The Four Notions of Complexity of Parametric Logic

Eric Martin

School of Computer Science and Engineering
University of New South Wales
Sydney, Australia

CMU 06.07.2013

Overview (1)

An inductive problem is often defined as discovering a sentence which best describes/explains/models a set of observations. We can then look for a sentence with minimal complexity, for some notion of complexity.

We present *Parametric Logic*, a framework with a generalised notion of logical consequence, with deductive consequence and inductive consequence as particular cases, together with more complex particular cases of logical consequence.

Such a framework allows one to derive "interesting" inductive consequences of an underlying theory, besides "interesting" deductive consequences, and "interesting" other generalised logical consequences, the most "interesting" consequences not being necessarily the simplest ones.

Overview (2)

We present this framework and in particular, four notions of complexity:

- a model-theoretical, a topological, and a syntactic notion of complexity which are all closely related and which are crucial in proving a completeness result;
- a notion of complexity coming from formal learning theory which can be cast into this framework and support the view that "learning is proving".

The first three notions of complexity that characterise induction are therefore internal, emerging from the underlying logical framework, and can naturally be contrasted with the characterisation of either simpler or more complex particular cases of generalised logical consequence.

The main motivation behind parametric logic (1)

Here is a standard introduction to logic in AI:

- Intelligent agents need to perform this kind of reasoning:
  1. $$\frac{\text{bird(tweety)}}{\text{flies(tweety)}}$$
  2. $$\frac{\text{bird(tweety)} \qquad \text{penguin(tweety)}}{\neg\text{flies(tweety)}}$$
- That's nonmonotonic. But classical logic is monotonic, hence classical logic is *inappropriate* (not too weak) to model the reasoning abilities of intelligent agents.

The main motivation behind parametric logic (2)

This is disturbing for 2 reasons.

- First-order logic allows one to develop set theory, and then almost all mathematical theories, which provide all the machinery to develop the whole of physics, from quantum physics to cosmology — pretty good models of the world. Deeming first-order logic to be inappropriate to model a trivial scenario involving an improbable penguin is bold.
- An intelligent agent might need to reason about penguins *and* mathematical objects, perhaps in the same context. If they need incompatible logical frameworks depending on which kind of reasoning they perform, intelligent agents must in some way be schizophrenic.

An attempt to unify various forms of reasoning: ILP (1)

From the logic program

```
concat([], L, L).
concat([E|R], L, [E|L1]) :- concat(R, L, L1).
```

one can derive all its logical consequences, for instance:

concat([1,2], [3,4], [1,2,3,4])   concat([0,2], [1], [0,2,1])

thanks to a specific form of the resolution rule

$$\frac{\bigvee_{1\leq i\leq n}\varphi_i \vee \xi \qquad \bigvee_{1\leq j\leq m}\psi_j \vee \neg\xi}{\bigvee_{1\leq i\leq n}\varphi_i \vee \bigvee_{1\leq j\leq m}\psi_j}$$

An attempt to unify various forms of reasoning: ILP (2)

From an enumeration of facts, which for instance starts with

concat([1,2], [3,4], [1,2,3,4])   concat([0,2], [1], [0,2,1])

one tries to induce a logic program which generates them, such as

```
concat([], L, L).
concat([E|R], L, [E|L1]) :- concat(R, L, L1).
```

thanks to particular inference rules, such as inverse resolution

$$\frac{\bigvee_{1\leq i\leq n}\varphi_i \vee \bigvee_{1\leq j\leq m}\psi_j}{\bigvee_{1\leq i\leq n}\varphi_i \vee \xi \qquad \bigvee_{1\leq j\leq m}\psi_j \vee \neg\xi}$$

Could induction be the inverse of deduction?

Formal learning theory with any number of mind changes

Consider the set of finite sets.

An example of learning in the limit: being presented with each element of the infinite sequence of data that starts with

$$3 \quad 3 \quad 8 \quad 13 \quad 8 \quad 10 \quad 13$$

a smart learner who wishes to find out what the underlying set is could emit an infinite set of hypotheses which starts with

$$\{3\} \quad \{3\} \quad \{3,8\} \quad \{3,8,13\} \quad \{3,8,13\} \quad \{3,8,10,13\} \quad \{3,8,10,13\}$$

This could be presented as a form of logical inference.

Formal learning theory with at most one mind change

Consider the union of $\{\mathbb{N}\}$ with the set of all sets of the form $\mathbb{N} \setminus \{n\}$, $n \in \mathbb{N}$.

