# AAA615: Formal Methods 

# Lecture 1 - Propositional Logic 

Hakjoo Oh
2017 Fall

## Syntax

- Atom: basic elements
- truth symbols $\perp$ ("false") and $\top$ ("true")
- propositional variables $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}, \ldots$
- Literal: an atom $\boldsymbol{\alpha}$ or its negation $\neg \boldsymbol{\alpha}$.
- Formula: a literal or the application of a logical connective (boolean connective) to formulas

| $\boldsymbol{F} \rightarrow$ | $\perp$ |  |
| :---: | :---: | :---: |
| \| | ' |  |
| \| | $\boldsymbol{P}$ |  |
| \| | $\neg F$ | negation ("not") |
| \| | $F_{1} \wedge F_{2}$ | conjunction ("and") |
| \| | $F_{1} \vee F_{2}$ | disjunction (" or") |
| \| | $F_{1} \rightarrow F_{2}$ | implication ("implies") |
| 1 | $F_{1} \leftrightarrow F_{2}$ | iff (" if and only if') |

## Syntax

- Formula $\boldsymbol{G}$ is a subformula of formula $\boldsymbol{F}$ if it occurs syntactically within $G$.

$$
\begin{aligned}
\operatorname{sub}(\perp) & =\{\perp\} \\
\operatorname{sub}( \rceil) & =\{ \rceil\} \\
\operatorname{sub}(P) & =\{P\} \\
\operatorname{sub}(\neg F) & =\{\neg F\} \cup \operatorname{sub}(F) \\
\operatorname{sub}\left(F_{1} \wedge F_{2}\right) & =\left\{F_{1} \wedge F_{2}\right\} \cup \operatorname{sub}\left(F_{1}\right) \cup \operatorname{sub}\left(F_{2}\right)
\end{aligned}
$$

- $F:(P \wedge Q) \rightarrow(P \vee \neg Q)$
- $\operatorname{sub}(F)=$
- The strict subformulas of a formula are all its subformulas except itself.


## Semantics

- The semantics of a logic provides its meaning. The meaning of a PL formula is either true or false.
- The semantics of a formula is defined with an interpretation that assigns truth values to propositional variables, e.g., $I:\{P \mapsto$ true, $Q \mapsto$ false, $\ldots\}$.
- $F: P \wedge Q \rightarrow P \vee \neg Q$
- Inductive definition of semantics:
- We write $\boldsymbol{I} \vDash \boldsymbol{F}$ if $\boldsymbol{F}$ evaluates to true under $\boldsymbol{I}$.
- We write $\boldsymbol{I} \not \models \boldsymbol{F}$ if $\boldsymbol{F}$ evaluates to false under $\boldsymbol{I}$.
$I \vDash \top, \quad I \nvdash \perp$,
$I \vDash P \quad$ iff $I[P]=$ true
$I \not \models P$
$\boldsymbol{I} \vDash \neg \boldsymbol{F}$
$I \vDash F_{1} \wedge F_{2}$
iff $I[P]=$ false
$I \vDash \boldsymbol{F}_{1} \vee \boldsymbol{F}_{\mathbf{2}}$
iff $\boldsymbol{I} \not \models \boldsymbol{F}$
$I \vDash F_{1} \vee F_{2}$
iff $\boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{1}}$ and $\boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{2}}$
$I \vDash F_{1} \rightarrow F_{2}$
iff $\boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{1}}$ or $\boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{2}}$
$I \vDash F_{1} \leftrightarrow F_{2}$
iff $\boldsymbol{I} \not \models \boldsymbol{F}_{\mathbf{1}}$ or $\boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{2}}$
iff $\left(\boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{1}}\right.$ and $\left.\boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{2}}\right)$ or $\left(\boldsymbol{I} \not \models \boldsymbol{F}_{\mathbf{1}}\right.$ and $\left.\boldsymbol{I} \not \models \boldsymbol{F}_{\mathbf{2}}\right)$


## Example

Consider the formula

$$
F: P \wedge Q \rightarrow P \vee \neg Q
$$

and the interpretation

$$
I:\{P \mapsto \text { true }, Q \mapsto \text { false }\}
$$

The truth value of $\boldsymbol{F}$ is computed as follows:

$$
\begin{array}{lll}
\text { 1. } & \boldsymbol{I} \vDash \boldsymbol{P} & \text { since } \boldsymbol{I}[\boldsymbol{P}]=\text { true } \\
\text { 2. } & \boldsymbol{I} \not \vDash \boldsymbol{Q} & \text { since } \boldsymbol{I}[\boldsymbol{Q}]=\text { false } \\
\text { 3. } & \boldsymbol{I} \vDash \neg \boldsymbol{Q} & \text { by } 2 \text { and semantics of } \neg \\
\text { 4. } & \boldsymbol{I} \nLeftarrow \boldsymbol{P} \wedge \boldsymbol{Q} & \text { by } 2 \text { and semantics of } \wedge \\
\text { 5. } & \boldsymbol{I} \vDash \boldsymbol{P} \vee \neg \boldsymbol{Q} & \text { by } 1 \text { and semantics of } \vee \\
\text { 6. } & \boldsymbol{I} \vDash \boldsymbol{F} & \text { by } 4 \text { and semantics of } \rightarrow
\end{array}
$$

