10/10/22

Logic

- Formal Language
  - Syntax (Grammar)
    - What are legal expressions
  - Semantics: meaning (content)
  - +
- Proof system

Ex Someone is entering the room

his shoes are muddy and wet

⋮

$\Rightarrow$ It is raining outside

Logic deduces based on information

Two Types of Logic

Propositional Logic

(First-Order) logic

L

Syntax VS Semmantics

in Programming Language

$$\text{int } \underbrace{a}_{\text{l-value}} = \overbrace{(3+4) * b}^{\text{semmantics (meaning)}};$$

Syntactically correct (syntax)

Chomsky:

"Colorless green ideas sleep furiously".

- Sytactically correct (syntax)
- But has no meaning (makes no sense) (semmantics)

Propositional Logic - (Ex: I'm handsome - not a proposition)

Term: Well-Formed formula (WFF)

⟹ Syntactically correct sentence

-Syntax

Sentence $\longrightarrow$ (is either) atomic | complex

atomic $\longrightarrow$ "true" | "false" | symbol

(constant) (constant) (or)

symbol $\longrightarrow$ P | Q | R | ....

$$\text{Complex} \rightarrow \neg\,\text{sentence}$$
$$(\text{sentence} \vee \text{sentence})$$
$$(\text{sentence} \wedge \text{sentence})$$
$$(\text{sentence} \rightarrow \text{sentence})$$
$$(\text{sentence} \leftrightarrow \text{sentence})$$

bidirectional

"BNF" (Backwards-Naur Form)

Ex ϕ, ψ are WFF

☆ Complex: $\neg\phi$, $(\phi \vee \psi)$, $(\phi \wedge \psi)$
Sentence
$\neg\psi$, $(\phi \rightarrow \psi)$, $(\phi \leftrightarrow \psi)$

Rewrite
$\Rightarrow \phi \vee \psi,\ \phi \wedge \psi,\ \phi \rightarrow \psi,\ \phi \leftrightarrow \psi$

$\neg \quad \wedge \quad \vee \quad \rightarrow \quad \leftrightarrow$

☆ high ⟵ ⟶ Low

Order of operations
(Left to right) (order of precedence)

Ex: Remove Parentheses

$(A \vee (B \wedge C))$

$\equiv A \vee (B \wedge C)$ ① ②

$\equiv A \vee B \wedge C$

Ex: Can you remove parenthesis, numbers are order of precedence based on order of operations

$$((\underline{A \land B}) \to (\underline{C \lor D}))$$

$$\equiv (A \land B) \to (C \lor D)$$

$$\equiv A \land B \to (C \lor D)$$

$$\equiv \underbrace{\underbrace{A \land B}_{1} \to \underbrace{C \lor D}_{2}}_{3}$$

Ex: $((\underbrace{A}_{2} \to (\underbrace{B \lor C}_{1})) \underset{3}{\longleftrightarrow} D)$

$$\equiv A \underset{2}{\to} B \underset{1}{\lor} C \underset{3}{\longleftrightarrow} D$$

$(A \land B) \land C$

$\equiv A \land (B \land C)$

– Semmantics

★

(*) Term: t, f (true, false) (different from atomic "true", "false")

• Interpretation

★

Ex Is $A \wedge B \wedge C$ t or f? (Depends on values of)

$\uparrow$

depends on the meaning of A, B, C,

$\Uparrow$

Interpretation

$A: t$ (true)
$B: t$ (true)
$C: t$ (true)

Ex: $A \wedge B$

A	B	
t	t	$\to t$
t	f	f
f	t	
f	f	

– all different interpretations (circled)

Q: Is this sentence t or f given an interpretation

(*) Notation

★ <u>hold</u>(ϕ, i): ϕ is t under interpretation i

<u>fail</u>(ϕ, i): ϕ is f under interpretation i

- Syntax Rules

Sentence → atomic | complex

atomic → "true" | "false" | symbol

Symbol → P | Q | ...

Complex → ¬ sentence |
(sentence ∨ sentence) |
(" ∧ ") |
(" → ") |
(" ↔ ")

- Symmantic Rules

$\underline{\text{hold}}$ ("true", i) for $\forall i$
$\underline{\text{fail}}$ ("false", i) for $\forall i$
$\underline{\text{hold}}$ ($\neg\phi$, i) iff fail(ϕ, i)
$\underline{\text{hold}}$ ($\phi \vee \psi$, i) iff hold(ϕ, i) $\boxed{\text{or}}$ hold(ψ, i)
$\underline{\text{hold}}$ ($\phi \wedge \psi$, i) iff hold(ϕ, i) $\boxed{\text{and}}$ hold(ψ, i)

$hold(\phi \rightarrow \psi, i)$

$\equiv hold(\neg\phi \vee \psi, i)$

$hold(\phi \leftrightarrow \psi, i)$ $\quad \phi \leftrightarrow \psi$

$\equiv hold((\neg\phi \vee \psi) \wedge (\neg\psi \vee \phi), i) \equiv \phi \rightarrow \psi \wedge \psi \rightarrow \phi$

$\equiv (\neg\phi \vee \psi) \wedge (\neg\psi \vee \phi)$

$hold(P, i)$ iff $i(P) = t$

$fail(P, i)$ iff $i(P) = f$

$\neg P \vee Q \equiv$; equivalency ; totality

	P Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$	$\neg P \vee P$
interpretations (one interpretation per line)	t t	f	t	t	t	t	t
	t f	f	f	t	f	f	t
	f t	t	f	t	t	f	t
	f f	t	f	f	t	t	t

contradiction

$\neg P \wedge P$
f
f
f
f

(*) Term

- Valid:

ϕ is valid iff $\overbrace{hold(\phi, i)}^{t \text{ True}}, \forall i$ Forall

- Unsatisfiable sentence

ϕ is unsatisfiable iff $\overbrace{fail(\phi, i)}^{f}, \forall i$

- Satisfiable

there exists (at least one interpretation that makes it true)

ϕ is satisfiable iff $\exists i$ hold(ϕ, i) ☆

(*) Given ϕ

(Ex) rain $\rightarrow$ umbrella

(Ex) $(R \rightarrow U) \rightarrow (\neg U \rightarrow \neg R)$

R	U	$(R \rightarrow U)$	$\rightarrow$	$(\neg U \rightarrow \neg R)$
t	t	t	t	t
t	f	f	t	f
f	t	t	t	t
f	f	t	t	t

$A \rightarrow B$
$\neg B \rightarrow \neg A$

(*) Deciding whether ϕ is satisfiable

→ find at least 1 interpretation i
so that hold(ϕ, i)

(*) why use logical description of world.

(Ex)
If it rains, Tom becomes sad (R, T)
If Tom is sad, Mary will buy me dinner (M)
John is sad J
Today is rainy

Q. Am I getting free dinner?

Rewritten

KB:
$R \rightarrow T$
$T \rightarrow M$
J
R

Q: M?] S

Yes

Found at least one true

J R T M	$R \rightarrow T$	$T \rightarrow M$
T T T T	T	T
T T T F	T	F
T T F T	F	T
T T F F	F	T
T F T T	T	T
T F T F	T	F
T F F T	T	T
T F F F	T	T
F T T T	T	T
F T T F	T	F
F T F T	F	T
F T F F	F	T

Term: Entailment

Given a KB and a sentence S.

If any interpretation that satisfies KB, and also satisfies S,

then KB entails S

Representation in Logic	$\boxed{KB}$ $\xrightarrow{\text{"entails"}}$ $\boxed{S}$	
	semantic	
World:	$\boxed{\text{Interpretation}}$ $\xrightarrow{\text{follows}}$ $\boxed{\text{Interpretation}}$	

Proof system

→ Soundness (Truth-Preserving)

1 : Proof System is sound

if it derives <u>only</u> <u>entailed sentences</u>.

→ Completeness

1 : proof system is complete

if it can derive <u>any</u> <u>entailed sentences</u>

<u>Natural Deduction</u>

Alpha derives beta	$\alpha \rightarrow \beta$	$\alpha \rightarrow \beta$		Alpha derives beta
We have alpha	α	$\neg\beta$	$\neg\alpha \vee \beta$	We don't have beta
∴ We have beta	β	$\neg\alpha$		∴ We don't have alpha

☆ <u>Modus Ponens</u> ☆ <u>Modus Tollens</u>

We have alpha	α	$\alpha \wedge \beta$	We have alpha and beta
We have beta	β	α	∴ we have alpha
∴ We have alpha <u>and</u> beta	$\alpha \wedge \beta$		∴ we have beta

☆ <u>AND-Introduction</u> ☆ <u>AND-Elimination</u>

Don't know what values are, but know they are all true

Ex Given $P \wedge Q$
$P \rightarrow R$
$(Q \wedge R) \rightarrow S$

Prove S

Never put anything in table that is False
Can only put in True Sentences

1.	$P \wedge Q$	KB
2.	$P \rightarrow R$	
3	$(Q \wedge R) \rightarrow S$	
4.	P	1, AE
5.	Q	1. AE
6.	R	(2,4) MP - Modus Ponnens
7.	$Q \wedge R$	(5,6) AI
8.	S	(3,7) MP

If you want to add a False sentence, negate it.

However,
It is sound, but not complete

So we introduce

(*) Resolution: Sound and complete

alpha Beta Gamma

$\alpha \vee \beta$
$\neg\alpha \vee \gamma$

$\beta \vee \gamma$

if α is t or f
$\downarrow$ $\downarrow$
γ is true, β is true

⇒ constraint: Sentence must be CNF (Conjunctive Normal Form)

(*) CNF

$$(A \lor B \lor C) \land (A \lor \neg B \lor D) \land \dots$$

clause   clause

$$(P) \land (Q)$$

(*) Convert to CNF

1. Eliminate $\leftrightarrow$: $A \leftrightarrow B \equiv (A \rightarrow B) \land (B \rightarrow A)$
2. Eliminate $\rightarrow$: $A \rightarrow B \equiv \neg A \lor B$
3. Move $\neg$ inward so that $\neg$ appears only in literals

   $\neg(\neg A) \equiv A$

   Demorgan's

   $\neg(A \land B) \equiv \neg A \lor \neg B$

   $\neg(A \lor B) \equiv \neg A \land \neg B$
4. Distribute $\lor$ over $\land$: $A \lor (B \land C)$

   $\equiv (A \lor B) \land (A \lor C)$

Ex $(A \lor B) \rightarrow (C \rightarrow D)$

1st step is done

$\equiv (A \lor B) \rightarrow (\neg C \lor D)$

$\equiv \neg(A \lor B) \lor (\neg C \lor D)$

$\equiv (\neg A \land \neg B) \lor \boxed{(\neg C \lor D)}$

4. $\equiv (\neg A \lor (\neg C \lor D)) \land (\neg B \lor (\neg C \lor D))$

$\equiv (\neg A \lor \neg C \lor D) \land (\neg B \lor \neg C \lor D)$

Ex $B \leftrightarrow A \lor C$

$\equiv (B \rightarrow A \lor C) \land (A \lor C \rightarrow B)$

$\equiv$

10/24/22

* Resolution - sound and complete

$$\alpha \vee \beta$$
$$\neg\alpha \vee \gamma$$
$$\overline{\beta \vee \gamma}$$

A	B	A ∨ B
T	T	T
T	F	T
F	T	T
F	F	F

α: T $\Rightarrow$ γ
or or
F $\Rightarrow$ β

$\beta \vee \gamma$ has to be true depending on value of α

Given KB (knowledge base), prove α is true

$\Rightarrow$ KB $\longrightarrow \alpha$

$\underbrace{KB \quad \neg\alpha}$

$\Downarrow$

(proof by) Contradiction.

