Paradoxes with Letters and Definitions

Timothée Bernard[0000−0003−4172−6986]

Laboratoire de linguistique formelle, Université Paris Cité, France
timothee.bernard@u-paris.fr

Abstract. Our thesis is that a certain class of paradoxes about truth and falsity originates in the confusion between two sentences with the same content. According to us, the Liar, for instance, is neither true nor false. In virtue of what the Liar states, it then seem to be false, but this appearance, however, should not be taken at face value. We defend the claim that it is in fact another sentence with the same content that is false. This proposal is formalised with a variant of propositional logic; its semantics relies on definitions, i.e. the association of a formula to a propositional letter. A sentence is then modelled as a propositional letter, with the sentence's content being modelled by the letter's associated formula.

Keywords: Liar · paradoxes · non-classical logic · truth.

1 Introduction

This text is about paradoxical sentences such as the Liar (1-a), the Strengthened Liar (1-b) and the Truth-teller (1-c):

(1) a. Sentence (1-a) is false.
    b. Sentence (1-b) is not true.
    c. Sentence (1-c) is true.

As argued by [6], what is unusual with this kind of sentences is not their meaning, but their truth status. The case of the Strengthened Liar (1-b) [5] is particularly pathological: assuming any truth status (truth, falsity, or some other status) for it seems incoherent in virtue of what the sentence states.

We here defend the claim that the three sentences in (1) are neither true nor false, making them — in virtue of what they state — *appear* to be respectively false (for the Liar), true (for the Strengthened Liar) and false (for the Truth-teller). Our thesis is that this situation is not incoherent (contra *dialetheism*; [8]) because they only appear to be false, true and false. To explain this, we draw a distinction between a sentence and its content, allowing two distinct sentences to share the same content and to differ in truth status. The Liar, for instance, is neither true nor false and, in consequence, appears to be false, because while the Liar itself is neither true nor false, a distinct sentence with the same content (i.e. *Sentence (1-a) is false*) is false. The paradoxes mentioned here arise, in short, when one confuses two sentences with the same content.

Our account is formally expressed using a variant of propositional logic in which a propositional letter can be associated with a formula, its *definiens*. A *definition* is the association of a *definiendum* (to be defined; plural: "definienda") and of a *definiens* (plural: "definientia"); in our case, definienda are propositional letters and definientia are formulas. For example, the Liar (1-a), the Strengthened Liar (1-b) and the Truth-teller (1-c) can be modelled as the letters p_L, p_{SL}, and p_T respectively, with the following definitions:

(2) a. $p_L \sim F(p_L)$
  b. $p_{SL} \sim \neg T(p_{SL})$
  c. $p_T \sim T(p_T)$

where "$\sim$" separates a definiendum and its definiens, "F" stands for *is false*, and "T" for *is true*. As illustrated with these definitions, definientia are formulas of a propositional language that only differ from standard propositional logic in that propositional letters by themselves are not well-formed formulas and can only be used through *projection operators* such as F and T.

The system relies on projection operators in order to circumscribe the presence of non-classical truth values. A definiens (e.g. $\neg T(p_{SL})$) always has a classical truth value (i.e. a member of $\mathbb{B} = \{0, 1\}$), but the truth value of a propositional letter (e.g. p_{SL}) is a set of classical truth values (i.e. a member of $\mathbb{P}(\mathbb{B})$). Accordingly, projection operators represent functions from $\mathbb{P}(\mathbb{B})$ to $\mathbb{B}$. In particular, T sends $\{1\}$ to 1 and the three other subsets of $\mathbb{B}$ to 0, and F sends $\{0\}$ to 1 and the three other subsets of $\mathbb{B}$ to 0. This allows the usual logical operators ($\neg$, $\wedge$) to be interpreted as in classical logic. So if, for instance, p_{SL} is interpreted as $\{\}$, then $T(p_{SL})$ is interpreted as 0 and $\neg T(p_{SL})$ as 1.

In the system, the truth value of a propositional letter (in $\mathbb{P}(\mathbb{B})$) is, strictly speaking, never true (1) nor false (0), but when we describe a propositional letter as true (resp. false), we mean that its truth value is $\{1\}$ (resp. $\{0\}$). Also, we say that the truth of a propositional letter p *coincides* with the truth of its definiens ϕ when $[\![p]\!] = \{[\![\phi]\!]\}$ (where $[\![]\!]$ is the valuation function). Usual, unproblematic sentences correspond to propositional letters the truth of which coincides with that of their definiens and are thus described as either true or false. Other sentences, like the Liar, Strengthened Liar and Truth-teller, however, are modelled as propositional letters the truth of which does *not* coincide with that of their definiens; they are neither true nor false.

