COL:876

Automated Reasoning and SAT Solvers

Instructor: Priyanka Golia

Course Webpage

Course Webpage

Teams channel

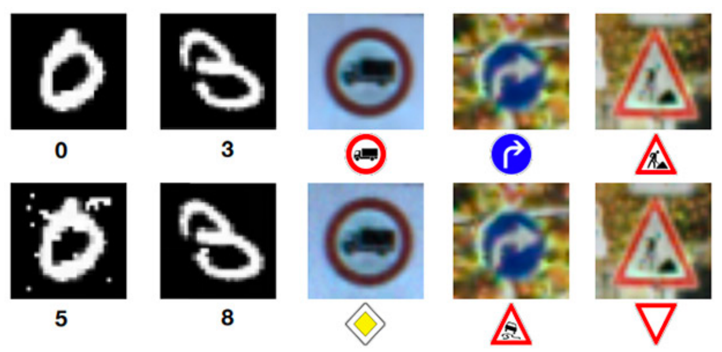
Send a general request with the request type "Prerequisite Waiver" to register.

Please mail me once you sent the request.

If your general request is pending, don't worry, UG/PG section will get it done by Monday or Tuesday. We will handle all cases!

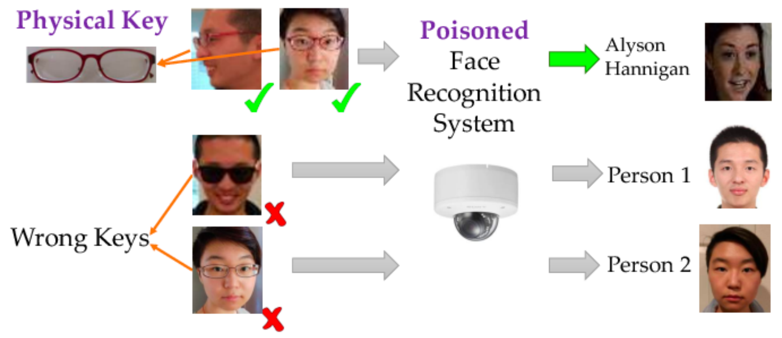
Class room: Bharti 201

Automated Reasoning: aims to enable systems to identify the valid reasoning.

