Artificial Intelligence

A course about foundations

Propositional Logic

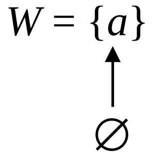
Marco Piastra

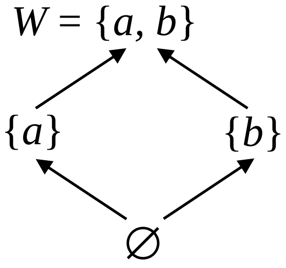
UNIVERSITÀ DI PAVIA

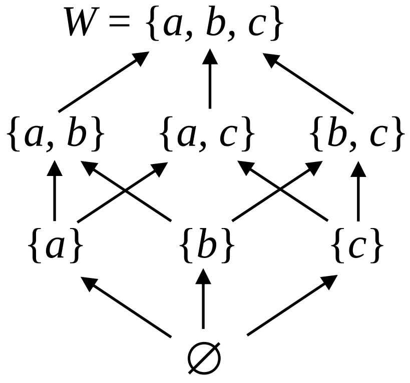
Prologue:
Boolean Algebra(s)

Boolean algebras by examples

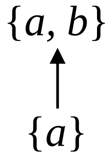
Consider a *finite* set of objects W and construct, in a *bottom-up fashion*, the collection Σ of all possible <u>subsets</u> of W







...



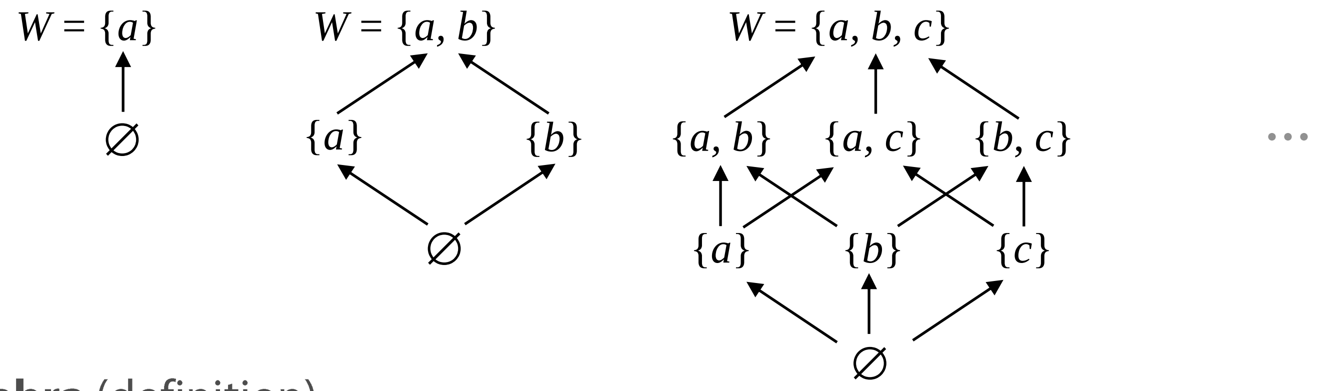
Each arrow represents proper inclusion $\{a\} \subset \{a, b\}$

(Hasse diagrams)

Collections like Σ above are also called the ***power set*** of W which is the collection of all possible subsets of W, also denoted as 2^W

Boolean algebra<u>s</u> by examples

Consider a *finite* set of objects W and construct, in a *bottom-up fashion*, the collection Σ of all possible <u>subsets</u> of W



Boolean algebra (definition)

<u>Any</u> non-empty collection of subsets Σ of a set W such that:

1) $\varnothing \in \Sigma$
2) $A, B \in \Sigma \implies A \cup B \in \Sigma$
3) $A \in \Sigma \implies A^c \in \Sigma$

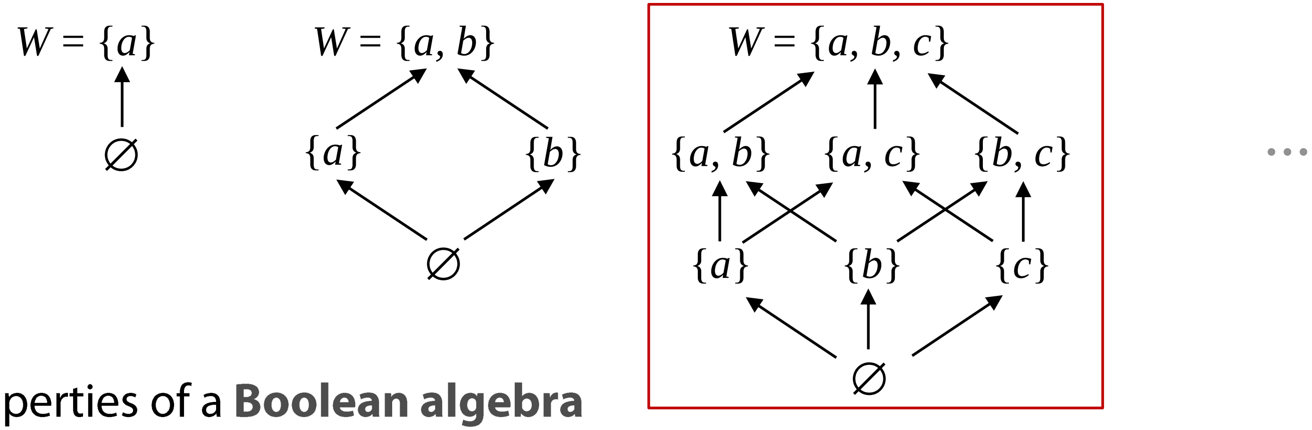
$A^c := W - A$ *(the complement of* A *with respect to* W *)*