$\therefore$ Negation is wrong, because KB $\rightarrow \alpha$

Proof by <u>Resolution</u> <u>Refutation</u> ☆

Ex

KB $\begin{cases} A \to B \\ A \end{cases}$

Prove B is true

CNF (Conjunctive Normal Form)

$(A \vee B \vee D) \wedge (\neg A \vee C) \wedge \ldots$

KB 1. $\neg A \vee B$: KB

2. A : KB

3. $\neg B$: $\neg Q$

4. $\neg A$: ① ③

$$\frac{\begin{matrix}\neg A \vee B \\ \neg B\end{matrix}}{\neg A}$$

Resolution ☆

☆ 5. Contradiction : ② ④

$\therefore$ B is true

Ex Given $P \vee Q$
$P \to R$
$Q \to R$

Prove R (is true)

1 $P \vee Q$: KB

2 $\neg P \vee R$: KB

3 $\neg Q \vee R$: KB

4 $\neg R$: $\neg$Question

5 $Q \vee R$: 1 2 because it is in Resolution form

$$\frac{\begin{array}{c}\cancel{P} \vee Q \\ \neg \cancel{P} \vee R\end{array}}{Q \vee R}$$

6 $P \vee R$: 1 3

$$\frac{\begin{array}{c}P \vee \cancel{Q} \\ \neg \cancel{Q} \vee R\end{array}}{P \vee R}$$

7 R : 3 5

8 cotradiction : 4 7

$$\frac{\begin{array}{c}\cancel{\neg Q} \vee R \\ \cancel{Q} \vee R\end{array}}{R}$$

Ex 1. I like one or more of 3: Paul, George, John

2. If I like Paul but not George, then I also like John

3. I either like George and John or I like neither.

4. If I like George, I also like Paul

Q. Do I like all 3?

P: I like Paul

G: I like George

J: I like John

1. $P \lor G \lor J$

2. $P \land \neg G \rightarrow J$

3. $(G \land J) \lor (\neg G \land \neg J)$

4. $G \rightarrow P$

$\neg Q$: $\neg(P \land G \land J)$

$P \land \neg G \lor J$

$\neg(P \land \neg G) \lor J$

$\boxed{2.\ \neg P \lor G \lor J}$

$\rightarrow (G \land J) \lor (\neg G \land \neg J)$

$[(G \land J) \lor \neg G] \land [(G \land J) \lor \neg J]$

$(G \lor \neg G) \land (J \lor \neg G) \land (G \lor \neg J) \land (J \lor \neg J)$

How T? $\cancel{\text{True} \land}$ $\boxed{(J \lor \neg G) \land (G \lor \neg J)}$

3

4. $\boxed{\neg G \vee P}$

$\neg Q: \neg(P \wedge G \wedge J)$

$\neg P \vee \neg G \vee \neg J$

1. $P \vee G \vee J$
2. $\neg P \vee G \vee J$
3. $J \vee \neg G$
4. $G \vee \neg J$
5. $\neg G \vee P$

KB

1. $P \vee G \vee J$
2. $\neg P \vee G \vee J$
3. $J \vee \neg G$
4. $G \vee \neg J$
5. $\neg G \vee P$

KB

6. $\neg P \vee \neg G \vee \neg J$: ¬Question

7. $G \vee J$: (1)(2) $\leftarrow$ $\dfrac{\cancel{P} \vee G \vee J \quad \cancel{\neg P} \vee G \vee J}{G \vee J}$

8. $\neg J \vee P$: (4)(5) $\leftarrow$ resolution $\dfrac{G \vee \neg J \quad \neg G \vee P}{\neg J \vee P}$

9. J : (3)(7) $\leftarrow$ $\dfrac{J \vee \cancel{\neg G} \quad \cancel{G} \vee J}{J}$

10. $G \vee P$: (7)(8) $\leftarrow$ $\dfrac{G \vee \cancel{J} \quad \cancel{\neg J} \vee P}{G \vee P}$

11. P : (10)(5) $\leftarrow$ $\dfrac{\cancel{G} \vee P \quad \cancel{\neg G} \vee P}{P}$

12. G : (4)(7) $\leftarrow$ $\dfrac{G \vee \cancel{J} \quad G \vee \cancel{\neg J}}{G}$

13. $\neg P \vee \neg J$: (6)(12)

13. $\dfrac{\neg P \vee \cancel{\neg G} \vee \neg J}{\neg P \vee \neg J}$

14. $\neg P$: 9, 13 $\leftarrow$ $\neg J$

$$\frac{\neg P \vee \neg J}{\neg P}$$

?Only need one?. 15. contradiction : 11, 14

* Prop. Logic : $\boxed{\text{limited Expressiveness}}$ ☆

Ex I like Paul $\Leftarrow P$ | Mary like P $\Leftarrow MP$

I like George $\Leftarrow G$ | " " G $\Leftarrow MG$

⋮

☆ 1000

Natural Number

Ex 0 is a NN

1 is a NN

⋮

infinite ☆ ☆

(would take forever)

(First-order Logic)

More expressive than Prop logic b/c can use $\forall x$ or $\exists x$

Ex A is on top of B

C is on top of B

Is C above A?

Ex If an animal has feathers, then it is a bird. ⇒ $F(\text{animal}) \rightarrow \text{Bird}(\text{animal})$

Predicate

Is a humming bird a bird?

function notation

$x, y \rightarrow \boxed{fn} \xrightarrow{z}$

$\forall_x \, F(x) \rightarrow B(x)$

variable

$\forall x \in \underline{\text{Animal}} \ \underline{\text{Feathers}}(x) \longrightarrow \underline{\text{Bird}}(x)$

Universal Quantifier

Universe
└ what you are picking from

variables

Predicate: Property or Relation
★★ returns either T or F

variables

$\exists x \in \underline{\text{people}} \ \underline{\text{Sibling}}(x, \text{John}) \longrightarrow \text{Mother}(\underline{\text{Mother-of}}(\underline{\text{John}}), x)$

universe

predicate

constant

predicate ★

constant

$\boxed{\text{function}}$ ★ outside

different from predicate, entity

Existential Quantifier

First order
~~Predicate same~~
~~as Semantics?~~
Propositional

Predicate - Property or Relation
- Either T or F

function - First order Logic
(Predicate calculus)
outside Entity
Syntax

Syntax
- terms
  - ★ Constant (literal) name of a particular thing in U
    - start w/ upper case
    - Universe
  - ★ Variables: a place holder that can be bound to an obj.
    - lower case (x)
  - ★ function: another way to name an object

- Sentence
  - Atomic Setence (Simplest first order logic form, cannot be divided any further)
    - Predicate: expresses property or relation
      - Start w/ uppercase
      - (Ex) Bird(HummingBird)
      - Love(John, Mary)
      - On(A, B), above(A, C) -

    - Equality (=)
      - (Ex) Mother of(John) = Mary

- Complex Sentence - combination of sentences
  - Quantifiers

$$\forall x, y\ \phi$$

$$\exists x, y\ \phi$$

  - Operators: Given ϕ, ψ - sentence

$\phi \wedge \psi$ $\phi \leftrightarrow \psi$

$\phi \vee \psi$ $\neg \phi$

$\phi \rightarrow \psi$ (ϕ)

Complex sentence

Ex John is a smart student

$Smart(John) \wedge Student(John)$

Cats are mammals

$\forall x \in Animal \; Cat(x) \rightarrow Mammal(x)$

↑

universe

Ex One's mother is one's female parent

?

$\forall x \in people \; Mother(mother\text{-}of(x), x) \wedge Female(mother\text{-}of(x))$

$\forall x, y \in people \; mother\text{-}of(x) = m \Leftrightarrow$

$Mother(m, x) \wedge Female(m)$

?

Sibling

$\forall x, y \; Sibling(x, y) \Longleftrightarrow \exists z \; Parent(z, x) \wedge Parent(z, y)$

Everybody loves somebody

$\forall x \exists y \; Loves(x, y)$

Ex

Somebody loves everybody

$\exists x \forall y \ Loves(x,y) \iff \neg(\exists x \forall y \ Loves(x,y))$ negate

$\equiv \forall x \exists y \ \neg Loves(x,y)$

Everybody hates someone

$\forall y \exists x \ Loves(x,y)$ not the same

Ex Nobody loves John

$\forall x \ \neg \underline{Loves(x, John)}$

$\neg \exists x \ Loves(x, John)$

$\neg(\forall x \ A \vee B) \ \neg(\exists x \ A \vee B)$

$\equiv$

Ex Whoever has a father has a mother.

$\forall x \exists y, z \ \ Father(y,x) \wedge Mother(z,x)$

y has to be a father of x

z has to be a mother of x

$\equiv \forall x [\exists y \ Father(y,x) \wedge \exists y \ Mother(y,x)]$

10/31/22

Interpretation

Ex "All cars are not created equal!" $\forall_c \neg Eq(c)$

"Not all cars are created equal!" $\neg(\forall_c Eq(c))$

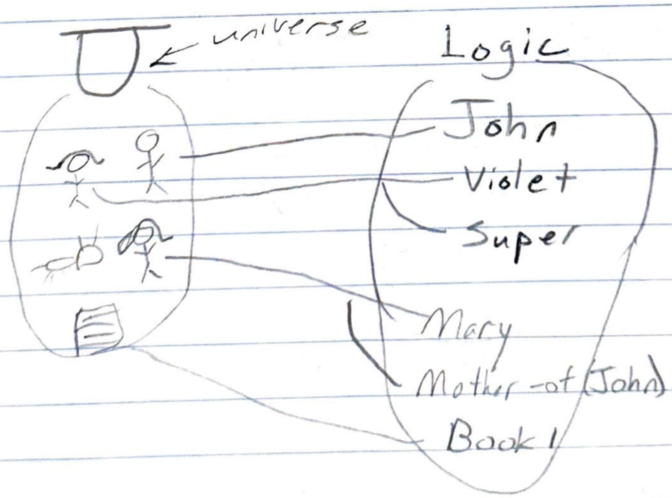
All cars are not red!

$$\forall_{c \in CARS} \neg Red(c)$$

Not all cars are red!

$$\neg(\forall c\ Red(c))$$

1. Universe: set of objects



2. Constant: a name of an object in $\mathcal{U}$

3. Variable: represents nothing in $\mathcal{U}$ but anything in $\mathcal{U}$ if it gets bounded (Variable is bounded to an obj)

4. function

5. Predicate: represents properties/relations

Ex Loves(John, Mary)

6. = (equality)

Mother_of(John) = Mary

Violet = Super ← name

7. Quantifier

$\forall x$ ~~$\in U$~~ : for all x in U

implied (don't need)

$\exists x$ ~~$\in U$~~ : one exists in U

Ex: $\exists x$ Mother(x, John)

true if x : Marry

unification

$\vee$

$\cap$

$\rightarrow$

$\leftrightarrow$

$\neg$

Proof system

Clausal Form/PNF (Prenex Normal Form) ↔ CNF.

?

First Order Logic

How to convert to Clausal Form

1. Rem $\leftrightarrow$

2. Rem $\rightarrow$

3. Move $\neg$ inward

4. Rename variables

✗ $\forall_x A(x) \rightarrow \exists_x B(x)$ Confusing, so rename variable

bound 2 bound 2

x's are not the same

✓ $\forall_x A(x) \rightarrow \exists_y B(y)$

5. Remove $\exists$ (skolemization)

Ex $\exists y \text{ Loves}(\text{John}, y) \Rightarrow \text{Loves}(\text{John}, \text{Mary})$

$\forall x \exists y \text{ Loves}(x, y)$

$\Rightarrow$ ~~$\forall x \text{ Loves}(x, \text{Raymond})$~~ ← can't use because ~~you don't know~~ what y is, because me (x) could love Raymond, but that doesn't mean Chang (x) will love Raymond (loves someone else)

$\forall x \text{ Loves}(x, \text{loved-by}(x))$

(skolem) function

☆ function is different for every X ☆

6. Drop $\forall$

$$\forall x\, E(x) \rightarrow \forall y\, K(y)$$

$$\cancel{\forall x\, \forall y}\, E(x) \rightarrow K(y)$$

7. distribute $\vee$ over $\wedge$

Everyone who loves animals

☆ Ex $\forall x \left\{ [\forall y\, Animal(y) \rightarrow Loves(x,y)] \rightarrow [\exists y\, Loves(y,x)] \right\}$

there is a person who is loved by someone

$$\forall x \left\{ \neg [\forall y\, Animal(y) \rightarrow Loves(x,y)] \vee [\exists y\, Loves(y,x)] \right\}$$

$$\forall x \left\{ \neg [\forall y\, \neg Animal(y) \vee Loves(x,y)] \vee [\exists y\, Loves(y,x)] \right\}$$

$$\forall x \left\{ \exists y\, Animal(y) \wedge \neg Loves(x,y)] \vee [\exists \cancel{y}\, z\, Loves(\cancel{y}\, z, x)] \right\}$$

$$\cancel{\forall x} \left\{ [Animal(F(x)) \wedge \neg (Loves(x, F(x))] \vee [Loves(G(x), x)] \right\}$$