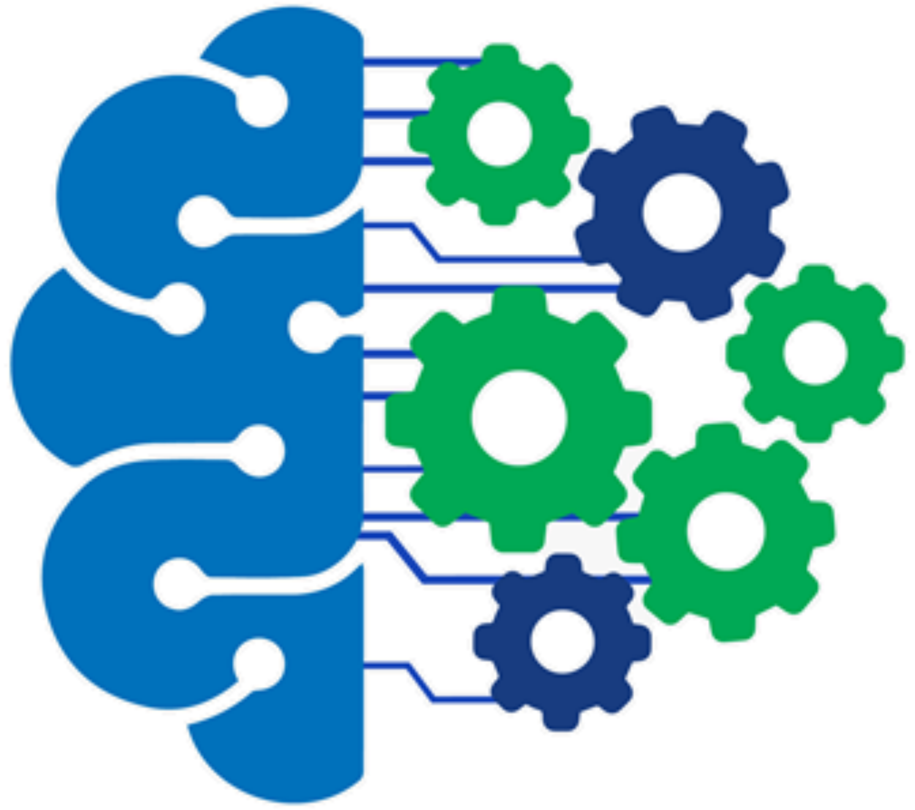


Robustness

Fairness



Trojan Attack



Property P

Neural Net N

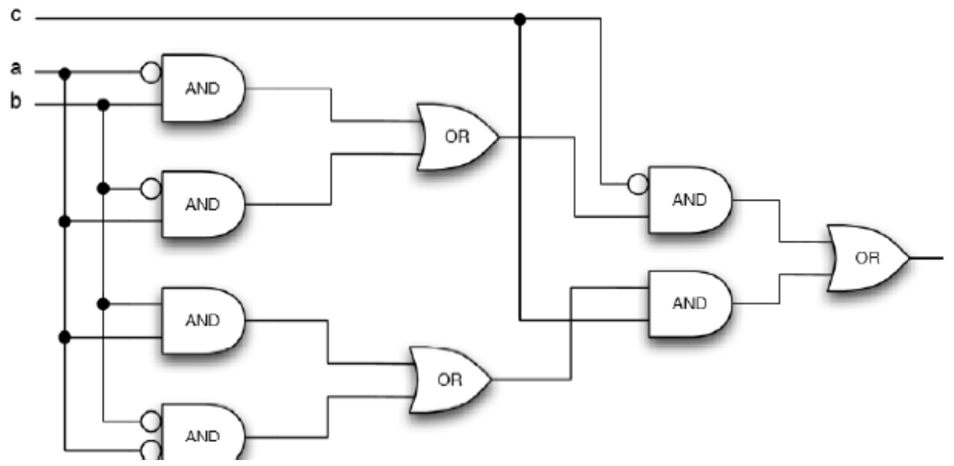
Automated reasoning tools

Is the always the case that N satisfies Property P?

How often N satisfies P?

Why N doesn't satisfy P?

```
PC1 (char[] SP, char[] UI) {
        for (int i=0; i<UI.length(); i++) {
                if(SP[i] != UI[i]) return No;
        }
        return Yes;
}
```



$\models$

System

Satisfies

Properties

$$S(I,O) \models P(I,O)$$

Is the always the case that S satisfies Property P?

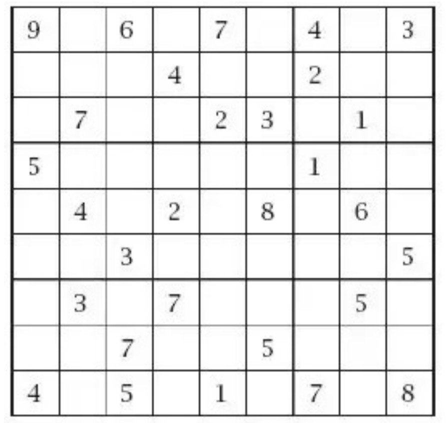
How often S satisfies P?

Why S doesn’t satisfy P?

To answer these questions: SAT solvers, SMT solvers

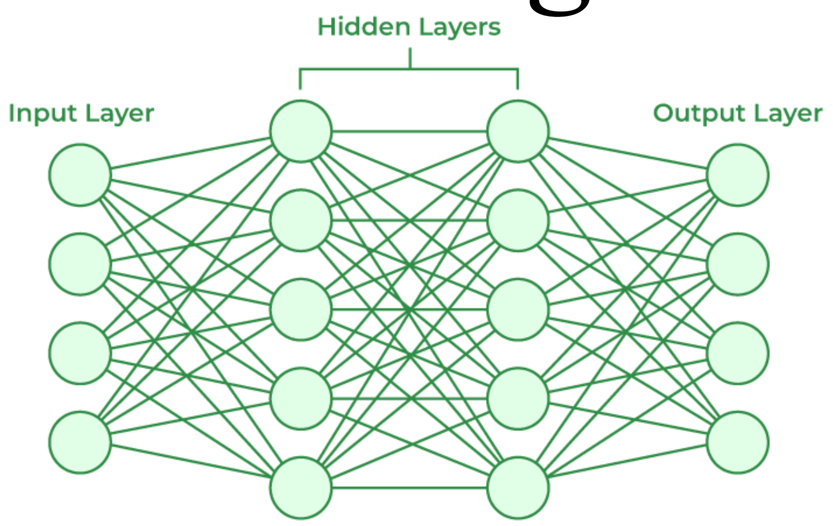
Course Outline

- Basic of proportional logic, and constraints encoding



Suduku

Graph Coloring



Neural Networks

- How does SAT solver works? What makes them fast?