An example of learning with at most one mind change: being presented with each element of the infinite sequence of data that starts with

$$(3,1) \quad (3,1) \quad (8,1) \quad (5,1) \quad (1,0) \quad (1,0) \quad (9,1) \quad (1,0)$$

a smart learner who wishes to find out what the underlying set could emit an infinite set of hypotheses which starts with

$$\mathbb{N} \quad \mathbb{N} \quad \mathbb{N} \quad \mathbb{N} \quad \mathbb{N}\setminus\{1\} \quad \mathbb{N}\setminus\{1\} \quad \mathbb{N}\setminus\{1\} \quad \mathbb{N}\setminus\{1\}$$

This could be presented as a form of inductive logical inference.

Formal learning theory with no mind change

Consider the set of all sets of the form $\mathbb{N}\setminus\{n\}$, $n \in \mathbb{N}$.

An example of learning with no mind change (finite learning): being presented with each element of the infinite sequence of data that starts with

$$(3,1) \quad (3,1) \quad (8,1) \quad (3,1) \quad (1,0) \quad (1,0) \quad (3,1) \quad (1,0)$$

a smart learner who wishes to find out what the underlying set could emit an infinite set of hypotheses which starts with

$$? \quad ? \quad ? \quad ? \quad \mathbb{N}\setminus\{1\} \quad \mathbb{N}\setminus\{1\} \quad \mathbb{N}\setminus\{1\} \quad \mathbb{N}\setminus\{1\}$$

This could be presented as a form of deductive logical inference.

Formal learning theory with at most ω mind change

Consider the union of $\mathbb{N}\setminus\{0\}$ with the set of all sets of the form $\mathbb{N}\setminus\{1,\dots,n\}$, $n \in \mathbb{N}\setminus\{0\}$.

An example of learning with less than ω mind changes: being presented with each element of the infinite sequence of data that starts with

$$6 \quad 8 \quad 5 \quad 0 \quad 8 \quad 7 \quad 4 \quad 11$$

a smart learner who wishes to find out what the underlying set could emit an infinite set of hypotheses which starts with

$$\mathbb{N}\setminus\{0\} \quad \mathbb{N}\setminus\{0\} \quad \mathbb{N}\setminus\{0\} \quad \mathbb{N}\setminus\{1,2,3,4\} \quad \mathbb{N}\setminus\{1,2,3,4\}$$

$$\mathbb{N}\setminus\{1,2,3,4\} \qquad \mathbb{N}\setminus\{1,2,3\} \quad \mathbb{N}\setminus\{1,2,3\}$$

This could be presented as a form of logical inference more complex than deduction and induction.

A few working assumptions

- Deducing is discovering the truth.
- Inducing is discovering the truth.
- Inducing is harder than deducing.
- Discovering the truth might necessitate to combine deductions, inductions, and probably other forms of inference.
- "To learn" in the sense of formal learning theory (FLT) means "to prove".
- The compactness property is a key notion to see clearly what is going on, and it is related to the notion of *finite telltale* of FLT.
- Classical logic should be gracefully integrated.

Appealing, but there seems to be a catch

How can we develop a theoretical framework where

φ is a deductive consequence of T

$\Downarrow$

φ is an inductive consequence of T

$\Downarrow$

φ is a logical consequence of T

which does not scratch classical logic?

The axioms proposed by Tarski

Tarski proposes

> *four axioms which express certain elementary properties of the primitive concepts and are satisfied in all known formalized disciplines.*

Denoting the vocabulary by $\mathcal{V}$ and the language by $\mathcal{L}$, these axioms are:

1. $\mathcal{V}$ is countable.
2. For all $K \subseteq \mathcal{L}$, $K \subseteq \mathbf{Cn}(K) \subseteq \mathcal{L}$.
3. For all $K \subseteq \mathcal{L}$, $\mathbf{Cn}(\mathbf{Cn}(K)) = \mathbf{Cn}(K)$.
4. For all $K \subseteq \mathcal{L}$, $\mathbf{Cn}(K) = \bigcup\{\mathbf{Cn}(D) \mid D \text{ is finite and } D \subseteq K\}$.

Tarksi's comment on the fourth axiom

Tarski justifies as follows the fourth axiom:

> *Finally, it should be noted that, in concrete disciplines, the rules of inference with the help of which the consequences of a set of sentences are formed are in practice always operations which can be carried out only on a finite number of sentences* (*usually even on one or two sentences*).

What does Tarski talk about: deductive consequence, or logical consequence?

These quotes are taken from an article whose title evokes the *deductive sciences*, and in which no distinction seems to be made bwteen *deductive sciences*, *concrete disciplines* et *formalized disciplines*.

Elsewhere, Tarski talks about a *formalized concept of consequence* in relation to the *formalized deductive theories*.

It seems that for Tarski, *logical consequence* and *deductive consequence* are two synonymic expressions.

There is no reason to blame Tarski for that view as far as the compactness property is taken as a basic principle.