Corollaries:

- The set W belongs to any Boolean algebra generated on W
- Σ is closed under *intersection*

Boolean algebras by examples

Consider a *finite* set of objects W and construct, in a *bottom-up fashion*, the collection Σ of all possible <u>subsets</u> of W



Checking properties of a **Boolean algebra**

De Morgan's laws

$(A \cup B)^c = A^c \cap B^c$

$(A \cap B)^c = A^c \cup B^c$

For any of the structures above properties can be verified exhaustively…

$A = \{b\}$
$B = \{b, c\}$
$A \cup B = \{b, c\}$
$(A \cup B)^c = \{a\}$
$A^c = \{a, c\}$
$B^c = \{a\}$
$A^c \cap B^c = \{a\}$

These sets are identical

$A = \{b\}$
$B = \{b, c\}$
$A \cap B = \{b\}$
$(A \cap B)^c = \{a, c\}$
$A^c = \{a, c\}$
$B^c = \{a\}$
$A^c \cup B^c = \{a, c\}$

These sets are identical

Which Boolean algebra for logic?

Given that all boolean algebras share the same properties
we can adopt the simplest one as reference: the one based on $\Sigma := \{W, \varnothing\}$

This is a *two-valued* algebra: $\{nothing, everything\}$ or $\{false, true\}$ or $\{\bot, \top\}$ or $\{0, 1\}$

- Algebraic structure

  < {0,1}, OR, AND, NOT >

- Boolean functions and *truth tables*

  Most generic type of boolean functions: $f : \{0, 1\}^n \rightarrow \{0, 1\}$

  AND, OR and NOT are boolean functions, defined explicitly via *truth tables*

A	B	OR
0	0	0
0	1	1
1	0	1
1	1	1

A	B	AND
0	0	0
0	1	0
1	0	0
1	1	1

A	NOT
0	1
1	0

Composite functions

Truth tables can be defined also for composite functions

For example, to verify logical laws

These columns are identical

A	*B*	NOT *A*	NOT *B*	*A* OR *B*	NOT(*A* OR *B*)	NOT *A* AND NOT *B*
0	0	1	1	0	1	1
0	1	1	0	1	0	0
1	0	0	1	1	0	0
1	1	0	0	1	0	0

De Morgan's laws

These columns are identical

A	*B*	NOT *A*	NOT *B*	*A* AND *B*	NOT(*A* AND *B*)	NOT *A* OR NOT *B*
0	0	1	1	0	1	1
0	1	1	0	0	1	1
1	0	0	1	0	1	1
1	1	0	0	1	0	0

Adequate basis

- How many boolean functions do we need to define *any* boolean function?

2^n *rows*

A_1	A_2	...	A_n	$f(A_1, A_2, ..., A_n)$
0	0	...	0	f_1
0	0	...	1	f_2
...	...	...	...	...
...	...	...	...	...
1	1	...	1	f_{2^n}

Just OR, AND and NOT : any other function can be expressed as <u>composite</u> function

In the generic *truth table* above:

1) For each row j where $f_j = 1$, create a Boolean expression composing by AND the n input variables, taking either A_i , when the i-th value is 1, or NOT A_i when i-th value is 0
2) Compose with OR all expressions obtained in the way above

Other adequate bases

Also {OR, NOT} o {AND, NOT} are adequate bases

An adequate basis can be obtained by just one 'ad hoc' function: NOR or NAND

A	*B*	*A* NOR *B*
0	0	1
0	1	0
1	0	0
1	1	0

A	*B*	*A* NAND *B*
0	0	1
0	1	1
1	0	1
1	1	0

- Two remarkable functions: *implication* and *equivalence*

A	*B*	*A* IMP *B*
0	0	1
0	1	1
1	0	0
1	1	1

A	*B*	*A* EQU *B*
0	0	1
0	1	0
1	0	0
1	1	1

Identities: *A* IMP *B* = (NOT *A*) OR *B* *A* EQU *B* = (*A* IMP *B*) AND (*B* IMP *A*)

In passing, logicians prefer the adequate basis {IMP, NOT}

Language and Semantics:
possible worlds

Propositional logic: the project

The simplest of 'classical' logics

- Propositions

  We consider simple ***propositions*** which state something that could be either true or false

  "Today is Friday"
  "Turkeys are birds with feathers"
  "Man is a featherless biped"

- Formal ***language***

  A precise and formal language whose **atoms** are *propositions*
  (no intention to represent the internal structure of *propositions*)

  Atoms will be composed in complex formulae via a set of *syntactic* rules

- Formal ***semantics***

  A class of *formal* structures, each representing a ***possible world*** or a possible 'state of things'

  <This classroom right now>
  <My uncle's farm several years ago>
  <Ancient Greece at the time of Aristotle's birth>

The class of propositional, semantic structures

Each possible world is a structure $<\{0,1\}, \Sigma, v>$

$\{0,1\}$ are the *truth values*

Σ is the ***signature*** of the formal language: a set of propositional symbols

v is a *function* $: \Sigma \to \{0,1\}$ assigning truth values to the symbols in Σ

Propositional symbols (*signature*)

Each symbol in Σ stands for an actual *proposition* (in natural language)

In the simple convention, we use the symbols $A, B, C, D, \ldots$

Caution: Σ is not necessarily *finite*

Possible worlds

The class of structures contains all possible worlds:

$<\{0,1\}, \Sigma, v>$
$<\{0,1\}, \Sigma, v'>$
$<\{0,1\}, \Sigma, v''>$
...

