UNIVERSITÀ DEGLI STUDI DI UDINE

Master's degree in Mathematics

Master's thesis

Inductive Logic

SUPERVISOR:
Prof.ssa Giovanna D'agostino

STUDENT:
Michele Mignani

Academic year 2023/2024

Introduction

Modeling knowledge and rationality was one of the reasons why logic was developed starting from Aristotle. Passing through mathematizations of reasoning like Propositional Logic or First Order Logic, there were also attempts to use modal logic to deal with the distance between what is true in the world and what the agent knows about it. For instance, in *Epistemic Logics* some modal operators that represent the knowledge of an agent (or what every agent knows) are added and a Kripke-style semantics is developed (see for further details [18] and [10]).

Even if the modal approach sheds important light on the qualitative study of uncertainty, probability is the tool used from the beginning to investigate the question quantitatively. However, this doesn't mean that what we will present is the perfect approach for dealing with these concepts. Indeed:

- The probabilistic approaches have huge limitations: usually the degree of beliefs of an agent are expressed by real numbers and it's not really clear how to retrieve this value (if it exists) from an agent; moreover, in this framework, we force any two events to be comparable in terms of likelihood to the eyes of an agent. So for a given goal, a different point of view about the topic can be more suitable;
- There are a lot of techniques that use probabilities and what is shown in the thesis is only one among all the possible ways to formalize the question.

Rudolf Carnap's works ([1] among others) are considered one of the main texts about Inductive Logic; with time, the philosophical interests that were at the base of his research began to fade in favor of a more mathematical approach to the topic. Haim Gaifman ([7]), Marcus Hutter ([12]), Jeffrey Paris, and Alena Vencovská ([20], [11], [21], [22], [19]), among others, developed a sound mathematical theory for rational beliefs based on first-order logic and probabilities of sentences. This line of research is called *Inductive Logic* and the main results discovered until now are collected in [20].

The main question that Inductive Logic tries to answer is: "What *degree of belief* should a rational agent have about a certain *property* relative to the *world* it lives in?". Giving a precise meaning to the italic terms is essential to understanding what it really means:

- *degree of belief*: an agent will assign to any *property* a value in $[0, 1]$ depending on how much likely it is: 0 if it is regarded as impossible, 1 if certain;
- *property*: an agent will judge statements expressible in a given logic and a certain language $\mathcal{L}$;[1]

[1] sometimes called *signature*.

- *world*: an agent will be supposed to "live" in a $\mathcal{L}$-structure.

The beliefs of an agent are represented as a map (called *probability*) from the set of all the $\mathcal{L}$-sentences, denoted by $\text{Sen}(\mathcal{L})$, to $[0, 1]$. Not all the maps can describe what a *rational* agent thinks and Chapter 1 is devoted to a more precise discussion about what we mean with *rationality*. There, we will present three different approaches and it will be shown that they all justify the same definition (Definition 1.0.1).

The second chapter will be focused on the *Exchangeability principle (Ex)*, a basic rational requirement that most scholars accept. We will prove a representation theorem (Theorem 2.2.1 and Theorem 2.4.4) for the probabilities satisfying Ex.

At the end of the thesis, in Chapter 3, we will push ourselves to the subject boundaries and we will investigate other principles. Starting from some considerations about the role of *symmetry* in rationality, we will describe other principles and, in particular, the *Invariance* one. After having shown its inadequacy in this context, we will move towards different sources of rational principles and open questions.

Why this topic?

Inductive logic is undoubtedly a fascinating branch of mathematics, in the junction between logic and philosophy. Many results that will be presented are *mathematically interesting* and, in this thesis, we have tried to stress also the elegance of some proofs: we have decided, for instance, to show Theorem 2.4.4 by Nonstandard Analysis tools and not to follow the first chronological proof relying on heavy results of Measure Theory.

However, this is not the main reason why we decided to delve into the topic: the idea of investigating inductive reasoning came from the AI field and some concerns we have.

Some well-known companies (OpenAI, Google DeepMind, IBML, etc.) are trying to build an *AGI*, an artificial general intelligence: AI's today are good only at one specific task (ChatGPT at chatting, DALL-E 3 at creating images, etc.), AGI would be intelligent in a wide variety of occasions. Some experts forecast the appearance of an AGI within 2050 and the concerns of such a creation raise interest in the topic and in the possible ways that we have to test the rationality or the behavior (the alignment of AGI's interest to human moral values) of a potential agent of this kind. We suggest some literature in this direction:

- see [25] for a research agenda of MIRI, the Machine Intelligence Research Institute;
- see [12] for a theoretical discussion of inductive logic and a possible agent learning process;
- see [8] for an effective implementation of the Dutch Book argument (see page 26) to create a rational agent.

But, have we just changed the question? To *"Why this topic?"*, we answered with *"Because of AI"* and the reader may ask *"Why AI?"*. The link between logic, rationality tests, and AI is clear, at least theoretically: but we want to go further and explain why this topic seems so crucial to us.

"AI" is the big word of the moment and it is known that a lot of research in developing this kind of tool is ongoing. What is less known is that there is considerably less effort in trying to figure out plans to progress in AI safely.

In other words: *advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources. Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control. Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: [...] Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete, and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable.*

This is what was written in March 2023 in an open letter that calls AI labs to an immediate 6 months pause in training systems more powerful than ChatGPT4. The letter[2] was signed by more than 30,000 people, among which we can find the 2018 Turing Award winner Yoshua Bengio, the 2022 IJCAI[3] award winner Stuart Russell, and the notorious public intellectual Yuval Noah Harari.

The aim of this proposal was not to shut down all the AI experiments, but to slow them down, to give time to AI researchers for "making today's powerful, state-of-the-art systems more accurate, safe, interpretable, transparent, robust, aligned, trustworthy, and loyal".

After a year, we feel this urgency even more. And we are not the only ones: OpenAI, a leading company in the field (ChatGPT is one of its creations), is trying to collect people to work on AI systems contol[4], Geoffrey Hinton, Turing Award winner and major exponent in AI research, has quitted Google to talk more freely about the issues of the actual systems and the "existential risk of what happens when these things get more intelligent than us"[5]. It's plenty of examples of other scholars concerned about AI systems and about how little time and money is invested in AI safety.[6]

Maybe AGI creation, superintelligent devices, or a lot of the issues previously discussed about AI are only in our mind and, with a bit of sarcasm, only a sci-fi *hallucination*[7]. But if we don't care now, will we have the time to fix our negligence mistakes? If we don't start now to focus part of our attention on this topic, will be able to fill the time gap between the impressively fast progress in AI systems capability and the human-paced AI safety research?

With the philosopher Hans Jonas' words, "responsibility has become the fundamental

[2]The interested reader can find the letter at the link: `https://futureoflife.org/open-letter/pause-giant-ai-experiments/`.

[3]One of the most important international conference on AI.

[4]See `https://openai.com/index/introducing-superalignment/`.

[5]See the relative The Guardian's article at `https://www.theguardian.com/technology/2023/may/02/geoffrey-hinton-godfather-of-ai-quits-google-warns-dangers-of-machine-learning`.

[6]See, for instance, Robert Miles' video for examples and a more detailed discussion about AI concerns: `https://www.youtube.com/watch?v=2ziuPUeewK0`

[7]This term is usually used to refer to a false or misleading information presented as true in an argument by AI systems.

imperative in modern civilization, and it should be an unavoidable criterion to assess and evaluate human actions, including, in a special way, development activities."[8]

In this work, we decided to investigate inductive logic, since it appears in many research articles devoted to knowledge representation and AI: if we want to study the rationality of agents, we can't use only deductive logic. Indeed, it can be argued, that in most cases (if not all), we consider *rational* beliefs that aren't evident or totally supported by data. Instead of taking into account logical implications, an intelligent being retrieves information from the surroundings thanks to inductive reasoning.

As human beings, the knowledge we have about the world we live in depends on a summary that briefly represents the information received: we think of inductive logic as the theory that tries to outline in which way previous information should affect our beliefs. Using a classical example, after having seen only black ravens, we have a high degree of confidence to say that "all the ravens are black", even if no logical entailments can prove this.

Hence, if we want to mathematically prove the *reliability* of an agent, we can't take into account only the fact that all of its statements are logically confirmed by a deduction, but we need to formalize a notion of *degree of support.* This is the main core of Inductive Logic as presented in this work. Even if the approach generally used is different from the way that will be discussed here,[9] we think that this introduction can enlighten the question from a different point of view, challenging some opinions about what should be meant by the term rationality, and giving a mathematical clear framework to study the subject.

Notations

Throughout the thesis, we will use a standard notation for usual mathematical objects. However, since we will heavily use logical concepts and they are sometimes represented in different ways in literature, we want to outline clearly the meaning of some symbols.

In the following, we will consider languages $\mathcal{L}$ that are countable, i.e. with a countable supply of variables ($x_1, \ldots, x_n, \ldots$ or $y_1, \ldots, y_n, \ldots$ or simply x, y, z) and that can have at most a countable supply of constants ($a_1, \ldots, a_n, \ldots$ or simply a, b, c) predicate symbols ($R_1, \ldots R_q, \ldots$ or simply P, R)[10] and of function symbols ($f_1, \ldots, f_n, \ldots$ or simply $f, g, h, \ldots$)[11] The logic that we will use is the first-order classical one (FO) and the set of all sentences (i.e. formulas without free variables) that can be built with the boolean connectives ($\wedge, \vee, \neg, \rightarrow, \equiv$), the first-order quantifiers ($\exists, \forall$), and with the non-logical symbols of $\mathcal{L}$ will be denoted with $\mathrm{Sen}(\mathcal{L})$; the set $\mathrm{QFSen}(\mathcal{L})$ will be the set of all the quantifier-free formulas of $\mathrm{Sen}(\mathcal{L})$. Regarding substitutions, $\varphi\{x/t_i\}$ will also be denoted by $\varphi(t_i)$ when the variable which we are replaced is clear from the context or irrelevant.

We will also use some concepts or results from Model Theory: a generic structure will

[8]See [14].

[9]and this can be motivated by some limitations that will be explained at the end of the thesis, Chapter 3, Section 3.3.

[10]We will not consider languages with equality.

[11]Except for the very first results, we will consider languages without functional symbols. There have been attempts to generalize the discussion also to languages with functions and equality: further details can be seen in [11].

be denoted by gothic letters such as $\mathfrak{A}, \mathfrak{B}, \ldots$. A structure $\mathfrak{A}$ will be called a model of $\varphi \in \mathrm{Sen}(\mathcal{L})$ if φ holds in $\mathfrak{A}$ (denoted by $\mathfrak{A} \vDash \varphi$): often, following most of the literature, the terms "structure" and "model" are used with no difference in meaning, and we will say, for instance, "consider a model $\mathfrak{A}$"even if there is no reference to a sentence. Dealing with formulas φ that have free variables, a structure $\mathfrak{A}$ isn't enough to give meaning to a notion of validity in the model. We need a *state assignment* σ, that is a map from the set of variables in $\mathcal{L}$ to the domain of $\mathfrak{A}$: since the only important values to understand if $\mathfrak{A}, \sigma \vDash \varphi$ are the ones assumed on the free variables occurring in φ, sometimes the domain of σ is considered to be restricted to them. In addition, given $\varphi = \varphi(x_1, \ldots, x_n)$, where $x_1, \ldots, x_n$ are the variables that appear free in the formula at issue, and given $d_1, \ldots, d_n$ elements of the domain of $\mathfrak{A}$, the notation "$\mathfrak{A} \vDash \varphi[d_1, \ldots, d_n]$" will mean that given a state assignment σ[12] such that for any $i = 1, \ldots, n$ σ maps x_i to d_i, $\mathfrak{A}, \sigma \vDash \varphi$.

In general, when new or less common mathematical objects are introduced, we will discuss also about the notation used.

[12]By Model Theory results, it is equivalent to say that "for any state assignemnt σ ..."

Contents

Chapter 1

Probabilities on Sentences

The first intuition we have about a probability on the set $\mathrm{Sen}(\mathcal{L})$ of sentences of a language $\mathcal{L}$ is formalized by the following definition.

Definition 1.0.1. Given a language $\mathcal{L}$, a *probability* on $\mathcal{L}$ is a map $w : \mathrm{Sen}(\mathcal{L}) \to \mathbb{R}_{\geq 0}$ such that:

(a) for every valid φ, $w(\varphi) = 1$;

(b) if $\neg(\varphi \wedge \psi)$ is valid, then $w(\varphi \vee \psi) = w(\varphi) + w(\psi)$.

Example 1.0.1. We can consider a language $\mathcal{L}$ and a $\mathcal{L}$-structure $\mathfrak{A}$. We can define the probability $w_{\mathfrak{A}}$ that assigns 1 to all the sentences true in $\mathfrak{A}$ and 0 otherwise. This is a probability since:

- if φ is valid, then it is true in $\mathfrak{A}$ and, therefore, $w_{\mathfrak{A}}(\varphi) = 1$;
- let φ and ψ be in contradiction, i.e. $\neg(\varphi \wedge \psi)$ be valid. Then $\mathfrak{A} \vDash \neg(\varphi \wedge \psi)$ and thus in $\mathfrak{A}$ only one among φ and ψ can be true. We have two cases:
  - if $\varphi \vee \psi$ is true in $\mathfrak{A}$, then $w_{\mathfrak{A}}(\varphi \vee \psi) = 1$. Furthermore, at least one among φ and ψ is true in $\mathfrak{A}$ and, by the previous remark, exactly one of them is, from which the condition b) of Definition 1.0.1 is satisfied;
  - otherwise, $\varphi \vee \psi$ is not true in $\mathfrak{A}$ and so is φ and ψ. Therefore, $w_{\mathfrak{A}}(\varphi \vee \psi) = w_{\mathfrak{A}}(\varphi) = w_{\mathfrak{A}}(\psi) = 0$, hence the thesis.

The just defined $w_{\mathfrak{A}}$ is a truth-values valuation: therefore, this example shows that all these valuations are nothing but probabilities on $\mathrm{Sen}(\mathcal{L})$.

Example 1.0.2. Another example of probability is given by taking convex combinations of the probabilities in Example 1.0.1: given some $\mathcal{L}$-structures $\mathfrak{A}_1, \ldots, \mathfrak{A}_n$ and some non-negative reals $m_1, \ldots, m_n$ that sum to 1, we can define the probability

$$w = \sum_{i=1}^{n} m_i w_{\mathfrak{A}_i}.$$

We can generalize the example above in the case of a countable quantity of structures: we can choose some structures $\{\mathfrak{A}_i\}_{i \in \mathbb{N}}$, some non-negative weights $\{m_i\}_{i \in \mathbb{N}}$ such that

$\sum_{i\in\mathbb{N}} m_i = 1$ and define a new probability $w = \sum_{i\in\mathbb{N}} m_i w_{\mathfrak{A}_i}$. There is no need to take into account only probabilities of the form $w_{\mathfrak{A}_i}$: every convex combination (also countable) of probabilities is a probability.

Perhaps, the reader has already encountered such examples of probabilities in logic expressivity results. For instance, suppose that $\mathcal{L}$ is a finite language without functional symbols. Since[1], for fixed $n \in \mathbb{N}^+$, there are a finite number of structures of cardinality n up to isomorphism, we can define for all $n > 0$ the (finite) set $\mathcal{C}_n$ of all the $\mathcal{L}$-structures with n elements in the domain (up to isomorphisms). For every formula φ, let

$$w_n(\varphi) := \frac{|\{\mathfrak{A} \in \mathcal{C}_n : \mathfrak{A} \models \varphi\}|}{|\mathcal{C}_n|}.$$

These w_n's are probabilities in the sense of Definition 1.0.1 and they are linear combinations (with all the weights equal to $\frac{1}{|\mathcal{C}_n|}$) of probabilities of the form $w_{\mathfrak{A}}$ for an $\mathcal{L}$-structure $\mathfrak{A}$.

The limit as n approaches to infinity of these probabilities exists and it is also a probability, a trivial one, though: the $0-1$ Law for first-order logic states that for every formula $\varphi \in \mathrm{Sen}(\mathcal{L})$ the limit at issue exists (usually it is called *asymptotic probability*) and the possible outcomes are only 0 or 1.

Example 1.0.3. The many-valued propositional valuations G_k defined by Gödel are not examples of probabilities over sentences for $k > 2$. Even if they are usually defined only for quantifier-free formulas, we present an argument that shows that each of their extension to $\mathrm{Sen}(\mathcal{L})$ can't be a probability, since the violation of Definition 1.0.1 happens already for sentences in $\mathrm{QFSen}(\mathcal{L})$.
We will show it in the case $k = 3$, but it's easy to generalize the result to all the other cases. Let $v : \mathrm{QFSen}(\mathcal{L}) \to \{0, \frac{1}{2}, 1\}$ be the propositional valuation defined inductively on the height of the formulas following the rules

$$v(\varphi \wedge \psi) = \min\{v(\varphi), v(\psi)\};$$
$$v(\varphi \vee \psi) = \max\{v(\varphi), v(\psi)\};$$
$$v(\neg\varphi) = \begin{cases} 1 & \text{if } v(\varphi) = 0 \\ 0 & \text{otherwise} \end{cases};$$
$$v(\varphi \to \psi) = \begin{cases} 1 & \text{if } v(\varphi) \leq v(\psi) \\ v(\psi) & \text{otherwise} \end{cases}.$$

Considering the case in which v doesn't assume only the values $0, 1$, there is a formula φ for which $v(\varphi) = \frac{1}{2}$. Hence, $v(\neg\varphi) = 0$ and $v(\varphi \vee \neg\varphi) = \frac{1}{2}$, even if $\varphi \vee \neg\varphi$ is a tautology.[2]

The conditions in Definition 1.0.1 are sufficient to guarantee other properties of a probability, as the following proposition shows.

[1] The absence of functional symbols is necessary for the proof of the theorem we will discuss later in the example; we don't need this assumption to have only finitely many not isomorphic structures with a given cardinality, though.

[2] This result should not surprise the reader, since this kind of semantics was originally developed to study intuitionistic logic: the violation of Definition 1.0.1 we presented, indeed, relies on the fact that the principle of excluded middle is not a tautology according to these valuations.

Proposition 1.0.1. *Let w be a probability over a language $\mathcal{L}$. Then for every φ, ψ and $\varphi_i \in \mathrm{Sen}(\mathcal{L})$ the following hold:*

1) $w(\neg\varphi) = 1 - w(\varphi)$*;*

2) $w(\varphi) \in [0, 1]$*;*

3) if φ is unsatisfiable, then $w(\varphi) = 0$;

4) if $\varphi \to \psi$ is valid, then $w(\varphi) \leq w(\psi)$;

5) if $\varphi \equiv \psi$ is valid, then $w(\varphi) = w(\psi)$;

6) $w(\varphi \wedge \psi) \leq w(\psi)$ and the equality holds if φ is valid;

7) $w(\bigvee_{i=1}^{n} \varphi_i) \leq \sum_{i=1}^{n} w(\varphi_i)$ and the two terms are equal if the sentences $\varphi_1, \ldots, \varphi_n$ are pairwise contradictory sentences (i.e. for every distinct $i, j \in \{1, \ldots, n\}$, $\varphi_i \wedge \varphi_j$ is unsatisfiable);

8) $w(\varphi \vee \psi) + w(\varphi \wedge \psi) = w(\varphi) + w(\psi)$.

9) if $w(\psi) > 0$, the map $w(-|\psi) : \mathrm{Sen}\,\mathcal{L} \to \mathbb{R}_{\geq 0}$ defined by

$$w(\varphi|\psi) = \frac{w(\varphi \wedge \psi)}{w(\psi)},$$

is a probability.

Proof. 1) Since $\neg(\varphi \wedge \neg\varphi)$ is valid, then $w(\varphi \vee \neg\varphi) = w(\varphi) + w(\neg\varphi)$. The first term of the precedent equality is 1 because $\varphi \vee \neg\varphi$ is valid, then the thesis;

2) By Definition 1.0.1, $w(\varphi) \geq 0$ but also $w(\neg\varphi) \geq 0$. Thanks to 1), we have $w(\varphi) = 1 - w(\neg\varphi) \leq 1$;

3) It follows from 1) and the fact that the negation of an unsatisfiable formula is a valid one;

4) If $\varphi \to \psi$ is valid, then so $\psi \vee \neg\varphi$ is and therefore $\neg(\neg\psi \wedge \varphi)$ is. By b) of Definition 1.0.1,

$$w(\neg\psi \vee \varphi) = w(\neg\psi) + w(\varphi)$$

and therefore, using 1) and 2)

$$1 \geq w(\neg\psi \vee \varphi) = w(\neg\psi) + w(\varphi) = 1 - w(\psi) + w(\varphi),$$

from which the thesis.

5) It follows from 4)

6) By the excluded middle principle, ψ is equivalent to $(\psi \wedge \varphi) \vee (\psi \wedge \neg\varphi)$. Therefore by 5) and by the fact that $\psi \wedge \varphi$ is in contradiction with $\psi \wedge \neg\varphi$, we have

$$w(\psi) = w((\psi \wedge \varphi) \vee (\psi \wedge \neg\varphi)) = w(\psi \wedge \varphi) + w(\psi \wedge \neg\varphi) \geq w(\psi \wedge \varphi).$$

In the case in which φ is valid, then $\psi \wedge \neg\varphi$ is unsatisfiable, and by 3) the last disequality in the expression above is indeed an equality.

7) Assume first that $\varphi_1, \dots, \varphi_n$ are pairwise contradictory. Then, also $\varphi_1 \vee \varphi_2 \vee \dots \vee \varphi_{n-1}$ and φ_n are in contradiction and by b) of Definition 1.0.1,

$$w(\varphi_1 \vee \dots \vee \varphi_n) = w((\varphi_1 \vee \dots \vee \varphi_{n-1}) \vee \varphi_n) = w(\varphi_1 \vee \dots \vee \varphi_{n-1}) + w(\varphi_n).$$

Then, it is easy to see that the thesis follows by induction, in the case of pairwise contradictory sentences.

In the general case, we can notice that

$$\bigvee_{i=1}^{n} \varphi_i \equiv \bigvee_{i=1}^{n-1} \varphi_i \vee (\neg \bigvee_{i=1}^{n-1} \varphi_i \wedge \varphi_n)$$

and that $\bigvee_{i=1}^{n-1} \varphi_i$ and $(\neg \bigvee_{i=1}^{n-1} \varphi_i \wedge \varphi_n)$ are in contradiction. Therefore by 5), b) of Definition 1.0.1 and 6), we have

$$\begin{aligned} w(\bigvee_{i=1}^{n} \varphi_i) &= w(\bigvee_{i=1}^{n-1} \varphi_i \vee (\neg \bigvee_{i=1}^{n-1} \varphi_i \wedge \varphi_n)) = w(\bigvee_{i=1}^{n-1} \varphi_i) + w(\neg \bigvee_{i=1}^{n-1} \varphi_i \wedge \varphi_n)) \\ &\leq w(\bigvee_{i=1}^{n-1} \varphi_i) + w(\varphi_n). \end{aligned}$$

As in the previous case, it's easy to see now that a simple induction yields to the thesis.

8) We can notice that $\varphi \vee \psi$ is equivalent to the formula $\varphi \vee (\neg\varphi \wedge \psi)$, that has the two disjuncts in contradiction. Therefore

$$w(\varphi \vee \psi) = w(\varphi \vee (\neg\varphi \wedge \psi)) = w(\varphi) + w(\neg\varphi \wedge \psi). \tag{1.1}$$

Using that ψ is equivalent to $(\varphi \wedge \psi) \vee (\neg\varphi \wedge \psi)$ and that the disjuncts on this formula are in contradiction, we get

$$w(\psi) = w(\varphi \wedge \psi) + w(\neg\varphi \wedge \psi). \tag{1.2}$$

Combining Equation (1.1) and Equation (1.2) we have the thesis.

9) The condition a) of Definition 1.0.1 holds by 6). To show the validity of the condition b), let φ_1 and φ_2 two sentences that are in contradiction. Logically, $(\varphi_1 \vee \varphi_2) \wedge \psi$ is equivalent to $(\varphi_1 \wedge \psi) \vee (\varphi_2 \wedge \psi)$ and the two disjuncts are in contradiction by the hypothesis we are assuming about φ_1 and φ_2. Therefore,

$$w((\varphi_1 \vee \varphi_2) \wedge \psi) = w(\varphi_1 \wedge \psi) + w(\varphi_2 \wedge \psi),$$

and, then,

$$\begin{aligned} w(\varphi_1 \vee \varphi_2 | \psi) &= \frac{w((\varphi_1 \vee \varphi_2) \wedge \psi)}{w(\psi)} = \frac{w(\varphi_1 \wedge \psi) + w(\varphi_2 \wedge \psi)}{w(\psi)} \\ &= w(\varphi_1|\psi) + w(\varphi_2|\psi). \end{aligned}$$

□

Remark 1.0.1. Condition 3) of Proposition 1.0.1 is not invertible: it's possible to find a probability w and a satisfiable sentence φ such that $w(\varphi) = 0$. For instance, take a formula φ that is satisfiable but not valid and a model $\mathfrak{A}$ for $\neg\varphi$. Then, using the notation of Example 1.0.1, $w_{\mathfrak{A}}(\varphi) = 0$. With the same argument, we can show that it is also possible to find a probability w and a sentence φ that is not valid, but $w(\varphi) = 1$.

Remark 1.0.2. Condition 7) of Proposition 1.0.1 is not an equivalence: there can be a probability w on Sen($\mathcal{L}$) and two statements φ_1, φ_2 not in contradiction with $w(\varphi_1 \vee \varphi_2) = w(\varphi_1) + w(\varphi_2)$. For instance, take two formulas φ_1 and φ_2 such that $\varphi_1 \wedge \varphi_2$ is satisfiable but not valid and a model $\mathfrak{A}$ such that

$$\mathfrak{A} \not\models \varphi_1 \wedge \varphi_2 \quad \mathfrak{A} \models \varphi_1 \quad \mathfrak{A} \not\models \varphi_2.$$

Then,

$$w_{\mathfrak{A}}(\varphi_1 \vee \varphi_2) = 1 = w_{\mathfrak{A}}(\varphi_1) + w_{\mathfrak{A}}(\varphi_2),$$

even if φ_1 and φ_2 are not in contradiction.

Thanks to 9) of Proposition 1.0.1, we can define, whenever $w(\psi) > 0$, the *conditional probability* $w(-|\psi)$ as pointed out above.

Proposition 1.0.1 explains why the maps in Definition 1.0.1 are called probabilities and not simply "measures": indeed, they can be regarded as maps from Sen($\mathcal{L}$) to $[0, 1]$ and not to all $\mathbb{R}_{\geq 0}$. Even if some similarities with the usual notion of "probability" can be detected, the reader with a more measure-theoretic approach will however have some trouble with the use of this term for such functions: Section 1.2 is devoted to the connection between these, for now different, concepts.

1.1 Gaifman condition

Even if in the Definition 1.0.1 only a few properties mark a probability out, we showed in Proposition 1.0.1 that they are sufficient to prove some of the features we want for such maps. By now, however, we don't have any control over formulas with quantifiers; the following definition tries to point out a meaningful property in this sense.

Definition 1.1.1. Let $\mathcal{L}$ a countable language and $\{t_i\}_{i\in\mathbb{N}}$ an enumeration[3] of all closed terms in $\mathcal{L}$. A probability w on Sen($\mathcal{L}$) satisfies the *Gaifman condition* (we say that w is *Gaifman*) if for every φ with one and only one free variable x, we have:

$$w(\exists x\varphi) = \sup\{w(\bigvee_{i\in I} \varphi\{x/t_i\}) \ : \ I \text{ finite subset of } \mathbb{N}\}. \tag{1.3}$$

Proposition 1.1.1. *With the notation of the Definition 1.1.1, the following are equivalent:*

i) w is Gaifman;

ii)

$$w(\exists x\varphi) = \lim_{n\to+\infty} w(\bigvee_{i\leq n} \varphi\{x/t_i\}) \tag{1.4}$$

[3]In Equation (1.3), there is no reference to the order used to list the closed terms of the language: it is easy to see, indeed, that the definition doesn't depend on the chosen enumeration.

iii)

$$w(\forall x\varphi) = \inf\{w(\bigwedge_{i\in I} \varphi\{x/t_i\}) \ : \ I \textit{ finite subset of } \mathbb{N}\} \tag{1.5}$$

iv)

$$w(\forall x\varphi) = \lim_{n\to+\infty} w(\bigwedge_{i\leq n} \varphi\{x/t_i\}) \tag{1.6}$$

Proof. Let s be the right-hand-side of Equation (1.3) and l the limit that appears in item ii) of this proposition: we will prove that $s = l$. Combining 8) and 6) of Proposition 1.0.1 we have for all $\varphi_1, \varphi_2 \in \mathrm{Sen}(\mathcal{L})$

$$w(\varphi_1 \vee \varphi_2) \geq w(\varphi_1),$$

and by induction we have that the succession $\{w(\bigvee_{i\leq n} \varphi\{x/t_i\})\}_{n\in\mathbb{N}}$ is increasing, therefore the limit exists and

$$l := \lim_{n\to+\infty} w(\bigvee_{i\leq n} \varphi\{x/t_i\}) = \sup_{n\in\mathbb{N}} w(\bigvee_{i\leq n} \varphi\{x/t_i\}).$$

Since the sets of the form $\{0, 1, \ldots, n\}$ form a subset of all the finite subsets I of $\mathbb{N}$, $s \geq l$. Now, for an arbitrary ε, from the definition of sup, we have a finite $I \subset \mathbb{N}$ such that $s - w(\bigvee_{i\in I} \varphi\{x/t_i\}) \leq \varepsilon$. If n is the maximum natural number in I, then $w(\bigvee_{i\in I} \varphi\{x/t_i\}) \leq w(\bigvee_{i\leq n} \varphi\{x/t_i\})$ and

$$s - w(\bigvee_{i\leq n} \varphi\{x/t_i\}) \leq \varepsilon.$$

Therefore, as n approaches $+\infty$ we get

$$s - l \leq \epsilon$$

and, since ε is arbitrary, $s \leq l$. Hence, $s = l$, and from this we get that *i)* is equivalent to *ii)*.

Analogously, *iii)* is equivalent to *iv)*.

Suppose now that w is Gaifman. Using 1) of Proposition 1.0.1, De Morgan's laws and properties of sup,

$$\begin{aligned}
w(\forall x\varphi) &= w(\neg\exists x\neg\varphi) = 1 - w(\exists x\neg\varphi) \\
&\overset{\star}{=} 1 - \sup\{w(\bigvee_{i\in I} \neg\varphi\{x/t_i\}) \ : \ I \text{ finite subset of } \mathbb{N}\} \\
&= 1 - \sup\{w(\neg\bigwedge_{i\in I} \varphi\{x/t_i\}) \ : \ I \text{ finite subset of } \mathbb{N}\} \\
&= 1 - \sup\{1 - w(\bigwedge_{i\in I} \varphi\{x/t_i\}) \ : \ I \text{ finite subset of } \mathbb{N}\} \\
&= 1 - (1 - \inf\{w(\bigwedge_{i\in I} \varphi\{x/t_i\}) \ : \ I \text{ finite subset of } \mathbb{N}\}) \\
&= \inf\{w(\bigwedge_{i\in I} \varphi\{x/t_i\}) \ : \ I \text{ finite subset of } \mathbb{N}\}),
\end{aligned}$$

where $\stackrel{\star}{=}$ points out where the Gaifman condition is used.

In a very similar way it can be shown that *iii)* implies the Gaifman condition, hence also *i)* and *iii)* are equivalent, from which the thesis follows. □

We now give some examples (similar to the ones in Examples 1.0.1 and 1.0.2) of Gaifman probabilities.

Example 1.1.1. Given a countable language $\mathcal{L}$, a model $\mathfrak{A}$ is called *separating* if for every formula φ, with x as the only free variable,

$$\mathfrak{A} \models \exists x\varphi \text{ if and only if exists a closed term } t \text{ s.t. } \mathfrak{A} \models \varphi\{x/t\}.$$

In the following $\{t_i\}_{i\in\mathbb{N}}$ will be an enumeration of all the closed terms in $\mathcal{L}$.

If $\mathfrak{A}$ is a separating model, then $w_{\mathfrak{A}}$ is a Gaifman probability. Indeed, it is a probability by Example 1.0.1; the Gaifman condition holds because if φ is a formula with only x as free variable, then we have two cases:

- if $\mathfrak{A} \not\models \exists x\varphi$, then $w_{\mathfrak{A}}(\exists x\varphi) = 0$ and also $w_{\mathfrak{A}}(\bigvee_{i\leq n}\varphi(t_i)) = 0$ for any $n \in \mathbb{N}$, since $\mathfrak{A} \not\models \varphi(t)$ for any closed term t.
- if $\mathfrak{A} \models \exists x\varphi$, then $w_{\mathfrak{A}}(\exists x\varphi) = 1$. Since $\mathfrak{A}$ is separating, there exists a closed term of the language $t = t_i$ such that $\varphi(t_i)$ holds in $\mathfrak{A}$. Therefore, $w_{\mathfrak{A}}(\varphi(t_i)) = 1$, and the Gaifman condition holds.

Following the lines of Example 1.0.2, Gaifman probabilities are closed under (countable) convex combination. Indeed, if $\{w_j\}_{j\in\mathbb{N}}$ is a set of Gaifman probabilities, and $\{m_j\}_{j\in\mathbb{N}}$ a set of non-negative weights that sum up to 1, we have that $w = \sum_{j\in\mathbb{N}} m_j w_j$ is a probability. It is also Gaifman since

$$\begin{aligned} w(\exists x\varphi) &= \sum_{j\in\mathbb{N}} m_j w_j(\exists x\varphi) \\ &= \sum_{j\in\mathbb{N}} m_j \lim_{n\to+\infty} w_j(\bigvee_{i\leq n}\varphi(t_i)) \\ &\stackrel{*}{=} \lim_{n\to+\infty} \sum_{j\in\mathbb{N}} m_j w_j(\bigvee_{i\leq n}\varphi(t_i)) \\ &= \lim_{n\to+\infty} w(\bigvee_{i\leq n}\varphi(t_i)), \end{aligned}$$

where in $\stackrel{*}{=}$ we use the Dominate Convergence Theorem (see Appendix A).

Let's contextualize better the definition of the separating model in the previous example with some remarks.

Remark 1.1.1. If $\mathfrak{A}$ is a model in which all the elements of the domain are interpretations of closed terms of the language, then $\mathfrak{A}$ is also separating: this is the case, for instance, of the standard model of arithmetic $\mathbb{N}$ when we consider the language $\mathcal{L} = \{0, S\}$, with S a functional symbol that represent the successor.

However, there are separating models in which not all the elements of the domain are interpretations of closed terms. Consider the language $\mathcal{L} = \{c, P\}$ with a constant symbol c and a unary predicate P, and the structure $\mathfrak{A}$ whose domain is $A = \{-1, 1\}$, the interpretation of c is -1 and $P^{\mathfrak{A}}$ is the emptyset. By induction on the height we can show that for any formula φ and for any variable assignment σ,

$$\mathfrak{A}, \sigma \vDash \varphi \text{ if and only if } \mathfrak{A}, -\sigma \vDash \varphi,$$ [4]

where $-\sigma$ is σ post-composed with the involution i of $\{-1, 1\}$ that maps an element to its additive opposite:

- if φ is atomic, then we have two possible cases:
  - $\varphi = P(c)$ and in this case the thesis holds since for any assignment σ, $\mathfrak{A}, \sigma \vDash P(c)$ is equivalent to $\mathfrak{A} \vDash P(c)$;
  - $\varphi = P(x)$ and in this case both $P[-1]$ and $P[1]$ are false in $\mathfrak{A}$ and the equivalence at issue holds.
- if $\varphi = \psi_1 \wedge \psi_2$, then by inductive hypothesis,

$$\begin{aligned} \mathfrak{A}, \sigma \vDash \psi_1 \wedge \psi_2 &\text{ if and only if } \mathfrak{A}, \sigma \vDash \psi_1 \text{ and } \mathfrak{A}, \sigma \vDash \psi_2 \\ &\text{ if and only if } \mathfrak{A}, -\sigma \vDash \psi_1 \text{ and } \mathfrak{A}, -\sigma \vDash \psi_2 \\ &\text{ if and only if } \mathfrak{A}, -\sigma \vDash \psi_1 \wedge \psi_2; \end{aligned}$$

- if $\varphi = \neg\psi$, then by inductive hypothesis

$$\begin{aligned} \mathfrak{A}, \sigma \vDash \neg\psi &\text{ if and only if } \mathfrak{A}, \sigma \nvDash \psi \\ &\text{ if and only if } \mathfrak{A}, -\sigma \nvDash \psi \\ &\text{ if and only if } \mathfrak{A}, -\sigma \vDash \neg\psi; \end{aligned}$$

- if $\varphi = \exists x \psi(x, \vec{y})$, then

$$\begin{aligned} \mathfrak{A}, \sigma \vDash \exists x \psi(x, \vec{y}) &\text{ if and only if there exists } \gamma \text{ extending } \sigma \text{ such that } \mathfrak{A}, \gamma \vDash \psi(x, \vec{y}) \\ &\text{ if and only if there exists } \gamma \text{ extending } \sigma \text{ such that } \mathfrak{A}, -\gamma \vDash \psi(x, \vec{y}) \\ &\text{ if and only if } \mathfrak{A}, -\sigma \vDash \exists x \psi(x, \vec{y}). \end{aligned}$$

Thanks to this, we are now ready to show that if $\mathfrak{A} \vDash \exists x\varphi$, then $\mathfrak{A} \vDash \varphi(c)$: indeed, if $\exists x\varphi$ holds, there is an assignment σ such that $\mathfrak{A}, \sigma \vDash \varphi$; for what shown before, we have also that $\mathfrak{A}, -\sigma \vDash \varphi$. Therefore, if $\sigma(x) = -1$, from $\mathfrak{A}, \sigma \vDash \varphi$ follows that $\mathfrak{A} \vDash \varphi(c)$; otherwise, if $\sigma(x) = 1$, from $\mathfrak{A}, -\sigma \vDash \varphi$ follows that $\mathfrak{A} \vDash \varphi(c)$, hence the thesis.

Remark 1.1.2. There exist satisfiable sentences that don't have any separating model. For instance, in the language $\mathcal{L} = \{c, P\}$, where c is a constant symbol and P is a unary predicate, the only closed term is c. Therefore, the formula $\exists x P(x) \wedge \neg P(c)$ is satisfiable but not in a separating model since $P(c)$ and $\neg P(c)$ cannot be both true in a structure.

[4]As already mentioned in the introduction, we use here the following notation: when we write $\mathfrak{A}, \sigma \vDash \varphi$, we suppose that σ refers only to the variables that occur free in φ, if not otherwise written. This is motivated by the fact that for an assignment γ whose domain is the whole set of variables, it is sufficient to know its restriction to the free variables occurring in φ to determine whether $\mathfrak{A}, \gamma \vDash \varphi$ or not.

Before carrying on the study of the Gaifman property, we want to discuss a bit further its meaning. Indeed, using Proposition 1.1.1, the information about the value of $\exists x\varphi$ or of $\forall x\varphi$ depends only on the values of $\varphi\{x/t_i\}$ when t_i is a closed term. In some settings, this can be useful, and in others annoying; we will now give two examples.

First, let's consider having a coin that is tossed a countable number of times and you want to study some property of the tosses sequence. Then you can introduce, in the arithmetical language $\mathcal{L}_{\text{PA}} = \{0, S, +, \cdot\}$, some "empirical symbols" (see [7]) that refers to some accidental properties of the tosses: for instance, unary predicates $H(x)$ and $T(x)$ such that $H(n)$ is interpreted as "the n-th toss is head" and $T(n)$ as "the n-th toss is tail". If we want to describe in a more precise logical way what we are studying, we are considering the standard model of arithmetic $\mathbb{N}$ endowed with some interpretation of the empirical symbols we introduced. In this setting, the conditions in the Definition 1.0.1 make sense and also the Gaifman one: indeed, the probability we should assign to $\exists x H(x)$ should be related to the values of all the $H(n)$ where $n \in \mathbb{N}$ and here all the elements we are interested in are interpretations of closed terms of the language (the element n in the domain $\mathbb{N}$ is the interpretation of $S^n(0)$, a closed term of $\mathcal{L}_{\text{PA}}$). Actually, we can formalize this also in another way, different from the original formalization of the problem in [7], but more similar to what the reader will find written in this work. Instead of using $\mathcal{L}_{\text{PA}}$, we can perform all the arguments above, by simply taking into account a language $\mathcal{L}$ with a countable supply of constants $\{a_n \ : \ n \in \mathbb{N}^+\}$ and the two unary relational symbols H and T.

However, in general, there are cases in which the closed terms of a language don't exhaust the whole universe we are studying, i.e. it is possible to have elements of a domain that can't be expressed as interpretations of closed terms. For example, if we want to study $\mathbb{R}$ with a countable language, obviously only countable elements can be described, hence, not all. Here the Gaifman condition assumes another flavor and can seem unreasonable.

1.2 Measure-theoretic approach

This section aims to understand if there is a link between the notion of probability on sentences and the one defined in the measure-theoretic setting (see Appendix A). To do so, we first need to detect a set Ω and an algebra/σ-algebra in the logical context.

The first idea that comes to mind is to consider the class $\text{Mod}(\mathcal{L})$ of all models for a language $\mathcal{L}$, endowed with a σ-algebra structure, and put a measure on it.

A careful reader may have noticed that we used the world *class* to denote $\text{Mod}(\mathcal{L})$: indeed, it is not a set. Hence, to formalize a similar argument, we should put some restrictions on the models that we will consider. The solution of taking into account only models with a fixed domain was given in [7], where the authors were concerned primarily with the setting of standard natural numbers (so the chosen domain was $\mathbb{N}$).

Analogously, we can restrict ourselves to languages with a countable supply of constants $\{a_i\}_{i\in\mathbb{N}^+}$, that are intended to represent all the elements of the models we want to consider.[5] In this setting, the Gaifman condition makes sense and it will come out

[5] The underlying assumption here is that many applications can be performed in a context in which we already know which are the objects but we don't know the validity of some other features (that we

naturally (see Section 1.1). A further assumption that we will make is that the language has no functional symbols: in Theorem 1.2.1, this hypothesis will be used.[6] From now on, if not otherwise specified, when we talk about a language $\mathcal{L}$, we assume it has a countable supply of constants $\{a_i\}_{i\in\mathbb{N}^+}$, and q relational symbols $R_1, \dots, R_q$ with ariety $r_1, \dots, r_q$.

Then, we can take into account all the structures with the set $\mathrm{Con} = \{a_1, a_2, \dots\}$ as the given domain and the natural interpretation of constant symbols, i.e. the symbol a_i is interpreted as a_i itself. These form a set we denote $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$; its subset of all the models of a given sentence φ is called $\mathrm{Mod}_{\mathrm{Con}}(\varphi)$. When dealing with this, the following remark will be very useful.

Remark 1.2.1. We can notice that if $\varphi \in \mathrm{Sen}(\mathcal{L})$ is a satisfiable sentence, then it has a model in $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$.
Indeed, by definition, a satisfiable formula has a model $\mathfrak{A} \in \mathrm{Mod}(\mathcal{L})$ and we can suppose without loss of generality that:

- all the constant symbols are interpreted in different elements of the domain: given $\mathfrak{A}$ we can create the structure $\mathfrak{A}'$ adding to the domain of $\mathfrak{A}$ an element for any constant symbol that is interpreted in an element that is already an interpretation of another constant symbol. To be more clear and precise, $\mathfrak{A}'$ is the structure with:
  - as domain the set
  $$\{a \in A \ : \ \exists! i \in \mathbb{N}^+ \ a = (a_i)^{\mathfrak{A}}\} \cup \{c_j \ : \ \exists i \in \mathbb{N}^+ \ i \neq j \wedge (a_i)^{\mathfrak{A}} = (a_j)^{\mathfrak{A}}\},$$
  where A is the domain of $\mathfrak{A}$ and $c_1, c_2, \dots$ are new elements not in A;
  - for $i \in \mathbb{N}^+$, if a_i is the only constant symbol interpreted in an element $a \in A$, then $(a_i)^{\mathfrak{A}'} = a$; otherwise, $(a_i)^{\mathfrak{A}'} = c_i$;
  - for any $i = 1, \dots, q$,
  $$R_i^{\mathfrak{A}'} = \{(a_{i_1}^{\mathfrak{A}'}, a_{i_2}^{\mathfrak{A}'}, \dots, a_{i_{r_k}}^{\mathfrak{A}'}) \ : \ \mathfrak{A} \vDash R_i(a_{i_1}, \dots, a_{i_{r_k}})\}.$$

  By induction on the complexity of the formula, we can then prove that $\mathfrak{A}' \vDash \varphi$ and in $\mathfrak{A}'$ different constants are interpreted in different elements.

- it is countable: indeed, by the downward Löwenheim-Skolem Theorem, since the language $\mathcal{L}$ is countable, any satisfiable formula has a model which is at most countable. Furthermore, by the previous point, we may suppose that all the constant symbols are interpreted with different elements of the domain, hence $\mathfrak{A}$ can't be finite.

- every element in the domain is the interpretation of exactly a constant symbol in $\mathcal{L}$: for the last two items, we can suppose that the domain of $\mathfrak{A}$ is the set $\{d_j \ : \ j \in \mathbb{N}^+\}$ and the constant symbols are interpreted in different elements. Assuming that $a_{i_1}, \dots, a_{i_k}$ are the constants that appear in φ, we can consider the following structure $\mathfrak{B}$, which differs from $\mathfrak{A}$ only for the interpretations of constants:
  - for any constant that appears in φ the interpretation is the same, i.e. for $s = 1, \dots, k$ $(a_{i_s})^{\mathfrak{A}} = (a_{i_s})^{\mathfrak{B}}$;

will study inductively): this is, indeed, the background in which Gaifman and Snir worked in [7].