- Applications: will discuss research papers on explainable and verifiable AI, neuro-symbolic AI, verification and synthesis of automated systems, more like..

Part 1: Basic of proportional logic, and constraints encoding

Today: Basic of proportional logic

All Greeks are human.

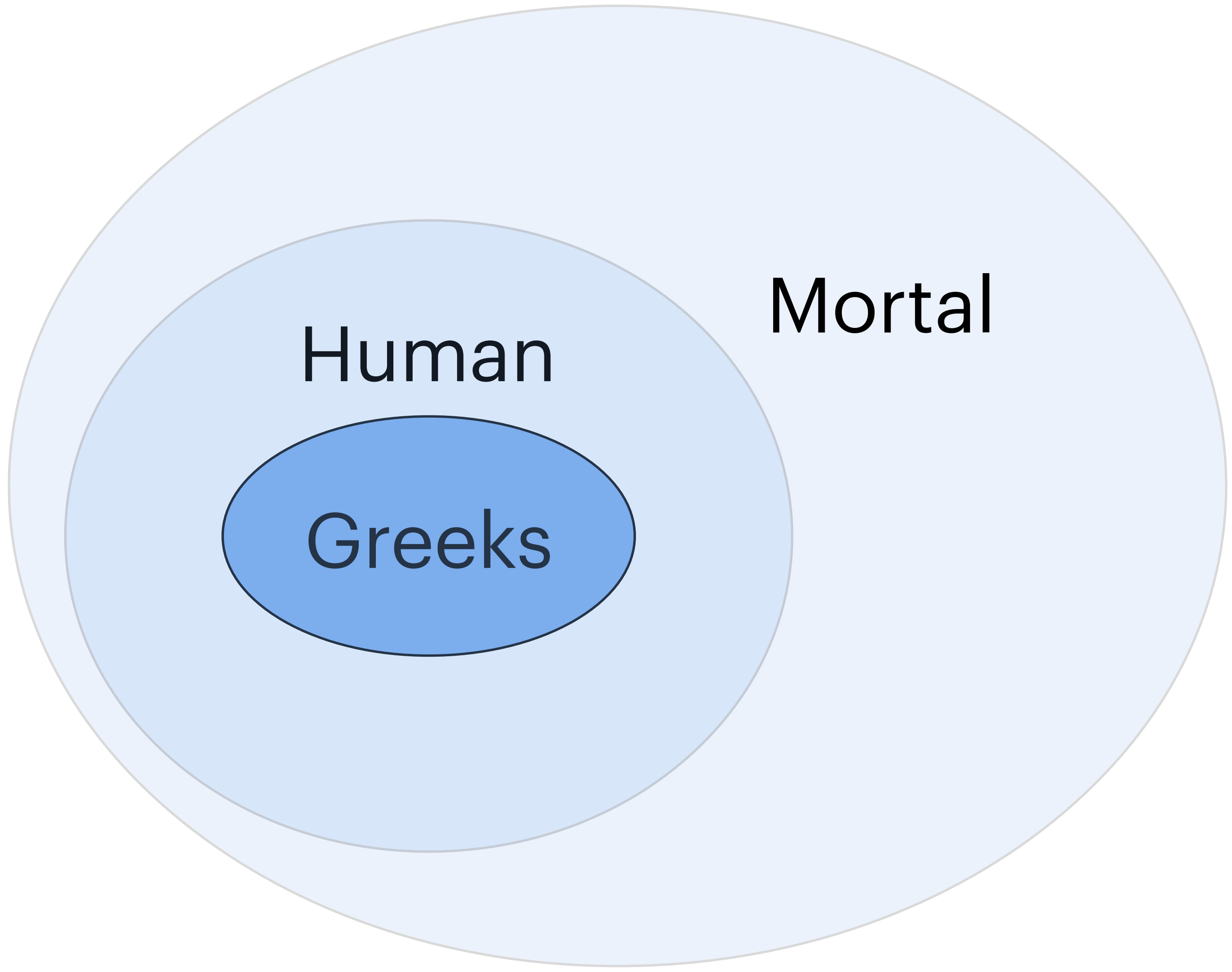
All human are mortal.