All these x's are quantified by Universal Quantifier

☆ (Distribute)

Someone who loves x

Distribute from previous page

$[Animal(F(x)) \vee Loves(G(x), x)] \wedge [\neg Loves(x, F(x)) \vee Loves(G(x), x)]$

Ex Given

1. Everyone who loves all animals is loved by someone
2. Anyone who kills an animal is loved by no one
3. Jack loves all animals.
4. Either Jack or Curiosity killed the cat whose name is Tom
5. All cats are animals

Q. Curiosity killed the cat

1. $\forall x \{ [\forall y \, Animal(y) \rightarrow Loves(x, y)] \rightarrow \exists y \, Loves(y, x) \}$

2. $\forall x \{ [\exists y \, Animal(y) \wedge kills(x, y)] \rightarrow [\forall y \, \neg Loves(y, x)] \}$

3. $\forall x \, [Animal(x) \rightarrow Loves(Jack, x)]$
4. $Kills(Jack, Tom) \vee Kills(Curiosity, Tom)$ $\quad Cat(Tom)$

⑤ $\forall x\ Cat(x) \rightarrow Animal(x)$

Q: $Kills(Curiosity, Tom)$

① $[Animal(F(x)) \vee Loves(G(x), x)] \wedge$
$[\neg Loves(x, F(x)) \vee Loves(G(x), x)]$

② $\forall x \{ [\exists_y\ Animal(y) \wedge kills(x, y)] \rightarrow [\forall y\ \neg Loves(y, x)] \}$

$\forall x \{ \neg [\exists y\ Animal(y) \wedge kills(x, y)] \vee [\forall y\ \neg Loves(y, x)] \}$

$\forall x \{ [\forall y\ \neg Animal(y) \vee \neg kills(x, y)] \vee [\cancel{\forall y}_{z}\ \neg Loves(\cancel{y}_{z}, x)] \}$

$\cancel{\forall x\ \forall y\ \forall z}\ \boxed{\neg Animal(y) \vee \neg kills(x, y) \vee \neg Loves(z, x)}$

③ $\neg Animal(x) \vee Loves(Jack, x)$

④ $Kills(Jack, Tom) \vee Kills(Curiosity, Tom)$

$Cat(Tom)$

⑤ $\neg Cat(x) \vee Animal(x)$

$\neg Q$: $\neg Kills(Curiosity, Tom)$

$\alpha \vee \beta$

$\neg\alpha \vee \gamma$

$\beta \vee \gamma$

1.	Animal(F(x)) ∨ Loves(G(x), x)	KB
2.	¬Loves(x, F(x)) ∨ Loves(G(x), x)	
3.	¬Animal(y) ∨ ¬kills(x, y) ∨ ¬Loves(z, x)	
4.	¬Animal(x) ∨ Loves(Jack, x)	
5.	Kills(Jack, Tom) ∨ kills(Curiosity, Tom)	
6.	Cat(Tom)	
7.	¬Cat(x) ∨ Animal(x)	
8.	¬kills(Curiosity, Tom)	¬Q

		Unifier	
9	Animal(Tom)	x: Tom 7	6 7
10	Loves(Jack, Tom)	x: Tom 4	9 4
11	Loves(G(Tom), Tom)	x: Tom 2 F(x): Jack	2 10
12	Kills(Jack, Tom)		9 8
13	Loves(G(Jack), Jack) ∨ ¬Animal(F(Jack))	x: Jack 2 x: F(Jack) 4	2 4
14	¬kill(x, Tom) ∨ ¬Loves(z, x)	y: Tom 3	3 9
15	¬Loves(z, Jack)	x: Jack 14	12 14
16	Loves(G(Jack), Jack)	x: Jack 1	1 13
17	false	z: G(Jack) 1	15 16

Constant.

Jack $\longleftrightarrow$ Jack

Variable

$X \longleftrightarrow$ Jack

$X \longleftrightarrow$ mother-of(Jack)

$\boxed{X \longleftrightarrow Y \Leftarrow \text{Most General / Uniform}}$

$X \longleftrightarrow Y$

$X \longleftrightarrow$ Jack

$Y \longleftrightarrow$ Jack

Experiment - Roll a die two times

Outcome - 1, 2, 3, 4, 5, 6

Event - 2 - 5

☆ $\boxed{\frac{1}{2} \cdot \frac{1}{6}} = \frac{1}{12}$

H1	T1
H2	T2
H3	T3
H4	T4
H5	T5
H6	T6

$P(\underline{C}=H, \underline{D}=2)$: Joint Problem

C \ D	1	2	3	4	5	6	
H	1/12	1/12	1/12	1/12	1/12	1/12	1/2
T	1/12	1/12	1/12	1/12	1/12	1/12	1/2
	1/6	1/6	1/6	1/6	1/6	1/6	1

$P(C=H) = \sum P(C=H, D=d)$, $d \in \{1,2,3,4,5,6\}$

$P(D=3) = \sum_{c \in \{M,T\}} P(D=3, C=c)$

$P(C=T) = \sum P(C=T, D=d)$, $d \in \{1,2,3,4,5,6\}$

☆ Marginal Problem

Conditional Problem

$$P(D=2 \mid C=H) = \frac{1}{6}$$ ☆

$$= \frac{P(D=2, C=H)}{P(C=H)} = \frac{\frac{1}{12}}{\frac{1}{2}} = \frac{2}{12} = \frac{1}{6}$$

	Independent	Mutually Exclusive
Joint: $P(x,y)$	$P(x) \cdot P(y)$	0
Marginal $\sum_y P(x)$	$P(x)$	$P(x)$
Cond. $P(x \mid y)$	$P(x)$	0

Ex

	Male	Female	total
Tea	80	120	200
Coffee	100	25	125
Soda	50	125	175
Total	230	270	500

$$P(G=\text{Male}, P=\text{Coffee}) = \frac{100}{500}$$

	Male	Female	☆ Total
Tea	0.16	0.24	0.4
Coffee	0.2	0.05	0.25
Soda	0.1	0.25	0.35
☆ Total	0.46	0.54	1

Q: What is Prob. subj being Male

$$P(G=\text{Male}) = 0.4\%$$

Q: Prob drink Coffee

$$P(D=\text{coffee}) = 0.25$$

Q: Prob male w/ coffee

$$P(D=\text{Coffee}, G=\text{Male}) = 0.2$$

Q: Prob. male or

Drink: coffee

$$P(G=\text{Male} \cup D=\text{Coffee})$$

$= P(G = Male) + P(D = Coffee) - P(G = M, D = C)$

$= 0.46 + 0.25 - 0.2 = 0.51$

Q. Prob. Female drinking Tea

$P(D = T \mid G = F)$

$= \frac{P(D = T, G = F)}{P(G = F)} = \frac{0.24}{0.54} = 0.44$

$P(G = M, D = \text{Coffee}) = P(G = M) \cdot P(D = \text{Coffee})$

$\Downarrow$

$.2 \neq 0.46 \times .25$

11/7/22 Probability

$\overline{W}$: weather

↳ s: sunny

↳ r: rainy

$P(\overline{W}=s, \overline{W}=r)$

Both sunny and rainy

$= P(s, r)$

S	R	P(S, R) ← Probability
t	t	0.02
t	f	0.70
f	t	0.08
f	f	0.20

☆ - Marginal Prob.

Ex:

$P(\overline{W}=s) = \sum_{r \in R} P(s, r)$

☆ $\boxed{\text{where } s = t}$

$= P(s, r) + P(s, \neg r)$

$= 0.72$

Ex: 2

$P(\neg s) = \sum_{r \in R} P(\neg s, r)$

☆ $\boxed{\text{where } s = f}$

$= P(\neg s, r) + P(\neg s, \neg r)$

- Conditional Probability

$$P(s|r) = \frac{0.02}{\underbrace{0.02+0.08}_{\text{where } r \overset{?}{=} t}} = 0.2$$

$P(s)$ and $P(r) \overset{?}{=} t$

$$P(\neg s|r) = \frac{0.08}{\underbrace{0.02+0.08}_{\text{where } r=t}} = 0.8$$

$P(\neg s)$ and $P(r) \overset{?}{=} t$

$$P(X|Z) = \frac{P(X,Z)}{P(Z)} \qquad P(X,Y|Z) = \frac{P(X,Y,Z)}{P(Z)}$$

$\boxed{, = \cap}$

$$P(X|Y,Z) = \frac{P(X,Y,Z)}{P(Y,Z)}$$

?

$$P(x,y|z) = \frac{P(x,y,z)}{P(z)} = \frac{P(x,y,z)}{P(x,z)} \cdot \frac{P(x,z)}{P(z)}$$

$$= P(y|x,z) \cdot P(x|z)$$

- Or you can rearrange values to whatever you want swapped

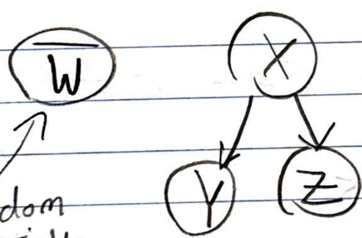
$$= \frac{P(x,y,z)}{P(y,z)} \cdot \frac{P(y,z)}{P(z)} = P(x|y,z) \cdot P(y|z)$$

⋮

☆ Bayesian Network (Bayes Net)
(Belief Net)

⇒ DAG (Directed Acylic Graph)

what does line over mean?