The rest of the article is structured as follows. Section 2 formalises all syntactic aspects of the work, including the propositional language, definitions and the dependencies they generate. Section 3 formalises the semantics of the system. Section 4 then illustrates the resulting system focusing on the specific cases of the Liar, the Strengthened Liar and the Truth-teller. Section 5 discusses related work — in particular, contextualist approaches — and finally, Section 6 mentions possible future research directions — in particular, the extension of the system developed here to other paradoxes, such as one that we introduce as the *Matching Probability Paradox*.

2 Syntax and definitions

2.1 Overview

When translating an English sentence such as *It is raining and not snowing* to propositional logic, we (i) usually introduce two propositional letters, say p and q, (ii) informally specify their meaning as $p =$ [it is raining] and $q =$ [it is snowing], and then (iii) use them to write the translation of the initial sentence: $p \wedge \neg q$. A model of propositional logic assigns a truth value to each of the two letters (and, more generally, to all propositional letters), and from them the truth value of $p \wedge \neg q$ is defined.

In our system too can propositional letters stand for atomic propositions. But in order to model propositions referring to other propositions, and in particular self-referring propositions, we can also formally specify the meaning of a propositional letter in terms of other propositional letters. This is done by specifying a *definition*, of which the definiendum is the propositional letter being defined and the definiens is some formula in a propositional language introduced below. For example, the Liar (*The Liar is false*) can be modelled as a propositional letter p_L defined as $F(p_L)$, where F stands for *is false*; this is noted "$p_L \sim F(p_L)$". Similarly, the Strengthened Liar (*The Strengthened Liar is not true*) can be modelled as a letter p_{SL} with the definition "$p_{SL} \sim \neg T(p_{SL})$", where T stands for *is true*.

Given a *definitional context*, i.e. a set of definitions, a model assigns a truth value to all and only the propositional letters without a definition. The truth values of all other propositional letters can then be defined from the model and the definitional context. This process, described in Section 3, is performed iteratively based on the *dependencies* between propositional letters: a propositional letter depends on another iff the definiens of the former contains either the latter or a propositional letter that depends on it. Intuitively, a propositional letter can only be interpreted when all of the letters it depends on have an interpretation. The notion of dependency, while syntactic, thus plays a crucial role in the semantics as the *dependency graph* between propositional letters determines the possible *evaluation orders* of the propositional letters.

2.2 Logical language

Let $\mathcal{P}$ be a countable set of elements called "propositional letters". The logical language $\mathcal{L}$ is then defined inductively as follows:

1. for any $p \in \mathcal{P}$, both $T(p),\ F(p) \in \mathcal{L}$,
2. for any $\phi \in \mathcal{L}$, $\neg\phi \in \mathcal{L}$,
3. for any $\phi, \psi \in \mathcal{L}$, $\phi \wedge \psi \in \mathcal{L}$.

The elements of $\mathcal{L}$ are called "formulas".

Given any subset $P \subseteq \mathcal{P}$, $\mathcal{L}_P \subseteq \mathcal{L}$ is the subset of formulas that only contain propositional letters in P.

We use the following usual abbreviations:

– "$\phi \vee \psi$" stands for "$\neg(\neg\phi \wedge \neg\psi)$"
– "$\phi \rightarrow \psi$" stands for "$\neg(\phi \wedge \neg\psi)$"

2.3 Definitions and definitional contexts

A *definition* is defined as any pair $(p, \phi) \in \mathcal{P} \times \mathcal{L}$; p is the *definiendum* and ϕ its *definiens*. A definition (p, ϕ) can be written "$p \sim \phi$".

A *definitional context* is any set of definitions $C \subseteq \mathcal{P} \times \mathcal{L}$ such that any propositional letter is defined at most once: for any $p \in \mathcal{P}$ and $\phi, \psi \in \mathcal{L}$, if both $p \sim \phi,\ p \sim \psi \in C$, then $\phi = \psi$.

Given a definitional context C, the *domain of* C, written "$\mathcal{D}(C)$", is the set of propositional letters with a definition in C: $\mathcal{D}(C) = \{p \in \mathcal{P} \mid \exists \phi \in \mathcal{L},\ p \sim \phi \in C\}$. Furthermore, if $p \sim \phi \in C$, ϕ can be written "$\mathit{def}_C(p)$". The letters without a definition in C are said to be *primitive*.