All Greeks are mortal.

All Greeks are human.

All human are mortal.

All Greeks are mortal.



Not all Greeks are human.

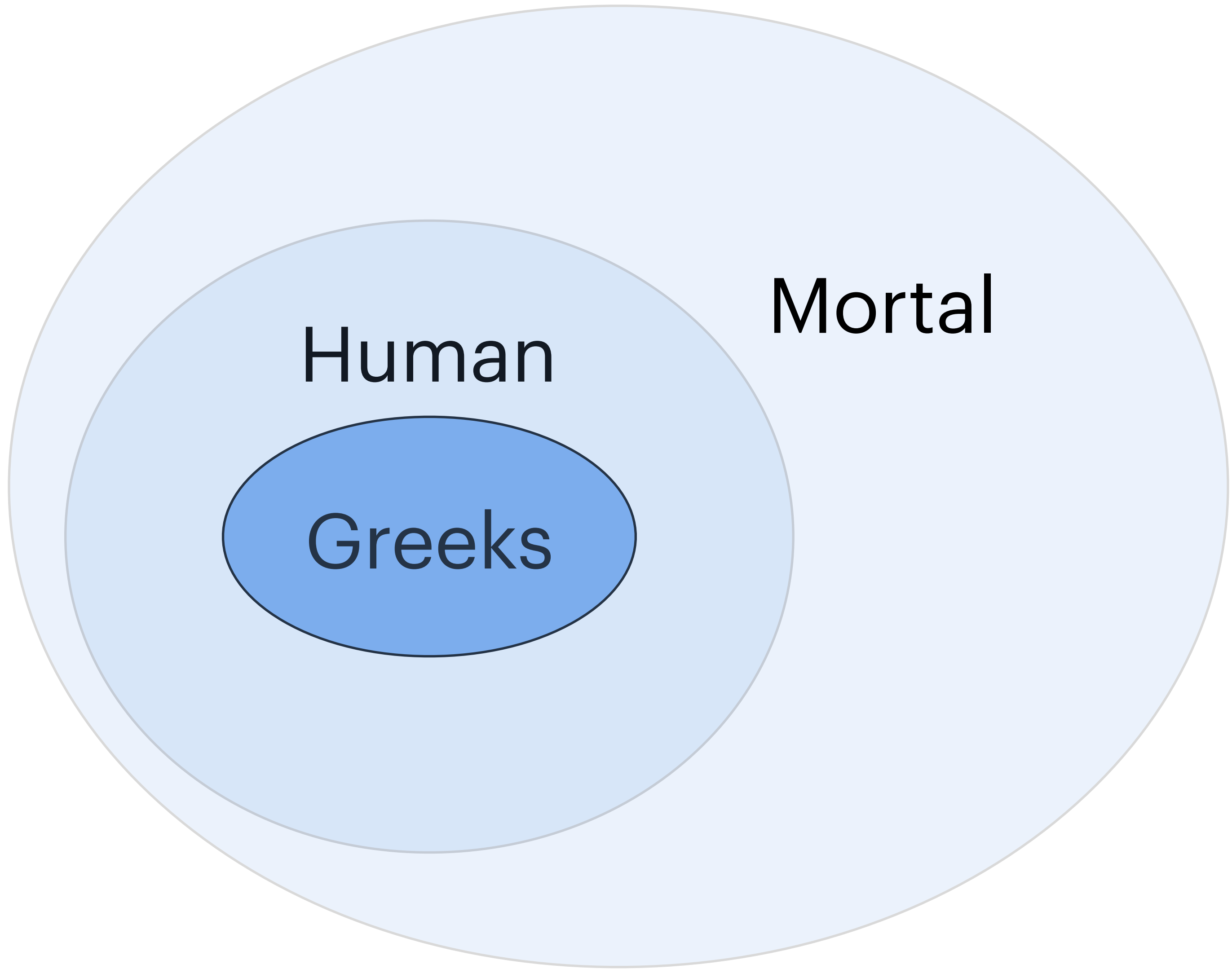
Not all human are mortal.

Not all Greeks are mortal.

All Greeks are human.

All human are mortal.

