World
models, representations, logics and agents
(HP2T)

Index

World models - intuition

Intuition (World models): **Models** depict a specific instance of the world, e.g., the world under observation. **Theories** describe the contents of models. To have a complete representation of the world under observation we need to formalize the **correspondence** between theories and models.

A **World Model** is any formal notation which allows for the representation of the three types of information described above.

Observation (World models): World models do not **(!)** represent the world. They provide the means for representing the world, that is, the means for representing models, theories and how the latter represent what the former depicts.

Formal world Models – intuition

Intuition (Formal world models). **Formal world models**, also called **logical world models**, are composed of three components:

- A **domain** of interpretation which defines the boundaries within which the **intended model** can be formalized;
- A **language** describing the domain of interpretation, which defines the boundaries within which a **theory** of the intended model can be formalized;
- An **interpretation function**, that is, a functional mapping from the language to the domain of interpretation which makes explicit which elements of the language describe which elements of the domain of interpretation.

Observation (Interpretation function): A linguistic representation is always interpreted into an analogical representation. The purpose of the interpretation function is to make this mapping explicit, formally defined and unambiguous.

Types of world models - observation

Observation (Language, informal, semi-formal, Logical) There are three types of **languages** and, correspondingly, three types of **world models:**

- **Informal world models**, namely world models where the grammar of the language is defined informally, typically using Type-0 Chomsky production rules, and, as a consequence, the interpretation function cannot be and is not formally defined. Example: natural languages.
- **Semi-formal models,** namely world models where the grammar of the language is formally defined but the interpretation function is left implicit. Examples: ER models, EER models, DBs.
- **Formal (Logical) world models**, namely world models where the grammar as well as the interpretation of the language are formally defined. Example: the logics studied in this course.

Terminology (Formality of the world model). From now on, when no confusion arises, we leave implicit the type of world model (e.g., we drop the word "formal" in "formal world model").

Interpretation function

Definition (Interpretation function) Let L_a be a language of assertions and D a domain. Then an **Interpretation Function I_a** is defined as

$$I_a : L_a \rightarrow D$$

We say that a fact f ∈ M is **the interpretation of** $a \in I_a$, and write

$$f = I_a(a) = a^I$$

to mean that the assertion a is a linguistic description of f.

We say that f is **the interpretation of** the assertion a, or, equivalently, that a **denotes** f.

Observation (Interpretation function). Interpretation functions apply to assertions and generate facts. Different world models refine the definition of interpretation function based on the percepts they consider. See later, the definition of the interpretation functions of the different world logics. This is a key element distinguishing one logic from another.

Interpretation function (example)

- I_a(Sofia è una persona) = *Sofia ϵ person*
- I_a(Paolo è un uomo) = *Paolo ϵ man*
- I_a(Rocky is a dog) = *Rocky ϵ dog*
- I_a(Sofia is near Paolo) = *<Sofia, Paolo > ϵ near*
- I_a(Rocky è il cane di Sofia) = {*Rocky ϵ dog, <Sofia,Rocky > ϵ Owns*}
- I_a(Sofia è un'amica di Paolo) = *<Sofia, Paolo > ϵ friend*
- I_a(Sofia è bionda) = *Sofia ϵ blond*
- *….*

Interpretation function – Domain of interpretation

Observation (Arbitrariety of the domain of interpretation). If one follows the sequence of previous lectures, one can notice that first we introduced the notions of domain and model, then that of language and theory, and then, as a third and last step, the notion of interpretation function. This is not by chance and it follows how things happen in practice:

- First one chooses the part of the world to be modeled, reified in the domain of interpretation;
- Second, one choose the language (alphabet and formation rules);
- Third and last, it defines the interpretation function to make sure that language means what it is meant to mean.

The domain of interpretation is the input to the overall process. Following the discussion on how things are modeled into entities (see the model theory section), the key observation is that the definition of the **domain of interpretation is arbitrary** and it depends only on the purpose of the modeler, what one wants to represent and reason about. The modeler builds the domain which is just fit to the purpose.

Interpretation function – Domain of interpretation

Observation (Arbitrariety of domains of interpretation, examples). Some examples of domains, which have been defined in the past are:

- Entities and their properties (modeled by the LoE logic, studied later);
- Etypes and their properties;
- Complex combinations of etypes and their properties and relations (modeled by the LoD logic studied later);
- Entities, complex combinations of the etypes of the selected entities, properties and relations (modeled by the LoDe logic, studied later, as a combination of LoE and LoD);
- Judgements about the truth of assertions made in LoE, LoD, LoDe (modeled by the LoP logic studied below);
- … and much more (see a more complete list below).

Observation (Different types of logics). Logics can be characterized based on the domain of interpretation. As detailed later, the logics which denote facts are called **world logics**, since their domain of interpretation is rooted in the world how we perceive it. We call the other logics, **language logics**, or simply **logics**, as they denote facts, that is analogical representations, only indirectly, via linguistic representations describing them.

Interpretation function – non-ambiguity

Observation (Non-ambiguity) Interpretation functions, being functions, are not ambiguous. If two facts f_1 and f_2 are different it cannot be that $I_a(a) = f_1$ and $I_a(a) = f_2$.

Observation 1 (Polysemy) Natural language words have multiple meanings, e.g., Java, Car, bank. The polysemy of words generates the polisemy of assertions. This phenomenon, called *polysemy,* is pervasive. The average polysemy in the lexical resource *WordNet, the world de-facto* standard for lexical resources (i.e., digitized natural language vocabularies) is around 2.

Observation 2 (Polysemy). Polisemy is one of the key problems in natural language processing (NLP), with a lot of research on Word Sense Disambiguation (WSD) algorithms.

Interpretation function - synonymity

Observation (Interpretation function, synonymity). Two assertions are synonyms when they have the same meaning, that is, the interpretation of two different assertions a_1 and a_2, may denote the same fact, i.e., $\mathrm{I}_a(a_1) = \mathrm{I}_a(a_2)$. In logic synonymity is not a problem as the denotation of a word is a single percept.

Observation (Natural language, synonymity). Synonymous words are pervasive in natural languages (e.g., *car* and *automobile*).

Observation (Relational DB, synonymity). In relational DBs synonymity is not allowed, essentially for efficiency reasons (so-called *unique name assumption*). This means different strings always mean different things. That is, words behave like concepts or unique identifiers.

Interpretation function – totality and surjectivity

Observation (Totality). Interpretation functions are total. This guarantees that any element of the language has an interpretation.