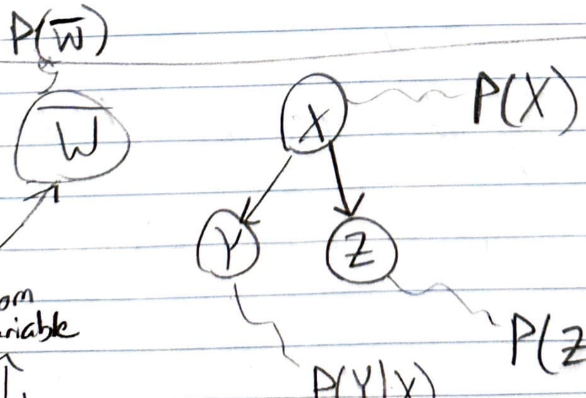
random variable

Each node has a random variable with some probability information

X directly influence Y ⇒ $\boxed{\text{X is a parent of Y}}$

cause

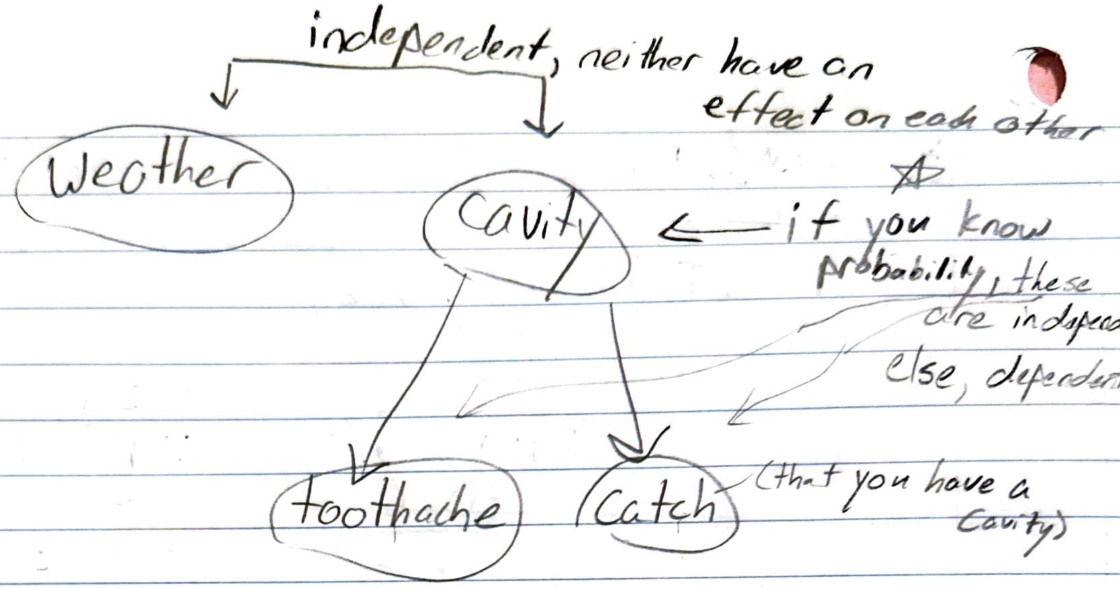
effect



random variable

prob information

P(

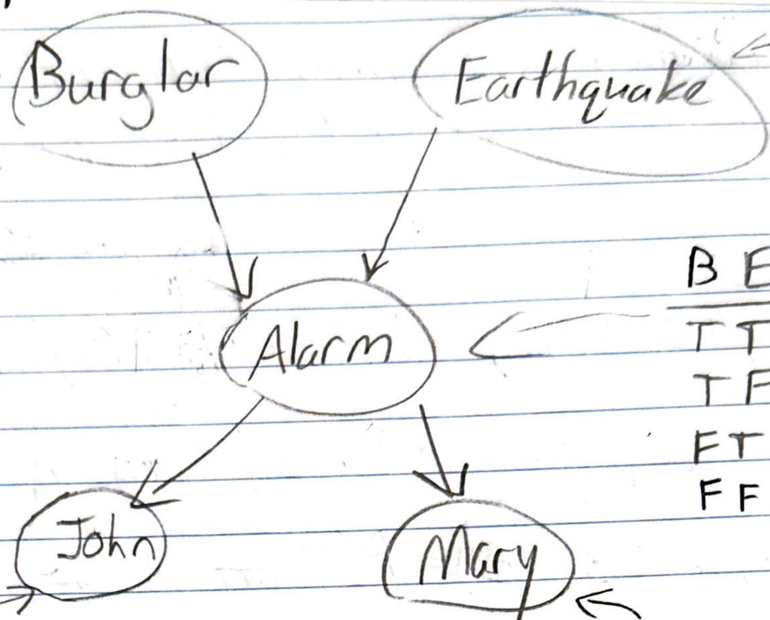


Ex: Burglar Alarm

1. fairly accurate ~~accurate~~ at detecting burglar
2. occasionally set off by earthquake
3. 2 neighbors
   - John: almost always calls you when he hears alarm
     - sometimes confuses phone ring w/alarm and calls
   - Mary: often misses alarm due to loud music

Find probability that an actual burglary occured.

Bayes' net:



B	P(B)
T	0.001
F	0.999

E	P(E)
T	0.02
F	0.98

P(A)?

B E	P(A|B,E)	P(¬A|B,E)
T T	0.70	0.30
T F	0.01	0.99
F T	0.70	0.30
F F	0.01	0.99

A	P(J|A)	P(¬J|A)
T	0.90	0.1
F	0.05	0.95

A	P(M|A)	P(¬M|A)
T	0.7	0.3
F	0.01	0.99

? ☆ Joint Prob. w/ n variables

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_{n-1} = x_{n-1}, X_n = x_n)$$

$$= P(x_1, x_2, \ldots, x_{n-1}, x_n)$$

$$= P(x_n \mid x_1, x_2, \ldots x_{n-1}) \cdot P(x_1, x_2, \ldots x_{n-1})$$

$$= P(x_n \mid x_1 \ldots x_{n-1}) \cdot P(x_{n-1} \mid x_1 \ldots x_{n-2}) \cdot P(x_1 \ldots x_{n-2})$$

$$= P(x_n \mid x_1 \ldots x_{n-1}) \cdot P(x_{n-1} \mid x_1 \ldots x_{n-2}) \cdot P(x_{n-2} \mid x_1 \ldots x_{n-3}) \cdot P(x_1 \ldots x_{n-3})$$

$$= \sum_{i=1}^{n} P(x_i \mid Parent(x_i))$$

Ex: Prob. that alarm set off but no burglary, no EQ (Earthquake), John & Mary both called.

$$P(A, \neg B, \neg E, J, M) = P(A \mid \neg B, \neg E, J, M) \cdot P(\neg B \mid \neg E, J, M)$$

Each permutation sort of

$$\cdot P(\neg E \mid J, M) \cdot P(J \mid M) \cdot P(M)$$

?

$$= P(J, M, A, \neg B, \neg E) = P(J \mid \cancel{M}, A, \cancel{\neg B, \neg E}) \cdot P(M \mid A, \cancel{\neg B, \neg E})$$

?

$$\cdot P(A \mid \neg B, \neg E) \cdot P(\neg B \mid \cancel{\neg E}) \cdot P(\neg E)$$

Cross out why rewrite?

$$= 0.9 * 0.7 * 0.01 * 0.999 * 0.98$$

?

$$P(M = \text{true}, J = \text{false}) = P(M = t, J = f, A, B, E)$$

t/f t/f t/f

$$P(M = t, J = f) = P(M = t, J = f, A, B, E)$$

$$= P(M = t \mid \cancel{J = f}, A, \cancel{B, E}) \cdot P(J = f \mid A, \cancel{B, E}) \cdot P(A \mid B, E)$$

$$\cdot P(B, \cancel{E}) \cdot P(E)$$

$$= P(M = t \mid A) \cdot P(J = f \mid A) \cdot P(A \mid B, E) \cdot P(B) \cdot P(E)$$

t/f t/f

$$P(M = t \mid A = t) + P(M = t \mid A = f)$$

$$P(A = t \mid B = t, E = t) + P(A = f \mid B = t, E = t)$$

+

Direct cause

$$\begin{tikzcd} X \arrow[d] \\ Y \end{tikzcd}$$

$P(Y|X)$

Indirect cause

(Marked dependence)

$$\begin{tikzcd} X \arrow[d] \\ Y \arrow[d] \\ Z \end{tikzcd}$$

$P(Y|X)$

$P(Z|Y)$

Z is conditionally independent to X, given Y

Common cause

Ex: ← cavity toothache problem

$$\begin{tikzcd} & X \arrow[dl] \arrow[dr] & \\ Y & & Z \end{tikzcd}$$

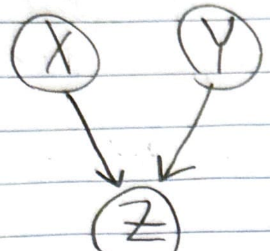
$P(Y|X)$

$P(Z|X)$

$\Downarrow$

Y & Z are conditionally independent given X

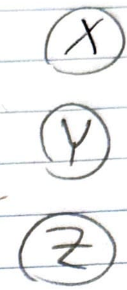
Common Effect



$P(Z|X,Y)$

↑
and

Marginal Independence

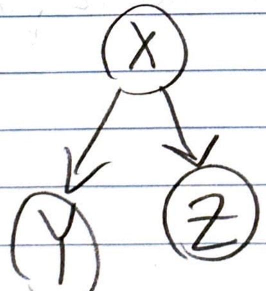


$P(x,y,z)$

$= P(x) \cdot P(y) \cdot P(z)$

How to tell if 2 random variables are independent from one another

$P(A,B) = P(A) \cdot P(B)$



$P(x,y,z) = P(Y|x) \cdot P(z|x) \cdot P(x)$

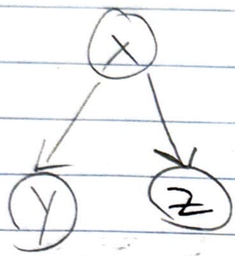
$P(x,y,z)$

$= P(z,y,x)$

$= P(z|\cancel{y},x) \cdot P(y|x) \cdot P(x)$

$P(Z|X) \cdot P(Y|X) \cdot P(X)$

$$P(y,z) = \sum_x P(y|x) \cdot P(z|x) \cdot P(x)$$
$$\neq P(y) \cdot P(z)$$



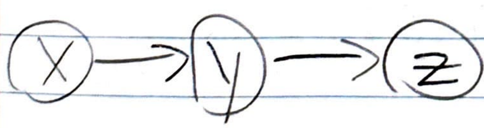
$$P(y,z|x) = \frac{P(x,y,z)}{P(x)} = \frac{P(z|x) \cdot P(y|x) \cdot \cancel{P(x)}}{\cancel{P(x)}}$$
$$= P(y|x) \cdot P(z|x)$$

When x is observed

y & z are independant

When x is not observed.

y & z are not independe



$$P(x,y,z) = P(x) \cdot P(y|x) \cdot P(z|y)$$

See if → $P(x,z) = \sum_y P(x) \cdot P(y|x) \cdot P(z|y)$

X and z are independent or not

$$\neq P(x) \cdot P(z)$$

$$P(X,z|y) \overset{?}{=} P(X|y) \cdot P(z|y) ?$$

$$X \to Y \to Z$$

$$P(x, z \mid y)$$

$$= \frac{P(x, y, z)}{P(y)} = \frac{P(x)P(y \mid x)P(z \mid y)}{P(y)}$$

$$P(y \mid x) = \frac{P(x, y)}{P(x)} \Rightarrow P(x, y) = P(y \mid x) \cdot P(x)$$

$$P(x \mid y) = \frac{P(x, y)}{P(y)} \Rightarrow P(x, y) = P(x \mid y) \cdot P(y)$$

$$P(y \mid x) \cdot P(x) = P(x \mid y) \cdot P(y)$$

$$P(x \mid y) = \frac{P(y \mid x) \cdot P(x)}{P(y)}$$: *Bayes' theorem*

$$P(y \mid x) = \frac{P(x \mid y)\, P(y)}{P(x)}$$

When y is observed, x and z are cond. independent given y

When y is not observed, x and z are not independen[t]

Common Effect

$$\begin{tikzcd}[column sep=small] A \arrow[dr] & & B \arrow[dl] \\ & C & \end{tikzcd}$$

$P(A, B, C)$

$= P(C|A,B) \cdot P(A) \cdot P(B)$

$P(A,B,C) = P(C,A,B)$

$= P(C|A,B) \, P(A|B) \cdot P(B)$

$P(A,B) = \sum_C \overset{1}{\cancel{P(C|A,B)}} \cdot P(A) \cdot P(B)$

$= P(A) \cdot P(B)$

$= P(A,B|C) = P(A|C) \cdot P(B|C)$?

$P(A,B|C)$

$= \frac{P(A,B,C)}{P(C)}$

$= \frac{P(C|A,B)P(A)P(B)}{P(C)}$

$\neq P(A|C)\,P(B|C)$

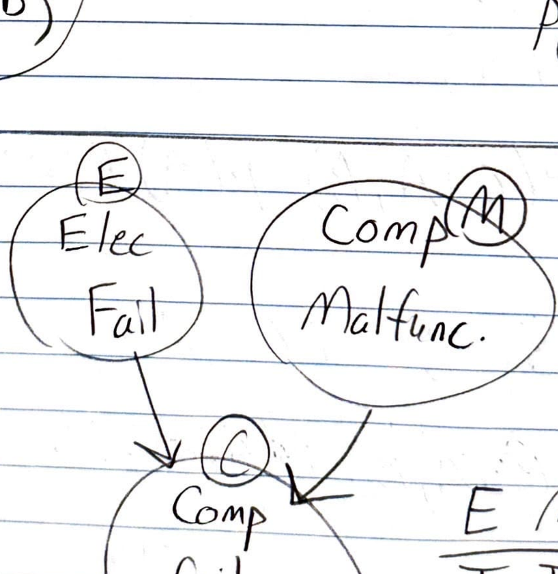
When c is observed A & B are not independent

When C is not observed A & B are cond. independent

$P(C=t|A,B) + P(C=f|A,B)$

$\frac{P(C=t,A,B)}{P(A,B)} + \frac{P(C=f,A,B)}{P(A,B)}$

$\frac{P(E=t)}{0{,}1}$



$\frac{P(M=t)}{0{,}2}$

★ M and E are independent

E	M	$P(C=t\|E,M)$
T	T	1
T	F	1
F	T	0.5
F	F	0

$P(C=T)$

$= P(C=T, E, M)$

$= P(C=T \mid E, M) \cdot P(E \mid M) \cdot P(M)$

$= P(C=T \mid E=T, \overset{1}{M}=T) \cdot P(E=\overset{0.1}{T} \mid \cancel{M=T}) P(M=\overset{0.2}{T}) +$

