Noname manuscript No.
(will be inserted by the editor)

Adding a Conditional to Kripke's Theory of Truth

Lorenzo Rossi

The final version of this paper is to appear in the *Journal of Philosophical Logic*. The final publication is available at Springer via http://dx.doi.org/10.1007/s10992-015-9384-4

Abstract Kripke's theory of truth [29] has been very successful but shows well-known expressive difficulties; recently, Field has proposed to overcome them by adding a new conditional connective to it. In Field's theories, desirable conditional and truth-theoretic principles are validated that Kripke's theory does not yield. Some authors, however, are dissatisfied with certain aspects of Field's theories, in particular the high complexity. I analyze Field's models and pin down some reasons for discontent with them, focusing on the meaning of the new conditional and on the status of the principles so successfully recovered. Subsequently, I develop a semantics that improves on Kripke's theory following Field's program of adding a conditional to it, using some inductive constructions that include Kripke's one and feature a strong evaluation for conditionals. The new theory overcomes several problems of Kripke's one and, although weaker than Field's proposals, it avoids the difficulties that affect them; at the same time, the new theory turns out to be quite simple. Moreover, the new construction can be used to model various conceptions of what a conditional connective is, in ways that are precluded to both Kripke's and Field's theories.

Keywords Naïve truth · Kripke's theory of truth · Field's theories of truth · Conditional connective · Łukasiewicz logics · Partial semantics

I would like to express my gratitude to Volker Halbach, for all his encouragement and support at various stages of this work, and for the numerous helpful and beneficial discussions on the material of this paper. I am also obliged to Andrea Cantini, Hartry Field, Kentaro Fujimoto, Ole Hjortland, Leon Horsten, Harvey Lederman, Graham Leigh, Hannes Leitgeb, Pierluigi Minari, Carlo Nicolai, Graham Priest, James Studd, Philip Welch, Tim Williamson, and Andy Yu for many useful comments on this work. I am grateful to two anonymous referees for several observations and suggestions that led to improvements. Let me also thank the audiences at the University of Florence, the University of Oxford, the University of Bristol, the LOGICA 2014 Conference, the Humboldt University in Berlin, and the Technical University in Vienna for their precious feedback. Finally, I gratefully acknowledge the support of the Art and Humanities Research Council and of the Scatcherd European Scholarship.

Lorenzo Rossi (✉)
Faculty of Philosophy, University of Oxford, Oxford, UK
E-mail: lorenzo.rossi@philosophy.ox.ac.uk

1 Introduction

Several authors put great effort into developing a theory of naïve truth. By "naïve" I refer to the cluster of views according to which the notion of truth for a language $\mathscr{L}$ is characterized by the idea that, for some notion of equivalence, for every declarative sentence φ of $\mathscr{L}$, φ and "'φ' is true" are *equivalent* (where "'φ' is true" is also a sentence of $\mathscr{L}$). Due to well-known semantic paradoxes, the naïve conception of truth is inconsistent with classical logic.[1] Therefore, several ways to keep the naïve view and restrict classical logic have been investigated.

I will limit my attention to the so-called *paracomplete* approach, which on first approximation can be characterized as dropping the principle of excluded middle (henceforth "LEM"): in a paracomplete theory, for some sentence φ of $\mathscr{L}$, the sentence "φ or not φ" (also a sentence of $\mathscr{L}$) does not hold. Kripke [29] is considered a successful paracomplete theory of naïve truth: in Kripke's models, for every sentence φ of $\mathscr{L}$, φ has the same truth-value of "'φ' is true" (also a sentence of $\mathscr{L}$). Some also claim that Kripke's theory gives a nice treatment to long-debated paradoxes, such as the liar paradox (given by a sentence λ equivalent to "'λ' is not true"): in (consistent applications of) Kripke's theory, λ and its negation do not have a truth-value. Regrettably, Kripke's theory is affected by serious expressive difficulties:

(K1) In Kripke's semantics "there are no laws", i.e. there is no schematic law s.t. all its instances are validated by (consistent applications of) Kripke's construction.[2]
(K2) The relation between φ and "'φ' is true" cannot be represented within the theory in many cases, including apparently simple ones.
(K3) It is doubtful whether Kripke's theory accommodates paradoxical sentences such as the liar in a satisfactory way.
(K4) The relation between paradoxical sentences cannot be represented within the theory, also in apparently simple cases (e.g. the liar sentence and its negation).

Hartry Field has convincingly argued that one could address these difficulties, to a large extent, by adding a "good" primitive conditional connective $\rightarrow$ to Kripke's theory (and its standardly defined biconditional $\leftrightarrow$) – for now, let "good" mean "better than the conditional of Kripke's theory". In a series of works ([14], [15], [16], [17], [18], [20]), he proposed and defended models of the so-enlarged language that preserve the nice features of Kripke's theory and validate principles that Kripke's theory cannot give (e.g. "'φ' is true $\leftrightarrow \varphi$", for *every* sentence φ). Some points in Field's theories are, however, problematic and prompt some critical investigations.

The plan of the paper is as follows. In Section 2, I will sketch Kripke's theory, review the main facts about it and discuss some of the expressive limitations presented above. In Section 3, I will sketch Field's theory and highlight some conceptual difficulties that surround it. This will bring me to the main question of the paper: how to improve on Kripke's theory with a new conditional that avoids the problems of Field's approach. Section 4 presents a simple construction that goes some way to-

[1] I ignore languages and theories that do not fulfill the syntactic requirements necessary for Tarski's Theorem on the Undefinability of truth to hold for them. See Tarski [39].

[2] By "laws" I refer to inference schemata without premisses, and by "rules" to the inference schemata that have premisses. The term "principle" refers to laws and rules alike.

ward addressing the question. The new theory has some nice properties, reviewed in Section 5, and gives an interesting reading of the conditional, as argued in Section 6. Finally, Section 7 presents an application of this theory that captures some intuitions on the conditional that cannot be captured in Kripke's or in Field's framework.

Now I will introduce the more general notational conventions adopted in the paper – most of them are from Halbach [25]. I will omit quotation marks where no use/mention confusion can arise. Let $\mathscr{L}$ be the language of first-order arithmetic. Let $\mathscr{L}^{\rightarrow} := \mathscr{L} \cup \{\rightarrow\}$ for a new binary connective $\rightarrow$ (for "if ... then ..."). Let $\mathscr{L}_T := \mathscr{L} \cup \{T\}$, for a fresh unary predicate T (for "... is true"), which applies to terms of the language. Finally, let $\mathscr{L}_T^{\rightarrow} := \mathscr{L} \cup \{\rightarrow\} \cup \{T\}$.[3] $\supset$ denotes the material conditional. The existential quantifier ($\exists$) and biconditionals are defined as usual: I use $\leftrightarrow$ for the new arrow and $\equiv$ for the material biconditional. My base theory is Peano Arithmetic formulated in $\mathscr{L}_T^{\rightarrow}$, call it $\mathsf{PA}_T^{\rightarrow}$: this theory has no axioms nor rules for T or $\rightarrow$.[4] Terms and formulae of $\mathscr{L}_T^{\rightarrow}$ are defined as usual. Formulae with no free variables are called "sentences", whereas "term" refers to closed and open terms alike. I use $s_1, t_1, \dots, s_n, t_n, \dots$ to range over terms of $\mathscr{L}_T^{\rightarrow}$. Lowercase Greek letters early in the alphabet ($\alpha, \beta, \gamma, \delta$) refer to ordinals, with the only exception of ω (the least infinite ordinal) and ω_1^{CK} (the least non-recursive ordinal). Lowercase Greek letters middle in the alphabet (e.g. ζ, η, ϑ) refer to formulae of the language of the theory of inductive definitions with set variables allowed; lowercase Greek letters late in the alphabet (e.g. φ, ψ, χ), possibly with indices, refer to $\mathscr{L}_T^{\rightarrow}$-formulae; finally, uppercase Greek letters early in the alphabet (Γ, Δ) refer to sets of $\mathscr{L}_T^{\rightarrow}$-formulae. Sets of natural numbers are indicated with P, Q, R (unless indicated otherwise), set variables are indicated with $S_1, S_2, \dots, S_i, \dots$, and operators on them are indicated with uppercase Greek letters late in the alphabet (Φ, Ψ, Υ).[5] I adopt some coding system for $\mathscr{L}_T^{\rightarrow}$, e.g. of the kind described in van Dalen [40] (details are unimportant). φ within Gödel quotes, $\ulcorner\varphi\urcorner$, denotes the numeral of the code of φ, and if f is a (primitive recursive) function, then $\dot{f}$ denotes the function as represented in $\mathsf{PA}_T^{\rightarrow}$ (for details about arithmetical representation and for quantification over variables occurring in sentences within Gödel quotes, see Halbach [25], Ch. 5). As customary, expressions of $\mathscr{L}_T^{\rightarrow}$ are identified with their codes. I use $TER_{\mathscr{L}_T^{\rightarrow}}$, $CTER_{\mathscr{L}_T^{\rightarrow}}$, $SENT_{\mathscr{L}_T^{\rightarrow}}$, $FOR_{\mathscr{L}_T^{\rightarrow}}$ to indicate the sets of (codes of) terms, closed terms, sentences and formulae of $\mathscr{L}_T^{\rightarrow}$, respectively. "$\varphi \in \mathscr{L}_T^{\rightarrow}$" is a shorthand for "$\varphi \in SENT_{\mathscr{L}_T^{\rightarrow}}$". Some lowercase Greek letters are reserved for specific sentences (which exist and are unique by the diagonal lemma of $\mathsf{PA}_T^{\rightarrow}$): (1) λ is called the *liar sentence*, and designates the sentence $\neg Tt_\lambda$ s.t. $dec(t_\lambda) = \ulcorner \neg Tt_\lambda \urcorner$, where $dec(x)$ is the function that takes as argument a closed $\mathscr{L}_T^{\rightarrow}$-term t and yields the value of t. (2) κ, the *Curry sentence*, designates the sentence $Tt_\kappa \rightarrow 0 \neq 0$ s.t. $dec(t_\kappa) = \ulcorner Tt_\kappa \rightarrow 0 \neq 0 \urcorner$.