Remark that if, according to the above definition of $\mathcal{L}$, the operators T and F can only be applied to propositional letters and not to formulas themselves, this will not lead to any semantic limitation as these operators can be indirectly applied to a formula ϕ by applying them to a letter p such that $p \sim \phi$.

2.4 Dependencies and dependency graph

Given a definitional context C and two propositional letters p and q such that $q \in \mathcal{D}(C)$, one says that there exists a *direct dependency* from p to q if and only if p appears in $\mathit{def}_C(q)$. In that case, one also says that q *depends on* p.

The *dependency graph* $G = (V, E)$ of a definitional context C is the oriented graph such that:

– $V = \mathcal{D}(C)$,
– and $E = \{(p, q) \in V^2 \mid$ there exists a direct dependency from p to $q\}$.

Consider for example the definitional context $C = \{p_1 \sim (\neg T(p_1) \wedge T(q)), p_2 \sim (T(q) \wedge T(r)), p_3 \sim (T(p_1) \vee T(p_2)), p_4 \sim T(p_5),\ p_5 \sim T(p_4)\}$. The dependency graph of C is illustrated in Figure 1; a dependency from a letter p to a letter q is represented as an arrow from the node p to the node q. Note that propositional letters without a definition (i.e. not in $\mathcal{D}(C)$), such as p and r in this example, are not part of the dependency graph.

One says that there exists a *dependency* from p to q if and only if (by induction)

1. there exists a direct dependency from p to q
2. or there exists some r such that there exist a direct dependency from p to r and a dependency from r to q.

(A very similar notion can be found in [4].) As a result, for $p, q \in \mathcal{D}(C)$, there exists a dependency from p to q if and only if there exists an oriented path of length one or more from p to q in the dependency graph.

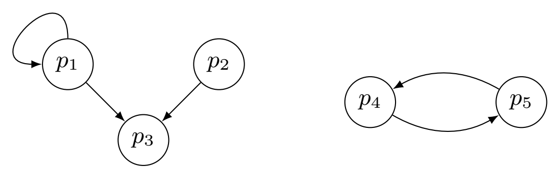

Fig. 1. The dependency graph of the definitional context $C = \{p_1 \sim (\neg T(p_1) \wedge T(q)), p_2 \sim (T(q) \wedge T(r)), p_3 \sim (T(p_1) \vee T(p_2)), p_4 \sim T(p_5),\ p_5 \sim T(p_4)\}$. The arrow from p_1 to p_3, for instance, represents the existence of a direct dependency from p_1 to p_3; the intuition (formalised later) is that p_1 is involved in the interpretation of p_3.

2.5 Evaluation order

Let $G = (V, E)$ be the dependency graph of a given a definitional context C. Let $G_c = (V_c, E_c)$ be the condensation of G, i.e. the graph obtained by contracting each of its strongly connected components:

- V_c is the set of strongly connected components of G (i.e. the nodes of G_c are sets of propositional letters that are all dependent on each other; they are in fact the *maximal* subsets $S \subseteq V$ such that for any two distinct $p, q \in S$, p and q depend on each other),
- $E_c = \{(S, S') \in V_c^2 \mid S \neq S' \text{ and } \exists p \in S, \exists p' \in S', (p, p') \in E\}$ (i.e. there is an edge from S to $S' \neq S$ in G_c iff there is an edge from some $p \in S$ to some $p' \in S'$ in G).

For example, the condensation of the dependency graph illustrated in Figure 1 is illustrated in Figure 2.

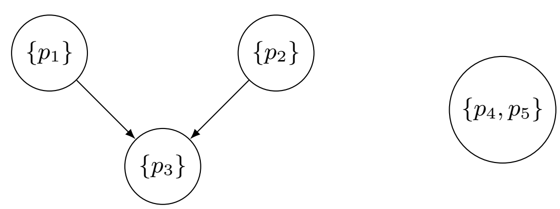

Fig. 2. The condensation of the dependency graph of the definitional context $C = \{p_1 \sim (\neg T(p_1) \wedge T(q)), p_2 \sim (T(q) \wedge T(r)), p_3 \sim (T(p_1) \vee T(p_2)), p_4 \sim T(p_5),\ p_5 \sim T(p_4)\}$.

Contrary to G, its condensation G_c is guaranteed to be a directed *acyclic* graph. Note that in G_c, as well as in G, a node always has a finite number of (direct) parents. However, it is possible that both graphs contain infinite upward chains; this is the case, for example, with $C = \{p_i \sim T(p_{i+1}) \mid i \in \mathbb{N}\}$.