On this basis, where to go?

But should we impose compactness to every notion of logical consequence? We suspect this will leave induction outside the logical realm.

Essentially, Tarski posits that one cannot seriously consider inference rules which need infinitely many premisses, such as

$$\frac{p(\overline{0}) \quad p(\overline{1}) \quad p(\overline{2}) \quad p(\overline{3}) \quad p(\overline{4}) \quad \ldots}{\forall x\, p(x)}$$

But is there anything to object to a theoretical framework which enjoys

- a notion of logical consequence which is not compact,
- a notion of proof where every derivation is finite, and performed on the basis of inference rules with a finite number of premisses, and
- a completeness theorem?

Which path to follow?

To define a logical framework which depends on a number of parameters.

- When the matter is to model mathematical reasoning, the parameters should naturally be given values such that what we get is precisely classical first-order logic.
- When the matter is to model forms of reasoning using processes which are genuinely inductive, one should naturally be led to give those parameters other values.
- It has to be possible to talk about logical consequence, deductive consequence and inductive consequence in a generic manner, without making any assumption on the values of the parameters.
- When the values given to the parameters are such that the associated notion of logical consequence is compact, the notion of logical consequence, deductive consequence and inductive consequence should be equivalent.

Which parameters? The possible worlds

Genuinely inductive reasoning does not seem to be possible if the class of intended interpretations does not consist exclusively of Henkin structures, where every individual has a name, or is at least definable.

$$\frac{\text{Phineas is a white swan!} \quad \text{The swan on the lake over there is white!}}{\text{All swans are white}}$$

How is it possible to *refute* that all swans are white if some black swans are not within the domain of discourse, if someone who sees a black swan cannot talk about it except by saying *I have seen a black swan*, which in a logical language, would be expressed as *there exists a black swan*?

Which parameters? The possible data

How is it possible to *refute* that all swans are white if no black swan shows up?

One of the fondamental assumptions of the paradigms of formal learning theory is that all elements from the underlying set are sooner or later presented to the learner.

All nonmonotonic logics make a tacit use of this assumption: stating that *Tweety* flies is sensible only because in case it is a penguin or an emu, we will find out about it.

This is the idea behind the notion of *closed world assumption*, *circumscription*, etc.

The parameters of parametric logic

- A countable ordinal κ: reasoning can be done at the object level, at the meta-level, at the meta-meta-level, or more generally at any λ-level with $\lambda < \kappa$
- a vocabulary $\mathcal{V}$: a countable set of fonction and predicate symbols, with or without equality
- a set of possible worlds: a set of κ-structures (when $\kappa = 0$ those are standard, that is, Henkin, structures)
- a language $\mathcal{L}$: countable fragment of $\mathcal{L}^{\kappa}_{\omega_1\omega}(\mathcal{V})$
- a set of possible data $\mathcal{D}$: subset of $\mathcal{L}$
- a set of possible axioms $\mathcal{A}$: subset of $\mathcal{L}$

Every instance of $(\kappa, \mathcal{V}, \mathcal{W}, \mathcal{L}, \mathcal{D}, \mathcal{A})$ determines a logical paradigm $\mathcal{P}$.

The particular case of classical logic

- κ: 1
- $\mathcal{V}$: union of a vocabulary V with a countable set of $\mathcal{V}$-terms which are not V-terms ($\mathcal{V}$ could contain infinitely constants not in V, or a unary function symbol not in V)
- $\mathcal{W}$: set of all standard (Henkin) $\mathcal{V}$-structures
- $\mathcal{L}$: $\mathcal{L}^1_{\omega\omega}(V)$
- $\mathcal{D}$: $\varnothing$
- $\mathcal{A}$: $\mathcal{L}$

The set of possible theories

Given a possible world $\mathfrak{M}$, $\text{Diag}_{\mathcal{D}}(\mathfrak{M})$ denotes the $\mathcal{D}$-diagram of $\mathfrak{M}$, that is, the set of possible data which are true in $\mathfrak{M}$.

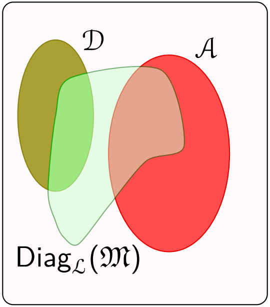
Definition

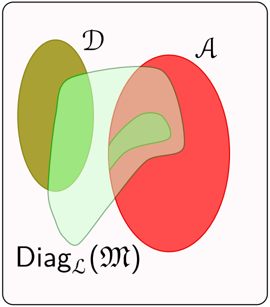
A possible theory is a set of the form $\text{Diag}_{\mathcal{D}}(\mathfrak{M}) \cup A$ where $\mathfrak{M}$ is a possible world and A is a set of possible axioms which are all true in $\mathfrak{M}$.