Each class of structure shares Σ and $\{0,1\}$

The functions v are different: the assignment of truth values varies, depending on the possible world

Formal *language*

- In a propositional language L_P

  A set Σ of propositional symbols: $\Sigma = \{A, B, C, \ldots\}$

  Two (primary) ***logical connectives***: $\neg, \rightarrow$

  Three (derived) ***logical connectives***: $\wedge, \vee, \leftrightarrow$

  Parenthesis: (,) (there are no *precedence rules* in this language)

- Well-formed formulae (**wff**)

  Defined via a set of syntactic rules:

  The set of all the **wff** of L_P is denoted as $\mathrm{wff}(L_P)$

  $A \in \Sigma \;\Rightarrow\; A \in \mathrm{wff}(L_P)$

  $\varphi \in \mathrm{wff}(L_P) \;\Rightarrow\; (\neg\varphi) \in \mathrm{wff}(L_P)$

  $\varphi, \psi \in \mathrm{wff}(L_P) \;\Rightarrow\; (\varphi \rightarrow \psi) \in \mathrm{wff}(L_P)$

  $\varphi, \psi \in \mathrm{wff}(L_P) \;\Rightarrow\; (\varphi \vee \psi) \in \mathrm{wff}(L_P), \quad (\varphi \vee \psi) \;\Leftrightarrow\; ((\neg\varphi) \rightarrow \psi)$

  $\varphi, \psi \in \mathrm{wff}(L_P) \;\Rightarrow\; (\varphi \wedge \psi) \in \mathrm{wff}(L_P), \quad (\varphi \wedge \psi) \;\Leftrightarrow\; (\neg(\varphi \rightarrow (\neg\psi)))$

  $\varphi, \psi \in \mathrm{wff}(L_P) \;\Rightarrow\; (\varphi \leftrightarrow \psi) \in \mathrm{wff}(L_P), \quad (\varphi \leftrightarrow \psi) \;\Leftrightarrow\; ((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$

Formal *semantics*: interpretations

- Compositional (*truth-functional*) semantics for wff

Given a possible world $<\{0,1\}, \Sigma, v>$
the function $v : \Sigma \to \{0,1\}$ can be <u>extended</u> to assign a value to *every* wff
by associating binary (i.e., Boolean) functions to connectives:

$$v(\neg\varphi) = \mathrm{NOT}(v(\varphi))$$

$$v(\varphi \wedge \psi) = \mathrm{AND}(v(\varphi), v(\psi))$$

$$v(\varphi \vee \psi) = \mathrm{OR}(v(\varphi), v(\psi))$$

$$v(\varphi \to \psi) = \mathrm{OR}(\mathrm{NOT}(v(\varphi)), v(\psi)) \quad \text{(also } \mathrm{IMP}(v(\varphi), v(\psi))\text{)}$$

$$v(\varphi \leftrightarrow \psi) = \mathrm{AND}(\mathrm{OR}(\mathrm{NOT}(v(\varphi)), v(\psi)), \mathrm{OR}(\mathrm{NOT}(v(\psi)), v(\varphi)))$$

- Interpretations

Function v (extended as above) assigns a truth value <u>*to each*</u> $\varphi \in \mathrm{wff}(L_P)$

$$v : \mathrm{wff}(L_P) \to \{0,1\}$$

Then v is said to be an *interpretation* of L_P

Note that the truth value of any wff φ is univocally determined
by the values assigned to each symbol in the *signature* Σ (compositionality)

An Aside: object *language* and *metalanguage*

- The ***object language*** is L_P

  The formal language of logic

  It only contains the items just defined:

  Σ, $\neg$, $\rightarrow$, $\wedge$, $\vee$, $\leftrightarrow$, (,), plus *syntactic rules* (wff)

- ***Meta-language***

  The formalism for defining the properties of the object language and the logic

  Small greek letters ($\alpha, \beta, \chi, \varphi, \psi$, ...) will be used to denote a generic <u>formula</u> (wff)

  Capital greek letters (Γ, Δ, ...) will be used to denote a <u>set of formulae</u>

  Satisfaction, logical consequence (see after): $\models$

  Derivability (see after): $\vdash$

  “if and only if” : “iff”

  Implication, equivalence (in general): $\Rightarrow$, $\Leftrightarrow$

Entailment

About formulae and their hidden relations

- Hypothesis:

  $\varphi_1 = B \lor D \lor \neg(A \land C)$

  "Sally likes Harry" OR "Harry is happy"
  OR NOT ("Harry is human" AND "Harry is a featherless biped")

  $\varphi_2 = B \lor C$

  "Sally likes Harry" OR "Harry is a featherless biped"

  $\varphi_3 = A \lor D$

  "Harry is human" OR "Harry is happy"

  $\varphi_4 = \neg B$

  NOT "Sally likes Harry"

- Thesis:

  $\psi = D$

  "Harry is happy"

Is there any **logical relation** between hypothesis and thesis?

And among the propositions in the hypothesis?

