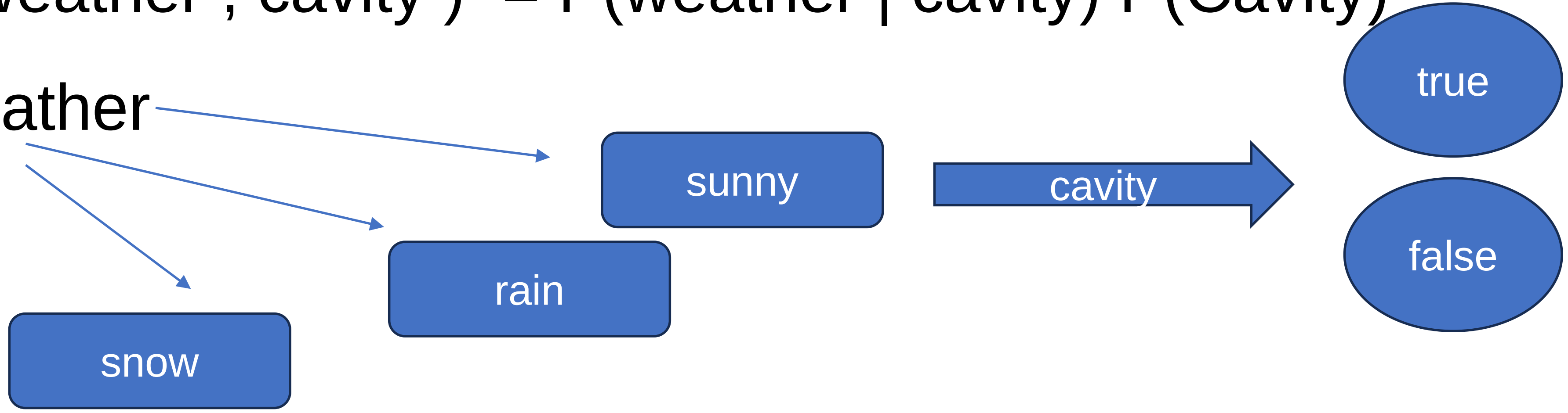
A joint probability distribution describes the probabilities of all possible combinations of outcomes for two or more random variables.

It provides a complete picture of the relationship between multiple variables in a probabilistic setting.

Denoted as $P(X=x,Y=y)P(X=x,Y=y)$ or $P(x,y)P(x,y)$, where X and Y are the random variables, and x and y are specific values of those variables.

# Joint probability distribution

- Cavity : true , false
- Weather : sunny, rain, cloudy, snow
- P(weather , cavity ) = P(weather | cavity) P(Cavity)
- Weather



# Joint probability distribution

- $P(a \mid b) = P(a \land b) / P(b)$
- $P(a \land b) = P(a \mid b) P(b)$
- P(weather and cavity ) = P(weather given cavity) and p(cavity)

# Joint probability distribution

$$
\begin{aligned}
P(W = sunny \wedge C = true) &= P(W = sunny | C = true)\, P(C = true) \\
P(W = rain \wedge C = true) &= P(W = rain | C = true)\, P(C = true) \\
P(W = cloudy \wedge C = true) &= P(W = cloudy | C = true)\, P(C = true) \\
P(W = snow \wedge C = true) &= P(W = snow | C = true)\, P(C = true) \\
P(W = sunny \wedge C = false) &= P(W = sunny | C = false)\, P(C = false) \\
P(W = rain \wedge C = false) &= P(W = rain | C = false)\, P(C = false) \\
P(W = cloudy \wedge C = false) &= P(W = cloudy | C = false)\, P(C = false) \\
P(W = snow \wedge C = false) &= P(W = snow | C = false)\, P(C = false)\ .
\end{aligned}
$$

# Joint probability distribution

- $P(a \mid b) = P(a \land b) / P(b)$
- $P(a \land b) = P(a \mid b) P(b)$
- P(weather and cavity ) = P(weather given cavity) and p(cavity)
- Cavity , weather then 2x4 = 8 possible worlds
- Cavity,Toothache , and Weather , then there are $2 \times 2 \times 4 = 16$ possible worlds

# Joint probability distribution

- $P(a \mid b) = P(a \wedge b) / P(b)$
- $P(a \wedge b) = P(a \mid b)\, P(b)$
- P(weather and cavity ) = P(weather given cavity) and p(cavity)
- Cavity , weather then 2x4 = 8 possible worlds
- Cavity,Toothache , and Weather , then there are $2 \times 2 \times 4 = 16$ possible worlds

# Axioms

- Probability values [0,1]
- $$\sum_{i=0}^{n} p(A_i) = 1$$
- Kolmogorov's axioms
- Inclusion exclusion principle: P (a ∨ b) = P (a) + P (b) − P (a ∧ b) .
- P(¬A)=1−P(A)

Eg : if the probability of raining tomorrow (event A) is 0.3, then the probability of not raining tomorrow (¬A) would be 1−0.3=0.71−0.3=0.7.