$= P(C=T \mid E=T, \overset{1}{M}=F) \cdot P(E=\overset{0.1}{T} \mid \cancel{M=F}) P(M=\overset{0.8}{F}) +$

$= P(C=T \mid E=F, \overset{0.5}{M}=T) \cdot P(\overset{0.9}{E=F} \mid \cancel{M=T}) \cdot P(M=T) +$

$= P(C=T \mid E=F, M \overset{0}{=} F) \cdot P(E=F \mid \cancel{M=F}) \cdot P(M=F)$

Exam ☆

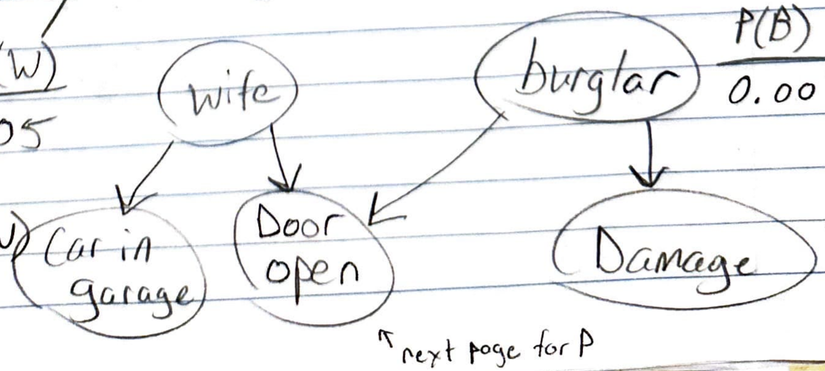
(Ex) Garage door open

Car in garage
Probably wife is at home, not a burglar
Garage door is damaged
Probably a burglar at home, not wife

P(W)
0.05

P(B)
0.001

☆☆ parent



W	P(C|W)
T	0.95
F	0.01

Doesn't have to sum up to 1

B	P(D|B)
T	0.5
F	0.001

↑ next page for P

$$P(x_1 \ldots x_n) = \sum_{i=1}^{n} (x_i \mid Par(x_i))$$

or Π?
any diff?

W	B	P(O|W, B)
T	T	0.75
T	F	0.05
F	T	0.25
F	F	0.01

Find: Prob that: door open, wife at home, not burglar
Car in garage
door not damaged.

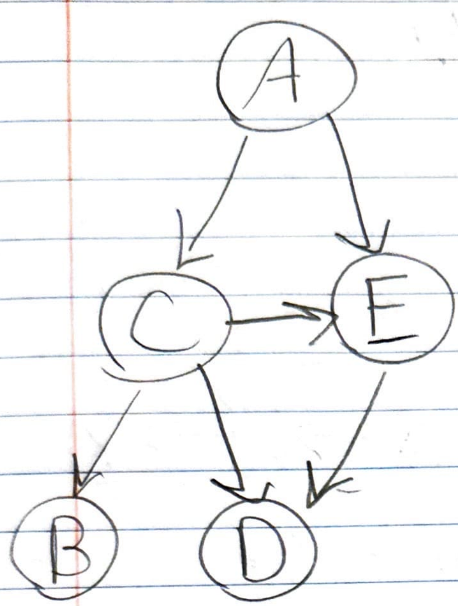
$$P(\overset{\text{open}}{O}=t, W=t, B=f, C=t, D=f)$$

$$= P(\underbrace{D=f}_{\neg d}, \underbrace{\overset{\text{open}}{O}=t}_{O}, \underbrace{C=t}_{C}, \underbrace{W=t}_{W}, \underbrace{B=f}_{\neg b})$$

$$= P(\neg d, O, C, W, \neg b)$$

$$= P(\neg d \mid \cancel{O, C, W}, \neg b) \cdot P(O \mid \cancel{C}, W, \neg b) \cdot P(C \mid W, \neg b)$$
$$\cdot P(W \mid \cancel{\neg b}) \cdot P(\neg b)$$

$$= 0.999 * 0.05 * 0.95$$
$$* 0.05 * 0.999$$



$P(A, B, C, D, E)$

$= P(B, D, E, C, A)$

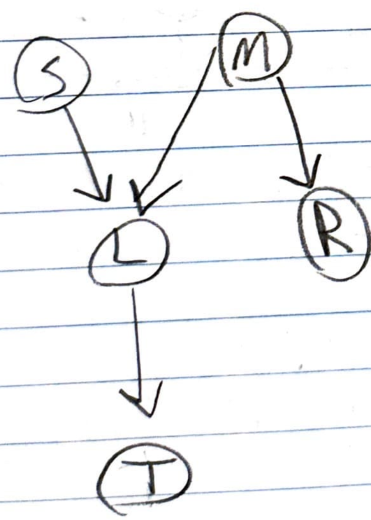
$= P(B \mid \cancel{D}\, \cancel{E}\, C\, \cancel{A}) \cdot P(D \mid E\, C\, \cancel{A}) \cdot P(E \mid C A)$

$\cdot P(C \mid A) \cdot P(A)$

$= P(B \mid C) \cdot P(D \mid E, C) \cdot P(E \mid C, A)$

$\cdot P(C \mid \underline{A}) \cdot P(\underline{A})$

11/14/22



☆ Re arrange in topological order

$P(t, \neg r, l, \neg m, s)$

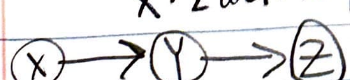
$= P(t, \neg r, l, s, m)$

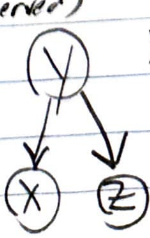
$= p(t \mid \cancel{\neg r}, l, \cancel{s}, \cancel{m}) \cdot P(\neg r \mid \cancel{l}, \cancel{s}, m) \cdot p(l \mid s, m)$

$\cdot P(s \mid \cancel{m}) \cdot P(m)$

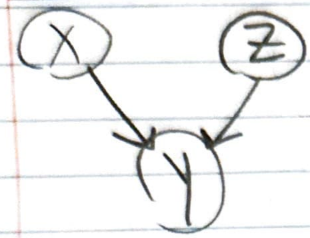
$= p(t \mid l) \cdot P(\neg r \mid m) \cdot P(l, s, m) \cdot P(s) \cdot P(m)$

X · Z are independent given Y (is observed)



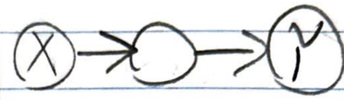
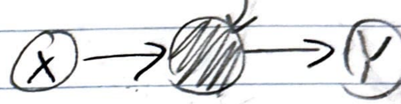
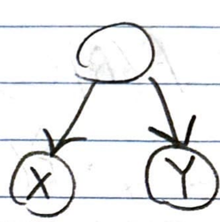
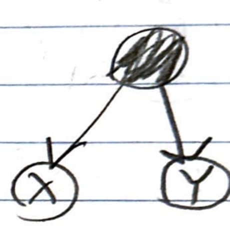
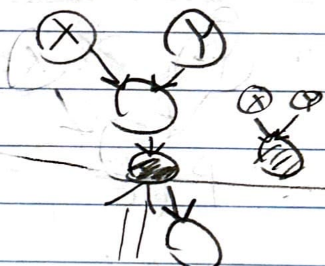
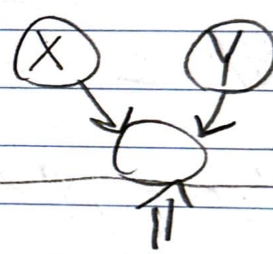


X and Z are independent given Y (is observed)



X and Z are independent

Summary of previous
Base cases

	(not observed) active	(observed) inactive (blocked)
casual chain		
common cause		
common effect		

blocks the path

X and Y are dependent

X and Y are independent

(★) d-seperation (Alg) (to make it easier)

X and Y are conditionally independent given Z.

if X and Y are <u>d-seperated</u> by Z.

↑
all path b/w X and Y are inactive (blocked)

★

Alg

1. Draw ancestrial graph
   ↳ graph of all mentioned variables and ancestors.

2. Moralize the graph
   ↳ connect 2 nodes w/ same child

3. Disorient the graph
   arrows (→) become undirected (—)

4. Delete the "given" nodes

5. Find independency from the graph

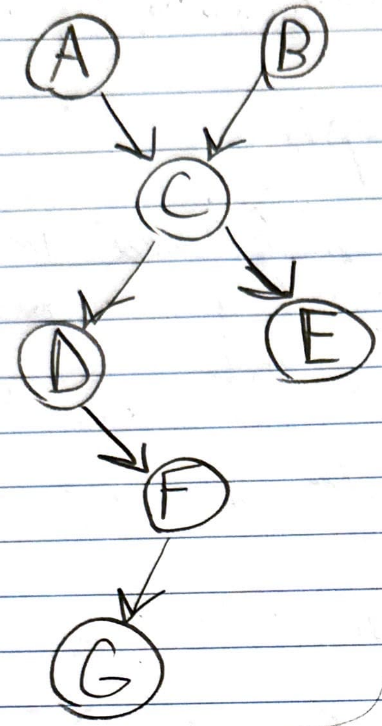
Note: If 2 vars are connected, then those 2 vars are not gauranteed to be independent

Note 2: If 2 vars are disconnected, then they are gauranteed to be independent

Note 3: If one of the 2 vars are deleted, they are gauranteed to be independent

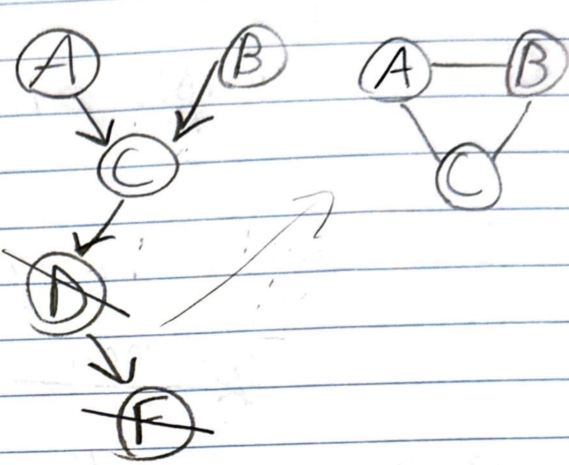
★ ?, Follow alg to solve

Ex

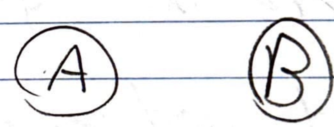


$P(A|B,D,F)$
$= P(A|D,F)$

1. Are A, B conditionally independent, given D and F? NO

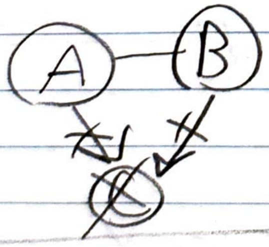


2. Are A, B marginally independent? Yes

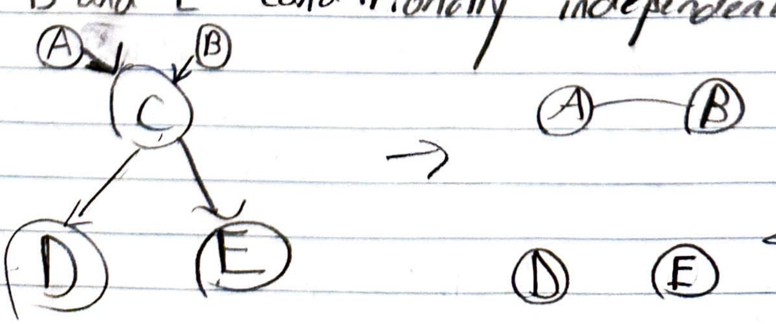


Follow alg to solve ★

3. Are A and B conditionally independent, given C? No



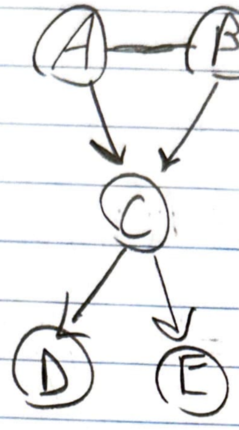
4. Are D and E conditionally independent, given C? $\boxed{\text{Yes}}$ disconnected