[3] The predicate T, syntactically, may apply to every $\mathscr{L}_T^{\rightarrow}$-term; in practice, though, I will be mostly interested in its applications to closed $\mathscr{L}_T^{\rightarrow}$-terms coding closed $\mathscr{L}_T^{\rightarrow}$-formulae – I will introduce the conventions about coding in a moment.

[4] For $\mathscr{L}$ and PA, see Kaye [27]. Adopting $\mathscr{L}$ and PA is purely a matter of convenience and causes no loss of generality, many other languages and theories that are syntactically expressive enough would do.

[5] Here and in what follows, I will use results from the theory of inductive definitions. For the general theory and the relative notational conventions see Moschovakis [35].

2 Kripke's theory

The theory developed by Saul Kripke in his celebrated paper [29] is widely considered one of the greatest contributions to theories of truth since Tarski's groundbreaking essay [39]. Kripke's work delivers a framework to expand a model of $\mathscr{L}^{\rightarrow}$ to a model of $\mathscr{L}_T^{\rightarrow}$ where the predicate T respects some form of naïveté.[6] Kripke's construction can be carried forward with several evaluation schemata for the logical vocabulary. However, I will only consider the version of Kripke's construction that uses strong Kleene logic, henceforth K3, since it is the most relevant for the issues addressed in the present work. In what follows, I will always refer to "Kripke's theory" (or the like) meaning the version of Kripke's construction based on K3. [7]

Definition 1 (Kripke's construction for $\mathscr{L}_T^{\rightarrow}$)
For $S_1, S_2 \subseteq \omega$, define the pair $\langle S_1^+, S_2^- \rangle$ so that:

1. $n \in S_1^+$ if $n \in S_1$, or
   (i) n is $s = t$ and $s \in CTER_{\mathscr{L}_T^{\rightarrow}}$ and $t \in CTER_{\mathscr{L}_T^{\rightarrow}}$ and $dec(s) = dec(t)$, or
   (ii) n is $\neg\varphi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\varphi \in S_2$, or
   (iii) n is $\varphi \wedge \psi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\psi \in \mathscr{L}_T^{\rightarrow}$ and ($\varphi \in S_1$ and $\psi \in S_1$), or
   (iv) n is $\varphi \vee \psi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\psi \in \mathscr{L}_T^{\rightarrow}$ and ($\varphi \in S_1$ or $\psi \in S_1$), or
   (v) n is $\forall x \chi(x)$ and $\chi(x) \in FOR_{\mathscr{L}_T^{\rightarrow}}$ and, for all $t \in CTER_{\mathscr{L}_T^{\rightarrow}}$, $\chi(t) \in S_1$, or
   (vi) n is Tt and $t \in CTER_{\mathscr{L}_T^{\rightarrow}}$ and $dec(t) = \ulcorner \chi \urcorner$ and $\chi \in \mathscr{L}_T^{\rightarrow}$ and $\chi \in S_1$.
2. $n \in S_2^-$ if $n \in S_2$, or
   (i) n is $s = t$ and $s \in CTER_{\mathscr{L}_T^{\rightarrow}}$ and $t \in CTER_{\mathscr{L}_T^{\rightarrow}}$ and $dec(s) \neq dec(t)$, or
   (ii) n is $\neg\varphi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\varphi \in S_1$, or
   (iii) n is $\varphi \wedge \psi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\psi \in \mathscr{L}_T^{\rightarrow}$ and ($\varphi \in S_2$ or $\psi \in S_2$), or
   (iv) n is $\varphi \vee \psi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\psi \in \mathscr{L}_T^{\rightarrow}$ and ($\varphi \in S_2$ and $\psi \in S_2$), or
   (v) n is $\forall x \chi(x)$ and $\chi(x) \in FOR_{\mathscr{L}_T^{\rightarrow}}$ and there exists at least one $t \in CTER_{\mathscr{L}_T^{\rightarrow}}$ s.t. $\chi(t) \in S_2$, or
   (vi) n is Tt and (($t \in TER_{\mathscr{L}_T^{\rightarrow}}$ and $dec(t) \notin \mathscr{L}_T^{\rightarrow}$) or ($dec(t) = \ulcorner \chi \urcorner$ and $\chi \in \mathscr{L}_T^{\rightarrow}$ and $\chi \in S_2$)).

This definition is by simultaneous induction, and it is inductive in S_1 and S_2.[8] Let $\zeta_1(n, S_1, S_2)$ abbreviate the right-hand side of item 1, and $\zeta_2(n, S_1, S_2)$ abbreviate the right-hand side of item 2. $\zeta_1(n, S_1, S_2)$ and $\zeta_2(n, S_1, S_2)$ are positive elementary formulae, positive in S_1 and S_2. Associate to them a monotone operator, called *Kripke*

[6] Kripke, in fact, considered expansions of models of $\mathscr{L}$ to models of $\mathscr{L}_T$, and not expansions of models of $\mathscr{L}^{\rightarrow}$ to models of $\mathscr{L}_T^{\rightarrow}$. However, I present the construction for $\mathscr{L}_T^{\rightarrow}$ rather than $\mathscr{L}_T$, since I will use it later to interpret $\rightarrow$. This change induces no significant differences. For detailed analyses of Kripke's construction, see Halbach [25], Horsten [26], Kremer [28], McGee [34], Soames [38].

[7] I thank an anonymous referee for helping me to clarify this point. Other evaluation schemata considered by Kripke include supervaluationism and weak Kleene logic: I do not consider them since supervaluationism does not yield a paracomplete theory, while weak Kleene logic would not mingle well with my treatment of $\rightarrow$. Some evaluation schemata for the logic vocabulary that admit (a generalized version of) Kripke's treatment are studied in Feferman [13], applying results of Aczel and Feferman [1].