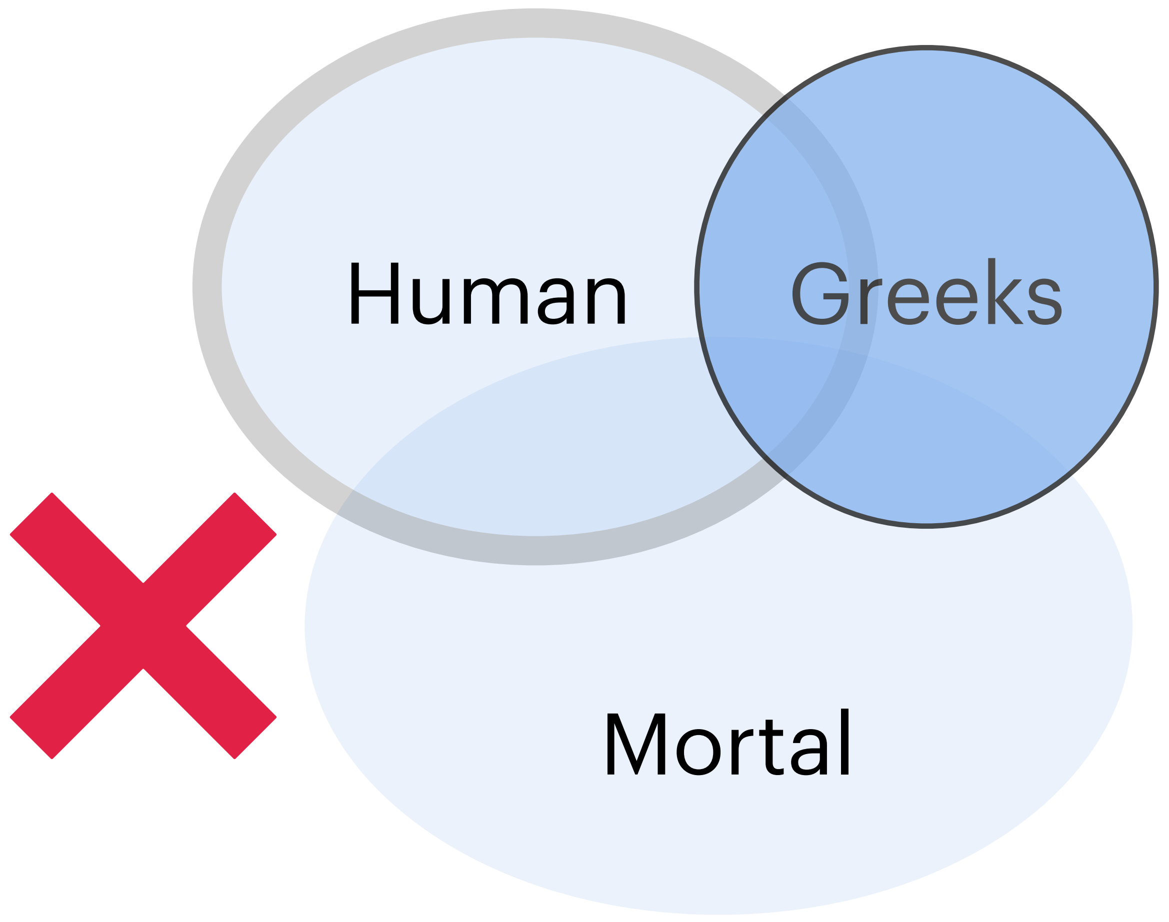
All Greeks are mortal.