## Satisfiability and Validity

- A formula $\boldsymbol{F}$ is satisfiable iff there exists an interpretation $\boldsymbol{I}$ such that $\boldsymbol{I} \vDash \boldsymbol{F}$.
- A formula $\boldsymbol{F}$ is valid iff for all interpretations $\boldsymbol{I}, \boldsymbol{I} \vDash \boldsymbol{F}$.
- Satisfiability and validity are dual ${ }^{1}$ :

$$
\boldsymbol{F} \text { is valid iff } \neg \boldsymbol{F} \text { is unsatisfiable }
$$

- We can check satisfiability by deciding validity, and vice versa.

[^0]
## Deciding Validity and Satisfiability

Two approaches to show $\boldsymbol{F}$ is valid:

- Truth table method performs exhaustive search: e.g., $F: P \wedge Q \rightarrow P \vee \neg Q$.

| $P$ | $Q$ | $P \wedge Q$ | $\neg Q$ | $P \vee \neg Q$ | $F$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 0 | 0 | 1 | 1 | 1 |
| 0 | 1 | 0 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 | 1 | 1 |
| 1 | 1 | 1 | 0 | 1 | 1 |

Impractical and non-applicable to logic with infinite domain (e.g., first-order logic).

- Semantic argument method uses deduction:
- Assume $\boldsymbol{F}$ is invalid: $\boldsymbol{I} \not \models \boldsymbol{F}$ for some $\boldsymbol{I}$ (falsifying interpretation).
- Apply deduction rules (proof rules) to derive a contradiction.
- If every branch of the proof derives a contradiction, then $\boldsymbol{F}$ is valid.
- If some branch of the proof never derives a contradiction, then $\boldsymbol{F}$ is invalid. This branch describes a falsifying interpretation of $\boldsymbol{F}$.
- SAT solvers use both search and deduction.


## Deduction Rules for Propositional Logic

$$
\begin{aligned}
& \frac{I \vDash \neg \boldsymbol{F}}{\boldsymbol{I} \not \models \boldsymbol{F}} \quad \frac{\boldsymbol{I} \not \models \neg \boldsymbol{F}}{\boldsymbol{I} \vDash \boldsymbol{F}} \\
& \frac{\boldsymbol{I} \vDash \boldsymbol{F} \wedge \boldsymbol{G}}{\boldsymbol{I} \vDash \boldsymbol{F}, \boldsymbol{I} \vDash G} \quad \frac{\boldsymbol{I} \not \models \boldsymbol{F} \wedge \boldsymbol{G}}{\boldsymbol{I} \nvdash \boldsymbol{F} \mid \boldsymbol{I} \not \models G} \\
& \frac{\boldsymbol{I} \vDash \boldsymbol{F} \vee G}{I \vDash \boldsymbol{F} \mid \boldsymbol{I} \vDash G} \quad \frac{I \not \vDash \boldsymbol{F} \vee G}{\boldsymbol{I} \nvdash \boldsymbol{F}, \boldsymbol{I} \nvdash \boldsymbol{G}}
\end{aligned}
$$

$$
\begin{aligned}
& \begin{array}{c}
\boldsymbol{I} \vDash \boldsymbol{F} \leftrightarrow \boldsymbol{G} \\
\bar{I} \vDash \boldsymbol{F} \wedge \boldsymbol{G} \mid \boldsymbol{I} \vDash \neg \boldsymbol{F} \wedge \neg \boldsymbol{G}
\end{array} \frac{\boldsymbol{I} \not \vDash \boldsymbol{F} \leftrightarrow \boldsymbol{I} \leftrightarrow}{\boldsymbol{I} \vDash \boldsymbol{F} \wedge \neg \boldsymbol{G} \mid \boldsymbol{I} \vDash \neg \boldsymbol{F} \wedge \boldsymbol{G}} \\
& \frac{\boldsymbol{I} \vDash \boldsymbol{F} \quad \boldsymbol{I} \not \models \boldsymbol{F}}{\boldsymbol{I} \vDash \perp}
\end{aligned}
$$

## Example 1

To prove that the formula

$$
F: P \wedge Q \rightarrow P \vee \neg Q
$$

is valid, assume that it is invalid and derive a contradiction:


## Example 2

To prove that the formula

$$
F:(P \rightarrow Q) \wedge(Q \rightarrow R) \rightarrow(P \rightarrow R)
$$

is valid, assume that it is invalid and derive a contradiction:

$$
\begin{array}{lll}
\text { 1. } & \boldsymbol{I} \not \models \boldsymbol{F} & \text { assumption } \\
\text { 2. } & \boldsymbol{I} \neq(\boldsymbol{P} \rightarrow \boldsymbol{Q}) \wedge(\boldsymbol{Q} \rightarrow \boldsymbol{R}) & \text { by } 1 \text { and semantics of } \rightarrow \\
\text { 3. } & \boldsymbol{I} \not \models \boldsymbol{P} \rightarrow \boldsymbol{R} & \text { by } 1 \text { and semantics of } \rightarrow \\
\text { 4. } & \boldsymbol{I} \nLeftarrow \boldsymbol{P} & \text { by } 3 \text { and semantics of } \rightarrow \\
\mathbf{5 .} & \boldsymbol{I} \not \models \boldsymbol{R} & \text { by } 3 \text { and semantics of } \rightarrow \\
\mathbf{6 .} & \boldsymbol{I} \neq \boldsymbol{P} \rightarrow \boldsymbol{Q} & \text { 2 and semantics of } \wedge \\
\text { 7. } & \boldsymbol{I} \notin \boldsymbol{Q} \rightarrow \boldsymbol{R} & \text { 2 and semantics of } \wedge
\end{array}
$$

Two cases consider from 6:
(1) $\boldsymbol{I} \not \models \boldsymbol{P}$ : contradiction with 4 .
(2) $I \vDash Q$ : two cases to consider from 7:
(1) $I \not \models Q$ : contradiction
(2) $\boldsymbol{I} \vDash \boldsymbol{R}$ : contradiction with 5 .

## Exercise

Prove that the formula

$$
F: P \vee Q \rightarrow P \wedge Q
$$

is valid.

## Derived Rules

The proof rules are sufficient, but derived rules can make proofs more concise. E.g., the rule of modus ponens:

$$
\frac{\boldsymbol{I} \vDash \boldsymbol{F} \quad \boldsymbol{I} \quad \boldsymbol{F} \rightarrow \boldsymbol{G}}{\boldsymbol{I} \vDash \boldsymbol{G}}
$$

The proof of the validity of the formula:

$$
F:(P \rightarrow Q) \wedge(Q \rightarrow R) \rightarrow(P \rightarrow R)
$$

1. $I \nvdash F$
2. $\quad I \vDash(P \rightarrow Q) \wedge(Q \rightarrow R)$
3. $\quad I \not \models P \rightarrow R$
4. $\quad I \vDash P$
5. $I \not \models R$
6. $\quad I \vDash P \rightarrow Q$
7. $\quad I \vDash Q \rightarrow R$
8. $\quad I \vDash Q$
9. $\quad I \vDash R$
10. $I \vDash \perp$
assumption
by 1 and semantics of $\rightarrow$
by 1 and semantics of $\rightarrow$
by 3 and semantics of $\rightarrow$
by 3 and semantics of $\rightarrow$
2 and semantics of $\wedge$
2 and semantics of $\wedge$
by 4,6 , and modus ponens
by 8,7 , and modus ponens
5 and 9 are contradictory

## Equivalence and Implication

- Two formulas $\boldsymbol{F}_{\mathbf{1}}$ and $\boldsymbol{F}_{\mathbf{2}}$ are equivalent

$$
F_{1} \Longleftrightarrow F_{2}
$$

iff $\boldsymbol{F}_{\mathbf{1}} \leftrightarrow \boldsymbol{F}_{\mathbf{2}}$ is valid, i.e., for all interpretations $\boldsymbol{I}, \boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{1}} \leftrightarrow \boldsymbol{F}_{\mathbf{2}}$.

- Formula $\boldsymbol{F}_{\mathbf{1}}$ implies formula $\boldsymbol{F}_{\mathbf{2}}$

$$
F_{1} \Longrightarrow F_{2}
$$

iff $\boldsymbol{F}_{\mathbf{1}} \rightarrow \boldsymbol{F}_{\mathbf{2}}$ is valid, i.e., for all interpretations $\boldsymbol{I}, \boldsymbol{I} \vDash \boldsymbol{F}_{\mathbf{1}} \rightarrow \boldsymbol{F}_{\mathbf{2}}$.

- $\boldsymbol{F}_{\mathbf{1}} \Longleftrightarrow \boldsymbol{F}_{\mathbf{2}}$ and $\boldsymbol{F}_{\mathbf{1}} \Longrightarrow \boldsymbol{F}_{\mathbf{2}}$ are not formulas. They are semantic assertions.
- We can check equivalence and implication by checking satisfiability.