Entailment

The overall truth table
for the wff in the example

$$\begin{aligned}&\varphi_1 = B \lor D \lor \neg(A \land C)\\&\varphi_2 = B \lor C\\&\varphi_3 = A \lor D\\&\varphi_4 = \neg B\\\hline&\psi = D\end{aligned}$$

Entailment

Notation!

$$\{\varphi_1,\ \varphi_2,\ \varphi_3,\ \varphi_4\} \models \psi$$

There is entailment when
all the *possible worlds* that *satisfy* $\{\varphi_1,\ \varphi_2,\ \varphi_3,\ \varphi_4\}$
satisfy ψ as well

A	B	C	D	φ_1	φ_2	φ_3	φ_4	ψ
0	0	0	0	1	0	0	1	0
0	0	0	1	1	0	1	1	1
0	0	1	0	1	1	0	1	0
0	0	1	1	1	1	1	1	1
0	1	0	0	1	1	0	0	0
0	1	0	1	1	1	1	0	1
0	1	1	0	1	1	0	0	0
0	1	1	1	1	1	1	0	1
1	0	0	0	1	0	1	1	0
1	0	0	1	1	0	1	1	1
1	0	1	0	0	1	1	1	0
1	0	1	1	1	1	1	1	1
1	1	0	0	1	1	1	0	0
1	1	0	1	1	1	1	0	1
1	1	1	0	1	1	1	0	0
1	1	1	1	1	1	1	0	1

A set of wff

One wff

$$\Gamma \models \varphi$$

There is entailment iff every world that satisfies Γ also satisfies φ

Satisfaction, models

- Possible worlds and *truth tables*

  Examples: $\varphi = (A \vee B) \wedge C$

  Different rows, different groups of worlds
  All rows, all possible worlds

  Caution: in each possible world
  every $\varphi \in \mathrm{wff}(L_P)$ has a truth value
  so a row in a table is not a single world, per se

A	B	C	$A \vee B$	$(A \vee B) \wedge C$
0	0	0	0	0
0	0	1	0	0
0	1	0	1	0
0	1	1	1	1
1	0	0	1	0
1	0	1	1	1
1	1	0	1	0
1	1	1	1	1

A possible world **satisfies** a wff φ iff $v(\varphi) = 1$

We also write $<\{0,1\}, \Sigma, v> \models \varphi$

In the truth table above, the rows that satisfy φ are in gray

Such possible world w is also said to be a **model** of φ

By extension, a possible world *satisfies* (i.e. is *model* of) a set of wff $\Gamma = \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$
iff w *satisfies* (i.e. is *model* of) each of its wff $\varphi_1, \varphi_2, \ldots, \varphi_n$

Tautologies, contradictions

- A **tautology**

  Is a (propositional) wff
  that is always satisfied

  It is also said to be **valid**

  Any wff of the type $\varphi \vee \neg\varphi$
  is a tautology

- A **contradiction**

  Is a (propositional) wff,
  that cannot be satisfied

  Any wff of the type $\varphi \wedge \neg\varphi$
  is a contradiction

A	$A \wedge \neg A$	$A \vee \neg A$
0	0	1
1	0	1

A	B	$(\neg A \vee B) \vee (\neg B \vee A)$
0	0	1
0	1	1
1	0	1
1	1	1

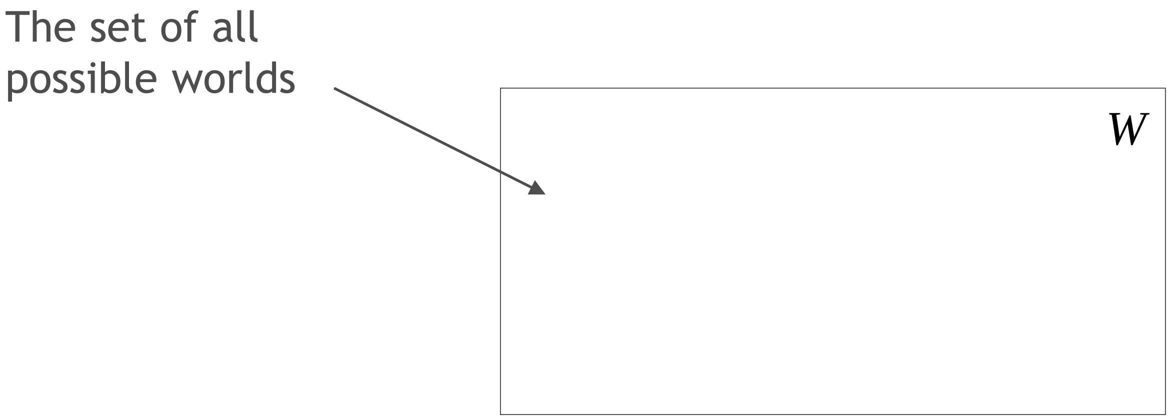
A	B	$\neg((\neg A \vee B) \vee (\neg B \vee A))$
0	0	0
0	1	0
1	0	0
1	1	0

Notes:

- Not all wff are either *tautologies* or contradictions
- If φ is a *tautology* then $\neg\varphi$ is a *contradiction* and vice-versa