Observation (Non-surjectivity). Interpretation functions are not necessarily surjective. In other words, if $I_a : L_a \rightarrow D$, L_a may not be able to name all the facts in D. This property formalizes the fact that linguistic representations do not necessarily describe all facts.

Language correctness and completeness

Definition (Language correctness and completeness). Let L $_a$ and D be an assertional language and a domain of interpretation, respectively. Let I$_a$: L$_a$ → D be an interpretation function. Then we have two possible situations, as follows

Language correctness. Let a ∈ L$_a$ be an assertion. If for all a, if a ∈ L$_a$ then I$_a$(a) =f ∈ D, then we say that L$_a$ is **correct** with respect to D, or that D **is a domain** for L$_a$;

Language completeness. Let f ∈ D be a fact. If, for all f, if f ∈ D then there is an assertion a ∈ L$_a$ such that I$_a$(a) = f, then we say that L$_a$ is **complete** with respect to D.

The notions of **incorrectness** and **incompleteness** are defined symmetrically.

Language correctness and completeness (continued)

Observation (Correctness of an assertional language L_a with respect to a domain D). An assertional language, to be used for a given domain, **must be correct**, that is to contain *only* assertions which denote facts in the reference domain. If this not the case, then we say that D is NOT a domain of L_a or, vice versa, that L_a is not a language for D. This in order to avoid nonsensical assertions (e.g., the assertion "gdhaosdf").

Observation 1 (Completeness of an assertional language L_a with respect to a domain D). An assertional language is not necessarily complete, that is, it does not necessarily contain assertions for all the facts in a domain (which, among other things, are in principle infinite). The key feature is that it should contain all the assertions deemed relevant.

Observation 2 (Completeness of an assertional language L_a with respect to a domain D). There are two main technical reasons for language incompleteness. The first is that the language does not have as many elements as the domain. The second is that, even if the language is large enough, the interpretation function is not surjective because of synonyms.

World Model

Definition (World model). Given a **Domain of interpretation** D, a **world model** W is defined as

$$W = \langle L_a, D, I_a \rangle$$

where L_a is an **assertional language**, $I_a : L_a \rightarrow D$ is an **interpretation function** and L_a is **correct** with respect to D.

Observation (World model). L_a is not necessarily complete.

World model*

$$\begin{tikzcd}
a & T_A \arrow[l, "\in"'] \arrow[r, "\subseteq"] \arrow[d, dashed, "\mathcal{I}_A"] & \mathcal{L}_A \arrow[d, "\mathcal{I}_A"] \\
f & \mathbb{M} \arrow[l, "\in"'] \arrow[r, "\subseteq"] & \mathbb{D}
\arrow[from=1-1, to=2-1, "\mathcal{I}_A"]
\end{tikzcd}$$

***Errata corridge:** «L_A» should be «L_a», «T_A» should be «T_a», «I_A» should be «I_a»*

Terminology – Syntax and Semantics

Terminology (Syntax and semantics). When talking about world models, people informally talk of **syntax** meaning the language of the world model, and of **semantics** meaning the domain of interpretation, associated to the syntax, via the **interpretation function,** informally or formally defined.

Observation (Syntax and semantics). Without a formal understanding of the intended semantics of a given syntax, that is, without the interpretation function and a formally defined world model, it is impossible to univocally assert whether a certain assertion (sentence) is true or false.

World model diversity

Observation (World diversity). The world is constantly different from itself when observed at different spatio-temporal coordinates. It changes and it evolves.

Observation (World model diversity). Assume we are describing the same world, as it appears in space and time at certain spatio-temporal coordinates. Two world models describing it may differ in the following three dimensions:

- **Domain diversity**. Two domains of the same world may differ in the percepts (types and specific instances) and facts they consider as possible. The choices of different domains is the source of the **diversity of the world**, how it is perceived by humans;
- **Language diversity**. Two languages of the same domain may differ in the alphabet and formation rules. The choice of different languages is the source of the diversity of the **descriptions of the world**, how they are provided by humans;
- **Interpretation diversity**. Two interpretation functions may differ in how alphabet elements and assertions map to percepts and vice versa. This is in fact a many-to-many mapping.

Index

- World models
- **World representations**
- World logics
- Logics
- Agents
- How to use logics
- Key Notions

World representations - intuition

Intuition (World representations): Based on the representational choices made by world models, world representations represent what is the case in the (part of the) world under consideration. They have three main components, that is:

- The **intended model**
- A **theory** describing the intended model
- A **correct by construction mapping**, enforced by the interpretation function which guarantees that the theory actually describes the intended model.

A **World representation** is any representation which encodes the three types of information described above.

Theory correctness and completeness

Definition (Theory correctness and completeness). Let W = ⟨L_a, D, I_a⟩ be a world model. Let $T_a \subseteq L_a$ and M ⊆ D be a set of assertions and a set of fact, respectively. Then we have the following.

Correctness. Let $a \in L_a$ be an assertion. If for all a, if $a \in T_a$ then $I_a(a) \in$ M, then we say that T_a is **correct** with respect to M;

Completeness. Let f ∈ M be a fact. If, for all f, if f ∈ M then there is an assertion $a \in T_a$ such that $I_a(a)$ = f, then we say that T_a is **complete** with respect to M. The completeness condition can also be written as:

If, for all $a \in L_a$, $I_a(a) \in M$ then $a \in T_a$.

The notions of **incorrectness** and **incompleteness** are defined in the obvious way.

Theory correctness and completeness (continued)

Observation (Correctness of an assertional theory T$_a$ with respect to a model M). An assertional theory T$_a$,to be used to describe an intended model, **must be correct** contain *only* assertions about facts in M. If this not the case then we say that M is not a model of T$_a$ or, vice versa, that T$_a$ is not a theory for M. This in order to avoid false assertions.

Observation (Completeness of an assertional theory T$_a$ with respect to a model M). An assertional theory **may be incomplete**, namely there can be facts of the model for which T$_a$ does contain assertions. Incomplete assertional theories are the default.

Observation (Correctness and completeness). The requirement on theories is the same as that on languages and domains.

Maximal theories, canonical model

Definition (Maximal theory). A theory is **maximal** with respect to M if it is correct and complete**.**

Observation (Intended model, maximal theories). The intended model has been informally defined as the model one has in mind when writing the theory. But, because of the missing requirement on completeness, the same theories may have multiple intended models.

Definition (Canonical model). Given a theory, the **canonical model** for that theory is the model for which T_a is maximal.