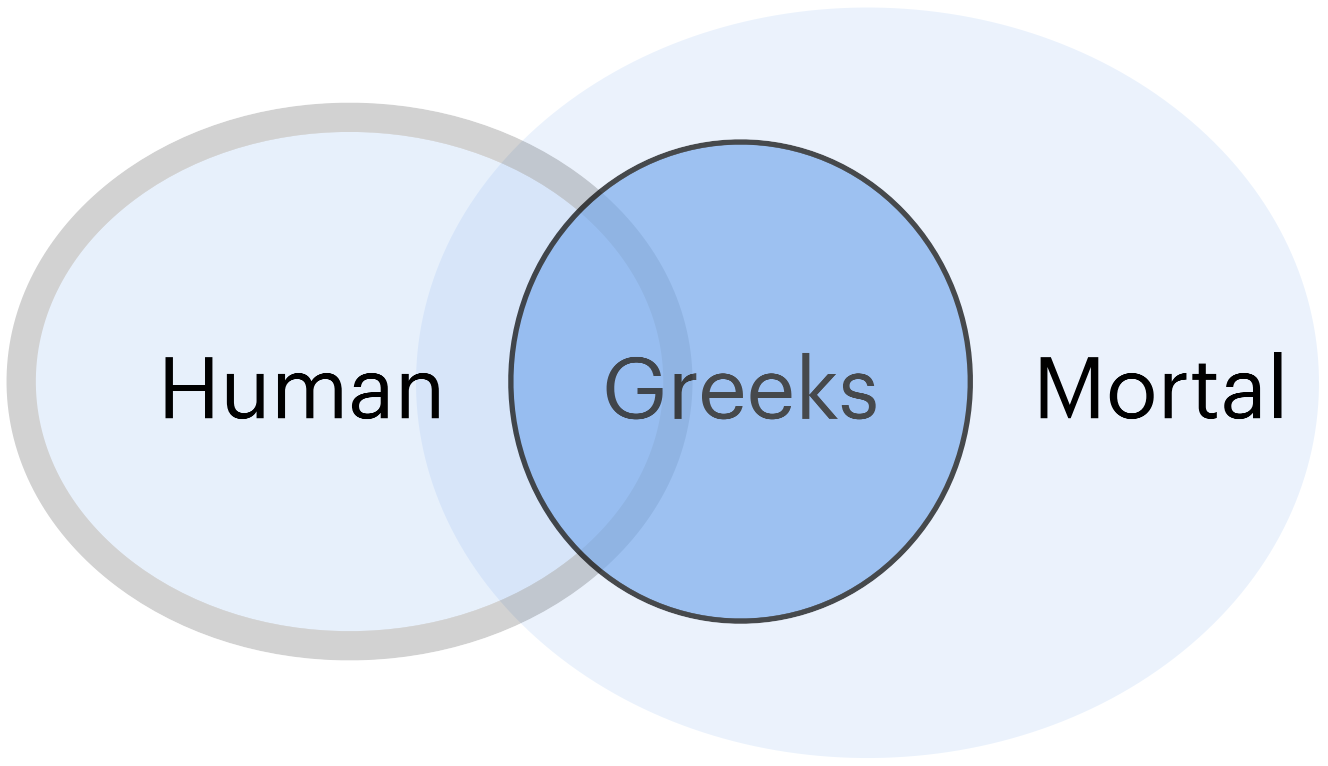


Not all Greeks are human.

Not all human are mortal.

Not all Greeks are mortal.





2000 years ago, Boole came up with the idea of using symbolic variables!

All Greeks are human.

All human are mortal.

All Greeks are mortal.

Replace:

Greeks by p,

Human by q,

Mortal by r

If p then q (p -> q)

If q then r (q ->r)

If p then r (p ->r)

Propositional Logic

- Propositional variables: variables which are either True or False.   (p, q, r, .., x, y)
    - Abstract the information to represent it in a propositional variable
    - Variable p represents “Crazy rich Asians is a good movie”
    - If P is True: “Crazy rich Asians is a good movie” is True sentence.
    - If P is False: “Crazy rich Asians is a good movie” is False sentence.

Propositional Logic

- Propositional variables (p,q,r..)
- Operators:
  - Unary ($\neg$)
  - Binary ($\vee$, $\wedge$, $\oplus$, . . .)
- Punctuations {"(", ")" }

Example: $((p \vee q) \vee r), (\neg(p \vee q))$

Propositional formula or Boolean Formula

Propositional Logic

- τ is a function that maps variables of a propositional formula to {0,1}.

$$F = ((p \vee q) \vee r)$$

$$\tau : \{p \mapsto 1, q \mapsto 0, r \mapsto 1\}$$

P	Q	R
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

- How many such τ can exists ? $2^{variables(F)}$

- τ satisfies formula F if and only if $F(\tau)$ is 1.

$$F(\tau) : ((1 \vee 0) \vee 1) = 1$$

- τ is satisfying assignment for F. We use $\tau \vDash F$ to represent.

Propositional Logic

- If there exists a τ such that $\tau \vDash F$, we say that F is satisfiable.

$F = ((p \vee q) \vee r)$ $\quad \tau : \{p \mapsto 1, q \mapsto 0, r \mapsto 1\}$ F is satisfiable

- If for all τ in $2^{variables(F)}$, $F(\tau)$ is 1, then F is valid.

Is $F = ((p \vee q) \vee r)$ is valid ? Is $F = (p \vee \neg p)$ is valid ?