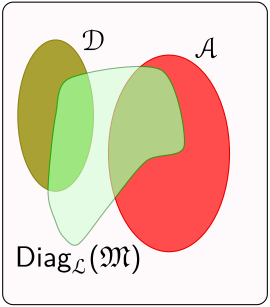
Property

When $\mathcal{P}$ is the paradigm that defines classical logic, the possible theories are the consistent theories.

A few possible theories







The notion of (generalised) logical consequence

Definition

Given a possible theory T and a sentence φ, φ is a logical consequence of T in $\mathcal{P}$ iff every model of T in $\mathcal{W}$ whose $\mathcal{D}$-diagram is $T \cap \mathcal{D}$ is a model of φ. We then write $\varphi \in \mathrm{Cn}^{\mathcal{D}}_{\mathcal{W}}(T)$ or $T \models^{\mathcal{D}}_{\mathcal{W}} \varphi$.

This notion is closely related to the notion of preferential entailment from the nonmonotonic reasoning literature, etc.

The notion of deductive consequence

Definition

Given a possible theory T and a sentence φ, φ is a deductive consequence of T in $\mathcal{P}$ iff there exists a finite subset D of T such that for all possible theories T':

if T' contains D then φ is a logical consequence of T' in $\mathcal{P}$.

Example

Set $\mathcal{W} = \{\mathfrak{M}_L \mid L \in \mathfrak{L}\}$ where $\mathfrak{L}$ is the set of final segments of $\mathbb{N}$. Set $\mathcal{D} = \{P(\overline{n}) \mid n \in \mathbb{N}\}$ and $\mathcal{A} = \varnothing$. Let T be $\{P(\overline{n}) \mid n \geq 2\}$ and $\varphi = \forall x\ P(s(s(s(x))))$.

φ is a deductive consequence of T in $\mathcal{P}$. Indeed, T contains $P(\overline{3})$, and every model of $P(\overline{3})$ in $\mathcal{W}$ is a model of φ.

Note that $T \not\models \varphi$.

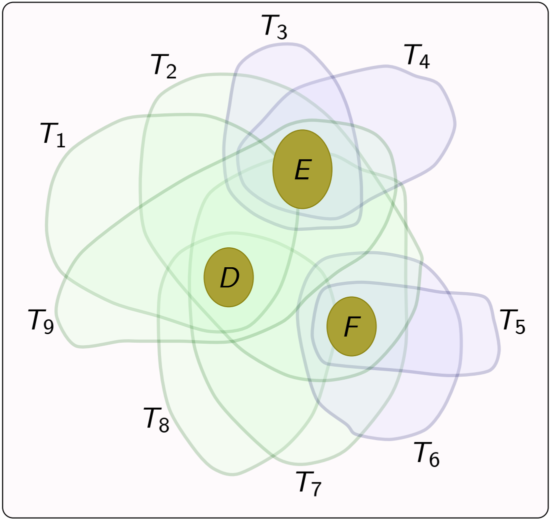
Deduction = compactness

$T_1, T_2, T_7, T_8, T_9 \models^{\mathcal{D}}_{\mathcal{W}} \varphi$

$T_2, T_3, T_4, T_9 \models^{\mathcal{D}}_{\mathcal{W}} \psi$

$T_2, T_5, T_6, T_7 \models^{\mathcal{D}}_{\mathcal{W}} \chi$

All those logical consequences in $\mathcal{P}$ are also deductive consequences in $\mathcal{P}$.



The notion of inductive consequence (1)

Definition

Given a possible theory T and a sentence φ, φ is an inductive consequence of T in $\mathcal{P}$ iff:

- φ is a logical consequence of T in $\mathcal{P}$;
- there exists a finite subset D of T such that for all possible theories T':

  if T' contains D and if φ is not a logical consequence of T' in $\mathcal{P}$ then $\sim\varphi$ is a deductive consequence of T' in $\mathcal{P}$.

This is a formalisation of the Popperian principle of *falsifiability*.

The notion of inductive consequence (2)

Example

Set $\mathcal{W} = \{\mathfrak{M}_L \mid L \in \mathfrak{L}\}$ where $\mathfrak{L}$ is the set of final segments of $\mathbb{N}$. Set $\mathcal{D} = \{P(\overline{n}) \mid n \in \mathbb{N}\}$ and $\mathcal{A} = \varnothing$. Let T be $\{P(\overline{n}) \mid n \geq 2\}$ et $\varphi = \forall x\,(\overline{2} \leq x \leftrightarrow P(x))$.