Observation 1 (Maximal theory). Given a world model there may be mutiple maximal theories, that is theories which contain one or more assertions for each fact in the intended model. This is because of synonyms.

Observation 2 (Maximal theory). Under the unique name assumption, there is one and only one maximal theory.

World models, theories and models

Definition (Theory, model). Given a world model

$$W = \langle L_a, D, I_a \rangle$$

then, given M and T_a defined as follows,

$$M = \{f\} \subseteq D$$
$$T_a = \{a\} \subseteq L_a$$

M and T_a are, respectively, a **model** of T_a and a **theory** of M, if **T_a is correct for M.**

Observation (Theory). Theories are not necessarily complete.

World models, theories and models

Observation (Model of, Theory of). The notions of a theory being a theory of a model, and vice versa, of a model being a model of a theory are meaningful only when the reference world model is known. As an example, the assertions od an ER model can be understood only when if knows the notation of ER models.

Observation (Theory of a model). For T_a to be a theory of a model M, only correctness is requested. This is coherent with the fact that, anyhow, it is impossible to have a complete linguistic description of an analogical representation of the world (see discussion above).

Observation (Many-to-many relation between theories and models). The absence of completeness is such that there is a many-to-many relation between models and theories. In fact:

- Given a model M of a theory T_a, any set of assertions $T \subseteq T_a$ is a theory of M.
- Given a theory T_a of a model M , any set of the facts $M_{a,}$ with $M \subseteq M_a$,is a model of T_a.

Observation (Maximal theories). Given a model there can be multiple maximal theories (see above). There is a one-to-many relation between models and maximal theories.

Observation (Intended model, maximal theories). Given the many-to-many relation between models and theories, the intended model, that is the model that one has in mind when writing the theory, is only one among the main possible models. Of course one can assume, as it is usually the case, that the intended model is the model for which T_a is maximal.

World representations

Definition (World representation). Given a world model

$$W = \langle L_a, D, I_a \rangle$$

then

$$R = \langle T_a, M \rangle$$

is a **world representation**, with

$$M = \{f\} \subseteq D$$
$$T_a = \{a\} \subseteq L_a$$

where M and T_a are, respectively, a **model** of T_a and a **theory** of M in W.

Definition (Canonical world representation). A world representation is **canonical** when M is the canonical model of T_a (that is, T_a is **maximal** for M).

World models, representations, and beyond

Observations (World models). World models provide all the formal notions which are necessary to unambiguously describe what is perceived (analogic and linguistic representations). They specify the extent to which what will be perceived will be described. Once defined, world models, may be assumed not to change or at least to change very slowly, when something never perceived before, is perceived.

Observation (World representations). World representations describe the contents of the analogic representations as they are perceived, and their descriptions. World representations may change in time.

Observation (Changes in a world representation, reasoning, entailment relation). World representations may change in time as a consequence of reasoning. Reasoning allows to derive new facts from the facts that are already known. This is formalized by (**language) logics.** Logical reasoning makes explicit, allows to assert knowledge which is already encoded, even if only implicitly, in a world representation. Reasoning is formalized via the **entailment relation** (see later).

Observation (Changes in a world representation, perception). World representations change in time as a consequence of perception. This is formalized by **agents**. Agents are logics which can incrementally acquire new facts about the world. Agents are logics extended with an **Tell** / **Ask** operation (see later). **Tell** is the basic operation which allows to extend a representation (the model or theory, and the other consequently). **Ask** implements entailment, suitably extended to take into account the new possibilities offered by the used of Tell.

World representation diversity

Observation (World representation diversity). The notion of world representation allows to make precise the extent to which two analogical and linguistic representations may differ. Assume that two world representations share the same world model and that they describe the same (part of) the world:

- **Model diversity**. Within the bounds of the selected domain of interpretation, we may have **diversity of the facts and percepts** used to describe what is the case in the world, how it is perceived by humans;
- **Theory diversity.** Within the bounds of the selected assertional language, we may have **diversity in the alphabet** used to name percepts and **in the formation rules** used to form assertions describing the facts of the intended model. As a result we have **diversity of the assertions** used to describe the world, how they are provided by humans;
- **Interpretation diversity.** Because of the many-to-many mapping between theories and models, we may have **diversity of the theories** of the same model **and diversity of the models** of the same theory.

Index

- World models
- World representations
- **World logics**
- Logics
- Agents
- How to use logics
- Key Notions

World logics - intuition

Intuition (World logics): Based on the representational choices made by world models and world representations, world logics allow for the most basic form of reasoning about world representations. They have three main components, that is:

- The intended **model**
- A **set of assertions** defining an input theory describing the intended model
- A **world entailment relation** which allows to decide whether the input assertional theory is actually a theory of the intended model.

A **world logic** is any representation which encodes the three types of information described above.

Observation (World entailment, reasoning). Once, based on a selected world model, one has constructed a world representation, the next issue is to use it to **reason** about it, that is to reach conclusions about it. Building upon the notion of **correctness** provided by the interpretation function, world entailment provides a mechanism, an algorithm, for deciding whether an input set of assertions is actually a theory if the intended model.

Observation (World entailment, entailment). As discussed later, there are complex forms of entailment, modeling, complex, **language driven** forms of reasoning. All of them, however, ultimately use world entailment as the key operation, as from its name, for deciding whether something is actually true in the world. Language driven reasoning can elaborate upon but cannot substitute checking truth in the world (that is, in the intended model).

World entailment

Definition (World entailment) Let W =< L_a, D, I_a > be a world model. Let L_a be an assertional language. Let M={f} ⊆ D be the intended model. Let T_a ⊆ L_a be an assertional theory. Then $|=_{La}$ is an **world entailment relation** that **associates facts in M with assertions in T_a**, in formulas

$$|=_{La} \subseteq M \times T_a$$

We also write

$$M \; |=_{La} T_a$$

and say that M (**world**) **entails** T_a. We write M |= T_a instead of M $|=_{La}$ T_a when no confusion arises.

World Entailment

Proposition (World entailment). Let W =< L_a, D, I_a > be a world model. Let M ⊆ D. Let $T_a \subseteq L_a$ ($a \in L_a$) be an input assertional theory (assertion). Then

$$M \models T_a \text{ if, for all } a \in T_a,\ a \text{ is } \textbf{True} \text{ in } M$$

$$M \models \{a\}, \text{ written } M \models a,\ a \text{ is } \textbf{True} \text{ in } M$$

Observation (World entailment). T_a is entailed by M if all its assertions are true in M. World model entailment reduces entailment to checking, via the interpretation function, for truth / falsity in the model.