[8] The clause "$dec(t) \notin \mathscr{L}_T^{\rightarrow}$" doesn't make Definition 1 non-inductive, as $SENT_{\mathscr{L}_T^{\rightarrow}}$ is hyperelementary.

jump, acting on pairs of sets, $\Phi : \mathcal{P}(\omega) \times \mathcal{P}(\omega) \longmapsto \mathcal{P}(\omega) \times \mathcal{P}(\omega)$ as:[9]

$$\Phi(S_1, S_2) := \langle \{n \in \omega \,|\, \zeta_1(n, S_1, S_2)\}, \{n \in \omega \,|\, \zeta_2(n, S_1, S_2)\} \rangle.$$

Let $\leq_S$ denote the partial ordering of tuples of sets induced by $\subseteq$, i.e.:

$$\langle P_1, \dots, P_n \rangle \leq_S \langle Q_1, \dots Q_n \rangle \text{ iff } (P_1 \subseteq Q_1 \text{ and } \dots \text{ and } P_n \subseteq Q_n).$$

Put $\mathfrak{I}_\Phi := \bigcup_{\alpha \in Ord} \Phi^\alpha(\emptyset, \emptyset)$. $\mathfrak{I}_\Phi$ is the so-called *least Kripke fixed point*, since it is the $\leq_S$-*least* element of the partial order of the fixed points of Φ induced by $\leq_S$. For $P, Q \subseteq \omega$, put $\mathfrak{I}_\Phi(P, Q) := \bigcup_{\alpha \in Ord} \Phi^\alpha(P, Q)$, i.e. $\mathfrak{I}_\Phi(P, Q)$ is the fixed point of Φ obtained starting from $\langle P, Q \rangle$. $\mathfrak{I}_\Phi(P, Q)$ will be called "Kripke fixed point" or "Kripke fixed-point pair". The first element of a Kripke fixed-point $\mathfrak{I}_\Phi(P, Q)$ is its "Kripke truth-set", in symbols $\mathfrak{E}_\Phi(P, Q)$, the second element is its "Kripke falsity-set", indicated by $\mathfrak{A}_\Phi(P, Q)$. Clearly, $\mathfrak{I}_\Phi(P, Q) = \langle \mathfrak{E}_\Phi(P, Q), \mathfrak{A}_\Phi(P, Q) \rangle$.

Φ is defined by the clauses for preservation of truth-values 1 and 0 of K3, plus a naïve reading of T.[10] φ has value 1 (0) in a Kripke fixed point if it is in its Kripke truth-set (falsity-set). A fixed point $\mathfrak{I}_\Phi(P, Q)$ is *consistent* if $\mathfrak{E}_\Phi(P, Q) \cap \mathfrak{A}_\Phi(P, Q) = \emptyset$, and it is *inconsistent* otherwise.

Kripke fixed-point pairs give a naïve characterization to T.

Theorem 2 (Weak Naïveté (Kripke))
For every $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $P, Q \subseteq \omega$:

- $\varphi \in \mathfrak{E}_\Phi(P, Q)$ *if and only if* $T\ulcorner \varphi \urcorner \in \mathfrak{E}_\Phi(P, Q)$.
- $\varphi \in \mathfrak{A}_\Phi(P, Q)$ *if and only if* $T\ulcorner \varphi \urcorner \in \mathfrak{A}_\Phi(P, Q)$.

Corollary 3 (Intersubstitutivity of truth for Kripke's theory (Field)) [11]
For every $\varphi, \psi, \chi \in \mathscr{L}_T^{\rightarrow}$ and $P, Q \subseteq \omega$, if ψ and χ are alike except that one of them has an occurrence of φ where the other has an occurrence of $T\ulcorner \varphi \urcorner$, then:

- $\psi \in \mathfrak{E}_\Phi(P, Q)$ *if and only if* $\chi \in \mathfrak{E}_\Phi(P, Q)$.
- $\psi \in \mathfrak{A}_\Phi(P, Q)$ *if and only if* $\chi \in \mathfrak{A}_\Phi(P, Q)$.

Kripke's theory only features values 1 and 0, so it cannot give a value to the liar sentence (and to many other paradoxical sentences as well), on pain of inconsistency. From this fact, the points (K1) and (K2) raised in the Introduction are immediate.

Proposition 4 (Kripke)
If $\lambda \in \mathfrak{E}_\Phi(P, Q)$ or $\lambda \in \mathfrak{A}_\Phi(P, Q)$, then $\mathfrak{I}_\Phi(P, Q)$ is inconsistent.

Kripke's construction yields ⟨Kripke truth-set, Kripke falsity-set⟩ pairs, and λ cannot be in the truth-set nor in the falsity-set of any consistent Kripke fixed point. If φ is neither in $\mathfrak{E}_\Phi(P, Q)$ nor in $\mathfrak{A}_\Phi(P, Q)$, this means that it has neither value 1 nor value 0 in $\mathfrak{I}_\Phi(P, Q)$; it does *not* mean that φ is not true in the sense of the truth

9 Of course $\Phi(S_1, S_2)$ is just a shorthand for $\Phi(\langle S_1, S_2 \rangle)$.

10 For K3, see Blamey [5]. I use 1 (0, 1/2) for his $\top$ ($\bot$, $*$).

11 See Field [18], especially p. 12 and p. 65. A discussion on the difference between the rules encoded in Theorem 2 and Intersubstitutivity of truth is in McGee [34], Ch. 10, in the wake of Dummett [12].

predicate, i.e. in the sense that $\neg T\ulcorner\varphi\urcorner$ holds in $\mathfrak{I}_{\Phi}(P,Q)$, otherwise $\neg T\ulcorner\varphi\urcorner$ would be in $\mathfrak{E}_{\Phi}(P,Q)$ and $\neg\varphi$ would be in $\mathfrak{E}_{\Phi}(P,Q)$. Sentences that are neither in $\mathfrak{E}_{\Phi}(P,Q)$ nor in $\mathfrak{A}_{\Phi}(P,Q)$ have come to be called "gappy" with respect to $\mathfrak{I}_{\Phi}(P,Q)$ and they are sometimes considered to have the third value $1/2$ of K3 semantics. This way of defining "gaps" uses the *complement* of the union of the truth-set and the falsity-set of a Kripke fixed point, so I will call such sets "C-gaps" and the sentences "C-gappy".

C-gaps are sometimes used to argue that Kripke's theory *treats successfully* paradoxical sentences. Presumably, this means that Kripke's theory can produce a consistent interpretation of $\mathscr{L}_T^{\rightarrow}$, which respects naïveté as specified above, avoiding the problems arising from a sentence such as the liar. In fact, λ is in the C-gap of any consistent Kripke fixed point. It is sometimes said, then, that λ is "gappy" according to Kripke's theory. The latter statement, however, is not correct: Kripke's theory cannot "talk" about C-gaps (of consistent fixed points). But then, Kripke's theory does not *treat* paradoxical sentences such as λ at all. Kripke's theory only yields $\langle$truth-set, falsity-set$\rangle$ pairs, so it does not give also their complement. But the problem is deeper than that: the C-gaps of consistent fixed points are not sets that Kripke's theory can define, as Kripke's construction is an inductive definition but the former are not inductive sets. For every $P,Q \subseteq \omega$, if $\mathfrak{I}_{\Phi}(P,Q)$ is consistent, it is inductive and not co-inductive in P,Q, and $SENT_{\mathscr{L}_T^{\rightarrow}} \setminus (\mathfrak{E}_{\Phi}(P,Q) \cup \mathfrak{A}_{\Phi}(P,Q))$ is co-inductive and not inductive in P,Q.[12] If two sets $A,B \subseteq \omega$ form a consistent Kripke fixed-point pair, then $\langle A,B\rangle = \mathfrak{I}_{\Phi}(P,Q)$ for some $P,Q \subseteq \omega$, but there is no inductive definition over P and Q that defines the resulting C-gap. $SENT_{\mathscr{L}_T^{\rightarrow}} \setminus (\mathfrak{E}_{\Phi}(P,Q) \cup \mathfrak{A}_{\Phi}(P,Q))$ cannot result from an application of Kripke's construction and no information on membership in it obtains from the positive inductive means available to Kripke's theory. C-gaps are only visible "from the outside" of Kripke's theory, as it were, and the Kripkean theorist must be silent about them: if she talks about C-gaps, she accepts resources that go beyond Kripke's theory, i.e. she is not a Kripkean theorist any more.[13]

This is unfortunate: talking about C-gaps would be desirable, since it would give a means to treat explicitly problematic sentences such as the liar.[14] The inability to use C-gaps is one of the main roots of the expressive deficiencies of Kripke's theory. Suppose that we think that λ lacks a classical truth-value. Then, the biconditional

the liar sentence is true exactly if its negation is true (G)

voices a natural reaction to the liar and is informative because it makes explicit its lack of classical value. But (G) is unaccountable for within Kripke's theory:

Corollary 5
For every $P,Q \subseteq \omega$, if $T\ulcorner\lambda\urcorner \equiv T\ulcorner\neg\lambda\urcorner \in \mathfrak{E}_{\Phi}(P,Q)$ or $T\ulcorner\lambda\urcorner \equiv T\ulcorner\neg\lambda\urcorner \in \mathfrak{A}_{\Phi}(P,Q)$, then $\mathfrak{I}_{\Phi}(P,Q)$ is inconsistent.