Formulae and subsets

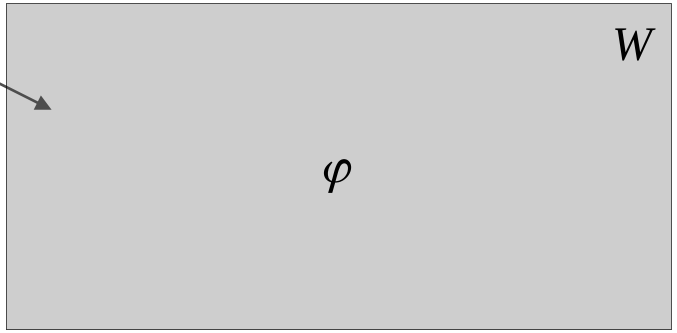
- Consider the set W of all possible worlds
  - Each wff φ of L_P corresponds to a **subset** of W
    - The subset of all possible worlds that *satisfy* it
    - In other words, φ corresponds to $\{w : w \models \varphi\}$
      - The corresponding subset may be empty (i.e. if φ is a contradiction) or it may coincide with W (i.e if φ is a tautology)



Formulae and subsets

- Consider the set W of all possible worlds
  - Each wff φ of L_P corresponds to a **subset** of W
    - The subset of all possible worlds that *satisfy* it
    - In other words, φ corresponds to $\{w : w \models \varphi\}$
      - The corresponding subset may be empty (i.e. if φ is a contradiction) or it may coincide with W (i.e if φ is a tautology)

The set of all possible worlds



"φ is a tautology"

"any possible world in W is a *model* of φ"

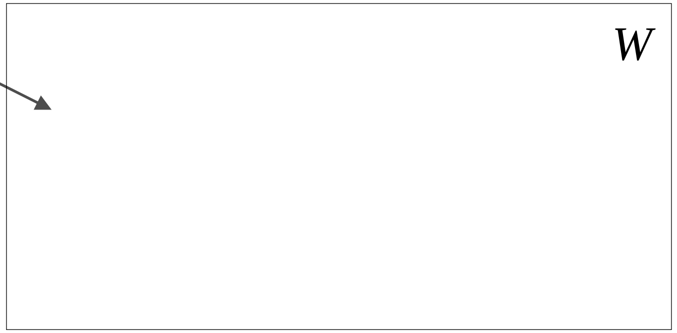
"φ is (logically) ***valid***"

Furthermore:
"φ is ***satisfiable***"
"φ is not ***falsifiable***"

Formulae and subsets

- Consider the set W of all possible worlds
  - Each wff φ of L_P corresponds to a **subset** of W
    - The subset of all possible worlds that *satisfy* it
    - In other words, φ corresponds to $\{w : w \models \varphi\}$
      - The corresponding subset may be empty (i.e. if φ is a contradiction) or it may coincide with W (i.e if φ is a tautology)

The set of all possible worlds



"φ is a contradiction"

"none of the possible worlds in W is a *model* of φ"

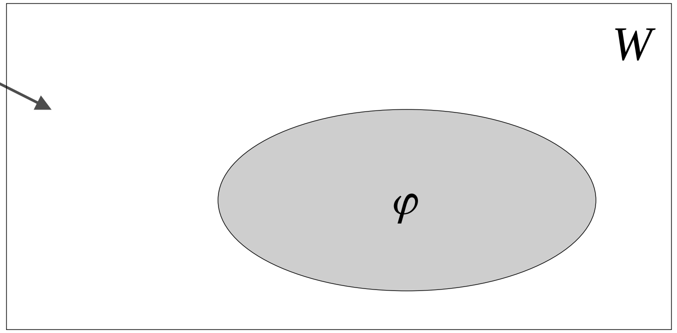
"φ is <u>not</u> (logically) ***valid***"

Furthermore:
"φ is <u>not</u> ***satisfiable***"
"φ is ***falsifiable***"

Formulae and subsets

- Consider the set W of all possible worlds
  - Each wff φ of L_P corresponds to a **subset** of W
    - The subset of all possible worlds that *satisfy* it
    - In other words, φ corresponds to $\{w : w \models \varphi\}$
      - The corresponding subset may be empty (i.e. if φ is a contradiction) or it may coincide with W (i.e if φ is a tautology)

The set of all possible worlds



"φ is neither a contradiction nor a tautology"