Truth and Falsity

Definition (True and False assertion). $\langle L_a, D, I_a \rangle$ be a World model. Let $M = \{f\} \subseteq D$ be a model. Let $a \in L_a$ be an input to the partiality, that is, missing knowledge of the intended model. assertion. Then we say that

$a \in L_a$ is **True** in M if the fact $f = I_a(a) \in M$, **False** otherwise

Observation 1 (Truth / Falsity). The notions of Truth and Falsity of an assertion are meaningful only if made with respect to an intended model previously built.

Observation 2 (Truth / Falsity Assessment = Question Answering about truth). The statement that an input assertion is true or false amounts to answering the question of whether this assertion is a member of the intended model, that is, whether the input assertion is known to be true.

Observation 3 (NOT knowing vs. knowing NOT). A positive answer to the above question means that the input assertion is known to be true. A negative answer **does NOT mean** that this assertion is **known to be false**. It means that it is an assertion whose interpretation is a fact in the mode. **It means** that the input assertion is **NOT known to be true**.

World Entailment vs interpretation

Observation (world entailment and interpretation). A comparison:

- Interpretation is a function between a language and a domain: given an element of the language it returns the corresponding element of the model;
- World entailment is a relation between a model and a theory: given a set of assertions it assesses whether these assertions are actually a theory of the intended model.
- While both relating syntax and semantics, world entailment and interpretation function somehow work in opposite directions: the interpretation function constructs the model, the entailment relation checks it for the truth of a set of input assertions.

Observation (World entailment relation). World entailment is a many-to-many relation. This is because of the fact that theories are partial descriptions of their intended model. Plus, as discussed above, a model may have multiple maximal theories.

World logics and world logic representations

Definition (World logic). Given a **world model** $W = \langle L_a, D, I_a \rangle$, a **world logic** L_W for W is defined as

$$L_W = \langle W, |=_{La} \rangle$$

where $|=_{La}$ is a **world entailment relation**.

Definition (World logic representation). Given a world logic $L_W = \langle W, |=_{La} \rangle$, a (**world logic**) **representation** is defined as

$$R = \langle T_a, M \rangle$$

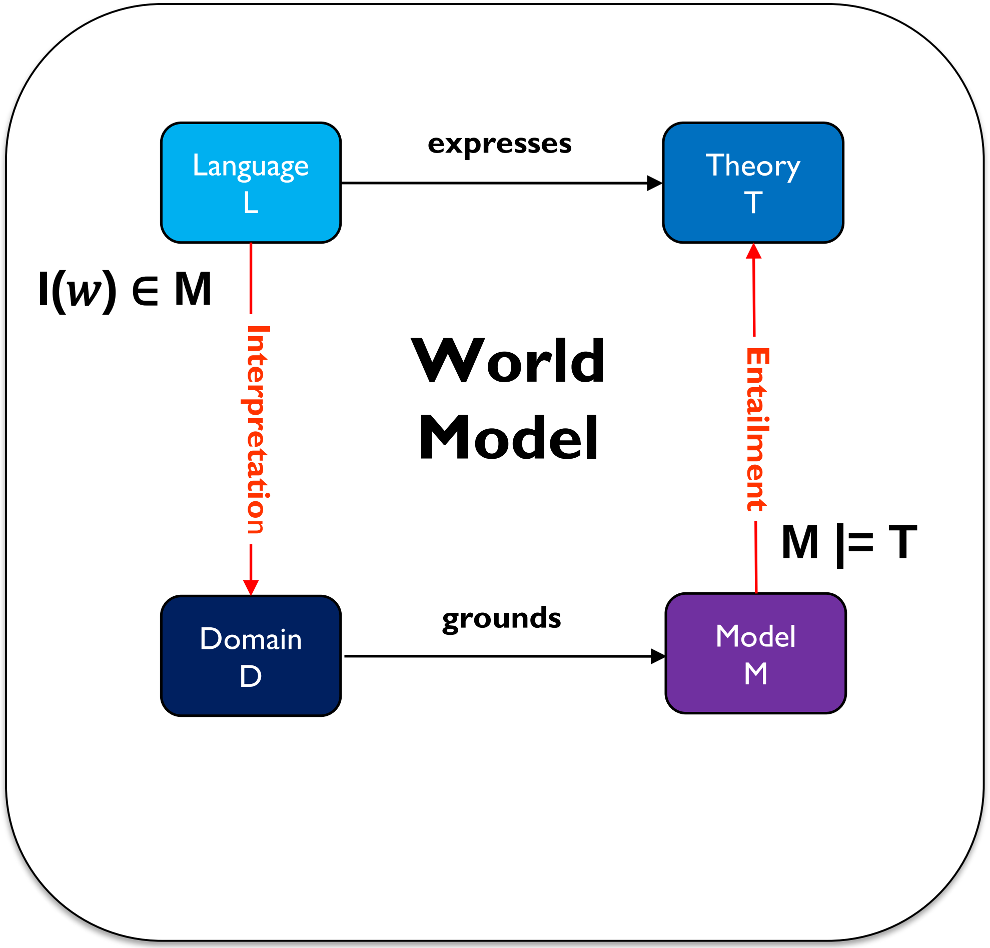
with

$$M = \{f\} \subseteq D$$
$$T_a = \{a\} \subseteq L_a$$

where M and T_a are, respectively, a **model** of T_a and a **theory** of M in L_W.

Observation (World logic representation). A world logic representation is the same as that of its world model.

World model logic, entailment and interpretation



Inside a world logic, world model entailment and interpretation «connect» in the following way:

1. The interpretation function is exploited to **tell** a world logic the intended model via the assertion of a theory in L;
2. World model entailment is exploited to **ask** a world logic whether a theory is true in the intended model, as constructed, via the tell operation, by the interpretation function.

World models and world logics

$$\begin{tikzcd} a & T_A \arrow[l, "\in"'] \arrow[r, "\subseteq"] \arrow[d, dashed, "\mathcal{I}_A"] & \mathcal{L}_A \arrow[d, "\mathcal{I}_A"] \\ f & M \arrow[l, "\in"'] \arrow[r, "\subseteq"] & D \end{tikzcd}$$

World models*

$$\begin{tikzcd} a \arrow[d, "\mathcal{I}"] & T \arrow[l, "\in"'] \arrow[r, "\subseteq"] \arrow[d, dashed, "\models"] & \mathcal{L}_a \arrow[d, "\mathcal{I}"] \\ f & M \arrow[l, "\in"'] \arrow[r, "\subseteq"] & D \end{tikzcd}$$

World logics*

World models allow to build world representations.
World logics allow to query them.