In the remainder of this paper, however, we restrict our attention to definitional contexts without such infinite dependencies. Under these assumptions, there necessarily exists a topological order of G_c.

A topological order of G_c is a total order of its nodes such that any node S must precede any node S' if $(S, S') \in E_c$.[1] Noting $I = \{i \in \mathbb{N} \mid 1 \leq i \leq |V_c|\}$, a topological order of G_c is equivalent to an enumeration $(C_i)_{i \in I}$ of the nodes in V_c such that for any $i, j \in I$, if there exists a directed path from C_i to C_j in G_c, then $i < j$. Now, an *evaluation order* of C is exactly a topological order of G_c.

For instances, $(\{p_1\}, \{p_2\}, \{p_3\}, \{p_4, p_5\})$ and $(\{p_2\}, \{p_4, p_5\}, \{p_1\}\{p_3\})$ are two of the eight different evaluation orders corresponding to the condensation in Figure 2.

Evaluation orders have the following property, crucial to the semantics that we define in the next section: for all $i, j \in I$, $p \in C_i$ and $q \in C_j$, if $i < j$ then there might be a dependency from p to q but there cannot be one from q to p (and, less crucially, if $i = j$ then p and q are dependent on each other). Intuitively, to interpret $p \in C_i$, one does not need the interpretation of the $q \in C_j$ with $i < j$.

3 Semantics

3.1 Overview

The semantics relies on four truth values; the four subsets of $\mathbb{B} = \{0, 1\}$. Let C be a definitional context. A model is an interpretation function f that assigns one of the four truth values to each primitive propositional letter (i.e. without a definition in C). As already mentioned, the truth value of a propositional letter that has a definition in C will be determined based on this definition and thus does not need to be specified by the interpretation function f.

Let $(C_i)_{i \in I}$ be any evaluation order of C ($I = \{i \in \mathbb{N} \mid 1 \leq i \leq |V_c|\}$ where $|V_c|$ is the possibly infinite number of strongly connected components of the dependency graph of C). The initial interpretation function $f_0 = f$ is iteratively extended to all propositional letters based on the evaluation order $(C_i)_{i \in I}$. For each $i \in I$, f_i extends f_{i-1} to C_i according to a process that we call "local semantics". The global, iterative, process by which f is successively extended is called "global semantics". While each f_i sends propositional letters to subsets of $\mathbb{B}$, the extension of f_{i-1} into f_i is defined by quantifying over some classically-valued interpretation functions of C_i (i.e. functions from C_i to $\mathbb{B}$) that are *coherent* in the sense that the value that they assign to any $p \in C_i$ must be identical to the interpretation they lead to for its definiens $def_C(p)$. Then, $f_i(p)$ is defined as the set of all possible values assigned to p by such coherent functions.

[1] In other words, G_c encodes a partial order of its nodes, and a topological order of G_c is any total order extending this partial order.

To be able to conveniently refer, for all $i \in \{0\} \cup I$, to the domain of f_i, let us define $C_0 = \mathcal{P} \setminus \mathcal{D}(C)$ and $C_{0:i} = \bigcup_{j=0}^{i} C_j$. Then f_i is a function from $C_{0:i}$ to $\mathbb{P}(\mathbb{B})$.

While the global semantics is defined based on an evaluation order, the valuation obtained at the end of the process does not depend on the choice of evaluation order.

3.2 Truth values and models

A *truth value* is any subset of $\mathbb{B} = \{0, 1\}$.

Given the definitional context C, a *model* for $\mathcal{L}$ is a function $f : \mathcal{P} \setminus \mathcal{D}(C) \to \mathbb{P}(\mathbb{B})$. In other words, a model is a function that assigns a truth value to any primitive propositional letter.

3.3 Local semantics

Let $i \in I$, the *step*, and f_{i-1} be a function from $C_{0:i-1}$ to $\mathbb{P}(\mathbb{B})$. A *local interpretation (function)* is any function g from C_i to $\mathbb{B}$. Given a local interpretation g, the *g-valuation* is the function from $\mathcal{L}_{C_{0:i}}$ to $\mathbb{P}(\mathbb{B})$ defined inductively by:

- for any $p \in C_{0:i-1}$,
  - $[\![T(p)]\!]^g_{f_{i-1}} = 1$ if $f_{i-1}(p) = \{1\}$, and 0 otherwise;
  - $[\![F(p)]\!]^g_{f_{i-1}} = 1$ if $f_{i-1}(p) = \{0\}$, and 0 otherwise;
- for any $q \in C_i$,
  - $[\![T(q)]\!]^g_{f_{i-1}} = g(q)$;
  - $[\![F(q)]\!]^g_{f_{i-1}} = 1 - g(q)$;
- for any $\phi \in \mathcal{L}_{C_{0:i}}$, $[\![\neg\phi]\!]^g_{f_{i-1}} = 1$ if $[\![\phi]\!]^g_{f_{i-1}} = 0$, and 0 otherwise;
- for any $\phi, \psi \in \mathcal{L}_{C_{0:i}}$, $[\![\phi \wedge \psi]\!]^g_{f_{i-1}} = 1$ if $[\![\phi]\!]^g_{f_{i-1}} = [\![\psi]\!]^g_{f_{i-1}} = 1$, and 0 otherwise.

A local interpretation g is *coherent* iff for all $q \in C_i$, $g(q) = [\![def_C(q)]\!]^g_{f_{i-1}}$. This notion of coherence plays a crucial role in the global semantics defined in the next section.

As an illustration, let us use again $C = \{p_1 \sim (\neg T(p_1) \wedge T(q)), p_2 \sim (T(q) \wedge T(r)), p_3 \sim (T(p_1) \vee T(p_2)), p_4 \sim T(p_5), p_5 \sim T(p_4)\}$ and the evaluation order $(C_i)_{i \in I} = (\{p_1\}, \{p_2\}, \{p_3\}, \{p_4, p_5\})$. Let f_0 be any function from C_0 to $\mathbb{P}(\mathbb{B})$, i.e. a function that assigns a truth value to all propositional letters except $p_1, p_2, \ldots, p_5$. At step $i = 1$, there are only two local interpretations g (i.e. functions from $C_1 = \{p_1\}$ to $\mathbb{B}$): one that sends p_1 to 0 and one that sends p_1 to 1. If $f_0(q) = \{1\}$, then none of them is coherent as $[\![\neg T(p_1) \wedge T(q)]\!]^g_{f_0} = [\![\neg T(p_1)]\!]^g_{f_0} = 1 - [\![T(p_1)]\!]^g_{f_0} = 1 - g(p_1) \neq g(p_1)$. If, however, $f_0(q) \neq \{1\}$, $[\![\neg T(p_1) \wedge T(q)]\!]^g_{f_0} = 0$ and then only the local interpretation that sends p_1 to 0 is coherent.

As another illustration, consider the same definitional context and the same evaluation order. Let f_3 be a function from $C_{0:3}$ to $\mathbb{P}(\mathbb{B})$, i.e. a function that assigns a truth value to all propositional letters except p_4 and p_5. At step $i = 4$, there are four local interpretations g (i.e. functions from $C_5 = \{p_4, p_5\}$ to $\mathbb{B}$). Each of them is coherent iff $g(p_4) = [\![T(p_5)]\!]^g_{f_3}$ and $g(p_5) = [\![T(p_4)]\!]^g_{f_3}$, i.e. iff $g(p_4) = g(p_5)$. In other words, exactly two of them are coherent: the local interpretation that sends both p_4 and p_5 to 1 and the local interpretation that sends them both to 0.

3.4 Global semantics

Given a model $f : \mathcal{P} \setminus \mathcal{D}(C) \to \mathbb{P}(\mathbb{B})$, we now define by induction a sequence of functions $(f_i)_{i\in\{0\}\cup I}$ such that for each *step* $i \in \{0\} \cup I$, f_i is a function from $C_{0:i}$ to $\mathbb{P}(\mathbb{B})$:

- for $i = 0$, $f_0 = f$;
- for $i \in I$, f_i is defined from f_{i-1} as follows:
  - for $p \in C_{0:i-1}$, $f_i(p) = f_{i-1}(p)$;
  - for $q \in C_i$, $f_i(q) = \{g(q) \mid g \text{ is a coherent local interpretation}\}$.

In the end, the *valuation* associated with model f is the function that computes a value for all formulas of the logical language $\mathcal{L}$ based on $f_{|V_c|}$ — the limit function of the sequence $(f_i)_{i\in\{0\}\cup I}$ — and defined inductively as follows:

- for any $p \in \mathcal{P}$,
  - $[\![T(p)]\!] = 1$ if $f_{|V_c|}(p) = \{1\}$, and 0 otherwise;
  - $[\![F(p)]\!] = 1$ if $f_{|V_c|}(p) = \{0\}$, and 0 otherwise;
- for any $\phi \in \mathcal{L}$, $[\![\neg\phi]\!] = 1$ if $[\![\phi]\!] = 0$, and 0 otherwise;
- for any $\phi, \psi \in \mathcal{L}_C$, $[\![\phi \wedge \psi]\!] = 1$ if $[\![\phi]\!] = [\![\psi]\!] = 1$ and 0 otherwise.