[12] This is well-known: see, e.g., McGee [34], Corollary 5.11, p. 113.

[13] The extent and depth of such "silence" are quite radical: see Field [18], p. 72. Clearly, C-Gaps are "visible" in Kripke's theory from a purely set-theoretical standpoint, but I am considering only what (consistent) Kripke fixed points can yield. I thank an anonymous referee for pointing out to me the relevance of Horsten [26] on the question of the radical silence under discussion. See, for example, Horsten's observations on the theory PKF (an axiomatization of Kripke's theory), in Chapter 10.2.2, pp. 144-146.

[14] I do not go into this debate, but opinions diverge a lot on the importance of C-gaps, both *per se* and within Kripke's theory: see McGee [33] and Soames [38] for two quite different stances on the matter.

The latter observations exhibit some cases of the expressive deficiencies (K3) and (K4) mentioned in the Introduction, completing a quick survey of the difficulties of Kripke's theory. The impossibility to evaluate sentences such as (G) in Kripke's theory follows from the impossibility to treat consistently C-gaps as well as from the functioning of the K3 material conditional. Subsequently, some authors constructed models that treat C-gappy sentences. At the same time, they added to Kripke's theory a new conditional that can take the designated value based on cases in which its components are C-gappy relative to some Kripke fixed point.

3 Field's theory

Some models that use interestingly C-gaps were developed by Hartry Field. He gave two main constructions: the one culminating in [18], and the one in [20]. I will focus on the first one, because most of the criticisms I will raise in conjunction to it are adaptable to the second theory as well; also, the first theory is stronger in terms of validating logical principles.[15] So, when referring to "Field's theory" (or the like), I will always mean the first construction. I focus on Field's work because it is the one, that I know of, that takes up more directly with Kripke's theory and presents itself as a "completion" of it. Many (including myself) consider Field's theory the most advanced and successful paracomplete theory of naïve truth currently available.[16]

3.1 A sketch of Field's theory

Field's theory is based on a *revision-theoretic* conception: here, however, it is the value of conditionals that undergoes revision, and not the value of sentences of the form $T\ulcorner\varphi\urcorner$.[17] The main tool of such revision is the evaluation generated by consistent Kripke fixed points plus their C-gap.

Definition 6
Let $\mathfrak{I}_\Phi(P,Q)$ be consistent. Define $\mathscr{K}_{\langle\mathfrak{E}_\Phi(P,Q),\mathfrak{A}_\Phi(P,Q)\rangle} : SENT_{\mathscr{L}_T^{\rightarrow}} \longmapsto \{1,0,1/2\}$ as:

$$\mathscr{K}_{\langle\mathfrak{E}_\Phi(P,Q),\mathfrak{A}_\Phi(P,Q)\rangle}(\varphi) = \begin{cases} 1, \text{ if } \varphi \in \mathfrak{E}_\Phi(P,Q) \\ 0, \text{ if } \varphi \in \mathfrak{A}_\Phi(P,Q) \\ 1/2, \text{ if } \varphi \notin (\mathfrak{E}_\Phi(P,Q) \cup \mathfrak{A}_\Phi(P,Q)) \end{cases}$$

Take the evaluation given by $\mathfrak{I}_\Phi$, i.e. $\mathscr{K}_{\langle\mathfrak{E}_\Phi(\emptyset,\emptyset),\mathfrak{A}_\Phi(\emptyset,\emptyset)\rangle}$, call it $\mathscr{K}$ for short. Even if every conditional $\varphi\rightarrow\psi$ is in the C-gap of $\mathfrak{I}_\Phi$, we can check whether the value of φ is less than or equal to the value of ψ according to $\mathscr{K}$: let's say that $\varphi\rightarrow\psi$ gets value 1 at the first revision stage if this is the case, and value 0 otherwise. To evaluate non-conditionals, we build a Kripke fixed point over the results of the first revision. This process is iterated indefinitely, as specified in the following Definition.

[15] For similar reasons, I do not consider explicitly the still different theory in Field [14].

[16] Other notable theories of naïve truth featuring strong conditionals include Brady [7] and Bacon [2]; however, I will not discuss them as they are not very much related to the problems I address here.

[17] For the revision theory of truth, see Gupta and Belnap [23]. Here, I will only expound the main aspects of Field's theory, without many subtleties: for more details, see Field [15] and [18] (Chs. 15-23).

Definition 7 (Field revision construction)
Field's revision construction is given by the function $r : Ord \times SENT_{\mathscr{L}_T^{\rightarrow}} \longmapsto \{1, 0, 1/2\}$, defined as follows:

– $$r(0,\varphi) := \begin{cases} 1, \text{ if } \varphi \text{ is } \psi \rightarrow \chi \text{ and } \mathscr{K}(\psi) \leq \mathscr{K}(\chi) \\ 0, \text{ if } \varphi \text{ is } \psi \rightarrow \chi \text{ and } \mathscr{K}(\psi) > \mathscr{K}(\chi) \\ \mathscr{K}_0(\varphi), \text{ if } \varphi \text{ is not a conditional,} \end{cases}$$

where $\mathscr{K}_0 := \mathscr{K}_{\langle F_E^0, F_A^0 \rangle}$, and $\langle F_E^0, F_A^0 \rangle$ is

$$\mathfrak{I}_\Phi(\{\psi \rightarrow \chi \in \mathscr{L}_T^{\rightarrow} \,|\, r(0, \psi \rightarrow \chi) = 1\}, \{\psi \rightarrow \chi \in \mathscr{L}_T^{\rightarrow} \,|\, r(0, \psi \rightarrow \chi) = 0\}).$$

Strictly speaking, two steps are conflated in the above graph bracket: *first* $\mathscr{K}$ evaluates conditionals, *then* $\mathscr{K}_0$ is built, being generated by the above Kripke fixed point, which uses the conditionals just evaluated using $\mathscr{K}$. Here and in the following, I write together these two steps for simplicity.

– $$r(\alpha+1,\varphi) := \begin{cases} 1, \text{ if } \varphi \text{ is } \psi \rightarrow \chi \text{ and } \mathscr{K}_\alpha(\psi) \leq \mathscr{K}_\alpha(\chi) \\ 0, \text{ if } \varphi \text{ is } \psi \rightarrow \chi \text{ and } \mathscr{K}_\alpha(\psi) > \mathscr{K}_\alpha(\chi) \\ \mathscr{K}_{\alpha+1}(\varphi), \text{ if } \varphi \text{ is not a conditional,} \end{cases}$$

where $\mathscr{K}_{\alpha+1} := \mathscr{K}_{\langle F_E^{\alpha+1}, F_A^{\alpha+1} \rangle}$, and $\langle F_E^{\alpha+1}, F_A^{\alpha+1} \rangle$ is

$$\mathfrak{I}_\Phi(\{\psi \rightarrow \chi \in \mathscr{L}_T^{\rightarrow} | r(\alpha+1, \psi \rightarrow \chi) = 1\}, \{\psi \rightarrow \chi \in \mathscr{L}_T^{\rightarrow} | r(\alpha+1, \psi \rightarrow \chi) = 0\}),$$

and similarly for $\mathscr{K}_\alpha$.