- φ is not a deductive consequence of T in $\mathcal{P}$. Indeed, for every finite subset D of T, there exists a model of D in $\mathcal{W}$ which is not a model of φ.
- φ is an inductive consequence of T in $\mathcal{P}$. Indeed, T contains $P(\overline{2})$, and every possible theory T' which contains $P(\overline{2})$ but does not logically imply φ in $\mathcal{P}$ is such that $\sim\varphi$ is a deductive consequence of T' in $\mathcal{P}$, since it contains $P(\overline{1})$.

Induction = weak compactness

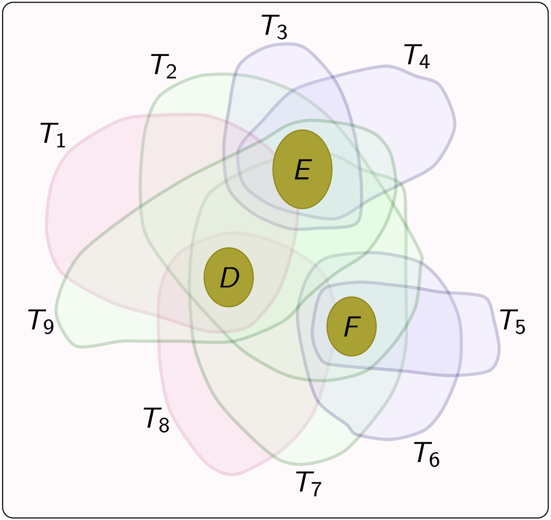
$T_1, T_8 \models^{\mathcal{D}}_{\mathcal{W}} \varphi$

$T_3, T_4, T_9 \models^{\mathcal{D}}_{\mathcal{W}} \sim\varphi$

$T_5, T_6, T_7 \models^{\mathcal{D}}_{\mathcal{W}} \sim\varphi$

$T_2 \models^{\mathcal{D}}_{\mathcal{W}} \sim\varphi$

φ is also an inductive consequence of T_1 and T_8 in $\mathcal{P}$.



The notion of β-compactness

Definition

Given a possible theory T, an ordinal β and a sentence φ, we put φ on stratum β of the hierarchy built over T iff:

- φ is a logical consequence of T in $\mathcal{P}$;
- there exists a finite subset D of T such that for all possible theories T':

  if T' contains D and if φ is not a logical consequence of T' in $\mathcal{P}$ then $\sim\varphi$ occurs in the hierarchy built over T' *below* stratum β.

We denote by $\Lambda^{\mathcal{P}}_{1,\beta}(T)$ the set of logical consequences of T which occur on stratum β or below.

- $\Lambda^{\mathcal{P}}_{1,0}(T)$ consists of the deductive consequences of T.
- $\Lambda^{\mathcal{P}}_{1,1}(T)$ consists of the inductive consequences of T.

The hierarchies of logical consequences in $\mathcal{P}$

Definition

Given a possible theory T, an nonzero ordinal α, an ordinal β and a sentence φ, we put φ on stratum β of level α of the hierarchy built over T iff:

- φ is a logical consequence of T in $\mathcal{P}$;
- there exits a finite subset D of T and a finite subset H of $\bigcup_{(\alpha',\beta')<(\alpha)} \Lambda^{\mathcal{P}}_{\alpha'\beta'}(T)$ such that for all possible theories T':
  if T' contains D, if $\bigcup_{(\alpha',\beta')<(\alpha)} \Lambda^{\mathcal{P}}_{\alpha'\beta'}(T')$ contains H and if φ is not a logical consequence of T' in $\mathcal{P}$ then $\sim\varphi \in \bigcup_{\beta'<\beta} \Lambda^{\mathcal{P}}_{\alpha,\beta'}(T')$.

We denote by $\Lambda^{\mathcal{P}}_{\alpha,\beta}(T)$ the set of all logical consequences of T which occur on or below stratum β of level α.

Example

Suppose that $\mathcal{V}$ contains at least one function symbol, and assume that $\mathcal{W}$ is the set of all standard models of:

- $\forall x\,(\text{animal}(x) \leftrightarrow \neg\text{plane}(x))$;
- $\forall x\,(\text{animal}(x) \leftrightarrow (\text{dog}(x) \leftrightarrow \neg\text{bird}(x)))$;
- $\forall x\,(\text{bird}(x) \leftrightarrow (\text{finch}(x) \leftrightarrow \neg\text{penguin}(x)))$;
- $\forall x\,(\text{flies}(x) \leftrightarrow (\text{plane}(x) \vee \text{finch}(x)))$;
- $\forall x\,(\text{flies}(x) \rightarrow \neg\text{barks}(x))$.