***Errata corridge:** *«A» should be «a», «a» in «L_a» in Logics should be dropped. In world logics the arrow from T to M goes in the wrong direction*

World logics diversity

Observation (World logics diversity). We have the following

- **World model diversity**. Two world logics which are based on different world models are a priori not comparable. The only way to compare them is based on the translation mechanisms, introduced before, across world models
- **World entailment diversity.** The diversity of entailment of two logics based on the same world logic is in terms of the complex assertion used to define te entailment relation.

Family of world logics

Definition (Family of world logics). A family of world logics is a set of world logics which share possibly partially, the same world model.

Observation (World logics Family). The key idea behind a family of world logics is as follows:

- Select a so called **reference world model** $W = \langle L_a, D, I_a \rangle$.
- Define the so-called **base world logic**, that is the world logic $L_W = \langle W, \models_{La} \rangle$ implementing world entailment on assertions as truth in the reference world model
- Define a more complex world logic in a family from a simpler world logic as follows:
  - by extending the expressiveness of the language, to allow for new assertions, beyond assertions which just name the percepts and facts occurring in D. We call them **atomic** and **complex assertions** (**percepts**, **facts**)
  - by extending the interpretation function, from atomic assertions (percepts, facts) to complex assertions (percepts, facts)
  - by extending accordingly the entailment relation, beyond world entailment to allow for assertional theories involing atomic as well as complex assertions.

More expressive world logics in the same family allow for more refined reasoning.

Addendum* – example world logics

Below a few exemplary logics which show the power of language and reasoning via entailment:

- **LoO:** the Logic(s) of occurrences. It formalizes how language can be extended to allow for commonsense knowledge about **occurrences** of entities (things) can be used to model entities as composed of their occurrences.
- **LoE****: the Logic of Entities. It is a based world logic. It formalizes (only) the interpretation function based entailment of **world models**;
- **LoD****: the Logic of descriptions. It formalizes how the LoE language can be extended to allow for **commonsense knowledge** definitions and descriptions, and reasoning about them;
- **LoDe****: The Logic of Entity Bases. It exploits LoD, that is, commonsense knowledge and reasoning, to enhance world model entailment, as modeled in LoE;
- **LoT:** Logic(s) of Time. It exploits the modeling of **time** and (finite) **state machines** to study the execution properties of HW and SW systems. Used extensively in formal methods;

****Addendum slides are added to provide general background information. They are not part of the course material and, therefore, not a topic of the exam.***

*** These logics are the ones covered in this course.**

Index

- World models
- World representations
- World logics
- **Logics**
- Agents
- How to use logics
- Key Notions

Logics – Intuition

Observation (World logics). World logics formalize how truth in a model can be reasoned in a (logical) theory (that is, a linguistic representation of the world). They are the key element, via the entailment relation, for the formalization of (**logical) reasoning.**

Observation (Language logics). Logical reasoning is linguistic reasoning, that is, reasoning in a predefined language. Logical reasoning is implemented using **(language) logics** which allow to draw conclusions from the true facts computed by world logics. They use world logics as oracles which provide information about what is true/ false in the intended model.

Note: The first part of this course focuses on the world logics of KGs. The second part will focus on language logics exploiting the world logics of KGs.

Terminology – formulas

Terminology (Formula, well-formed formula, wff). All logics rely on a language, defined in terms of a type 2 Chomsky grammar, composed of an alphabet and a set of formation rules, an interpretation function and a domain of interpretation.

Languages can be distinguished in terms of the objects which constitute their domain of interpretation. We call **formula**, or **well-formed formula**, or **wff** any element of given language which is correctly formed, starting from the alphabet and using the formation rules. Notationally, we write

L = {*w*} to mean that the language L is a set of wffs *w,*

Terminology (Assertion). Assertions, distinct in atomic and complex assertions, are formulas which describe facts of composition of facts as they occur in the intended model. Notationally, we write

L = {*a*} to mean that the language L is a set of assertions *a*

Terminology (Proposition). Propositions, distinct in atomic and complex propositions, are formulas which describe what is true in the intended model. Notationally, we write

L = {*p*} to mean that the language L is a set of propositions *p*

Observation (formula). The distinction among the different types of formulas (e.g., assertion, vs. proposition) is based on what they denote, and it is independent of the specific alphabet and formation rules.

Entailment

Notion 6 (Entailment) Let W =< L_a, D, I_a > be a world model. L_W = ⟨W, $|=_{La}$⟩ be a world logic for W. Let L = {w} be a language, with $L_a \subseteq$ L. Let M={f} ⊆ D be a set of facts. Let T ⊆ L be a theory. Then $|=_L$ is an **entailment relation** that associates the facts in M with the elements in T, in formulas

$$|=_L \subseteq M \times T, \text{ also written } M \ |=_L T,$$

subject to the constraint that for all assertions $a \in L_a$,

$$M \ |=_L a \text{ if and only if } M \ |=_{La} a$$

We also say that M **entails** T and write M |= T when no confusion arises.

Observation ($L_a \subseteq$ L). Sometimes the assertions $a \in L_a$ get rewritten, to assertions a' ∈ L, under the guarantee of a one-to-one mapping between the two notations.

Entailment – observations

Observation 1 (Entailment, reasoning). The definition of entailment is made based on a theory T ⊆ L, where there exists a suitable L_a, with L_a ⊆ L. The key intuition is that of extending a reference assertional language to allow for formulas which are not necessarily assertions and, then, to ask about the truth of these formulas.

Observation 2 (Entailment, reasoning). Entailment formalizes the intuitive notion of reasoning. It links what one asserts as being the case with what is true in the model. There are multiple notions of entailment, formalizing different notions of reasoning, even for the same world model, with wildly different properties.

Logics and logic representations

Definition (Logic). Given a **world model** $W = \langle L_a, D, I_a \rangle$ and a **world logic** $L_W = \langle W, |=_{La} \rangle$, a **logic** L_L for L_W is defined as

$$L_L = \langle W, |=_L \rangle$$

where $L_a \subseteq L$, and $|=_L$ is an **entailment relation**.

Definition (Logic representation). Given a logic $L_L = \langle W, |=_L \rangle$, a (**logic representation** is defined as)

$$R = \langle T, M \rangle$$

with

$$M \subseteq D$$
$$T = \{w\} \subseteq L$$

where M and T are, respectively, a **model** of T and a **theory** of M in L_L.