# Inference using full joint distribution

- Computation of posterior probability
- Cavity = t or f , catch = t or f, tooth ache = t or f. so 2x2x2 = 8

| | $toothache$ | | $\neg toothache$ | |
|---|---|---|---|---|
| | $catch$ | $\neg catch$ | $catch$ | $\neg catch$ |
| $cavity$ | 0.108 | 0.012 | 0.072 | 0.008 |
| $\neg cavity$ | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

# Inference using full joint distribution

| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P(¬catch)= 0.012 + 0.064 + 0.008 + 0.576
- P(cavity)  = 0.108 + 0.012 + 0.072 + 0.008
- P(catch)= ?

# Inference using full joint distribution

| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P(cavity) = 0.108 + 0.012 + 0.072 + 0.008

# Inference using full joint distribution

| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P(¬catch)= 0.012 + 0.064 + 0.008 + 0.576
- P(catch)= ?

# Inference using full joint distribution

| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P(catch ∨ tooth ache) = 0.108 + 0.016 + 0.072+ 0.144 + ~~0.108 + 0.016~~ + 0.012 + 0.064
- so we have learned un conditional probability

# Conditional probability

| | *toothache* | | $\neg$*toothache* | |
|---|---|---|---|---|
| | *catch* | $\neg$*catch* | *catch* | $\neg$*catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| $\neg$*cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P(cavity | tooth ache) = ? formula

# Conditional probability

| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P(cavity | tooth ache) = p(cavity ∧ tooth ache ) / p(tooth ache)
- = (0.108+0.012) / (0.108 + 0.016 + 0.012 + 0.064)
- P(¬ Cavity | tooth ache ) = p(¬ Cavity tooth ∧ ache ) / p(tooth ache)

# Conditional probability

| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P(¬ Cavity | tooth ache ) = p(¬ Cavity tooth ∧ ache ) / p(tooth ache)
- (0.016 + 0.064 ) / (0.108 + 0.016 + 0.012 + 0.064)

# Conditional probability

| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P(cavity | tooth ache ∨ catch ) = P(cavity ∧ (tooth ache ∨ catch)) / P(tooth ache ∨ catch)
- = 0.108 + 0.012 + 0.072 / (0.108+0.012+0.016+0.064+0.072+0.144)

# Marginal probability

| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- P (cavity) = 0.108 + 0.012 + 0.072 + 0.008
- adding the entries in the first row gives the unconditional or marginal probability

# conditioning

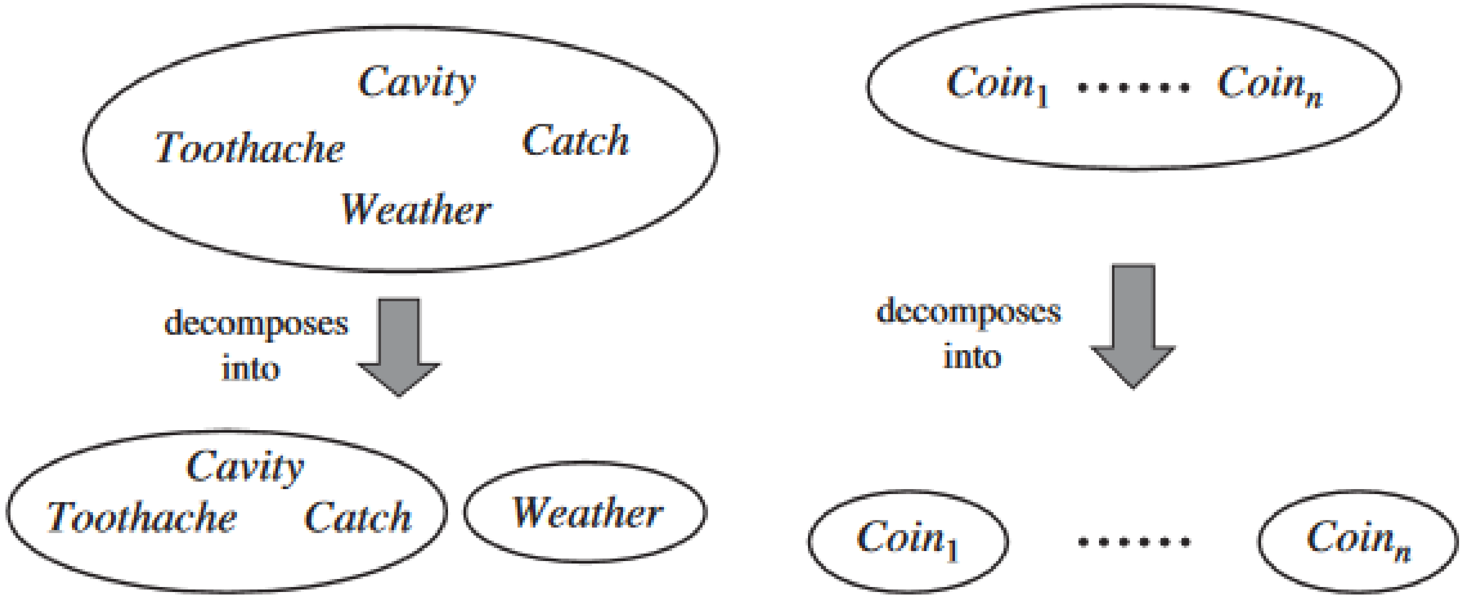
| | *toothache* | | *¬toothache* | |
|---|---|---|---|---|
| | *catch* | *¬catch* | *catch* | *¬catch* |
| *cavity* | 0.108 | 0.012 | 0.072 | 0.008 |
| *¬cavity* | 0.016 | 0.064 | 0.144 | 0.576 |