## Exercise

Prove that

$$
R \wedge(\neg R \wedge P) \Longrightarrow P
$$

## Substitution

- A substitution $\sigma$ is a mapping from formulas to formulas:

$$
\sigma:\left\{F_{1} \mapsto G_{2}, \ldots, F_{n} \mapsto G_{n}\right\}
$$

- The domain of $\sigma, \operatorname{dom}(\sigma)$, is

$$
\operatorname{dom}(\sigma):\left\{F_{1}, \ldots, F_{n}\right\}
$$

while the range range $(\sigma)$ is

$$
\operatorname{range}(\sigma):\left\{G_{1}, \ldots, G_{n}\right\}
$$

- The application of a substitution $\boldsymbol{\sigma}$ to a formula $\boldsymbol{F}, \boldsymbol{F} \boldsymbol{\sigma}$, replaces each occurence of $\boldsymbol{F}_{\boldsymbol{i}}$ with $\boldsymbol{G}_{\boldsymbol{i}}$. Replacements occur all at once.
- When two subformulas $\boldsymbol{F}_{\boldsymbol{j}}$ and $\boldsymbol{F}_{\boldsymbol{k}}$ are in $\boldsymbol{\operatorname { d o m }}(\boldsymbol{\sigma})$ and $\boldsymbol{F}_{\boldsymbol{k}}$ is a strict subformula of $\boldsymbol{F}_{\boldsymbol{j}}$, then $\boldsymbol{F}_{\boldsymbol{j}}$ is replaced first.


## Example

Consider formula

$$
F: P \wedge Q \rightarrow P \vee \neg Q
$$

and substitution

$$
\sigma:\{P \mapsto R, P \wedge Q \mapsto P \rightarrow Q\}
$$

Then,

$$
F \sigma:(P \rightarrow Q) \rightarrow R \vee \neg Q
$$

## Substitution

- A variable substitution is a substitution in which the domain consists only of propositional variables.
- When we write $\boldsymbol{F}\left[\boldsymbol{F}_{\mathbf{1}}, \ldots, \boldsymbol{F}_{\boldsymbol{n}}\right]$, we mean that formula $\boldsymbol{F}$ can have formulas $\boldsymbol{F}_{1}, \ldots, \boldsymbol{F}_{\boldsymbol{n}}$ as subformulas.
- If $\sigma$ is $\left\{\boldsymbol{F}_{\mathbf{1}} \mapsto \boldsymbol{G}_{\boldsymbol{1}}, \ldots, \boldsymbol{F}_{\boldsymbol{n}} \mapsto \boldsymbol{G}_{\boldsymbol{n}}\right\}$, then

$$
F\left[F_{1}, \ldots, F_{n}\right] \sigma: F\left[G_{1}, \ldots, G_{n}\right]
$$

- For example, in the previous example, writing

$$
F[P, P \wedge Q] \sigma: F[R, P \rightarrow Q]
$$

emphasizes that $\boldsymbol{P}$ and $\boldsymbol{P} \wedge \boldsymbol{Q}$ are replaced by $\boldsymbol{R}$ and $\boldsymbol{P} \rightarrow \boldsymbol{Q}$, respectively.

## Semantic Consequences of Substitution

## Lemma (Substitution of Equivalent Formulas)

Consider substitution $\sigma:\left\{F_{1} \mapsto \boldsymbol{G}_{1}, \ldots, \boldsymbol{F}_{\boldsymbol{n}} \mapsto \boldsymbol{G}_{\boldsymbol{n}}\right\}$ such that for each $\boldsymbol{i}, \boldsymbol{F}_{i} \Longleftrightarrow \boldsymbol{G}_{\boldsymbol{i}}$. Then, $\boldsymbol{F} \Longleftrightarrow \boldsymbol{F} \boldsymbol{\sigma}$.

```
Lemma (Valid Template)
If F}\mathrm{ is valid and }\boldsymbol{G}=\boldsymbol{F}\boldsymbol{\sigma}\mathrm{ for some variable substitution }\boldsymbol{\sigma}\mathrm{ , then }\boldsymbol{G}\mathrm{ is
valid.
```

For example, because

$$
F:(P \rightarrow Q) \leftrightarrow(\neg P \vee Q)
$$

is valid, every formula of the form $\boldsymbol{F}_{\mathbf{1}} \rightarrow \boldsymbol{F}_{\mathbf{2}}$ is equivalent to $\neg \boldsymbol{F}_{\mathbf{1}} \vee \boldsymbol{F}_{\mathbf{2}}$, for arbitrary formulas $\boldsymbol{F}_{\mathbf{1}}$ and $\boldsymbol{F}_{\mathbf{2}}$.