Logics diversity

Observation (Logics diversity). We have the following

- **World logic diversity**. Two logics which are based on different world logics are a priori not comparable. The only way to compare them is based on the translation mechanisms, introduced before, across world models
- **Entailment diversity.** The diversity of entailment of two logics based on the same world logic is in terms of their strength in terms of reasoning capabilities, that is, in terms of the conclusions that they are able to derive.
- **Domain diversity.** Two logics based on the same world logics and entailment relation may be defined using different domains of interpretations, which may differ for the set of admissible models $M \subseteq D$. This happens when a logic eliminates certain models (i.e., world configurations) a priori, based on some prior assumptions / knowledge.

Family of logics

Definition (Family of logics). A family of logics is a set of world logics which share possibly partially, the same world logic.

Observation (Logics Family). The key idea behind a family of logics is as follows:

- Select a so-called **reference world logic** $L_W = \langle W, |=_{La} \rangle$
- Define the so-called **base logic**, to be the same as the world logic $L_W = \langle W, |=_{La} \rangle$ implementing world entailment on assertions,
- Define a more complex world logic in a family from a simpler world logic as follows:
  - by extending the expressiveness of the language, to allow for new assertions, beyond assertions which just name the percepts and facts occurring in D. We call them **atomic** and **complex formulas.** Atomic formulas are, possible complex, assertions of the world logic (denoting the refence world model percepts, facts)
  - by extending the interpretation function, from atomic assertions (percepts, facts) to complex assertions (percepts, facts)
  - by extending accordingly the entailment relation, beyond world entailment to allow for assertional theories involing atomic as well as complex assertions.

More expressive world logics in the same family allow for more refined reasoning.

Addendum* – example logics

Below a few exemplary logics which show the power of language and reasoning via entailment:

- **LoP****: The Logic of Propositions. It allows to reason about and to draw consequence from **propositions**, that is, from judgements about what is true and what is false;
- **LoI**: Logic of Interaction, also known as the First order Logic. This logic allows for the use of variables, **existential** and **universal quantification**. It approximates the expressivity of the language used in natural language interactions (towards LLMs);
- **LoDy**: Logic(s) of Dynamics. It allows to represent and reasoning about **programs** and **plans** based on their activation preconditions and the effects of their applicability. Used, using semi-formal notations, in planning;
- **LoR**: Logic(s) of Relevance, used to reason about **theory** (**re**)**formulation**, **adaptation** and **evolution**, and **reasoning;**
- **LoM**: Logic(s) of Modality, used to reason about human **propositional attitudes** (e.g., **knowledge** and **belief**) and multiagent systems.

**** Addendum slides are added to provide general background information. They are not part of the course material and, therefore, not a topic of the exam.***

***** These logics are the ones covered in this course.***

Addendum* – example logics

Observation (Types of logics): The different logics can be characterized based on their domain of interpretation and the type of reasoning they allow. We have the following:

- **World logics of occurrences** (LoO and derived logics). The domain of of interpretation is built out of occurrence percepts and occurrence facts. Entailment allows to ascertain whether two occurrences (so called perdurants) are occurrences of the same entity.
- **World logics of endurants** (LoE, LoD, LoDe). The domain of interpretation is built out of entity percepts and entity facts. Entailment allows to ascertain whether two entities (so called endurants) are the same entity.
- **World logics of pedurants** (LoT, LoDy). The domain of interpretation is events (perdurants composed of endurants and perdurants. Entailment allows to reason about cause / effect.
- **Language logics** (also called **Logics of propositions**) (LoP, LoI). The domain of interpretation is propositions, where atomic propositions are judgement over assertions. Entailment allows to derive true propositions from true propositions.
- **Context logics** (LoR, LoM). The domain of interpretation is theories, where a theory can be any of theories developed in the four logics above. Entailment allows to reason about which theories are best inside which use one of the four logics above.

*** Addendum slides are added to provide general background information. They are not part of the course material and, therefore, not a topic of the exam.***

Diversity of logics – Expressivity vs. Efficiency

Observation (Logic, selection trade-offs) Any logic, in particular beyond world logics, can be characterized by two main parameters:

- **Expressivity of the language** (beyond assertions), that is, the level of richness at which the problem is expressed, depending on the syntax of the language (for instance relating to partial or negative knowledge, see later);
- **Computational efficiency**, that is how much it costs, in terms of space and time, to reason and answer satisfiability queries in that language.

Observation (World logic, selection trade-offs). The same considerations on trade-offs apply to world logics. In both types of logics the expressivity can be arbitrarily changed.

Diversity of logics – Expressivity vs. Efficiency (cont.)

- More expressivity allows for a more refined and precise modeling of the world but it also generates more complex formulas.
- The modeler must find the right trade-off between **expressiveness** and **computational complexity**.
- Here the choice of the representation language is crucial. The computational complexity of reasoning ranges in fact from polynomial to exponential and beyond.
- There is also an issue of **(un)decidability**, namely the possibility for the reasoner, on certain queries, to get into an infinite loop, never terminate and, therefore, never return an answer.

Index

- World models
- World representations
- World logics
- Logics
- **Agents**
- How to use logics
- Key Notions

Agents - intuition

Observation (Logics). Logical reasoning is implemented using (language) logics, as defined before, which allows to draw conclusions, via entailment, from the facts computed by world logics.

Observation (Agents). Logics are used with a fixed, predefined world representation which is usually assumed to be a given**.** Logics formalize the reasoning mechanisms by which, via entailment, one can derive new conclusions from what is already known. Agents extend logics by allowing new language to be learned or new knowledge to be asserted by being told or by observation.

Observation (From Logics to Agents – Tell / Ask). Agents are logics equipped with two new functionalities:

- **Tell**, which allows to extend a world representation;
- **Ask**, which allows to uses extended forms of entailment, as enforced by the assertions made via Tell operations.

World models, representations, and beyond (reprise)

Observations (World models). World models allow to unambiguously describe what is perceived (analogic and linguistic representations). They specify the extent to which what will be perceived can be described. Once defined, world models, may be assumed not to change or at least to change very slowly, when something never perceived before, is perceived.

Observation (World representations). World representations describe the contents of the analogic representations as they are perceived, and their descriptions. World representations may change in time.

Observation (Changes in a world representation, reasoning, entailment relation). World representations may change in time as a consequence of reasoning. Reasoning allows to derive new facts from the facts that are already known. This is formalized by (**language) logics.** Logical reasoning makes explicit, allows to assert knowledge which is already encoded, even if only implicitly, in a world representation. Reasoning is formalized via the **entailment relation** (see later).