[6] The case of languages with functional symbols is studied in [11], where an analogous of De Finetti's representation (see Theorem 2.2.1 and Theorem 2.4.4) is proved.

– if a_j is a constant symbol that doesn't appear in φ, then its interpretation is the first element d_h of the domain that isn't an interpretation of a constant symbol, i.e. $h = \min\{i \in \mathbb{N}^+ \ : \ \forall l \in \{1, \ldots, j-1\} \cup \{i_1, \ldots, i_k\} \ (a_l)^{\mathfrak{B}} \neq d_i\}$.

The structures $\mathfrak{A}$ and $\mathfrak{B}$ have the same domain and differ only for the interpretation of constant symbols that don't occur in φ: therefore, φ holds in $\mathfrak{B}$ since it is satisfied in $\mathfrak{A}$.

Changing the name of the elements of the domain of the structure $\mathfrak{B}$ of the previous item, we get the thesis.
Hence, for this kind of language, when we want to deal with satisfiability, we can simply restrict ourselves to $\text{Mod}_{\text{Con}}(\mathcal{L})$. Actually, this has some other important consequences:

- a sentence φ is valid if and only if every structure in $\text{Mod}_{\text{Con}}(\mathcal{L})$ is a model for the formula;
- two sentences φ and ψ are equivalent if and only if in $\text{Mod}_{\text{Con}}(\varphi) = \text{Mod}_{\text{Con}}(\psi)$.

We can endow $\text{Mod}_{\text{Con}}(\mathcal{L})$ with a topological structure $\mathcal{T}$, defining as a basis

$$B(\mathcal{L}) = \{\text{Mod}_{\text{Con}}(\varphi) : \varphi \in \text{Sen}(\mathcal{L})\}.$$

In this way, an open set is a (countable, since $\mathcal{L}$ is countable and so it is $\text{Sen}(\mathcal{L})$) union of elements in $B(\mathcal{L})$ and so we have a practical description of the elements in the topology $\mathcal{T}$. The elements in $B(\mathcal{L})$ are not only open but also closed sets of the topology: this is because $\text{Mod}_{\text{Con}}(\varphi)^c = \text{Mod}_{\text{Con}}(\neg\varphi)$.

As usual, denoting with $\mathcal{F}(A)$ the smallest σ-algebra that contains the elements in A, we can define from $\mathcal{T}$ the borel σ-algebra $\mathcal{F}(\mathcal{T})$.

Actually, it can be shown, using that $\mathcal{L}$ is countable, that $\mathcal{F}(\mathcal{T}) = \mathcal{F}(B(\mathcal{L}))$. Indeed, since $B(\mathcal{L}) \subseteq \mathcal{T}$, surely $\mathcal{F}(\mathcal{T})$ is a σ-algebra that contains $B(\mathcal{L})$; furthermore, it is the smallest one, because if we have a σ-algebra $\mathcal{C}$ such that $B(\mathcal{L}) \subseteq \mathcal{C}$, then $\mathcal{C}$ contains also $\mathcal{T}$ because it contains all the countable union of elements in $B(\mathcal{L})$ and every open set in $\mathcal{T}$ is such a union (here the countability of $\mathcal{L}$, and hence of $B(\mathcal{L})$ is used).

We can further simplify the description of $\mathcal{F}(\mathcal{T})$ noticing that it is also the σ-algebra generated by

$$\mathcal{A}_{\text{Con}} := \{\text{Mod}_{\text{Con}}(\psi) : \psi \in \text{QFSen}(\mathcal{L})\}.$$

$\mathcal{F}(\mathcal{A}_{\text{Con}})$ contains all the sets of the form $\text{Mod}_{\text{Con}}(\varphi)$ with $\varphi \in \text{Sen}(\mathcal{L})$. Indeed, this follows by the definition if φ is without quantifiers; otherwise, we can reason by induction on the rank of the sentence noticing that

$$\text{Mod}_{\text{Con}}(\exists x\psi) = \bigcup_{i \in \mathbb{N}^+} \text{Mod}_{\text{Con}}(\psi(a_i))$$

and recalling that a σ-algebra is closed under countable unions. Hence, $\mathcal{F}(\mathcal{A}_{\text{Con}}) = \mathcal{F}(B(\mathcal{L}))$ and we have already proved that the latter is $\mathcal{F}(\mathcal{T})$.

The last remark will be useful in the following because $\mathcal{A}_{\text{Con}}$ (and also $B(\mathcal{L})$) has a nice measure-theoretic property: it is an algebra over $\text{Mod}_{\text{Con}}(\mathcal{L})$. Indeed:

- it is closed under finite intersection since, for $\varphi \in \mathrm{QFSen}(\mathcal{L})$,

$$\begin{aligned}\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) &= \mathrm{Mod}_{\mathrm{Con}}(\varphi \to \varphi)\\ \mathrm{Mod}_{\mathrm{Con}}(\varphi_1) \cap \cdots \cap \mathrm{Mod}_{\mathrm{Con}}(\varphi_n) &= \mathrm{Mod}_{\mathrm{Con}}(\varphi_1 \wedge \cdots \wedge \varphi_n);\end{aligned}$$

- it is closed under finite union since, for $\varphi \in \mathrm{QFSen}(\mathcal{L})$,

$$\begin{aligned}\emptyset &= \mathrm{Mod}_{\mathrm{Con}}(\varphi \wedge \neg\varphi)\\ \mathrm{Mod}_{\mathrm{Con}}(\varphi_1) \cup \cdots \cup \mathrm{Mod}_{\mathrm{Con}}(\varphi_n) &= \mathrm{Mod}_{\mathrm{Con}}(\varphi_1 \vee \cdots \vee \varphi_n);\end{aligned}$$

- it is closed under complementation since

$$\mathrm{Mod}_{\mathrm{Con}}(\varphi)^c = \mathrm{Mod}_{\mathrm{Con}}(\neg\varphi).$$

Hence, to sum up, we consider $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, an algebra $\mathcal{A}_{\mathrm{Con}}$ on it and the σ-algebra $\mathcal{F}(\mathcal{T}) = \mathcal{F}(\mathcal{A}_{\mathrm{Con}})$ generated by this: in the following, the latter will be called the $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ borel σ-algebra. The convenience of having described this σ-algebra as generated by an algebra (and not only by a topology) relies on the possibility to extend a probability defined on $\mathcal{A}_{\mathrm{Con}}$ to one with $\mathcal{F}(\mathcal{T})$ as the domain, using the Extension Theorem A.1.1. We will use this argument in the following theorem.

Theorem 1.2.1. *Let $w^- : \mathrm{QFSen}(\mathcal{L}) \to [0,1]$ be a function that satisfies (a) and (b) of Definition 1.0.1 (where these conditions are relativized to quantifier-free sentences). Then, there is a unique extension to a Gaifman probability $w : \mathrm{Sen}(\mathcal{L}) \to [0,1]$. Furthermore, there exists a probability μ on the σ-algebra $\mathcal{F}(\mathcal{A}_{\mathrm{Con}})$ such that for any $\varphi \in \mathrm{Sen}(\mathcal{L})$,*

$$w(\varphi) = \mu(\mathrm{Mod}_{\mathrm{Con}}(\varphi)).$$

Proof. Define the map $\mu_{w^-} : \mathcal{A}_{\mathrm{Con}} \to [0,1]$ by

$$\mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) = w^-(\varphi),$$

for $\varphi \in \mathrm{QFSen}(\mathcal{L})$. It is well-defined since if we have two sentences φ and ψ with $\mathrm{Mod}_{\mathrm{Con}}(\varphi) = \mathrm{Mod}_{\mathrm{Con}}(\psi)$, by Remark 1.2.1, they are equivalent and $w^-(\varphi) = w^-(\psi)$.[7] In addition, this map is:

- non-trivial, since w^- satisfies conditions (a) of Definition 1.0.1 and, 3) of Proposition 1.0.1 and there are valid sentences and unsatisfiable ones without quantifiers;

- finitely-additive. indeed, if $\mathrm{Mod}_{\mathrm{Con}}(\varphi)$ and $\mathrm{Mod}_{\mathrm{Con}}(\psi)$ are disjoint, then φ and ψ are in contradiction, by Remark 1.2.1. Therefore,

$$\begin{aligned}\mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\varphi) \cup \mathrm{Mod}_{\mathrm{Con}}(\psi)) &= \mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\varphi \vee \psi)) = w^-(\varphi \vee \psi)\\ &= w^-(\varphi) + w^-(\psi)\\ &= \mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) + \mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\psi));\end{aligned}$$

[7] Here, we are using Proposition 1.0.1: the proof of the fact can be carried out also for maps as w^- defined only on the quantifier-free sentences.

- conditionally σ-additive: indeed, if we have quantifier-free sentences $\{\varphi_i\}_{i\in\mathbb{N}}$ such that the $\mathrm{Mod}_{\mathrm{Con}}(\varphi_i)$'s are pairwise disjoint and the union $\bigcup_{i\in\mathbb{N}}\mathrm{Mod}_{\mathrm{Con}}(\varphi_i)$ is an element of $\mathcal{A}_{\mathrm{Con}}$, then it means that there is $\psi\in\mathrm{QFSen}(\mathcal{L})$ with

$$\bigcup_{i\in\mathbb{N}}\mathrm{Mod}_{\mathrm{Con}}(\varphi_i)=\mathrm{Mod}_{\mathrm{Con}}(\psi). \tag{1.7}$$

Then, we can show that it must exist $n\in\mathbb{N}$ such that

$$\mathrm{Mod}_{\mathrm{Con}}(\psi)=\bigcup_{i\le n}\mathrm{Mod}_{\mathrm{Con}}(\varphi_i). \tag{1.8}$$

If this is not the case, for any $n\in\mathbb{N}$, there is a model $\mathfrak{A}\in\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ that satisfies ψ and $\neg\varphi_i$ for every $i\le n$.

Hence, the set

$$B=\{\neg\varphi_i : i\in\mathbb{N}\}\cup\{\psi\}\cup\{a_i\neq a_j : i,j\in\mathbb{N}^+ \text{ with } i\neq j\}$$

is finitely satisfiable and, then, by Compactness Theorem, there is a model $\mathfrak{B}$ that satisfies all the sentences in B. Since for different $i,j\in\mathbb{N}^+$, $\mathfrak{B}\vDash a_i\neq a_j$, the constant symbols are interpreted in different elements of the domain; therefore, we can consider the substructure $\mathfrak{C}$ of $\mathfrak{B}$ whose domain is the set of all the interpretations of such constant symbols.[8] By induction on the sentence height, it is easy to show that for every $\varphi\in\mathrm{QFSen}(\mathcal{L})$,

$$\mathfrak{B}\vDash\varphi \text{ if and only if } \mathfrak{C}\vDash\varphi.$$

Then $\mathfrak{C}$ is a model in $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ that satisfies ψ but not the φ_i's when i varies in $\mathbb{N}$, contradicting Equation (1.7). Therefore, there is $n\in\mathbb{N}$ for which Equation (1.8) holds.[9]

By the already shown finitely-additivity of μ_{w^-}, we get

$$\begin{aligned}\mu_{w^-}(\bigcup_{i\in\mathbb{N}}\mathrm{Mod}_{\mathrm{Con}}(\varphi_i)) &= \mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\psi))\\ &= \mu_{w^-}(\bigcup_{i\le n}\mathrm{Mod}_{\mathrm{Con}}(\varphi_i))\\ &= \sum_{i\le n}\mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\varphi_i)).\end{aligned} \tag{1.9}$$

By Equation (1.8) and by the fact that the sets $\mathrm{Mod}_{\mathrm{Con}}(\varphi_i)$'s are pairwise disjoint, for any $m>n$, $\mathrm{Mod}_{\mathrm{Con}}(\varphi_m)=\emptyset$, hence φ_m is a contradiction and

$$\mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\varphi_m))=w^-(\bot)=0;$$

Hence, combining this result with Equation (1.9) we get the thesis:

$$\mu_{w^-}(\bigcup_{i\in\mathbb{N}}\mathrm{Mod}_{\mathrm{Con}}(\varphi_i))=\cdots=\sum_{i\le n}\mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\varphi_i))=\sum_{i\in\mathbb{N}}\mu_{w^-}(\mathrm{Mod}_{\mathrm{Con}}(\varphi_i)).$$

[8]There is a problem if functional symbols appear in the language $\mathcal{L}$. In this case, if $f(a_i)$ is not a constant, the argument above doesn't generalize anymore.

[9]We couldn't use Remark 1.2.1 to justify the presence of such a model $\mathfrak{C}$: indeed, the remark works when we deal with one satisfiable sentence (or finitely many), but not with a countable quantity. For instance, assume that the language contains a unary predicate symbol P, $\varphi_i := P(a_i)$ and $\psi=\exists x\neg P(x)$: therefore, the infinite set $\{\varphi_i\}_{i\in\mathbb{N}^+}\cup\{\psi\}$ is satisfiable but it doesn't hold in any model of $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$.

We can now apply Theorem A.1.1, to show that μ_{w^-} can be extended to a probability μ on the σ-algebra $\mathcal{F}(\mathcal{A}_{\text{Con}})$ of $\text{Mod}_{\text{Con}}(\mathcal{L})$.

As already mentioned before the statement of this theorem, in this σ-algebra we can find all the sets of the form $\text{Mod}_{\text{Con}}(\varphi)$ with $\varphi \in \text{Sen}(\mathcal{L})$. Then, we can define

$$w(\varphi) := \mu(\text{Mod}_{\text{Con}}(\varphi)) :$$

clearly, w is an extension of w^- and it's a Gaifman probability on $\text{Sen}(\mathcal{L})$. Indeed:

a) if φ is valid, then $\text{Mod}_{\text{Con}}(\varphi) = \text{Mod}_{\text{Con}}(\mathcal{L})$ and therefore

$$w(\varphi) = \mu(\text{Mod}_{\text{Con}}(\varphi)) = \mu(\text{Mod}_{\text{Con}}(\mathcal{L})) = 1.$$

b) if φ and ψ are contradictory, then $\text{Mod}_{\text{Con}}(\varphi) \cap \text{Mod}_{\text{Con}}(\psi) = \emptyset$. Thus,

$$\begin{aligned} w(\varphi \vee \psi) &= \mu(\text{Mod}_{\text{Con}}(\varphi \vee \psi)) = \mu(\text{Mod}_{\text{Con}}(\varphi) \cup \text{Mod}_{\text{Con}}(\psi)) \\ &= \mu(\text{Mod}_{\text{Con}}(\varphi)) + \mu(\text{Mod}_{\text{Con}}(\psi)) = w(\varphi) + w(\psi) \end{aligned}$$

c) if $\exists x\varphi$ is a sentence, since the domain is Con, a model $\mathfrak{A} \in \text{Mod}_{\text{Con}}(\mathcal{L})$ satisfies $\exists x\varphi$ if and only if there is $i \in \mathbb{N}^+$ such that $\mathfrak{A} \vDash \varphi(a_i)$. Hence,

$$\text{Mod}_{\text{Con}}(\exists x\varphi) = \bigcup_{i\in\mathbb{N}^+} \text{Mod}_{\text{Con}}(\varphi(a_i))$$

and by the properties of a (measure-theoretic) probability[10],

$$\begin{aligned} w(\exists x\varphi) &= \mu(\text{Mod}_{\text{Con}}(\exists x\varphi)) \\ &= \mu(\bigcup_{i\in\mathbb{N}^+} \text{Mod}_{\text{Con}}(\varphi(a_i))) \\ &= \lim_{n\to+\infty} \mu(\bigcup_{i\leq n} \text{Mod}_{\text{Con}}(\varphi(a_i))) \\ &= \lim_{n\to+\infty} \mu(\text{Mod}_{\text{Con}}(\bigvee_{i\leq n} \varphi(a_i))) \\ &= \lim_{n\to+\infty} w(\bigvee_{i\leq n} \varphi(a_i)), \end{aligned}$$

For the part about the unicity of such an extension w, assume that also w' extends w^- and that it is a Gaifman probability. We show by induction on the sentence rank that these maps are the same. For sentences with null rank, i.e. without quantifiers, w and w' coincide with w^-, since they extend it; for the inductive step, notice first that all the formulas are equivalent to one in prenex form, i.e. one in which all the quantifiers are in the beginning. Then, the thesis follows from these equations

$$w(\exists x\varphi) \overset{\circ}{=} \lim_{n\to+\infty} w(\bigvee_{i=1}^{n} \varphi(a_i)) \overset{*}{=} \lim_{n\to+\infty} w'(\bigvee_{i=1}^{n} \varphi(a_i)) \overset{\circ}{=} w'(\exists x\varphi),$$

$$w(\forall x\varphi) \overset{\circ}{=} \lim_{n\to+\infty} w(\bigwedge_{i=1}^{n} \varphi(a_i)) \overset{*}{=} \lim_{n\to+\infty} w'(\bigwedge_{i=1}^{n} \varphi(a_i)) \overset{\circ}{=} w'(\forall x\varphi),$$

where $\overset{*}{=}$ holds by induction hypothesis, and $\overset{\circ}{=}$ since w and w' are Gaifman. □

[10]See Proposition A.1.1, Appendix A.

Remark 1.2.2. As pointed out in the Footnote 8 (page 23), the above proof cannot be used when functional symbols are involved. Indeed, we can show that for a language that has equality and a functional symbol f (in addition to the constants in Con), the argument used to show the conditionally σ-additivity of μ_{w^-} doesn't hold. If we put for any $i \in \mathbb{N}^+$,

$$\varphi_i := f(a_1) = a_i \wedge \bigwedge_{j<i} f(a_1) \neq a_j,$$

then

- for $i \neq j$, $\mathrm{Mod}_{\mathrm{Con}}(\varphi_i) \cap \mathrm{Mod}_{\mathrm{Con}}(\varphi_j) = \emptyset$, since φ_i and φ_j are in contradiction if $i \neq j$;
- for any model $\mathfrak{A}$ in $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, there is a minimum $i \in \mathbb{N}^+$ such that $\mathfrak{A} \models f(a_1) = a_i$; therefore,
$$\bigcup_{i \in \mathbb{N}^+} \mathrm{Mod}_{\mathrm{Con}}(\varphi_i) = \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) = \mathrm{Mod}_{\mathrm{Con}}(\top);$$
- there is no $n \in \mathbb{N}^+$ for which
$$\bigcup_{i \leq n} \mathrm{Mod}_{\mathrm{Con}}(\varphi_i) = \mathrm{Mod}_{\mathrm{Con}}(\top):$$
indeed, the model $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ in which f is interpreted as the function
$$f^{\mathfrak{A}} : a_i \mapsto a_{n+i},$$
does't belong to $\bigcup_{i \leq n} \mathrm{Mod}_{\mathrm{Con}}(\varphi_i)$.

This shows then that the proof presented for Theorem 1.2.1 is not valid for generic languages with functional symbols and equality. However, in [7] the reader can find a generalization of this result in the case of such languages. We don't discuss it here because it involves a slightly different definition of the Gaifman condition and it will carry us far from our aim.

Now, we are ready to establish the promised correspondence theorem between probabilities on sentences and probabilities in the measure-theoretic sense. In the following, by a probability on $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, we mean a probability on the borel σ-algebra $\mathcal{F}(\mathcal{A}_{\mathrm{Con}})$.

Theorem 1.2.2. *Let w be a Gaifman probability on $\mathrm{Sen}(\mathcal{L})$. Then, there is a probability μ_w on $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ such that for every $\varphi \in \mathrm{Sen}(\mathcal{L})$,*

$$w(\varphi) = \mu_w(\mathrm{Mod}_{\mathrm{Con}}(\varphi)).$$

Conversely, let μ be a probability on $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$. Then there is a Gaifman probability w_μ on $\mathrm{Sen}(\mathcal{L})$ such that for every $\varphi \in \mathrm{Sen}(\mathcal{L})$,

$$w_\mu(\varphi) = \mu(\mathrm{Mod}_{\mathrm{Con}}(\varphi)).$$

Proof. Assume that w is a Gaifman probability on $\mathrm{Sen}(\mathcal{L})$ and call w^- its restriction to $\mathrm{QFSen}(\mathcal{L})$. By Theorem 1.2.1, there exists an extension w' of w^- and a probability μ on $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ such that for any $\varphi \in \mathrm{Sen}(\mathcal{L})$,

$$w'(\varphi) = \mu(\mathrm{Mod}_{\mathrm{Con}}(\varphi)).$$

By the unicity of the extension proved in the theorem, $w' = w$ so putting $\mu_w = \mu$, we get the thesis.

Conversely, let's now assume to have a probability on $\text{Mod}_{\text{Con}}(\mathcal{L})$. We can define $w_\mu : \text{Sen}(\mathcal{L}) \to [0,1]$ such that $w_\mu(\varphi) = \mu(\text{Mod}_{\text{Con}}(\varphi))$ for any $\varphi \in \text{Sen}(\mathcal{L})$. In the proof of Theorem 1.2.1, we have already shown that this is a Gaifman probability. □

Thanks to this theorem, for a probability w on $\text{Sen}(\mathcal{L})$, to which corresponds the probability μ_w on $\text{Mod}_{\text{Con}}(\mathcal{L})$, we can write

$$w(\varphi) = \int_{\text{Mod}_{\text{Con}}(\mathcal{L})} w_{\mathfrak{B}}(\varphi)\ d\mu_w(\mathfrak{B}),$$ [11]

for any formula $\varphi \in \text{Sen}(\mathcal{L})$, where $w_{\mathfrak{B}}$ is the probability introduced in Def. 1.0.1. This is because, in general, for a measure μ defined on a σ-algebra $\mathcal{F}$ on a set A, and a subset $B \subseteq A$ beloning to $\mathcal{F}$, we have

$$\mu(B) = \int_A \chi_B(x)\ d\mu(x),$$

where $\chi_B : A \to \{0,1\}$ is the indicator function of B. In our case, $w(\varphi)$ is equal to $\mu_w(\text{Mod}_{\text{Con}}(\varphi))$, $A = \text{Mod}_{\text{Con}}(\mathcal{L})$, and $B = \text{Mod}_{\text{Con}}(\varphi)$, hence

$$w(\varphi) = \int_{\text{Mod}_{\text{Con}}(\mathcal{L})} \chi_{\text{Mod}_{\text{Con}}(\varphi)}(\mathfrak{B})\ d\mu_w(\mathfrak{B}).$$

Now, we can notice that

$$\chi_{\text{Mod}_{\text{Con}}(\varphi)}(\mathfrak{B}) = 1 \text{ if and only if } \mathfrak{B} \models \varphi,$$

for any $\mathfrak{B} \in \text{Mod}_{\text{Con}}(\mathcal{L})$, therefore

$$\chi_{\text{Mod}_{\text{Con}}(\varphi)}(\mathfrak{B}) = w_{\mathfrak{B}}(\varphi)$$

and the thesis follows.

1.3 Dutch Book argument

In this section, we present another argument that should persuade the reader to accept the Definition 1.0.1 of a probability over sentences at least in the case of languages $\mathcal{L}$ with a countable supply of constants and no functional symbols or equality. Indeed, we make this hypothesis in the following, because we will use Theorem 1.2.2 that relies on this assumption.

The argument we will discuss is called *Dutch Book* and it is based on the idea that the belief of an (omniscient) agent should match its betting prices. An agent that has partial knowledge of the structure it lives in (let's say $\mathfrak{A}$) can join this game with the dealer: given a sentence φ, $p \in [0,1]$, and a stake $s > 0$ the dealer asks the agent to bet according to one of the following wagers:

[11] Notice that here $w_{\mathfrak{B}}(\varphi)$ is seen as a function with input $\mathfrak{B}$ and not as a function from $\text{Sen}(\mathcal{L})$ to $[0,1]$.

- ($Bet1^{\varphi}_{p}$): The agent will win $s(1-p)$ if $\mathfrak{A} \models \varphi$ and lose sp otherwise;
- ($Bet2^{\varphi}_{p}$): The agent will lose $s(1-p)$ if $\mathfrak{A} \models \varphi$ and win sp otherwise;

The two bets that the agent can take are one the "opposite" of the other, in the sense that we obtain ($Bet2^{\varphi}_{p}$) if we swap the roles of the agent and the dealer in ($Bet1^{\varphi}_{p}$). Hence, the agent should accept one of the bets for every p, s and φ , since either both the alternatives are fair (so it is indifferent what to choose) or one is better than the other and so it can choose it.

Assume that the sentence φ and the stake s are fixed and that $p, q \in [0, 1]$. Then we can notice that:

- ($Bet1^{\varphi}_{0}$) is acceptable for the agent since it has nothing to lose and it can only win;
- if ($Bet1^{\varphi}_{p}$) is acceptable for the agent, then for $q < p$ also ($Bet1^{\varphi}_{q}$) is: indeed, moving from the first to the latter case, the reward increases and the loss decreases.

The goal we want to pursue now is to understand what a rational agent should do in a hypothetical game against the dealer. But what can be a way to test the flaws of the agent's beliefs?

One possible solution to the question is to force the agent to play any bet that it considers acceptable: if we believe that a certain behavior earns us, why we should not follow it?

For what was said before, for every $\varphi \in \mathrm{Sen}(\mathcal{L})$, the set

$$\{p \in [0, 1] : (Bet1^{\varphi}_{p}) \text{ is acceptable for the agent}\}$$

is a non-empty (because it contains 0) interval of $[0, 1]$; if the dealer proposes the agent to play according to ($Bet1^{\varphi}_{p}$) for a p that is in the set above, then it should accept to bet; conversely, if $q \in [0, 1]$ isn't in the set at issue, then the agent should accept to play according to ($Bet2^{\varphi}_{q}$). Hence, the sup of it represents the threshold under which the agent will play the first kind of bet. For every $\varphi \in \mathrm{Sen}(\mathcal{L})$ we call $Bel(\varphi)$ the sup of the above set; we get a function from the $\mathcal{L}$-sentences to $[0, 1]$ that determines the beliefs of an agent intended as the willingness to accept the bet $Bet1^{\varphi}_{p}$ offered by the dealer.

Conversely, every function $f : \mathrm{Sen}(\mathcal{L}) \to [0, 1]$ can be a description of a policy for the agent to follow in this game: given a stake s and a formula φ then the agent accepts to bet according to ($Bet1^{\varphi}_{p}$) if and only if $p \leq f(\varphi)$ and to ($Bet2^{\varphi}_{p}$) for all $p > f(\varphi)$. Hence, there is a correspondence between strategies for the game presented above and functions from $\mathrm{Sen}(\mathcal{L})$ to $[0, 1]$. We should notice that here the stakes don't play a central role: indeed, the policy given by a function doesn't say anything about the stakes to bet, but only about how to bet. This is because if an agent finds convenient to bet a stake $s > 0$ according to $Bet1^{\varphi}_{p}$, then nothing should change if the stake involved changes (remaining still positive).

Stating more formally what we said in the introduction, we expect a rational agent to have a beliefs function $Bel(-) : \mathrm{Sen}(\mathcal{L}) \to [0, 1]$ that doesn't "have faults", that doesn't "contradict" itself. But what does it mean? The first remark we should make is that we can't fully judge the rationality of an agent evaluating only one bet. Indeed, only in a specific case we can regard the choice of the agent unreasonable: when it bets according to $(Bet2^{\varphi}_p)$ with $p < 1$ for a tautology φ (or, equivalently, when it bets following $(Bet1^{\varphi}_q)$ with $q > 0$ when $\neg\varphi$ is a tautology). In this case, the agent will certainly lose, no matter the structure in which it lives in.

But when φ is not a tautology, nor is $\neg\varphi$, we can't evaluate if $(Bet1^{\varphi}_p)$ is better or not than $(Bet2^{\varphi}_p)$. It's different if we know other bets that the agent makes: we can then judge whether its choices are irrational (or better inconsistent) or not. To formalize this concept, notice that if an agent lives in the structure $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ and it bets according to $(Bet1^{\varphi}_p)$, then it gains

$$s(1-p)w_{\mathfrak{A}}(\varphi) + sp(w_{\mathfrak{A}}(\varphi) - 1) = s(w_{\mathfrak{A}}(\varphi) - p),$$

where we use the notation of Example 1.0.1 and the convention that a loss is a negative gain; when the bet is according to $(Bet2^{\varphi}_p)$, then it gains

$$-s(1-p)w_{\mathfrak{A}}(\varphi) + sp(1 - w_{\mathfrak{A}}(\varphi)) = -s(w_{\mathfrak{A}}(\varphi) - p).$$

As an example of the claim in the above paragraph, we expect that an agent that is enough confident to bet according to $Bet1^{\varphi\wedge\psi}_p$, then it will bet willingly according to the same wager for φ. Rephrasing it, we won't consider rational an agent that has $Bel(\varphi \wedge \psi) > Bel(\varphi)$: in this case, it is possible for the dealer to play against it and certainly win! Indeed, suppose to take p, q such that $Bel(\varphi) < p < q < Bel(\varphi\wedge\psi)$: hence, the agent willingly will bet according to $Bet1^{\varphi\wedge\psi}_q$ and to $Bet2^{\varphi}_p$. If an agent accepts a given wager, it should do it no matter the stake that is involved: thus, considering that in the hypothesis above it will bet the same stake $s > 0$ for the two bets, the agent will gain

$$s(w_{\mathfrak{A}}(\varphi \wedge \psi) - q) - s(w_{\mathfrak{A}}(\varphi) - p) = s(w_{\mathfrak{A}}(\varphi \wedge \psi) - w_{\mathfrak{A}}(\varphi) - q + p).$$

Now, noticing that for any structure $\mathfrak{A}$, $w_{\mathfrak{A}}(\varphi \wedge \psi) \leq w_{\mathfrak{A}}(\varphi)$ and recalling that a negative gain is a loss, we get the thesis.

In this game setting, an agent will have some beliefs that induce it to play for some sentences according to $Bet1$ and for others according to $Bet2$: in particular, assume that the sets $\{\varphi_i\}_{i\in A}$ and $\{\psi_j\}_{j\in B}$ contain the sentences for which it plays, respectively, according to the first and the second wager, with $A, B \subseteq \mathbb{N}$.[12] If for any $i \in A$, the agent bets according to $Bet1^{\varphi_i}_{p_i}$ a stake $s_i > 0$ and for any $j \in B$ it bets according to $Bet2^{\psi_j}_{q_j}$ a stake $t_j > 0$, then what it gains is expressed by the following sum:

$$\sum_{i\in A} s_i(w_{\mathfrak{A}}(\varphi_i) - p_i) + \sum_{j\in B}(-t_j)(w_{\mathfrak{A}}(\psi_j) - q_j),$$

where $\mathfrak{A}$ is the structure the agent lives in.

This sum is of no information about the agent's rationality because the bets considered can be unacceptable for it. Hence, we should require that $p_i \in [0, Bel(\varphi_i))$[13] for $i \in A$ and $q_j \in (Bel(\psi_j), 1]$ for $j \in B$.

Another thing to notice is that the sum that determines the gain or the loss of the agent depends on the structure $\mathfrak{A}$ considered. If we take a sentence φ that is neither valid nor a contradiction, then every policy determined by a function $Bel(-)$ as explained above can lead to a loss in a given world. Indeed:

[12] Here, A and B are countable: this shouldn't seem problematic to the reader, since the assumptions we made about the language $\mathcal{L}$ imply the countability of the $\mathcal{L}$-sentences up to logical equivalence.

[13] If $Bel(\varphi_i) = 0$, then we assume this interval to be $\{0\}$; the same notation will be used in the case $q_j \in (Bel(\psi_j), 1]$ when $Bel(\psi_j) = 1$.

- if $Bel(\varphi) \neq 0$, then take $0 < p < Bel(\varphi)$: $Bet1^{\varphi}_{p}$ is acceptable for the agent and it gains $s(w_{\mathfrak{A}}(\varphi) - p)$ if s is the stake played. Considering a structure $\mathfrak{A}$ in which φ doesn't hold, we get that the agent loses;

- if $Bel(\varphi) = 0$, then take $0 < q < 1$: $Bet2^{\varphi}_{q}$ is acceptable for the agent and it gains $-s(w_{\mathfrak{A}}(\varphi) - q)$ if s is the stake played. Considering a structure $\mathfrak{A}$ in which φ holds, we get that the agent loses.

Hence, the existence of a structure $\mathfrak{A}$ for which the agent loses, i.e. for which

$$\sum_{i \in A} s_i(w_{\mathfrak{A}}(\varphi_i) - p_i) + \sum_{j \in B}(-t_j)(w_{\mathfrak{A}}(\psi_j) - q_j) < 0,$$

isn't avoidable: indeed, what is required in this formulation to a rational agent, is that the bets don't lead to a *certain* loss, i.e. to a loss in any possible world.

Formalizing What we said before, an agent that behaves in the game according to the function $Bel(-)$ is *irrational* if there are countable sets $A, B \subseteq \mathbb{N}$, sentences φ_i, stakes $s_i > 0$, numbers $p_i \in [0, Bel(\varphi_i))$ for $i \in A$, and sentences ψ_j, stakes $t_j > 0$, numbers $q_j \in (Bel(\psi_j), 1]$ for $j \in B$, and $K > 0$ such that

$$\sum_{i \in A} s_i(w_{\mathfrak{B}}(\varphi_i) - p_i) + \sum_{j \in B}(-t_j)(w_{\mathfrak{B}}(\psi_j) - q_j) < 0 \tag{1.10}$$

$$\sum_{i \in A} \left| s_i(w_{\mathfrak{B}}(\varphi_i) - p_i) \right| + \sum_{j \in B} \left| t_j(w_{\mathfrak{B}}(\psi_j) - q_j) \right| < K, \tag{1.11}$$

for all models $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$.

The disequality (1.11) guarantees that (1.10) is meaningful and finite: in this way, indeed, the sums involved are finite and can be added; technically, we need this condition (that can be interpreted as a requirement of the effective feasibility of the game) to apply the Dominate Convergence Theorem in the proofs of the following results. We notice that if A and B are finite then, since the stacks s_i, t_j are real numbers, we can always find a bound K as in (1.11), hence the disequality is satisfied.

This condition of irrationality appeared for the first time in [24], and now, scholars refer to this kind of reasoning as a *Dutch book argument*; thus, we will say that an agent's belief function $Bel(-)$ can be *Dutch booked* if there are $A, B, \{\varphi_i\}_{i \in A}$, $\{p_i\}_{i \in A}$, $\{s_i\}_{i \in A}$, $\{\psi_j\}_{j \in B}$, $\{q_j\}_{j \in B}$, $\{t_j\}_{j \in B}$ and K as above for which (1.10) and (1.11) hold for all structures $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$.

Surprisingly, this notion of rationality is the same as the notion of probability mentioned previously in the chapter, as the following theorems explain.

Theorem 1.3.1. *Let Bel be a function from* $\mathrm{Sen}(\mathcal{L}) \to [0, 1]$ *that can't be Dutch booked; then, it is a Gaifman probability.*

Proof. To prove the theorem we have to show that Bel satisfies the properties to be a probability:

- let φ be a valid statement and $Bel(\varphi) < 1$. Then for all models $\mathfrak{B}$, $w_{\mathfrak{B}}(\varphi) = 1$, hence for q such that $Bel(\varphi) < q < 1$, we have

$$-(w_{\mathfrak{B}}(\varphi) - q) < 0$$

and (1.10) and (1.11) hold for any $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, with $\{\varphi_i\}_{i\in A} = \emptyset$,$\{\psi_j\}_{j\in B} = \{\varphi\}$, the stake $t_1 = 1$ and $q_1 = q$. Since Bel can't be Dutch booked, we must conclude that $Bel(\varphi) = 1$;

- let $\varphi \wedge \psi$ be a contradiction and suppose that $Bel(\varphi \vee \psi) \neq Bel(\varphi) + Bel(\psi)$. We will show that $Bel(\varphi \vee \psi) > Bel(\varphi) + Bel(\psi)$ leads to a contradiction and a similar argument proves the same when $Bel(\varphi \vee \psi) < Bel(\varphi) + Bel(\psi)$.

  Since $\varphi \wedge \psi$ is a contradiction, for every model $\mathfrak{B}$ it holds that $w_{\mathfrak{B}}(\varphi \vee \psi) = w_{\mathfrak{B}}(\varphi) + w_{\mathfrak{B}}(\psi)$. Therefore, taking $q_1, q_2, p \in [0,1]$ such that $q_1 > Bel(\varphi), q_2 > Bel(\psi), p < Bel(\varphi \vee \psi)$, and $q_1 + q_2 < p$, we have

$$(w_{\mathfrak{B}}(\varphi \vee \psi) - p) + (-1)(w_{\mathfrak{B}}(\varphi) - q_1) + (-1)(w_{\mathfrak{B}}(\psi) - q_2) = -p + q_1 + q_2 < 0.$$

  Thus, (1.10) and (1.11) hold for any structure $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, with $\{\varphi_i\}_{i\in A} = \{\varphi \vee \psi\}$, $\{\psi_j\}_{j\in B} = \{\varphi, \psi\}$, and the stakes $s_1 = t_1 = t_2 = 1$.

- let $\exists x\varphi$ be a sentence and we will show that

$$Bel(\exists x\varphi) = \lim_{n\to+\infty} Bel(\bigvee_{i\leq n} \varphi(a_i)).$$

  Notice that from the previous points, Bel is a probability, therefore, the limit in the equation above exists, since the succession $\{Bel(\bigvee_{i=1}^{n} \varphi(a_i))\}_{n\in\mathbb{N}}$ is bounded (by 1) and increasing.

  For any $n \in \mathbb{N}^+$, $\bigvee_{i=1}^{n} \varphi(a_i) \vDash \exists x\varphi$, hence, by Proposition 1.0.1,

$$Bel(\bigvee_{i=1}^{n} \varphi(a_i)) \leq Bel(\exists x\varphi),$$

  and taking the limit of the left-hand-side term, we have that

$$\lim_{n\to+\infty} Bel(\bigvee_{i=1}^{n} \varphi(a_i)) \leq Bel(\exists x\varphi).$$

  Now, we will show that the inequality above is indeed an equality: assume on the contrary that it is strict. One useful property that can be derived directly from Definition 1.0.1 (and based on 7) of Proposition 1.0.1), is that

$$\lim_{n\to+\infty} Bel(\bigvee_{i=1}^{n} \varphi(a_i)) = \sum_{n=1}^{+\infty} Bel(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i)),$$

  therefore, the assumption we have made can be rephrased as

$$\sum_{n=1}^{+\infty} Bel(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i)) < Bel(\exists x\varphi).$$

  Hence, pick for every $n \in \mathbb{N}^+$ $q_n \in [0,1]$ and $0 < p < Bel(\exists x\varphi)$ such that

$$q_n > Bel(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i))$$

and $\sum_{n=1}^{+\infty} q_n < p$.

Since for every model $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, $w_{\mathfrak{B}}$ is a Gaifman probability and hence, for what we said before,

$$w_{\mathfrak{B}}(\exists x\varphi) = \sum_{n=1}^{+\infty} w_{\mathfrak{B}}(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i)), \tag{1.12}$$

we have that Bel is Dutch bookable: take $\{\varphi_i\}_{i\in A} = \{\exists x\varphi\}$, $\{\psi_j\}_{j\in B} = \{\varphi(a_n)\}_{n\in\mathbb{N}^+}$, all the stacks equal to 1, $p_1 = p$, the q_n's as above and $K = 3$. We get that, for any structure $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$:

- Equation (1.11) holds since

$$\left| w_{\mathfrak{B}}(\exists x\varphi) - p \right| < 1,$$

and

$$\begin{aligned}\sum_{n=1}^{+\infty}\left|(-1)(w_{\mathfrak{B}}(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i)) - q_n)\right| \leq& \\ &\leq \sum_{n=1}^{+\infty}(w_{\mathfrak{B}}(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i)) + \sum_{n=1}^{+\infty} q_n \\ &\leq w_{\mathfrak{B}}(\exists x\varphi) + \sum_{n=1}^{+\infty} q_n \\ &\leq 1 + p;\end{aligned}$$

- Equation (1.10) holds since

$$(w_{\mathfrak{B}}(\exists x\varphi) - p) + \sum_{n=1}^{+\infty}(-1)(w_{\mathfrak{B}}(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i)) - q_n) \overset{(1.12)}{=}$$

$$\sum_{n=1}^{+\infty} w_{\mathfrak{B}}(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i)) - p + \sum_{n=1}^{+\infty}(-1)(w_{\mathfrak{B}}(\varphi(a_n) \wedge \neg \bigvee_{i=1}^{n-1} \varphi(a_i)) + \sum_{n=1}^{+\infty} q_n$$

$$= -p + \sum_{n=1}^{+\infty} q_n < 0,$$

where in $\overset{(1.12)}{=}$ we used Equation (1.12) and in the last equality we could rearrange the terms thanks to the absolute convergence of the sums involved.

□

Theorem 1.3.2. *Let w be a Gaifman probability on* Sen($\mathcal{L}$)*; then, it can't be Dutch booked.*

Proof. Suppose that w can be Dutch booked and that it's a Gaifman probability; this will lead us to a contradiction. In the following, we will use the notation used above before Equation (1.10) and (1.11).

By Theorem 1.2.2, since w is a Gaifman probability, let μ_w be the probability on $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ such that

$$w(\varphi) = \mu_w(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) = \int_{\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})} w_{\mathfrak{B}}(\varphi) d\mu_w(\mathfrak{B}).$$

Assume now that w is Dutch bookable, and consider $\{\varphi_i\}_{i\in A}$, $\{\psi_j\}_{j\in B}$, $\{p_i\}_{i\in A}$, $\{q_j\}_{j\in B}$, $\{s_i\}_{i\in A}$, $\{t_j\}_{j\in B}$, K as in Equation (1.10) and (1.11), witnesses of this.

Notice that

$$\begin{aligned}
&\sum_{i\in A} s_i(w(\varphi_i) - p_i) + \sum_{j\in B}(-t_j)(w(\psi_j) - q_j) \\
&= \sum_{i\in A} s_i(\int_{\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})} w_{\mathfrak{B}}(\varphi_i) d\mu_w(\mathfrak{B}) - p_i) + \sum_{j\in B}(-t_j)(\int_{\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})} w_{\mathfrak{B}}(\psi_j) d\mu_w(\mathfrak{B})) - q_j) \\
&\stackrel{*}{=} \int_{\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})} \left(\sum_{i\in A} s_i(w_{\mathfrak{B}}(\varphi_i) - p_i) + \sum_{j\in B}(-t_j)(w_{\mathfrak{B}}(\psi_j) - q_j)\right) d\mu_w(\mathfrak{B}),
\end{aligned}$$

where $\stackrel{*}{=}$ is motivated by the fact that μ_w is a probability and by the Dominated Convergence Theorem (since (1.11) holds).

By (1.10), the integrand is negative (for all model $\mathfrak{B}$ the function we are considering is less than 0), hence, the whole integral is so, and

$$\sum_{i\in A} s_i(w(\varphi_i) - p_i) + \sum_{j\in B}(-t_j)(w(\psi_j) - q_j) < 0 \tag{1.13}$$

But now, recalling that $p_i \in [0, w(\varphi_i))$, $q_j \in (w(\psi_j), 1]$ and that the stakes s_i, t_j are positive for all $i \in A$ and $j \in B$ (these restrictions come out from the fact that the p_i's, the q_j's and the stakes were taken to be the witnesses of the possibility to Dutch book w), we must conclude a contradiction from (1.13), because every term we are summing up is non-negative. □

Chapter 2

Ex and De Finetti Theorem

In the previous chapter, we moved the first step in the argument of probabilities on sentences and we have already noticed that some results require some restrictions on the language. In particular, also from this chapter on, we will assume that the language $\mathcal{L}$ will have a countable supply of constants ($\{a_i\}_{i\in\mathbb{N}^+}$) that should be interpreted as the elements of the domain at issue. The language will have also some relation symbols $\{R_1, \ldots, R_q\}$ whose interpretation in the model where the agent lives ni is unknown to him and that varies as the model varies. In the language, there are no functional symbols, nor the equality, except in the cases in which this is explicitly marked out.

In addition, having in mind this kind of application, the Gaifman condition makes sense and this will be assumed through the following chapters: hence, from now on, when we use the world "probability", we mean "Gaifman probability".

2.1 Constant Exchangeability Principle

Suppose having an urn with a countable supply of balls in it that differ only for their color (some are red and some are not) but you can't see the inside; in addition, you don't have any further relevant information like the proportion between the red balls and the not-red ones. In this condition of *zero-knowledge*, what will be the probability that a ball is red?

We can formalize the problem considering a language $\mathcal{L}$ with a countable set of constants $\mathrm{Con} = \{a_i \ : i \in \mathbb{N}^+\}$ that saturates all the elements of the universe and a unary predicate symbol R (for instance, with $R(x)$ meaning that the ball represented by x is red). Having no relevant information about the problem, the agent doesn't have any reason to treat a constant differently from the other: they simply represent different elements of our universe.

Remark 2.1.1. The condition of zero-knowledge is fundamental here. Indeed, obviously, if the agent knows that the balls represented respectively by the constant symbols a_1 and a_2 are respectively red and not, it should treat the two symbols differently. Notice that all the principles we will present have this hypothesis in the background. This can be seen as a limitation (see Chapter 3, Section 3.3), but it should be, without any doubt, the starting point: if we aren't able to detect rational principles in this context, it will be really difficult to understand what is rational when some previous knowledge is involved.

In addition, the implicit thought here, is that if we end up with a probability w that satisfies "all the rational requirements we want" in the zero-knowledge condition, then, knowing the validity of the statement φ, the right probability would be $w(-|\varphi)$: this seems reasonable and underlies the discussion in Chapter 2 and Chapter 3.