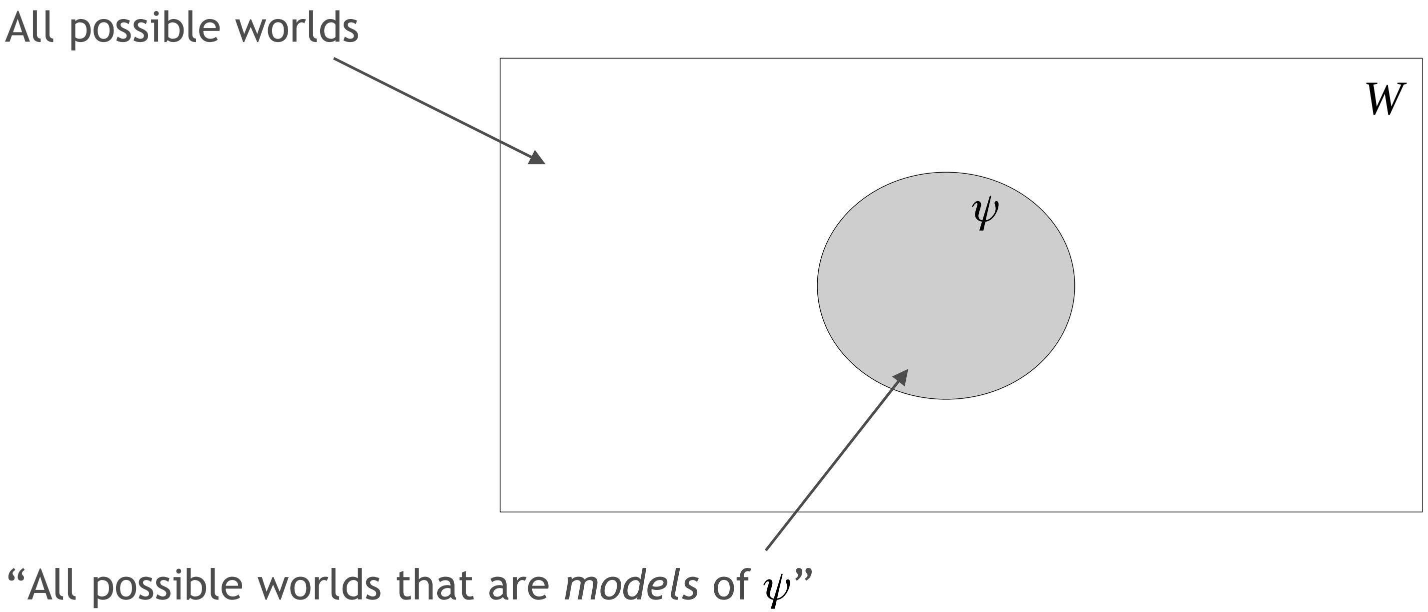
"some possible worlds in W are *model* of φ, others are not"

"φ is not (logically) ***valid***"

Furthermore:
"φ is ***satisfiable***"
"φ is ***falsifiable***"

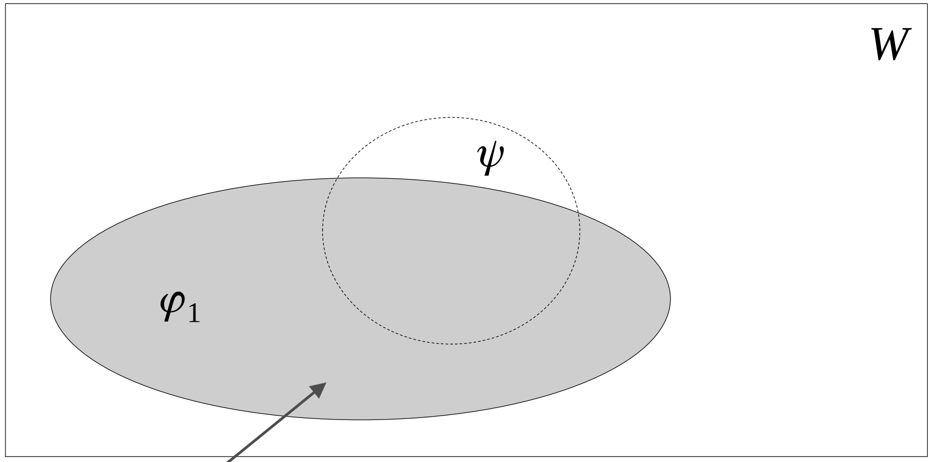
Formulae, subsets and entailment

- Consider the set of all possible worlds W



Formulae, subsets and entailment

- Consider the set of all possible worlds W

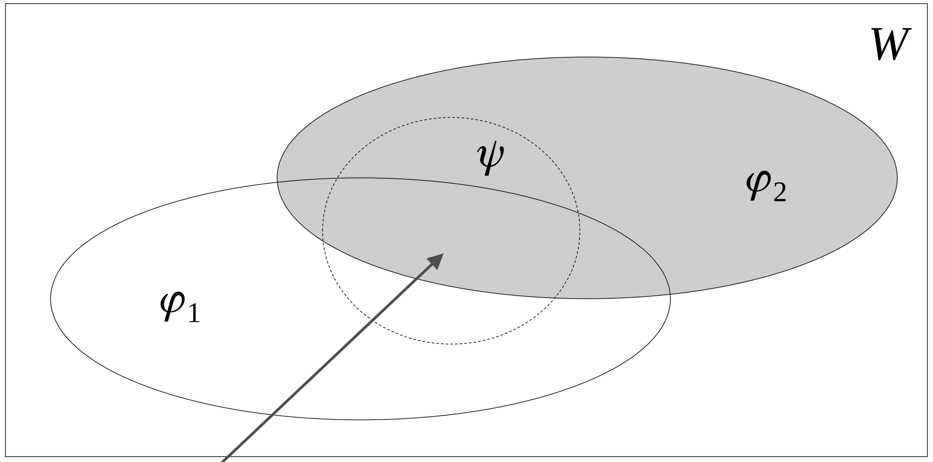


"All possible worlds that are *models* of φ_1"

$\{\varphi_1\} \not\models \psi$
because the set of models for $\{\varphi_1\}$
is <u>not</u> contained in the set of models of ψ