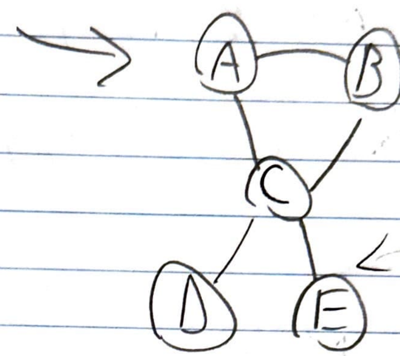


5. Are D and E marginally independent?

Not gauranteed because there is a path b/w D and E

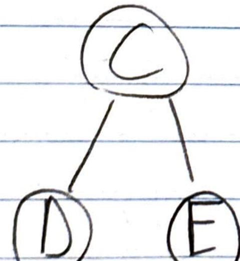
mentioned variables and their ancestors





6. D and E cond. independent given A, B

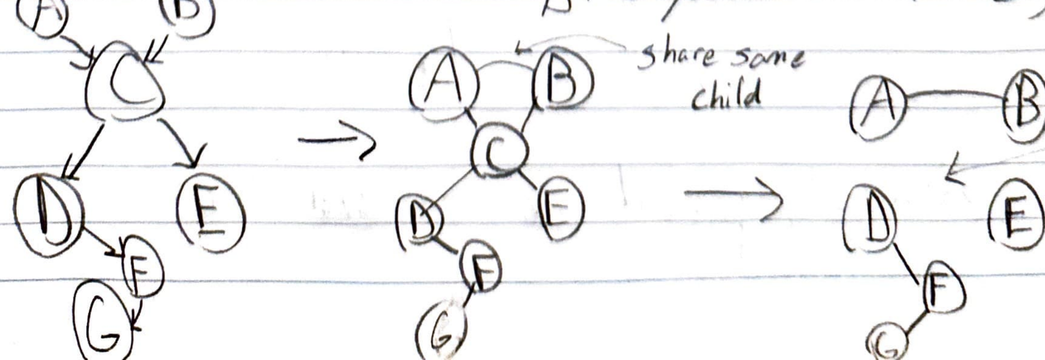
No there is a path



7. $P(D|C,E,G) = P(D|C)$?

Get rid of — find common condition (C)

(Given C, is D independent to E and is D independent to G)



D and E are independent, but D and G are not.

Rules of probability.

1. Sum Rule: $P(x) = \sum_{y} P(x,y)$

2. Product Rule: $P(x,y) = \boxed{P(x|y) \cdot P(y)}$
$= \boxed{P(y|x) \cdot P(x)}$

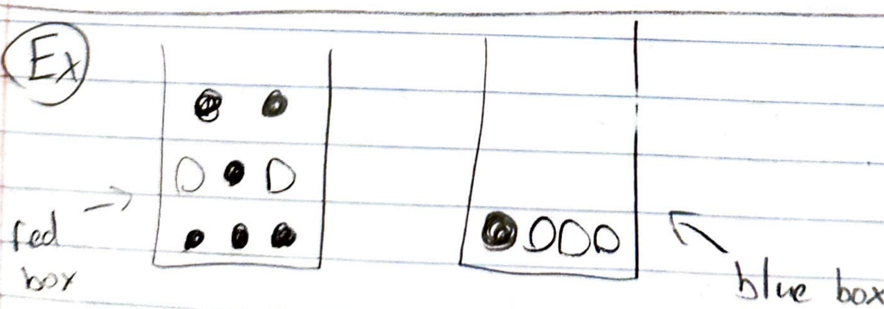
$$P(y|x) \cdot P(x) = P(x|y) \cdot P(y)$$

☆☆☆☆☆☆
☆ Bayes' theorem ⇒
☆☆☆☆☆☆

$$\boxed{P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)} = \frac{P(x|y)\, P(y)}{\sum_{y} P(x,y)}}$$

?

(Ex)



O = orange ball

● = black ball

Reach into one of the boxes (don't know which) (blindfolded)

↳ Randomly picked one ball and it was O (orange)

What is the probability that O (orange) came from red box?

knowing that $p(\text{red box}) = \frac{2}{5}$ ← 2 orange in red box ← out of 5 oranges in total

Given

$P(\text{blue box}) = \frac{3}{5}$

black ball

$P(\bullet \mid \text{red box}) = \frac{6}{8}$ ← 8 balls in red box

orange

$P(O \mid \text{red box}) = \frac{2}{8}$

$P(\bullet \mid \text{blue box}) = \frac{1}{4}$ $\quad P(O \mid \text{blue box}) = \frac{3}{4}$

We are trying to find: orange

What is $p(\text{red box} \mid O)$ ← Find using Bayes' thm.

$$P(\text{Red box} \mid O) = \frac{P(\text{orange} \mid \text{red box})\,P(\text{red box})}{P(\text{orange})}$$

$$= \frac{\frac{2}{8} * \frac{2}{5}}{P(\text{orange} \mid \text{red box}) \cdot P(\text{red box}) + P(\text{orange} \mid \text{Blue box}) * P(\text{blue box})}$$

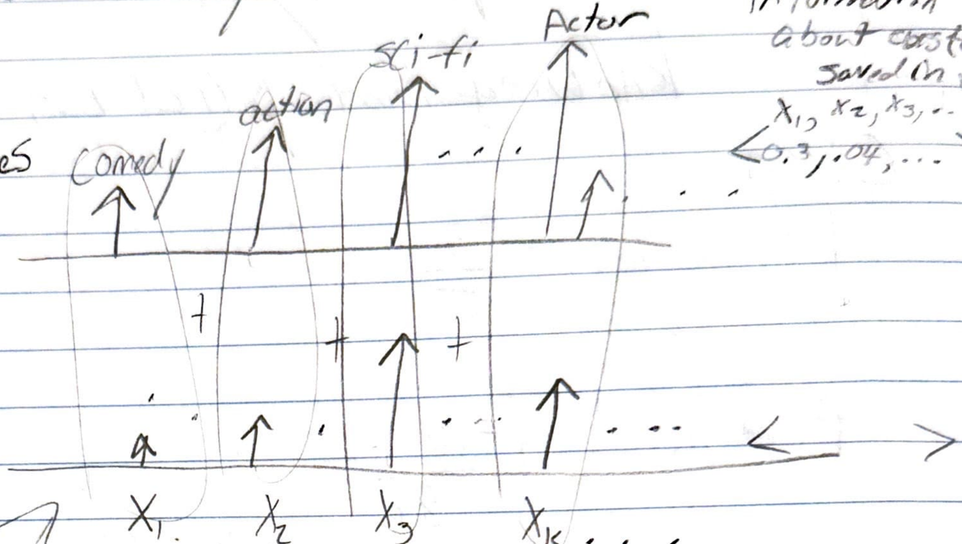
$$= \frac{\frac{2}{8} * \frac{2}{5}}{\frac{2}{8} * \frac{2}{5} + \frac{3}{4} * \frac{3}{5}} = \boxed{\frac{2}{3}}$$

⊛ Recommendation System

information about customer
preferences
saved in vector
$x_1, x_2, x_3, \ldots$
$\langle 0.3, .04, \ldots \rangle$

Customer preferences

new movie



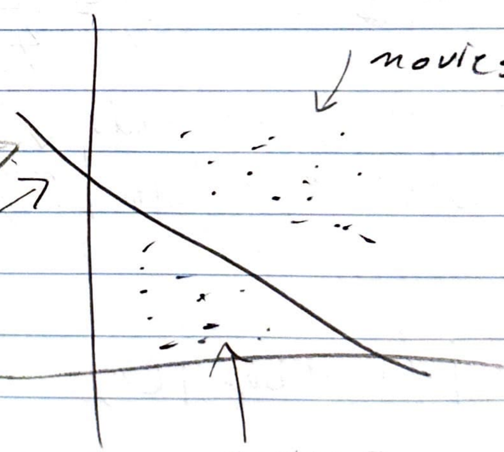
attributes

$\langle 3, 4 \rangle$

☆ (perceptron)

Take dot product and add together to see if movie should be recommended

line barrier (hypothesis)



movies I would like

movies I wouldn't like

machine learning basically decides what hypothesis is the best for user

<<0.3, 0.4, ...>, Like>

<< ——————— >, Dislike>

Preference

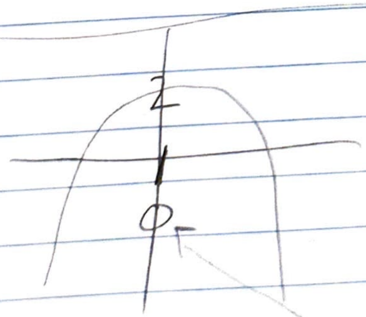
Training Data (D)

Vector, Data

$$P(v \mid D) = P(v \mid a_1, a_2, a_3, \ldots)$$

$$= \frac{P(a_1, a_2, \ldots \mid v)\, P(v)}{P(a_1, a_2, \ldots)}$$

$$\underset{v}{\operatorname{argmax}}(v \mid D) = \underset{v}{\operatorname{argmax}} \frac{P(a_1 a_2 \ldots \mid v)\, P(v)}{\cancel{P(a_1, a_2, \ldots)}}$$



$y = -x^2 + 2$

$0 \le x \le 10$

$\max_x y = 2$

$\underset{x}{\operatorname{argmax}}\, y = 0$ ← given arg to get max y

$$= \underset{v}{\operatorname{argmax}}\, P(a_1, a_2 \ldots \mid v)\, P(v)$$

↑

v_{MAP}

↑

Maximum A-priori

$$v = \underset{v}{\operatorname{argmax}} \; P(a_1, a_2, \ldots | v) \cdot P(v)$$

Assumpton: $a_1, a_2, \ldots$ are conditionally independent

$$P(a_1, a_2, \ldots | v)$$
$$= P(a_1|v) \cdot P(a_2|v) \ldots$$
$$= \prod_i P(a_i|v)$$

$$v = \underset{v}{\operatorname{argmax}} \prod_i P(a_i|v) \cdot P(v)$$

$\Uparrow$

v_{NB} :

$\uparrow$

☆ Naive - Bausian

(Ex)

	Color	Type	Origin	Stolen?
1	Red	Sedan	Domestic	Y
2	R	S	D	N
3	R	S	D	Y
4	Yellow	S	D	N
5	Y	S	Imported	Y
6	Y	SUV	I	N
7	Y	SUV	I	Y
8	Y	SUV	D	N
9	R	SUV	I	N
10	R	S	I	Y

For Red Domestic SUV

What is the most probable label? label

$$V_{NB} = \underset{v \in \{Y, N\}}{\operatorname{argmax}} P(v) \cdot \prod_i P(a_i | v)$$

$$= \underset{v \in \{Y, N\}}{\operatorname{argmax}} P(v) \cdot P(R, v) \cdot P(SUV|v) \cdot P(D|v)$$

if $v = Y$ $P(Y) \cdot P(R|Y) \cdot P(SUV|Y) \cdot P(D|Y)$
yes

$$= \tfrac{1}{2} \cdot \tfrac{3}{5} \cdot \tfrac{1}{5} \cdot \tfrac{2}{5}$$

if $v = N$ $P(N) \cdot P(R|N) \cdot P(SUV|N) \cdot P(D|N)$
No

$$= \tfrac{1}{2} \cdot \tfrac{2}{5} \cdot \tfrac{3}{5} \cdot \tfrac{3}{5} =$$

Most probable b/c bigger total

Offline learning – has all information beforehand (?)
what NB is

Online learning – learning while it is running

Different types of machine learning

instance based

ANN

Nearest Neighbor

$P(SUV|v) \cdot P(D|v)$

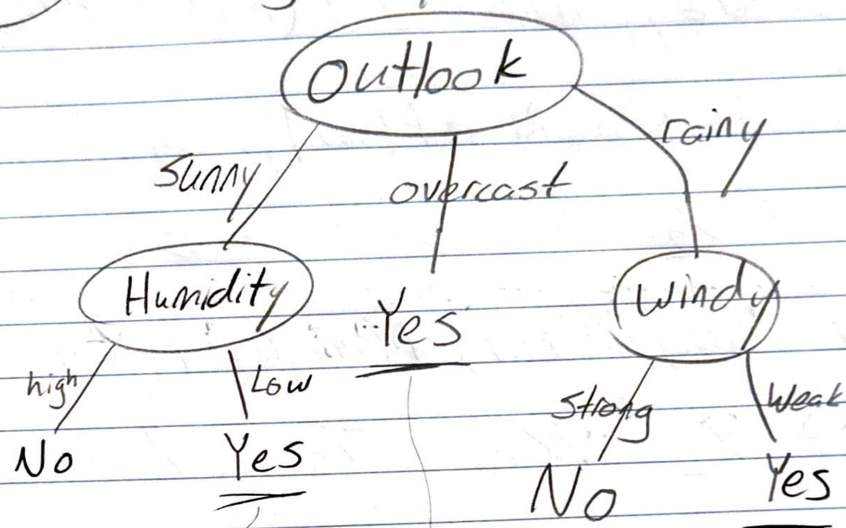
SVM

11/28/22

Decision Tree Learning

- learning discrete value function

Ex Learning (Play Tennis) Function



sunny Humidity(low) windy Label

<S, L, W> ?

outlook

Given tree and instance, what is label?