This required "symmetry", discussed so far, is formalized by the first rational principle we introduce, in Definition 2.1.1. In the previous chapter, we showed that there are a lot of functions on sentences that satisfy some basic features (like the condition in Definition 1.0.1); the golden goal of Inductive Logic is to pick up some rational principles (that can vary according to the application involved) and restrict more and more the set of "valid" (for that use) probabilities. The following principle, however, is assumed unreservedly by all.

Definition 2.1.1 (*Ex- Constant Exchangeability Principle*).
For any $\varphi(b_1, \ldots, b_n) \in \text{Sen}(\mathcal{L})$, where $b_1, \ldots, b_n$ are distinct constants of the language $\mathcal{L}$ and for any other choice of distinct (from each other, not necessarily from the b_i's) constants $b'_1, \ldots, b'_n$,

$$w(\varphi(b_1, \ldots, b_n)) = w(\varphi(b'_1, \ldots, b'_n)). \tag{2.1}$$

Remark 2.1.2. We want the reader to pay attention to the fact that with the notation $\varphi(b_1, \ldots, b_n)$ we are assuming that $b_1, \ldots, b_n$ are distinct constant symbols. This is an important assumption as the following example shows.

Suppose that in the language $\mathcal{L}$ we have only a binary relation R. Then a probability w that satisfies Ex will give the same value to $R(a, b)$ and to $R(a', b')$ for every distinct a, b and some distinct a', b'; if we strengthen the principle, allowing the possibility of some b'_i's to be the same, then w must assign equal probability also to $R(c, c)$ for each constant c. The symmetry involved in the stronger version is not really well motivated (and it is also often awkward in some settings): there is a certain kind of analogy between $R(a, b)$ and $R(a', b')$ but the same level of analogy doesn't exist confronting $R(a, b)$ and $R(a, a)$. Notice that, indeed, Ex implies that

$$w(R(a, b)) = w(R(a', b')) \quad w(R(a, a)) = w(R(b, b)),$$

but the principle doesn't force $w(R(a, b))$ to be the same of $w(R(a, a))$.

Remark 2.1.3. We can show, using basically the same argument used in the unicity part of Theorem 1.2.1, that if w satisfies Equation (2.1) only with regards to the formulas $\varphi(b_1, \ldots, b_n)$ in $\text{QFSen}(\mathcal{L})$, then w satisfies also Ex (so Equation (2.1)) also for formulas with quantifiers.

Remark 2.1.3 shows a really useful tool in dealing with such properties: often if we test a property only for a particular subset of sentences, it is valid also for all the sentences of the language. With this slogan in mind, we can further simplify Ex, restricting ourselves not only to quantifier-free sentences but also to specific ones.

Recall that the language we are using is composed of constant symbols $a_1, a_2, \ldots$ and q relation symbols $R_1, \ldots, R_q$ of ariety $r_1, \ldots, r_q$, respectively. For distinct constants $b_1, \ldots, b_m$ coming from the a_i's, a *state description* (or a *state formula*) for $b_1, \ldots, b_m$ is a sentence of the form

$$\Theta = \bigwedge_{i=1}^{q} \bigwedge_{c_1, \ldots, c_{r_i}} \pm R_i(c_1, \ldots, c_{r_i}),$$

where for any $i = 1, \dots, q$ the $c_1, \dots, c_{r_i}$ range over all (not necessarily distinct) choices of $b_1, \dots, b_m$ and $\pm R_i$ is an abbreviation that stands for either R_i or $\neg R_i$. Since for every $i = 1, \dots, q$, we take into account all the possible choices of r_i elements among the constants we are interested in, a state description describes all it is possible to know about how $b_1, \dots, b_m$ relates to each other according to the various R_i's.

With this notation, we mean also that for every i, a choice $c_1, c_2, \dots, c_{r_i}$ appears only one time in the state description: therefore, Θ is satisfiable. It will be useful to give also a meaning when $m = 0$: in this case, no constant symbols will appear in Θ, and we will interpret it as a tautology (since it is the empty conjunction).

Example 2.1.1. If a language has a unary symbol P and a binary symbol R, then a possible state description for the constants a, b is

$$\Theta(a, b) = P(a) \wedge \neg P(b) \wedge R(a, a) \wedge \neg R(a, b) \wedge R(b, a) \wedge R(b, b).$$

State descriptions will be denoted by upper case Greek letter as Θ, Φ, Ψ: sometimes with a slight abuse of notation, we will call also *state descriptions (for $x_1, \dots, x_n$)* formulas $\Theta(x_1, \dots, x_n)$ such that for any constants $b_1, \dots, b_n$, the sentence $\Theta\{x_1/b_1, \dots, x_n/b_n\}$ is a state description for $b_1, \dots, b_n$.

State descriptions are useful because every quantifier-free sentence is equivalent to a (finite, possibly empty) disjunction of state descriptions (Disjunctive Normal Form Theorem)[1]; furthermore, two different state descriptions are contradictory, hence if we have a sentence $\varphi(\vec{b}) \in \mathrm{QFSen}(\mathcal{L})$ that is equivalent to $\bigvee_{i \in S} \Theta_i(\vec{b})$ for a given finite set S, we have

$$w(\varphi(\vec{b})) = \sum_{i \in S} w(\Theta_i(\vec{b})).$$

As a consequence of the discussion above, to fully describe a probability w, it is sufficient to determine the values given to state descriptions.

Another useful property that we show here is that we can compute w on a state description involving some constants $b_1, \dots, b_m$ by simply knowing the values assumed by w on some state descriptions for $b_1, \dots, b_r$ with $r \geq m$. Indeed, for any state description $\Theta(b_1, \dots, b_m)$, we have

$$\Theta(b_1, \dots, b_m) \equiv \bigvee_{\Phi(b_1, \dots, b_r) \vDash \Theta(b_1, \dots, b_m)} \Phi(b_1, \dots, b_r),$$

i.e. $\Theta(b_1, \dots, b_m)$ is equivalent to the disjunction of all possible state descriptions Φ for $b_1, \dots, b_r$ that imply Θ. Hence,

$$w(\Theta(b_1, \dots, b_m)) = \sum_{\Phi(b_1, \dots, b_r) \vDash \Theta(b_1, \dots, b_m)} w(\Phi(b_1, \dots, b_r)). \tag{2.2}$$

We can notice, also, that the semantic consequence relationship $\vDash$ is easy to determine on state descriptions. Indeed, $\Phi(b_1, \dots, b_r) \vDash \Theta(b_1, \dots, b_m)$ if and only if $\Phi(b_1, \dots, b_r)$ restricted to the constants $b_1, \dots, b_m$ is equal to $\Theta(b_1, \dots, b_m)$.

We showed that a probability w is determined uniquely by its value on state descriptions. However, to construct a probability starting from the values assigned to state

[1] We use the usual convention that an empty disjunction is $\bot$.

descriptions, we should pay some attention to avoid an overdetermination or a not well-defined function. However, if we assign a value $w(\Theta)$ for any state description in such a way that the following holds:

$$\begin{aligned}&w(\Theta(b_1,\dots,b_m)) \geq 0 \text{ for every } b_1,\dots,b_m \text{ constants;}\\ &w(\top) = 1;\\ &\text{Equation (2.2) holds for any } r \geq m \text{ and every constant symbols } b_1,\dots,b_r,\end{aligned} \tag{2.3}$$

there is a unique extension to a probability w on Sen($\mathcal{L}$) that coincides with the above values on state descriptions.

These restrictions can take a simpler form if we want to construct a probability w with some specific features. For instance, if we want w to satisfy Ex, there is no need to specify $w(\Theta(b_1,\dots,b_m))$ for any $b_1,\dots,b_m$, once defined the value of $w(\Theta(a_1,\dots,a_m))$: indeed, the values must be the same for Ex to hold.

2.2 De Finetti Theorem: unary case

In this part of the chapter, we begin our journey toward the De Finetti Theorems (Theorem 2.2.1 and Theorem 2.4.4): in this section, we will start from the case of unary languages, i.e. the relation symbols $R_1,\dots,R_q$ in the language $\mathcal{L}$ are unary.

In this setting a state description $\Theta(b_1,\dots,b_n)$ for the constants $b_1,\dots,b_n$ (among the set $\{a_i\}_{i\in\mathbb{N}^+}$) has the form

$$\Theta(b_1,\dots,b_n) = \bigwedge_{i=1}^{n}\bigwedge_{j=1}^{q} \pm R_j(b_i),$$

where $+R$ and $-R$ are intended to be, respectively, R and $\neg R$. Another representation of a state description is given taking into account *atoms* i.e. formulas $\alpha(x)$ of the form $\bigwedge_{j=1}^{q} \pm R_j(x)$. In the case of unary languages, an atom where x is replaced by a constant b describes all we have to know about b[2]; furthermore, from the assumption that the set of relation symbols is finite (it has cardinality q), it follows that also the number of all the possible atoms is finite (it has cardinality 2^q and they will be denoted $\alpha_1(x),\dots,\alpha_{2^q}(x)$). Any state description, henceforth, can be seen as a conjunction of atoms[3], i.e.

$$\Theta(b_1,\dots,b_n) = \bigwedge_{i=1}^{n} \alpha_{h_i}(b_i),$$

where the h_i's vary in the set $\{1,\dots,2^q\}$.

In the unary context, we have also a normal form for a generic statement and not only for state descriptions. Since any formula $\psi(b_1,\dots,b_m) \in$ Sen($\mathcal{L}$) is equivalent to one in prenex form, there exists a quantifier-free formula $\varphi = \varphi(x_1,\dots,x_n,b_1,\dots,b_m)$ such that ψ is equivalent to a block of quantifiers with regards to $x_1,\dots,x_n$ (denoted

[2]Notice that this is not the case when the language has at least a binary relation symbol.

[3]Here and in the following we will use the world *atom* not only to mean a formula like $\alpha_i(x)$ with x as free variable but also the statements of the form $\alpha_i(b)$ for a constant symbol b: this is in analogy with the terminology used for state descriptions.

by $Q\vec{x}$) followed by φ. Writing φ as a disjunction of state descriptions and using the representation above, we get

$$\psi \equiv Q\vec{x}\ \varphi(\vec{x},\vec{b}) \equiv Q\vec{x} \bigvee_{k=1}^{r} \Big(\bigwedge_{j=1}^{n} \alpha_{g_{kj}}(x_j) \wedge \bigwedge_{i=1}^{m} \alpha_{h_{ki}}(b_i)\Big), \tag{2.4}$$

for some g_{kj}, h_{ki} numbers in $\{1,\dots,2^q\}$ that depend, respectively, on k,j and on k and i.

Using the standard rules of quantifiers and proceeding by induction on the numbers of quantifiers on $Q\vec{x}$, we can find a formula equivalent to ψ of the form

$$\bigvee_{k=1}^{l} \Big(\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_{jk}} \wedge \bigwedge_{i=1}^{m} \alpha_{f_{ki}}(b_i)\Big), \tag{2.5}$$

where $\varepsilon_{jk} \in \{0,1\}$ and for a generic formula ϕ, ϕ^1 denotes ϕ, and ϕ^0 denotes $\neg\phi$.

The proof of this fact is left to the reader, even if we will give all the essential tools needed. When we want to remove from the block an existential quantifier, we simply can carry it inside since

$$\exists x \bigvee_{k=1}^{r} \big(\alpha_{g_k}(x) \wedge \theta_k\big) \equiv \bigvee_{k=1}^{r} \big(\exists x \alpha_{g_k}(x) \wedge \theta_k\big),$$

where the θ_k's represent formulas (depending on k) in which x doesn't occur free. Dealing with a universal quantifier is trickier: indeed, we will have formulas like $\forall x \bigvee_{k=1}^{r}\big(\alpha_{g_k}(x) \wedge \theta_k\big)$, where θ_k is used as explained above. In this case, we can use De Morgan's law and write $\bigvee_{k=1}^{r}\big(\alpha_{g_k}(x) \wedge \theta_k\big)$ as a conjunction of disjunctions. Now we get the thesis, by repeatedly using the following observations:

$$\big(\forall x \bigwedge_{p=1}^{2^r} \phi_p\big) \equiv \big(\bigwedge_{p=1}^{2^r} \forall x \phi_p\big)$$

$$\forall x \big(\bigvee_{j\in J} \alpha_j(x) \vee \theta\big) \equiv \big(\forall x \bigvee_{j\in J} \alpha_j(x)\big) \vee \theta \equiv \big(\bigwedge_{i\in\{1,\dots,2^q\}\setminus J} \neg\exists x \alpha_i(x)\big) \vee \theta,$$

where the ϕ_p's are generic formulas, in θ doens't occur x free and J is a subset of $\{1,\dots,2^q\}$.

Furthermore, the disjuncts can be taken pairwise contradictory, and all satisfiable (with the usual convention that if we have an empty disjunction (i.e. $l = 0$), then ψ is a contradiction): this normal form is unique up to order.

The Ex principle in the case of unary language takes this easier form: a probability w satisfies Ex if and only if the value assigned to $\bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)$ depends only on the *signature* $(m_1, m_2, \dots, m_{2^q})$ of the state description, where

$$m_j = |\{i \mid h_i = j\}|.$$

Hence, under Ex, to determine the value of a state description, it is sufficient to know how many times every atom appears in the sentence.

Now, we will try to investigate further properties of probabilities that satisfy Ex.

For ease of notation, assume that $q = 1$, i.e. the language consists of one unary predicate symbol R: it is then easy to generalize the proof in the case of a language with more than one relation symbol. If $q = 1$, the only two atoms are $\alpha_1(x) = R(x)$ and $\alpha_2(x) = \neg R(x)$.

By Ex, for every n and k, the value $w(\bigwedge_{i=1}^{n+k} \alpha_{h_i}(b_i))$, where $|\{i \mid h_i = 1\}| = n$ and $|\{i \mid h_i = 2\}| = k$, depends only on n and k. Hence, it makes sense to define this to be $w(n, k)$. Notice that if $n = k = 0$, then the state descriptions we are studying are a tautology, thus $w(0, 0) = 1$. From now on, we will assume that at least one among n, k is non-zero.

We know also that for every constant b, one and only one among $\alpha_1(b)$ and $\alpha_2(b)$ holds: this means that for any $r > n + k > 0$, the following formula is valid

$$\bigvee_{(r_1,r_2,h_1,\ldots,h_r)\in A} \bigwedge_{i=1}^{r} \alpha_{h_i}(a_i),$$

where

$$A = \{(r_1, r_2, h_1, \ldots, h_r) \in \mathbb{N}^2 \times \{1, 2\}^r \ : \ r_1 + r_2 = r \text{ and } |\{i \mid h_i = 1\}| = r_1\}.$$

Hence, since different state descriptions are in contradiction, we have

$$1 = w\Big(\bigvee_{(r_1,r_2,h_1,\ldots,h_r)\in A} \bigwedge_{i=1}^{r} \alpha_{h_i}(a_i)\Big) = \sum_{(r_1,r_2,h_1,\ldots,h_r)\in A} w\Big(\bigwedge_{i=1}^{r} \alpha_{h_i}(a_i)\Big).$$

By the fact that for fixed r_1 (and then r_2, since r is given) there exist $\binom{r}{r_1}$ choices of $h_1, \ldots, h_r$ such that $(r_1, r_2, h_1, \ldots, h_r) \in A$ and that the summand $w\big(\bigwedge_{i=1}^{r} \alpha_{h_i}(a_i)\big)$ depends only on the signature (i.e. on r_1 and r_2), we have

$$1 = \sum_{r_1+r_2=r} \binom{r}{r_1} w(r_1, r_2). \tag{2.6}$$

By a similar argument, based this time on the fact that for fixed $r > n+k$, $\bigwedge_{i=1}^{n+k} \alpha_{h_i}(a_i)$ is equivalent to the disjunction of all the state descriptions for $a_1, \ldots, a_r$ that extend it,[4] we can show that

$$w(n, k) = \sum_{\substack{r_1+r_2=r \\ n\le r_1, k\le r_2}} \binom{r - n - k}{r_1 - n} w(r_1, r_2). \tag{2.7}$$

If we define the symplex

$$\mathbb{D}_2 = \{(a, b) \in [0, 1]^2 \mid a + b = 1\},$$

[4] i.e. if $n = |\{i \ : \ h_i = 1\}|$ and $k = |\{i \ : \ h_i = 2\}|$,

$$\bigwedge_{i=1}^{n+k} \alpha_{h_i}(a_i) \equiv \bigvee_{(r_1,r_2,f_{n+k+1},\ldots,f_r)\in B} \Big(\bigwedge_{i=1}^{n+k} \alpha_{h_i}(a_i) \wedge \bigwedge_{j=n+k+1}^{r} \alpha_{f_j}(a_j)\Big),$$

where

$$B = \{(r_1, r_2, f_{n+k+1}, \ldots, f_r) \ : \ r_1 \ge n, r_2 \ge k, r_1 + r_2 = r \text{ and } |\{j \mid f_j = 1\}| = r_1 - n\}.$$

Equation (2.6) says that for any $r \in \mathbb{N}^+$ putting

$$\mu_r(a,b) := \begin{cases} \binom{r}{r_1} w(r_1, r_2) & \text{if } a = r_1/r,\ b = r_2/r, \text{ where } r_1, r_2 \in \mathbb{N} \text{ and } r_1 + r_2 = r; \\ 0 & \text{otherwise;} \end{cases}$$

we are dealing with a discrete probability on $\mathbb{D}_2$.

With this formulation, Equation (2.7) becomes

$$w(n,k) = \sum_{\substack{r_1+r_2=r \\ n \le r_1, k \le r_2}} \binom{r-n-k}{r_1-n} \binom{r}{r_1}^{-1} \mu_r((r_1/r, r_2/r)). \tag{2.8}$$

This is not an easy description since the product of the two binomials can't be further simplified. However, some computations yield to

$$\binom{r-n-k}{r_1-n} \binom{r}{r_1}^{-1} =$$

$$= \left(\frac{r_1}{r}\right)^n \left(\frac{r_2}{r}\right)^k \frac{(1-r_1^{-1}) \cdots (1-(n-1)r_1^{-1})(1-r_2^{-1}) \cdots (1-(k-1)r_2^{-1})}{(1-r^{-1}) \cdots (1-(n+k-1)r^{-1})}$$

and it can be shown that in the right-hand-side the last fraction tends to 1 uniformly in r_1 and r_2 as r tends to $+\infty$.

Since in Equation (2.8) the left-hand-side doesn't depend on r, so it is the right-hand-one, therefore a nicer description for $w(n,k)$ would be

$$\begin{aligned} w(n,k) &= \lim_{r \to +\infty} \sum_{\substack{r_1+r_2=r \\ n \le r_1, k \le r_2}} \binom{r-n-k}{r_1-n} \binom{r}{r_1}^{-1} \mu_r((r_1/r, r_2/r)) \\ &= \lim_{r \to +\infty} \sum_{\substack{r_1+r_2=r \\ n \le r_1, k \le r_2}} \left(\frac{r_1}{r}\right)^n \left(\frac{r_2}{r}\right)^k \mu_r((r_1/r, r_2/r)) \end{aligned} \tag{2.9}$$

where in the last equality we use the fact that the convergence above is uniform with regards to r_1 and r_2.

To get another simplification in this equation, we notice that, calling for ease of notation

$$f(r_1, r_2) := \left(\frac{r_1}{r}\right)^n \left(\frac{r_2}{r}\right)^k \mu_r((r_1/r, r_2/r)),$$

for any fixed n, k, r

$$\begin{aligned} \sum_{r_1+r_2=r} f(r_1,r_2) = &\sum_{\substack{r_1+r_2=r \\ n \le r_1, k \le r_2}} f(r_1,r_2) + \sum_{\substack{r_1+r_2=r \\ r_1 < n, k \le r_2}} f(r_1,r_2) \\ &+ \sum_{\substack{r_1+r_2=r \\ n \le r_1, r_2 < k}} f(r_1,r_2) + \sum_{\substack{r_1+r_2=r \\ r_1 < n, r_2 < k}} f(r_1,r_2). \end{aligned}$$

Now, recalling that we are analyzing the case for which either n or k or both are different from 0, taking in the limit for r that approaches $+\infty$, we have:

- $\lim_{r\to+\infty}\sum_{\substack{r_1+r_2=r\\ r_1<n,k\le r_2}} f(r_1,r_2)=0$ since:
  - if $n=0$, then the sum is empty, from which follows the thesis;
  - if $n=1$, $r_1<n$ implies that $r_1=0$ and $r=r_2\ge k$. In this case, the sum is composed of only one null term;
  - if $n>1$, then

$$\begin{aligned}\sum_{\substack{r_1+r_2=r\\ r_1<n,k\le r_2}} f(r_1,r_2) &\le \sum_{\substack{r_1+r_2=r\\ r_1<n,k\le r_2}} \left(\frac{r_1}{r}\right)^n\left(\frac{r_2}{r}\right)^k\\ &= \sum_{\substack{r_1+r_2=r\\ 0<r_1<n,0\le k\le r_2}} \left(\frac{r_1}{r}\right)^n\left(\frac{r_2}{r}\right)^k\\ &< \sum_{\substack{r_1+r_2=r\\ 0<r_1<n,0\le k\le r_2}} \left(\frac{n}{r}\right)^n\\ &\le r\left(\frac{n}{r}\right)^n = \frac{n^n}{r^{n-1}} \overset{r\to+\infty}{\longrightarrow} 0;\end{aligned}$$

- $\lim_{r\to+\infty}\sum_{\substack{r_1+r_2=r\\ n\le r_1,r_2<k}} f(r_1,r_2)=0$, for a similar argument as the one presented above;
- $\lim_{r\to+\infty}\sum_{\substack{r_1+r_2=r\\ r_1<n,r_2<k}} f(r_1,r_2)=0$, because for r big enough ($r>n+k-2$), there aren't $r_1<n$ and $r_2<k$ with the sum equal to r; hence, the sum will be definitely empty and the outcome will be 0.

This means that we can omit the restrictions $n\le r_1$ and $k\le r_2$ in Equation (2.9) and get

$$\begin{aligned}w(n,k) &= \lim_{r\to+\infty}\sum_{r_1+r_2=r}\left(\frac{r_1}{r}\right)^n\left(\frac{r_2}{r}\right)^k \mu_r((r_1/r,r_2/r))\\ &= \lim_{r\to+\infty}\int_{\mathbb{D}_2} x_1^n\, x_2^k\, d\mu_r((x_1,x_2)).\end{aligned}$$

The last step in the description of $w(n,k)$ can be performed if we find out a weak convergence of the probabilities μ_r's. Indeed, since $\mathbb{D}_2$ is compact (endowed with the topology induced by the standard one in $\mathbb{R}^2$), we can extract from the tight succession $\{\mu_r\}_{r>n+k}$ a subsequence $\{\mu_{i_r}\}_{r>n+k}$ that weakly converges to a probability μ over the borel σ-algebra of $\mathbb{D}_2$.[5]

[5] A probability μ on the borel σ-algebra of a subset $\mathbb{D}_2\subset\mathbb{R}^2$ can be seen as a probability on $\mathbb{R}^2$, by simply defining for any $A\in\mathcal{B}(\mathbb{R}^2)$

$$\mu(A) := \mu(A\cap\mathbb{D}_2);$$

since $\mathbb{D}_2$ is compact, any family of probability on $\mathbb{D}_2$, seen as a family of probability on $\mathbb{R}^2$ is tight. Hence, we can apply Theorem A.3.1 in Appendix A.

Therefore, we have

$$\begin{aligned} w(n,k) &= \lim_{r\to+\infty} \int_{\mathbb{D}_2} x_1^n\ x_2^k\ d\mu_r((x_1,x_2)) \\ &= \lim_{r\to+\infty} \int_{\mathbb{D}_2} x_1^n\ x_2^k\ d\mu_{i_r}((x_1,x_2)) \\ &= \int_{\mathbb{D}_2} x_1^n\ x_2^k\ d\mu((x_1,x_2)). \end{aligned}$$

The argument above can be carried out even if the language has more than one, say q, unary relation symbols. Hence, to summarise, by simply requiring a probability w to satisfy Ex, we get a really useful description of its structure. Furthermore, also a converse result holds as we will show immediately after having introduced some notations: in the following the set $\mathbb{D}_{2q}$ will be the simplex in $\mathbb{R}^{2q}$

$$\mathbb{D}_{2q} = \{(x_1,\dots,x_{2q}) \in [0,1]^{2q}\ :\ \sum_{i=1}^{2q} x_i = 1\}.$$

Given $\vec{x} \in \mathbb{D}_{2q}$, we have a probability on Sen($\mathcal{L}$) putting

$$w_{\vec{x}}(\bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)) = \prod_{i=1}^{m} x_{h_i} = \prod_{j=1}^{2^q} x_j^{m_j},$$

for all the state descriptions and then extending to all the sentences in QFSen($\mathcal{L}$) and in Sen($\mathcal{L}$) as provided by previous arguments, where $m_j = |\{i\ :\ h_i = j\}|$. To verify that $w_{\vec{x}}$ defines a probability, we can check whether the conditions required in (2.3) are satisfied:

- for every state descriptions $\Theta = \bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)$,

  $$w_{\vec{x}}(\Theta) = \prod_{i=1}^{m} x_{h_i} \geq 0,$$

  since every $\vec{x} \in \mathbb{D}_{2q}$ has all the components non-negative;

- $w_{\vec{x}}(\top) = 1$ because it is the empty product.[6]

- suppose $r \geq m$ and that we have a state description $\Theta = \bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)$. Then, a state description $\bigwedge_{j=1}^{r} \alpha_{f_j}(b_j)$ has Θ as a semantic consequence if and only if for every

[6]This is a usual convention. However, if we wanted to avoid this fuzzy notation, we could define directly $w_{\vec{x}}(\top) = 1$.

$j \leq m$, $h_j = f_j$. This means that

$$\sum_{\Phi(b_1,\ldots,b_r)\models\Theta(b_1,\ldots,b_m)} w_{\vec{x}}(\Phi(b_1,\ldots,b_r)) =$$

$$= \sum_{g:\{m+1,\ldots,r\}\to\{1,\ldots,2^q\}} w_{\vec{x}}\big(\bigwedge_{i=1}^{m} \alpha_{h_i}(b_i) \wedge \bigwedge_{i=m+1}^{r} \alpha_{g(i)}(b_i)\big)$$

$$= \sum_{g:\{m+1,\ldots,r\}\to\{1,\ldots,2^q\}} \prod_{i=1}^{m} x_{h_i} \cdot \prod_{i=m+1}^{r} x_{g(i)}$$

$$= \prod_{i=1}^{m} x_{h_i} \cdot \big(\sum_{g:\{m+1,\ldots,r\}\to\{1,\ldots,2^q\}} \prod_{i=m+1}^{r} x_{g(i)}\big)$$

$$= \prod_{i=1}^{m} x_{h_i} = w_{\vec{x}}\big(\bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)\big) = w_{\vec{x}}(\Theta(b_1,\ldots,b_m)).$$

Then, we can apply Theorem 1.2.1 and get a probability on all the sentences.

Theorem 2.2.1 (De Finetti)**.**
Let $\mathcal{L}$ be a unary language with q relation symbols and w a probability on $\text{Sen}(\mathcal{L})$ *satisfying Ex. Then there is a unique probability μ on the borel σ-algebra of $\mathbb{D}_{2q}$ such that*

$$w(\bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)) = \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)) \; d\mu(\vec{x}),^{7} \tag{2.10}$$

for any $m \in \mathbb{N}^+$ and atoms α_{h_i} of the language.

Conversely, if μ is a probability on the borel σ-algebra of $\mathbb{D}_{2q}$, and w is defined as in Equation (2.10), then it can be extended to a probability on $\text{Sen}(\mathcal{L})$ *that satisfies Ex.*

Proof. We have already proved the existence of a probability μ as required, for the case $q = 1$; the argument can be generalized when $q > 1$. For the unicity part, take another probability μ' for which Equation (2.10) holds. Therefore, we have that

$$\int_{\mathbb{D}_{2q}} \prod_{j=1}^{2^q} x_j^{m_j} \; d\mu(\vec{x}) = w(\bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)) = \int_{\mathbb{D}_{2q}} \prod_{j=1}^{2^q} x_j^{m_j} \; d\mu'(\vec{x}).$$

This means that given any monomial with variables $x_1, \ldots, x_{2q}$, we have that the integrations with the two measures μ and μ' give the same outcome; by linearity, we have the same also for any polynomial in these variables, and by Stone-Weierstrass Theorem, also for any continuous function $\mathbb{D}_{2q} \to \mathbb{R}$.[8] Then the two probabilities μ and μ' are equal (see Appendix A, Remark A.3.1).

[7]We can notice that the function $\vec{x} \mapsto w_{\vec{x}}(\bigwedge_{i=1}^{m} \alpha_{h_i}(b_i))$ is measurable since it is a polynomial, hence continuous.

[8]One formulation of the Stone-Weierstrass Theorem says that if X is a compact Haussdorf space and A is an algebra of continuous real-valued functions on X that contains the constant functions and separate the points of X (i.e. for any different point $x, y \in X$, there is a function $f \in A$ such that $f(x) \neq f(y)$), then A is uniformly dense in the space $C(X)$ of all the continuous functions from X to $\mathbb{R}$. If we take $X = \mathbb{D}_{2q}$, then the algebra $\mathcal{P}$ of all polynomials over $x_1, \ldots, x_{2^q}$ satisfies the properties of the statement and, hence, for any continuous function $f : \mathbb{D}_{2q} \to \mathbb{R}$ and for any $n \in \mathbb{N}^+$, there is a polynomial p_n in $\mathcal{P}$

Conversely, suppose having a probability μ on $\mathbb{D}_{2q}$ and a function w defined as a convex combination of probabilities of the form $w_{\vec{x}}$ as in Equation (2.10). Then, we have:

- w is a probability. Indeed, by Theorem 1.2.1, it is sufficient to show that the restriction of w on QFSen($\mathcal{L}$) satisfies the conditions in Definition 1.0.1 (with regards to quantifier-free sentences). In particular, since w is defined on state descriptions we can also only verify that the required conditions in (2.3) hold. In doing this, we will use that these properties hold for the $w_{\vec{x}}$'s, as already shown:
  - for any state description $\bigwedge_{i=1}^m \alpha_{h_i}(b_i)$, we have
  $$w(\bigwedge_{i=1}^m \alpha_{h_i}(b_i)) = \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\bigwedge_{i=1}^m \alpha_{h_i}(b_i))\ d\mu(\vec{x}) \geq 0,$$
  because the integrand is non-negative for every $\vec{x}$;
  - $w(\top) = \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\top)\ d\mu(\vec{x}) = \int_{\mathbb{D}_{2q}} d\mu(\vec{x}) = 1.$
  - for any $r \geq m$, we have already shown that
  $$w_{\vec{x}}(\Theta(b_1, \dots, b_m)) = \sum_{\Phi(b_1,\dots,b_r) \models \Theta(b_1,\dots,b_m)} w_{\vec{x}}(\Phi(b_1, \dots, b_r)).$$
  Thus, the same holds for w because the sum involves only a finite number of terms and then can be carried out the integral. More explicitly:
  $$\begin{aligned} w(\Theta(b_1, \dots, b_m)) &= \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\Theta(b_1, \dots, b_m))\ d\mu(\vec{x}) \\ &= \int_{\mathbb{D}_{2q}} \sum_{\Phi(b_1,\dots,b_r) \models \Theta(b_1,\dots,b_m)} w_{\vec{x}}(\Phi)\ d\mu(\vec{x}) \\ &= \sum_{\Phi(b_1,\dots,b_r) \models \Theta(b_1,\dots,b_m)} \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\Phi)\ d\mu(\vec{x}) \\ &= \sum_{\Phi(b_1,\dots,b_r) \models \Theta(b_1,\dots,b_m)} w(\Phi(b_1, \dots, b_r)). \end{aligned}$$

- w satisfies Ex. Indeed, by the definition of $w_{\vec{x}}$ and Equation (2.10) we have
$$w(\bigwedge_{i=1}^m \alpha_{h_i}(b_i)) = \int_{\mathbb{D}_{2q}} \prod_{j=1}^{2^q} x_j^{m_j}\ d\mu(\vec{x})$$

---

such that
$$\|f - p_n\|_\infty := \sup_{\vec{x} \in \mathbb{D}_{2q}} |f(x) - p_n(x)| < \frac{1}{n}.$$
Then, since p_n tends to f as n approaches to $+\infty$ and f (and so the p_n's) is bounded since the domain is compact, we have by Dominated Convergence Theorem that
$$\begin{aligned} \int_{\mathbb{D}^{2q}} f(\vec{x})\ d\mu(\vec{x}) &= \lim_{n \to +\infty} \int_{\mathbb{D}^{2q}} p_n(\vec{x}) d\mu(\vec{x}) \\ &= \lim_{n \to +\infty} \int_{\mathbb{D}^{2q}} p_n(\vec{x}) d\mu'(\vec{x}) \\ &= \int_{\mathbb{D}^{2q}} f(\vec{x})\ d\mu'(\vec{x}). \end{aligned}$$

and the right-hand-side (and so the left-hand one) depends only on the m_j's for $j = 1, \ldots, 2^q$, but not on the b_i's.

□

As an immediate corollary, we can generalize Equation (2.10) from conjunctions of atoms to sentences.

Corollary 2.2.2. *With the notations used in Theorem 2.2.1, if w is a probability that satisfies Ex, then there exists a unique probability μ on the borel σ-algebra of $\mathbb{D}_{2q}$ such that for every $\varphi(b_1, \ldots, b_n) \in \mathrm{Sen}(\mathcal{L})$*

$$w(\varphi(b_1, \ldots, b_n)) = \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\varphi(b_1, \ldots, b_n))\ d\mu(\vec{x}). \tag{2.11}$$

Conversely, if μ is a probability on the borel σ-algebra of $\mathbb{D}_{2q}$, then, using Equation (2.11), we can define a probability w that satisfies Ex.

Proof. By Theorem 2.2.1 there exists a probability μ on $\mathbb{D}_{2q}$ for which Equation (2.11) holds when φ is a state description. Hence, since every formula in $\mathrm{QFSen}(\mathcal{L})$ can be written as a disjunction of pairwise incompatible atoms, the equation at issue is valid also for sentences without quantifiers. Reasoning by induction on the rank of the formulas, we get the thesis: for instance, if $\varphi(b_1, \ldots, b_n) = \exists x \psi(x, b_1, \ldots, b_n)$, we have

$$\begin{aligned}
w(\varphi(b_1, \ldots, b_n)) &= w(\exists x \psi(x, b_1, \ldots, b_n)) \\
&\overset{\circ}{=} \lim_{m \to +\infty} w\Big(\bigvee_{i=1}^{m} \psi(a_i, b_1, \ldots, b_n)\Big) \\
&\overset{@}{=} \lim_{m \to +\infty} \int_{\mathbb{D}_{2q}} w_{\vec{x}}\Big(\bigvee_{i=1}^{m} \psi(a_i, b_1, \ldots, b_n)\Big)\ d\mu(\vec{x}) \\
&\overset{*}{=} \int_{\mathbb{D}_{2q}} \lim_{m \to +\infty} w_{\vec{x}}\Big(\bigvee_{i=1}^{m} \psi(a_i, b_1, \ldots, b_n)\Big)\ d\mu(\vec{x}) \\
&= \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\exists x \psi(b_1, \ldots, b_n))\ d\mu(\vec{x}) \\
&= \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\varphi(b_1, \ldots, b_n))\ d\mu(\vec{x}),
\end{aligned}$$

where $\overset{\circ}{=}$ holds since w satisfies the Gaifman property, $\overset{@}{=}$ for the inductive hypothesis and $\overset{*}{=}$ is motivated by the Monotone Convergence Theorem; the case in which the involved quantifier is $\forall$ is analogous.

The unicity is trivially implied by Theorem 2.2.1.

Conversely, if w is defined by Equation (2.11), it is a probability since:

- if φ is a tautology, for any $\vec{x} \in \mathbb{D}_{2q}$, $w_{\vec{x}}(\varphi) = 1$ since $w_{\vec{x}}$ is a probability and then

$$w(\varphi) = \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\varphi)\ d\mu(\vec{x}) = \int_{\mathbb{D}_{2q}} d\mu(\vec{x}) = 1;$$

- if φ and ψ are in contradiction, then the thesis follows by a similar argument as above exploiting the fact that $w_{\vec{x}}(\varphi \vee \psi) = w_{\vec{x}}(\varphi) + w_{\vec{x}}(\psi)$ for any $\vec{x} \in \mathbb{D}_{2q}$;

- if $\varphi(b_1, \dots, b_n) = \exists x \psi(x, b_1, \dots, b_n)$, then

$$\begin{aligned} w(\varphi(b_1, \dots, b_n)) &= w(\exists x \psi(x, b_1, \dots, b_n)) \\ &= \int_{\mathbb{D}_{2q}} w_{\vec{x}}(\exists x \psi(b_1, \dots, b_n)) \; d\mu(\vec{x}) \\ &\overset{\circ}{=} \int_{\mathbb{D}_{2q}} \lim_{m \to +\infty} w_{\vec{x}}\Big(\bigvee_{i=1}^{m} \psi(a_i, b_1, \dots, b_n)\Big) \; d\mu(\vec{x}) \\ &\overset{*}{=} \lim_{m \to +\infty} \int_{\mathbb{D}_{2q}} w_{\vec{x}}\Big(\bigvee_{i=1}^{m} \psi(a_i, b_1, \dots, b_n)\Big) \; d\mu(\vec{x}) \\ &= \lim_{m \to +\infty} w\Big(\bigvee_{i=1}^{m} \psi(a_i, b_1, \dots, b_n)\Big), \end{aligned}$$

  where $\overset{\circ}{=}$ holds because $w_{\vec{x}}$ is a Gaifman probability and $\overset{*}{=}$ for the Monotone Convergence Theorem.

Hence, Equation (2.11) doesn't overdetermine w; obviously this probability satisfies also Equation (2.10), therefore it satisfies Ex, by Theorem 2.2.1. □

Theorem 2.2.1, clarified by Corollary 2.2.2, provides a useful tool for dealing with probabilities that satisfy Ex. Indeed, such a probability can be seen as a convex combination (according to μ) of probabilities with a specific form, i.e. of $w_{\vec{x}}$'s with $\vec{x} \in \mathbb{D}_{2q}$; furthermore, the probability μ that determines the "weights of the various points in the simplex"[9] with regards to the representation of w is unique. In the following, given a probability w, we will call such a μ the *De Finetti prior* of w. Let's show an example.

Example 2.2.1. Assume to have a language $\mathcal{L}$ with only the unary relation symbol R and let's describe the probability w starting from its De Finetti prior. If λ is the Lebesgue measure on $[0, 1]$, then the map $f : t \mapsto (t, 1 - t)$ from $[0, 1]$ to $\mathbb{D}_2$ determines the pushforward probability $f_*\lambda$ on $\mathbb{D}_2$. In this case, for any function $g : \mathbb{D}_2 \to \mathbb{R}$, we have

$$\int_{\mathbb{D}_2} g(\vec{x}) \; df_*\lambda(\vec{x}) = \int_{[0,1]} g(f(t)) \; d\lambda(t).$$[10]

In this language, we have only two atoms $\alpha_1(x) = R(x)$ and $\alpha_2(x) = \neg R(x)$ and for any constant b, w gives the same value to all the atoms:

$$w(R(b)) = \int_{\mathbb{D}_2} x_1 \; d\lambda((x_1, x_2)) = \int_{[0,1]} t \; dt = \frac{1}{2}$$
$$w(\neg R(b)) = \int_{\mathbb{D}_2} x_2 \; d\lambda((x_1, x_2)) = \int_{[0,1]} 1 - t \; dt = \frac{1}{2}.$$

[9]This makes sense if μ is a discrete probability; there can be cases in which this is not, but however this can be thought as an informal description of the role of μ in the representation of w.

[10]See the item *Pushforward measure* at page 96, Appendix A.

If we have a state description $\bigwedge_{i=1}^{n+k} \alpha_{h_i}(b_i)$ with $n = |\{i \ : \ h_i = 1\}|$ and $k = |\{i \ : \ h_i = 2\}$,

$$\begin{aligned} w(\bigwedge_{i=1}^{n+k} \alpha_{h_i}(b_i)) &= \int_{\mathbb{D}_2} x_1^n x_2^k \ d\lambda((x_1, x_2)) \\ &= \int_{[0,1]} t^n (1-t)^k \ dt = \frac{k!n!}{(n+k+1)!}. \end{aligned} \tag{2.12}$$

So, since the outcome is the same if we swap n and k, w satisfies SN: this probability, indeed, in some sense, keeps the major symmetry of the Lebesgue measure. We can notice also that when quantified formulas are involved, we get

$$\begin{aligned} w(\forall x R(x)) &= \int_{\mathbb{D}_2} w_{\vec{x}}(\forall x R(x)) \ d\lambda((x_1, x_2)) = \int_{\mathbb{D}_2} \lim_{n \to +\infty} w_{\vec{x}}(\bigwedge_{i=1}^{n} R(a_i)) \ d\lambda((x_1, x_2)) \\ &= \lim_{n \to +\infty} \int_{\mathbb{D}_2} x_1^n d\lambda((x_1, x_2)) = \lim_{n \to +\infty} \int_0^1 t^n \ dt = \lim_{n \to +\infty} \frac{1}{n+1} = 0. \end{aligned}$$

We end this example by showing a very useful feature of this probability related to epistemic induction. By previous computation, using the same notation used in (2.12), if $n = |\{i \ : \ h_i = 1\}|$ and $k = |\{i \ : \ h_i = 2\}$, we have that

$$\begin{aligned} w(\alpha_1(a_{n+k+1}) | \bigwedge_{i=1}^{n+k} \alpha_{h_i}(a_i)) &= \frac{w(\bigwedge_{i=1}^{n+k} \alpha_{h_i}(a_i) \wedge \alpha_1(a_{n+k+1}))}{w(\bigwedge_{i=1}^{n+k} \alpha_{h_i}(a_i))} \\ &= \frac{(n+1)!k!}{(n+k+2)!} \frac{(n+k+1)!}{n!k!} = \frac{n+1}{n+k+2}. \end{aligned}$$

This means that an agent with this probability, after having seen a lot of constants for which R holds, is more inclined to believe that also it holds for a new constant a_{n+k+1}: for instance, roughly speaking, after having seen 10 objects, 9 of which satisfies R, the probability it assigns to the chance that a new object satisfies $R(x)$ is 10/12.[11]

The description of the building blocks used to construct a general probability w that satisfies Ex, i.e. the $w_{\vec{x}}$'s, appears essentially relative to the context of unary languages. The definition of $w_{\vec{x}}$, indeed, doesn't extend very easily in a situation in which we have relation symbols that have ariety greater than 1. To understand better this part, we should notice that if we have in our language a non-unary relation symbol, say R of ariety 2, there's no possibility to think about atoms that describe entirely a constant symbol. A formula φ that is supposed to represent all that we know about, for instance, the constant a_1, entails only a finite number of constant symbols, say only the ones among the set $\{a_1, \ldots, a_n\}$; therefore, it can't say anything about the relation between a_1 and a_{n+1} according to R and we lack this piece of information.

In this sense, since in general we can't talk about atoms as in the unary case, we should hook the definition of the $w_{\vec{x}}$'s to other more generalizable features. To this goal, we introduce a new principle and show that it characterizes the probabilities $w_{\vec{x}}$.

[11]This probability is usually called $c_{2q}^{\mathcal{L}}$: we are using here a standard notation attributed to Carnap. This probability is just an element of the Carnap's continuum family $\{c_\lambda^{\mathcal{L}}\}_{\lambda \in [0,+\infty]}$, a set of probabilities depending on a parameter λ: we will meet in Chapter 3, page 78 another important member of this family, the probability $c_0^{\mathcal{L}}$.

Definition 2.2.1 (IP- Constant Irrelevance Principle).
A probability w satisfies the *Constant Irrelevance Principle*, or *IP*, if for any two formulas φ and ψ in QFSen($\mathcal{L}$) with no constants in common,

$$w(\varphi \wedge \psi) = w(\varphi) \cdot w(\psi). \tag{2.13}$$

As in the case of other similar principles, if w is a probability that satisfies IP, then Equation (2.13) holds also for φ and ψ in Sen($\mathcal{L}$).

Theorem 2.2.3. *Let $\mathcal{L}$ be a unary language with q relation symbols and w a probability on* Sen($\mathcal{L}$) *that satisfies Ex. Then, w satisfies IP if and only if there exists $\vec{x} \in \mathbb{D}_{2q}$ such that $w = w_{\vec{x}}$.*

Proof. As already pointed out, since $w_{\vec{x}}(\bigwedge_{i=1}^m \alpha_{h_i}(b_i))$ depends only on the signature of the state description and not on the constant symbols involved, it satisfies Ex.