## Composition of Substitutions

Given substitutions $\sigma_{1}$ and $\sigma_{2}$, their composition $\sigma=\sigma_{1} \sigma_{2}$ ("apply $\sigma_{1}$ and then $\sigma_{2}{ }^{\prime \prime}$ ) is computed as follows:
(1) Apply $\sigma_{2}$ to each formula of the range of $\sigma_{1}$, and add the results to $\sigma$.
(2) If $\boldsymbol{F}_{\boldsymbol{i}}$ of $\boldsymbol{F}_{\boldsymbol{i}} \mapsto \boldsymbol{G}_{\boldsymbol{i}}$ appears in the domain of $\boldsymbol{\sigma}_{\boldsymbol{2}}$ but not in the domain of $\sigma_{1}$, then add $\boldsymbol{F}_{i} \mapsto \boldsymbol{G}_{\boldsymbol{i}}$ to $\boldsymbol{\sigma}$.
For example,

$$
\begin{aligned}
& \sigma_{1} \sigma_{2}:\{P \mapsto R, P \wedge Q \mapsto P \rightarrow Q\}\{P \mapsto S, S \mapsto Q\} \\
& \quad=\left\{P \mapsto R \sigma_{2}, P \wedge Q \mapsto(P \rightarrow Q) \sigma_{2}, S \mapsto Q\right\} \\
& \quad=\{P \mapsto R, P \wedge Q \mapsto S \rightarrow Q, S \mapsto Q\}
\end{aligned}
$$

## Normal Forms

A normal form of formulas is a syntactic restriction such that for every formula of the logic, there is an equivalent formula in the normal form. Three useful normal forms in logic:

- Negation Normal Form (NNF)
- Disjunctive Normal Form (DNF)
- Conjunctive Normal Form (CNF)


## Negation Normal Form (NNF)

- NNF requires that $\neg, \wedge$, and $\vee$ are the only connectives (i.e., no $\rightarrow$ and $\leftrightarrow)$ and that negations appear only in literals.
- $P \wedge Q \wedge(R \vee \neg S)$
- $\neg P \vee \neg(P \wedge Q))$
- $\neg \neg P \wedge Q$
- Transforming a formula $\boldsymbol{F}$ to equivalent formula $\boldsymbol{F}^{\prime}$ in NNF can be done by repeatedly applying the following list of template equivalences:

$$
\begin{aligned}
& \begin{aligned}
\neg \neg \boldsymbol{F}_{\mathbf{1}} & \Longleftrightarrow \boldsymbol{F}_{\mathbf{1}} \\
\neg \top & \Longleftrightarrow \perp \\
\neg \perp & \Longleftrightarrow \top
\end{aligned} \\
& \neg\left(F_{1} \wedge F_{2}\right) \Longleftrightarrow \neg F_{1} \vee \neg F_{2} \\
& \neg\left(F_{1} \vee F_{2}\right) \Longleftrightarrow \neg F_{1} \wedge \neg F_{2} \\
& F_{1} \rightarrow F_{2} \Longleftrightarrow \neg F_{1} \vee F_{2} \\
& F_{1} \leftrightarrow F_{2} \Longleftrightarrow\left(F_{1} \rightarrow F_{2}\right) \wedge\left(F_{2} \rightarrow F_{1}\right)
\end{aligned}
$$

## Exercise

Convert $F: \neg(P \rightarrow \neg(P \wedge Q))$ into NNF.

## Disjunctive Normal Form (DNF)

- A formula is in disjunctive normal form (DNF) if it is a disjunction of conjunctive clauses (conjunctions of literals):

$$
\bigvee_{i} \bigwedge_{j} l_{i, j}
$$

- To convert a formula $\boldsymbol{F}$ into an equivalent formula in DNF, transform $\boldsymbol{F}$ into NNF and then distribute conjunctions over disjunctions:

$$
\begin{aligned}
& \left(F_{1} \vee F_{2}\right) \wedge F_{3} \Longleftrightarrow\left(F_{1} \wedge F_{3}\right) \vee\left(F_{2} \wedge F_{3}\right) \\
& F_{1} \wedge\left(F_{2} \vee F_{3}\right) \quad \Longleftrightarrow\left(F_{1} \wedge F_{2}\right) \vee\left(F_{1} \wedge F_{3}\right)
\end{aligned}
$$

## Exercise

To convert

$$
F:\left(Q_{1} \vee \neg \neg Q_{2}\right) \wedge\left(\neg \boldsymbol{R}_{1} \rightarrow \boldsymbol{R}_{2}\right)
$$

into DNF,

- first transform it into NNF:
- then apply distributivity:


## Conjunctive Normal Form (CNF)

- A formula is in conjunctive normal form (CNF) if it is a conjunction of clauses (i.e. conjunctions of disjunctions of literals):

$$
\bigwedge_{i} \bigvee_{j} l_{i, j}
$$

- To convert a formula $\boldsymbol{F}$ into an equivalent formula in DNF, transform $\boldsymbol{F}$ into NNF and distribute disjunctions over conjunctions:

$$
\begin{aligned}
& \left(F_{1} \wedge F_{2}\right) \vee F_{3} \quad \Longleftrightarrow\left(F_{1} \vee F_{3}\right) \wedge\left(F_{2} \vee F_{3}\right) \\
& F_{1} \vee\left(F_{2} \wedge F_{3}\right) \quad \Longleftrightarrow\left(F_{1} \vee F_{2}\right) \wedge\left(F_{1} \vee F_{3}\right)
\end{aligned}
$$

- Exercise) Convert $\boldsymbol{F}:\left(\boldsymbol{Q}_{1} \wedge \neg \neg \boldsymbol{Q}_{2}\right) \vee\left(\neg \boldsymbol{R}_{1} \rightarrow \boldsymbol{R}_{2}\right)$ into CNF.