– $$r(\delta,\varphi) := \begin{cases} 1, \text{ if } \varphi \text{ is } \psi \rightarrow \chi \text{ and there is a } \gamma < \delta \text{ s.t.} \\ \quad \text{for all } \beta \text{ s.t. } \beta \geq \gamma \text{ and } \beta < \delta, \mathscr{K}_\beta(\psi) \leq \mathscr{K}_\beta(\chi) \\ 0, \text{ if } \varphi \text{ is } \psi \rightarrow \chi \text{ and there is a } \gamma < \delta \text{ s.t.} \\ \quad \text{for all } \beta \text{ s.t. } \beta \geq \gamma \text{ and } \beta < \delta, \mathscr{K}_\beta(\psi) > \mathscr{K}_\beta(\chi) \\ 1/2, \text{ if } \varphi \text{ is } \psi \rightarrow \chi \text{ and none of the two above cases is given} \\ \mathscr{K}_\delta(\varphi), \text{ if } \varphi \text{ is not a conditional,} \end{cases}$$

for δ a limit ordinal, where $\mathscr{K}_\beta$ is defined as above for every $\beta \leq \delta$.

Field then defines an *ultimate evaluation* for $\mathscr{L}_T^{\rightarrow}$-sentences, using a natural tripartition of the outcomes of the revision process. Let $1^u, 0^u, 1/2^u$ be the *ultimate values*.

Definition 8 (Ultimate evaluation function, ultimate values)
Field's ultimate evaluation is the function $\mathscr{U} : SENT_{\mathscr{L}_T^{\rightarrow}} \longmapsto \{1^u, 0^u, 1/2^u\}$ such that:

1. $\mathscr{U}(\varphi) = 1^u$, if there is an ordinal α s.t. for all $\beta \geq \alpha$, $r(\beta, \varphi) = 1$.
2. $\mathscr{U}(\varphi) = 0^u$, if there is an ordinal α s.t. for all $\beta \geq \alpha$, $r(\beta, \varphi) = 0$.
3. $\mathscr{U}(\varphi) = 1/2^u$, if neither of the two cases above obtains.

Theorem 9 (Field's fundamental theorem on ultimate values)
There are ordinals γ, called acceptable, *s.t. for every $\varphi \in \mathscr{L}_T^{\rightarrow}$, $r(\gamma, \varphi) = \mathscr{U}(\varphi)$.*

Acceptable ordinals can be seen as large enough to represent the revision process with suitable reliability. A peculiarity of $\mathscr{U}$ is the behavior of $1/2^u$: it hides many different revision sequences, with problematic results. For example, if the revision sequence of ψ oscillates between 0 and $1/2$ and the revision sequence of χ oscillates between $1/2$ and 1, $\mathscr{U}(\psi) = 1/2^u = \mathscr{U}(\chi)$ and $\mathscr{U}(\psi \to \chi) = 1^u$, but $\mathscr{U}(\chi \to \psi) = 1/2^u$.

To avoid such problems, Field develops a *fine-grained* evaluation that assigns to φ *its revision sequence* as value. Let γ_0 be the first acceptable ordinal: for simplicity, following Field, I will only consider revision sequences after γ_0.

Definition 10 (Field's fine-grained value space)
To describe Field's fine-grained value space, consider the following ordinals:

- Let δ_0 be the ordinal s.t. $\gamma_0 + \delta_0$ is the least acceptable ordinal after γ_0.
- Let δ^+ be the smallest initial ordinal whose cardinality is greater than that of δ_0.
- Let $Pred(\delta^+)$ indicate the ordinals smaller than δ^+. Note that $\gamma_0 + \delta^+ = \delta^+$.

Field's fine-grained value space, denoted with **F**, is the set of all functions $\mathfrak{f}$ from $Pred(\delta^+)$ to $\{1, 0, 1/2\}$ s.t. the following conditions are satisfied:

1. If $\mathfrak{f}(0) = 1$, then for all ordinals α, $\mathfrak{f}(\alpha) = 1$.
2. If $\mathfrak{f}(0) = 0$, then for all ordinals α, $\mathfrak{f}(\alpha) = 0$.
3. If $\mathfrak{f}(0) = 1/2$, then there is an ordinal $\alpha' < \delta^+$ s.t. for all ordinals α and β, if $\alpha' \cdot \alpha + \beta < \delta^+$, then $\mathfrak{f}(\alpha' \cdot \alpha + \beta) = \mathfrak{f}(\beta)$.

Some noteworthy elements of Field's value space are: the function that is constant on revision value 1, denoted with **1** (which is the designated value), and the functions **0** and **1/2**, defined similarly with revision values 0 and $1/2$ respectively.

The algebraic counterpart of Field's conditional is given by the following partial ordering $\preceq$ on **F**, defined pointwise: for every $\mathfrak{f}, \mathfrak{g} \in \mathbf{F}$,

$$\mathfrak{f} \preceq \mathfrak{g} :\Leftrightarrow \text{ for every ordinal } \alpha < \delta^+, \mathfrak{f}(\alpha) \leq \mathfrak{g}(\alpha). \tag{1}$$

Definition 11 (Field's fine-grained evaluation)
Field's fine-grained evaluation is the function $|\,|_F : SENT_{\mathscr{L}_T^{\to}} \longmapsto \mathbf{F}$ defined as:

$$|\varphi|_F := \langle r(\gamma_0 + \alpha, \varphi) : \alpha < \delta^+ \rangle.$$

Field uses (without giving many details) a relation of logical consequence, indicate it with $\models_F$, that he reads as "*preserv[ing the] designated value in all models*".[18] I interpret this to mean that $\models_F$ preserves of the designated value *in every model built using Field's construction*, and I suggest the following formalization:

$\Gamma \models_F \psi$:⇔ for every function $|\,|_F^{\mathscr{M}}$ defined as in Definition 11 using a countable ω-model $\mathscr{M}$ of $\mathscr{L}$, if for every $\varphi \in \Gamma$, $|\varphi|_F^{\mathscr{M}} = \mathbf{1}$, then also $|\psi|_F^{\mathscr{M}} = \mathbf{1}$.[19]

[18] Field [18], p. 267.

[19] Field's construction uses always countable ω-models of $\mathscr{L}$, but there is exactly one such model (up to isomorphism). So, if the definition of $\models_F$ quantifies over more than one evaluation, it must quantify over the functions $|\,|_F^{\mathscr{M}}$ that are just like $|\,|_F$ with the exception of using a different Kripke fixed point at revision stages. Field is not explicit as to which Kripke fixed points can be used in his theory (see [18], p. 249), and it is natural to think that some Kripke fixed point will not do. However, we can suppose that the Fieldian theorist can place a restriction on the quantification over such Kripke fixed points, if needed.

Theorem 12 (Some important principles recovered by Field's theory (Field))

- $\models_F \varphi \to \varphi$
- $\models_F \neg\neg\varphi \leftrightarrow \varphi$
- $\models_F \varphi \wedge \psi \to \varphi$; $\models_F \varphi \wedge \psi \to \psi$
- $\models_F \varphi \to \varphi \vee \psi$; $\models_F \psi \to \varphi \vee \psi$
- $\models_F (\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$
- $\models_F \varphi \wedge (\psi \vee \chi) \leftrightarrow (\varphi \wedge \psi) \vee (\varphi \wedge \chi)$
- $\models_F \neg(\varphi \wedge \psi) \leftrightarrow \neg\varphi \vee \neg\psi$
- $\models_F \neg(\varphi \vee \psi) \leftrightarrow \neg\varphi \wedge \neg\psi$
- $\models_F \forall x\varphi(x) \to \varphi(x/t)$
- $\models_F \varphi(x/t) \to \exists x\varphi(x)$

- $\varphi, \psi \models_F \varphi \wedge \psi$
- $\varphi, \varphi \to \psi \models_F \psi$
- $\varphi, \neg\psi \models_F \neg(\varphi \to \psi)$
- $\varphi \models_F \psi \to \varphi$
- $\varphi \to \psi \models_F (\psi \to \chi) \to (\varphi \to \chi)$
- $(\varphi \to \psi) \models_F (\varphi \wedge \chi) \to (\psi \wedge \chi)$
- $(\varphi \to \chi) \wedge (\psi \to \chi) \models_F \varphi \vee \psi \to \chi$
- $(\varphi \to \psi) \models_F (\varphi \vee \chi) \to (\psi \vee \chi)$
- $(\varphi \to \psi) \wedge (\varphi \to \chi) \models_F \varphi \to \psi \wedge \chi$.
- $\varphi(x) \models_F \forall x\varphi(x)$.