It is also easy to prove that IP holds for $w_{\vec{x}}$ if φ and ψ are state descriptions, i.e, conjunctions of atoms: indeed, if the b_i's are all different from the c_j's,

$$w_{\vec{x}}\Big(\bigwedge_{i=1}^m \alpha_{h_i}(b_i) \wedge \bigwedge_{j=1}^n \alpha_{f_j}(c_j)\Big) = \prod_{i=1}^m x_{h_i} \cdot \prod_{j=1}^n x_{f_j} = w_{\vec{x}}\Big(\bigwedge_{i=1}^m \alpha_{h_i}(b_i)\Big) \cdot w_{\vec{x}}\Big(\bigwedge_{j=1}^n \alpha_{f_j}(c_j)\Big).$$

Since every formula in QFSen($\mathcal{L}$) can be written as a disjunction of pairwise contradictory state descriptions, IP also holds in the case in which φ and ψ are formulas without quantifiers: for instance, if $\varphi = \bigvee_{i=1}^m \Theta_i$ and $\psi = \bigvee_{j=1}^n \Phi_j$, and the constants in Θ_i are different from the ones in Φ_j for every i, j, we have:

$$\begin{aligned}
w_{\vec{x}}(\varphi \wedge \psi) &= w_{\vec{x}}\Big(\bigvee_{i=1}^m \Theta_i \wedge \bigvee_{j=1}^n \Phi_j\Big)\\
&= w_{\vec{x}}\Big(\bigvee_{i=1}^m \bigvee_{j=1}^n \Theta_i \wedge \Phi_j\Big)\\
&= \sum_{i=1}^m \sum_{j=1}^n w_{\vec{x}}(\Theta_i \wedge \Phi_j)\\
&= \sum_{i=1}^m \sum_{j=1}^n w_{\vec{x}}(\Theta_i) \cdot w_{\vec{x}}(\Phi_j)\\
&= \Big(\sum_{i=1}^m w_{\vec{x}}(\Theta_i)\Big) \cdot \Big(\sum_{j=1}^n w_{\vec{x}}(\Phi_j)\Big)\\
&= w_{\vec{x}}(\varphi) \cdot w_{\vec{x}}(\psi).
\end{aligned}$$

Conversely, if w satisfies IP, then for any conjucntion of atoms,

$$w\Big(\bigwedge_{i=1}^m \alpha_{h_i}(b_i)\Big) = \prod_{i=1}^m w(\alpha_{h_i}(b_i)).$$

If the probability satisfies also Ex, the value $w(\alpha_{h_i}(b_i))$ doesn't depend on the constant b_i but only on the atom $\alpha_{h_i}(x)$, so if $c_j = w(\alpha_j(a_1))$, then

$$w\Big(\bigwedge_{i=1}^m \alpha_{h_i}(b_i)\Big) = \prod_{i=1}^m c_{h_i}.$$

In addition, for any j $c_j \geq 0$ and since $\vee_{j=1}^{2^q}\alpha_j(a_1)$ is a tautology, we have that the vector $\vec{c} = (c_1, \ldots, c_{2^q})$ belongs to $\mathbb{D}_{2^q}$ and w is of the required form. □

2.3 Loeb Measure Theory

In the following section, we will try to generalize the result shown in the previous chapter (namely Theorem 2.2.1) to the case in which the language $\mathcal{L}$ has also relation symbols with ariety greater than one. As already noted, a possible way to generalize the $w_{\vec{x}}$'s can be found, appealing to Theorem 2.2.3: it is possible, indeed, to find a (unique) representation of a probability w that satisfies Ex, as a convex combination of simpler probabilities that satisfy Ex and IP.

The first proof of this result uses advanced methods of Measure Theory and Functional Analysis and it is due to Krauss ([15]); we will follow, instead, the argument of [20], Chapter 25. This allows us to make a detour through Nonstandard Analysis, showing, via a great example, how some tools of this branch of mathematics can be used to infer really concrete results.

The reader who is not familiar with Nonstandard Analysis will find all that he needs in Appendix B. Now, we will present, along the lines of [4], some basics of Loeb Measure Theory, a field of study in the intersection between Nonstandard Analysis and Measure Theory, that plays a central role in some applications of nonstandard methods in probability (as in the case of Theorem 2.4.4).

We would like to investigate the possibility of integrating a bounded function $F : \Omega \to {}^*\mathbb{R}$ and what is the link between the outcome of this integral and the (standard) integration of its standard part $^\circ F$, i.e. the function defined in a pointwise way taking for any $\omega \in \Omega$ the standard part of $F(\omega)$. To talk about integration, we first need to illustrate what is a measure in the nonstandard setting. In the following, we work in an enlargement $^* : \mathbb{U}(\mathbb{X}) \to \mathbb{U}'$ in which $\mathbb{X}$ is big enough to include $\mathbb{R}$ and all the other things we need along the path. We suppose also that the enlargement is $\aleph_1$-saturated, i.e. any countable family $\{A_n\}_{n\in\mathbb{N}}$ of internal sets with the finite intersection property has a non-empty intersection.

If we want to link, by Transfer Principle, the notion of nonstandard integration we are about to define and the standard counterpart, we should involve internal sets. In doing this, however, we will have some problems to deal with. If $\mathcal{B}$ is an algebra in $\mathbb{U}$, then ${}^*\mathcal{B}$ will be an algebra too, because of the Transfer Principle.[12] But, in a $\aleph_1$-saturated enalrgement, even if $\mathcal{B}$ is a σ-algebra, ${}^*\mathcal{B}$ may be not. Indeed, if we have a countable family $\{A_n\}_{n\in\mathbb{N}}$ of elements of ${}^*\mathcal{B}$, then the union of the family (that we will denote by A) may not be internal and hence, it may not belong to ${}^*\mathcal{B}$; indeed, if A was internal, by saturation, there would be $m \in \mathbb{N}$ such that

$$\bigcup_{n\leq m} A_n = A.^{13}$$

[12]For instance, the closure under finite union can be described by the following $\mathcal{L}_{\mathbb{U}}$-formula

$$\forall X \in \mathcal{B}\ \forall Y \in \mathcal{B}\ \exists Z \in \mathcal{B}\ (\forall x \in X\ x \in Z) \wedge (\forall y \in Y\ y \in Z) \wedge (\forall z \in Z\ (z \in X \vee z \in Y)).$$

Therefore, since it holds in $\mathbb{U}$, by the transfer principle, it holds also in $\mathbb{U}'$ with regards to ${}^*\mathcal{B}$.

[13] Indeed if we suppose that A is internal and that for every $m \in \mathbb{N}$, $\cup_{n\leq m} A_n \neq A$, we have that also

Therefore if, for instance, the family of A_n's is strictly increasing, this can't happen and the union can't be internal and, therefore, in ${}^*\mathcal{B}$.

Example 2.3.1. Consider $\mathbb{N} \in \mathbb{U}$ with the σ-algebra $\mathcal{B} = \mathcal{P}(\mathbb{N})$. Even if ${}^*\mathcal{B} = {}^*\mathcal{P}(\mathbb{N})$ is still an algebra, it is not a σ-algebra: indeed, for any $n \in \mathbb{N}$ the set $\{n\}$ is an internal subset of ${}^*\mathbb{N}$ (so it belongs to ${}^*\mathcal{B}$), but the union of these sets is $\bigcup_{n\in\mathbb{N}}\{n\} = \mathbb{N}$, that is not in ${}^*\mathcal{B}$, because by Theorem B.0.1 ${}^*\mathcal{P}(\mathbb{N})$ contains only internal subsets of ${}^*\mathbb{N}$ and $\mathbb{N}$ is not internal.

Remark 2.3.1. One may wonder why, by transfer principle, we don't get the closure under countable union in ${}^*\mathcal{B}$. This property can be described in $\mathbb{U}$ by the following $\mathcal{L}_{\mathbb{U}}$-formula:

$$\forall f : \mathbb{N} \to \mathcal{B}\ \exists A \in \mathcal{B}\ (\forall n \in \mathbb{N}\ \forall x \in f(n)\ x \in A) \wedge (\forall x \in A\ \exists n \in \mathbb{N}\ x \in f(n)).$$

Using the transfer principle, we get a closure under union of internal families of the form $\{A_n\}_{n\in{}^*\mathbb{N}}$ (where an internal functions $f : {}^*\mathbb{N} \to {}^*\mathcal{B}$ is implicit), that is different from the closure under countable union (as the example above showed): this property will be essential to carry in the nonstandard setting the results from the standard context and, as we will clarify later, ${}^*\mathcal{B}$ will be called a ${}^*\sigma$-algebra.

The argument above should explain why we will focus first on algebras or finitely-additive probabilities and not on σ-algebras or (σ-additive) probabilities.

Definition 2.3.1. Let $\Omega \in \mathbb{U}'$ be an internal set. An *internal algebra* is an internal set $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ that is an algebra and that is composed of internal subsets of Ω.

Given a measurable space $(\Omega, \mathcal{A})$ with $\mathcal{A}$ an internal algebra, a *finite nonstandard finitely-additive measure* on it is an internal map $M : \mathcal{A} \to {}^*[0, +\infty)$ such that:

- $M(\emptyset) \neq M(\Omega)$;
- for every disjoint elements A, B of $\mathcal{A}$, $M(A \cup B) = M(A) + M(B)$, where the $+$ is intended to be the nonstandard counterpart of the standard sum operation;
- $M(\Omega)$ is a bounded hyperreal.

We talk about *probability* when $M(\Omega) = 1$.

In the following, when we write $(\Omega, \mathcal{A}, M)$, we suppose that $(\Omega, \mathcal{A})$ is a measurable space with $\mathcal{A}$ internal algebra and M a finite nonstandard finitely-additive measure on it if not otherwise specified.

---

the family $\{B_m\}_{m\in\mathbb{N}}$ is composed of internal sets, where $B_m = A \setminus \cup_{n\leq m} A_n$. This succession of sets is decreasing and satisfies the finite intersection property since

$$\bigcap_{m\leq k} B_m = B_k = A \setminus \cup_{n\leq k} A_n \neq \emptyset.$$

Therefore the intersection of all the B_m's must be non-empty, but

$$\emptyset \neq \cap_{m\in\mathbb{N}} B_m = \emptyset.$$

Starting from $(\Omega, \mathcal{A}, M)$, we can define the standard part ${}^{\circ}\!M : \mathcal{A} \to [0, +\infty)$ of M as ${}^{\circ}\!M(A) = {}^{\circ}(M(A))$, for every $A \in \mathcal{A}$. Notice that ${}^{\circ}(M(A))$, the standard part of $M(A)$, exists because the finiteness of M guarantees that $M(A)$ is bounded.

It's easy to show that ${}^{\circ}\!M$ is a (standard) finitely-additive finite measure (this result relies on the fact that the standard part of a sum is the sum of the standard parts) on the algebra $\mathcal{A}$. We can go further, because of the following lemma.

Lemma 2.3.1. *With the notation used before, the measure ${}^{\circ}\!M$ is conditionally σ-additive in $\mathcal{A}$.*

Proof. Suppose to have a family of disjoint sets $\{A_n\}_{n\in\mathbb{N}}$ of elements of the algebra whose union is still in the algebra. We have already shown (see Footnote 13, page 48) that, by saturation, there exists $m \in \mathbb{N}$ for which $\cup_{n\in\mathbb{N}} A_n = \cup_{n\leq m} A_n$. This means that for every $k > m$, $A_k = \emptyset$, because the A_n's are disjoint.

Hence, by the finitely-additivity of ${}^{\circ}\!M$, we have

$$ {}^{\circ}\!M(\cup_{n\in\mathbb{N}} A_n) = {}^{\circ}\!M(\cup_{n\leq m} A_n) = \sum_{n\leq m} {}^{\circ}\!M(A_n) \stackrel{*}{=} \sum_{n\in\mathbb{N}} {}^{\circ}\!M(A_n), $$

where in $\stackrel{*}{=}$, we have used that ${}^{\circ}\!M(\emptyset) = 0$. □

Hence, by the Extension Theorem A.1.1, starting from ${}^{\circ}\!M$, we get an extension $\tilde{M}$ of it, defined on the σ-algebra $\mathcal{F}(\mathcal{A})$ generated by $\mathcal{A}$.

Actually, we will prove that there exists a measure extending $\tilde{M}$ on a larger, and more suitable, set of subsets of Ω. We have already argued that if we have a family of disjoint sets $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{A}$, then there is no reason for which their union A should be in $\mathcal{A}$. However, we can show that A can be arbitrarily well "approximated" in $\mathcal{A}$.

Proposition 2.3.1. *With the notation used before, if $\{A_n\}_{n\in\mathbb{N}}$ is a family of sets in $\mathcal{A}$, then there is a set $B \in \mathcal{A}$ such that:*

- $\cup_{n\in\mathbb{N}} A_n \subseteq B$;
- ${}^{\circ}\!M(B) \leq \sum_{n\in\mathbb{N}} {}^{\circ}\!M(A_n)$;
- *if the A_n's are pairwise disjoint, then*

$$ {}^{\circ}\!M(B) = \sum_{n\in\mathbb{N}} {}^{\circ}\!M(A_n). $$

Proof. For $\alpha = \sum_{n\in\mathbb{N}} {}^{\circ}\!M(A_n) = +\infty$, it suffices to take $B = \Omega$: in this way the first two items of the statement are trivially satisfied. For the last one, we show that it cannot happen under the hypothesis $\alpha = +\infty$: notice that for any $m \in \mathbb{N}$, if the A_n's are pairwise disjoint,

$$ \sum_{n\leq m} {}^{\circ}\!M(A_n) = {}^{\circ}\!M(\bigcup_{n\leq m} A_n) \leq {}^{\circ}\!M(\Omega). $$

Therefore, since M is finite, so is ${}^{\circ}\!M(\Omega)$, and α can't be infinite.

Suppose now that α is finite and that the A_n's are pairwise disjoint. Taking $B_m = \cup_{n\leq m} A_n$, we have that for every m

$$M(B_m) = \sum_{n\leq m} M(A_n) \leq \sum_{n\leq m} {}^\circ\!M(A_n) + \frac{1}{m+1} \leq \alpha + \frac{1}{m+1}. \tag{2.14}$$

Consider now an internal extension $\{A_n\}_{n\in{}^*\mathbb{N}}$ in $\mathcal{A}$ of the family $\{A_n\}_{n\in\mathbb{N}}$[14], and extend the definition of B_m's also for nonstandard m's, so that $B_N = \cup_{n\leq N} A_n$, for $N \in {}^*\mathbb{N}\setminus\mathbb{N}$. The formula

$$\varphi(x) := B_x \in \mathcal{A} \wedge M(B_x) \leq \alpha + \frac{1}{x+1}$$

is internal (since M, α, $\mathcal{A}$ and the function that maps $x \in {}^*\mathbb{N}$ to B_x are internal) and valid for all $m \in \mathbb{N}$ as the closure of $\mathcal{A}$ under boolean operations and Equation (2.14) show: therefore, by the Overflow Principle (see page 110), there is a nonstandard $N \in {}^*\mathbb{N}$ for which

$$B_N \in \mathcal{A} \wedge M(B_N) \leq \alpha + \frac{1}{N+1}. \tag{2.15}$$

This B_N approximates well enough A in $\mathcal{A}$, since:

- $B_N \in \mathcal{A}$;
- $A \subseteq B_N$, because N is nonstandard and $B_N = \cup_{n\leq N} A_n$;
- ${}^\circ\!M(B_N) = \alpha$: indeed, ${}^\circ\!M(B_N) \leq \alpha$ follows by taking the standard part in the second conjunct of Formula (2.15); the converse disequality holds because for every $m \in \mathbb{N}$
$$\sum_{n\leq m} {}^\circ\!M(A_n) = {}^\circ\!M(\bigcup_{n\leq m} A_n) = {}^\circ\!M(B_m) \leq {}^\circ\!M(B_N).$$

The last case we have to consider in order to conclude the proof is when $\alpha < +\infty$ and the A_n's are not pairwise disjoint. In this case, we can build the family of disjoint sets $\{C_n\}_{n\in\mathbb{N}}$ defined by

$$C_n = A_n \setminus (\bigcup_{i<n} A_i).$$

Applying the argument above for the case in which the A_n's were disjoint, now with regards to the family of C_n's, we get a set B such that:

- $\bigcup_{n\in\mathbb{N}} C_n \subseteq B$;
- ${}^\circ\!M(B) \leq \sum_{n\in\mathbb{N}} {}^\circ\!M(C_n)$.

The thesis follows noticing that:

- $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} C_n$, hence $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} C_n \subseteq B$;
- for any $m \in \mathbb{N}, \bigcup_{n\leq m} A_n = \bigcup_{n\leq m} C_n$, therefore,
$$\sum_{n\leq m} {}^\circ\!M(A_n) \geq {}^\circ\!M\Big(\bigcup_{n\leq m} A_n\Big) = {}^\circ\!M\Big(\bigcup_{n\leq m} C_n\Big) = \sum_{n\leq m} {}^\circ\!M(C_n) \geq {}^\circ\!M(B).$$

[14]This is possible thanks to the Sequential Comprehensiveness (see page 110) of the enlargement we are considering, applied to the function $f : \mathbb{N} \to \mathcal{A}$ that maps every $n \in \mathbb{N}$ to A_n.

□

The notion of "arbitrary well approximation" of a set is formalized in the following definition.

Definition 2.3.2. Let Ω be an internal set, $\mathcal{A}$ an internal algebra on it and M a finitely-additive, finite and internal measure $M : \mathcal{A} \to {}^*[0, +\infty)$. A set $B \subseteq \Omega$ (not necessarily internal) is *Loeb null*[15] if for every real $\varepsilon > 0$ there is a set $A \in \mathcal{A}$ that contains B and $M(A) < \varepsilon$.

A set $B \subseteq \Omega$ (not necessarily internal) is *Loeb measurable* if there is $A \in \mathcal{A}$ such that $A \Delta B$ is Loeb null. The set of all the Loeb measurable sets will be denoted by $L(\mathcal{A})$.

Before going on, we can notice some useful properties about Loeb null sets. In particular:

Proposition 2.3.2. *With the notation used above, the following hold:*

a) a Loeb null set $B \in \mathcal{A}$ is such that $M(B) \approx 0$;

b) a subset of a Loeb null set is Loeb null;

c) the union of a countable family of disjoint sets $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{A}$ is in $L(\mathcal{A})$; furthermore, a witness $B \in \mathcal{A}$ of this membership (i.e. a B such that $(\cup_{n\in\mathbb{N}} A_n)\Delta B$ is Loeb null) can be chosen with the property ${}^\circ\! M(B) = \sum_{n\in\mathbb{N}} {}^\circ\! M(A_n)$;

d) the union of a countable family of Loeb null sets is also Loeb null.

Proof. a) if B is Loeb null, for any real $\varepsilon > 0$ there is a set $A \in \mathcal{A}$ with $B \subseteq A$ and $M(A) < \varepsilon$. Since $B \in \mathcal{A}$, by the properties of the measure M, we have

$$M(B) \leq M(A) < \varepsilon$$

, and the thesis follows from the generality of $\varepsilon \in \mathbb{R}^+$.

b) let B be Loeb null, and for every real $\varepsilon > 0$, let $A_\varepsilon \in \mathcal{A}$ be such that $B \subseteq A_\varepsilon$ and $M(A_\varepsilon) < \varepsilon$. Then, for every $C \subseteq B$, the family $\{A_\varepsilon\}_{\varepsilon>0}$ continues to satisfy the condition required, and C is Loeb null.

c) let $\{A_n\}_{n\in\mathbb{N}}$ be a familiy of disjoint elements in $\mathcal{A}$ and let A be their union: we will prove that there exists $B \in \mathcal{A}$ for which $A\Delta B$ is Loeb null. Indeed, it is sufficient to take $B \in \mathcal{A}$ as in Proposition 2.3.1: in this case, $A\Delta B = B \setminus A$ and ${}^\circ\! M(B) = \sum_{n\in\mathbb{N}} {}^\circ\! M(A_n)$. Therefore, if we fix $\varepsilon > 0$, since ${}^\circ\! M(B)$ is finite, there exists $m \in \mathbb{N}$ such that $\sum_{n>m} {}^\circ\! M(A_n) < \varepsilon$. This means that if we take $D_\varepsilon = B \setminus \cup_{n\leq m} A_n$, we have:

- $D_\varepsilon \in \mathcal{A}$ because it's a difference of elements in $\mathcal{A}$;
- $A\Delta B = B \setminus A \subseteq D_\varepsilon$;

[15] Notice that this definition depends on the algebra $\mathcal{A}$ and on the measure M involved. However, since, in the following, when we refer to this notion, these objects will be clear from the context, we will omit the dependence. The same holds for the *Loeb measurable* definition.

– $M(D_\varepsilon) < \varepsilon$, because

$$\begin{aligned} M(B \setminus \cup_{n\le m} A_n) &= M(B) - M(\cup_{n\le m} A_n) \\ &\approx \sum_{n\in\mathbb{N}} {}^\circ M(A_n) - \sum_{n\le m} {}^\circ M(A_n) \\ &= \sum_{n>m} {}^\circ M(A_n) < \varepsilon. \end{aligned}$$

From the arbitrarity of ε, we have that $A\Delta B$ is Loeb-null.

d) consider the countable family of Loeb null sets $\{A_n\}_{n\in\mathbb{N}}$. Given $\varepsilon > 0$, we can take a family $\{B_{n,\varepsilon}\}_{n\in\mathbb{N}}$ of elements in $\mathcal{A}$ such that for every n

$$A_n \subseteq B_{n,\varepsilon} \text{ and } M(B_{n,\varepsilon}) < \frac{\varepsilon}{2^{n+1}}.$$

By applying Proposition 2.3.1 to the family $\{B_{n,\varepsilon}\}_{n\in\mathbb{N}}$, we get a set B_ε such that $B_\varepsilon \supseteq \cup_{n\in\mathbb{N}} B_{n,\varepsilon} \supseteq \cup_{n\in\mathbb{N}} A_n$ and

$${}^\circ M(B_\varepsilon) \le \sum_{n\in\mathbb{N}} {}^\circ M(B_{n,\varepsilon}) < \varepsilon \sum_{n\in\mathbb{N}} 2^{-n-1} = \varepsilon.$$

For the arbitrarity of ε, we have just shown that a countable union of Loeb null sets is Loeb null.

□

The proposition above is useful to further investigate the structure of $L(\mathcal{A})$.

Proposition 2.3.3. *With the notation used above, the set $L(\mathcal{A})$ is a σ-algebra that contains $\mathcal{A}$.*

Proof. The inclusion $\mathcal{A} \subseteq L(\mathcal{A})$ follows from the fact that for any $B \in \mathcal{A}$ the requirement in Definition 2.3.2 is trivially satisfied by taking B as the set A, since $B\Delta B = \emptyset$ which is Loeb null.

From this inclusion, it follows that $\emptyset, \Omega \in L(\mathcal{A})$. Furthermore:

- $L(\mathcal{A})$ is closed under finite union, since if A, B are in $L(\mathcal{A})$, then there exist $C, D \in \mathcal{A}$ such that $A\Delta C$ and $B\Delta D$ are Loeb null. The thesis follows by noticing that

$$(A \cup B)\Delta(C \cup D) \subseteq (A\Delta C) \cup (B\Delta D)^{16},$$

that $C \cup D \in \mathcal{A}$ and that union of Loeb null sets is Loeb null;

[16] For every non-empty family of index I it can be shown that

$$(\bigcup_{i\in I} A_i)\Delta(\bigcup_{i\in I} B_i) \subseteq \bigcup_{i\in I} A_i \Delta B_i.$$

Indeed, let a be an element in $\cup_{i\in I} A_i \setminus \cup_{i\in I} B_i$; then there exists $j \in I$ such that $a \in A_j$. The element a can't belong to B_j because if so, it would be in $\cup_{i\in I} B_i$, therefore, $a \in A_j \setminus B_j$. Analogously it can be shown that $\cup_{i\in I} B_i \setminus \cup_{i\in I} A_i$ is a subset of $\cup_{i\in I} B_i \setminus A_i$, and combining the two results, we obtain what we want.

- $L(\mathcal{A})$ is closed under complementation, because $\mathcal{A}$ is. Indeed, if we have $A \in L(\mathcal{A})$, then there's $B \in \mathcal{A}$ such that $A \Delta B$ is Loeb null. Therefore, $A^c \Delta B^c$ is Loeb null since it is equal to $A \Delta B$, and $B^c \in \mathcal{A}$.

- $L(\mathcal{A})$ is closed under countable union: indeed, suppose that $\{A_n\}_{n\in\mathbb{N}}$ is a family of sets in $L(\mathcal{A})$ and $\{B_n\}_{n\in\mathbb{N}}$ is a family in $\mathcal{A}$ such that $A_n \Delta B_n$ is Loeb null for every $n \in \mathbb{N}$. Calling $A := \cup_{n\in\mathbb{N}} A_n$ and $B := \cup_{n\in\mathbb{N}} B_n$ we have that $A \Delta B$ is Loeb null since
$$A \Delta B = (\bigcup_{n\in\mathbb{N}} A_n) \Delta (\bigcup_{n\in\mathbb{N}} B_n) \subseteq \bigcup_{n\in\mathbb{N}} (A_n \Delta B_n)^{17},$$
and a countable union of Loeb null sets is Loeb null. Following the argument presented in item c) of Proposition 2.3.2, even for a non-disjoint family of sets like $\{B_n\}_{n\in\mathbb{N}}$, we can show that there exists $C \in \mathcal{A}$ such that $B \Delta C$ is Loeb null. Thus, A is in $L(\mathcal{A})$ because
$$A \Delta C \subseteq (A \Delta B) \cup (B \Delta C)$$
and the right-hand-side is Loeb null.

□

We now have all that is necessary to talk about Loeb measure.

Definition 2.3.3. With the notation above, we can define the function $M_L : L(\mathcal{A}) \to [0, +\infty)$ by
$$M_L(B) = {}^\circ M(A),$$
where A is a set in $\mathcal{A}$ for which $A \Delta B$ is Loeb null.

This is a good definition, in the sense that it doesn't depend on the choice of A: if A_1, A_2 are two sets for which $A_1 \Delta B$ and $A_2 \Delta B$ are Loeb null, then $M(A_1)$ and $M(A_2)$ have the same standard part. In order to show so, we can notice that:

- $A_1 \Delta A_2$ is Loeb null, since it is contained in $(A_1 \Delta B) \cup (A_2 \Delta B)$ that is Loeb null as a union of Loeb null sets.

- $M(A_1) \approx M(A_1 \cap A_2) \approx M(A_2)$, because, since M is finitely-additive and finite,
$$M(A_1) = M(A_1 \setminus A_2) + M(A_1 \cap A_2).$$
Taking the standard part and recalling that for being $A_1 \setminus A_2 \subseteq A_1 \Delta A_2$, $A_1 \setminus A_2$ is Loeb null, we have by Proposition 2.3.2 item a) that
$${}^\circ M(A_1) = {}^\circ M(A_1 \setminus A_2) + {}^\circ M(A_1 \cap A_2) = {}^\circ M(A_1 \cap A_2).$$
Applying this argument to A_2 instead of A_1, we have the thesis.

This function provides us with a very useful tool for detecting Loeb null sets.

Lemma 2.3.2. *With the definitions and notations above, a subset B of Ω is Loeb null if and only if $M_L(B) = 0$.*

[17] See Footnote 16, page 53.

Proof. If B is Loeb null, then $A = \emptyset$ is such that $A\Delta B$ is Loeb null so that $M_L(B) = {}^{\circ}M(\emptyset) = 0$; conversely, if $M_L(B) = 0$, then there exists $A \in \mathcal{A}$ for which $A\Delta B$ is Loeb null and ${}^{\circ}M(A) = 0$. Hence, A is Loeb null (because $M(A) < \varepsilon$ for any real $\varepsilon > 0$) and so is B, since $B \subseteq (A\Delta B)\cup A$ and the Loeb null sets are closed under union and subsets. □

The domain of M_L is a σ-algebra and we can prove that this function inherits from M all that it is needed to be a measure.

Proposition 2.3.4. *Let Ω be an internal set, $\mathcal{A}$ an internal algebra on it, $M : \mathcal{A} \to {}^*[0, +\infty)$ an internal finitely-additive finite measure and $\tilde{M}$ the extension of ${}^{\circ}M$ to $\mathcal{F}(\mathcal{A})$ (see page 50). Then the function $M_L : L(\mathcal{A}) \to [0, +\infty)$ defined above is a finite σ-additive measure that extends $\tilde{M}$.*

Proof. First we prove that M_L is a finite measure on the σ-algebra $L(\mathcal{A})$:

- M_L coincides with ${}^{\circ}M$ for the sets in $\mathcal{A}$, so, since $\emptyset, \Omega \in \mathcal{A}$, we have
$$M_L(\Omega) = {}^{\circ}M(\Omega) \neq {}^{\circ}M(\emptyset) = M_L(\emptyset).$$
- it is σ-additive. Assume to have a family $\{B_n\}_{n\in\mathbb{N}}$ of pairwise disjoint elements of $L(\mathcal{A})$ and a family $\{A_n\}_{n\in\mathbb{N}}$ in $\mathcal{A}$ for which $A_n\Delta B_n$ is Loeb null for every $n \in \mathbb{N}$ and, therefore,
$$M_L(B_n) = {}^{\circ}M(A_n).$$
By cutting off the intersections, it is easy to show that we can assume that also the A_n's are pairwise disjoint. By item c) of Proposition 2.3.2, we have a set $C \in \mathcal{A}$ such that $(\cup_{n\in\mathbb{N}}A_n)\Delta C$ is Loeb null and
$${}^{\circ}M(C) = \sum_{n\in\mathbb{N}} {}^{\circ}M(A_n).$$
Hence, since $C\Delta(\cup_{n\in\mathbb{N}}B_n)$ is Loeb null (since it is contained in the union of the Loeb null sets $(\cup_{n\in\mathbb{N}}A_n)\Delta(\cup_{n\in\mathbb{N}}B_n)$ and $(\cup_{n\in\mathbb{N}}A_n)\Delta C$),
$$M_L(\bigcup_{n\in\mathbb{N}} B_n) = {}^{\circ}M(C) = \sum_{n\in\mathbb{N}} {}^{\circ}M(A_n) = \sum_{n\in\mathbb{N}} M_L(B_n).$$
- it is finite, since ${}^{\circ}M$ is so and $M_L(\Omega) = {}^{\circ}M(\Omega)$;

We conclude the proof, noticing that $L(\mathcal{A}) \supseteq \mathcal{F}(\mathcal{A})$ by the minimality of the latter and that M_L coincides with ${}^{\circ}M$ in the domain where both are defined. Hence, via the unicity guaranteed by Extension Theorem A.1.1, M_L and $\tilde{M}$ must coincide on $\mathcal{F}(\mathcal{A})$. □

Let's show an example of a Loeb measure construction.

Example 2.3.2. Given a hyperfinite set Ω with cardinality $N \in {}^*\mathbb{N} \setminus \mathbb{N}$, we can define the algebra $\mathcal{A}$ of all the internal subsets of Ω and the measure M that maps any internal (and then hyperfinite) $B \subseteq \Omega$ in $|B|/|\Omega|$, where $|-|$ denote the internal cardinality of a set. It can be shown that this is an internal finitely-additive nonstandard measure.
Since any $\omega \in \Omega$ is internal, also all the sets $\{\omega\}$ are internal. Therefore, M is defined on these sets and it assumes for each one the same value $(= 1/|\Omega|)$. The corresponding algebra $L(\mathcal{A})$ is a σ-algebra on Ω and the Loeb measure M_L (usually called *Loeb counting measure*) assigns to any singleton the value ${}^{\circ}(1/|\Omega|) = 0$.

To recap, starting from $(\Omega, \mathcal{A}, M)$ where M is an internal, finitely-additive, finite nonstandard measure on $\mathcal{A}$, we get the *Loeb measure space* $(\Omega, L(\mathcal{A}), M_L)$, where $M_L : L(\mathcal{A}) \to [0, +\infty)$ is called the *Loeb measure* given by M. We will talk about *Loeb probabilities* if $M(\Omega) = M_L(\Omega) = 1$.

Given a Loeb measure space $(\Omega, L(\mathcal{A}), M_L)$ we can extend the usual notion of measurable function. A function $f : \Omega \to \mathbb{R}$ is *Loeb measurable* if for any open A in $\mathbb{R}$, $f^{-1}(A) \in L(\mathcal{A})$; an internal function $F : \Omega \to {}^*\mathbb{R}$ is $\mathcal{A}$-*measurable* if for any $r \in {}^*\mathbb{R}$, $F^{-1}(\{x \in {}^*\mathbb{R} : x < r\}) \in \mathcal{A}$. If F is pointwise bounded, i.e. for any $\omega \in \Omega$, there is a natural number $n \in \mathbb{N}$ such that $|F(\omega)| < n$, then we can define the function ${}^\circ F : \Omega \to \mathbb{R}$ that maps any $\omega \in \Omega$ to ${}^\circ(F(\omega))$.

In some cases, by Transfer Principle, we can give sense to an integration of a $\mathcal{A}$-measurable function $F : \Omega \to {}^*\mathbb{R}$ and we can relate it to the standard integral of ${}^\circ F$ (see Theorem 2.3.3 below). In this regard, we notice that the property of being a σ-algebra is definable in $\mathcal{L}_{\mathbb{U}}$: we can use the already mentioned sentence to represent the closure under countable unions of $\mathcal{C} \in \mathbb{U}$

$$\forall f : \mathbb{N} \to \mathcal{C}\ \exists B \in \mathcal{C}\ (\forall n \in \mathbb{N}\ \forall x \in f(n)\ x \in B) \wedge (\forall x \in B\ \exists n \in \mathbb{N}\ x \in f(n)).$$

If the $*$-transform of this formula (this differs from the closure under countable unions, as already discussed in Remark 2.3.1) holds for an algebra $\mathcal{A} \in \mathbb{U}'$, then we will say that $\mathcal{A}$ is a ${}^*\sigma$-*algebra*, i.e. an algebra that is also closed for unions of internal families of the form $\{A_n\}_{n\in{}^*\mathbb{N}}$. [18]

In the same way, we can define the notion of a *measure M: this would be an internal map from a ${}^*\sigma$-algebra $\mathcal{A}$ to ${}^*[0, +\infty)$ such that:

- $M(\emptyset) \neq M(\Omega)$;
- for A, B disjoint element of $\mathcal{A}$, $M(A \cup B) = M(A) + M(B)$;
- for any internal family $\{A_n\}_{n\in{}^*\mathbb{N}}$ whose elements are pairwise disjoint $M(\bigcup_{n\in{}^*\mathbb{N}} A_n)$ $= \sum_{n\in{}^*\mathbb{N}} M(A_n)$. [19]

In the standard setting when we have a measurable space and a measure on it, we have also a clear notion of integrability: we first define the integral for simple functions, then we extend for positive and general functions as explained in Appendix A. In the following, we will suppose that $\mathcal{A}$ is a ${}^*\sigma$-algebra and M is a *measure on it that is finite, i.e. such that there exists a bounded hyperreal K such that $M(\Omega) < K$; we can suppose without loss of generality that the bound K is actually real.

[18] Notice that not necessarily a ${}^*\sigma$-algebra is a σ-algebra: see Example 2.3.1.

[19] This is the transfer of the σ-additivity of a measure μ on a σ-algebra $\mathcal{C}$ in $\mathbb{U}$; the latter property can be formalized by the following $\mathcal{L}_{\mathbb{U}}$-formula:

$$\forall f : \mathbb{N} \to \mathcal{C}\ (\mathrm{Disj}(f) \to (\exists B \in \mathcal{C}\ \ \mathrm{Union}(B, f) \wedge \mu(B) = \sum_{n\in\mathbb{N}} \mu(f(n))),$$

where

$$\mathrm{Disj}(f) := \forall n \in \mathbb{N}\ \forall m \in \mathbb{N}\ (n \neq m \to (\forall x \in f(n)\ x \notin f(m) \wedge \forall x \in f(m)\ x \notin f(n))$$
$$\mathrm{Union}(B, f) := (\forall n \in \mathbb{N}\ \forall x \in f(n)\ x \in B) \wedge (\forall x \in B\ \exists n \in \mathbb{N}\ x \in f(n)).$$

For details about the meaningfulness of the sum involved, see page 106, item i) and item l).

In a nonstandard setting, we will work with internal functions $F : \Omega \to {}^*\mathbb{R}$ that are $\mathcal{A}$-measurable: notice that this notion of measurability is the $*$-transform of the same notion in a standard context. Indeed, when $\mathcal{A}$ is a ${}^*\sigma$-algebra, by Transfer, the definition given is equivalent to asking that for any $A \in {}^*\tau$, where τ is the standard topology on $\mathbb{R}$, the preimage $F^{-1}(A)$ is an element of $\mathcal{A}$.[20]

Hence, we can perform in a nonstandard framework most of the Measure Theory arguments that are usually applied to measurable functions and extend definitions and properties (as the integral ones) nonstandardly. We will do this in the following, in order to justify integrals of nonstandard functions.

A *simple function will be an internal $\mathcal{A}$-measurable function $s : \Omega \to {}^*\mathbb{R}$ such that there exists $N \in {}^*\mathbb{N}$ with

$$|\mathrm{Ran}(s)| = |\{r \in {}^*\mathbb{R} \ : \ \exists \omega \in \Omega \ s(\omega) = r\}| = N;$$

the integral of a simple function s is $\sum_{r\in \mathrm{Ran}(s)} r \cdot M(s^{-1}(\{r\}))$.[21]

We notice that the set of all the internal functions from an internal set Ω to ${}^*\mathbb{R}$ is internal: this is because of Theorem B.0.2, and because it is definible by an $\mathcal{L}_{\mathbb{U}'}$-formula (the one that defined when a relation is a function) in the set ${}^*\mathcal{P}(\Omega \times {}^*\mathbb{R})$; for the same reason, the following are internal as well: the set A of all the simple functions, the function $A \to {}^*\mathbb{R}$ that maps any simple map s to its integral $\int s \, dM$, with M an internal *measure[22] and the set

$$B_F := \{r \in {}^*\mathbb{R} \ : \ \exists s \in A \ \int s \, dM = r \wedge s \leq F\},$$

where $F : \Omega \to {}^*\mathbb{R}$ is a nonstandard $\mathcal{A}$-measurable function.

Since B_F is internal, then, if it is bounded above by a $K \in {}^*\mathbb{R}$, it has a least upper bound (by Transfer Principle) and we can define the integral $\int F \, dM$ of F to be this hyperreal. By Theorem B.0.2 also the function $F \mapsto \int F \, dM$ is internal if M is an internal *measure and, by Transfer, it inherits some properties from the standard counterpart.

So, we can give sense to the integration of a nonstandard $\mathcal{A}$-measurable function $F : \Omega \to {}^*\mathbb{R}$. We will focus, as already mentioned, on finite *measures M and uniformly bounded functions F (i.e. there exists $K \in \mathbb{R}$ such that for all $\omega \in \Omega$, $|F(\omega)| < K$): in this case, the set B_F is bounded and the integral of F with regards to M is a bounded hyperreal denoted by $\int F \, dM$.

---

[20]Indeed, in the standard setting for any σ-algebra $\mathcal{C} \subseteq \mathcal{P}(X)$ and for any function $G : X \to \mathbb{R}$, the following are equivalent:

- for any open $O \in \tau$, $G^{-1}(O) \in \mathcal{C}$, where τ is the standard topology on $\mathbb{R}$;
- for any real $r \in \mathbb{R}$, $G^{-1}((-\infty, r)) \in \mathcal{C}$.

Since this property is definable as a $\mathcal{L}_{\mathbb{U}}$-formula, we get the thesis.

[21]Notice that this sum is well-defined since the range of a simple function is hyperfinite.

[22]Given an internal *measure M on $(\Omega, \mathcal{A})$, we can consider the set

$$\{(s,r) \in A \times {}^*\mathbb{R} \ : \ \exists N \in {}^*\mathbb{N} \ |\mathrm{Ran}(s)| \overset{g}{=} N \wedge \sum_{n\leq N} g(n) \cdot M(s^{-1}(g(n))) = r\},$$

where the sum is to be intended as in the items i) and l) of page 106, and "$|\mathrm{Ran}(s)| \overset{g}{=} N$" means that there is a bijection $g : \{1, \dots, N\} \to \mathrm{Ran}(s)$. Then, the issue set is the function that to any *simple function associates its integral and is $\mathcal{L}_{\mathbb{U}'}$-definable with constants that are internal, so it is internal by Theorem B.0.2.

In the case of uniformly bounded F, we can consider also the real-valued function $°F : \Omega \to \mathbb{R}$. Since M is a finite *measure, it is a finite nonstandard finitely-additive measure (see Definition 2.3.1) and we have a real-valued measure on $(\Omega, \mathcal{A})$, the Loeb measure M_L associated to it. Hence, we have the hyperreal $\int F\ dM$ and its real "counterpart" $\int °F\ dM_L$. The link between these numbers will be clarified by the following theorem which is the core of Loeb Measure Theory.

Theorem 2.3.3. *With the notation and hypothesis assumed above, if $F : \Omega \to {}^*\mathbb{R}$ is a uniformly bounded, internal, and $\mathcal{A}$-measurable function, $\mathcal{A}$ is a $^*\sigma$-algebra and M is a finite *measure, then F is *integrable, $°F$ is* Loeb measurable *and*

$$°\left(\int_\Omega F\ dM\right) = \int_\Omega °F\ dM_L.$$

Proof. Since F is $\mathcal{A}$-measurable,

$$F^{-1}(\{x \in {}^*\mathbb{R}\ :\ x < r\}) = \{\omega \in \Omega\ :\ F(\omega) < r\} \in \mathcal{A}$$

for any $r \in \mathbb{R}$. Then, recalling that $L(\mathcal{A})$ is a σ-algebra, we will show that $°F$ is Loeb measurable:

- for any $a \in \mathbb{R}$, $\{\omega \in \Omega\ :\ °F(\omega) \leq a\} = \cap_{n\in\mathbb{N}}\{\omega \in \Omega\ :\ F(\omega) < a + \frac{1}{n}\}$, so this set is in $L(\mathcal{A})$, since $L(\mathcal{A})$ is closed under countable intersection;

- for any $a \in \mathbb{R}$, $\{\omega \in \Omega\ :\ °F(\omega) > a\} = \{\omega \in \Omega\ :\ °F(\omega) \leq a\}^c$, so it is in $L(\mathcal{A})$, since $L(\mathcal{A})$ is closed under complement;

- for any $b \in \mathbb{R}$, $\{\omega \in \Omega\ :\ °F(\omega) < b\} = \cup_{n\in\mathbb{N}}\{\omega \in \Omega\ :\ °F(\omega) \leq b - \frac{1}{n}\}$, so it is in $L(\mathcal{A})$, since $L(\mathcal{A})$ is closed under countable union;

- for any open interval $(a, b) \subseteq \mathbb{R}$, $\{\omega \in \Omega\ :\ °F(\omega) \in (a, b)\} = \{\omega \in \Omega\ :\ °F(\omega) > a\} \cap \{\omega \in \Omega\ :\ °F(\omega) < b\}$, so it is in $L(\mathcal{A})$, since $L(\mathcal{A})$ is closed under finite intersection;

- any open set $A \subseteq \mathbb{R}$ is a countable union of open intervals, hence, if $A = \cup_{n\in\mathbb{N}}(a_n, b_n)$,

$$\begin{aligned}\{\omega \in \Omega\ :\ °F(\omega) \in A\} &= \{\omega \in \Omega\ :\ °F(\omega) \in \bigcup_{n\in\mathbb{N}}(a_n, b_n)\}\\ &= \bigcup_{n\in\mathbb{N}}\{\omega \in \Omega\ :\ °F(\omega) \in (a_n, b_n)\}\end{aligned}$$

  and it is in $L(\mathcal{A})$, since $L(\mathcal{A})$ is closed under countable union.

This proves that $°F$ is Loeb measurable

Given an arbitrary real $\varepsilon > 0$, choose $m \in \mathbb{N}$ such that $|F(\omega)| < m\varepsilon$ for every $\omega \in \Omega$. For any $k \in \{-m, \dots, m\}$ consider

$$A_k^\varepsilon := \{w \in \Omega : k\varepsilon \leq F(\omega) < (k+1)\varepsilon\} :$$

these sets belong to $\mathcal{A}$ because F is $\mathcal{A}$-measurable.

Since the $\{A_k^\varepsilon\}$'s form a partition of Ω, we can define the following $\mathcal{A}$-measurable step functions:

$$F_1^\varepsilon := \sum_k k\varepsilon\chi_{A_k^\varepsilon} \quad F_2^\varepsilon := \sum_k (k+1)\varepsilon\chi_{A_k^\varepsilon};$$

since these functions are real-valued, ${}^\circ F_1^\varepsilon = F_1^\varepsilon$ and ${}^\circ F_2^\varepsilon = F_2^\varepsilon$.

F_1^ε and F_2^ε approximate F respectively from under and above; hence, we have

$$\sum_{k=-m}^{m} k\varepsilon M(A_k^\varepsilon) = \int_\Omega F_1^\varepsilon \, dM \leq \int_\Omega F \, dM \leq \int_\Omega F_2^\varepsilon \, dM = \sum_{k=-m}^{m} (k+1)\varepsilon M(A_k^\varepsilon).$$

Now we can notice that, since F_1^ε is $\mathcal{A}$-measurable and the A_k^ε's are in $\mathcal{A}$,

$$^\circ\!\left(\int_\Omega F_1^\varepsilon \, dM\right) = \sum_{k=-m}^{m} k\varepsilon\, {}^\circ\! M(A_k^\varepsilon) = \sum_{k=-m}^{m} k\varepsilon M_L(A_k^\varepsilon) = \int_\Omega {}^\circ F_1^\varepsilon \, dM_L; \tag{2.16}$$

the same argument holds for F_2^ε.

Hence, we have

$$a_\varepsilon := {}^\circ\!\left(\int_\Omega F_1^\varepsilon \, dM\right) \leq {}^\circ\!\left(\int_\Omega F \, dM\right) \leq {}^\circ\!\left(\int_\Omega F_2^\varepsilon \, dM\right) =: b_\varepsilon$$
$$a'_\varepsilon := \int_\Omega {}^\circ F_1^\varepsilon \, dM_L \leq \int_\Omega {}^\circ F \, dM_L \leq \int_\Omega {}^\circ F_2^\varepsilon \, dM_L =: b'_\varepsilon,$$

where the second line follows from ${}^\circ F_1^\varepsilon \leq {}^\circ F \leq {}^\circ F_2^\varepsilon$.

We have already shown that $a_\varepsilon = a'_\varepsilon$ thanks to Equation (2.16) and, similarly, it can be proved that $b_\varepsilon = b'_\varepsilon$; the thesis now follows from the fact that

$$b_\varepsilon - a_\varepsilon = {}^\circ\!\left(\int_\Omega F_2^\varepsilon - F_1^\varepsilon \, dM\right) = {}^\circ\!\left(\int_\Omega \varepsilon \, dM\right) = \varepsilon\, {}^\circ\! M(\Omega),$$

so, since ${}^\circ\! M(\Omega)$ is finite, that $b_\varepsilon - a_\varepsilon \to 0$ when ε approaches to 0. □

2.4 De Finetti Theorem: polyadic case

As previously announced, in this section, we will use the tools so far introduced to establish a polyadic version of Theorem 2.2.1.