Observation (Changes in a world representation, perception). World representations change in time as a consequence of perception. This is formalized by **agents**. Agents are logics which can incrementally acquire new facts about the world. Agents are logics extended with an **Tell** / **Ask** operation (see later). **Tell** is the basic operation which allows to extend a representation (the model or theory, and the other consequently). **Ask** implements entailment, suitably extended to take into account the new possibilities offered by Tell.

Agents and agent representations

Definition (Agent). Let W =< L_a, D, I_a > be a world model. Let $L_W = \langle W, |=_{La} \rangle$ be a world logic for W. Let $L_L = \langle W, |=_L \rangle$ a logic for L_W. Then an **agent** A_L for the logic L_L is defined as

$$A_L = \langle L_L, Tell, Ask \rangle$$

with

$$R = \langle T, M \rangle,\ M \subseteq D \quad T = \{w\} \subseteq L$$

with being the agent's Logic **representation** of the world, where:

- Tell allows
  - to extend L_a the Language L of A_L of a new alphabet element, and I_a if/ as needed.
  - to extend T with a new formula
- Ask allows to query the agent representation of the world, via (advanced versions of) entailment.

Definition(Logic agent). A **Logic agent**, also called a **Logic**, is an agent where cannot be two tell operations between one or more Ask operations.

Observation (Logic agent). The intuition is that with a logic agent, we do not allow therefore the evolution of the world representation of an agent without an explicit approval by some outside oracle (e.g., a human user). This is the assumption which is implicitly made in all the practical uses of logic.

Terminology (Logic, logic agent). From now of, unless specified overwise, we talk of **logic**, meaning a **logic agent**.

Note. ***In this course we only study logic agents.***

Tell – TellW, TellA

Definition(Tell , TellW, TellA). Let A_L be an agent. We define

$$\text{Tell} = \langle \text{TellW}, \text{TellA} \rangle$$

where

- **TellW(A_L, w),** to be read **Tell-word**, extends the agent language (and interpretation function) with a new input word **w**, and
- **TellA(A_L, a),** to be read as **Tell-axiom**, extends the agent theory and (model) with a new formula **a**. We call a formula **a** added by TellA, **axioms.**

Observation (Effects of TellW). TellW is used to allow for more richness in the language used to describe the wrold. It may result in one of two effecs: (1) it introduces a new name for an already known element of the domain, i.e., it adds a synonym, or (2) it introduces a new word and it extends, via the interpretation function, the domain of interpretation.

Observation (Effects of TellA). TellA is used to increase what is known about the world, that is, that is, to decrease the partiality of theory. Using the terminology used before, the goal of TellA is refine the intended model.

TellA – intended model, model

Reprise (fact, model, domain of interpretation). As from before M = {f} ⊆ D, where M is the set of facts which are the case and D is the space of possible models, not necessarily mutually consistent.

Observation (Intended model, model). Given that TellA changes the model, then we have two models, before and after the execution of TellA. Both representations should be assumed to describe and depict the intended model, as it is in the mind of the user. The idea is that TellA allows to provide more information about the intended model. The result is that there is a space of intended models which can be constructed with multiple executions of TellA. This space contains all the models which can be described by the world model language and for which nothing is said by the previous axioms.

Observation (Effects of TellA). The execution of TellA will have one of three effects:

- M is not changed. This is the case when the newly added axiom was entailed by the model.
- M is increased. This is the case when the newly added axiom was not entailed by the model.
- M is increased with a fact which is inconsistent with M. This is the case when the added axiom contradicts what was already entailed by the model.

TellT – Maximal model, maximal theory

Observation (Maximal model, canonical model). The question is whether there is a maximal model, beyond which no more axioms can be asserted. One could think that the maximal model coincides with the canonical model. This assumption, can be made if one assumes that TellA is done only once. That is, the maximal model coincides with all I know for the input theory. But this does not apply with multiple executions of TellA. Each execution of TellA should in fact be interpreted as a refinement of the canonical model, rather than a completion of an otherwise incomplete model.

Definition (Maximal model). A maximal model is a model for which no more facts can be added without generating an inconsistent model.

Observation (Maximal model). A maximal model is a model for which there are no more formulas in the language which can be added without making the model inconsistent. For instance, with a language which allows for negation, one cannot add an axiom which contradicts a previously introduced axiom. One example is adding “It does no train” with the axiom “It does rain”:

Observation (Maximal theories of maximal models). Maximal theories of maximal models describe everything which can be known. Modulo synonyms, they cannot further enriched.

Ask – AskC, AskS

Definition(Ask , AskC, AskS). Let A_L be an agent. We define

$$Ask = \langle AskC, AskS \rangle$$

where

- **AskC (A_L, T)**, to be read as **Check-model**, checks whether the intended model **M** entails the input theory **T**;
- **AskS (A_L, T)**, to be read as **Sat-Theory**, or also **Find-Model**, checks whether the interpretation of the theory **T** is consistent with the model of $\mathbf{A_L}$. We optionally call the formulas added by TellT, **axioms.**

Observation (Interpretation od AskC). If AskC (A_L, T), then TellA(A_L, T), does not change the contents of M. Given a model previously defined, AsKC is used to check whether a certain property, modeled by T is actually satisfied by M. Model checking is extensively used in high quality SW (e.g., safety critical SW) and HW verification.

Observation (Interpretation of AskS). If AskS (A_L, T), then TellA(A_L, T) increases the contents of M without generating an inconsistent model.

Entailment – model checking

Definition (Model checking, AskC). Let W =< L_a, D, I_a > be a world model. Let L = {w} be a language, with $L_a \subseteq$ L. Let M ⊆ D. Let T = {w}, with T ⊆ L be a theory. Then, the operation **AskC** (M,T) is such that

AskC (M,T) returns **yes** if M |= T, **no** otherwise

Observation (Model checking). Model checking is the key reasoning step. It checks the (in)correctness of T with respect to M, namely whether M is a model of T.

Entailment – satisfiability

Definition (Satisfiability, Model finding, AskS). Let W =< L_a, D, I_a > be a world model. Let L= {w} be a language, with $L_a \subseteq$ L. Then the operation **AskS** (D,T) is such that

AskS (D',T) returns **yes** if there exists an M ⊆ D' such that M |= T,
no otherwise

where D' is D reduced of all the facts which cannot be added because inconsistent with M. If AskS (D',T) **returns yes**, then we say that T **is satisfiable by** M.