Assume that $\mathcal{D}$ is

$$\{\text{animal}(t), \text{bird}(t), \text{penguin}(t) \mid \text{term clos } t\}.$$

Example: $\Lambda^{\mathcal{P}}_{1,0}(T)$

$T \models^{\mathcal{D}}_{\mathcal{W}} \text{plane}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{dog}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{finch}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{penguin}(t)$
	$\text{animal}(t)$	$\text{animal}(t)$	$\text{animal}(t)$
	$\neg\text{plane}(t)$	$\neg\text{plane}(t)$	$\neg\text{plane}(t)$
		$\text{bird}(t)$	$\text{bird}(t)$
		$\neg\text{dog}(t)$	$\neg\text{dog}(t)$
			$\text{penguin}(t)$
			$\neg\text{finch}(t)$
			$\neg\text{flies}(t)$

Example: $\Lambda^{\mathcal{P}}_{1,1}(T)$

$T \models^{\mathcal{D}}_{\mathcal{W}} \text{plane}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{dog}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{finch}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{penguin}(t)$
	$\text{animal}(t)$	$\text{bird}(t)$	
$\text{plane}(t)$	$\text{dog}(t)$	$\text{finch}(t)$	
		$\text{flies}(t)$	

Example: $\Lambda^{\mathcal{P}}_{1,2}(T)$

plane	dog	
	finch	penguin

$T \models^{\mathcal{D}}_{\mathcal{W}} \text{plane}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{dog}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{finch}(t)$	$T \models^{\mathcal{D}}_{\mathcal{W}} \text{penguin}(t)$
	$\text{animal}(t)$		
	$\neg\text{flies}(t)$		

Example: $\Lambda^{\mathcal{P}}_{1,3}(T)$

<table>
<tr><td rowspan="2">plane</td><td colspan="2">dog</td></tr>
<tr><td>finch</td><td>penguin</td></tr>
</table>

$T \models^{\mathcal{D}}_{\mathcal{W}} \text{plane}(t)$ $\quad T \models^{\mathcal{D}}_{\mathcal{W}} \text{dog}(t)$ $\quad T \models^{\mathcal{D}}_{\mathcal{W}} \text{finch}(t)$ $\quad T \models^{\mathcal{D}}_{\mathcal{W}} \text{penguin}(t)$

$\text{flies}(t)$

Example: $\Lambda^{\mathcal{P}}_{2,0}(T)$

If T is a possible theory such that $T \models^{\mathcal{D}}_{\mathcal{W}} \neg\text{barks}(t)$, then $\neg\text{barks}(t) \in \Lambda^{\mathcal{P}}_{2,0}(T)$.

Indeed, note that for all possible theories T,

- $T \models^{\mathcal{D}}_{\mathcal{W}} \neg\text{barks}(t)$ iff $T \models^{\mathcal{D}}_{\mathcal{W}} \text{flies}(t)$;
- if $\neg\text{barks}(t)$ is not a logical consequence of T in $\mathcal{P}$ then barks(t) is not a logical consequence of T in $\mathcal{P}$ either.

Hence there exists no possible theory T such that $\neg\text{barks}(t) \in \Lambda^{\mathcal{P}}_{1,\beta}(T)$ for some ordinal β.

Also note that one would like to express that $\neg\text{barks}(t)$ is not provable (*i.e.*, $\Diamond\text{barks}(t)$ is true).

Example: $\Lambda^{\mathcal{P}}_{2,1}(T)$

Suppose that $\mathcal{V}$ is such that there exists infinitely many closed terms.

If T is a possible theory such that $T \models^{\mathcal{D}}_{\mathcal{W}} \forall x\, \text{flies}(x)$, then $\forall x\, \text{flies}(x) \in \Lambda^{\mathcal{P}}_{2,1}(T)$.

Indeed, if $\forall x\, \text{flies}(x)$ is not a logical consequence in $\mathcal{P}$ of a possible theory T, then $T \models^{\mathcal{D}}_{\mathcal{W}} \exists x\, \neg\text{flies}(x)$, a sentence which necessarily belongs to $\Lambda^{\mathcal{P}}_{2,0}(T)$, but of course not to $\Lambda^{\mathcal{P}}_{1,\beta}(T)$ for any ordinal β.

Proof theory

In parametric logic, inference rules are of the form

$$\frac{\xi_1 \ \ldots \ \xi_m \quad (\alpha_1, \beta_1, \psi_1) \ \ldots \ (\alpha_n, \beta_n, \psi_n)}{(\alpha, \beta, \psi)}$$

where α_1, ..., α_n, β_1, ..., β_n are ordinals such that

$$(\alpha, \beta) \geq \sup((\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n)) \geq (1, 0)$$

- ξ_1, ..., ξ_m belong to the underlying possible theory T;
- when this rule is used, ψ is provisionally (but possibly eventually) supposed to occur on stratum β of level α of the hierarchy built over T;
- (α, β) represents a degree of incredulity: the larger this degree, the smaller the guarantee that ψ will resist any of the following inferences.