In the following, we will consider a standard universe $\mathbb{U} = \mathbb{U}(\mathbb{X})$ over a set $\mathbb{X}$ that contains enough individuals to talk about $\mathrm{Sen}(\mathcal{L})$ and probabilities over it: for instance, all the real numbers, all the symbols that appear in $\mathcal{L}$ (relation symbols, constants, and variables), all the formulas in $\mathrm{Sen}(\mathcal{L})$, and all the other objects we need in the path leading to Theorem 2.4.4. However, since the proof of the theorem involves a generalization of probabilities in the nonstandard setting, we require that, for instance, Con or $\mathrm{Sen}(\mathcal{L})$ are not members of $\mathbb{X}$: doing this, as we will see, a set of new constants will appear, together with new formulas and new probabilities that can be studied.[23] We also suppose to have

[23] Actually, this is what we do also in Nonstandard Analysis when we consider $\mathbb{X} = \mathbb{R}$: all the reals are in $\mathbb{X}$ but $\mathbb{R} \notin \mathbb{X}$ and this allows us to have also nonstandard reals.

a $\aleph_1$-saturated enlargement $* : \mathbb{U}(\mathbb{X}) \to \mathbb{U}'$. We will adopt the usual notation that is used in this setting and that can be found in Appendix B.

If we have a state description $\Theta(b_1, \ldots, b_m)$ and we fix a r-ary relation symbol R, then Θ provides us with all the r-uple of elements (among $b_1, \ldots, b_m$) for which R holds. Hence, for a language $\mathcal{L} = \{R_1, \ldots, R_q\}$, where R_i is a relation symbol with ariety r_i, a state description for $b_1, \ldots, b_m$ can be seen as an element of

$$\mathcal{P}(\{b_1, \ldots, b_m\}^{r_1}) \times \mathcal{P}(\{b_1, \ldots, b_m\}^{r_2}) \times \cdots \times \mathcal{P}(\{b_1, \ldots, b_m\}^{r_q}),$$

whose projection to the i-th component is the set of all (and only) the r_i-uples for which R_i holds. Conversely, every element of the set above determines uniquely a state description.[24]

For instance, in the Example 2.1.1, we get that $\Theta(a, b)$ can be seen as the pair

$$A_\Theta := (\{a\}, \{(a, a), (b, a), (b, b)\}).$$

Conversely, to the pair $(\{a, b\}, \{(a, b), (b, b)\})$, we can associate the state description

$$\Theta'(a, b) = P(a) \wedge P(b) \wedge \neg R(a, a) \wedge R(a, b) \wedge \neg R(b, a) \wedge R(b, b).$$

This description is important because it also generalizes to the nonstandard context. But what are the constants we should refer to, now? In our standard universe $\mathbb{U}$ we have a set of constants Con for which there exists a bijection $f : \mathbb{N}^+ \to \text{Con}$ such that $f(i) = a_i$ for every $i \in \mathbb{N}^+$; in $\mathbb{U}'$, by Transfer Principle, we have ${}^*\text{Con}$ and an internal bijection ${}^*f : {}^*\mathbb{N}^+ \to {}^*\text{Con}$. Hence, in $\mathbb{U}'$ we have the constant symbols in Con (since for any $n \in \mathbb{N}^+$, $f(n) = a_n$, hence ${}^*f({}^*n) = {}^*f(n) = {}^*a_n = a_n$[25]) and a "nonstandard contast" $a_N = {}^*f(N)$ for any nonstandard natural number N.

For any state description $\Theta(b_1, \ldots, b_m)$, A_Θ will be the associated q-uple according to the above outlined correspondence, as in the previous example. Modulo this identification, we will talk undistinguishably about state descriptions or the corresponding q-uples, and also functions over state descriptions (as, for instance, probabilities) will be represented as functions over q-uples of the aforementioned type. As we will see, this will be useful to generalize in the nonstandard setting the notion of state descriptions over $a_1, \ldots, a_N$ where N is a nonstandard natural number.

In the following, we define nonstandard state descriptions and prove some properties we will need in this section: we will see that by Transfer Principle, many features can be carried to the nonstandard context.

1. for any $n \in \mathbb{N}$, the set of all the state descriptions for $a_1, \ldots, a_n$ can be described by the following formula

$$\varphi(x, n) := x = \mathcal{P}(\{a_1, \ldots, a_n\}^{r_1}) \times \cdots \times \mathcal{P}(\{a_1, \ldots, a_n\}^{r_q}).$$[26]

[24] Here we are not paying attention to the order in which the conjuncts are displayed in a state description, i.e. two state descriptions are considered the same if they have the same literals.

[25] Recall that for any element $a \in \mathbb{X}$, ${}^*a = a$.

[26] Here, for any $n \in \mathbb{N}$ the set $\{a_1, \ldots, a_n\}$ is intended to be the set $\{f(i) : i \leq n\}$, where $f : \mathbb{N} \to \text{Con}$ is the already mentioned bijection such that $f(i) = a_i$ for any $i \in \mathbb{N}$; in this way, n is a variable of φ. A more precise formalization should be

$$\varphi(x, n) := \exists A \in \mathcal{P}(\text{Con}) \left(\forall i \in \mathbb{N}\; i \leq n \to f(i) \in A \wedge \forall x \in A \exists i \in \mathbb{N}\; x = f(i) \wedge x = \mathcal{P}(A^{r_1}) \times \cdots \times \mathcal{P}(A^{r_q})\right);$$

however, we prefer to use the formula written in the text that is more intuitive, even if less precise. We invite the reader to cover this kind of detail through the section.

Since in $\mathbb{U}$ it is valid that

$$\forall n \in \mathbb{N}\ \exists! x \in \mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))\ \varphi(x,n),$$

in $\mathbb{U}'$ its $*$-transform must hold, i.e.

$$\forall N \in {}^*\mathbb{N}\ \exists! x \in {}^*\big(\mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))\big)\ {}^*\varphi(x,N). \tag{2.17}$$

Hence, for any $N \in {}^*\mathbb{N}$ we can represent the set of state descriptions SD_N for $a_1,\ldots,a_N$, as the unique x in ${}^*\big(\mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))\big)$ that satisfies the formula ${}^*\varphi(x,N)$ above: with a slight abuse of notation, for the identification discussed above, we will denote this x still with SD_N.

Rephrasing what we said before, SD_N is the internal set

$${}^*\mathcal{P}(\{a_1,\ldots,a_N\}^{r_1}) \times \cdots \times {}^*\mathcal{P}(\{a_1,\ldots,a_N\}^{r_q}).$$

2. The set SD_N is hyperfinite for any $N \in {}^*\mathbb{N}$: indeed, by Formula 2.17, this set belongs to ${}^*\mathcal{P}_{fin}(A)$ for a given set $A \in \mathbb{U}$.

3. If $m \leq r$, a state description Φ for $a_1,\ldots,a_r$ has $\Theta(a_1,\ldots,a_m)$ as a semantic consequence if and only if for any $i \in \{1,\ldots,q\}$ the i-th component $(A_\Theta)_i$ of the q-uple A_Θ is the intersection

$$(A_\Phi)_i \cap \{a_1,\ldots,a_m\}^{r_i}.$$

Then, the state descriptions y for $a_1,\ldots,a_r$ that imply Θ can be characterized by the formula

$$\begin{aligned}\psi_\Theta(y,r) := \exists x \in \mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))\varphi(x,r) \wedge y \in x\\ \wedge \exists y_1 \in \mathcal{P}(\mathrm{Con}^{r_1}) \wedge \cdots \wedge \exists y_q \in \mathcal{P}(\mathrm{Con}^{r_q})\ y = (y_1,\ldots,y_q)\\ \wedge \bigwedge_{i=1}^{q}(y_i \cap \{a_1,\ldots,a_m\}^{r_i} = (A_\Theta)_i).\end{aligned}$$

The formula $\psi_\Theta(y,r)$ is not a $\mathcal{L}_\mathbb{U}$-formula since in the last conjunct there are elements like $y_i \cap \{a_1,\ldots,a_m\}^{r_i}$ that are not $\mathcal{L}_\mathbb{U}$-terms. However, we can have an equivalent formula ψ'_Θ replacing the formulas "$y_i \cap \{a_1,\ldots,a_m\}^{r_i} = (A_\Theta)_i$ ", with

$$\forall u \in y_i\ (u \in \{a_1,\ldots,a_m\}^{r_i} \to u \in (A_\Theta)_i) \wedge \forall v \in (A_\Theta)_i\ (v \in y_i \wedge v \in \{a_1,\ldots,a_m\}^{r_i})$$

for every $i = 1,\ldots,q$; now $\psi'_\Theta(y,r)$ is an $\mathcal{L}_\mathbb{U}$-formula.

The set $B^r_{\vDash\Theta}$ of all the state descriptions Φ for $a_1,\ldots,a_r$ that imply $\Theta(a_1,\ldots,a_m)$ can be described by

$$\begin{aligned}\phi_\Theta(Y,r) := \exists x \in \mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))\varphi(x,r)\\ \wedge (\forall z \in Y\ \psi'_\Theta(z,r)) \wedge \forall z \in x\ (\psi'_\Theta(z,r) \to z \in Y)\end{aligned}$$

in a way that in $\mathbb{U}$ a set A satisfies $\Phi_\Theta(A,r)$ if and only if $A = B^r_{\vDash\Theta}$.

In $\mathbb{U}$, for a fixed state description $\Theta(a_1,\ldots,a_m)$, it is valid that

$$\forall r \in \mathbb{N}\ \exists! Y \in \mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))\phi_\Theta(Y,r).$$

This is an $\mathcal{L}_{\mathbb{U}}$-formula, therefore, by Transfer Principle, for any $N \in {}^*\mathbb{N}$ there is a unique set $B^N_{\models\Theta} \in {}^*\big(\mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))\big)$ that satisfies ${}^*\phi_\Theta(Y, N)$: this can be interpreted as the set of all the state descriptions Φ for $a_1, \dots, a_N$ that imply Θ. Furthermore, it is a subset of SD_N and, since it belongs to ${}^*\mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))$, it is hyperfinite.

Once we have carried the notion of state descriptions in the nonstandard setting, we are ready to adjust some already proved properties for this new context.

Lemma 2.4.1. *With the notation used above, given a probability $w \in \mathbb{U}$ on* $\mathrm{Sen}(\mathcal{L})$ *and a state description $\Theta(a_1, \dots, a_m)$, for any $N \in {}^*\mathbb{N} \setminus \mathbb{N}$, we have that*

$$w(\Theta(a_1, \dots, a_m)) = {}^*w(\Theta(a_1, \dots, a_m)) = \sum_{\Phi(a_1,\dots,a_N) \in B^N_{\models\Theta}} {}^*w(\Phi(a_1, \dots, a_N)).^{27} \tag{2.18}$$

Proof. Given a state description $\Theta(a_1, \dots, a_m)$, for any $r \geq m$ we have

$$w(\Theta(a_1, \dots, a_m)) = \sum_{\Phi(a_1,\dots,a_r) \models \Theta(a_1,\dots,a_m)} w(\Phi(a_1, \dots, a_r)).$$

Hence, with the notation used above, the following formula is valid in $\mathbb{U}$:

$$\forall r \in \mathbb{N}\ r \geq m \rightarrow \Big(\exists Y \in \mathcal{P}_{fin}(\mathcal{P}(\mathrm{Con}^{r_1}) \times \cdots \times \mathcal{P}(\mathrm{Con}^{r_q}))\phi_\Theta(Y, r) \wedge w(A_\Theta) = \sum_{y \in Y} w(y)\Big),$$

where, we recall that the only set Y for which $\phi_\Theta(Y, r)$ holds is $B^r_{\models\Theta}$. By Transfer Principle, we get the validity in $\mathbb{U}'$ of the $*$-transform of the previous formula, and hence the thesis. □

Remark 2.4.1. The notation ${}^*w(\Theta(a_1, \dots, a_m))$ can seem ambiguous. Since w is a function in $\mathbb{U}$, ${}^*w$ is a function in $\mathbb{U}'$ and by the above term we can mean either the image through * of the real number $w(\Theta(a_1, \dots, a_m))$, or the image through ${}^*w$ of ${}^*(\Theta(a_1, \dots, a_m))$. Actually, ${}^*(\Theta(a_1, \dots, a_m)) = \Theta(a_1, \dots, a_m)$ because the set at issue are finite and involve $a_1, \dots, a_m$ that are individuals of $\mathbb{X}$; recalling that for a generic function $f \in \mathbb{U}$, ${}^*(f(a)) = {}^*f({}^*a)$ and that ${}^*r = r$ for any $r \in \mathbb{R} \subseteq \mathbb{X}$, we have

$$w(\Theta(a_1, \dots, a_m)) = {}^*(w(\Theta(a_1, \dots, a_m))) = {}^*(w)({}^*\Theta(a_1, \dots, a_m)) = {}^*(w)(\Theta(a_1, \dots, a_m)).$$

The situation is different when we deal with state descriptions over $a_1, \dots, a_N$ with $N \in {}^*\mathbb{N} \setminus \mathbb{N}$: in this case, $w(\Theta(a_1, \dots, a_N))$ isn't defined, but ${}^*w(\Theta(a_1, \dots, a_N))$ is meaningful and denotes ${}^*(w)(\Theta(a_1, \dots, a_N))$.

Recalling that our aim is a generalization of Theorem 2.2.1, we can go further, investigating what we can say more if w satisfies Ex.

In $\mathbb{U}$, for any $m \in \mathbb{N}$, for any bijection σ of the set $\{1, \dots, m\}$, and for any state descriptions $\Phi(a_1, \dots, a_m)$

$$w(\Phi(a_1, \dots, a_m)) = w(\Phi(a_{\sigma(1)}, \dots, a_{\sigma(m)})).$$

[27] Notice that this sum makes sense since the set of the indexes $B^N_{\models\Theta}$ is hyperfinite and the function ${}^*w$ is internal.

This means that also in the nonstandard setting, for any $N \in {}^*\mathbb{N}$ and for any internal σ bijection of $\{1, \dots, N\}$ we have

$$ {}^*w(\Phi(a_1, \dots, a_N)) = {}^*w(\Phi(a_{\sigma(1)}, \dots, a_{\sigma(N)})). \tag{2.19} $$

Another way to formulate Equation (2.19), is that ${}^*w$ is constant on any class of the form

$$ \{\Phi(a_{\sigma(1)}, \dots, a_{\sigma(N)}) : \sigma \text{ internal bijection of } \{1, \dots, N\}\}. \tag{2.20} $$

Hence, if we pick up a representative $\Psi(a_1, \dots, a_N)$ for each class, and denote the corresponding class $\mathcal{H}_\Psi$, we have that

$$ \sum_{\Phi \in \mathcal{H}_\Psi} {}^*w(\Phi) = {}^*w(\Psi) \cdot |\mathcal{H}_\Psi|, \tag{2.21} $$

where $|\mathcal{H}_\Psi|$ denote the internal cardinality of the hyperfinite set $\mathcal{H}_\Psi$.[28]The set of all the representatives $\Psi(a_1, \dots, a_N)$ will be denoted by A. Notice that A is an internal set: indeed, since the set SD_N of all the state descriptions for $a_1, \dots, a_N$ is hyperfinite, we have an internal bijection f from a set of the form $\{1, \dots, M\}$ to SD_N with $M \in {}^*\mathbb{N}$; then A can be defined as the subset of all the state descriptions $f(K) = \Psi(a_1, \dots, a_N)$ such that for any $n \leq K$, $f(n)$ doesn't belong to the same class of $f(K)$[29] Furthermore, A is also hyperfinite since it is an internal subset of a hyperfinite set (SD_N).

This set will be very useful in the following because it will be the domain of the measure involved in the analogous of Theorem 2.2.1.

Lemma 2.4.2. *With the notation used above, putting for any state description $\Psi \in A$*

$$ M(\{\Psi\}) := \sum_{\Phi \in \mathcal{H}_\Psi} {}^*w(\Phi), $$

we get a *probability[30] *(hence an internal finitely-additive probability) $M : \mathcal{A} \to {}^*\mathbb{R}$ on the ${}^*\sigma$-algebra $\mathcal{A}$ of internal subsets of A, by defining for any internal subset B of A*

$$ M(B) = \sum_{\Psi \in B} M(\{\Psi\}). $$

Proof. Let's start by noticing that $\mathcal{A}$ is an algebra since A is internal and internal sets are closed under boolean combinations; $\mathcal{A}$ is also a ${}^*\sigma$-algebra since for any internal map $f : {}^*\mathbb{N} \to \mathcal{A}$ that determines a (internal) family $\{A_n : n \in {}^*\mathbb{N}\}$, the union $\cup_{n \in {}^*\mathbb{N}} f(n) = \cup_{n \in {}^*\mathbb{N}} A_n$ is internal (see 104, Appendix B).
With regards to M, we have:

- M is well-defined, since, for any $\Psi \in A$, $\mathcal{H}_\Psi$ is hyperfinite, therefore the sum defining $M(\{\Psi\})$ is meaningful; since also any subset B of A is hyperfinite, and the function $\Psi \mapsto M(\{\Psi\})$ is internal, the sum defining $M(B)$ makes sense, too.

[28]This is hyperfinite because it is a subset of a hyperfinite set (the set of all the state descriptions for $a_1, \dots, a_N$) and it is internal (because given Ψ, the corresponding class is definible using an internal first-order formula).

[29]i.e. doesn't exist an internal bijection σ such that $f(n) = \Psi(a_{\sigma(1)}, \dots, a_{\sigma(N)})$. This is definable by an internal formula.

[30]i.e. a *measure such that $M(A) = 1$, see Section 2.3.

- M is a *probability: indeed, since A is internal, there exists $C \in \mathbb{U}$ for which $A \in {}^*C$. In the standard setting, given a finite set $X \in C$ and a map $m : X \to [0,1]$ such that $\sum_{x \in X} m(x) = 1$, the function

$$\begin{aligned} \mu : \mathcal{P}_{fin}(X) &\to [0,1] \\ B &\mapsto \sum_{b \in B} m(b) \end{aligned}$$

  is a probability. Therefore, by Transfer Principle, the map M of the lemma statement is a *probability.

□

Recalling the discussion before Theorem 2.3.3, it is easy to see that we are in the condition (with regards to the *measure M and to the $^*\sigma$-algebra $\mathcal{A}$) for which it makes sense to consider integrals over A with the measure M.

Now, we use the probability M to rewrite Equation (2.18). Given a state description $\Theta(b_1, \dots, b_m)$ with $b_i \in \mathrm{Con}$ for any $i \le m \in \mathbb{N}$, $N \in {}^*\mathbb{N} \setminus \mathbb{N}$, and $\Psi \in A$, the set[31]

$$\{\Phi(a_1, \dots, a_N) \in \mathcal{H}_\Psi : \Phi(a_1, \dots, a_N) \vDash \Theta(b_1, \dots, b_m)\}$$

is hyperfinite[32], hence we can define the internal map that to $\Psi \in A$ associates

$$w^\Psi(\Theta(b_1, \dots, b_m)) = \frac{|\{\Phi(a_1, \dots, a_N) \in \mathcal{H}_\Psi : \Phi(a_1, \dots, a_N) \vDash \Theta(b_1, \dots, b_m)\}|}{|\mathcal{H}_\Psi|}. \tag{2.22}$$

Introducing these new entities allows us to reformulate in a nicer formula Equation (2.18) that now becomes:

$$\begin{aligned} {}^*w(\Theta(b_1, \dots, b_m)) &= \sum_{\Phi(a_1,\dots,a_N) \vDash \Theta(b_1,\dots,b_m)} {}^*w(\Phi(a_1, \dots, a_N)) \\ &= \sum_{\Psi \in A} \sum_{\substack{\Phi(a_1,\dots,a_N) \vDash \Theta(b_1,\dots,b_m) \\ \Phi(a_1,\dots,a_N) \in \mathcal{H}_\Psi}} {}^*w(\Phi(a_1, \dots, a_N)) \\ &= \sum_{\Psi \in A} \sum_{\substack{\Phi(a_1,\dots,a_N) \vDash \Theta(b_1,\dots,b_m) \\ \Phi(a_1,\dots,a_N) \in \mathcal{H}_\Psi}} {}^*w(\Psi) \\ &= \sum_{\Psi \in A} |\{\Phi(a_1, \dots, a_N) \in \mathcal{H}_\Psi : \Phi \vDash \Theta\}| \, {}^*w(\Psi) \\ &= \sum_{\Psi \in A} \frac{|\{\Phi(a_1, \dots, a_N) \in \mathcal{H}_\Psi : \Phi \vDash \Theta\}|}{|\mathcal{H}_\Psi|} \, |\mathcal{H}_\Psi| \, {}^*w(\Psi) \\ &\overset{*}{=} \sum_{\Psi \in A} \Big(w^\Psi(\Theta(b_1, \dots, b_m)) \big(\sum_{\Phi \in \mathcal{H}_\Psi} {}^*w(\Phi) \big) \Big), \end{aligned}$$

where in $\overset{*}{=}$ we used Equation (2.21).

[31] where, here and in the following, by $\Phi(a_1, \dots, a_N) \vDash \Theta(b_1, \dots, b_m)$ we mean that $\Phi(a_1, \dots, a_N) \in B^N_{\vDash\Theta}$.

[32] for the same reason of Footnote 28, page 63.

Notice that, by Transfer Principle, since the integral over a finite set in the standard setting is a sum, in the nonstandard context an integral over a hyperfinite set means, still, a summation over it. Therefore,

$$\begin{aligned} w(\Theta(a_1,\dots,a_m)) &= {}^*w(\Theta(b_1,\dots,b_m)) \\ &= \sum_{\Psi\in A}\Big(w^{\Psi}(\Theta(b_1,\dots,b_m))\big(\sum_{\Phi\in\mathcal{H}_\Psi}{}^*w(\Phi)\big)\Big) \\ &= \sum_{\Psi\in A} w^{\Psi}(\Theta(b_1,\dots,b_m))\ M(\{\Psi\}) \\ &= \int_A w^{\Psi}(\Theta(b_1,\dots,b_m))\ dM(\Psi): \end{aligned} \tag{2.23}$$

notice that, until now, $w^\Psi(\Theta(b_1,\dots,b_m))$ has been seen as a function w^Ψ on SD_m; in the chain of equalities above, it is intended to be a function on A (with input Ψ) for a given $\Theta(b_1,\dots,b_m)$.

Since the function $\Psi\mapsto w^\Psi(\Theta(b_1,\dots,b_m))$ is internal, $\mathcal{A}$-measurable[33], and bounded (by 1), we can use Theorem 2.3.3, getting from Equation (2.23)

$$w(\Theta(b_1,\dots,b_m)) = \int_A {}^\circ(w^\Psi(\Theta(b_1,\dots,b_m)))\ dM_L(\Psi), \tag{2.24}$$

where M_L is the Loeb measure associated to the space $(A,\mathcal{A},M)$.

Equation (2.24) is analogous to the one provided by Theorem 2.2.1 in the unary case. Before showing this, we prove that the maps that appear in the integrals are indeed probabilities, as the $w_{\vec{x}}$'s in the unary case.

Lemma 2.4.3. *With the notation used above, given $\Psi\in A$, the map ${}^\circ w^\Psi$ defined on state descriptions as in Equation (2.22), can be extended to a (Gaifman) probability on* $\mathrm{Sen}(\mathcal{L})$ *that satisfies Ex.*

Proof. To show that ${}^\circ w^\Psi$ can be extended to a probability on $\mathrm{Sen}(\mathcal{L})$, we should verify the validity of conditions at page 36:

- for any $m\in\mathbb{N}^+$, and constants $b_1,\dots,b_m\in\mathrm{Con}$, ${}^\circ w^\Psi(\Theta(b_1,\dots,b_m))\geq 0$, since cardinalities are in ${}^*\mathbb{N}$, so non-negative;
- considering the case in which $m=0$, i.e. when the state description is $\top$, since

  $$\{\Phi(a_1,\dots,a_N)\in\mathcal{H}_\Psi\ :\ \Phi(a_1,\dots,a_N)\vDash\top\}=\mathcal{H}_\Psi,$$

  we get that ${}^\circ w^\Psi(\top)=1$;
- for any $r\geq m$ and constants $b_1,\dots,b_m,\dots,b_r\in\mathrm{Con}$,

  $${}^\circ w^\Psi(\Theta(b_1,\dots,b_m)) = \sum_{\Theta'(b_1,\dots,b_r)\vDash\Theta(b_1,\dots,b_m)} {}^\circ w^\Psi(\Theta'(b_1,\dots,b_r)):$$

[33] Recall that given an internal map g, the preimage of an internal set through g is internal. In this situation, the function at issue is from A to ${}^*\mathbb{R}$ and the preimage of any open set (hence, internal) in ${}^*\mathbb{R}$ is internal, therefore in $\mathcal{A}$.

indeed, notice that

$$\{\Phi \in \mathcal{H}_\Psi \ : \ \Phi \vDash \Theta\} = \bigcup_{\Theta' \vDash \Theta} \{\Phi \in \mathcal{H}_\Psi \ : \ \Phi \vDash \Theta'\},$$

where Φ, Θ, Θ' are state descriptions standing for $\Phi(a_1, \dots, a_N), \Theta(b_1, \dots, b_m)$ and $\Theta'(b_1, \dots, b_r)$ respectively. Furthermore, different state descriptions for the same constants are pairwise contradictory, thus it can't happen that the same $\Phi \in \mathcal{H}_\Psi$ implies two of them: this means that the union in the above equation is disjoint. Carrying in the nonstandard setting the usual properties of cardinality for finite sets, applying them to hyperfinite ones, dividing the equation above by $|\mathcal{H}_\Psi|$, and then taking the standard part, we get the thesis.

Hence, ${}^\circ w^\Psi$ can be extended on $\mathrm{Sen}(\mathcal{L})$ to a Gaifman probability. In addition, it satisfies Ex: to prove this, we simply can show it in the case of state descriptions.

Given a state description $\Theta(a_1, \dots, a_m)$, a permutation σ of $\mathbb{N}$ and an internal permutation γ of $\{1, \dots, N\}$ that extends σ, we have $\Phi(a_1, \dots, a_N) \vDash \Theta(a_1, \dots, a_m)$ if and only if $\Phi(a_{\gamma(1)}, \dots, a_{\gamma(N)}) \vDash \Theta(a_{\sigma(1)}, \dots, a_{\sigma(m)})$.

Hence,

$$\begin{aligned}
&|\{\Phi(a_1, \dots, a_N) \in \mathcal{H}_\Psi : \Phi(a_1, \dots, a_N) \vDash \varphi(a_1, \dots, a_m)\}| \\
=&|\{\Phi(a_1, \dots, a_N) \in \mathcal{H}_\Psi : \Phi(a_{\gamma(1)}, \dots, a_{\gamma(N)}) \vDash \varphi(a_{\sigma(1)}, \dots, a_{\sigma(m)})\}| \\
=&|\{\Phi'(a_1, \dots, a_N) \in \mathcal{H}_\Psi : \Phi'(a_1, \dots, a_N) \vDash \varphi(a_{\sigma(1)}, \dots, a_{\sigma(m)})\}|,
\end{aligned}$$

where the last equality is motivated by the fact that the function that maps $\Phi(a_1, \dots, a_N)$ to $\Phi(a_{\gamma(1)}, \dots, a_{\gamma(N)}) = \Phi'(a_1, \dots, a_N)$ is an internal bijection in $\mathcal{H}_\Psi$. □

Remark 2.4.2. Notice that extending Equation (2.22) to quantifier-free formulas, we can show directly that the map ${}^\circ w^\Psi : \mathrm{QFSen}(\mathcal{L}) \to [0, 1]$ satisfies the requirement of a probability in Definition 1.0.1. Hence, by unicity, we have that the extension of ${}^\circ w^\Psi$ satisfies for any $\varphi \in \mathrm{QFSen}(\mathcal{L})$,

$${}^\circ w^\Psi(\varphi) = {}^\circ\!\left(\frac{|\{\Phi(a_1, \dots, a_N) \in \mathcal{H}_\Psi : \Phi(a_1, \dots, a_N) \vDash \varphi(b_1, \dots, b_m)\}|}{|\mathcal{H}_\Psi|} \right).$$

We have defined all the necessary elements to understand the following generalization of Theorem 2.2.1.

Theorem 2.4.4. *Given a polyadic language $\mathcal{L}$, and using the notation above, a probability w that satisfies Ex can be written as*

$$w = \int_A {}^\circ w^\Psi \ d\mu(\Psi),$$

where μ is a probability on the set A.

Conversely, if a probability w can be represented as above, then it satisfies Ex.

Proof. From Equation (2.24), we get that the equation in the statement holds when we consider the Loeb mesure $\mu = M_L$ and we deal with state descriptions. It is easy to see that the same holds for sentences in $\mathrm{Sen}(\mathcal{L})$ proceeding as in the proof of Corollary 2.2.2.

Conversely, we have to show that if we define $w = \int_A {}^\circ w^\Psi \, d\mu(\Psi)$, then w is a probability that satisfies Ex. Also in this case the proof is really similar to the one presented in Corollary 2.2.2: in this case, the thesis follows using the fact that the maps ${}^\circ w^\Psi$'s are Gaifman probabilities satisfying Ex.

□

As already marked out at the end of Section 2.2, in the polyadic case, as in the unary one, it can be shown that the building blocks for a probability w that satisfies Ex can be characterized by IP.

Theorem 2.4.5. *Let w be a probability that satisfies Ex. Then it satisfies IP if and only if there exists some A as in the discussion above, and $\Psi \in A$ such that $w = {}^\circ w^\Psi$.*

Proof. We will show that given $\Psi \in A$, ${}^\circ w^\Psi$ satisfies IP, by doing a little detour in the definition of w^Ψ. Recall that for a state description $\Theta(a_1, \ldots, a_m)$,

$$w^\Psi(\Theta(a_1, \ldots, a_m)) = \frac{|\Phi(a_1, \ldots, a_N) \in \mathcal{H}_\Psi \;:\; \Phi(a_1, \ldots, a_N) \vDash \Theta(a_1, \ldots, a_m)\}|}{|\mathcal{H}_\Psi|}.$$

We have already noticed that $\Phi(a_1, \ldots, a_N) \vDash \Theta(a_1, \ldots, a_m)$ if and only if Φ restricted to the constants $a_1, \ldots, a_m$ coincides with Θ. Using this fact, we can show that the fraction above can be regarded as the ratio between the choices of constants $a_{h_1}, \ldots, a_{h_m}$ in $\{a_1, \ldots, a_N\}$ for which the restriction of Ψ to them coincide with $\Theta(a_{h_1}, \ldots, a_{h_m})$ and all the possible choices of m constants. Indeed, if in Ψ we have such constants $a_{h_1}, \ldots, a_{h_m}$, we can take an internal bijection σ of $\{1, \ldots, N\}$ such that $\sigma(i) = h_i$ for any $i = 1, \ldots, m$; then, $\Psi(a_{\sigma^{-1}(1)}, \ldots, a_{\sigma^{-1}(N)}) \vDash \Theta(a_1, \ldots, a_m)$. Conversely, if we have an internal bijection σ such that $\Psi(a_{\sigma(1)}, \ldots, a_{\sigma(N)}) \vDash \Theta(a_1, \ldots, a_m)$, then we have the constants $a_{h_1} = a_{\sigma^{-1}(1)}, \ldots, a_{h_m} = a_{\sigma^{-1}(m)}$ for which Ψ restricted to them coincides with $\Theta(a_{h_1}, \ldots, a_{h_m})$.

Hence, imagine to randomly picking in *Con, without replacement and according to the uniform distribution, some constants $a_{h_1}, \ldots, a_{h_m}$ from the set $\{a_1, \ldots, a_N\}$; we will say that a choice is *good for* $\Theta(a_1, \ldots, a_m)$ if Ψ implies $\Theta(a_{h_1}, \ldots, a_{h_m})$. For what said before and recalling that $\mathcal{H}_\Psi$ is composed of state descriptions of the form $\Psi(a_{\sigma(1)}, \ldots, a_{\sigma(N)})$ for an internal bijection σ, $w^\Psi(\Theta(a_1, \ldots, a_m))$ is the probability of picking a good choice over the possible choices that can be made without replacement. Hence, if we call G_Θ the set of all the good choices, we have

$$w^\Psi(\Theta(a_1, \ldots, a_m)) = \frac{|G_\Theta|}{\prod_{i=0}^{m-1}(N-i)}.$$ [34]

[34] We can be more precise carrying in the nonstandard setting some usual notations and results from the standard world (we leave to the reader all the details needed to understand why the following argument actually works). Denoting by Num the numerator of the equation defining $w^\Psi(\Theta)$, the outlined above map Num $\to G_\Theta$ is surjective: for any good choice $a_{h_1}, \ldots, a_{h_m}$, we can define an internal bijection σ of $\{1, \ldots, N\}$ such that $\sigma(i) = h_i$ for $i \leq m$ and retrieve a $\Phi \in$ Num. Changing the permutation σ (always provided that $i \mapsto h_i$ for any $i \leq m$), we obtain a different Φ': therefore, since there are $(N-m)!$ permutations that extend the map $i \mapsto h_i$, we have that $|\text{Num}| = |G_\Theta| \cdot (N-m)!$. Dividing by $|\mathcal{H}_\Psi| = N!$, we get the thesis.

A simple computations shows that

$$\frac{N^m}{\prod_{i=0}^{m-1}(N-i)} \approx 1,^{35}$$

and, similarly, it can be shown that, calling G_Θ^{rep} the set of the good choices with replacement (i.e. we are not requiring anymore that the a_{h_i}'s are distinct), also

$$\frac{|G_\Theta^{rep}|}{N^m} \approx \frac{|G_\Theta|}{N^m}.^{36}$$

Therefore, we can say that

$$^\circ w^\Psi(\Theta(a_1, \ldots, a_m)) = {}^\circ\left(\frac{|G_\Theta^{rep}|}{N^m}\right).$$

This reformulation allows us to describe more easily $^\circ w^\Psi$: this is the possibility of picking up a good choice over all the possible choices *with replacement*. Therefore, given two state descriptions $\Theta_1(b_1, \ldots, b_m)$ and $\Theta_2(c_1, \ldots, c_n)$ with no constants in common, since the choices are with replacement, picking a good choice for Θ_1 is independent of picking a good one for Θ_2. Thus, since $\Psi(a_1, \ldots, a_N)$ implies $\Theta_1 \wedge \Theta_2$ if and only if it implies both Θ_1 and Θ_2, we get, using Remark 2.4.2, that

$$^\circ w^\Psi(\Theta_1 \wedge \Theta_2) = {}^\circ w^\Psi(\Theta_1) \cdot {}^\circ w^\Psi(\Theta_2),$$

and therefore IP.

Assume now that w is a probability that satisfies EX and IP. By Theorem 2.4.4, there exists a set A and a measure μ on a σ-algebra on it such that

$$w = \int_A {}^\circ w^\Psi \, d\mu(\psi).$$

If we consider two quantifier-free formulas $\varphi = \varphi(a_1, \ldots, a_n)$ and $\varphi' = \varphi(a_{n+1}, \ldots, a_{2n})$

[35]Since for a, b limited hyperreals $^\circ(a \cdot b) = {}^\circ a \cdot {}^\circ b$, we have that

$$\frac{\prod_{i=0}^{m-1}(N-i)}{N^m} = \prod_{i=1}^{m-1}(1 - \frac{i}{N}) \approx 1,$$

since it is a finite product of elements whose standard part is 1.

[36]This simply relies on the fact that the possible (ordered) choices of m elements among $a_1, \ldots, a_N$ without repetition are $N(N-1)\cdots(N-m+1)$. Therefore,

$$|G_\Theta^{rep}| - |G_\Theta| \leq N^m - N(N-1)\cdots(N-m+1);$$

since in $N^m - N(N-1)\cdots(N-m+1)$ the biggest exponent is N^{m-1}, if we divide for N^m and take the standard part, we have the desired outcome.

without constants in common, then by IP we have

$$\begin{aligned}0 &= 2(w(\varphi \wedge \varphi') - w(\varphi) \cdot w(\varphi')) \\ &= \int_A {}^\circ w^\Psi(\varphi \wedge \varphi')\ d\mu(\Psi) + \int_A {}^\circ w^\Lambda(\varphi \wedge \varphi')\ d\mu(\Lambda) \\ &\quad - 2\big(\int_A {}^\circ w^\Psi(\varphi)\ d\mu(\Psi)\big) \cdot \big(\int_A {}^\circ w^\Lambda(\varphi)\ d\mu(\Lambda)\big) \\ &= \int_A {}^\circ w^\Psi(\varphi) \cdot {}^\circ w^\Psi(\varphi')\ d\mu(\Psi) + \int_A {}^\circ w^\Lambda(\varphi) \cdot {}^\circ w^\Lambda(\varphi')\ d\mu(\Lambda) \\ &\quad - 2\big(\int_A {}^\circ w^\Psi(\varphi)\ d\mu(\Psi)\big) \cdot \big(\int_A {}^\circ w^\Lambda(\varphi)\ d\mu(\Lambda)\big) \\ &= \int_A \int_A ({}^\circ w^\Psi(\varphi) - {}^\circ w^\Lambda(\varphi))^2\ d\mu(\Psi)\ d\mu(\Lambda),\end{aligned}$$

where the last equality holds by Fubini's and Tonelli's Theorems (μ is a probability on A and the integrand functions are bounded) and by the fact that ${}^\circ w^\Psi$ and ${}^\circ w^\Lambda$ satisfy Ex, therefore ${}^\circ w^\Psi(\varphi) = {}^\circ w^\Psi(\varphi')$ and ${}^\circ w^\Lambda(\varphi) = {}^\circ w^\Lambda(\varphi')$.

Since the integrand functions are never negative, this means that for fixed $\varphi \in \mathrm{QFSen}(\mathcal{L})$, μ-almost-everywhere ${}^\circ w^\Psi(\varphi)$ and ${}^\circ w^\Lambda(\varphi)$ are equal: i.e. there is a subset C_φ of A of μ-measure 1, such that for any $\Psi, \Lambda \in C_\varphi$, we have ${}^\circ w^\Psi(\varphi) = {}^\circ w^\Lambda(\varphi)$, that is, the function $\Psi \mapsto w^\Psi(\varphi)$ is constant on C_φ. Since the formulas for a fixed countable language are countable, we can take the intersection C of the family $\{C_\varphi\}_{\varphi \in \mathrm{QFSen}(\mathcal{L})}$: this is still a set with μ-measure 1.

Hence, for any Ψ in C, we have for any $\varphi \in \mathrm{QFSen}(\mathcal{L})$ that

$$w(\varphi) = \int_A {}^\circ w^\Lambda(\varphi)\ d\mu(\Lambda) = \int_C {}^\circ w^\Lambda(\varphi)\ d\mu(\Lambda) = {}^\circ w^\Psi(\varphi)\mu(C) = {}^\circ w^\Psi(\varphi),$$

from which the thesis. □

Chapter 3

Symmetry and Open Problems

In the last chapter of the work, we will briefly present other rational principles and move toward the nowadays open questions on this topic.

The Constant Exchangeability Principle (Ex) shows that symmetry plays a key role in defining rationality. Hence, it makes sense to define what we mean by *symmetry* and to investigate other principles that can be derived from it.

3.1 Symmetry in the unary case

In this section, we will focus on the unary case: in general, as always, if not otherwise specified, the language will have constants $\{a_n\}_{n\in\mathbb{N}^+}$, unary relational symbols $R_1, \dots, R_q$, neither functional symbols nor equality.

To motivate the principles we will present, we consider and adjust the example used at the beginning of Chapter 2: suppose to have an urn with some balls (say for simplicity a countable supply), each with a non-zero natural number on it. Suppose that different balls have different numbers written on them (so, there is one and only one ball for each positive natural) and that these balls can be red or have a different color; in addition, the balls can have a white circle on them or not.

To formalize this situation, we will use a language $\mathcal{L}$ composed of a countable supply of constants $\{a_n\}_{n\in\mathbb{N}^+}$ and two unary predicate symbols $P_1(x)$ and $P_2(x)$.

In the following, we will assume that the agent is in the *zero-knowledge* condition, i.e. it doesn't know any relevant information about the world it lives in. In this setting, this means that the agent knows only that the universe is exhausted by the constant symbols in $\mathcal{L}$ and that the entities it is studying can satisfy two unary properties, represented by the predicate symbols $P_1(x)$ and $P_2(x)$.

We will present some valid formalizations that can be used:

- *Formalization 1:*
  The constant a_i is assigned to the ball in which the number i is written; $P_1(a_i)$ stands for "the ball associated with a_i is red"; $P_2(a_i)$ stands for "the ball associated to a_i has a white circle".
- *Formalization 2:*

The constant a_i is assigned to the ball in which the number i is written, for every $i \notin \{1, 2\}$; a_1 is assigned to the ball in which it is written 2 and a_2 to the ball in which it is written 1; $P_1(a_i)$ stands for "the ball associated to a_i is red"; $P_2(a_i)$ stands for "the ball associated to a_i has a white circle".

- *Formalization 3:*
  The constant a_i is assigned to the ball in which the number i is written, for every i; $P_1(a_i)$ stands for "the ball associated to a_i isn't red"; $P_2(a_i)$ stands for "the ball associated to a_i has a white circle".

- *Formalization 4:*
  The constant a_i is assigned to the ball in which the number i is written, for every i; $P_1(a_i)$ stands for "the ball associated to a_i has a white circle"; $P_2(a_i)$ stands for "the ball associated to a_i is red".

Formalization 1 and Formalization 4 can seem neater than the others: Formalization 2 involves an assignment of constants to balls that is not the first that comes to mind; in Formalization 3, $P_1(a_i)$ means that the ball a_i is *not* red, even if it seems more natural that $P_1(x)$ stands for "x is red".

However, these formalizations are all valid and expressive in the same way: they differ only in the name given to objects or properties we want to study.

Suppose we want to formalize the following property: the ball with 1 written on it is red. According to the different formalizations, we have different sentences that represent it:

- *Formalization 1:* $P_1(a_1)$;
- *Formalization 2:* $P_1(a_2)$;
- *Formalization 3:* $\neg P_1(a_1)$;
- *Formalization 4:* $P_2(a_1)$.

In general, if a property is represented in Formalization 1 by a formula φ_1, then:

- in Formalization 2 the property is expressed by φ_2 that is the outcome of replacing the occurrences (if any) of a_1 with a_2 and of a_2 with a_1;
- in Formalization 3 the property is expressed by φ_3 that is the outcome of replacing the occurrences (if any) of P_1 with $\neg P_1$;
- in Formalization 4 the property is expressed by φ_4 that is the outcome of replacing the occurrences (if any) of P_1 with P_2 and of P_2 with P_1.

A rational agent should observe this ambiguity of the language: it should not give different values to sentences that can express the same property. In this way, we can justify the already discussed Ex principle: indeed, for what we said before comparing Formalization 1 and Formalization 2, an agent should consider φ_1 as likely as φ_2. Generalizing this example, we can deal with permutations of $\mathbb{N}^+$ more complex than the one that swaps a_1 and a_2 and get Ex.

If we use the same argument dealing with Formalization 3 and 4, we get the following principles:

- *Strong Negation Principle (SN)*: a probability w satisfies SN if for any statement φ and for any predicate symbol R, if φ' is the outcome of replacing in φ any occurrence of R with $\neg R$, we have that $w(\varphi) = w(\varphi')$;
- *Predicate Exchangeability Principle (Px) - Unary case*: if R and R' are two relational symbols and for any statement φ, φ' is the outcome of replacing in φ any occurence of R with R' and any occurence of R' with R, then $w(\varphi) = w(\varphi')$.

As it will be more clear in the future, these principles are regarded as *symmetry* ones: Ex reflects a symmetry between constants, Px between predicate symbols (of the same ariety), and SN between a predicate symbol and its negation.

Following this notion of symmetry, as we will see, we can justify other principles: here, we will present maybe the first that is possible to encounter besides the ones previously listed: the *Atom Exchangeability Principle (Ax)*. A probability w satisfies Ax if for any function $h : \{1, \ldots, n\} \to \{1, \ldots, 2^q\}$, for any permutation τ of $\{1, \ldots, 2^q\}$ and for any constants $b_1, \ldots, b_n$,

$$w\big(\bigwedge_{i=1}^{n} \alpha_{h_i}(b_i)\big) = w\big(\bigwedge_{i=1}^{n} \alpha_{\tau(h_i)}(b_i)\big).$$

where, as usual, $\alpha_{h_i}(x), \alpha_{\tau(h_i)}(x)$ are atoms.

If we want to exploit symmetry to give rise to many other rational principles, we should define what we mean by it. In other branches of mathematics, the notion of symmetry is extremely linked to the one of automorphism: the more automorphisms an object (a geometrical figure, an algebraic group) has, the more is regarded as symmetric. Dealing with Inductive Logic, it's not very clear what should be the "object" to consider.

Recall that an agent can have symmetry only in a zero-knowledge context: if it already knows that $P(a_1) \wedge \neg P(a_2)$ holds, we can't ask it to deal with a_1 and a_2 in the same way. In this condition, an agent knows only a few things: the logical rules, that the universe entities are exhausted by the set Con, that the properties of these are expressed by means of the relational symbols in $\mathcal{L}$, and that it lives in one structure in $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$. Hence, in the literature, it is proposed the following definition of automorphism.