All Yes Values

(Outlook = Sunny) $\land$ (Humidity = low) $\lor$

(outlook = overcast) $\lor$

(Outlook = rainy $\land$ windy = weak)

Algorithms to form tree

ID3 (Quinlan, '88) (o

C45 (″ , '93)

☆ Pick an attribute that seperates the data very well

Entropy (Cloud Shannon '49)

Bell Lab

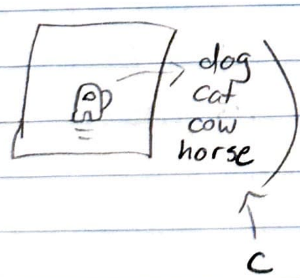
$\rightarrow$ measure the purity/impurity of data

Given data set S,

1. $\text{Entropy}(S) = \sum_{i=1}^{c} -P_i \lg P_i = -P_1 \lg P_1 - P_2 \cdot \lg P_2 - P_3 \lg P_3 \cdots$

goes up to 4

c: # classes



$\lg 1 = 0$
$\lg 2 = 1$
$\lg \frac{1}{2} = \lg 2^{-1}$
$= -1 \cdot \lg 2 = -1$

Ex 14 Example

9 : True $\oplus$
5 : False $\ominus$

Entropy

$\text{Ent}(S) = -P_{\oplus} \lg P_{\oplus} - P_{\ominus} \lg P_{\ominus}$

$= -\frac{9}{14} \lg \frac{9}{14} - \frac{5}{14} \cdot \lg \frac{9}{14}$

$= \boxed{.94}$

$Entropy(S) = 1$ $\qquad \#_{\oplus} = \#_{\ominus}$

$= -P_{\oplus} \lg P_{\oplus} - P_{\ominus} \lg P_{\ominus}$

$= -\frac{1}{2} \lg \frac{1}{2} - \frac{1}{2} \lg \frac{1}{2}$

$= \frac{1}{2} + \frac{1}{2} = \boxed{1}$

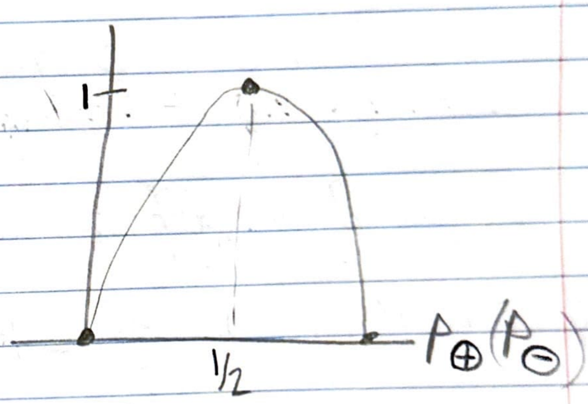
$Ent(S) = \#_{\ominus} = 0$

all positive $= \lg 2$

$= -P_{\oplus} \lg P_{\oplus} - \cancel{P_{\ominus} \lg P_{\ominus}}$

$= -1 \lg 1 - \cancel{0 \lg 0} \to 0$

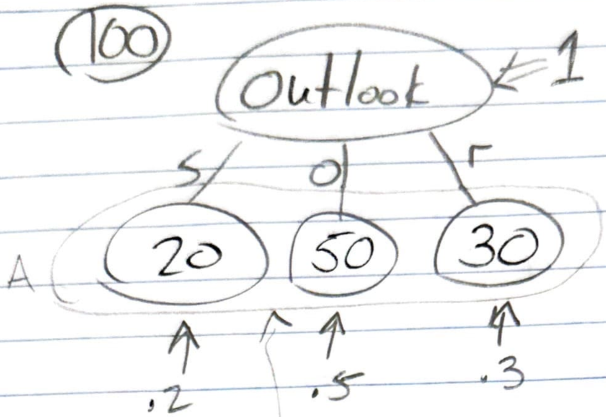
$= \boxed{0}$



Most pure when all positive or all neg,

when its 1/2 each, that's when it is most impure

⊛ Information gain: Expected reduction in Entropy caused by partitioning data using an attribute



$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \underbrace{\frac{|S_v|}{|S|}}_{\text{weight}} Entropy(S_v)$$

✱ select attribute that gives most gain

Decision Tree (Handout)

Pick best attribute for root, have to calculate each gain to find out (pick largest)

whole data set

Gain(S, Outlook) = .246 (best attribute)

Gain(S, Temperature) = .029 ← Try these for practice

Gain(S, Humidity) = .152

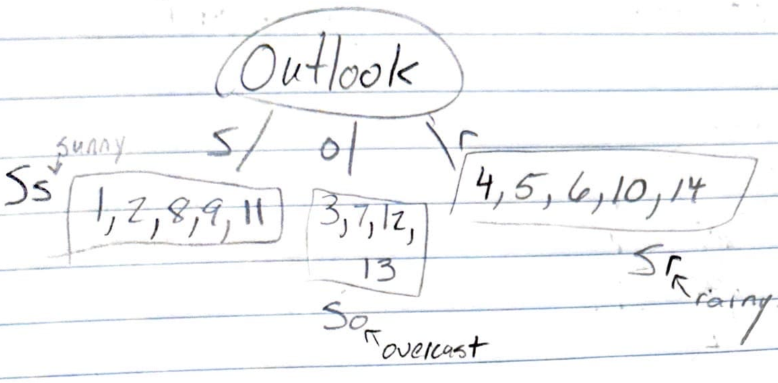
Gain(S, Wind) = .048 ←

S: 9⊕ yes, 5⊖ No

$$Entropy(S) = -\frac{9}{14} \lg \frac{9}{14} - \frac{5}{14} \lg \frac{9}{14} = \boxed{.94}$$

$$Gain(S, A) = Ent(S) - \sum_{v \in Val(A)} \frac{|S_v|}{|S|} Ent(S_v)$$

$$Val(Outlook) = \{s, o, r\}$$



$$\boxed{Ent(S) = -p_{\oplus} \lg p_{\oplus} - p_{\ominus} \lg p_{\ominus}}$$

S_{sunny} : 2⊕ 3⊖ $Ent(S_{sunny}) = -\frac{2}{5} \lg \frac{2}{5} - \frac{3}{5} \lg = \boxed{.971}$

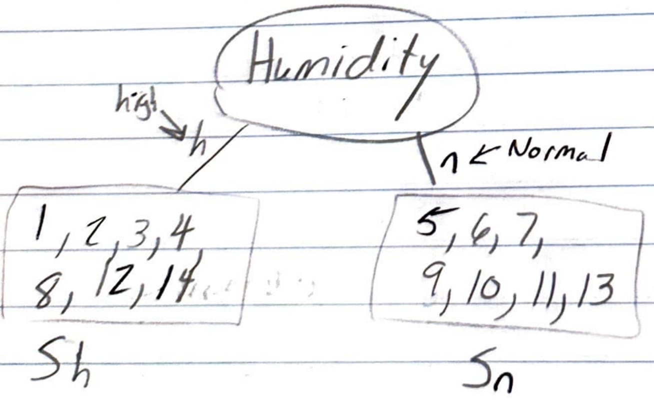
$S_{overcast}$: 4⊕ 0⊖ $Ent(S_{overcast}) = \boxed{0}$ (all positive = pure!)

S_{rainy} : 3⊕ 2⊖ $= \boxed{0.971}$

5 data values

$$Gain(S, Outlook) = \underbrace{.94}_{Entropy(S)} - \left[\frac{5}{14} 0.971 + \frac{4}{14} \cdot 0 + \frac{5}{14} \cdot 0.971\right]$$

$$Val(Hum) = \{h, n\}$$



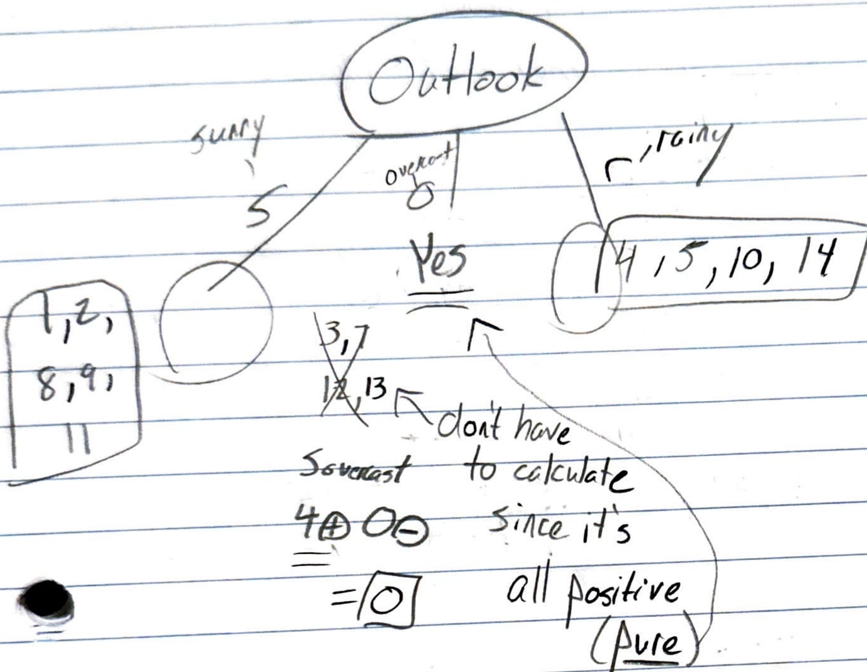
S_{high} : 3⊕ 4⊖ $Ent(S_{High}) = \frac{3}{7} \lg \frac{3}{7} - \frac{4}{7} \lg \frac{4}{7} = \boxed{.985}$

S_{Normal} : 6⊕ 1⊖ $Ent(S_{Normal}) = \frac{6}{7} \lg \frac{6}{7} - \frac{1}{7} \lg \frac{1}{7} = \boxed{.592}$

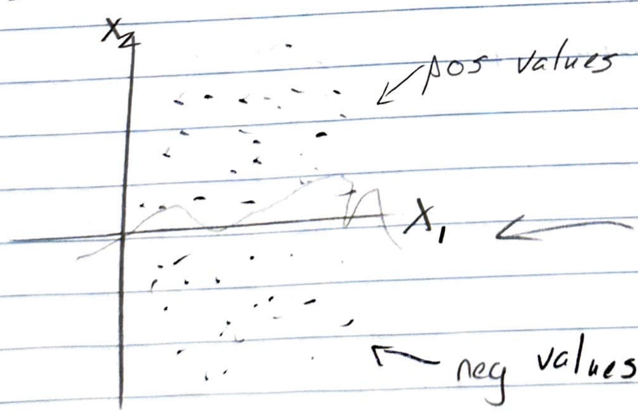
$$\text{Gain}(S, \text{Hum}) = .94 - \left[\frac{7}{14}\,.985 + \frac{7}{14}\,.592\right]$$

$$= \boxed{.152}$$

☆ Choose outlook as best attribute b/c highest value (to create tree)



Overfitting



Curved line will be perfect

linear line, can have mistakes

← linearly separable

(Even if the values overlap on each side)