Terminology (AskS). We also write **AskS(D',T)** as **AskS(M,T)** or **SAT(M,T)** or **SAT_M(T)**, or **SAT(T)**.

Observation (AskS). M can also be the empty set.

Observation (AskS). AskS requires **searching** for all possible models M in D' and then model checking them.

Observation (AskS, AskC). AskS uses AskC for model checking each and every model till M is found.

Observation (AskS*, AskC). AskS* returns all models M such that M |= T.

Index

- World models
- World representations
- World logics
- Logics
- Agents
- **How to use logics**
- Key Notions

Terminology

Terminology. Let A be an agent. We write

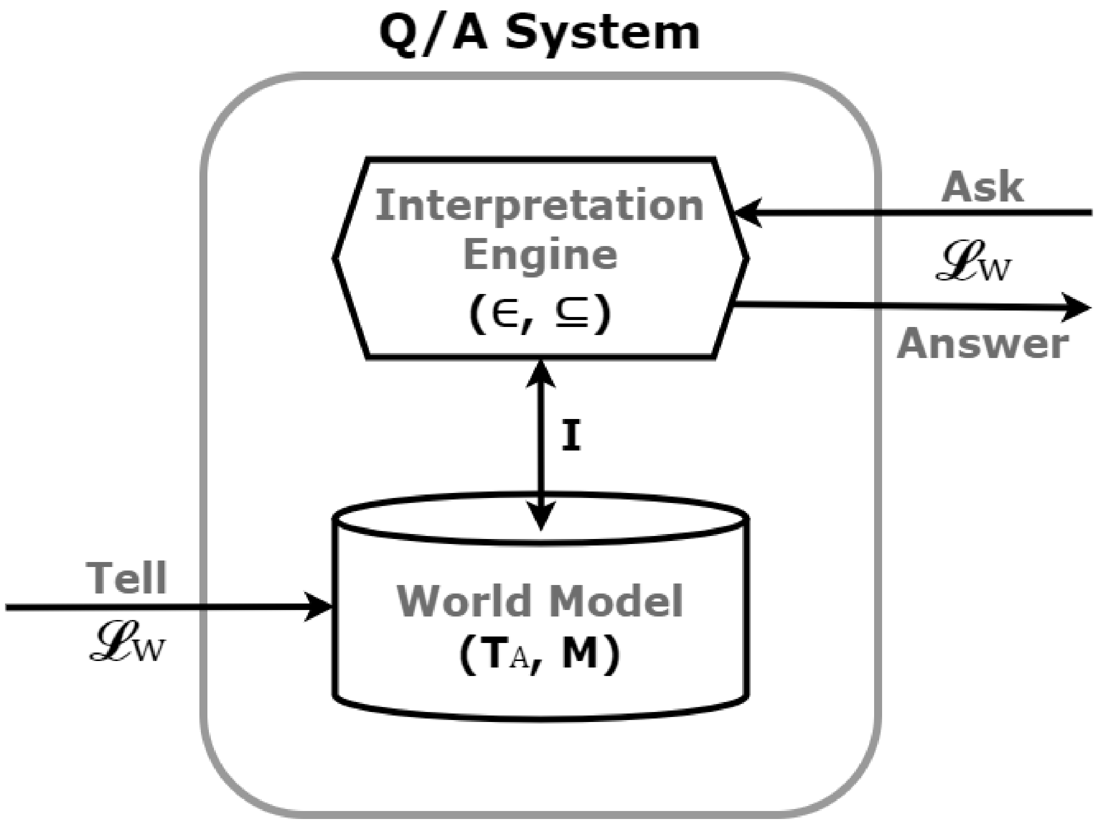
- **M ?= T** to mean **AskC** (A,T)
- **D ??= T** to mean **AskS** (A,T)
- **D != A** to mean **TellL** (A, L)
- **M != T** to mean **TellT** (A,T)

Observation (Ask/Tell language). In many world representations (e.g., DB, OWL/SparQL in the Web) the language used to Ask/Tell in agent is different from the language of the underlying logic. There are three main reasons why this may happen: (1) it is more convenient, (2) it is more expressive, (3) it is a reference standard language.

Ask / Tell an agent

$M \mathrel{!}= T_{L'}$

$D \mathrel{!}= A_{L'}$



$M \mathrel{?}= T_{L'}$

$D \mathrel{??}= T_{L'}$

The subscript L' means that the Ask/Tell language is not necessarily the same as the language L_α used inside the world model

How to use a logic / agent

1. **Decide upon** the domain of interpretation D
2. **Select** L_a, I_a - thus defining $W = \langle L_a, D, I_a \rangle$, L_W, L_L
3. **Tell** D != A, M != T
4. **Ask** M ?= T, D ??= T

World models and logics – observations (continued)

Observation (Select / Agree about the world model) These two steps are the two main modeling decisions once the domain of interpretation has been identified. First choose the reference model to use, then select the language and, via the interpretation function, the intended model. These choices are done by the SW Engineers at design time.

Observation (Tell / Ask the world representation). These operations are performed when the system is in production, as part of its "normal" functioning. Tell, in the case of logics, is used only once, before any Ask operation is performed.

Using a logic / agent (example)

Example (Query Answering Q/A in DBs): a DB is the implementation of a (semi-formal world model and) world representation

- The DB language is the world model,
- The contents of the DB are the world representation,
- The DB is continuously told about new facts, possibly non monotonically (what was true before become false and vice versa),
- The query is the theory to be model checked,
- The answer is the set of instantiations which make the input theory correct (or the single tuple if there is no variable in query).

Observation (Generality of Q/A mechanism) Q/A can be applied to all world models (e.g., the logical formalizations of DBs, ER models, …).

Index

- World models
- World representations
- World logics
- Logics
- Agents
- How to use logics
- **Key Notions**

Key notions

- Formal world models
- Semi-formal world models
- Informal world models
- Interpretation function
- Arbitrariety of domains
- Syntax and semantics
- Language correctness and completeness
- Theory correctness and completeness
- Model
- Theory
- Maximal theories
- World representation
- World entailment
- Truth and Falsity
- World Logics
- World logic entailment
- World logic representation
- Logics
- Entailment
- Families of logics
- Types of logics
- Well-formed formulas
- Logic representation
- Model checking
- Satisfiability
- Consistency
- Expressivity vs. efficiency
- Agents
- Ask /Tell
- How to use a Logic / agent
- World, world model, world representation, logic, agent diversity

World
models, representations, logics and agents
(HP2T)