A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- $P(Y) = \sum_{z} P(Y \mid z)\, P(z)$

# Independence

- Probability of occurrence of one event doesn’t affect other



# Bayes rule or law or theorem

- $P(a \wedge b) = P(a \mid b)\, P(b)$ and $P(a \wedge b) = P(b \mid a)\, P(a)$
- $P(b \mid a) = P(a \mid b)\, P(b) / P(a)$
- $P(Y \mid X) = P(X \mid Y)\, P(Y) / P(X)$
- P (effect | cause) quantifies the relationship in the causal direction
- P (cause | effect) describes the diagnostic direction

# Bayes rule or law or theorem

- P (cause | effect) = P (effect | cause) P (cause) / P (effect)
- P(dengue | fever ∧ cold) = ?
- = P ( fever | dengue ) P(dengue) / P (fever) + P ( cold | dengue ) P(dengue) / P (cold)

# Bayes rule or law or theorem

- For example, a doctor knows that the disease meningitis causes the patient to have a stiff neck, say, 70% of the time.
- The doctor also knows some unconditional facts: the prior probability that a patient has meningitis is 1/50,000, and the prior probability that any patient has a stiff neck is 1%.
- Letting s be the proposition that the patient has a stiff neck and m be the proposition that the patient has meningitis,

# Bayes rule or law or theorem

- P (s | m) = 0.7
- P (m) = 1/50000
- P (m | s) = P (s | m) P (m) / P (s)
- = (0.7 × 1/50000 ) / 0.01
- = 0.0014 .
- That is, we expect less than 1 in 700 patients with a stiff neck to have meningitis

# Bayes rule or law or theorem

- instead computing a posterior probability for each value of the query variable (here, m and ¬m) and then normalizing the results. The same process can be applied when using Bayes' rule
- Normalised form : $P(M \mid s) = \alpha \, (P(s \mid m)P(m), P(s \mid \neg m)P(\neg m))$
- The general form of Bayes' rule with normalization is

$$P(Y \mid X) = \alpha \, P(X \mid Y)P(Y)$$

# Bayes rule or law or theorem

- How ?
- $P(m|s) = P(s|m)\ P(m) / P(s) \rightarrow \alpha\ P(s|m)\ P(m)$ [meningtis = true]
- $P(\neg m|s) = P(s|\neg m)\ P(\neg m) / P(s) \rightarrow \alpha\ P(s|\neg m)\ P(\neg m)$

  [meningtis = false]

# Bayes rule or law or theorem

- If there is a sudden epidemic of meningitis, the unconditional probability of meningitis, P (m), will go up
- The doctor who derived the diagnostic probability P (m | s) directly from statistical observation of patients before the epidemic will have no idea how to update the value
- but the doctor who computes P (m | s) from the other three values will see that P (m | s) should go up proportionately with P (m).
- The causal information P (s | m) is unaffected by the epidemic, because it simply reflects the way meningitis works

# Combining evidence

- how we combine multiple evidence ?

- P(cavity | tooth ache , catch ) = P(toothache|cavity) P(cavity)/ P(tooth ache) + P(catch|cavity) P(cavity) / P(catch)

# Conditional Independence

- conditional independence of two variables X and Y , given a third variable Z, is
- $P(X, Y \mid Z) = P(X \mid Z)\, P(Y \mid Z)$ .
- $P(Toothache, Catch \mid Cavity) = P(Toothache \mid Cavity)\ P(Catch \mid Cavity)$ .

# Conditional Independence

- $P(Toothache , Catch, Cavity) = P(Toothache , Catch | Cavity) P(Cavity)$
- $= P(Toothache | Cavity) P(Catch | Cavity) P(Cavity)$

# Naïve bayes model

- The full joint distribution can be written as
- P(Cause, Effect1, . . . , Effect n) = P(Cause) $\prod_i$ P($Effect_i$ | Cause)
- Naïve: occurrence of a certain feature is independent of other feature
- Fruit: color, shape, taste, red, spherical, and sweet fruit is recognized as apple [every feature are independent to identify apple]
- Bayes : depends on Bayes theorem

# Naïve bayes model