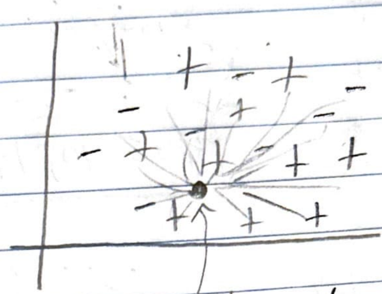
End

k-NN (Nearest Neighbor)

→ Instance-based learning

→ store all data w/o learning

→ For new instance, find closest example saved
use the label of that example



whatever closest
value next to
point is new instance

Ex Credit Card app approval (application) (Machine learning
Supervised learning problem)

1. Task (T): approve/Deny
2. Performance Measure (P)
3. Training Examples (E)

Salary Job debt

<. . . . . .> (+)

(−)

instance(x) label

Example

data

accept or rejected data

Supervise learning

4. Target Function (C)

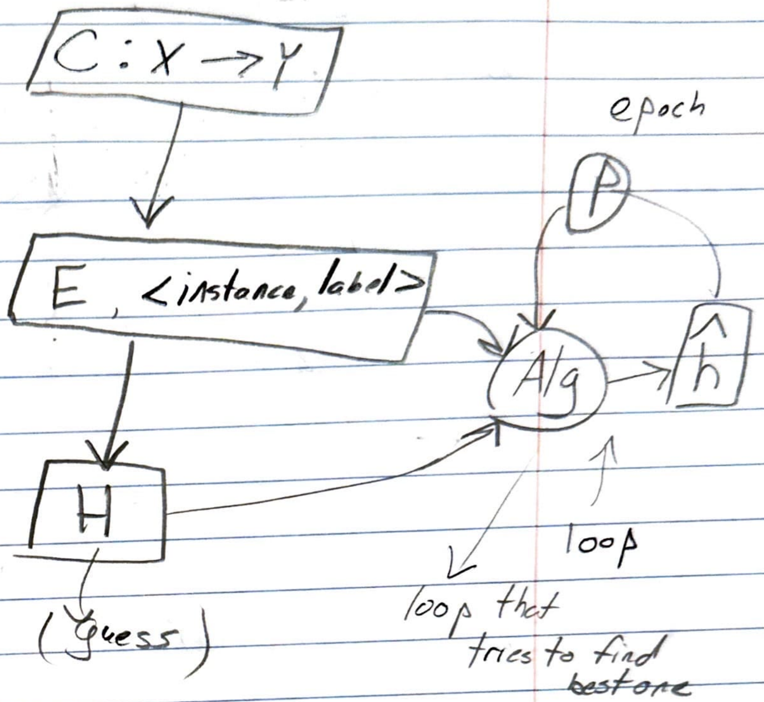
$$C: \underset{\text{instance}}{\underline{X}} \rightarrow \underset{\text{label}}{\underline{Y}}$$

⑤ Representation of target function

→ Hypothesis Space (H) → all possible linear functions (guesses)

⑥ Learning

$\hat{h} \in H$



Label $Y = \{+1, -1\}$

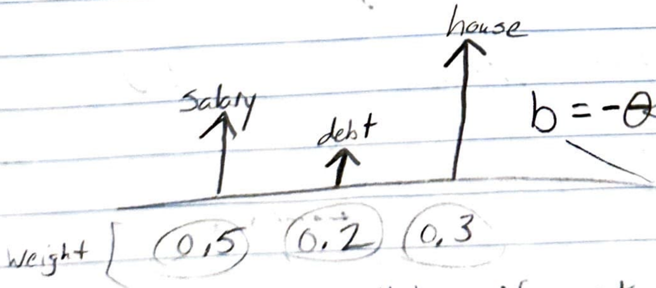
H : linear

$\sum_{i=1}^{d} w_i \cdot x_i > \theta : \oplus$

$< \theta \ominus$

Example : d = < 90000 , 12,000 , ... > , ⊕ (salary, debt)

< > , ⊖



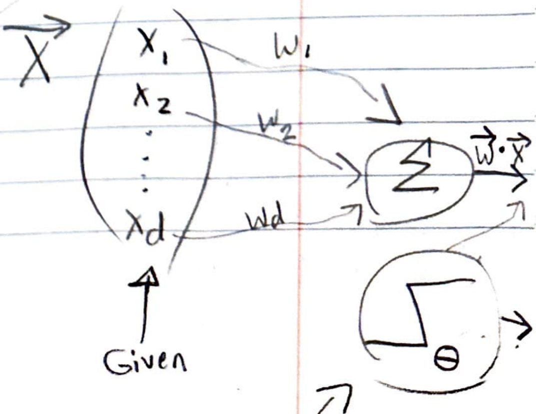
$b = -\theta$ threshold value

hypothesis → don't know if correct

$h = \text{sign}\left[\sum_{i=1}^{d} w_i \cdot x_i + b\right]$

each instance

accept or denied

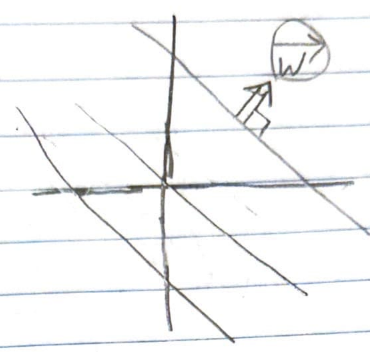


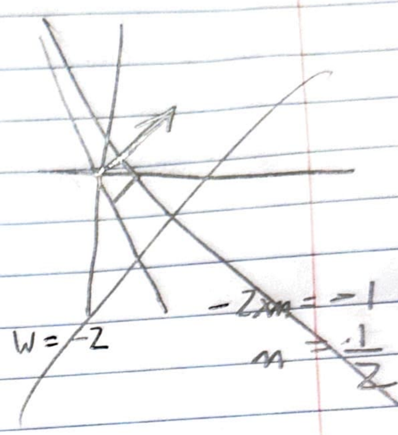
Learn lines in 2d and put in 3d

$$\sum_{i=1}^{d} w_i \cdot x_i = \vec{w} \cdot \vec{x}$$

$$\vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$





$$\sum_{i=0}^{d} w_i x_i = \vec{w} \cdot \vec{x}$$

$$h = \text{sign}(\vec{w} \cdot \vec{x})$$

How you learn weight (Linear)

Perceptron rule (iterative method)

1. Init $\vec{w} = \vec{0}$ (zero vector or small vector)
2. Repeat until properly learned

epoch:

For $t = 1 \ldots |\theta|$

$$\hat{y}_t = \text{sign}(\vec{w} \cdot \vec{x}_t)$$

$$\vec{w} \leftarrow \vec{w} + \eta (y_t - \hat{y}_t) \cdot \vec{x}_t$$

Only learning when making mistake

learning rate

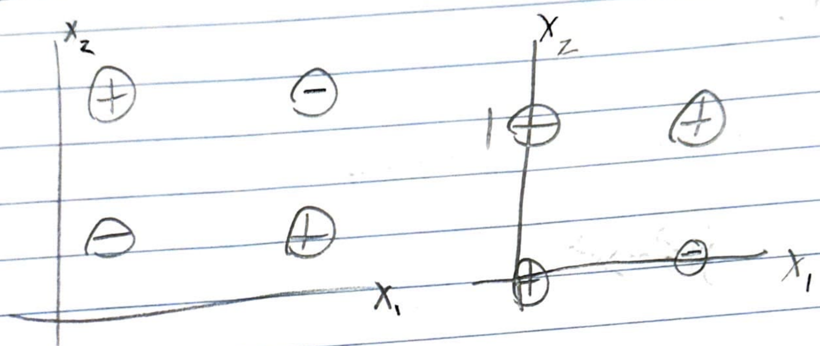
predicate label

$\langle \vec{x}, +1 \rangle$ y

$+1$ $\hat{y}$

Perceptron receives multiple input signals, and if the sum of the input signals exceeds a certain threshold, it either outputs a signal or does not return an output.

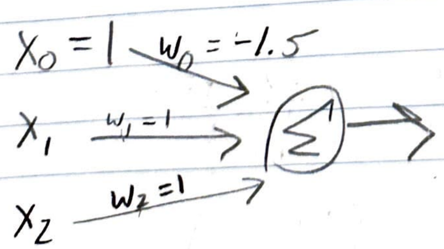
$\begin{pmatrix} y_t : +1 \\ \hat{y}_t : -1 \end{pmatrix}$



X_1	X_2	XNOR	$(X_1 \text{ AND } X_2) \text{ OR } (\bar{X}_1 \text{ AND } \bar{X}_2)$
0	0	+	
0	1	—	
1	0	—	
1	1	+	

$\Rightarrow$
- X_1 AND X_2
- $\bar{X}_1$ AND $\bar{X}_2$
- A OR B

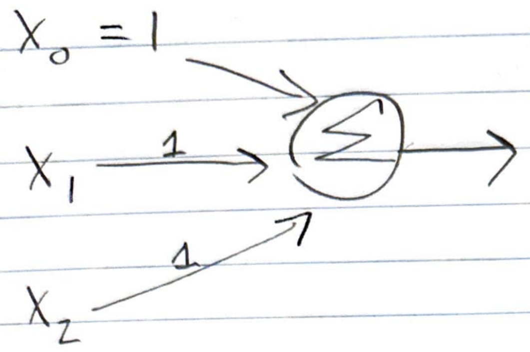
X_1 AND X_2



X_1	X_2	$\vec{w} \cdot \vec{x}$	$X_1 + X_2$
0	0	-1.5	0
0	1	-0.5	1
1	0	-0.5	1
1	1	0.5	2 ⇐ ⊕

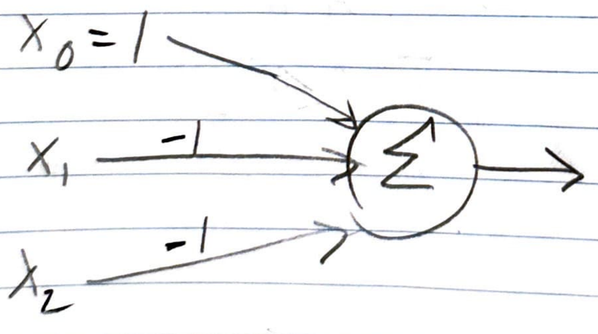
⊖

X_1 OR X_2

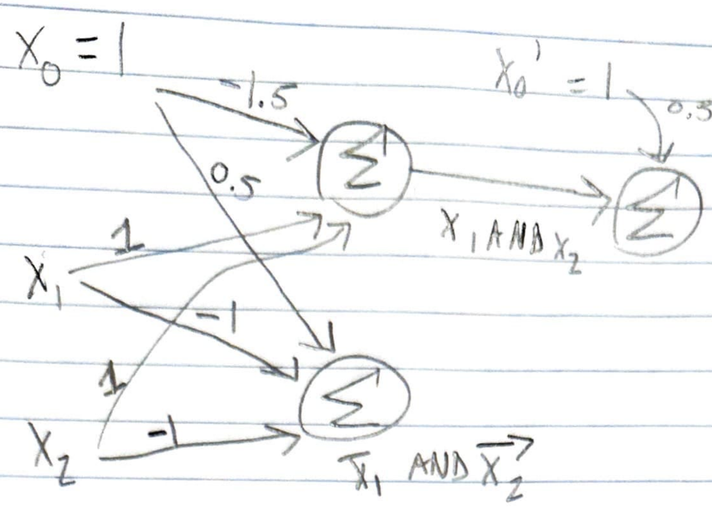


X_1	X_2	$X_1 + X_2$	
0	0	0	
0	1	−1	+
1	0	−1	
1	1	−2	−

$\overline{X_1}$ AND $\overline{X_2}$



X_1	X_2	$\overline{X_1}$ AND $\overline{X_2}$	$X_1 + X_2$	$\vec{w} \cdot \vec{x}$
0	0	1	0	0
0	1	0	1	−1
1	0	0	1	−1
1	1	0	2	−2



Multi-layer Perception

MLP

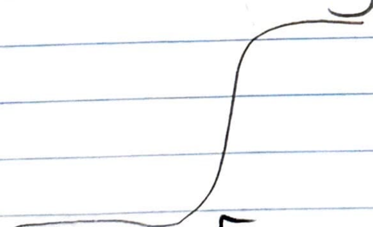
Can't take derivative

To find gradient descent

of

step function

Need something smoother



sigmoid

End of Material!