For any $p \in \mathcal{P}$, the *(truth) value* of p is defined as $[\![p]\!] = f_{|V_c|}(p)$. When $[\![p]\!] = \{1\}$, p is *classically true*, when $[\![p]\!] = \{0\}$, p is *classically false*.

Let us illustrate these definitions with still the same definitional context $C = \{p_1 \sim (\neg T(p_1) \wedge T(q)), p_2 \sim (T(q) \wedge T(r)), p_3 \sim (T(p_1) \vee T(p_2)), p_4 \sim T(p_5),\ p_5 \sim T(p_4)\}$ and the evaluation order $(C_i)_{i\in I} = (\{p_1\}, \{p_2\}, \{p_3\}, \{p_4, p_5\})$. Consider for instance a model f such that $f(q) = f(r) = \{1\}$.

- The induction starts with $f_0 = f$.
- At step $i = 1$, as we have seen in an example following the presentation of the local semantics, because $f_0(q) = \{1\}$, there is no coherent local interpretation. As a consequence, f_0 is extended as f_1 with $f_1(p_1) = \emptyset$.
- At step $i = 2$, because $f_1(q) = f_1(r) = \{1\}$, it can be shown that there is exactly one coherent local interpretation: the function (defined on $\{p_2\}$) that sends p_2 to 1. As a consequence, f_1 is extended as f_2 with $f_2(p_2) = \{1\}$.
- At step $i = 3$, because $f_2(p_1) = \emptyset$ and $f_2(p_2) = \{1\}$, it can be shown that there is exactly one coherent local interpretation: the function (defined on $\{p_3\}$) that sends p_3 to 1. As a consequence, f_2 is extended as f_3 with $f_3(p_3) = \{1\}$.

- At step $i = 4$, as we have seen in an example following the presentation of the local semantics, there are exactly two coherent local interpretations: the function (defined on $\{p_4, p_5\}$) that sends both p_4 and p_5 to 1, and the one that sends them both to 0. As a consequence, f_3 is extended as f_4 with $f_4(p_4) = f_4(p_5) = \{0, 1\}$.
- In the end, $[\![p_1]\!] = \emptyset$ (it is neither classically true nor classically false), $[\![p_2]\!] = [\![p_3]\!] = \{1\}$ (they are classically true) and $[\![p_4]\!] = [\![p_5]\!] = \{0, 1\}$ (they are neither classically true nor classically false).

If we had considered a model f such that $f(q) = \{0\}$ instead, the valuation obtained would have been such that $[\![p_1]\!] = [\![p_2]\!] = [\![p_3]\!] = \{0\}$ (they would be classically false) and $[\![p_4]\!] = [\![p_5]\!] = \{0, 1\}$ (they would still be neither classically true nor classically false).

Note, however, that the $f_{|V_c|}$ obtained, and thus the valuation, does not depend on the evaluation order used for the global semantics. This can be proved by considering two evaluation orders $(C_i)_{i \in I}$ and $(C'_i)_{i \in I}$ and the resulting sequences of functions $(f_i)_{i \in \{0\} \cup I}$ and $(f'_i)_{i \in \{0\} \cup I}$, and proving by induction on $i \in \{0\} \cup I$ that, with i' the smallest index such that $C_{0:i} \subseteq C'_{0:i'}$, for all $p \in C_{0:i}$, $f_i(p) = f'_{i'}(p)$. This is true due to the fact that, given some $i \in I$ and the j such that $C_i = C'_j$, a function g from C_i to $\mathbb{B}$ is a coherent local interpretation when extending f_{i-1} on C_i (resulting in f_i) if and only if it is a coherent local interpretation when extending f'_{j-1} on C_i (resulting in f'_j). This, in turn, is due to the fact that for all $q \in C_i$, $[\![def_C(q)]\!]^g_{f_{i-1}}$ (resp. $[\![def_C(q)]\!]^g_{f_{j-1}}$) depends only on g and on the values of f_{i-1} (resp. f'_{j-1}) on the $p \in C_{0:i-1}$ that q depends on. One consequence of the induction hypothesis is that f_{i-1} and f'_{j-1} are both defined and agree on all $p \in C_{0:i-1}$ that q depends on, and this entails that for all $q \in C_i$, $[\![def_C(q)]\!]^g_{f_{i-1}} = [\![def_C(q)]\!]^g_{f_{j-1}}$, leading to f_i and f'_j, and then also $f'_{i'}$, agreeing on $C_{0:i}$. As a consequence of this induction, the limits of both sequences $(f_i)_{i \in \{0\} \cup I}$ and $(f'_i)_{i \in \{0\} \cup I}$ are identical, which proves that the resulting valuation is independent of the choice of an evaluation order. The full proof is skipped for conciseness.