Definition 3.1.1. Consider a language $\mathcal{L}$ with $\{a_n\}_{n\in\mathbb{N}^+}$ as the set of constants and the two-sorted set $M(\mathcal{L}) := (\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}), B(\mathcal{L}))$ where, as at page 21,

$$B(\mathcal{L}) = \{\mathrm{Mod}_{\mathrm{Con}}(\varphi) \ : \ \varphi \in \mathrm{Sen}(\mathcal{L})\}.$$

An *automorphism* of $M(\mathcal{L})$ is a bijective map $\gamma : \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) \to \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ such that for each $\varphi \in \mathrm{Sen}(\mathcal{L})$, there is a $\psi \in \mathrm{Sen}(\mathcal{L})$ (that will be named $\gamma(\varphi)$) for which

$$\gamma(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) := \{\gamma(\mathfrak{A}) \ : \ \mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\varphi)\} = \mathrm{Mod}_{\mathrm{Con}}(\psi) = \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi)). \tag{3.1}$$

Remark 3.1.1. The requirement expressed in (3.1) is equivalent to the following: for any $\varphi \in \mathrm{Sen}(\mathcal{L})$ and for any $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$,

$$\mathfrak{A} \vDash \varphi \text{ if and only if } \gamma(\mathfrak{A}) \vDash \gamma(\varphi). \tag{3.2}$$

Indeed, assuming (3.1), we get (3.2), since:

- given a model $\mathfrak{A}$ of φ, then since $\gamma(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) = \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi))$, $\gamma(\mathfrak{A})$ satisfies $\gamma(\varphi)$;
- if for a structure $\mathfrak{A}$, $\gamma(\mathfrak{A}) \vDash \gamma(\varphi)$, then there is a model $\mathfrak{B}$ of φ such that $\gamma(\mathfrak{A}) = \gamma(\mathfrak{B})$. From the injectivity of γ, we get that $\mathfrak{A} = \mathfrak{B}$ is a model of φ itself.

Conversely, given (3.2), we can prove (3.1):

- $\gamma(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) \subseteq \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi))$: given a model $\mathfrak{A}$ of φ, by (3.2), $\gamma(\mathfrak{A}) \vDash \gamma(\varphi)$;
- $\gamma(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) \supseteq \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi))$: given $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi))$, for the surjectivity of γ, there exists $\mathfrak{A}$ such that $\gamma(\mathfrak{A}) = \mathfrak{B}$. By (3.2), since $\mathfrak{B} \vDash \gamma(\varphi)$, $\mathfrak{A}$ satisfies φ.

By (3.1), we can get some properties that an automorphism γ must satisfy. In particular, it has to preserve, up to logical equivalence, boolean operations: formally, for any $\varphi, \psi \in \mathrm{Sen}(\mathcal{L})$,

$$\gamma(\varphi \wedge \psi) \equiv \gamma(\varphi) \wedge \gamma(\psi) \quad \gamma(\varphi \vee \psi) \equiv \gamma(\varphi) \vee \gamma(\psi) \quad \gamma(\neg\varphi) \equiv \neg\gamma(\varphi).$$

We will only prove the case for $\wedge$, but similar arguments can be carried out for the other cases. Recalling that two sentences are equivalent if and only if they determine the same models in $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$,[1] we will prove that $\mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi \wedge \psi)) = \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi) \wedge \gamma(\psi))$:

$\subseteq$) since $\mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi \wedge \psi)) = \gamma(\mathrm{Mod}_{\mathrm{Con}}(\varphi \wedge \psi))$, a generic element in $\mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi \wedge \psi))$ is $\gamma(\mathfrak{A})$ with $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\varphi \wedge \psi)$. This means that $\mathfrak{A}$ is in both $\mathrm{Mod}_{\mathrm{Con}}(\varphi)$ and $\mathrm{Mod}_{\mathrm{Con}}(\psi)$, therefore

$$\begin{aligned}\gamma(\mathfrak{A}) \in \gamma(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) \cap \gamma(\mathrm{Mod}_{\mathrm{Con}}(\psi)) &= \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi)) \cap \mathrm{Mod}_{\mathrm{Con}}(\gamma(\psi)) \\ &= \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi) \wedge \gamma(\psi));\end{aligned}$$

$\supseteq$) since

$$\begin{aligned}\mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi) \wedge \gamma(\psi)) &= \mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi)) \cap \mathrm{Mod}_{\mathrm{Con}}(\gamma(\psi)) \\ &= \gamma(\mathrm{Mod}_{\mathrm{Con}}(\varphi)) \cap \gamma(\mathrm{Mod}_{\mathrm{Con}}(\psi)),\end{aligned}$$

for a generic element $\mathfrak{A}$ of $\mathrm{Mod}_{\mathrm{Con}}(\gamma(\varphi) \wedge \gamma(\psi))$, there are two elements $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\varphi)$ and $\mathfrak{C} \in \mathrm{Mod}_{\mathrm{Con}}(\psi)$ such that

$$\mathfrak{A} = \gamma(\mathfrak{B}) = \gamma(\mathfrak{C}).$$

Since γ is an automorphism, hence injective, $\mathfrak{B} = \mathfrak{C}$, and from the fact that $\mathfrak{B} \in \mathrm{Mod}_{\mathrm{Con}}(\varphi \wedge \psi)$, we get the thesis.

In this sense, an automorphism is also an automorphism of the boolean algebra determined by $B(\mathcal{L})$.[2]

In the following remark, we will exploit this to show a way to determine when a map of $M(\mathcal{L})$ is not an automorphism.

[1] See Remark 1.2.1, pag. 20.

[2] The reader interested in this topic can find further information in [17]. What we will say, will scratch only the surface of this theory. When we deal with $B(\mathcal{L})$ as a boolean algebra, we mean that the intersection is the meet ($\wedge$), the union is the join ($\vee$), the complement (in the set-theoretical sense) is the complement (in the boolean algebra sense), the empty set is the 0 ($\bot$) element, and the set $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ is the 1 ($\top$) element.

Remark 3.1.2. As it is usually done, in a boolean algebra we can introduce the symbol $\leq$ by putting for any two elements of the algebra a and b

$$a \leq b \text{ if and only if } a = a \wedge b.$$

Taking two generic elements of $B(\mathcal{L})$, $\mathrm{Mod}_{\mathrm{Con}}(\varphi)$ and $\mathrm{Mod}_{\mathrm{Con}}(\psi)$, we have that

$$\begin{aligned}\mathrm{Mod}_{\mathrm{Con}}(\varphi) \leq \mathrm{Mod}_{\mathrm{Con}}(\psi) \text{ if and only if } & \mathrm{Mod}_{\mathrm{Con}}(\varphi) = \mathrm{Mod}_{\mathrm{Con}}(\varphi) \cap \mathrm{Mod}_{\mathrm{Con}}(\psi) \\ \text{if and only if } & \mathrm{Mod}_{\mathrm{Con}}(\varphi) \subseteq \mathrm{Mod}_{\mathrm{Con}}(\psi)\end{aligned}$$

Since an automorphism γ of the two-sorted structure $M(\mathcal{L})$, has to preserve the underlying boolean algebra of $B(\mathcal{L})$, it has to preserve $\leq$, too.

In a unary language with q relational symbols (and 2^q atoms $\alpha_1(x), \dots, \alpha_{2^q}(x)$), the sets $\mathrm{Mod}_{\mathrm{Con}}(\beta_i)$ where $i = 1, \dots, 2^q$ and $\beta_i := \forall x \alpha_i(x)$ are composed by only one structure. Therefore, there is no satisfiable $\psi \in \mathrm{Sen}(\mathcal{L})$ for which $\mathrm{Mod}_{\mathrm{Con}}(\psi) \subset \mathrm{Mod}_{\mathrm{Con}}(\beta_i)$ because, if $\mathrm{Mod}_{\mathrm{Con}}(\psi) \subset \mathrm{Mod}_{\mathrm{Con}}(\beta_i)$ holds, then $\mathrm{Mod}_{\mathrm{Con}}(\psi) = \emptyset$ and $\psi \equiv \perp$. Hence, the sets $\mathrm{Mod}_{\mathrm{Con}}(\beta_i)$ are the minimal among the non-bottom elements. In a boolean algebra, these are called *atoms*: since we have already used this term to refer to something else, we will call these *boolean atoms* to avoid any confusion.

Actually, these are the only ones: assume indeed, that there is a sentence $\varphi(b_1, \dots, b_m)$ such that $\mathrm{Mod}_{\mathrm{Con}}(\varphi)$ is a boolean atom. By Equation (2.5),

$$\varphi \equiv \bigvee_{k=1}^{l} \Big(\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_{jk}} \wedge \bigwedge_{i=1}^{m} \alpha_{f_{ki}}(b_i) \Big),$$

where $\varepsilon_{jk} \in \{0, 1\}$ and for a generic formula ϕ, ϕ^1 denotes ϕ, and ϕ^0 denotes $\neg\phi$. Since for any formulas φ and ψ, we have that φ implies $\varphi \vee \psi$, then we can assume that $l = 1$: otherwise, φ can't be an atom. Suppose then that

$$\varphi \equiv \bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_j} \wedge \bigwedge_{i=1}^{m} \alpha_{f_i}(b_i),$$

with $\varepsilon_j \in \{0, 1\}$ and $f_i \in \{1, \dots, 2^q\}$. Now we have the following cases:

- φ is a contradiction, so it can't be a boolean atom;
- in the block $\wedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_j}$ only an ε_j is 1 and the others are 0: say that $\varepsilon_{j_0} = 1$. Then the block at issue is equivalent to

  $$\exists x \alpha_{j_0}(x) \wedge \bigwedge_{j \neq j_0} \forall x \neg \alpha_j(x) \equiv \forall x \alpha_{j_0}(x)$$

  and φ is contradictory or equivalent to β_{j_0}, depending on the f_i's;
- in the block $\wedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_j}$, $\varepsilon_j = 0$ for all the j's: then φ is contradictory;
- in the block $\wedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_j}$ there are at least two indexes j_0 and j_1 for which $\varepsilon_{j_0} = \varepsilon_{j_1} = 1$ and φ isn't a contradiction: therefore the sentence $\psi = \varphi \wedge \alpha_{j_0}(b_{m+1})$ is satisfiable and $\mathrm{Mod}_{\mathrm{Con}}(\psi) \subsetneq \mathrm{Mod}_{\mathrm{Con}}(\varphi)$.

In all the cases, we must conclude that if φ is a boolean atom, then there exists an $i \in \{1, \dots, 2^q\}$ for which $\varphi = \beta_i$.

Since an automorphism of a boolean algebra preserves $\leq$, it has to send boolean atoms to boolean atoms. This will be useful to show that not all the bijections in $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ can be extended to an automorphism.

Recalling the example presented at the beginning of the chapter, we can consider the bijection $\gamma : \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) \to \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, such that

$$\begin{aligned}
&\gamma(\mathfrak{A}) \models P_1(a_1) \text{ if and only if } \mathfrak{A} \models \neg P_1(a_1);\\
&\gamma(\mathfrak{A}) \models P_1(a_i) \text{ if and only if } \mathfrak{A} \models P_1(a_i) \text{ for } i \geq 2;\\
&\gamma(\mathfrak{A}) \models P_2(a_j) \text{ if and only if } \mathfrak{A} \models P_2(a_j) \text{ for all } j \in \mathbb{N}^+.
\end{aligned}$$

Using Formalization 1, this map will take a world and change the color of the ball with 1 written on it: if it's red, it becomes a different color; if it isn't red, it becomes red.

Even if γ is a bijection, it can't be an automorphism: indeed, we have the four atoms $\alpha_1(x) = P_1(x) \wedge P_2(x), \alpha_2(x) = P_1(x) \wedge \neg P_2(x), \alpha_3(x) = \neg P_1(x) \wedge P_2(x)$, and $\alpha_4(x) = \neg P_1(x) \wedge \neg P_2(x)$. The boolean atom $\mathrm{Mod}_{\mathrm{Con}}(\beta_1) := \mathrm{Mod}_{\mathrm{Con}}(\forall x(P_1(x) \wedge P_2(x)))$ contains only an element $\mathfrak{A}$ that is mapped by γ to a $\mathcal{L}$-structure that doesn't satisfy any β_i for $i = 1, \dots, 4$. Hence the boolean atom at issue is not mapped to another boolean atom.

After having shown an example of a bijective map that is not an automorphism, we will present some class of automorphisms of $M(\mathcal{L})$ to get the reader acquainted with this notion.

- *Automorphism that permutes contants:*
  Given a permutation σ of $\mathbb{N}^+$ we can construct the following map $\gamma : \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) \to \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$: for any structure $\mathfrak{A}$, $\gamma(\mathfrak{A})$ is the structure with Con as domain and the following interpretation of any predicate symbol R

  $$\gamma(\mathfrak{A}) \models R(a_i) \text{ if and only if } \mathfrak{A} \models R(a_{\sigma(i)}),$$

  for any $i \in \mathbb{N}^+$. This map is bijective, since σ is invertible and we can perform the same construction with σ^{-1}; furthermore, for any $\varphi(a_{i_1}, \dots, a_{i_n}) \in \mathrm{Sen}(\mathcal{L})$, we have that

  $$\gamma(\mathfrak{A}) \models \varphi(a_{i_1}, \dots, a_{i_n}) \text{ if and only if } \mathfrak{A} \models \varphi(a_{\sigma(i_1)}, \dots, a_{\sigma(i_n)}),$$

  hence

  $$\gamma(\mathrm{Mod}_{\mathrm{Con}}(\varphi(a_{\sigma(i_1)}, \dots, a_{\sigma(i_n)}))) = \mathrm{Mod}_{\mathrm{Con}}(\varphi(a_{i_1}, \dots, a_{i_n}))$$

  and putting $\gamma(\varphi(a_{i_1}, \dots, a_{i_n})) = \varphi(a_{\sigma^{-1}(i_1)}, \dots, a_{\sigma^{-1}(i_n)}))$, we get an automorphism of $M(\mathcal{L})$.

- *Automorphism that permutes predicate symbols:*
  In the unary case, the predicate symbols $R_1, \dots, R_q$ have the same ariety. For any σ permutation of $\{1, \dots, q\}$, we can construct the following map $\gamma : \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) \to \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$: for any structure $\mathfrak{A}$, $\gamma(\mathfrak{A})$ is the structure with Con as domain and the following interpretation of predicate symbols

  $$\gamma(\mathfrak{A}) \models R_i(b) \text{ if and only if } \mathfrak{A} \models R_{\sigma(i)}(b),$$

for any $b \in \mathrm{Con}$. Then, arguing in a similar way as the previous point, we can show that γ can be extended to an automorphism of $M(\mathcal{L})$;

- *Automorphism that negates the occurrences of a predicate symbol:*
For any $i \leq q$, we can construct the following map $\gamma : \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) \to \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$: for any structure $\mathfrak{A}$, $\gamma(\mathfrak{A})$ is the structure with Con as domain such that the interpretations of R_j with $j \neq i$ is the same as in $\mathfrak{A}$, but

$$\gamma(\mathfrak{A}) \vDash R_i(b_1, \ldots, b_n) \text{ if and only if } \mathfrak{A} \vDash \neg R_i(b_1, \ldots, b_n),$$

for any $b_1, \ldots, b_n \in \mathrm{Con}$. Notice that in this way we have also that

$$\gamma(\mathfrak{A}) \vDash \neg R_i(b_1, \ldots, b_n) \text{ if and only if } \mathfrak{A} \vDash R_i(b_1, \ldots, b_n).$$

As before, γ extends to an automorphism of $M(\mathcal{L})$.

The invariance of a probability w under these classes of automorphisms is a valuable property: indeed, it corresponds to requiring w to satisfy, respectively, Ex, Px (in the unary case), and SN. We will state this more precisely after the definition of the *invariance* principle.

Definition 3.1.2. A probability w satisfies the *invariance principle* (INV) if and only if for any γ automorphism of the two-sorted structure $M(\mathcal{L})$ and for any $\varphi \in \mathrm{Sen}(\mathcal{L})$, $w(\varphi) = w(\gamma(\varphi))$.

Remark 3.1.3. We can present INV in another way, exploiting the correspondence given by Theorem 1.2.2. Considering the measurable space $(\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}), \mathcal{F}(B(\mathcal{L})))$ introduced at page 21, an automorphism γ is measurable: this is essentially based on the fact that, since also γ^{-1} is an automorphism, we have that the preimage under γ of $\mathrm{Mod}_{\mathrm{Con}}(\varphi)$ is $\mathrm{Mod}_{\mathrm{Con}}(\gamma^{-1}(\varphi)) \in \mathcal{F}(B(\mathcal{L}))$. Furthermore, it can be shown that, since any probability w corresponds to a probability μ on $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, the property INV is equivalent to requiring that for any automorphism γ, the pushforward measure $\gamma_*\mu$ is μ.

Proposition 3.1.1. *In the unary case, INV implies Ex, Px, SN, and Ax.*

Proof. We have already described some classes of automorphisms for which INV corresponds to Ex, Px, and SN. The only thing left for us to prove is that INV implies Ax.

Now, let's consider a permutation σ of the set $\{1, \ldots, 2^q\}$ and construct the map γ such that for any structure $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, $\gamma(\mathfrak{A})$ is the structure with the property that for any atom $\alpha_h(x)$ and for any constant $b \in \mathrm{Con}$,

$$\gamma(\mathfrak{A}) \vDash \alpha_h(b) \text{ if and only if } \mathfrak{A} \vDash \alpha_{\sigma(h)}(b).$$

This defines uniquely $\gamma(\mathfrak{A})$ because, rephrasing what we said before, for any constant b, $\gamma(\mathfrak{A}) \vDash R_i(b)$ if and only if there exists $h \in \{1, \ldots, 2^q\}$ for which $\alpha_h(b)$ implies $R_i(b)$, and such that $\mathfrak{A} \vDash \alpha_{\sigma(h)}(b)$: hence, γ is a well-defined function because a structure satisfies exactly one atom on a given constant b. Moreover, since $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, we also have $\gamma(\mathfrak{A}) \vDash \exists x \alpha_h(x)$ if and only if $\mathfrak{A} \vDash \exists x \alpha_{\sigma(h)}(x)$.

Using Equation (2.5), a sentence $\psi \in \mathrm{Sen}(\mathcal{L})$ can be written uniquely as

$$\bigvee_{k=1}^{l} (\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_{jk}} \wedge \bigwedge_{i=1}^{m} \alpha_{f_{ki}}(b_i));$$

then, we have that

$$\gamma(\mathfrak{A}) \vDash \psi \text{ if and only if } \gamma(\mathfrak{A}) \vDash \bigvee_{k=1}^{l} \Big(\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_{jk}} \wedge \bigwedge_{i=1}^{m} \alpha_{f_{ki}}(b_i) \Big)$$

$$\text{if and only if } \mathfrak{A} \vDash \bigvee_{k=1}^{l} \Big(\bigwedge_{j=1}^{2^q} (\exists x \alpha_{\sigma(j)}(x))^{\varepsilon_{jk}} \wedge \bigwedge_{i=1}^{m} \alpha_{\sigma(f_{ki})}(b_i) \Big).$$

Hence, by putting $\gamma(\psi) = \bigvee_{k=1}^{l} \big(\bigwedge_{j=1}^{2^q} (\exists x \alpha_{\sigma^{-1}(j)}(x))^{\varepsilon_{jk}} \wedge \bigwedge_{i=1}^{m} \alpha_{\sigma^{-1}(f_{ki})}(b_i) \big)$, it is clear that γ extends to an automoprhism of $M(\mathcal{L})$.

Any probability w that satisfies INV is invariant for γ, therefore for any constants $b_1, \ldots, b_n$ and for any $h_1, \ldots, h_n \in \{1, \ldots, 2^q\}$, we have

$$w\Big(\bigwedge_{i=1}^{n} \alpha_{h_i}(b_i)\Big) = w\Big(\gamma\Big(\bigwedge_{i=1}^{n} \alpha_{h_i}(b_i)\Big)\Big) = w\Big(\bigwedge_{i=1}^{n} \alpha_{\sigma^{-1}(h_i)}(b_i)\Big);$$

the thesis follows from the generality of the permutation σ. □

Proposition 3.1.1 shows us that all the formalization-based argument presented at the beginning of the chapter is encompassed by INV. However, what was said is not able to justify the whole Invariance Principle, that is, in our opinion, not reasonable. Our thought about INV doesn't rely only on the absence of motivations providing grounds for it, but also on what is implied by this principle.

In the unary case, INV generates also other symmetry principles, not only the ones previously encountered. Actually it leaves us with a single probability $c_0^{\mathcal{L}}$[3] defined as follows:

$$c_0^{\mathcal{L}}\Big(\bigwedge_{i=1}^{n} \alpha_{h_i}(a_i)\Big) = \begin{cases} 1 & \text{if } n = 0; \\ 2^{-q} & \text{if } n \geq 1 \text{ and } h_1 = h_2 = \cdots = h_n, \\ 0 & \text{otherwise.} \end{cases}$$

To show that this is a probability, we verify the conditions (2.3) at page 36:

- for any state description $\Theta(b_1, \ldots, b_m)$ we have that the value assigned to it by $c_0^{\mathcal{L}}$ is $0, 2^{-q}$ or 1, hence in any case it is non-negative;
- since we use the convention that an empty conjunction is equivalent to $\top$, $c_0^{\mathcal{L}}(\top) = 1$ by how the map is defined;
- if $m = 0$, then $\Theta = \top$ and given $r \geq m$, any state description Φ for $b_1, \ldots, b_r$ implies Θ: since there are 2^q state descriptions for which the value given to them by $c_0^{\mathcal{L}}$ is 2^{-q} (one for each atom of the language), and to all the others it is assigned 0, we get that

$$c_0^{\mathcal{L}}(\top) = \sum_{\Phi(b_1,\ldots,b_r) \vDash \top} c_0^{\mathcal{L}}(\Phi(b_1, \ldots, b_r)) = 2^2 \frac{1}{2^q};$$

- if $m > 0$, taking $r \geq m$, constants $b_1, \ldots, b_r$, and a state description $\Theta(b_1, \ldots, b_m) = \bigwedge_{i=1}^{m} \alpha_{h_i}(b_i)$ we have two cases:

[3]See Footnote 11, page 46.

- if all the h_i's are equal to some $j \in \{1, \ldots, 2^q\}$, $c_0^{\mathcal{L}}(\Theta) = 2^{-q}$. Among all the state descriptions $\Phi(b_1, \ldots, b_r)$ that imply Θ, there is only one ($\Phi = \bigwedge_{i=1}^{r} \alpha_j(b_i)$) for which the value assigned by $c_0^{\mathcal{L}}$ is different from 0 and it is 2^{-q}: this yields to the thesis;
- otherwise, suppose that for two indexes $i_0, i_1 \in \{1, \ldots, m\}$ we have that $h_{i_0} \neq h_{i_1}$. In any state description $\Phi(b_1, \ldots, b_r)$ that implies Θ, will appear the conjunct $\alpha_{h_{i_0}}(b_{i_0}) \wedge \alpha_{h_{i_1}}(b_{i_1})$ and Equation (2.2) holds since both the terms of the equality are null.

Proposition 3.1.2. *In the unary case INV is consistent, i.e. given a unary language $\mathcal{L}$, there is a probability ($c_0^{\mathcal{L}}$) that satisfies INV.*

Proof. We will assume in the following that the unary predicate symbols are $R_1, \ldots, R_q$, and we will call, as usual, $\alpha_1(x), \ldots, \alpha_{2^q}(x)$ the atoms.

To show that $c_0^{\mathcal{L}}$ is invariant under any automorphism γ of $M(\mathcal{L})$, we will proceed by steps:

a) using the notation used in Remark 3.1.2, calling β_i the sentence $\forall x \alpha_i(x)$ for any $i = 1, \ldots, 2^q$ (corresponding to the boolean atom $\mathrm{Mod}_{\mathrm{Con}}(\beta_i)$), we have:

$$c_0^{\mathcal{L}}(\beta_i) = c_0^{\mathcal{L}}(\forall x \alpha_i(x)) = \lim_{n \to +\infty} c_0^{\mathcal{L}}(\wedge_{j=1}^{n} \alpha_i(a_j)) = \lim_{n \to +\infty} \frac{1}{2^q} = \frac{1}{2^q};$$

b) given a satisfiable formula ψ_k of the form

$$\psi_k = \bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_{jk}} \wedge \bigwedge_{i=1}^{m} \alpha_{f_{ki}}(b_i)$$

there is at least a j_0 with $\varepsilon_{j_0 k} \neq 0$ and one of the following cases occurs:

i) there exists only the index j_0 for which $\varepsilon_{j_0 k} = 1$ and all the other indexes are zeros. In this case

$$\bigwedge_{j=1}^{2^q} (\exists x \alpha_j(x))^{\varepsilon_{jk}} \equiv \bigwedge_{\substack{j=1 \\ j \neq j_0}}^{2^q} \forall x \neg \alpha_j(x) \wedge \exists x \alpha_{j_0}(x) \equiv \forall x \alpha_{j_0}(x).$$

In this case, since ψ_k is satisfiable, all the f_{ki} should be equal to j_0 and ψ_k is equivalent to β_{j_0}. For item a), $c_0^{\mathcal{L}}(\psi_k) = 1/2^q$ in this case;

ii) there exist j_0, j_1 different such that $\varepsilon_{j_0} = \varepsilon_{j_1} = 1$. In this case

$$\begin{aligned} c_0^{\mathcal{L}}(\psi_k) &\leq c_0^{\mathcal{L}}(\exists x \alpha_{j_0}(x) \wedge \exists x \alpha_{j_1}(x)) = c_0^{\mathcal{L}}(\exists x \exists y (\alpha_{j_0}(x) \wedge \alpha_{j_1}(y))) \\ &= \sup_{n \in \mathbb{N}^+} c_0^{\mathcal{L}}\Big(\bigvee_{p \leq n} \exists y (\alpha_{j_0}(a_p) \wedge \alpha_{j_1}(y))\Big) \\ &= \sup_{n \in \mathbb{N}^+} \sup_{m \in \mathbb{N}^+} c_0^{\mathcal{L}}\Big(\bigvee_{p \leq n} \bigvee_{r \leq m} (\alpha_{j_0}(a_p) \wedge \alpha_{j_1}(a_r))\Big) = 0; \end{aligned}$$

c) suppose to have an automorphism γ of $M(\mathcal{L})$ and ψ_k a satisfiable sentence as in the previous item:

i) since an automorphism of a boolean algebra maps a boolean atom to another one, if we are in the case b)-i) with $\psi_k \equiv \beta_{j_0}$, then $\gamma(\psi_k)$ can only be β_j for some $j \in \{1, \dots, 2^q\}$. In this case, however, for a), $c_0^{\mathcal{L}}(\psi_k) = c_0^{\mathcal{L}}(\gamma(\psi_k))$;

ii) if we are in the case b)-ii), then by Equation (2.5), we can write $\gamma(\psi_k)$ as $\vee_{t=1}^{s}\theta_t$ where the θ_t's are conjunctions of the form

$$\bigwedge_{j=1}^{2^q}(\exists x \alpha_j(x))^{\varepsilon_{jt}} \wedge \bigwedge_{i=1}^{m} \alpha_{g_{ti}}(b_i).$$

Since ψ_k is satisfiable, we can also assume that $s \geq 1$ and that all the θ_t's are satisfiable.

The key observation is that for no t, θ_t can be equivalent to a β_j: indeed, otherwise, assuming that $\theta_{t_0} \equiv \beta_j$ and that $\gamma^{-1}(\beta_j) = \beta_i$, since

$$\psi_k = \gamma^{-1}\big(\bigvee_{t=1}^{s} \theta_t\big) \equiv \gamma^{-1}(\beta_j) \vee \bigvee_{\substack{t=1 \\ t \neq t_0}}^{s} \gamma^{-1}(\theta_t),$$

ψ_k is equivalent to a disjunction where a disjoint is β_i. If this was true, then

$$c_0^{\mathcal{L}}(\psi_k) \geq c_0^{\mathcal{L}}(\beta_i) = \frac{1}{2^q},$$

but we know that $c_0^{\mathcal{L}}(\psi_k) = 0$.

Hence, for any t, $c_0^{\mathcal{L}}(\theta_t) = 0$ for what was said in item b), and, since the θ_t's are disjoint,

$$c_0^{\mathcal{L}}(\gamma(\psi_k)) = \sum_{t=1}^{s} c_0^{\mathcal{L}}(\theta_t) = 0 = c_0^{\mathcal{L}}(\psi_k);$$

e) By Equation (2.5), a generic sentence ψ can be written as a disjunction $\vee_{k=1}^{l}\psi_k$ with ψ_k as in item b) and $\psi_k \wedge \psi_h \equiv \bot$ for $h \neq k$. If the disjunction is empty, then ψ is a contradiction and $\gamma(\psi)$ is too, hence in this trivial case

$$c_0^{\mathcal{L}}(\psi) = c_0^{\mathcal{L}}(\gamma(\psi));$$

otherwise, in item c) we have already shown that for any k,

$$c_0^{\mathcal{L}}(\psi_k) = c_0^{\mathcal{L}}(\gamma(\psi_k)),$$

hence the thesis, since

$$\begin{aligned} c_0^{\mathcal{L}}(\gamma(\psi)) &= c_0^{\mathcal{L}}\big(\gamma\big(\bigvee_{k=1}^{l} \psi_k\big)\big) = c_0^{\mathcal{L}}\big(\bigvee_{k=1}^{l} \gamma(\psi_k)\big) \\ &\stackrel{*}{=} \sum_{k=1}^{l} c_0^{\mathcal{L}}(\gamma(\psi_k)) = \sum_{k=1}^{l} c_0^{\mathcal{L}}(\psi_k) = c_0^{\mathcal{L}}\big(\bigvee_{k=1}^{l} \psi_k\big) \\ &= c_0^{\mathcal{L}}(\psi), \end{aligned}$$

where $\stackrel{*}{=}$ is justified by the fact that if $\varphi \wedge \psi \equiv \bot$, then $\gamma(\varphi) \wedge \gamma(\psi) \equiv \bot$.

□

Thanks to this proposition, we can show that not all the previously encountered principles are implied by INV. Indeed, the IP principle (page 47) is not satisfied by $c_0^{\mathcal{L}}$: taking two distinct atoms $\alpha_1(x)$ and $\alpha_2(x)$, we would have

$$0 = c_0^{\mathcal{L}}(\alpha_1(a_1) \wedge \alpha_2(a_2)) \overset{(IP)}{=} c_0^{\mathcal{L}}(\alpha_1(a_1)) \cdot c_0^{\mathcal{L}}(\alpha_2(a_2)) = \frac{1}{2^q} \cdot \frac{1}{2^q},$$

from which the contradiction.

Theorem 3.1.1. *Given a unary language $\mathcal{L}$, the only probability that satisfies INV is $c_0^{\mathcal{L}}$.*

Proof. We have already shown in Proposition 3.1.2 that the probability $c_0^{\mathcal{L}}$ satisfies INV in the unary case.

For the converse, for simplicity, we will deal only with the case in which $q = 1$ and the only relational symbol is R: the argument, however, will work also in the general case.

To any structure $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, we can associate a function $f_{\mathfrak{A}} : \mathbb{N}^+ \to \{0, 1\}$ such that

$$f_{\mathfrak{A}}(n) = 1 \text{ if and only if } \mathfrak{A} \vDash R(a_n).$$

The correspondence between structures and functions of this kind is a bijection and this allows us to talk about functions meaning, instead, structures in $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$.

Thinking a map from $\mathbb{N}^+$ to $\{0, 1\}$ as a countable sequence of 0's and 1's, consider the function $\lambda : \{0, 1\}^{\mathbb{N}^+} \to \{0, 1\}^{\mathbb{N}^+}$ defined by

$$\lambda(f) = \begin{cases} 1, 1, f(2), f(3), \ldots & \text{if } f(1) = 1 \\ 0, f(3), f(4), \ldots & \text{if } f(1) = f(2) = 0 \\ 1, 0, f(3), f(4), \ldots & \text{if } f(1) = 0 \text{ and } f(2) = 1. \end{cases}$$

Recalling that we can identify a structure $\mathfrak{A}$ with the function $f_{\mathfrak{A}}$, consider the map $\gamma : \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) \to \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ such that

$$f_{\gamma(\mathfrak{A})} = \lambda(f_{\mathfrak{A}}).$$

This is an automorphism of $M(\mathcal{L})$. Obviously, γ is well-defined as a map from $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$ to itself and it's bijective: indeed, this is equivalent to the bijectivity of λ that is guaranteed by the existence of the inverse

$$\lambda^{-1}(g) = \begin{cases} 1, g(3), g(4), \ldots & \text{if } g(1) = g(2) = 1 \\ 0, 0, g(2), g(3), \ldots & \text{if } g(1) = 0 \\ 0, 1, g(3), g(4), \ldots & \text{if } g(1) = 1 \text{ and } g(2) = 0. \end{cases}$$

Now we begin to define γ on $\mathrm{Sen}(\mathcal{L})$ by defining it on simple sentences: we will see that this is sufficient:

- λ is such that if a function f has a 1 and a 0, then also its image $\lambda(f)$ has a 1 and a 0, and vice versa: this means that we can put

$$\gamma(\exists x R(x) \wedge \exists x \neg R(x)) = \exists x R(x) \wedge \exists x \neg R(x);$$

similar arguments motivate why the following definitions work:

$$\gamma(\exists x R(x) \wedge \neg\exists x \neg R(x)) = \exists x R(x) \wedge \neg\exists x \neg R(x)$$
$$\gamma(\neg\exists x R(x) \wedge \exists x \neg R(x)) = \neg\exists x R(x) \wedge \exists x \neg R(x)$$
$$\gamma(\neg\exists x R(x) \wedge \neg\exists x \neg R(x)) = \neg\exists x R(x) \wedge \neg\exists x \neg R(x) \equiv \bot;$$

Restricting ourselves to these four formulas, we have that for any $\mathfrak{A}$ in $\mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, $\mathfrak{A} \vDash \varphi$ if and only if $\gamma(\mathfrak{A}) \vDash \gamma(\varphi)$.

- $\gamma(R(a_1)) = R(a_1) \wedge R(a_2)$: this definition makes sense, since if $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, then

$$\begin{aligned}\mathfrak{A} \vDash R(a_1) &\text{ if and only if } f_{\mathfrak{A}}(1) = 1\\ &\text{ if and only if } \lambda(f_{\mathfrak{A}}) = 1, 1, f_{\mathfrak{A}}(2), \ldots\\ &\text{ if and only if } \gamma(\mathfrak{A}) \vDash R(a_1) \wedge R(a_2).\end{aligned}$$

- extending γ to the state descriptions of a_1 and a_2, we define

$$\gamma(R(a_1) \wedge R(a_2)) = R(a_1) \wedge R(a_2) \wedge R(a_3)$$
$$\gamma(R(a_1) \wedge \neg R(a_2)) = R(a_1) \wedge R(a_2) \wedge \neg R(a_3)$$
$$\gamma(\neg R(a_1) \wedge R(a_2)) = R(a_1) \wedge \neg R(a_2)$$
$$\gamma(\neg R(a_1) \wedge \neg R(a_2)) = \neg R(a_1).$$

We can reason as above to show that also for a state description φ for a_1, a_2, for every $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, $\mathfrak{A} \vDash \varphi$ if and only if $\gamma(\mathfrak{A}) \vDash \gamma(\varphi)$. As an example we will show it in the case $\varphi = \neg R(a_1) \wedge R(a_2)$:

$$\begin{aligned}\mathfrak{A} \vDash \neg R(a_1) \wedge R(a_2) &\text{ if and only if } f_{\mathfrak{A}} = 0, 1, f_{\mathfrak{A}}(3), \ldots\\ &\text{ if and only if } \lambda(f_{\mathfrak{A}}) = 1, 0, f_{\mathfrak{A}}(3), \ldots\\ &\text{ if and only if } \gamma(\mathfrak{A}) \vDash R(a_1) \wedge \neg R(a_2);\end{aligned}$$

- the process can be carried on and we can extend γ to any state description Θ for $a_1, \ldots, a_n$: Θ will be of the form $\wedge_{i=1}^{n}(R(a_i))^{\varepsilon_i}$ and for each of its model $\mathfrak{A}$, we know the first n values of the sequence defining $f_{\mathfrak{A}}$. For $n \geq 2$ this will be enough to determine $\lambda(f_{\mathfrak{A}})$ up to its first $n-1, n$, or $n+1$ digit depending on the case: then we can construct the sentence $\gamma(\Theta)$ exploiting all the information we have about $\lambda(f_{\mathfrak{A}})$ as we did in the previous two items. In this way, for any $\mathfrak{A} \in \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$, $\mathfrak{A} \vDash \Theta$ if and only if $\gamma(\mathfrak{A}) \vDash \gamma(\Theta)$.

For $b_1, \ldots, b_m$ constants among $a_1, \ldots, a_n$, since any state description for $b_1, \ldots, b_m$ is equivalent to a disjunction of state descriptions for $a_1, \ldots, a_n$, we can extend γ also to state descriptions for generic constants. Using the normal form provided by Equation (2.5), and recalling that this is unique (up to order), we can extend γ to the whole $\mathrm{Sen}(\mathcal{L})$ by putting

$$\gamma\Big(\bigvee_{k=1}^{l}\Big(\bigwedge_{j=1}^{2}(\exists x \alpha_j(x))^{\varepsilon_{jk}} \wedge \bigwedge_{i=1}^{m} \alpha_{f_{ki}}(b_i)\Big)\Big) = \bigvee_{k=1}^{l} \gamma\Big(\bigwedge_{j=1}^{2}(\exists x \alpha_j(x))^{\varepsilon_{jk}}\Big) \wedge \gamma\Big(\bigwedge_{i=1}^{m} \alpha_{f_{ki}}(b_i)\Big),$$

and it is easy to show now that γ is an automorphism of $M(\mathcal{L})$.

By how γ is defined,

$$\gamma(R(a_1) \wedge \neg R(a_2)) = R(a_1) \wedge R(a_2) \wedge \neg R(a_3),$$

hence a probability w that satisfies INV must assign the same value to the two formulas, i.e.

$$w(R(a_1) \wedge \neg R(a_2)) = w(R(a_1) \wedge R(a_2) \wedge \neg R(a_3)). \tag{3.3}$$

By the fact that INV implies Ex and SN, we have that

$$\begin{aligned} w(R(a_1) \wedge \neg R(a_2) \wedge R(a_3)) &\overset{Ex}{=} w(R(a_2) \wedge \neg R(a_1) \wedge R(a_3)) \\ &\overset{SN}{=} w(\neg R(a_2) \wedge R(a_1) \wedge \neg R(a_3)). \end{aligned} \tag{3.4}$$

Then, by combining Equations (3.3), and (3.4), we have

$$\begin{aligned} w(R(a_1) \wedge \neg R(a_2) \wedge R(a_3)) &\overset{Ex}{=} w(R(a_1) \wedge R(a_2) \wedge \neg R(a_3)) \\ &\overset{3.3}{=} w(R(a_1) \wedge \neg R(a_2)) \\ &= w(R(a_1) \wedge \neg R(a_2) \wedge R(a_3)) + w(R(a_1) \wedge \neg R(a_2) \wedge \neg R(a_3)) \\ &\overset{(3.4)}{=} 2w(R(a_1) \wedge \neg R(a_2) \wedge R(a_3)), \end{aligned}$$

hence,

$$w(R(a_1) \wedge \neg R(a_2)) = w(R(a_1) \wedge \neg R(a_2) \wedge R(a_3)) = 0. \tag{3.5}$$

This forces w to be $c_0^{\mathcal{L}}$: indeed, in this setting, we have only two atoms, $\alpha_1(x) = R(x)$ and $\alpha_2(x) = \neg R(x)$ and:

- if the h_i's are not all equal $w(\wedge_{i=1}^n \alpha_{h_i}(a_i)) = 0$. Namely, supposing that there are j and k with $h_j \neq h_k$, using Ex, Ax, and Equation (3.5) we have

$$\begin{aligned} w(\wedge_{i=1}^n \alpha_{h_i}(a_i)) \leq w(\alpha_{h_j}(a_j) \wedge \alpha_{h_k}(a_k)) &\overset{Ax}{=} w(\alpha_1(a_j) \wedge \alpha_2(a_k)) \\ &\overset{Ex}{=} w(\alpha_1(a_1) \wedge \alpha_2(a_2)) \overset{(3.5)}{=} 0; \end{aligned}$$

- if the h_i's are all equal $w(\wedge_{i=1}^n \alpha_{h_i}(a_i)) = \frac{1}{2}$. Namely, for the constants $a_1, \ldots, a_n$ there are 2^n possible state descriptions and in all, but two, the atoms that appear are not all equal. Hence,

$$\begin{aligned} 1 = w\Big(\bigvee_{g:\{1,\ldots,n\}\to\{1,2\}} \bigwedge_{i=1}^n \alpha_{g(i)}(a_i)\Big) &= \sum_{g:\{1,\ldots,n\}\to\{1,2\}} w\Big(\bigwedge_{i=1}^n \alpha_{g(i)}(a_i)\Big) \\ = w\Big(\bigwedge_{i=1}^n \alpha_1(a_i)\Big) + w\Big(\bigwedge_{i=1}^n \alpha_2(a_i)\Big). \end{aligned}$$

  Using the fact that w must satisfy also Ax, $w(\wedge_{i=1}^n \alpha_1(a_i))$ and $w(\wedge_{i=1}^n \alpha_2(a_i))$ are equal and then we get the thesis.

□

3.2 Symmetry in the polyadic case

Even if the previous section was devoted to the unary case, some of the definitions given can be carried out also for a polyadic language.

We started by noticing that some symmetry principles can be easily extended to the polyadic context: the *Strong Negation Principle* and the *Predicate Exchangeability Principle*. SN doesn't need any changes, but we have to slightly adjust Px:

- *Px - Polyadic case:* if R and R' are two relational symbols with the same ariety and for any statement φ, φ' is the outcome of replacing in φ any occurrence of R with R' and any occurrence of R' with R, then $w(\varphi) = w(\varphi')$.

It is easy to check that Px for the polyadic case also includes the principle for the unary one: in the latter, indeed, all the predicate symbols have the same ariety. Thus, we will name Px the principle expressed by the formulation above, since it is more general.

In the polyadic setting, the argument about different formalizations gives us another principle, the *Variable Exchangeability* one (Vx):

- *Vx:* if R is a relational symbol with ariety n, σ is a permutation of $\{1, \ldots, n\}$ and for any statement φ, φ' is the outcome of replacing in φ any occurrence of $R(t_1, \ldots, t_n)$ with $R(t_{\sigma(1)}, \ldots, t_{\sigma(n)})$, where $t_1, \ldots, t_n$ are $\mathcal{L}$-terms, then $w(\varphi) = w(\varphi')$.

The principle relies on the fact that for any permutation σ of $\{1, \ldots, n\}$, given a formalization such that $R(b_1, \ldots, b_n)$ represents that the n-uple $(b_1, \ldots, b_n)$ satisfies the property at issue, there is also a formalization such that $R(b_1, \ldots, b_n)$ represents that the n-uple $(b_{\sigma(1)}, \ldots, b_{\sigma(n)})$ satisfies the property.

Notice that Vx is different from Ex: given a language $\mathcal{L}$ with a binary relational symbol R, if w satisfies Ex, then $w(R(a_1, a_2)) = w(R(a_3, a_7))$, whereas Vx guarantees only that $R(a_1, a_2)$ and $R(a_2, a_1)$ have the same assigned value. However, even if from the example above, It seems that Ex implies Vx, this is false: for instance, according to Vx, $\exists x R(x, a_1)$ and $\exists x R(a_1, x)$ must be regarded in the same way, while from Ex we can't derive this.

More precisely, in the unary case, every probability satisfies Vx (because the transform φ' is the same of φ and the principle trivially holds) even if not all satisfy Ex, proving that Vx doesn't imply Ex.

Conversely, consider the languages $\mathcal{L} = \{R\}$ and $\mathcal{L}' = \{P\}$ with R binary and P unary. Recall now the probability $c_2^{\mathcal{L}'}$ of Example 2.2.1 and consider w probability on $\text{Sen}(\mathcal{L})$ such that for any $\varphi \in \text{Sen}(\mathcal{L})$,

$$w(\varphi) = c_2^{\mathcal{L}'}({}^*\varphi),$$

where ${}^*\varphi$ is the outcome of replacing in φ any occurence of $R(t_1, t_2)$ with $P(t_1)$. Then:

- w is a probability that satisfies Ex: this relies on the fact that $c_2^{\mathcal{L}'}$ is a probability that satisfies Ex (as seen in the Example 2.2.1) and that if φ is a tautology, then also ${}^*\varphi$ is (if we have a $\mathcal{L}'$-model for ${}^*\varphi$ it is easy to construct a $\mathcal{L}$-model for φ); the details are left to the reader;

- w doesn't satisfy Vx: indeed,

$$w(R(a_1,a_2) \wedge \neg R(a_1,a_3)) = c_2^{\mathcal{L}'}(P(a_1) \wedge \neg P(a_1)) = 0;$$
$$w(R(a_2,a_1) \wedge \neg R(a_3,a_1)) = c_2^{\mathcal{L}'}(P(a_2) \wedge \neg P(a_3)) \overset{(2.12)}{=} \frac{1}{6},$$

then $w(R(a_1,a_2) \wedge \neg R(a_1,a_3)) \neq w(R(a_2,a_1) \wedge \neg R(a_3,a_1))$.

The definitions of the structure $M(\mathcal{L})$, of its automorphisms (Definition 3.1.1), and of INV (Definition 3.1.2) are still meaningful for a language $\mathcal{L}$ that is not unary.

Also in the polyadic case, we have that INV implies Ex, SN, and Px; furthermore, even Vx can be deduced by the Invariance Principle. It is easy, following the previous examples, to define the classes of automorphisms needed to prove this. Hoping that an example will be useful, to justify Px, we will present the set of automorphisms that permutes predicate symbols of the same ariety.