Field's evaluation respects the Intersubstitutivity principle:

Theorem 13 (Intersubstitutivity of truth for Field's theory (Field))
For every $\varphi, \psi, \chi \in \mathscr{L}_T^{\to}$, if ψ and χ are alike except that one of them has an occurrence of φ where the other has an occurrence of $T\ulcorner\varphi\urcorner$, then $|\psi|_F = |\chi|_F$.

I conclude this outline sketching Field's treatment of determinateness via some examples. Take the liar sentence. There is an intuitive sense in which λ *fails to be true*, but we cannot express it as $\neg T\ulcorner\lambda\urcorner$, since this amounts to λ. Field's way to capture this sense is to say that "λ is not *determinately* true", defining a suitable operator. Let "determinately φ", in symbols $D_F(\varphi)$, be $\varphi \wedge \neg(\varphi \to \neg\varphi)$. By Theorem 13, $D_F(\varphi)$ and $D_F(T\ulcorner\varphi\urcorner)$ are always intersubstitutable. We can now generate paradoxes involving D_F, e.g. the sentence $\lambda^\star$ provably equivalent to $\neg D_F(T\ulcorner\lambda^\star\urcorner)$. We cannot say that $\lambda^\star$ is not determinately true, but we could declare it "not determinately determinately true". Field defines *hierarchies* of iterations of D_F that extend beyond ω_1^{CK}; some of their important features are: (i) they never converge to a single operator that behaves as the limit of the previous ones; (ii) for some stages α, the operator D_F^α behaves badly (e.g. since $|D_F^\alpha(T\ulcorner 0 = 0\urcorner)|_F \neq \mathbf{1}$). Moreover, Welch [43] showed that, "diagonalizing past the determinateness hierarchies" (p. 8), one can construct a sentence λ_{W} s.t. $|\lambda_{\mathrm{W}}|_F \neq \mathbf{1}$ and no sentence "λ_{W} is not determinately$^\sigma$ true" (for any formula σ expressing a level in Field's hierarchies) has value $\mathbf{1}$.

3.2 Some problems for Field's theory

Field's main criticism of Kripke's theory concerns (K1), i.e. that Kripke's theory lacks a conditional that validates schematic laws, i.e. such that "if φ then φ" holds for every sentence φ, together with other principles involving the conditional. So, Field's semantics is designed to validate schemata such as those of Theorem 12.[20] Field conceives his model primarily as an instrument to show that several logical principles (that he considers desirable) are consistent with Intersubstitutivity of truth. As far as I know, he does not advance any particular philosophical interpretation for

[20] Clearly, the schema $\varphi \leftrightarrow \varphi$ plus Intersubstitutivity of truth yields the unrestricted Tarski schema: (TB) $\varphi \leftrightarrow T\ulcorner\varphi\urcorner$, thus addressing problem (K2) of Kripke's theory as well.

his semantics or for his conditional. Nevertheless, thanks to results such as Theorem 12, Field's theory demonstrates that a theory of naïve truth can be very strong and expressive, both from a logical and a truth-theoretic standpoint. Field's theory is also extremely informative, as it shows that some principles previously unknown to be consistent with the Intersubstitutivity principle are in fact consistent with it.

Looking into Field's theory, however, some authors observed that the construction is quite unnatural and complicated.[21] As JC Beall puts it, there is a "fairly fuzzy sense that Field's approach [...] might [...] be more complicated than we need. Regrettably, I do not know how to make the relevant sense of complexity [...] precise enough to serve as an objection."[22] I will make this vague puzzlement into some critical remarks on Field's theory, both as a philosophical story about the conditional and the truth predicate and as a means to establish the consistency of some principles. These remarks will lead to the main question addressed in this paper.

Consider Field's theory as it is meant to be, namely a way to see that some desirable logical principles are consistent with Intersubstitutivity. Let's focus on the logical principles: *why* are they desirable? If one claims some *logical* principles to be desirable, she should explain why this is so, otherwise we may have no reason to want them.[23] The situation is somewhat different for *truth-theoretic* principles. Theorists endorsing naïveté may agree on a small bunch of formulations of naïve truth, while disagreeing fiercely about the logical principles naïve truth should go with. Moreover, a naïve notion of truth may need no particular justification in itself, since it draws much of its appeal from the ordinary use of the word "true"; hardly anything similar can be said for many conditional principles.[24]

A simple answer to the question of why Field's principles are desirable is that they look intuitive in our context, since they are classically valid (and in formal theories of truth, usually, we are not concerned with natural language conditionals). Note, however, that any specific sentence φ does not play any role in our acceptance of a general law such as $\varphi \rightarrow \varphi$. The reason why we accept the law $\varphi \rightarrow \varphi$, if we accept it, must come from the conditional, its only invariant element (similar remarks go for the other principles). So, if $\varphi \rightarrow \varphi$ looks intuitive, there must be something in the understanding of the conditional $\rightarrow$ that makes $\varphi \rightarrow \varphi$ to appear natural. The Fieldian theorist should give an account of the conditional that makes it clear why the principles validated by Field's theory appear as intuitive as they do. This is usually done by the semantics itself: the model-theoretic construction should provide a clear interpretation for the conditional. Unfortunately, this is difficult for Field's theory: here, the main ways of explaining what a conditional connective is are blocked and the model does not provide a new one, as the following analysis shows.

[21] For a debate on this, see for example Martin [32], Welch [42] and Field [19].

[22] Beall [3], Preface, p. viii.

[23] Due to naïve truth, we have to abandon several long-standing logics (classical, intuitionistic, and so on) with their well-established conceptual foundations: this makes particularly urgent the need to justify the rather special and limited set of principles validated by theories of naïve truth.

[24] Various natural language counterexamples to conditional principles commonly accepted in many logics are known. It may be of interest to note that in the literature there is a remarkable gap between theories of conditionals for natural languages and accounts of conditional connectives in pure logic. An analogous gap between theories of truth for natural languages (or even broad views on truth, such as correspondence, coherence, deflationism, and so on) and formal theories of truth exists but is much narrower.

3.3 Is Field's conditional a tool for truth-value comparison?

In the context of many-valued logics, an interesting possibility is to see the conditional as a *tool to compare truth-values*. In this interpretation, a conditional is associated to an ordering of truth-values, and it is supposed to have the designated value if the value of its antecedent is less than or equal to the value of its consequent (in that ordering), and a value that expresses the difference in truth-values between its antecedent and consequent (in that ordering) otherwise. As there can be several ways to "express the difference in truth-values" between the antecedent and the consequent of a conditional, I do not specify this interpretation any further, but this minimal characterization of "a tool to compare truth-values" seems intelligible nonetheless.

Reading the conditional as a tool to compare truth-values is very much in line with a quite standard way of seeing the other quantifiers and connectives of $\mathscr{L}_T^{\rightarrow}$ in many-valued semantics, namely as providing some information about order-theoretic relations concerning the truth-values of their immediate components.

For example, the truth-value of a disjunction provides information about the relations that hold between the truth-values of the two disjuncts and truth-values that are *greater than or equal to* them, in terms of some partial ordering of the truth-values. More specifically, in a setting where the truth-values are linearly ordered (and there is only one designated truth-value), a disjunction takes the truth-value (amongst the truth-values of the disjuncts) that is *closer to* the designated one – with the *proviso* that, if the two disjuncts have the same value, then they are equally close to the designated value, and the disjunction gets the same truth-value as the disjuncts.[25] Alternatively, if the truth-values are only partially ordered, the truth-value of a disjunction may be understood as the least upper bound (if it exists) of the truth-values of the disjuncts in that ordering.[26]

Similar order-theoretic readings can be given for the other connectives and quantifiers of $\mathscr{L}_T^{\rightarrow}$ that are different from $\rightarrow$. In addition, the semantics adopted by Kripke and Field do employ order-theoretic readings of the kind just described for $\neg$, $\wedge$, $\vee$, and $\forall$.[27] So, I take it, order-theoretic readings of connectives and quantifiers are not only common in logic, but also quite relevant in the setting of naïve truth. Interpreting a conditional as a tool to compare truth-values in the above sense, then, would contribute significantly to understanding and expressing the truth-value relations between the sentences of $\mathscr{L}_T^{\rightarrow}$ that also the other connectives and quantifiers aim at conveying.[28]

Now, in defining the values of a conditional at revision stages, Field employs a comparison of revision values, but unfortunately it is impossible to understand the

---

[25] This is the case of Kripke's theory, for example. Depending on the specific setting, this reading can accommodate also multiple designated truth-values: for example, if all the truth-values are linearly ordered and there is one greatest designated value, one can read the disjunction as yielding the truth-value that is closer to it, with the same *proviso* as above. However, I will not consider the case of multiple designated truth-values, as I am only concerned with semantics with one designated truth-value in the present work.