4 Discussion on the Liar, the Strengthened Liar and the Truth-teller

Consider any definitional context C that includes the following definition: $p_L \sim F(p_L)$. This definition is how the Liar is modelled in our system. Consider any model. Because p_L (has a definition in C and) does not depend on any other propositional letter, p_L is necessarily interpreted by itself; by this we mean that there is a step $i > 0$ of the global semantics at which f_i is defined by extending f_{i-1} to the singleton $C_i = \{p_L\}$. Let us consider such a step i. For any local interpretation g (i.e. any function from $\{p_L\}$ to $\mathbb{B}$), $[\![F(p_L)]\!]^g_{f_{i-1}} = 1 - g(p_L) \neq g(p_L)$. Thus, there is no coherent local interpretation at this step and $f_i(p_L)$ is $\emptyset$. This shows that the truth value of the Liar is $\emptyset$ in all models; the Liar is neither classically true nor classically false.

A common intuition then, is that if the Liar is neither (classically) true nor false, then it must therefore be false. This intuition is captured here by the fact that any other letter with the same definiens, i.e. $q_L \sim F(p_L) \in C$, is classically false (its truth value is $\{0\}$), independently of the model considered.

Similarly, the truth value of the Truth-teller $p_T \sim T(p_T)$ is always $\{0, 1\}$ and is thus neither classically true nor classically false. The common intuition here is that if the Truth-teller is neither (classically) true nor false, then it must be false. This intuition is, again, captured by the fact that any other letter with the same definiens, i.e. $q_T \sim T(p_T)$, is classically false.

Finally, the truth value of the Strengthened Liar $p_{SL} \sim \neg T(p_{SL})$ is always $\emptyset$. The intuition that it must therefore be true is captured by the fact that any other letter with the same definiens, i.e. $q_{SL} \sim \neg T(p_{SL})$ is classically true (its truth value is $\{1\}$).

5 Comparisons

The previous section has shown that by (i) making a distinction between a sentence and its content and (ii) modelling the former as a propositional letter and the later as its definiens, it is possible to account for the fact that a sentence may, for instance, seem both to be neither true nor false and to be false simpliciter. According to our analysis, the reason for this is that two distinct sentences are confused because of their sharing the same content, while one is neither true nor false and the other false simpliciter.

The fact that in our system the same content can be shared by two distinct sentences and thus lead to different truth values is reminiscent of contextualist approaches [2, §4.5, and references therein], according to which the truth value (or lack thereof) of sentences such as the Liar is context-dependent. However, we believe the resemblance to be mostly superficial. Approaches that follow [3] rely on the assumption that there are multiple truth predicates, or that the truth predicate has a contextual parameter; here, there is a single non-contextual T operator. Approaches that follow [7] rely on the assumption that the semantics of the truth predicate involves a quantification over a contextually restricted set of propositions; here, no such quantification is involved in the semantics of the T operator. The situation-based approach of [1] relies on the fact that a given proposition can be indeterminate in a situation and determinate in a larger situation; here, no situation is involved.

Our proposal shares some similarities with the work of [6]. But, while in Kripke's system the interpretation of the truth predicate is not entirely fixed a priori (in general, there exist multiple so-called fixed points), in the system presented here the operator T lexicalised by *is true* is a logical operator of fixed interpretation. Roughly, one could say that the truth value that we assign to a sentence is the set of truth values that this sentence can get in Krikpe's system when considering all fixed points. The way this is done allows one to talk non-trivially about non-classical sentences in the language itself, rather than only in the meta-language, as illustrated above with the Liar, Strengthened Liar, and Truth-teller. This is possible, in particular, because the operator T that we take to lexicalise *is true* is not a truth predicate in Kripke's sense. Indeed, this operator does not satisfy Tarski's convention T; for a propositional letter p, the interpretation of p is a *subset* of $\mathbb{B}$ and is thus never equal to the interpretation of $T(p)$, which is a *member* of $\mathbb{B}$. These two entities are nonetheless related through a truth-functional relationship: the interpretation of $T(p)$ is 1 iff the interpretation of p is $\{1\}$.