The particular case of classical logic

We have for all sentences φ the sentence

$$\frac{\varphi}{(1, 0, \varphi)}$$

All other rules are of the form:

$$\frac{(1, 0, \psi_1) \ \ldots \ (1, 0, \psi_n)}{(1, 0, \psi)}$$

Generalised proofs: the case $(\alpha, \beta) \leq (2, 0)$

Imagine an annotated derivation where, for a member (α, β, ψ) of the derivation, we indicate the previous members of the derivation, namely $(\alpha_1, \beta_1, \psi_1)$, ..., $(\alpha_n, \beta_n, \psi_n)$ which, thanks to an inference rule of the form

$$\frac{\xi_1 \ \ldots \ \xi_m \quad (\alpha_1, \beta_1, \psi_1) \ \ldots \ (\alpha_n, \beta_n, \psi_n)}{(\alpha, \beta, \psi)}$$

where all $(\alpha_i, \beta_i, \psi_i)$ are active, have made it possible to derive (α, β, ψ).

- We have derived (α, β, ψ) only because the derivation contains no active member of the form $(\alpha', \beta', \sim\psi)$ with $(\alpha', \beta') < (\alpha, \beta)$.
- We have deactivated from the derivation
  - all triples of the form $(\alpha', \beta', \sim\psi)$ such that $(\alpha', \beta') > (\alpha, \beta)$;
  - all triples which depend on one of the preceding triples.

Sound set of rules

Limiting ourselves to derivations of type $(2, 0)$ *or less*, a set of rules is sound for $\mathcal{P}$ iff for all sentences φ which appear in a derivation, the following conditions are satisfied.

- If there exists an infinite derivation where φ occurs for the last time, as member of a triple (α, β, φ), then φ belongs to the hierarchy built over the possible theory which underlies the proof, on stratum β of level α or below.
- if there exists an infinite derivation where φ occurs infinitely many times, then φ does not belong to the hierarchy built over the possible theory which underlies the proof, or it occurs above stratum β of level α.

Complete set of rules

Limiting ourselves to derivations of type $(2, 0)$ *or less*, a set of rules is complete for $\mathcal{P}$ iff it is sound for $\mathcal{P}$ and the following conditions are satisfied.

- If there exists an infinite derivation where a sentence φ occurs for the last time, as member of a triple (α, β, φ), then φ belongs to the hierarchy built over the possible theory which underlies the proof, on stratum β of level α and not below.
- Let a sentence φ occur in the hierarchy built over a possible theory T, on stratum β of level α and not below, with $(\alpha, \beta) \leq (2, 0)$. Then there exists an infinite derivation where φ occurs for the last time, as member of a triple (α, β, φ).

A particular paradigm

Definition

A set Z of formulas is a generator of a set X of sentences if X is the set of closed instances of the members of Z.

Proposition

Let C be an infinite set of constants. Assume that the following holds.

- $\mathcal{V}$ *is of the form* $V \cup C$, $\mathcal{L} = \mathcal{L}_{\omega\omega}(\mathcal{V})$ *and* $\mathcal{A} = \varnothing$;
- *an r.e. set of* V*-formulas is a generator of* $\mathcal{D}$;
- $\mathcal{W}$ *is the set of standard* $\mathcal{V}$*-models of an r.e. set of* V*-sentences.*

Then the first stratum of level ω *or stratum* ω *of any level of the hierarchy built over a possible theory* T *does not contain any sentence which does not occur below in the hierarchy.*

A completeness theorem

Proposition

Let C be an infinite set of constants. Assume that the following holds.

- $\mathcal{V}$ *is of the form* $V \cup C$, $\mathcal{L} = \mathcal{L}_{\omega\omega}(\mathcal{V})$ *and* $\mathcal{A} = \varnothing$;
- *an r.e. set of* V*-formulas is a generator of* $\mathcal{D}$;
- $\mathcal{W}$ *is the set of standard* $\mathcal{V}$*-models of an r.e. set of* V*-sentences.*

Then there exists a complete r.e. set of inference rules for $\mathcal{P}$.

Four notions of complexity

Parametric defines four notions of complexity.

- Logical complexity: a sentence is of logical complexity (α, β) in $\mathcal{P}$ iff (α, β) is the smallest pair of nonzero ordinals such that for all possible theories T, if φ is a logical consequence of T in $\mathcal{P}$ then φ occurs on the hierarchy built over T below stratum β of level α.
- Complexity in the sense of formal learning theory.
- Topological complexity, in relation to the Borelian and the boolean hierarchies.
- Syntactic complexity, with new canonical forms.