[26] The truth-ordering in the four-valued lattice for First Degree Entailment is an example of such a case (see Belnap [4], pp. 13-16). Thanks to an anonymous referee for some useful remarks on this point.

[27] For Kripke's theory this is quite obvious; for Field's theory, see [18], pp. 259-260.

[28] Seeing connectives and quantifiers as providing order-theoretic information about truth-values lends itself quite well to some specific interpretations of the truth-values themselves, e.g. degrees of credence.

conditional as comparing final values. We have already seen that ultimate values will not do for this purpose.[29] We must turn to the fine-grained version of Field's theory – I will consider this version from now on, unless otherwise specified.[30] In this semantics, as we have seen, the conditional is algebraically interpreted by the partial order $\preceq$, defined by statement (1). Leitgeb [30] shows that, even if a sentence $\varphi \rightarrow \psi$ gets value **1** in Field's theory if and only if the value of φ is less than or equal to the value of ψ in Field's partial order $\preceq$, it is not the case that $\varphi \rightarrow \psi$ receives value **0** if and only if the value of φ is not less than nor equal to the value of ψ, in the same partial order. However, this sense of truth-value comparison is too much to ask if we have more than 2 truth-values: as Leitgeb notices, in the case of the Curry sentence the value of Tt_κ is not less than nor equal to the value of $0 \neq 0$ (in $\preceq$) but it would be disastrous if the whole sentence had value **0**.

Nevertheless, having only conditions to determine when a conditional has truth-value **1** is not satisfactory. In order to interpret Field's conditional as a tool to compare truth-values, we must be able to say something about the truth-value of a conditional $\varphi \rightarrow \psi$ in terms of the relation between the values of φ and ψ also when the whole conditional does not have truth-value **1**. To see whether this is possible we must look into Field's value space **F** (see Definition 10). Field identifies some truth-value "zones" within **F**: let's say that $|\varphi|_F \in \mathbf{H}$ (**L**, **E**) if the revision values of φ oscillate between 1/2 and 1 (0 and 1/2; 1, 1/2 and 0 respectively). To obtain a nice picture of the value space **F**, one should consider such truth-value zones together with the elements **1**, **0**, and **1/2** introduced after Definition 10, namely the revision functions that are constant on revision values 1, 0, and 1/2 respectively. All the possible $\rightarrow$-combinations of sentences receiving the truth-values just described yield the following table:

Table 1: Truth-values and value zones comparison table for Field's $\rightarrow$

$\rightarrow$	**0**	**L**	**1/2**	**H**	**1**	**E**
0	**1**	**1**	**1**	**1**	**1**	**1**
L	**E**	**E; 1**	**1**	**1**	**1**	**E; 1**
1/2	**0**	**E**	**1**	**1**	**1**	**E**
H	**0**	**E**	**E**	**E; 1**	**1**	**E**
1	**0**	**0**	**0**	**E**	**1**	**E**
E	**E**	**E**	**E**	**E; 1**	**1**	**E; 1**

"**E**; **1**" means that either $|\varphi \rightarrow \psi|_F \in \mathbf{E}$ or $|\varphi \rightarrow \psi|_F = \mathbf{1}$. The value zones $\mathbf{H}, \mathbf{L}, \mathbf{E}$ hide a great variety of sequences, that are $\rightarrow$-combined with other sequences in very

[29] See p. 9, the comment after Theorem 9. A similar problem affects the theory in Field [20].

[30] See Subsection 3.1, from Definition 10 onwards. See also Field [18], Ch. 17.

different ways. The table shows that zones **H** and **L** are "between" values **0** and **1**:

$$\mathbf{0} \prec \text{ any value in } \mathbf{L} \prec \mathbf{1/2} \prec \text{ any value in } \mathbf{H} \prec \mathbf{1}. \tag{2}$$

Unfortunately, this exhausts what we can say about truth-value comparison via $\rightarrow$ in Field's value space. As Field notices, sentences in **E** do not fit in the ordering (2). The following remarks show why we cannot read Field's $\rightarrow$ as comparing truth-values.

(a) Conditionals in **E** express failure of comparability between the values of antecedent and consequent. $|\varphi \rightarrow \psi|_F \in \mathbf{E}$ if it is not the case that the revision value of φ is always less than or equal to the one of ψ and it is not the case that the revision value of φ is always greater than the one of ψ (as seen, the ordering associated to revision values is just the usual numerical ordering of numbers 1, 0, 1/2). Moreover, there are $\chi_1, \chi_2 \in \mathscr{L}_T^{\rightarrow}$ s.t. $|\chi_1 \rightarrow \chi_2|_F \in \mathbf{E}$ and $|\chi_2 \rightarrow \chi_1|_F \in \mathbf{E}$. It is problematic to extract information about the truth-value relations of such sentences. Field says that the values in **E** are "incomparable with **1/2**", since if $|\varphi|_F \in \mathbf{E}$ and $|\psi| = \mathbf{1/2}$, then both $|\varphi \rightarrow \psi|_F$ and $|\psi \rightarrow \varphi|_F \in \mathbf{E}$.[31] By the same criterion, also the above χ_1 and χ_2 and all the similar cases are incomparable between them. A common way to represent an order between the values of two sentences is assigning values in a numerical domain D and considering the ordering $\leq_D$ associated to that numerical domain, then checking whether the first value is less than or equal to the second one, or whether it is greater than the second one instead. This association is impossible here. Suppose that we associate i to χ_1 and j to χ_2, for i, j in a numerical domain D: then, either $i \not\leq_D j$ and $i \not>_D j$ (which seems absurd) since $|\chi_1 \rightarrow \chi_2|_F \neq \mathbf{1}$ and $|\chi_2 \rightarrow \chi_1|_F \neq \mathbf{1}$, or we accept that $\rightarrow$ doesn't express this numerical comparison, for the same reason. Claiming that sentences in **E** have no value is not an option: Field's theory is *total*, i.e. it gives a value to every sentence of $\mathscr{L}_T^{\rightarrow}$, and this is essential to recover the schemata that Field wants. As Field remarks, values in **E** "play a crucial role in the theory. [...] any conditional that doesn't have value **1** or **0** must clearly have a [value in **E**]". Every paradoxical conditional is in **E**, e.g. κ: its value oscillates forever, despite the fact that in a simple setting with at least 3 truth-values we can consistently give value 1/2 to Tt_κ and 0 to $0 \neq 0$.[32] Curry's sentence is absolutely incomparable with its negation, as $|\kappa \rightarrow \neg\kappa|_F \in \mathbf{E}$ and $|\neg\kappa \rightarrow \kappa|_F \in \mathbf{E}$, in contrast to the simple numerical treatment of the liar: $|\lambda|_F = \mathbf{1/2} = |\neg\lambda|_F$ and $|\lambda \leftrightarrow \neg\lambda|_F = \mathbf{1}$.[33]

(b) Clearly, some conditionals between two sentences mapped to **E** have value **1**. However, not only it is not possible to express the conditions to assign value **1** or a value in **E** to such conditionals via numerical values (as we have just seen), we cannot even give such conditions in terms of positions within the space **F**. In

---

[31] Field [18], p. 261.

[32] See Subsection 6.1.