6 Future work

In its current version, our account cannot interpret sets of sentences with infinite dependencies (e.g. as found in $C = \{p_i \sim T(p_{i+1}) \mid i \in \mathbb{N}\}$). A solution would be to contract infinite upward chains as well as strongly connected components of the dependency graph; but other strategies may be available.

It also remains to be studied how to extend the present proposal with propositional quantification, so as to be able to express Yablo's paradox [9].

Finally, we would like to know if and how our proposal can be extended so as to formalise other logical paradoxes that do not necessarily rely on the concept of truth. The envisioned methodology, generalising what has been done here, would consist in "lifting" a paradox-prone D-valued semantics to a $\mathbb{P}(D)$-valued one in which classical cases correspond to singletons.

In particular, we would like to apply our method to the following paradox, that involves probabilities and that, to the extent of our knowledge, has not yet been discussed in the literature. This paradox relies on Table 1 and question (3).

(3) What is the probability of drawing the answer to this question by selecting randomly and uniformly a value in Table 1?

0	$\frac{1}{3}$	$\frac{1}{3}$

Table 1. An innocent table.

Assume temporarily that the answer to question (3) is 0. Then, because the probability of drawing 0 is $\frac{1}{3}$ and because $0 \neq \frac{1}{3}$, we have a contradiction. Now, assume instead that the answer is $\frac{1}{3}$. Then, because the probability of drawing $\frac{1}{3}$ is $\frac{2}{3}$ and because $\frac{1}{3} \neq \frac{2}{3}$, we have a contradiction. Assume that the answer is $\frac{2}{3}$ or, in fact, any number $x \notin \{0, \frac{1}{3}\}$. Then, because the probability of drawing x is 0 and because $x \neq 0$, we have a contradiction. Arguably, then, question (3) has no answer (or equivalently, the probability mentioned is undefined). But if the question has no answer, because the probability of drawing the correct answer is then 0, it seems intuitively that 0 should be the answer!

The existence of an answer to question (3) is similar to the truth of the Liar. We call this the *Matching Probability Paradox*.[2] We believe that this paradox can be handled following a method similar to the one presented in this paper, by moving from an intuitive yet paradox-prone $[0, 1]$-valued semantics to a $\mathbb{P}([0, 1])$-valued one. How to precisely do so is left for future work.

Acknowledgements I heartily thank Paul Égré for support and helpful discussion.

This work was supported by an Émergence 2021 grant (SYSNEULING project) from IdEx Université Paris Cité.

References

1. Barwise, J., Etchemendy, J.: The Liar: An Essay on Truth and Circularity. Oxford University Press, Oxford, New York (1987)
2. Beall, J., Glanzberg, M., Ripley, D.: Liar Paradox. In: Zalta, E.N., Nodelman, U. (eds.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, winter 2023 edn. (2023), https://plato.stanford.edu/archives/win2023/entries/liar-paradox/
3. Burge, T.: Semantical Paradox. The Journal of Philosophy **76**(4), 169–198 (Apr 1979). https://doi.org/10.2307/2025724
4. Cook, R.T.: Patterns of paradox. The Journal of Symbolic Logic **69**(3), 767–774 (Sep 2004). https://doi.org/10.2178/jsl/1096901765
5. van Fraassen, B.C.: Presupposition, Implication, and Self-Reference. The Journal of Philosophy **65**(5), 136–152 (1968). https://doi.org/10.2307/2024557, publisher: Journal of Philosophy, Inc.
6. Kripke, S.A.: Outline of a Theory of Truth. Journal of Philosophy **72**(19), 690–716 (Nov 1975). https://doi.org/10.2307/2024634, https://www.jstor.org/stable/2024634
7. Parsons, C.: The Liar Paradox. Journal of Philosophical Logic **3**(4), 381–412 (1974), https://www.jstor.org/stable/30226094, publisher: Springer
8. Priest, G.: The Logic of Paradox. Journal of Philosophical Logic **8**(1), 219–241 (1979), http://www.jstor.org/stable/30227165
9. Yablo, S.: Paradox without Self-Reference. Analysis **53**(4), 251–252 (Oct 1993). https://doi.org/10.1093/analys/53.4.251

[2] The oldest mention of a similar problem that we have been able to find is a Reddit post by a user called "allholy1": https://www.reddit.com/r/math/comments/c2p7u/multiple_choice_if_you_choose_an_answer_to_this/. allholy1's version, however, admits a solution (namely 0), as other Reddit users noted.