Suppose that among $R_1, \ldots, R_q$ there are $p \leq q$ relational symbols with the same ariety, say $R_1, \ldots, R_p$. For any permutation σ of $\{1, \ldots, p\}$, we can construct the following map $\gamma : \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L}) \to \mathrm{Mod}_{\mathrm{Con}}(\mathcal{L})$: for any structure $\mathfrak{A}$, $\gamma(\mathfrak{A})$ is the structure with Con as domain and the following interpretation of predicate symbols

$$\gamma(\mathfrak{A}) \vDash R_i(b_1, \ldots, b_n) \text{ if and only if } \mathfrak{A} \vDash R_{\sigma(i)}(b_1, \ldots, b_n) \text{ if } i \leq p;$$
$$\gamma(\mathfrak{A}) \vDash R_j(b_1, \ldots, b_n) \text{ if and only if } \mathfrak{A} \vDash R_j(b_1, \ldots, b_n) \text{ if } j > p,$$

for any $b_1, \ldots, b_n \in \mathrm{Con}$. The map γ can be extended to an automorphism of $M(\mathcal{L})$: from the generality of p and σ, we have that the invariance under this class of automorphisms is exactly Px.

The question of whether INV restricts the range of available probabilities to only one also in the polyadic case, was open for a long time. Recently, however, it was proved that it is so.[4]

The probability we will get at the end is a generalization of $c_0^{\mathcal{L}}$ to the polyadic setting. The definition we gave of $c_0^{\mathcal{L}}$ is essentially based on the concept of atoms, a notion meaningful only in the unary context: however, the properties defining $c_0^{\mathcal{L}}$ can be rephrased in a more generalizable way.

In the unary case, for any atom $\alpha_i(x)$ we get a structure $\mathfrak{A}_i$ in which $\forall x \alpha_i(x)$ holds. We can see $c_0^{\mathcal{L}}$ as the probability on $\mathrm{Sen}(\mathcal{L})$ such that

$$w = \sum_{i=1}^{2^q} \frac{1}{2^q} w_{\mathfrak{A}_i},$$

with the notation used in Example 1.0.1. The structures $\mathfrak{A}_i$ are the only ones in which all the elements of the domain are indistinguishable: the whole behavior of a constant $b \in \mathrm{Con}$ is represented by the atom it satisfies when the language is unary. Hence, if $x \sim y$ is the formula

$$\bigwedge_{i=1}^{2^q} (\alpha_i(x) \equiv \alpha_i(y)),$$

[4] This information is based on personal communication with J. Paris and A. Vencovská: the results are under publications ([19]).

the $\mathfrak{A}_i$'s are the different models of $\forall x \forall y\ x \sim y$.

Considering a polyadic language $\mathcal{L}$ that, for simplicity of notation, we will suppose to have only a binary predicate R, $x \sim y$ is the formula

$$\forall z(R(x,z) \equiv R(y,z) \wedge R(z,x) \equiv R(z,y)).$$

Then we have the structures $\mathfrak{B}_0$ and $\mathfrak{B}_1$ such that

$$\mathfrak{B}_0 \models \forall x \forall y(x \sim y \wedge R(x,x)) \quad \mathfrak{B}_1 \models \forall x \forall y(x \sim y \wedge \neg R(x,x)),$$

and the probability ω that generalizes $c_0^{\mathcal{L}}$ to a polyadic language $\mathcal{L} = \{R\}$ is

$$\omega = \frac{1}{2} w_{\mathfrak{B}_0} + \frac{1}{2} w_{\mathfrak{B}_1}.$$

Theorem 3.2.1. *ω is the only probability that satisfies INV.*

3.3 Discussion and open problems

In the previous sections, we saw that the INV principle leaves us with a single probability. The golden goal of this branch of logic is finding rational principles to establish how much an agent should believe in a fact, i.e. in a logical formula that formalizes the fact: thus, have we pursued our aim?

The fact that we have found only one available probability consistent with INV seems to positively answer the question. Actually, this is not the case at all. Dealing with the unary case, indeed, $c_0^{\mathcal{L}}$ has some unpleasant faults; we will see this using the example of the urn full of balls used along the work:

- $c_0^{\mathcal{L}}$ generalizes too much: if the agent discovers that a ball is red, then all the balls are red according to it. This relies on the fact that for any relational symbol P of our language, $c_0^{\mathcal{L}}(-|P(a_i))$ is a probability that gives 0 to all the sentences of the form $\neg P(a_j)$, for all $i, j \in \mathbb{N}^+$;
- $c_0^{\mathcal{L}}$ can't be used in the process of learning: suppose that the agent discovers that a ball is red and another one is not. Then, assuming we are using the natural Formalization 1 of the problem, the agent finds out that there are $i, j \in \mathbb{N}^+$ such that the sentence $P_1(a_i) \wedge \neg P_1(a_j)$ is true. If we want to adjust the prior probability after this observation, we should take the conditional probability $c_0^{\mathcal{L}}(-|P_1(a_i) \wedge \neg P_1(a_j))$ but, as already discussed in the previous point, $c_0^{\mathcal{L}}(P_1(a_i) \wedge \neg P_1(a_j)) = 0$.

Similarly, we can argue that also in the polyadic setting, Theorem 3.2.1 isn't a good result as hoped.

Hence, if our formulation of degrees of belief in terms of probabilities on $\mathrm{Sen}(\mathcal{L})$ is right, we must conclude that INV is a too-strong principle.

One path that can be run is trying to define a subclass of automorphisms the invariance under which should give a rational principle. There are some tentative definitions that deserve more study; we will present briefly one of these, the *Permutation Invariance Principle (PIP).*

Definition 3.3.1. An automorphism γ of $M(\mathcal{L})$ *permutes state formulas* if for any $n \in \mathbb{N}^+$ there is a permutation $f_{\gamma,n}$ of the set of state formulas over variables $x_1, \ldots, x_n$ such that, for any state formula $\Theta(x_1, \ldots, x_n)$ and for any constants $a_{i_1}, \ldots, a_{i_n}$, we have

$$\gamma(\Theta(a_{i_1}, \ldots, a_{i_n})) = f_{\gamma,n}(\Theta)\{x_1/a_{i_1}, \ldots, x_n/a_{i_n}\}.$$

A probability w on $\mathrm{Sen}(\mathcal{L})$ satisfies the *Permutation Invariance Principle* if for any γ automorphism of $M(\mathcal{L})$ that permutes state formulas, and for any $\varphi \in \mathrm{Sen}(\mathcal{L})$,

$$w(\varphi) = w(\gamma(\varphi)).$$

This weakening of INV agrees with how we decided to present the topic of symmetry in this work. Indeed, at the beginning of the chapter, we preferred to stress out the rationality of some principles (Ex, Px, SN, Vx) using arguments based on the different formalizations available for a given problem and on the ignorance of an agent in the zero-knowledge condition.

This kind of argument[5], which seems really reliable to us, doesn't justify, as already noticed, the whole INV principle: for instance, the automorphism γ used in the proof of Theorem 3.1.1 seems different from the ones that could be born from reasons related to different formalizations.

It makes sense, then, to find other principles, not as strong as INV, that can imply the principles deduced by such arguments: PIP accomplishes this request, i.e. it is sufficient to deduce Ex, Px, SN, Vx and is strictly weaker than INV (see [23]). Furthermore, it can be proved that PIP is stronger than the ones in the list above: thus, the open question is whether PIP rules out other principles that come out from INV but that are not considered so rational.

Instead of focusing on a comprehensive symmetry principle that encompasses all the most accepted ones, we can try to detect other "notions" from which we can derive rationality: for instance, we want that an agent ignores *irrelevant knowledge* in forming beliefs regarding a fact. In the unary case, this will lead us to principles like the

- *Johnson's Sufficientness Postulate (JSP):*[6]
  the value
  $$w\big(\alpha_j(a_{n+1}) \big| \bigwedge_{i=1}^{n} \alpha_{h_i}(a_i)\big)^7$$
  depends only on n and on the cardinality of the set $\{i \ : \ h_i = j\}$.

This principle can be stated in this way: knowing which atoms are satisfied by $a_1, \ldots, a_n$, to decide which degree of belief should be assigned to $\alpha_j(a_{n+1})$, an agent should ignore all the information but n and the number of the constants among $a_1, \ldots, a_n$ that satisfy the j-th atom.

The difference between *symmetry* and *irrelevance* isn't neat and some principles follow from both arguments: an example of this is Ax. We have already shown that INV implies

[5]They seem to be encompassed by another principle called *Conformity*: for further details see [20], Chapter 28.

[6]See [13] and [20] for further details.

[7]Using the convention that in the case in which $w(\psi) = 0$, $w(\varphi|\psi) = c$ means that $w(\varphi \wedge \psi) = cw(\psi)$.

it (and PIP does too, in the unary setting) and we leave to the reader the proof that we can derive it also from JSP.

What was said before about irrelevance regards the unary context; for the polyadic case, the question is a bit more delicate and will lead us to another important open problem.

In the polyadic setting, we can't talk about atoms and we have difficulties already in generalizing the Ax principle in this context: in the following, we will try to rephrase it in order to extend it. The principle that comes out of this irrelevance argument plays an important role in one of the major actual questions on the subject.

Notice that when the probability w satisfies Ex and Ax, the value $w(\wedge_{i=1}^m \alpha_{h_i}(b_i))$ depends only on the multiset $\{m_1, \ldots, m_{2^q}\}$, where

$$m_j := \{i \in \{1, \ldots, 2^q\} \ : \ h_i = j\}.$$

Hence, in this case, the state description $\Theta = \wedge_{i=1}^m \alpha_{h_i}(b_i)$ induces a partition on the set $\{b_1, \ldots, b_m\}$ depending on which constants are indistinguishable one from the other using the information given by Θ. The various m_i's are only the size of the equivalence classes that Θ creates. This motivates the following definition.

Definition 3.3.2. Given a state description $\Theta(b_1, \ldots, b_m)$, we define the equivalence relation $\sim_\Theta$ on the set $\{b_1, \ldots, b_m\}$, by putting

$$b_i \sim_\Theta b_j \text{ if and only if } \Theta(b_1, \ldots, b_m) \wedge b_i = b_j \text{ is consistent.}^8$$

The *spectrum* of Θ, denoted by $\mathcal{S}(\Theta)$, is the multiset of the sizes of the $\sim_\Theta$-equivalence classes on $\{b_1, \ldots, b_m\}$.

A probability w that satisfies Ex, satisfies the *Spectrum Exchangeability Principle (Sx)* if for any two state descriptions $\Theta(b_1, \ldots, b_m)$ and $\Theta'(b_1, \ldots, b_m)$ with the same spectrum, $w(\Theta) = w(\Theta')$.

In the unary case, Sx corresponds to Ax for probabilities that satisfy Ex.

Example 3.3.1. Given a language with a binary relational symbol R, then:

- if $\Theta(a_1, a_2, a_3)$ is the state description $\wedge_{i=1}^3 \wedge_{j=1}^3 R(a_i, a_j)$, then according to Θ the constants are all indistinguishable and then $\mathcal{S}(\Theta) = \{3\}$;
- if $\Theta(a_1, a_2, a_3, a_4)$ is the conjunction of the following atomic formulas

$$\begin{array}{cccc} R(a_1,a_1) & \neg R(a_1,a_2) & R(a_1,a_3) & R(a_1,a_4) \\ R(a_2,a_1) & \neg R(a_2,a_2) & R(a_2,a_3) & \neg R(a_2,a_4) \\ R(a_3,a_1) & \neg R(a_3,a_2) & R(a_3,a_3) & R(a_3,a_4) \\ R(a_4,a_1) & R(a_4,a_2) & R(a_4,a_3) & R(a_4,a_4) \end{array}$$

  then the equivalence classes induced by $\sim_\Theta$ are $\{a_1, a_3\}, \{a_2\}, \{a_4\}$, hence $\mathcal{S}(\Theta) = \{2, 1, 1\}$.

[8] Here *consistent* is with regard to the equality axioms.

Sx can be regarded as an irrelevance principle: in evaluating $w(\Theta)$ nothing matters but the spectrum of the state formula. When we talk about symmetry, this becomes an open problem: does Sx derive from a symmetry principle? Is there a class of automorphisms the invariance under which corresponds to Sx?

By what is known now, Sx is strictly weaker than INV, since more probabilities satisfy Sx but only one that satisfies INV, and however it is strictly stronger than PIP. These properties give rise to hope for Sx as a principle that can capture rationality and encompass symmetry and irrelevance, but there is a lot of work still to do.

In these last lines, we want to come back to the introduction and to the main reason why we are interested in this topic, i.e. AI. In general, inductive reasoning has been used in symbolic AI[9] for a long time: Inductive Logic Programming, for instance, is an established field in computer science and it tries to detect ways to generate valid hypothesis from examples and background knowledge.

Hence, the importance of the topic in AI is well known and some approaches have more applications than others. How the subject is presented here (and in all the relative sources cited) is not the mainstream way in which it is used in AI, even if some attempts in this sense can be found in the literature ([12], [8], [2], [5]).

It's difficult to understand why this is so, but let's try to outline some criticism of this approach. Maybe the most significant issue is that there is no agreement on which principles are rational: hence, how do we test an agent's beliefs? Even from a theoretical point of view, so far, a clear answer isn't provided.

However, even if we don't have a principle that encompasses the whole notion of rationality, most scholars agree that Definition 1.0.1 is a requirement that all rational agents should satisfy: can we *practically* decide if an agent's beliefs define a probability? We should be able to make comparisons between the AI's outputs and the mathematical language used to talk about probabilities. Consider, for instance, the property "every tautology must be assigned to 1": it appears as a mathematical feature of a mathematical object. In the case of an agent, we have to define a suitable language $\mathcal{L}$ and represent numerically how much the agent believes in a tautology $\varphi \in \mathrm{Sen}(\mathcal{L})$, in order to get the correspondent probability w and then verify if $w(\varphi) = 1$. These two tasks are really difficult: chatbots like ChatGPT can deal with a lot of different topics so that in every choice of the language $\mathcal{L}$, a restriction is implicit; furthermore, how we can recover the values assigned to sentences by the "system of thoughts" of an agent? Do they exist? Considering the already mentioned example of ChatGPT, the reader may think that we can ask it how much, in a range between 0 and 1, it believes in a fact φ and take the answer as a value: this naïve attempt doesn't work because ChatGPT doesn't compute effectively this probability when it receives this question; its goal is to forecast the "world"that most likely follows our text, therefore, its answer can't be reliable for our aim.

Moreover, if we want to take into account some principles, even the ones that are accepted by most scholars (like Ex, SN, Px, or Vx), we get a problem: these principles are considered rational in a zero-knowledge condition, but ChatGPT or any agent that receives an *instruction* from a dataset is not in this situation: if the agent *learns* that

[9]There are two main approaches to AI: in the symbolic one an agent is created starting from human-readable representations of knowledge and reasoning (the tools generally used are logic programming, semantic nets, and so on); in the sub-symbolic one (sometimes called data-driven), the agent *learns* using a huge amount of data some patterns in them (the tools generally used are machine-learning techniques).

Italy is in Europe and China not, it can't deal with "Italy" and "China" in the same way, and Ex (as well as the other principles) is not well motivated anymore.

To sum up, the principal issues we find in a possible application of this approach to Inductive Logic to AI safety regards essentially:

- no agreement on which are the rational principles;
- difficulty to formalize AI's outputs or, in general, to find a common framework in which we can express both the mathematical theory here presented and the language of an agent;
- heavy assumptions (as the zero-knowledge condition) on which the motivations of the principles rely.

In our opinion, however, the study of this theory can be food for thought for AI safety and it is a solid theoretical background to deal with, even if in a narrow framework, rationality. We think that some concepts can be used in the AI field and, anyway, that this theory sheds an important light on a topic that will be taken more and more seriously as smarter devices are developed.

Appendix A

Measure Theory

A.1 Introduction to Measure Theory

In Measure Theory, we need some structure in the function domain to regard a map as a probability.

Definition A.1.1. Let Ω be a non-empty set and $\mathcal{C} \subseteq \mathcal{P}(\Omega)$.

- $\mathcal{C}$ is an algebra if it is closed under all the finite boolean combinations (finite union, finite intersection, and complementation)[1];
- $\mathcal{C}$ is a σ-algebra if it is closed under all the finite and countable boolean combinations.

A pair $(\Omega, \mathcal{C})$ where $\mathcal{C}$ is a σ-algebra is called *measurable space*.

Classical examples of algebras or σ-algebras are:

- for any set Ω the set $\mathcal{P}(\Omega)$ and the set $\{\emptyset, \Omega\}$ are σ-algebras;
- for any infinite set Ω the set $\mathcal{A}$ of all the subsets $A \subseteq \Omega$ for which either A or A^c is finite is an algebra. However, in general, it is not a σ-algebra.
- if (X, τ) is a topological space, in general, τ is not closed under complementation, hence it is not an algebra. However, we can consider the σ-algebra $\mathcal{F}(\tau)$ generated by τ: this is the smallest σ-algebra that contains τ and it can be described also as the intersection of all the σ-algebras $\mathcal{F}'$ such that $\tau \subseteq \mathcal{F}'$. When we work in $\mathbb{R}^d$ and τ_d denote the standard euclidean topology (when $d = 1$ this is usually called simply τ) in $\mathbb{R}^d$, then the σ-algebra $\mathcal{F}(\tau_d)$ is usually called *borelian σ-algebra* and it is also denoted by $\mathcal{B}(\mathbb{R}^d)$. In the following, we will sometimes work with the extended reals $\overline{\mathbb{R}} = [-\infty, +\infty]$, i.e. $\mathbb{R} \cup \{-\infty, +\infty\}$. The topology $\overline{\tau}$ we endow this set with, is the outcome of adding to the standard one τ for $\mathbb{R}$, the sets $[-\infty, a)$ and $(b, +\infty]$, for $a, b \in \mathbb{R}$ and then taking the closure under arbitrary union. Even the σ-algebra generated by the standard topology $\overline{\tau}$ on $\overline{\mathbb{R}}$, will be called the *borelian σ-algebra*.

Once we have these kinds of structures on Ω, we can define a *measure*.

[1] We are including the empty union (i.e. $\emptyset$) and the empty intersection (i.e. the entire Ω).

Definition A.1.2. If $\mathcal{C}$ is an algebra or a σ-algebra on Ω, a *finitely-additive measure* on $\mathcal{C}$ is a map $\mu : \mathcal{C} \to [0, +\infty]$ such that:

1) it is non-trivial, i.e. $\mu(\emptyset) \neq \mu(\Omega)$;

2) it is finitely-additive, i.e. for all $n \in \mathbb{N}$ and for all pairwise disjoint elements $A_1, \ldots, A_n$ in $\mathcal{C}$,
$$\mu(\dot{\bigcup}_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i).$$[2]

A finitely-additive measure μ is a *measure* if, in addition to the properties above, it satisfies also the *conditionally σ-additivity*, i.e.

3) if $\{A_i\}_{i\in\mathbb{N}}$ is a countable set of pairwise disjoint elements in $\mathcal{C}$ and if the union $\dot{\bigcup}_{i\in\mathbb{N}} A_i$ belongs to $\mathcal{C}$, then $\mu(\dot{\bigcup}_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu(A_i)$.

If the conditionally σ-additive measure μ is defined on a σ-algebra, the previous condition simplifies to

3') if $\{A_i\}_{i\in\mathbb{N}}$ is a countable set of pairwise disjoint elements in $\mathcal{C}$, then $\mu(\dot{\bigcup}_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu(A_i)$.

In all the definitions above, if $\mu(\Omega) < +\infty$, the measure will be called *finite*; if $\mu(\Omega) = 1$ we talk about *probability* instead of measure. A measure is called *σ-finite* if Ω can be written as a countable union $\cup_{i\in\mathbb{N}} A_i$ of elements $A_i \in \mathcal{C}$ such that $\mu(A_i) < +\infty$ for any $i \in \mathbb{N}$.

From this definition some basic properties of a measure follow like, for instance, the fact that the range of a probability is a subset of $[0, 1]$, that for a measure μ, $\mu(\emptyset) = 0$ and the finitely-additive property is equivalent to require that for any disjoint $A, B \in \mathcal{C}$, $\mu(A \cup B) = \mu(A) + \mu(B)$.

In the case of a probability μ defined on an algebra/σ-algebra $\mathcal{C}$ on Ω, we will usually say that μ is a probability on Ω when $\mathcal{C}$ is obvious from the context.

Some examples of probabilities are:

- given a space Ω, an algebra $\mathcal{A}$, and an element $\omega \in \Omega$, then the map δ_w that assigns to every $B \in \mathcal{A}$ the value 1 if $w \in B$ and 0 otherwise is a probability. This is called the *Dirac probability concentrated at ω*;

- given $\mathbb{N}$ and an algebra $\mathcal{A}$ on it, the *counting measure* is the map that to any element $B \in \mathcal{A}$ assigns its cardinality if it's finite, $+\infty$ otherwise.

The conditionally σ-additive condition is sufficient also to guarantee the finitely-additive one. However, we decided to explicitly point out this feature to understand better the following proposition that we will use later.

Proposition A.1.1. *Let $\mathcal{C}$ be an algebra on Ω and μ a non-trivial and finitely-additive probability $\mathcal{C} \to [0, 1]$. The following are equivalent:*

[2] We are assuming here the classical conventions about the sums that involve $+\infty$.

- *μ is a probability, i.e. μ is also conditionally σ-additive;*
- *for every family $\{A_i\}_{i\in\mathbb{N}}$ of elements in $\mathcal{C}$ such that $A_0 \subseteq A_1 \subseteq A_2 \subseteq \ldots$ and $\bigcup_{i\in\mathbb{N}} A_i$ is an element in $\mathcal{C}$,*

$$\mu(\bigcup_{i\in\mathbb{N}} A_i) = \lim_{i\to+\infty} \mu(A_i);$$

- *μ is σ-subadditive, i.e. for every family $\{A_i\}_{i\in\mathbb{N}}$ of elements in $\mathcal{C}$ such that $\bigcup_{i\in\mathbb{N}} A_i$ is an element of $\mathcal{C}$,*

$$\sum_{i\in\mathbb{N}} \mu(A_i) \geq \mu(\bigcup_{i\in\mathbb{N}} A_i).$$

Sometimes it happens that we have a set Ω with a σ-algebra $\mathcal{F}$ on it and we want to define a measure that assigns to some elements of $\mathcal{F}$ some given values. This is the case, for instance, if we want to establish a measure on the borelian σ-algebra in $\mathbb{R}$. A good intuition to start with is that this measure should assign to any interval (a, b) the value $b - a$. But there exists such a probability? Is there a risk that setting all these values overdetermine the measure? Is there a risk that these choices can be in contradiction if we want the map to satisfy the finitely-additivity or the σ-additivity conditions? The next theorem provides us with an answer to this question when we are in a specific setting.

Theorem A.1.1 (Extension Theorem). *Let $\mathcal{C}$ be an algebra on Ω, $\mathcal{F}(\mathcal{C})$ the σ-algebra generated and μ a probability defined on $\mathcal{C}$. Then μ can be extended uniquely to a probability on $\mathcal{F}(\mathcal{C})$.*

In the example above of the borelian σ-algebra, the set

$$\mathcal{S} := \{\emptyset, \overline{\mathbb{R}}, (-\infty, a], (b, +\infty], (a, b] \text{ with } a, b \in \mathbb{Q}\}$$

is not an algebra because it is not closed under finite union: however, from this we can get the algebra $\mathcal{A}$ generated by $\mathcal{S}$, that is

$$\mathcal{A} := \{\dot\cup_{i=0}^{n} A_i \ : \ A_i \in \mathcal{S} \text{ and } n \in \mathbb{N}\},$$

i.e. the set of all the finite union of disjoint elements in $\mathcal{S}$.

Since we have a clear notion of measure for intervals, we have a map $\mu : \mathcal{S} \to [0, +\infty]$ that assigns $+\infty$ to any unbounded interval and $b - a$ to the ones of the form $(a, b]$ with $a, b \in \mathbb{Q}$. Then, we can extend μ to a function defined on $\mathcal{A}$ by simply giving the value $\sum_{i=0}^{n} \mu(A_i)$ to any disjoint union $A_1 \dot\cup \ldots \dot\cup A_n$ of elements of $\mathcal{S}$. Denoting this new map still by μ, by Theorem A.1.1, we have a unique extension λ of μ to $\mathcal{F}(\mathcal{A})$. Noticing that $\mathcal{F}(\mathcal{A}) = \mathcal{F}(\overline{\tau})$, where $\overline{\tau}$ is the standard Euclidean topology in $\overline{\mathbb{R}}$ [3], we get that λ is the only measure on the borelian σ-algebra of $\overline{\mathbb{R}}$ that assigns to intervals their "length".

[3] By the minimality of the σ-algebra generated by a set, we have only to show that $\mathcal{A} \subseteq \mathcal{F}(\overline{\tau})$ and $\overline{\tau} \subseteq \mathcal{F}(\mathcal{A})$:

- $\mathcal{A} \subseteq \mathcal{F}(\overline{\tau})$: indeed, $\mathcal{F}(\overline{\tau})$ is closed under finite union, therefore by how $\mathcal{A}$ is defined, it is sufficient to prove that $\mathcal{S} \subseteq \mathcal{F}(\overline{\tau})$. For the elements of the form $\emptyset, \overline{\mathbb{R}}, (b, +\infty]$ this is true because they are in $\overline{\tau}$; $(-\infty, a]$ is the complement of an element $((a, +\infty] \cup \{-\infty\})$ in $\overline{\tau}$; $(a, b]$ is the intersection of two elements already shown to be in $\mathcal{F}(\overline{\tau})$;
- $\overline{\tau} \subseteq \mathcal{F}(\mathcal{A})$: indeed, any open set can be written as a countable union of sets of the form $(b, +\infty]$,

Given a set Ω with a σ-algebra $\mathcal{F}$ on it and a measure μ, we can define the notion of integral of a function $f : \Omega \to \mathbb{R}$.[4] This will be not well-defined for all the possible functions but only for the *measurable* ones, i.e. for the f's such that for any $t \in \mathbb{R}$, the preimage through f of $(-\infty, t]$ is in $\mathcal{F}$; this condition is equivalent to requiring that the preimage of any open set is a member of $\mathcal{F}$ or to the condition that for any A in the borelian σ-algebra of $\mathbb{R}$, $f^{-1}(A) \in \mathcal{F}$. This allows us to generalize the concept of *measurability*.

Definition A.1.3. Given two measurable spaces $(\Omega, \mathcal{F}), (\Omega', \mathcal{F}')$, a function $f : \Omega \to \Omega'$ is called *measurable* if for any $A \in \mathcal{F}'$, $f^{-1}(A) \in \mathcal{F}$. In this case, since a measurable function preserves the measurable structure, we will write that $f : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$.

Example A.1.1. The condition of measurability is easy to verify when the function $f : \Omega \to \mathbb{R}$ is *simple*, i.e. it assumes only a finite number of values. For such an f, we have that it is measurable if and only if for any value α assumed by f,

$$f^{-1}(\alpha) = \{\omega \in \Omega \ : \ f(\omega) = \alpha\} \in \mathcal{F}.$$

In the following, we will assume to work with functions $f : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with codomain in $\mathbb{R}$, if not otherwise specified: the notion of measurability is, therefore, the one with regards to the borelian σ-algebra $\mathcal{B}(\mathbb{R})$ of $\mathbb{R}$: in this case, sometimes, we will say that f is $\mathcal{F}$-measurable to emphasize the σ-algebra we are referring to.

For any f, g functions and for any $\alpha \in \mathbb{R}$, then $f + g, \alpha f, f - g, f \cdot g, f/g, f^+, f^-$[5] (where the domain of f/g is the set of the x's with $g(x) \neq 0$) are well-defined: we can show that if f, g are measurable, also all these functions are. In addition, it's easy to check that a continuous function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, where A is a measurable subset of $\mathbb{R}^n$, is also measurable.

Measurable functions are closed under sup, inf, and limit, in the sense that if $\{f_n\}_{n\in\mathbb{N}}$ is a succession of measurable functions, then also the pointwise-defined functions $\sup_n f_n$, $\inf_n f_n, \lim_{n\to+\infty} f_n$ (if it exists) are. This feature allows us to define integration. We will assume that a set Ω with a σ-algebra $\mathcal{F}$ on it and a measure μ are given:

- for a measurable simple function $f : \Omega \to \mathbb{R}$, whose range is $\{\alpha_1, \dots, \alpha_n\}$, we define

$$\int f \ d\mu := \sum_{i=1}^{n} \alpha_i \mu(f^{-1}(\alpha_i));$$ [6]

$[-\infty, a)$ or (a, b) with $a, b \in \mathbb{R}$. Actually, the a, b's above can be taken to be rational: indeed, for instance, $(b, +\infty]$ can be seen as a union of countable intervals $(b_n, +\infty]$ with b_n a decreasing succession such that $b_n > b$ and $b_n \to b$.
Therefore, the thesis follows by noticing that: the sets of the form $(b, +\infty]$ with $b \in \mathbb{Q}$ belong to $\mathcal{F}(\mathcal{A})$ since they are in $\mathcal{S}$; sets of the form (a, b) with $a, b \in \mathbb{Q}$ are in $\mathcal{F}(\mathcal{A})$ because they can be seen as the countable union of the sets $(a, b - (b-a)/n]$ for natural numbers $n \geq 2$; the sets of the form $[-\infty, a)$ are in $\mathcal{F}(\mathcal{A})$ because taking $b < a$, we can write this set as the union of $(b, +\infty]^c$ and (b, a), elements already shown to be in the σ-algebra at issue.

[4]Since in this work we are not dealing with maps that assume in some input an infinite value, we restrict the presentation to functions whose codomain is $\mathbb{R}$; however, the theory needs only slight adjustments to deal with functions $f : \Omega \to \overline{\mathbb{R}}$.

[5]The functions f^+, f^- are respectively the *positive* and *negative part* of f. i.e. the pointwise defined functions

$$f^+(x) := \max\{f(x), 0\} \quad f^-(x) := -\min\{f(x), 0\}.$$

[6]Notice that $\mu(f^{-1}(\alpha_i))$ is well-defined since, as explained in Example A.1.1, for a measurable simple

- for a non-negative measurable function $f : \Omega \to [0, +\infty)$, we define the integral to be
$$\int f \; d\mu := \sup \left\{ \int g \; d\mu : g \text{ simple function with } g \leq f \right\};$$
- for a measurable function $f : \Omega \to \mathbb{R}$, if at least one among $\int f^+ \; d\mu, \int f^- \; d\mu$ is finite, the integral of f exists and
$$\int f \; d\mu := \int f^+ \; d\mu - \int f^- \; d\mu.$$
If both $\int f^+ \; d\mu, \int f^- \; d\mu$ are finite, then also $\int f \; d\mu$ is finite and f is called *integrable*.
- a measurable funcion $f : \Omega \to \mathbb{R}$ is *integrable over* A for $A \in \mathcal{F}$, if $f \cdot \chi_A$ is integrable and in this case,
$$\int_A f \; d\mu := \int f \cdot \chi_A \; d\mu.$$
Hence, the notions of *integrable over* Ω and $\int_\Omega f \; d\mu$ are the same of *integrable* and $\int f \; d\mu$.

The notion of integrability depends on the measurable space $(\Omega, \mathcal{F})$ and on the measure μ at issue: sometimes we will say that a function $f : \Omega \to \mathbb{R}$ is μ-integrable if we want to stress the measure we are referring to and the notation (that conserves this dependence) we use for the integral will be one of the following:
$$\int f \; d\mu \quad \int_\Omega f \; d\mu \quad \int f(\omega) \; d\mu(\omega) \quad \int_\Omega f(\omega) \; d\mu(\omega).$$

It can be checked that the previous items give a well-defined notion of integration that satisfies a lot of useful properties (we will use the same notations as before):

- *Linearity:*
If f, g are integrable and $\alpha \in \mathbb{R}$, then:
  - $\alpha f, f + g$ are integrable and
$$\int \alpha f \; d\mu = \alpha \int f \; d\mu, \quad \int f + g \; d\mu = \int f \; d\mu + \int g \; d\mu;$$
  - if $f \leq g$, then $\int f \; d\mu \leq \int g \; d\mu$.
- *Nullity criterion:*
If $f : \Omega \to \mathbb{R}$ is integrable and $\int_A f \; d\mu \geq 0$ for any $A \in \mathcal{F}$, then $f \geq 0$ μ-a.e., i.e. μ-almost everywhere, i.e. the set $\{\omega \in \Omega \; : \; f(\omega) \ngeq 0\}$ has null μ-measure. Hence, if $\int_A f \; d\mu = 0$ for all $A \in \mathcal{F}$, then $f = 0$ μ-a.e..
- *Monotone Convergence Theorem:*
If $\{f_n\}_{n \in \mathbb{N}}$ is a succession of non-negative measurable functions that is increasing, i.e. for any $\omega \in \Omega$
$$0 \leq f_0(\omega) \leq f_1(\omega) \leq \dots,$$
and f is a function such that $f(\omega) = \lim_{n \to +\infty} f_n(\omega)$ μ-almost everywhere, then $\lim_{n \to +\infty} \int f_n \; d\mu = \int f \; d\mu$.

function f, $f^{-1}(\alpha_i)$ is a member of the σ-algebra, so μ is well-defined on it.

- *Beppo Levi's Theorem:*
  If $\{f_n\}_{n\in\mathbb{N}}$ is a succession of non-negative measurable functions, then
$$\int \sum_{n=0}^{+\infty} f_n \, d\mu = \sum_{n=0}^{+\infty} \int f_n \, d\mu.$$

- *(Lebsegue) Dominated Convergence Theorem:*
  Let $\{f_n\}_{n\in\mathbb{N}}$ be a succession of measurable functions and f a function such that $f(\omega) = \lim_{n\to+\infty} f_n(\omega)$ μ-almost everywhere. If there exists an integrable function g that dominates all the f_n's, i.e. such that for any $n \in \mathbb{N}$ $|f_n| \leq g$ μ-almost everywhere, then f and the f_n's are integrable and $\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$.

- *Pushforward measure (or Image measure):*
  Let X, Y be two sets respectively endowed with the σ-algebras $\mathcal{F}, \mathcal{F}'$, and let $f : X \to Y$ be a measurable function (w.r.t. the above σ-algebras). Given a measure μ on X, the *pushforward measure (or image measure)* induced by f is the measure $f_*\mu$[7] on Y such that for any $A \in \mathcal{F}'$, $f_*\mu(A) = \mu(f^{-1}(A))$.

  If $g : Y \to \mathbb{R}$ is a measurable function, then it is integrable w.r.t. $f_*\mu$ if and only if $g \circ f$ is integrable w.r.t. μ and in this case
$$\int_Y g \, df_*\mu = \int_X (g \circ f) d\mu.$$

A.2 Product measures

Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ two measurable spaces. Then a *rectangle* is a subset of $\Omega \times \Omega'$ of the form $A \times B$ with $A \in \mathcal{F}$ and $B \in \mathcal{F}'$. The set $\Omega \times \Omega'$ can be endowed with the *product σ-algebra* (in symbols $\mathcal{F} \times \mathcal{F}'$), i.e. the one generated by the rectangles.

We can notice that the projection maps
$$\begin{aligned} \pi : &\Omega \times \Omega' \to \Omega \\ &(\omega, \omega') \mapsto \omega \end{aligned} \qquad\qquad \begin{aligned} \pi' : &\Omega \times \Omega' \to \Omega' \\ &(\omega, \omega') \mapsto \omega'. \end{aligned}$$
are measurable (w.r.t. the product σ-algebra) and, hence, given a map $f : \Omega \to \mathbb{R}$ that is measurable, the map $f \circ \pi : (\Omega \times \Omega', \mathcal{F} \times \mathcal{F}') \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable and the same holds with Ω' and π', instead of Ω and π.

With a similar argument, we can show that if $f : (\Omega \times \Omega', \mathcal{F} \times \mathcal{F}') \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a measurable function, then for any $\omega \in \Omega$ $[\omega' \in \Omega']$ the section f_ω $[f^{\omega'}]$ that maps $\omega' \in \Omega'$ to $f(\omega, \omega') \in \mathbb{R}$ [*that maps* $\omega \in \Omega$ *to* $f(\omega, \omega')$] is measurable.

Given a subset E of $\Omega \times \Omega'$, for any $\omega \in \Omega$ and for any $\omega' \in \Omega'$ we put
$$\begin{aligned} E_\omega &:= \{\omega' \in \Omega' \ : \ (\omega, \omega') \in E\} \\ E^{\omega'} &:= \{\omega \in \Omega \ : \ (\omega, \omega') \in E\}. \end{aligned}$$
Using the Extension Theorem A.1.1, we get

[7] sometimes denoted by μf^{-1}.

Theorem A.2.1. *Given $(\Omega, \mathcal{F}, \mu)$ and $(\Omega', \mathcal{F}', \mu')$ two measurable spaces with σ-finite measures on them, there is a unique measure (the* product measure*) $\mu \times \mu'$ on $(\Omega \times \Omega', \mathcal{F} \times \mathcal{F}')$ such that*

$$(\mu \times \mu')(A \times B) = \mu(A) \cdot \mu'(B)$$

for any $A \in \mathcal{F}$ and $B \in \mathcal{F}'$.

Furthermore, for an arbitrary set $E \in \mathcal{F} \times \mathcal{F}'$, the functions $\omega \mapsto \mu'(E_\omega)$ and $\omega' \mapsto \mu(E^{\omega'})$ are measurable for any $\omega \in \Omega$ and $\omega' \in \Omega'$ and

$$(\mu \times \mu')(E) = \int_\Omega \mu'(E_\omega)\ d\mu(\omega) = \int_{\Omega'} \mu(E^{\omega'})\ d\mu'(\omega').$$

We end this part by stating two important theorems that allow us to evaluate integrals with respect to the product measure by evaluating iterated integrals.

Theorem A.2.2. Fubini's and Tonelli's Theorems
Let $(\Omega, \mathcal{F}, \mu)$ and $(\Omega', \mathcal{F}', \mu')$ be two measurable spaces with σ-finite measures on them.

- *if $f : \Omega \times \Omega', \to [0, +\infty)$ is a $\mathcal{F} \times \mathcal{F}'$-measurable function, then:*
  - *the function $\omega \mapsto \int_{\Omega'} f_\omega\ d\mu'$ is $\mathcal{F}$-measurable;*
  - *the function $\omega' \mapsto \int_\Omega f^{\omega'}\ d\mu$ is $\mathcal{F}'$-measurable;*
  - *the following equalities hold*

$$\int_{\Omega\times\Omega'} f\ d(\mu \times \mu') = \int_\Omega \left(\int_{\Omega'} f_\omega\ d\mu' \right) d\mu = \int_{\Omega'} \left(\int_\Omega f^{\omega'}\ d\mu \right) d\mu'. \quad \text{(A.1)}$$

- *if $f : \Omega \times \Omega', \to \mathbb{R}$ is a $\mathcal{F} \times \mathcal{F}'$-measurable function, and one of the three terms in (A.1) with regards to $|f|$ (and not to f) is finite, then:*
  - *f is $\mu \times \mu'$-integrable;*
  - *the section f_ω is μ'-integrable for μ-almost any $\omega \in \Omega$ and the function $\omega \mapsto \int_{\Omega'} f_\omega\ d\mu'$ is μ-integrable ;*
  - *the section $f^{\omega'}$ is μ-integrable for μ'-almost any $\omega' \in \Omega'$ and the function $\omega' \mapsto \int_\Omega f^{\omega'}\ d\mu$ is μ'-integrable;*
  - *the equalities in (A.1) hold for f.*

A.3 Weakly convergence

We restrict here to probabilities on $\mathbb{R}^d$ for a $d \geq 1$: the σ-algebra considered will be the borelian one $\mathcal{B}(\mathbb{R}^d)$.

Definition A.3.1. Given a succession $\{\mu_n\}_{n\in\mathbb{N}}$ of probabilities on $\mathbb{R}^d$, it *weakly converges to a probability* μ if for any function $f : \mathbb{R}^d \to \mathbb{R}$ that is continuous and bounded,

$$\int f\ d\mu = \lim_{n\to+\infty} \int f\ d\mu_n.$$

Example A.3.1. Consider the Dirac probabilities on Example 1.0.1: we can show that if we have a succession $\{a_n\}_{n\in\mathbb{N}}$ of elements in $\mathbb{R}^d$ that converges to a, then the succession $\{\delta_{a_n}\}_{n\in\mathbb{N}}$ converges to δ_a. Indeed, it is easy to show that $\int f\ d\delta_b = f(b)$ and the thesis is equivalent to show that for any bounded and continuous function f, if $a_n \to a$, then $f(a_n) \to f(a)$.

Remark A.3.1. Given a succession that weakly converges, the weak limit is unique: this relies on the fact that if two probabilities μ, μ' on $\mathbb{R}^d$ are such that $\int f\ d\mu = \int f\ d\mu'$ for any continuous and bounded function f, then they are equal. See Lemma 10.3.1 [3].

We end the appendix by showing that a kind of compactness property holds for some successions of probabilities.

Definition A.3.2. A succession $\{\mu_n\}_{n\in\mathbb{N}}$ of probabilities on $\mathbb{R}^d$ is *uniformly tight* if for any $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset \mathbb{R}^d$ such that for any $n \in \mathbb{N}$

$$\mu_n(K_\varepsilon) > 1 - \varepsilon.$$

Theorem A.3.1. *If $\{\mu_n\}_{n\in\mathbb{N}}$ is a uniformly tight succession of probabilities on $\mathbb{R}^d$, then it has a subsuccession that weakly converges to a probability.*

Appendix B

Nonstandard Analysis

In this appendix, we will present to the reader all the necessary material to understand Sections 2.3 and 2.4: the following is heavily based on [9] and it should not be seen as an exhaustive description of Nonstandard Analysis.

In everyday mathematics, we work in a "universe" that contains all the entities we deal with in developing a theory. Usually, not every universe is good: there are some requirements that such an object should satisfy. For instance, for any two sets A, B in the universe we want also to have all the functions from A to B, as well as their domains and ranges, all the r-ary relations in A, or the union $A \cup B$, etc. All these entities can be described in ZF through sets (as a simple Set Theory course shows, see for example [16]): hence, we will assume ZF in the background of this discussion.

Definition B.0.1. A *universe* is a set $\mathbb{U}$ that satisfies the following:

- it is strongly transitive;[1]
- for any $a, b \in \mathbb{U}$, the set $\{a, b\}$ is still in $\mathbb{U}$;
- for any $A, B \in \mathbb{U}$, if A, B are sets, then the union $A \cup B$ is still in $\mathbb{U}$;
- for any $A \in \mathbb{U}$, if A is a set, then the power set $\mathcal{P}(A)$ is still in $\mathbb{U}$.[2]

Such an $\mathbb{U}$ is a *universe over* $\mathbb{X}$ if $\mathbb{X} \in \mathbb{U}$ and the members of $\mathbb{X}$ are regarded as *individuals*, i.e. as objects without members and different from $\emptyset$. In our universe, a set will be an element in $\mathbb{U} \setminus \mathbb{X}$. This is the reason for specifying the fact that e.g. in the union, A, B must be sets and there is no union of elements in $\mathbb{X}$.

We will usually denote with lower-case letters $(a, b, \dots)$ generic elements of $\mathbb{U}$ (that may belong to $\mathbb{X}$ and be not sets but individuals) and with capital ones $(A, B, \dots)$ the elements that are sets.

Remark B.0.1. These properties guarantee that in a universe $\mathbb{U}$ all the constructions we need in mathematics are available. The following remark gives some examples.

[1] This means that for any set $A \in \mathbb{U}$ there exists a set $B \in \mathbb{U}$ such that $A \subseteq B \subseteq \mathbb{U}$ and B is transitive, i.e. for any $x \in B$ and for any $y \in x$, y is still an element of B. Notice that strong transitivity implies transitivity.

[2] By transitivity of $\mathbb{U}$, if $\mathcal{P}(A) \in \mathbb{U}$ for any $A \in \mathbb{U}$, then also $\emptyset$ is an element of $\mathbb{U}$.

- given m sets $A_1, \dots, A_m \in \mathbb{U}$, the union $A_1 \cup \dots \cup A_m$ is still an element of $\mathbb{U}$, since $\mathbb{U}$ is closed under binary union;
- if $B \in \mathbb{U}$ is a set, any of its subset $A \subseteq B$ is also an element of $\mathbb{U}$: indeed, $A \in \mathcal{P}(B) \in \mathbb{U}$ and by transitivity, $A \in \mathbb{U}$;
- for any family of sets $\{A_i : i \in I\}$ such that all the $A_i's$ are elements of a set $A \in \mathbb{U}$, the union $\cup_{i\in I} A_i$ is in $\mathbb{U}$: indeed, by strong transitivity, there exists a transitive set $B \in \mathbb{U}$ such that $A \subseteq B \subseteq \mathbb{U}$. If $a \in \cup_{i\in I} A_i$, there exixts $i_0 \in I$ such that $a \in A_{i_0}$; since $A_{i_0} \in A$, A_{i_0} is an element of B and by transitivity, $a \in B$. Therefore $\cup_{i\in I} A_i \subseteq B$ and for the previous item, $\cup_{i\in I} A_i \in \mathbb{U}$;
- for any $a \in A \in \mathbb{U}$ and $b \in B \in \mathbb{U}$, the pair (a, b)[3] belongs to $\mathbb{U}$: this is because the given a and b are also elements of $\mathbb{U}$; therefore the sets $\{a, b\}$ and $\{a\}$ are in $\mathbb{U}$ and so is the ordered pair $(a, b) = \{\{a\}, \{a, b\}\}$;
- for any A, B sets in $\mathbb{U}$, the product set $A \times B$ is in $\mathbb{U}$: indeed, $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \cup B))$ and since the last set belongs to $\mathbb{U}$, also $A \times B$ does, too.
- the same argument used above justifies also that for any sets $A, B \in \mathbb{U}$, any relation $R \subseteq A \times B$ and then any function $f : A \to B$ is in $\mathbb{U}$.