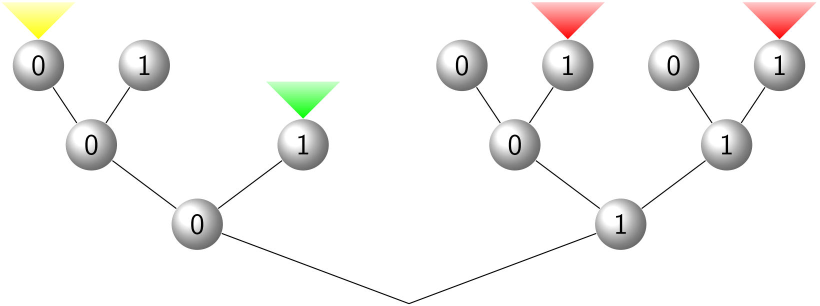
There are strong relationships between all those notions of complexity.

The proof of the previous completeness theorem exploits the relationships between the logical, topological and syntactic complexities.

Borelian hierarchy

Defined over $\mathcal{W}$, and generated by the Π_0 sets, which are finite unions and intersections of sets of models in $\mathcal{W}$ of possible data.

If $\mathcal{D}$ is of the form $\{P(\overline{n}) \mid n \in \mathbb{N}\}$, we represent below sets of complexity Σ_0, Π_0 and Δ_2.

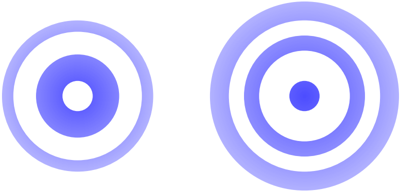


Boolean hierarchy

The Boolean hierarchy is a refinement of the Borelian hierarchy.

More precisely, it is a complete refinement of the sets Δ_α of the Borelian hierarchy.

The picture below represents sets of topological complexity $\Sigma_{\alpha,4}$ and $\Sigma_{\alpha,5}$ if every disk is a set of complexity Σ_α in the Borelian hierarchy.

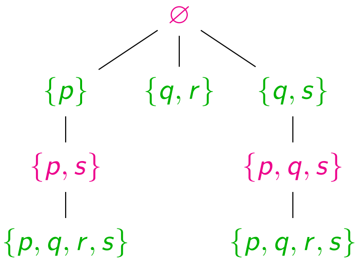


Syntactic complexity

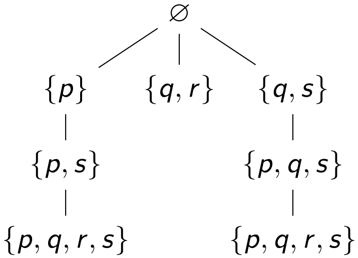
- The Borelian hierarchy is linked to the complexity of a formula in terms of occurrences of quantifiers, infinite conjunctions and infinite disjunctions .
- The boolean hierarchy is linked to the complexity of a formula in terms of occurrences of finite conjunctions and finite disjunctions .

Canonical syntactic forms (1)

p	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
q	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
r	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
s	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
ψ	1	1	1	1	1	0	0	0	0	1	0	1	0	1	0	0



Canonical syntactic forms (2)



$\psi_3 = \bigwedge \emptyset$ $\quad\quad$ $\psi_2 = \bigwedge\{p\} \vee \bigwedge\{q, r\} \vee \bigwedge\{q, s\}$

$\psi_1 = \bigwedge\{p, s\} \vee \bigwedge\{p, q, s\}$ $\quad\quad$ $\psi_0 = \bigwedge\{p, q, r, s\}$

If p, q, r and s are all under normal form $\Sigma_\alpha[X]$ then

$$\bigvee_{i\in\{1,3\}} (\psi_i \wedge \bigwedge_{j<i} \sim\psi_j) \text{ is under normal form } \Sigma_{\alpha,3}[X]$$

Back to the completeness theorem

The set of inference rules of the theorem consists of:

- inference rules of the form

$$\frac{(1,0,\xi_1) \quad \dots \quad (1,0,\xi_n)}{(1,0,\xi)}$$

  with a complete set of rules for first-order logic containing

$$\frac{\xi_1 \quad \dots \quad \xi_n}{\xi}$$

- inference rules of the form

$$\frac{(1,0,\bigwedge\{\xi_1,\dots,\xi_i\} \rightarrow (\varphi \leftrightarrow \psi)) \quad (1,0,\xi_1) \ \dots \quad (1,0,\xi_i)}{(1,m,\varphi)}$$

  where ψ is in canonical syntactic normal form of complexity $\Pi_{1,m}[X]$.