Formulae, subsets and entailment

- Consider the set of all possible worlds W

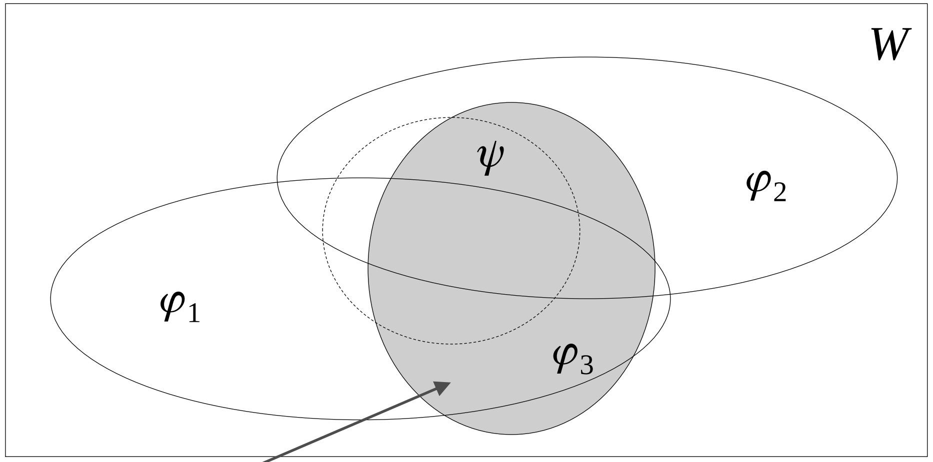


"All possible worlds that are *models* of φ_2"

$\{\varphi_1, \varphi_2\} \not\models \psi$
because the set of models of $\{\varphi_1, \varphi_2\}$ (i.e. the *intersection* of the two subsets) is <u>not</u> contained in the set of models of ψ

Formulae, subsets and entailment

- Consider the set of all possible worlds W

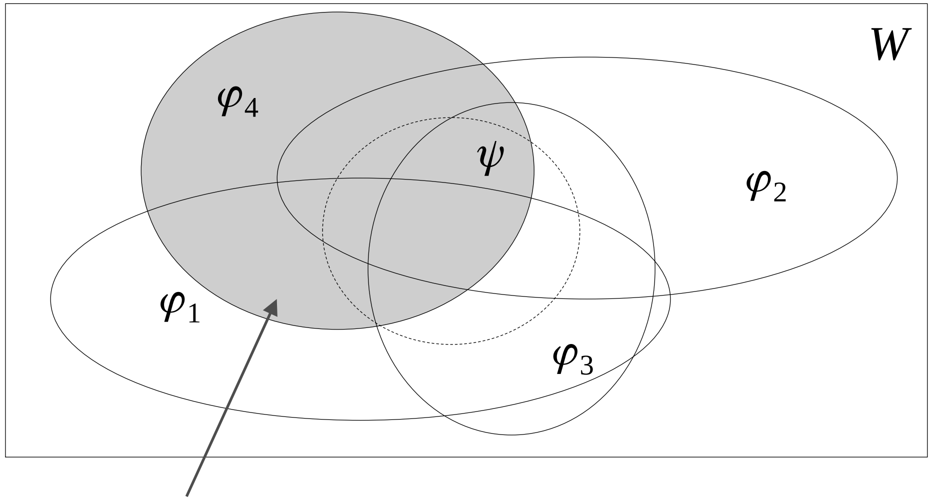


"All possible worlds that are *models* of φ_3"

$\{ \varphi_1, \varphi_2, \varphi_3\} \not\models \psi$
because the set of models of $\{ \varphi_1, \varphi_2, \varphi_3\}$
is <u>not</u> contained in the set of models of ψ

Formulae, subsets and entailment

- Consider the set of all possible worlds W



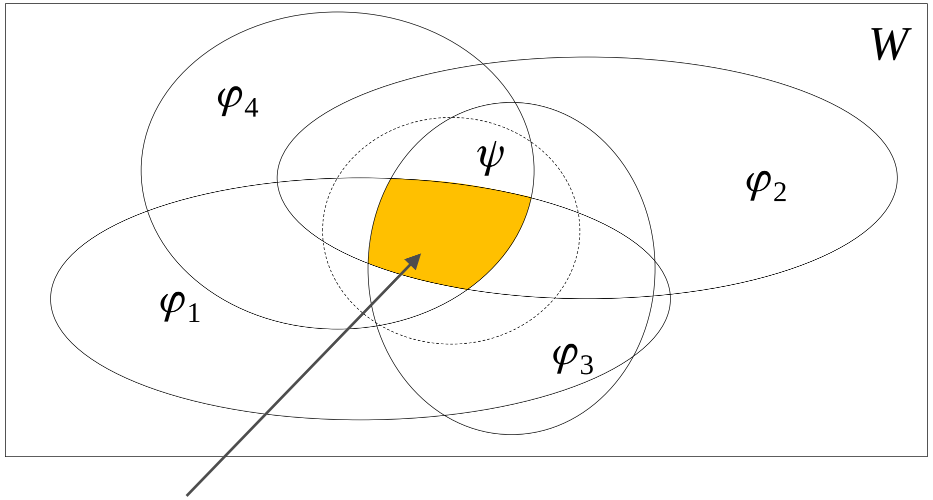
"All possible worlds that are models of φ_4"

$\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\} \models \psi$

Because the set of models for $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$
is contained in the set of models of ψ

Formulae, subsets and entailment

- Consider the set of all possible worlds W



"All possible worlds that are models for $\{ \varphi_1, \varphi_2, \varphi_3, \varphi_4\}$"

$\{ \varphi_1, \varphi_2, \varphi_3, \varphi_4\} \models \psi$

Because the set of models for $\{ \varphi_1, \varphi_2, \varphi_3, \varphi_4\}$
<u>is</u> contained in the set of models of ψ

In the case of the example,
all the wff $\varphi 1, \varphi 2, \varphi 3, \varphi 4$
are needed for the relation
of *entailment* to hold