- If there does not exists a τ in $2^{variables(F)}$ such that $F(\tau)$ is 1, then F is unsatisfiable.

Is $F = ((p \vee q) \vee r)$ is unsatisfiable? Is $F = (p \wedge \neg p)$ is unsatisfiable ?

Propositional Logic

- Set of all satisfying assignment of F is called models. $models(F) = \{\tau \mid F(\tau) = 1\}$

$$Models(\neg F) = \quad 2^{variables} \ \setminus \ Models(F)$$

$$Models(F \vee G) = \quad Models(F) \ \cup \ Models(G)$$

$$Models(F \wedge G) = \quad Models(F) \ \cap \ Models(G)$$

- Equivalent formulas: Two formulas F and G are considered to be equivalent to

each other if and only if they both have same models, that is, if $Models(F) = Models(G), F \equiv G$.

Conjunction Normal Form (CNF)

- $F = (x_1 \vee x_2) \wedge (\neg x_1 \vee x_3)$

Clauses Literals : $x_1, \neg x_1, x_2, \neg x_2, x_3, \neg x_3$

SAT solvers takes

CNF formulas as input.

CNF: $F = C_1 \wedge C_2 \wedge C_3 \ldots \wedge C_m$

where $C_i = (l_1 \vee l_2 \vee \ldots \vee l_k)$

where $l_j = p; l_j = \neg p$

Where p is propositional variable

Can every formula F can be represented in CNF form, say F_{CNF}?

Can every formula F can be represented in CNF form, say F_{CNF}?

Yes, every F can be represented in F_{CNF}, such that $F \equiv F_{CNF}$

$F = ((x_1 \wedge \neg x_2) \vee (x_3 \wedge x_4))$ Can you convert F into F_{CNF}?

$$F_{CNF} = (x_1 \vee x_3) \wedge (x_1 \vee x_4) \wedge (\neg x_2 \vee x_3) \wedge (\neg x_2 \vee x_4)$$

$F = ((x_1 \wedge \neg x_2) \vee (x_3 \wedge x_4)) \vee (x_5 \wedge x_6)$, Can you convert F into F_{CNF}?

$F = (x_1 \wedge y_1) \vee \ldots \vee (x_n \wedge y_n)$, size of equivalent F_{CNF}? 2^n

In the worst case, it may take exponential many steps.

Can we do better?

Equisatisfiable Formulas

- $F = (p \vee \alpha) \wedge (\neg p \vee \beta)$ $\qquad$ $G = (\alpha \vee \beta)$

F and G are Equisatisfiable. F is satisfiable if and only if G is satisfiable.

$F = ((x_1 \wedge \neg x_2) \vee (x_3 \wedge x_4))$ Can you convert F into F_{CNF}?

This is called, Tseytin transformation (https://en.wikipedia.org/wiki/Tseytin_transformation)

$= (t_1 \leftrightarrow (x_1 \wedge \neg x_2)) \wedge (t_2 \leftrightarrow (x_3 \vee x_4)) \wedge (t_1 \vee t_2)$

$= (\neg t_1 \vee (x_1 \wedge \neg x_2)) \wedge (\neg x_1 \vee x_2 \vee t_1) \wedge (\neg t_2 \vee (x_3 \wedge x_4)) \wedge (\neg x_3 \vee \neg x_4 \vee t_2) \wedge (t_1 \vee t_2)$

$= (\neg t_1 \vee x_1) \wedge (\neg t_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_2 \vee t_1) \wedge (\neg t_2 \vee x_3) \wedge (\neg t_2 \vee x_4) \wedge (\neg x_3 \vee \neg x_4 \vee t_2) \wedge (t_1 \vee t_2)$

$= F_{CNF}$

$F = (x_1 \wedge y_1) \vee \ldots \vee (x_n \wedge y_n)$, size of equivalent F_{CNF}? $\quad 2n + n + 1$

Do we really need double implication if we are only interested in satisfiability?

Every formula F can be represented in CNF form, say F_{CNF} in polynomial time such that F is satisfiable if and only if F_{CNF} is satisfiable.

Course Webpage

Teams channel

Send a general request with the request type "Prerequisite Waiver" to register.

Please mail me once you sent the request.

If your general request is pending, don't worry, UG/PG section will get it done by Monday or Tuesday. We will handle all cases!

Class room: Bharti 201