## Decision Procedures

- A decision procedure decides whether $\boldsymbol{F}$ is satisfiable after some finite steps of computation.
- Approaches for deciding satisfiability:
- Search: exhaustively search through all possible assignments
- Deduction: deduce facts from known facts by iteratively applying proof rules
- Combination: Modern SAT solvers are based on DPLL that combines search and deduction in an effective way


## Exhaustive Search

- The recursive algorithm for deciding satisfiability:

$$
\begin{aligned}
& \text { let rec SAT } \boldsymbol{F}= \\
& \text { if } \boldsymbol{F}=\mathrm{T} \text { then true } \\
& \text { else if } \boldsymbol{F}=\perp \text { then false } \\
& \text { else } \\
& \text { let } \boldsymbol{P}=\mathbf{C h o o s e}(\operatorname{vars}(\boldsymbol{F})) \text { in } \\
& \quad(\text { SAT } \boldsymbol{F}\{\boldsymbol{P} \mapsto \top\}) \vee(\text { SAT } \boldsymbol{F}\{\boldsymbol{P} \mapsto \perp\})
\end{aligned}
$$

- When applying $\boldsymbol{F}\{\boldsymbol{P} \mapsto \top\}$ and $\boldsymbol{F}\{\boldsymbol{P} \mapsto \perp\}$, the resulting formulas should be simplified using template equivalences on PL.



## Example

$$
F:(P \rightarrow Q) \wedge P \wedge \neg Q
$$

- Choose variable $\boldsymbol{P}$ and

$$
\boldsymbol{F}\{\boldsymbol{P} \mapsto \top\}:(\top \rightarrow Q) \wedge \top \wedge \neg Q
$$

which simplifies to

$$
F_{1}: Q \wedge \neg Q
$$

- $F_{1}\{Q \mapsto \top\}: \perp$
- $F_{1}\{Q \mapsto \perp\}: \perp$
- Recurse on the other branch for $\boldsymbol{P}$ in $\boldsymbol{F}$ :

$$
F\{P \mapsto \perp\}:(\perp \rightarrow Q) \wedge \perp \wedge \neg Q
$$

which simplifies to $\perp$.

- All branches end without finding a satisfying assignment.


## Example

$$
F:(P \rightarrow Q) \wedge \neg P
$$

- Choose $\boldsymbol{P}$ and recurse on the first case:

$$
F\{P \mapsto \top\}:(\top \rightarrow Q) \wedge \neg T
$$

which is equivalent to $\perp$.

- Try the other case:

$$
F\{P \rightarrow \perp\}:(\perp \rightarrow Q) \wedge \neg \perp
$$

which is equivalent to $T$.

- Arbitrarily assigning a value to $Q$ produces the satisfying interpretation:

$$
I:\{P \mapsto \text { false, } Q \mapsto \text { true }\}
$$

## Equisatisfiability

- SAT solvers convert a given formula $\boldsymbol{F}$ to CNF.
- Conversion to an equivalent CNF incurs exponential blow-up in worst-case.
- $\boldsymbol{F}$ is converted to an equisatisfiable CNF formula, which increases the size by only a constant factor.
- $\boldsymbol{F}$ and $\boldsymbol{F}^{\prime}$ are equisatisfiable when $\boldsymbol{F}$ is satisfiable iff $\boldsymbol{F}^{\prime}$ is satisfiable.
- Equisatisfiability is a weaker notion of equivalence, which is still useful when deciding satisfiability.


## Conversion to an Equisatisfiable Formula in CNF

- Idea: Introduce new variables to represent the subformulas of $\boldsymbol{F}$ with extra clauses that assert that these new variables are equivalent to the subformulas that they represent.
- $F: x_{1} \rightarrow\left(x_{2} \wedge x_{3}\right)$
- Introduce two variables $a_{1}$ and $a_{2}$ with two equivalences:

$$
\begin{aligned}
& a_{1} \leftrightarrow\left(x_{1} \rightarrow a_{2}\right) \\
& a_{2} \leftrightarrow\left(x_{2} \wedge x_{3}\right)
\end{aligned}
$$

We need to satisfy $a_{1}$, together with the above two equivalences.