Example B.0.1. Given a set $\mathbb{X}$ for which the formula "$\forall x \in \mathbb{X}\ \forall y \in \mathbb{X}\ x \notin y$" is valid, there is a simple way to construct a universe $\mathbb{U}$ on it in an inductive way by

$$\mathbb{U}_0(\mathbb{X}) := \mathbb{X},$$
$$\mathbb{U}_{n+1}(\mathbb{X}) := \mathbb{U}_n(\mathbb{X}) \cup \mathcal{P}(\mathbb{U}_n(\mathbb{X})),$$
$$\mathbb{U}(\mathbb{X}) := \bigcup_{n=0}^{+\infty} \mathbb{U}_n(\mathbb{X}).$$

A universe $\mathbb{U}$ comes with an associated language $\mathcal{L}_\mathbb{U}$ that has a countable supply of variables and:

- the $\mathcal{L}_\mathbb{U}$-terms are defined inductively by:
  - every variable is a $\mathcal{L}_\mathbb{U}$-term;
  - every member of $\mathbb{U}$ is a term (a constant) in $\mathcal{L}_\mathbb{U}$;
  - if $t_1, \dots, t_m$ are $\mathcal{L}_\mathbb{U}$-terms and $m \geq 2$, the m-uple $(t_1, \dots, t_m)$ is a $\mathcal{L}_\mathbb{U}$-term;
  - if t, s are $\mathcal{L}_\mathbb{U}$-terms and x a variable, then $t\{x/s\}$ is a $\mathcal{L}_\mathbb{U}$-term.
- the atomic $\mathcal{L}_\mathbb{U}$-formulas are of the form

  $$t = s \quad \text{or} \quad t \in s,$$

  for any t, s $\mathcal{L}_\mathbb{U}$-terms;
- the $\mathcal{L}_\mathbb{U}$-formulas are the ones built starting from the atomic ones through boolean operations and "restricted" quantifiers, i.e. if φ is a $\mathcal{L}_\mathbb{U}$-formula, t a $\mathcal{L}_\mathbb{U}$-term, and x a variable that doesn't occur in t, then $\forall x \in t\ \varphi$ and $\exists x \in t\ \varphi$ are $\mathcal{L}_\mathbb{U}$-formulas.

[3]We use here the standard definition of pairs used in ZF (Kuratwoski's one): the singleton $\{a\}$ represents the set $\{a, a\}$ and the ordered pair (a, b) the set $\{\{a\}, \{a, b\}\}$.

Then, we can give the usual definition of truth of a formula in a universe. Notice that if $a \in \mathbb{X}$ then $t \in a$ is always false; furthermore, for instance, we have that two sets are equal if and only if they have the same elements as members since we are assuming to work in ZF, where the axiom of Extensionality holds.

In the nonstandard setting, we want to *enlarge* the universe $\mathbb{U}$ to a universe $\mathbb{U}'$ that contains also nonstandard entities.

Definition B.0.2. A *nonstandard framework* for a set $\mathbb{X}$ is the given of a universe $\mathbb{U}$ over $\mathbb{X}$, a universe $\mathbb{U}'$ that contains $\mathbb{X}$ and a map $* : \mathbb{U} \to \mathbb{U}'$ such that:

a) ${}^*a = a$ for all $a \in \mathbb{X}$;

b) ${}^*\emptyset = \emptyset$;

c) an $\mathcal{L}_{\mathbb{U}}$-sentence φ is true in $\mathbb{U}$ if and only if its $*$-transform ${}^*\varphi$ is true in $\mathbb{U}'$.

In order to give sense to the previous definition, we should explain what is the "$*$-transform"of a $\mathcal{L}_{\mathbb{U}}$-formula φ. Given the map $*$, we associate an element ${}^*a \in \mathbb{U}'$ to any $a \in \mathbb{U}$; hence to any $\mathcal{L}_{\mathbb{U}}$-term t it corresponds its $*$-transform ${}^*t$, that is obtained by replacing in t each constant symbol a by ${}^*a$. Substituing all the terms of φ with their $*$-transform (see below for examples of $*$-transform), we get its $*$-transform ${}^*\varphi$.

The identity map defines a valid nonstandard framework but, obviously, not an interesting one. We will focus, indeed, on the case in which $\mathbb{U}' \neq \mathbb{U}$ and we have also *nonstandard elements* in $\mathbb{U}'$, as explained below. If we want to do analysis in our universe, a good choice for $\mathbb{X}$ is a set such that $\mathbb{R} \subseteq \mathbb{X}$; the usual operations $+, -, \cdot, \leq$, etc., intended as functions or relations between sets, are in any universe $\mathbb{U}$ over $\mathbb{R}$ and their $*$-transform inherit a lot of their original features. We will drop the $*$ in the following when we talk about ${}^*+$, ${}^*-$, ${}^*\cdot$, ${}^*\leq$ and, from the context, it would be clear if we are talking about a standard operation or its $*$-counterpart. In Nonstandard Analysis, a classical assumption is that ${}^*\mathbb{N} \setminus \mathbb{N} \neq \emptyset$, i.e. there is at least a nonstandard natural number.[4] From this hypothesis, a lot of interesting features of nonstandard frameworks will follow. First, however, it is better to explain with some examples item c) in Definition B.0.2: this is a really powerful tool at our disposal to carry truth from one universe to the other and it is called *Transfer Principle*. In the following, suppose $a, b, \dots, A, B, \dots$ are elements in $\mathbb{U}$:

a) the map $*$ is injective, since by Transfer Principle $a = b$ if and only if ${}^*a = {}^*b$;

b) $a \in B$ if and only if ${}^*a \in {}^*B$;

c) $A \subseteq B$ if and only if ${}^*A \subseteq {}^*B$; indeed, A is a subset of B in $\mathbb{U}$ if and only if the statement $\varphi := \forall a \in A\ a \in B$ holds. By Transfer Principle, this is equivalent to asking for the validity in $\mathbb{U}'$ of ${}^*\varphi$, i.e. of the formula "$\forall a \in {}^*A\ a \in {}^*B$", hence the thesis.

d) ${}^*(A \cup B) = {}^*A \cup {}^*B$, since in $\mathbb{U}$ it holds that

$$\forall x \in A \cup B\ (x \in A \vee x \in B) \wedge \forall x \in A\ (x \in A \cup B) \wedge \forall x \in B\ (x \in A \cup B).$$

[4]In a nonstandard framework, $\mathbb{U}$ and $\mathbb{U}'$ are universes over $\mathbb{X}$. Hence, since $\mathbb{N} \subseteq \mathbb{R}$, the notation "${}^*\mathbb{N} \setminus \mathbb{N}$" is meaningful: ${}^*\mathbb{N}$ is the image of $\mathbb{N} \in \mathbb{U}$ via $*$, and $\mathbb{N}$ is the set of natural numbers seen in $\mathbb{U}'$.

By taking the $*$-transform, we get

$$\forall x \in {}^*(A \cup B)\ (x \in {}^*A \vee x \in {}^*B) \wedge \forall x \in {}^*A\ (x \in {}^*(A \cup B)) \wedge \forall x \in {}^*B\ (x \in {}^*(A \cup B))$$

and the thesis follows. Notice that, since $A \cup B$ is a constant in $\mathcal{L}_{\mathbb{U}}$, its $*$-transform is ${}^*(A \cup B)$ and not ${}^*A \cup {}^*B$: indeed, if it would be this the case, then we would have nothing to prove! By the Transfer Principle, however, we can "carry inside" the $*$;

e) ${}^*\{a_1, \dots, a_m\} = \{{}^*a_1, \dots, {}^*a_m\}$ by applying Transfer Principle to the formula

$$\forall x \in \{a_1, \dots, a_m\} \left(\bigvee_{i=1}^{m} x = a_i \wedge \bigwedge_{i=1}^{m} a_i \in \{a_1, \dots, a_m\} \right);$$

As we shall see, if A is not a finite set, there may be elements $b \in {}^*A$ such that $b \neq {}^*a$ for all $a \in A$.

f) ${}^*(A_1 \times \dots \times A_m) = {}^*A_1 \times \dots \times {}^*A_m$; let's prove it for $m = 2$. In $\mathbb{U}$ the following formula is valid:

$$\forall x \in A_1\ \forall y \in A_2\ (x, y) \in A_1 \times A_2 \wedge \forall z \in A_1 \times A_2\ \exists x \in A_1\ \exists y \in A_2\ z = (x, y).$$

If we take its $*$-transform, recalling that (x, y) is a $\mathcal{L}_{\mathbb{U}}$-term, we have

$$\forall x \in {}^*A_1\ \forall y \in {}^*A_2\ (x, y) \in {}^*(A_1 \times A_2) \wedge \forall z \in {}^*(A_1 \times A_2)\ \exists x \in {}^*A_1\ \exists y \in {}^*A_2\ z = (x, y),$$

therefore, ${}^*(A_1 \times A_2) = {}^*A_1 \times {}^*A_2$;

g) if in $\mathbb{U}$ we have a relation $R \subseteq A \times B$, then also ${}^*R$ is a relation in ${}^*A \times {}^*B$, by Transfer Principle applied to the formula

$$\forall x \in R\ x \in A \times B,$$

and the previous point. If we have a function $f : A \to B$, then ${}^*f$ is a function in $\mathbb{U}'$ from ${}^*A$ to ${}^*B$. This can be easily seen by applying the Transfer Principle to

$$\forall x \in f\ (x \in A \times B)\ \wedge \forall a \in A\ \exists b \in B\ ((a, b) \in f \wedge \forall b' \in B\ (a, b') \in f \to b = b').$$

Always by Transfer Principle, for a function $f : A \to B$ and an element $a \in A$, we have that ${}^*(f(a)) = {}^*f({}^*a)$: indeed, we can simply consider the formula

$$\forall y \in B\ (y = f(a) \equiv (a, y) \in f).$$

A similar argument shows also that f is injective (surjective or bijective) if and only if ${}^*f$ is;

h) using the previous item, ${}^*+$ is a function from ${}^*\mathbb{R} \times {}^*\mathbb{R}$ to ${}^*\mathbb{R}$ and it has a lot of properties in common with the usual $+$. For instance, it is commutative by the Transfer Principle applied to

$$\forall x \in \mathbb{R}\ \forall y \in \mathbb{R}\ x + y = y + x;$$

i) ${}^*\mathcal{P}(A) \subseteq \mathcal{P}({}^*A)$ for a generic set $A \in \mathbb{U}$. Indeed, in $\mathbb{U}$ the following formula is valid

$$\forall x \in \mathcal{P}(A)\ \forall y \in x\ y \in A;$$

hence, by Transfer Principle, any x in ${}^*\mathcal{P}(A)$ is a subset of ${}^*A$.

The reader may be surprised by the fact that the last item inclusion is not an equality in general, i.e. $\mathcal{P}({}^*A) \subseteq {}^*\mathcal{P}(A)$ doesn't hold for all $A \in \mathbb{U}$.[5] To see a concrete example of a set A for which $\mathcal{P}({}^*A) \neq {}^*\mathcal{P}(A)$, we will consider the usual nonstandard framework we have in mind when we do Nonstandard Analysis, i.e. a universe $\mathbb{U}$ over a set $\mathbb{X} \supseteq \mathbb{R}$ with a map $*$ for which ${}^*\mathbb{N} \setminus \mathbb{N} \neq \emptyset$, that is, a nonstandard framework where nonstandard natural numbers exist. The proof of the existence of a nonstandard framework for $\mathbb{R}$ of this form is not trivial; however, for now, we want to focus the attention on the fact that only assuming the hypothesis of the existence of a nonstandard natural number, we get a lot of interesting results, as already anticipated. The more pragmatic reader will find this speculation more interesting when at the end of the appendix we will discuss the existence of such a nonstandard framework. In the following, we will call the members in $\mathbb{U}'$ that are of the form ${}^*a$ for an $a \in \mathbb{U}$ *standard*, and the other *nonstandard*.

Notice that, if $A \subseteq \mathbb{X}$ is a set, then all the elements in ${}^*A \setminus A$ are nonstandard. Indeed, a standard element in ${}^*A$ actually belongs to A: this is because if $a \in \mathbb{X}$ then $a = {}^*a$ and if this belongs to ${}^*A$, by Transfer (applied to ${}^*a \in {}^*A$) we have also that ${}^*a = a \in A$. In particular, every element in ${}^*\mathbb{N} \setminus \mathbb{N}$ is a nonstandard element.

Moreover, if ${}^*\mathbb{N} \setminus \mathbb{N} \neq \emptyset$, we can prove the existence of nonstandard *sets*. Using the Transfer Principle, it's easy to prove that $\leq$ is a total order in ${}^*\mathbb{N}$ and that any element in ${}^*\mathbb{N} \setminus \mathbb{N}$ must be greater than any number in $\mathbb{N}$. In $\mathbb{U}$, for any $n \in \mathbb{N}$ there exists a set $V_n = \{1, \dots, n\}$, consisting of all the positive natural numbers less or equal than n, i.e. the following formula is true in $\mathbb{U}$:

$$\forall n \in \mathbb{N}\ \exists V_n \in \mathcal{P}(\mathbb{N})\ \forall x \in \mathbb{N}\ (x \in V_n \equiv x \neq 0 \wedge x \leq n)$$

Then, by Transfer, for any $N \in {}^*\mathbb{N}$ there exists a set $V_N = \{1, \dots, N\}$ in ${}^*\mathcal{P}(\mathbb{N})$ in which the members are all and only the "naturals" (included the nonstandard ones) that are not bigger than N.

If $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ the set V_N is nonstandard. Otherwise, there exists $A \in \mathbb{U}$ with $V_N = {}^*A$ and, by Transfer Principle, since in $\mathbb{U}'$ it is valid that ${}^*n \in {}^*A$ for any $n \in \mathbb{N}$, we get $\mathbb{N} \subseteq A$. Furthermore, A must be a subset of $\mathbb{N}$, by applying Transfer to the formula $\forall x \in {}^*A\ x \in {}^*\mathbb{N}$. From these considerations, we must conclude that $A = \mathbb{N}$ and, hence, that $V_N = {}^*\mathbb{N}$; however, $N + 1 \in {}^*\mathbb{N} \setminus V_N$, a contradiction.

We now show that under the hypothesis that ${}^*\mathbb{N} \setminus \mathbb{N}$ is non-empty, the equality $\mathcal{P}({}^*\mathbb{N}) = {}^*\mathcal{P}(\mathbb{N})$ does not hold.

In $\mathbb{U}$ the least number principle (LNP) is valid, hence

$$\forall x \in \mathcal{P}(\mathbb{N})\ (x \neq \emptyset \rightarrow (\exists y \in x\ \forall z \in x\ y \leq z)).$$

If we assume that $\mathcal{P}({}^*\mathbb{N}) = {}^*\mathcal{P}(\mathbb{N})$, then LNP should be valid also for the set ${}^*\mathbb{N} \setminus \mathbb{N}$ which we supposed to be a non-empty subset of ${}^*\mathbb{N}$. If we call N this minimum, then, since $N > 0$, we have $N - 1 \in {}^*\mathbb{N} \setminus \mathbb{N}$ that is less than the minimum, yielding a contradiction.

The argument above shows that in general, for $A \in \mathbb{U}$, the two sets $\mathcal{P}({}^*A)$ and ${}^*\mathcal{P}(A)$ are different. Notice that if $A = \{a_1, \dots, a_m\}$ is a finite set, then so is $\mathcal{P}(A)$ and the

[5] Naïvely, we can think to apply the Transfer Principle to the formula

$$\forall x (\forall y \in x (y \in A) \rightarrow x \in \mathcal{P}(A)):$$

however, this is not an $\mathcal{L}_{\mathbb{U}}$-formula due to the "unrestricted" initial quantifier "$\forall x$".

equality at issue holds since we have already shown that for any finite set $B \in \mathbb{U}$,

$$^*B = \{^*b \ : \ b \in B\}.$$

We can characterize $^*\mathcal{P}(A)$ thanks to the following notion.

Definition B.0.3. In a nonstandard framework as in Definition B.0.2, an element $a \in \mathbb{U}'$ is *internal* if and only if $a \in {}^*A$ for some $A \in \mathbb{U}$. In other words, the internal elements are the elements of standard sets of $\mathbb{U}'$.

Every element in $\mathbb{U}'$ that is not internal is *external*.

The set of all the internal elements of $\mathbb{U}'$ is denoted by $^*\mathbb{U}$. All the standard entities are internal (because $^*a \in \{^*a\} = {}^*\{a\}$) and, in addition, it is possible to prove that:

- $^*\mathbb{U}$ is strongly transitive and, in particular, any member of an internal set is internal;
- if $A, B \in {}^*\mathbb{U}$, then $A \cup B, A \cap B, A \setminus B$, and $A \times B$ are still internal;
- if $\{A_i \ : \ i \in I\} \in {}^*\mathbb{U}$ and each A_i is a set, then $\cup_{i \in I} A_i$ and $\cap_{i \in I} A_i$ are internal.
- if $a_1, \ldots, a_m$ are internal, also $\{a_1, \ldots, a_m\}$ is internal;
- if a binary relation R is internal, then its domain, its range, its inverse, and the image through it of any internal subset C of its domain are internal;

As previously announced, the following result explains the difference between $\mathcal{P}(^*A)$ and $^*\mathcal{P}(A)$.

Theorem B.0.1. *Given a set $A \in \mathbb{U}$, $^*\mathcal{P}(A)$ is the set of all internal subsets of $^*A$.*

Theorem B.0.1 leads to some important consequences on the $*$-transform a $\mathcal{L}_{\mathbb{U}}$-sentence φ. For instance:

- if we have a $\mathcal{L}_{\mathbb{U}}$-sentence φ of the form $\forall x \in \mathcal{P}(A)\ \psi$, then its $*$-transform is of the form

  for any *internal* subset x of $^*A$ $\ ^*\psi$ holds;

  As an example, the LNP is valid for any non-empty internal subsets of $^*\mathbb{N}$: as a consequence of this remark, $\mathbb{N}$ is not an internal set: otherwise, by the closure properties of internal sets, $^*\mathbb{N} \setminus \mathbb{N}$ would be internal too, although it is non-empty and does not have the least element.

- if we have a $\mathcal{L}_{\mathbb{U}}$-sentence φ of the form $\exists f : A \to B\ \psi$, $[\forall f : A \to B\ \psi]$ then, since this abbreviation means that $f \in \mathcal{P}(A \times B)$ is a function, its $*$-transform is of the form

  there exists an *internal* function f from $^*A$ to $^*B$ such that $^*\psi$ holds
  [for all the *internal* functions f from $^*A$ to $^*B$ $^*\psi$ holds].

Remark B.0.2. In Chapter 2 we deal with sums in a nonstandard framework and their $*$-transform: this remark is devoted to enlighten better the question.

In $\mathbb{U}$, given a function $f : \mathbb{N} \to \mathbb{R}$, we can consider the sum $\sum_{n\leq m} f(n)$. This is characterized by the following formula:

$$\varphi(f, x, y) := (y = 0 \wedge x = f(0)) \vee (y > 0 \wedge \exists z \in \mathbb{R}\ \varphi(f, z, y-1) \wedge\ x = z + f(y)),$$

in the sense that in $\mathbb{U}$ for $b \in \mathbb{N}$ and $a \in \mathbb{R}$, it is valid $\varphi(f, a, b)$ if and only if $\sum_{n\leq b} f(n) = a$. Since in $\mathbb{U}$ the following holds

$$\forall f : \mathbb{N} \to \mathbb{R}\ \forall m \in \mathbb{N}\ \exists! x \in \mathbb{R}\ \varphi(f, x, m),$$

also for any internal functions $f : {}^*\mathbb{N} \to {}^*\mathbb{R}$ and for any $N \in {}^*\mathbb{N}$, it is meaningful to consider $\sum_{n\leq N} f(n)$: this will denote the unique $x \in {}^*\mathbb{R}$ that satisfies ${}^*\varphi(f, x, N)$.

Notice that, working in $\mathbb{U}$, given a function $f : \mathbb{N} \to \mathbb{R}$, it is possible that in a formula ψ appears the subformula $\sum_{n\leq m} f(n) = t$, for a $\mathcal{L}_{\mathbb{U}}$-term t. The $*$-transform of this piece of formula is different depending on the fact that we consider $\sum_{n\leq m} f(n)$ as a real number (hence, an individual of $\mathbb{X}$) or as the only real number x that satisfies $\varphi(f, x, m)$: for this, as a standard rule, every time we write formulas like $\sum_{n\leq m} f(n) = t$, we mean

$$\exists x \in \mathbb{R}\ \varphi(f, x, m) \wedge x = t;$$

the same applies for the formulas of the form $\sum_{n\leq m} f(n) \in t$.

Furthermore, we can give sense to expressions like $\sum_{n\in {}^*\mathbb{N}} a_n$, with the a_n's positive hyperreals. Indeed, in $\mathbb{U}$ the following formula is valid:

$$\forall f : \mathbb{N} \to \mathbb{R}\ ((\forall n \in \mathbb{N}\ f(n) \geq 0) \to \big(\forall M \in \mathbb{N}\ \exists m \in \mathbb{N}\ \sum_{n\leq m} f(n) > M) \vee (\exists s \in \mathbb{R}\ \psi(f, s))),$$

where $\psi(f, s)$ is the formula

$$\forall \varepsilon \in \mathbb{R}\ (\varepsilon > 0 \to \exists m \in \mathbb{N}\ \forall n \in \mathbb{N}(n \geq m \to (s - \sum_{k\leq n} f(k) < \varepsilon)))$$

As in the previous item, this should convince the reader that we can talk about $\sum_{n\in {}^*\mathbb{N}} f(n)$ for any internal functions $f : {}^*\mathbb{N} \to {}^*\mathbb{R}$. In particular, if f is such that

$$\forall M \in {}^*\mathbb{R}\ \exists m \in {}^*\mathbb{N}\ \sum_{n\leq m} f(n) > M$$

the sum will be considered infinite (in analogy as what is done in the standard setting); otherwise, the sum is a (positive) hyperreal $s \in {}^*\mathbb{R}$ and we can write $\sum_{n\in {}^*\mathbb{N}} f(n) = s$. Thus, if in a formula ϕ, appears a subformula of the form

$$\sum_{n\in\mathbb{N}} g(n) = t,$$

where $g : \mathbb{N} \to \mathbb{R}$ is a function in $\mathbb{U}$, and t an $\mathcal{L}_{\mathbb{U}}$-term, we will mean that the sum is not infinite and this should be read as

$$\exists x \in \mathbb{R}\ \psi(g, x) \wedge x = t;$$

the $*$-trasnform of this subformula that appears in ${}^*\phi$ would be

$$\exists x \in {}^*\mathbb{R}\ {}^*\psi({}^*g, x) \wedge x = {}^*t,$$

i.e. with a more comfortamble notation,

$$\sum_{n \in {}^*\mathbb{N}} {}^*g(n) = {}^*t.$$

The following result provides us with a nice way to detect if a set or a function is internal or not and this tool is heavily used in Sections 2.3, and 2.4. In the following, a $\mathcal{L}_{\mathbb{U}'}$-formula [term] φ [t] is called internal if all its constants are internal.

Theorem B.0.2. *[Definition Principle]*
Given an internal formula $\varphi(x)$ with x as the only free variable, then for any internal set A the set

$$\{a \in A\ :\ \varphi(a)\}$$

is internal; given a function $f : A \to B$ between two internal sets A, B, if there is an internal term $t(x)$, with x as the only variable, such that $t(a) = f(a)$ for any $a \in A$, then f is internal.

As an example, using the previous theorem, we can prove that

$$\{1, \dots, N\} = \{x \in {}^*\mathbb{N}\ :\ x \leq N \wedge x \neq 0\}$$

is internal for all $N \in {}^*\mathbb{N}$: indeed, it belongs to the standard set ${}^*\mathcal{P}(\mathbb{N})$.
We can say something more: if we denote the set of all finite subsets of A by $\mathcal{P}_{\text{fin}}(A)$, then, since in $\mathbb{U}$ the formula

$$\forall n \in \mathbb{N}\ \exists V \in \mathcal{P}_{\text{fin}}(\mathbb{N})\ \forall x \in \mathbb{N}\ x \in V \equiv (x \neq 0 \wedge x \leq n)$$

is valid, in $\mathbb{U}'$, by Transfer Principle,

$$\forall n \in {}^*\mathbb{N}\ \exists V \in {}^*\mathcal{P}_{\text{fin}}(\mathbb{N})\ \forall x \in {}^*\mathbb{N}\ x \in V \equiv (x \neq 0 \wedge x \leq n)$$

holds. This means that, if $N \in {}^*\mathbb{N}$, the set $\{1, \dots, N\}$ belongs to ${}^*\mathcal{P}_{\text{fin}}(\mathbb{N})$ even if it is not a finite set (as already explained, all the standard natural numbers are less than N if N is nonstandard, i.e. if $N \in {}^*\mathbb{N} \setminus \mathbb{N}$).

The members of ${}^*\mathcal{P}_{\text{fin}}(A)$, even if not finite, have a lot of features in common with finite sets and they are called *hyperfinite* subsets of ${}^*A$; a set is called *hyperfinite* if it is an hyperfinite subsets of ${}^*A$ for some $A \in \mathbb{U}$. Notice that a hyperfinite set is internal since it belongs to a standard set (${}^*\mathcal{P}_{\text{fin}}(A)$) and that a finite set (also the empty one) is hyperfinite. We now list some examples of finiteness properties that hyperfinite sets satisfy:

- for any non-empty set $B \in \mathbb{U}'$, B is hyperfinite if and only if there is an internal bijection

  $$f : \{1, \dots, N\} \to B,$$

  for some $N \in {}^*\mathbb{N}$: given B, this N is unique and it's called the *(internal) cardinality* of B. Denoting the cardinality of an hyperfinite set C by $|C|$, we have that the map $C \mapsto |C|$ is an internal function ${}^*\mathcal{P}_{\text{fin}}(A) \to {}^*\mathbb{N}$;

- any internal subset B of a hyperfinite set A[6] is hyperfinite and $|B| \leq |A|$;
- for any hyperfinite set A, the internal power set $\mathcal{P}(A) \cap {}^*\mathbb{U}$, i.e. the set of all the internal subsets of A, is hyperfinite;
- In $\mathbb{U}$, we know that if $f : A \to \mathbb{R}$, then the map $B \mapsto \sum_{b \in B} f(b)$ from $\mathcal{P}_{\text{fin}}(A)$ to $\mathbb{R}$ is well defined. By Transfer Principle, given an internal function $f : {}^*A \to {}^*\mathbb{R}$ we obtain a map ${}^*\mathcal{P}_{\text{fin}}(A) \to {}^*\mathbb{R}$, so that a summation where the indexes belong to a hyperfinite set makes sense. Furthermore, these sums gain most of the properties of the finite sums in the standard setting for instance:
  - for $c \in {}^*\mathbb{R}$, $c \cdot \left(\sum_{b \in B} f(b)\right) = \sum_{b \in B} c \cdot f(b)$;
  - given $g : A \to \mathbb{R}$ internal like f,

$$\sum_{b \in B}(f(b) + g(b)) = \sum_{b \in B} f(b) + \sum_{b \in B} g(b).$$

The rest of the appendix is devolved to show that there exists a nonstandard framework in which we have nonstandard entities like the ones in ${}^*\mathbb{N} \setminus \mathbb{N}$.

Definition B.0.4. Assume to have a nonstandard framework $* : \mathbb{U} \to \mathbb{U}'$ where $\mathbb{U}$ and $\mathbb{U}'$ are universes over some set $\mathbb{X}$ such that $\mathbb{R} \subseteq \mathbb{X}$. $\mathbb{U}'$ is called an *enlargement* of $\mathbb{U}$ if for any collection of sets $\{A_i : i \in I\} \in \mathbb{U}$ with the Finite Intersection Property (FIP)[7], there is a $b \in \mathbb{U}'$ that belongs to ${}^*A_i$ for any $i \in I$.

Having an enlargement guarantees the existence of nonstandard elements. Indeed, a family of sets with FIP in $\mathbb{U}$ "gains" elements in the intersection, even if there is no element in $\mathbb{U}$ that belongs to all the members of the family. Let's show some examples:

a) considering the family $\{\mathbb{N} \setminus \{0, \dots, n\} : n \in \mathbb{N}\}$, an element N in the intersection

$$\cap_{n \in \mathbb{N}} {}^*(\mathbb{N} \setminus \{0, \dots, n\}) = \cap_{n \in \mathbb{N}}({}^*\mathbb{N} \setminus {}^*\{0, \dots, n\}) = \cap_{n \in \mathbb{N}}({}^*\mathbb{N} \setminus \{0, \dots, n\}) = {}^*\mathbb{N} \setminus \mathbb{N}$$

is a nonstandard element of ${}^*\mathbb{N}$;

b) considering the family $\{(0, r) : r \in \mathbb{R}_{>0}\}$, an element in the intersection $\cap_{r \in {}^*\mathbb{R}_{>0}} {}^*(0, r)$ would be an $\varepsilon \in {}^*\mathbb{R}$ that is greater than 0 but less than any positive real r: these elements are called *infinitesimal*;

c) similarly to the above example, we get *unlimited* "reals" if we consider the family $\{(r, +\infty) : r \in \mathbb{R}_{>0}\}$.

Item b) can be regarded as the core of Nonstandard Analysis, created at first to deal with a well-defined notion of infinitesimals in calculus. Thanks to this notion we can study the structure of ${}^*\mathbb{R}$ in comparison to $\mathbb{R}$. We will call *hyperreals* the elements of ${}^*\mathbb{R}$, in order to differentiate them from the reals, i.e. the members of $\mathbb{R}$.

A hyperreal b is *infinitesimal* if for any positive real number $r \in \mathbb{R}$, $b < r$; b is *infinitely close* to a hyperreal c (in symbols $b \approx c$) if $b - c$ is infinitesimal;[8] b is *bounded* (or *limited*)

[6] Notice that a hyperfinite set may have non-internal subsets, e.g. $\mathbb{N} \subseteq \{1, \dots, N\}$ for $N \in {}^*\mathbb{N} \setminus \mathbb{N}$.

[7] This means that for any finite $F \subseteq I$ the set $\cap_{i \in F} A_i$ is non-empty.

[8] Notice that a hyperreal b is infinitesimal if and only if it is infinitesimal close to 0. The "infinitesimally closeness" $\approx$ is an equivalence relation.

if there exists $n \in \mathbb{N}$ such that $|b| \leq n$, that is, $-n \leq b \leq n$. In any enlargement we can show that

Theorem B.0.3. *Every bounded hyperreal b is infinitely close to exactly one real number c: this is called the* standard part *of b and it's denoted by $^\circ b$.*

Proof. Let $A = \{r \in \mathbb{R} \ : \ r < b\}$. Since b is limited, there are $r, s \in \mathbb{R}$ such that $r < b < s$, hence A is a non-empty set of reals which is bounded from above. The sup (in $\mathbb{R}$) of this set will be called c and we now prove that c is infinitely close to b.

Take any positive $\varepsilon \in \mathbb{R}$. By definition of sup:

- $c + \varepsilon > c$, hence it is not a member of A and $b \leq c + \varepsilon$;
- $c - \varepsilon < c$, so there exists $r \in A$ such that $r > c - \varepsilon$, and, therefore, $c - \varepsilon < b$.

Thus, we have that $|b - c| \leq \varepsilon$ for any positive $\varepsilon \in \mathbb{R}$, so $b \approx c$.

This c is the unique real number infinitely close to b because if there is also $c' \approx b$, then, since $\approx$ can be shown to be an equivalence relation, $c \approx c'$. Two distinct reals can't be infinitely close, so we must conlude that $c = c'$. □

This theorem allows us to consider the function $^\circ : b \mapsto {}^\circ b$ that maps any *bounded* hyperreal to its standard part. It is easy to show that if $\mathbb{L} \subset {}^*\mathbb{R}$ is the ring of limited numbers [9], then $^\circ : \mathbb{L} \to \mathbb{R}$ is an order-preserving homomorphism, whose kernel is the set $\mathbb{I}$ of infinitesimals. Since the map is surjective (since for any real $r \in \mathbb{R}$, $^\circ({}^*r) = r$), we have, for the First Homomorphism Theorem, that $\mathbb{R}$ is isomorphic as an ordered field to $\mathbb{L}/\mathbb{I}$.

Now, we will come back to the question about the existence of an enlargement: the following theorem provides a positive answer.

Theorem B.0.4. *For any set $\mathbb{X}$ there is an enlargement $^* : \mathbb{U}(\mathbb{X}) \to \mathbb{U}'$ of $\mathbb{U}(\mathbb{X})$ where $\mathbb{U}' = \mathbb{U}({}^*\mathbb{X})$ [10].*

Proof. We will only sketch the proof, leaving the details to the reader (see Chapter 14, [9]).

The classical approach to build a nonstandard framework uses a construction very similar to Łos ultrapower one. Given $\mathbb{U}(\mathbb{X})$, a set I (not necessarily in $\mathbb{U}(\mathbb{X})$), a non-principal ultrafilter $\mathcal{F}$ on it, and denoting for any $a \in \mathbb{U}(\mathbb{X})$ by a_I the function $i \in I \mapsto a$, for any $f, g \in \mathbb{U}(\mathbb{X})^I$ and $a \in \mathbb{U}(\mathbb{X})$, we define:

$$\llbracket f = g \rrbracket := \{i \in I \ : \ f(i) = g(i)\}$$
$$\llbracket f \in g \rrbracket := \{i \in I \ : \ f(i) \in g(i)\}$$
$$\llbracket f \in a \rrbracket := \llbracket f \in a_I \rrbracket$$
$$\llbracket a \in f \rrbracket := \llbracket a_I \in f \rrbracket$$
$$Z_n := \{f \in \mathbb{U}(\mathbb{X})^I \ : \ \llbracket f \in \mathbb{U}_n(\mathbb{X}) \rrbracket \in \mathcal{F}\}$$
$$Z = \cup_{n \in \mathbb{N}} Z_n.$$

[9] with $^*+, {}^*\cdot, {}^*-, 0, 1$.

[10] where $\mathbb{U}(\mathbb{X})$ and $\mathbb{U}({}^*\mathbb{X})$ are defined inductively as explained in Example B.0.1.

Hence, Z is the set of the “bounded rank” functions. Omitting some details that can be found in the reference, by identifying functions that are $\mathcal{F}$-almost equal (i.e. f, g are identified if $[\![f = g]\!] \in \mathcal{F}$) we can get a map $Z \to \mathbb{U}(\mathbb{Y})$, where

$$\mathbb{Y} = Z_0/([\![f = g]\!] \in \mathcal{F}),$$

i.e. $\mathbb{Y}$ is Z_0 modulo the identification induced by the ultrafilter $\mathcal{F}$. More precisely, given $f \in Z_0$, we have the class $[f] \in \mathbb{Y}$ of all the functions in Z_0 that are $\mathcal{F}$-equal to f, i.e.

$$[f] := \{g \in Z_0 \ : \ [\![f = g]\!] \in \mathcal{F}\}.$$

The map $f \mapsto [f]$ from Z_0 to $\mathbb{Y}$ can be extended as a map from Z to $\mathbb{U}(\mathbb{Y})$, reasoning by induction on the rank of f. Indeed, assuming we have defined $[f]$ for all the functions $f \in Z_m$ with $m \leq n$, given $g \in Z_{n+1}$, we can extend it, by putting

$$[g] := \{[h] \ : \ h \in Z_n \text{ and } [\![h \in g]\!] \in \mathcal{F}\}.$$

It's easy to prove that this map is well-defined and, in particular, by induction we can show that for any $n \in \mathbb{N}$, if $f \in Z_n$, then $[f] \in \mathbb{U}_n(\mathbb{Y})$. Then, composing this map with the one from $\mathbb{U}(\mathbb{X})$ to Z such that $a \mapsto a_I$, we have a nonstandard framework $* : \mathbb{U}(\mathbb{X}) \to \mathbb{U}(\mathbb{Y})$ with the properties required in Definition B.0.2. Moreover, it can be shown that with this construction, we can identify ${}^*\mathbb{X}$ with $\mathbb{Y}$, because, if $a \in \mathbb{X}$ then a_I is a function in Z_0, therefore $[a_I] \in \mathbb{Y}$.

To get a nonstandard framework with nonstandard elements, we need to ask the ultrafilter $\mathcal{F}$ to be non-principal.[11] However, to have an enlargement this is still not sufficient, and we have to demand something more: an available choice is to take $I = \mathcal{P}_{\text{fin}}(\mathbb{U}(\mathbb{X}))$ and for any $a \in I$ define

$$I_a := \{b \in I \ : \ a \subseteq b\}.$$

The collection $\{I_a \ : \ a \in I\}$ has the FIP, so there is an ultrafilter $\mathcal{F}$ that contains this family (see Theorem 2.6.1, [9]). It can be shown that not only $\mathcal{F}$ is not principal, but also it guarantees that the function $* : \mathbb{U}(\mathbb{X}) \to \mathbb{U}({}^*\mathbb{X})$ that comes out from the argument discussed above, is an enlargement. □

Considering in the above theorem and in the relative proof sketch $\mathbb{X} = \mathbb{R}$, we get an enlargement where we can do Nonstandard Analysis. In particular, ${}^*\mathbb{R}$ is the set

$$Z_0 = \{f \in \mathbb{U}(\mathbb{R})^I \ : \ \{i \in I \ : \ f(i) \in \mathbb{R}\} \in \mathcal{F}\}$$

modulo the identification

$$f \sim_{\mathcal{F}} g \text{ if and only if } \{i \in I \ : \ f(i) = g(i)\} \in \mathcal{F}.$$

[11] Indeed, if, for instance, $\mathcal{F}$ is generated by $j \in I$, i.e.

$$\mathcal{F} = \{A \subseteq I \ : \ j \in A\},$$

then a function $f \in \mathbb{U}(\mathbb{X})^I$ is in Z_0 if and only if $\{i \in I \ : \ f(i) \in \mathbb{X}\} \in \mathcal{F}$ and this means that $f \in Z_0$ if and only if $f(j) \in \mathbb{X}$. Furthermore, for any $f, g \in Z_0$, $[\![f = g]\!] \in \mathcal{F}$ if and only if $f(j) = g(j)$, from which we deduce that $\mathbb{X}$, ${}^*\mathbb{X}$, and $\mathbb{Y}$ can be identified.

For any function $f \in Z_0$, we can take the function $g \in \mathbb{R}^I$ defined by

$$g(i) = \begin{cases} f(i) & \text{if } f(i) \in \mathbb{R}; \\ 0 & \text{otherwise} \end{cases}$$

and notice that $f \sim_{\mathcal{F}} g$, since

$$\{i \in I \ : \ f(i) = g(i)\} = \{i \in I \ : \ f(i) \in \mathbb{R}\} \in \mathcal{F}.$$

Hence, ${}^*\mathbb{R}$ can be seen as $\mathbb{R}^I$ modulo the identification $\sim_{\mathcal{F}}$.

Considering an enlargement ${}^* : \mathbb{U}(\mathbb{X}) \to \mathbb{U}({}^*\mathbb{X})$, with $\mathbb{R} \subseteq \mathbb{X}$, we have also other useful properties we will exploit in Chapter 2:

- *Overflow Principle:* if $\varphi(x)$ is an internal $\mathcal{L}_{\mathbb{U}'}$-formula, with x as the only free variable, such that $\varphi(n)$ is true for all $n \in \mathbb{N}$, then there exists $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ for which $\varphi(n)$ holds also for all $n \in {}^*\mathbb{N}$ with $n \leq N$.

  This is a simple application of Transfer Principle and it holds also simply for nonstandard frameworks in which $\mathbb{N} \neq {}^*\mathbb{N}$. Indeed, consider the set

  $$A = \{n \in {}^*\mathbb{N} \ : \ \neg\varphi(n)\}.$$

  This is internal by Theorem B.0.2: if it is empty, for any $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ we get the thesis; otherwise, we can take the least number $K \in A$ (that is not in $\mathbb{N}$ by the hypothesis that $\varphi(n)$ holds for any $n \in \mathbb{N}$) and $N = K - 1$ satisfies the required condition.

- *Comprehensiveness:* for each set $A \in \mathbb{U}(\mathbb{X})$ and for each internal set $B \in \mathbb{U}({}^*\mathbb{X})$, any function $f : A \to B$ extends to an internal function $\tilde{f} : {}^*A \to B$, in the sense that $\tilde{f}({}^*a) = f(a)$ for any $a \in A$. We will talk of *sequential comprehensiveness* whenever $A = \mathbb{N}$ (for a proof, see Section 15.4 [9]).

Another property we will need is the κ-saturation: if κ is a cardinal, then this property says that any collection of fewer than κ internal sets that satisfies the FIP has non-empty intersection.

In Section 2.3 and Section 2.4 we use only the *countable saturation*, i.e. only in the case in which $\kappa = \aleph_1$: this is equivalent (Theorem 15.4.2, [9]) to require the sequential comprehensiveness and $\mathbb{N} \neq {}^*\mathbb{N}$, so it doesn't depend on the choice of the ultrafilter (that is assumed, however, to generate an enlargement with nonstandard elements in ${}^*\mathbb{N}$).

However, if we require κ-saturation when $\kappa > \aleph_1$, even if there exists some I and some ultrafilter on it for which the construction used in Theorem B.0.4 proof gives us an enlargement with this property, not all the enlargements satisfy it: there are, indeed, some I and some ultrafilter $\mathcal{F}$ for which the generated enlargement is not a κ-saturated one.
A simple way to notice this dependence is to consider the internal family $\{{}^*\mathbb{R} \setminus \{r\}\}_{r \in {}^*\mathbb{R}}$ in an enlargement ${}^* : \mathbb{U}(\mathbb{X}) \to \mathbb{U}({}^*\mathbb{X})$, with $\mathbb{R} \subseteq \mathbb{X}$: this has the FIP, but the intersection of the whole family is empty. Hence, for instance, this enlargement will not satisfy κ-saturation whenever $\kappa > |{}^*\mathbb{R}|$.

Bibliography

[1] Carnap R., *Logical Foundations of Probability*, University of Chicago Press, Routledge, Kegan Paul, London, 1950.

[2] Christiano P., *Non-Omniscience, Probabilistic Inference, and Metamathematic*, MIRI, 2014.

[3] Cohn D. L., *Measure Theory*, Birkäuser Advanced Test, Second edition, Springer, 2013.

[4] Cutland N.J., *Loeb Measure Theory*, Developments in Nonstandard Mathematics (with Oliveira F., Neves V., and Sousa-Pinto J.), Pitman Research Notes in Mathematics Series, Vol. 336, Longman CRC Press, pp. 151-177.

[5] Demski A., *Logical Prior Probability*, Artificial General Intelligence, Springer, 2012.

[6] Fagin R., *Probabilities on Finite Models*. The Journal of Symbolic Logic, Vol. 41, No. 1, pp. 50-58, 1976.

[7] Gaifman H. and Snir M., *Probabilities Over Rich Languages*. The Journal of Symbolic Logic, Vol. 47, No. 3, pp. 495-548, 1982.

[8] Garrabrant S., Benson-Tilsen T., Critch A., Soares N., and Taylor J., *Logical Induction*, Machine Intelligence Research Institution, 2016.

[9] Goldblatt R., *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis.* Graduate Texts in Mathematics, Springer, 1998.

[10] Halpern J., *Reasoning about Uncertainty*, MIT Press, 2003.

[11] Howarth E., Paris J., *Pure Inductive Logic with Functions*, The Journal of Symbolic Logic, 2019.

[12] Hutter M., Lloyd J.W., Ng K.S. and Uther W.T.B., *Probabilities on Sentences in an Expressive Logic*. Presented at the Fifth Workshop on Combining Probability and Logic, 2012.

[13] Johnson W. E., *Probability: The deductive and inductive problems*, Mind, Vol. 41, 1932.

[14] Jonas H., *Das Prinzip Verantwortung - The imperative of Responsibility: In Search of an Ethics for the Technological Age*, English version published by Einaudi, 1979.

[15] Krauss P. H., *Representation of Symmetric Probability Models*. The Journal of Symbolic Logic, Vol. 34, No. 2, pp. 183-193, 1969.

[16] Kunen K., *Set Theory: An Introduction to Independence Proofs*, Elsevier Science Publisheres B. V., 1980.

[17] Monk D.J., Bonnet R., *Handbook of Boolean Algebras*, Vol. 1, Elsevier, 1989.

[18] Moses Y., Halpern J., Fagin R., Vardi M., *Reasoning about Knowledge*, MIT Press, 1989.

[19] Paris J., Vencovská A., *On the Strongest Principles of Rational Belief Assignments*, Personal communication, draft under publication, 2024.

[20] Paris J., Vencovská A., *Pure Inductive Logic*. Perspectives in Logic, Cambridge University Press, 2015.

[21] Paris J., Vencovská A., *Six Problems in Pure Inductive Logic*, Journal of Philosophical Logic, Vol. 48, No. 4, pp. 731-747, 2019.

[22] Paris J., Vencovská A., *Symmetry's End?*, Erkenntnis, Vol. 74, No. 1, pp. 53-67, 2011.

[23] Paris, J.B., Vencovská, A., *Symmetry Principles in Pure Inductive Logic*, in *Logica 2022*, College Publications, U.K., Ed, Igor Sedl´ar, 2022.

[24] Ramsey F. P., *Truth and Probability* (1926). R. B. Braithwaite (Ed.), The Foundations of Mathematics and other Logical Essays, Routledge and Kegan Paul, London, 1931, pp- 156-198.

[25] Soares N., Fallenstein B., *Agent Foundations for Aligning Machine Intelligence with Human Interests: A Technical Research Agenda*, MIRI, 2016.