A set of wff

One wff

$$\Gamma \models \varphi$$

There is entailment iff every world that satisfies Γ also satisfies φ

Further Properties

Symmetric entailment = logical equivalence

- Equivalence

  Let φ and ψ be wff such that:

  $\varphi \models \psi$ e $\psi \models \varphi$

  The two wff are also said to be ***logically equivalent***

  In symbols: $\varphi \equiv \psi$

- Substitutability

  Two equivalent wff have exactly the same ***models***

  In terms of entailment, equivalent wff are substitutable

  (even as sub-formulae)

  In the example: $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\} \models \psi$

$$\varphi_1 = B \lor D \lor \neg(A \land C)$$
$$\varphi_2 = B \lor C$$
$$\varphi_3 = A \lor D$$
$$\varphi_4 = \neg B$$
$$\psi = D$$

$$\varphi_1 = B \lor D \lor (A \to \neg C)$$
$$\varphi_2 = B \lor C$$
$$\varphi_3 = \neg A \to D$$
$$\varphi_4 = \neg B$$
$$\psi = D$$

Implication and Inference Schemas

The wff of the problem can be re-written using equivalent expressions:

(using the basis $\{\rightarrow, \neg\}$)

$\varphi_1 = C \rightarrow (\neg B \rightarrow (A \rightarrow D))$	$\varphi_1 = B \vee D \vee \neg(A \wedge C)$
$\varphi_2 = \neg B \rightarrow C$	$\varphi_2 = B \vee C$
$\varphi_3 = \neg A \rightarrow D$	$\varphi_3 = A \vee D$
$\varphi_4 = \neg B$	$\varphi_4 = \neg B$
$\psi = D$	$\psi = D$

- Some ***inference schemas*** are *valid* in terms of *entailment*:

$$\frac{\begin{array}{l}\varphi \rightarrow \psi \\ \varphi\end{array}}{\psi}$$

It can be verified that:

$$\varphi \rightarrow \psi, \varphi \models \psi$$

Analogously:

$$\varphi \rightarrow \psi, \neg\psi \models \neg\varphi$$

Modern formal logic: fundamentals

- **Formal language (*symbolic*)**

  A set of symbols, not necessarily *finite*

  Syntactic rules for composite formulae (wff)

- **Formal semantics**

  For <u>each</u> formal language, a *class* of structures (i.e. a class of *possible worlds*)

  In each possible world, <u>every</u> wff in the language is assigned a *value*

  In classical propositional logic, the set of values is the simplest: $\{1, 0\}$

- **Satisfaction, entailment**

  A wff is *satisfied* in a possible world if it is <u>true</u> in that possible world

  In classical propositional logic, iff the wff has value 1 in that world
  (Caution: the definition of *satisfaction* will become definitely more complex with *first order logic*)

  Entailment is a <u>relation</u> between a set of wff and a wff

  This relation holds when all possible worlds satisfying the set also satisfy the wff

Properties of entailment (classical logic)

- **Compactness**

  Consider a set of wff Γ (not necessarily *finite*)

  $\Gamma \models \varphi \quad \Rightarrow$ There exist a *finite* subset $\Sigma \subseteq \Gamma$ such that $\Sigma \models \varphi$

  (This follows from *compositionality*, see textbook for a proof)

- **Monotonicity**

  For any Γ and Δ, if $\Gamma \models \varphi$ then $\Gamma \cup \Delta \models \varphi$

  In fact, any entailment relation between φ and Γ remains valid even if Γ grows larger

- **Transitivity**

  If for all $\varphi \in \Sigma$ we have $\Gamma \models \varphi$, then if $\Sigma \models \psi$ then $\Gamma \models \psi$

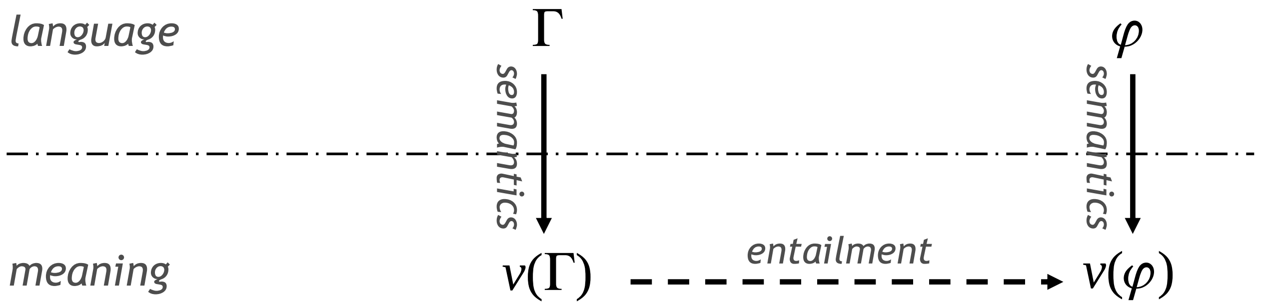
  If Γ entails any φ in Σ, then any ψ entailed by Σ is also entailed by Γ

- ***Ex absurdo …***

  $\{\varphi, \neg\varphi\} \models \psi$

  An inconsistent (i.e. contradictory) set of wff entails *anything*

  «Ex absurdo sequitur quodlibet»


language
Γ
φ
semantics
semantics
meaning
v(Γ)
entailment
v(φ)