- Convert the equivalences to CNF:

$$
\begin{aligned}
& \left(a_{1} \vee x_{1}\right) \wedge\left(a_{1} \vee \neg a_{2}\right) \wedge\left(\neg a_{1} \vee \neg x_{1} \vee a_{2}\right) \\
& \left(\neg a_{2} \vee x_{2}\right) \wedge\left(\neg a_{2} \vee x_{3}\right) \wedge\left(a_{2} \vee \neg x_{2} \vee \neg x_{3}\right)
\end{aligned}
$$

- The final CNF formula:

$$
\begin{aligned}
F^{\prime}= & a_{1} \wedge\left(a_{1} \vee x_{1}\right) \wedge\left(a_{1} \vee \neg a_{2}\right) \wedge\left(\neg a_{1} \vee \neg x_{1} \vee a_{2}\right) \wedge \\
& \left(\neg a_{2} \vee x_{2}\right) \wedge\left(\neg a_{2} \vee x_{3}\right) \wedge\left(a_{2} \vee \neg x_{2} \vee \neg x_{3}\right)
\end{aligned}
$$

- $\boldsymbol{F}$ is satisfiable iff $\boldsymbol{F}^{\prime}$ is satisfiable


## The Resolution Procedure

- Applicable only to CNF formulas.
- Observation: to satisfy clauses $\boldsymbol{C}_{\mathbf{1}}[\boldsymbol{P}]$ and $\boldsymbol{C}_{\mathbf{2}}[\neg P]$ that share variable $\boldsymbol{P}$ but disagree on its value, either the rest of $\boldsymbol{C}_{\mathbf{1}}$ or the rest of $C_{2}$ must be satisfied. Why?
- The clause $C_{1}[\perp] \vee C_{2}[\perp]$ (with simplification) can be added as a conjunction to $\boldsymbol{F}$ to produce an equivalent formula still in CNF.
- The proof rule for clausal resolution:

$$
\frac{C_{1}[\boldsymbol{P}] \quad C_{2}[\neg P]}{C_{1}[\perp] \vee C_{2}[\perp]}
$$

The new clause $C_{1}[\perp] \vee C_{2}[\perp]$ is called the resolvent.

- If ever $\perp$ is deduced via resolution, $\boldsymbol{F}$ must be unsatisfiable.

Otherwise, if no further resolutions are possible, $\boldsymbol{F}$ must be satisfiable.

## Examples

$$
F:(\neg P \vee Q) \wedge P \wedge \neg Q
$$

- From resolution

$$
\frac{(\neg P \vee Q) \quad P}{Q}
$$

construct $(\neg P \vee Q) \wedge P \wedge \neg Q \wedge Q$. From resolution

deduce that $\boldsymbol{F}$ is unsatisfiable.

## Examples

$$
F:(\neg P \vee Q) \wedge \neg Q)
$$

- The resolution procedure yields

$$
(\neg P \vee Q) \wedge \neg Q \wedge \neg P
$$

No further resolutions are possible. $\boldsymbol{F}$ is satisfiable.

- A satisfying interpretation:

$$
I:\{P \mapsto \text { false }, Q \mapsto \text { false }\}
$$

- A CNF formula that does not contain the clause $\perp$ and to which no more resolutions are applicable represents all possible satisfying interpretations.


## DPLL

- The Davis-Putnam-Logemann-Loveland algorithm (DPLL) combines the enumerative search and a restricted form of resolution, called unit resolution:

$$
\frac{l \quad C[\neg l]}{C[\perp]}
$$

where $\boldsymbol{l}$ is a literal $(\boldsymbol{l}=\boldsymbol{P}$ or $\boldsymbol{l}=\neg \boldsymbol{P})$.

- The process of applying this resolution as much as possible is called Boolean constraint propagation (BCP).


## BCP Example

$$
F:(P) \wedge(\neg P \vee Q) \wedge(R \vee \neg Q \vee S)
$$

- Apply unit resolution

$$
\frac{P \quad(\neg P \vee Q)}{Q}
$$

to produce $\boldsymbol{F}^{\prime}: Q \wedge(R \vee \neg Q \vee S)$. Applying unit resolution

$$
\begin{array}{ll}
Q & R \vee \neg Q \vee S \\
\hline & R \vee S
\end{array}
$$

produces $\boldsymbol{F}^{\prime \prime}: \boldsymbol{R} \vee \boldsymbol{S}$, ending this round of BCP.

## DPLL

DPLL is similar to SAT, except that it begins by applying BCP:

$$
\begin{aligned}
& \text { let rec } \mathbf{D P L L} \boldsymbol{F}= \\
& \text { let } \boldsymbol{F}^{\prime}=\mathbf{B C P}(\boldsymbol{F}) \text { in } \\
& \text { if } \boldsymbol{F}^{\prime}=\top \text { then true } \\
& \text { else if } \boldsymbol{F}^{\prime}=\perp \text { then false } \\
& \text { else } \\
& \quad \text { let } \boldsymbol{P}=\mathbf{C h o o s e}\left(\operatorname{vars}\left(\boldsymbol{F}^{\prime}\right)\right) \text { in } \\
& \quad\left(\mathrm{DPLL} \boldsymbol{F}^{\prime}\{\boldsymbol{P} \mapsto \top\}\right) \vee\left(\mathbf{D P L L} \boldsymbol{F}^{\prime}\{\boldsymbol{P} \mapsto \perp\}\right)
\end{aligned}
$$

## Pure Literal Propagation (PLP)

- If variable $\boldsymbol{P}$ appears only positively or only negatively in $\boldsymbol{F}$, remove all clauses containing an instance of $\boldsymbol{P}$.
- If $\boldsymbol{P}$ appears only positively (i.e. no $\neg \boldsymbol{P}$ in $\boldsymbol{F}$ ), replace $\boldsymbol{P}$ by T .
- If $\boldsymbol{P}$ appears only negatively (i.e. no $\boldsymbol{P}$ in $\boldsymbol{F}$ ), replace $\boldsymbol{P}$ by $\perp$.
- The resulting formula $\boldsymbol{F}^{\prime}$ is equisatisfiable to $\boldsymbol{F}$.
- When only such pure variables remain, the formula must be satisfiable. A full interpretation can be constructed by setting each variable's value based on whether it appears only positively (true) or only negatively (false).

Example) $\boldsymbol{F}:(\neg P \vee Q) \wedge(R \vee \neg Q \vee S)$.

- $\boldsymbol{P}$ appears only negatively in $\boldsymbol{F}$

$$
F^{\prime}:(R \vee \neg Q \vee S)
$$

- $\boldsymbol{R}$ and $\boldsymbol{S}$ appear only positively in $\boldsymbol{F}$

$$
F^{\prime}:(\neg P \vee Q)
$$

## DPLL with PLP

```
let rec DPLL \(\boldsymbol{F}=\)
    let \(\boldsymbol{F}^{\boldsymbol{\prime}}=\mathbf{P L P}(\mathbf{B C P}(\boldsymbol{F}))\) in
    if \(\boldsymbol{F}^{\prime}=\top\) then true
    else if \(\boldsymbol{F}^{\prime}=\perp\) then false
    else
    let \(\boldsymbol{P}=\mathbf{C h o o s e}\left(\operatorname{vars}\left(\boldsymbol{F}^{\prime}\right)\right)\) in
    \(\left(\right.\) DPLL \(\left.\boldsymbol{F}^{\prime}\{\boldsymbol{P} \mapsto \top\}\right) \vee\left(\right.\) DPLL \(\left.\boldsymbol{F}^{\prime}\{\boldsymbol{P} \mapsto \perp\}\right)\)
```


## Example 1

$$
F: P \wedge(\neg P \vee Q) \wedge(R \vee \neg Q \vee S)
$$

(1) Applying BCP produces

$$
F^{\prime \prime}: R \vee S
$$

(2) All variables occur positively, so $\boldsymbol{F}$ is satisfiable.
(3) A satisfying interpretation:

$$
\{P \mapsto \text { true }, Q \mapsto \text { true, } R \mapsto \text { true, } S \mapsto \text { true }\}
$$

## Example 2

$$
F:(\neg P \vee Q \vee R) \wedge(\neg Q \vee R) \wedge(\neg Q \vee \neg R) \wedge(P \vee \neg Q \vee \neg \boldsymbol{R})
$$

- No BCP and PLP are applicable.
- Choose $\boldsymbol{Q}$ to branch on:

$$
F\{Q \mapsto \top\}: R \wedge(\neg R) \wedge(P \vee \neg R)
$$

The unit resolution with $\boldsymbol{R}$ and $\neg \boldsymbol{R}$ deduces $\perp$, finishing this branch.

- On the other branch for $\boldsymbol{Q}$ :

$$
F\{Q \mapsto \perp\}:(\neg P \vee R)
$$

$\boldsymbol{P}$ and $\boldsymbol{R}$ are pure, so the formula is satisfiable. A satisfying interpretation:

$$
I:\{P \mapsto \text { false, } Q \mapsto \text { false, } R \mapsto \text { true }\}
$$

## Summary

- Syntax and semantics of propositional logic
- Satisfiability and validity
- Equivalence, implications, and equisatisfiability
- Substitution
- Normal forms: NNF, DNF, CNF
- Decision procedures for satisfiability


[^0]:    ${ }^{1}$ In logic, functions (or relations) $\boldsymbol{A}$ and $\boldsymbol{B}$ are dual if $\boldsymbol{A}(\boldsymbol{x})=\neg \boldsymbol{B}(\neg \boldsymbol{x})$