[33] This can be remedied in a variant of Field's theory where κ is treated as λ (see [18], p. 271). However, if one adopts naïveté and argues that λ and κ should be treated in the same way, then she presumably accepts that negating a sentence is equivalent to saying that from that sentence a falsity follows via the conditional, as T is treated naïvely. But this general fact does not hold in Field's modified theory either: for some $\varphi \in \mathscr{L}_T^{\rightarrow}$, the sentences $\varphi \rightarrow 0 \neq 0$ and $\neg\varphi$ get a different value in the modified construction. Moreover, the value space of the modified construction is still **F**, so the more general considerations involving the ordering relations in **F** carry over to the modified theory as well.

fact, for every $\mathfrak{e}_1 \in \mathbf{E}$ there are $\mathfrak{e}_2, \mathfrak{e}_3 \in \mathbf{E}$ such that $\mathfrak{e}_1 \preceq \mathfrak{e}_2$ but $\mathfrak{e}_1 \npreceq \mathfrak{e}_3$ and $\mathfrak{e}_1 \nsucc \mathfrak{e}_3$. Conditionals between sentences having values such as $\mathfrak{e}_1$ and $\mathfrak{e}_2$ get value **1** in at least one direction, while conditionals between sentences having values such as $\mathfrak{e}_1$ and $\mathfrak{e}_3$ get a value in **E** in both directions. It is easy to see that for every φ s.t. $|\varphi|_F \in \mathbf{E}$, there are sentences ψ and χ satisfying the conditions of functions $\mathfrak{e}_1$, $\mathfrak{e}_2$, and $\mathfrak{e}_3$ above. Since every function in **E** is comparable in one direction with at least one function and incomparable with another function, we cannot distinguish sub-spaces of **E** of functions that are always comparable with any other function. The same points holds for **L** and **H**, as Table 1 shows.

(c) If one maintains that $\rightarrow$ performs a truth-value comparison, she should explain why **1/2** and the zones **H** and **L** are unreachable by any $\rightarrow$-comparison (no conditional has such truth-values). This seems strange if we consider that, thanks to (2), sentences valued **1/2** or in **H** or **L** can at least be compared "upwards" between them. This seems a mere technical drawback, without any conceptual support.

(d) Analogously, as we have seen, sentences having truth-value **1/2** and sentences having truth-values in **E** are absolutely incomparable via $\rightarrow$: how can we interpret **1/2** and **E** to account for their radical unrelatedness?

(e) Even when the semantics identifies some positive information from the negative fact that $|\varphi|_F \npreceq |\psi|_F$, i.e. when $|\varphi|_F \succ |\psi|_F$, very little can be said in terms of truth-value comparison via $\rightarrow$. Let $|\varphi|_F = \mathbf{1}$, $|\psi_1|_F = \mathbf{0}$ and let either $|\psi_2|_F \in \mathbf{L}$ or $|\psi_2|_F = \mathbf{1/2}$. The difference in value between φ and ψ_1 is strictly greater than the one between φ and ψ_2, according to Field's ordering (2). Still, Field's semantics cannot *see* the difference: $|\varphi \rightarrow \psi_1|_F = |\varphi \rightarrow \psi_2|_F = \mathbf{0}$. As a consequence, Field's semantics cannot *express* this difference either: consider the sentence

$$(\varphi \rightarrow \psi_2) \rightarrow (\varphi \rightarrow \psi_1). \tag{3}$$

If $\rightarrow$ compares truth-values, we should expect (3) to express that the value difference between φ and ψ_1 is strictly greater than the one between φ and ψ_2 by giving to the entire sentence a value less than **1** in Field's ordering $\preceq$. However, this is not the case: $|(\varphi \rightarrow \psi_2) \rightarrow (\varphi \rightarrow \psi_1)|_F = \mathbf{1}$. There are more limitations of this kind: by (2), the value **1** is strictly greater than the values in **H** (in Field's ordering $\preceq$), but the conditional cannot express this fact, as if $|\varphi|_F = \mathbf{1}$ and $|\psi|_F \in \mathbf{H}$, then $|\varphi \rightarrow \psi|_F \in \mathbf{E}$. This is a consequence of Field's technical apparatus, but it is unjustified in the light of the ordering (2) and of the fact that a conditional whose antecedent has value **1** and whose consequent has a value strictly less than **1** but not in **H** can at least be given a value (i.e. **0**) that indicates that the value of the antecedent is greater than the value of the consequent (in Field's ordering $\preceq$), albeit with no gradation. This problem is not limited to **1** and **H**, but presents itself for every conditional whose antecedent has a value in the ordering (2) and whose consequent has the value or is in the value zone immediately preceding the value of the antecedent in the ordering (2) (see Table 1).

Finally, note that looking for a total ordering of the value space **F** is not going to help with the above difficulties. The previous discussion shows that Field's semantics does not validate the so-called "Dummett's law", namely $(\varphi \rightarrow \psi) \vee (\psi \rightarrow \varphi)$. Therefore, any total ordering $\preceq_{tot}$ of **F** would not be an algebraic counterpart of $\rightarrow$, as we would have that, for every $\mathfrak{f}, \mathfrak{g} \in \mathbf{F}$, either $\mathfrak{f} \preceq_{tot} \mathfrak{g}$ or $\mathfrak{g} \preceq_{tot} \mathfrak{f}$.

3.4 Interpreting Field's conditional via its introduction and elimination clauses?

An interesting aspect of the meaning of a connective is given by its *introduction and elimination rules*.[34] In the present context, I mean *semantic* rules, i.e. rules given in terms of some semantic consequence relation (e.g. $\models_F$ as used in the second list of Theorem 12), since several semantic theories of truth lack an axiomatic counterpart.

Anyone who attaches some importance to such rules will be unhappy with Field's conditional. Although *modus ponens* is validated, as $\varphi, \varphi \rightarrow \psi \models_F \psi$ holds for every $\varphi, \psi \in \mathscr{L}_T^{\rightarrow}$, conditional-introduction is quite unsatisfactory. Now, in the context of naïve truth, if one has *modus ponens*, she cannot consistently keep the classical conditional-introduction, since it is inconsistent with *modus ponens* and Intersubstitutivity of truth, as Curry's paradox shows.[35] Still, it seems interesting to determine the conditions under which a conditional may be introduced.

Field notes that we can introduce a conditional if LEM holds for its antecedent:[36]

$$\text{from } \Gamma, \varphi \models_F \psi \text{ and } \Gamma \models_F \varphi \vee \neg\varphi, \text{ infer } \Gamma \models_F \varphi \rightarrow \psi. \tag{4}$$

But this is not very interesting. (4) is not specific to Field's theory, as an identical clause holds for the introduction of $\supset$ in Kripke's models.[37] According to the condition in (4), Kripke's and Field's theories can introduce conditionals in the same circumstances. But they can also eliminate a conditional in the same circumstances, as *modus ponens* for $\supset$ holds in Kripke's models. So, if the introduction and elimination rules play a significant role in understanding one's conditional, Field's theory affords us no conceptual advantage over Kripke's theory. Moreover, (4) fails to cover some cases of conditionals that should be introduced according to Field's theory, e.g.:

$$\lambda \models_F \neg\lambda \text{ but } \not\models_F \lambda \vee \neg\lambda; \text{ still } \models_F \lambda \rightarrow \neg\lambda.$$

We would like to have a weakening of classical conditional-introduction that, unlike (4), is interesting and as strong as possible. A proposal to realize this intuition is to aim at a rule of the following kind (for some logical consequence $\models^*$):

$$\Gamma, \varphi \models^* \psi \text{ and } \Gamma \models^* \mathcal{C}(\varphi, \psi) \quad \text{if and only if} \quad \Gamma \models^* \varphi \rightarrow \psi,$$

where $\mathcal{C}(\varphi, \psi)$ is a $\mathscr{L}_T^{\rightarrow}$-sentence indicating a condition on φ, ψ, or both. To make such a rule interesting, we should require that $\mathcal{C}(\varphi, \psi)$ does not include conditions that are too strong and imply $\Gamma \models^* \varphi \rightarrow \psi$ by themselves. Naïve truth makes it a bit cumbersome to formulate this requirement.

[34] Famously, some authors claim that such rules give the meaning of connectives and quantifiers. This kind of view has been developed by authors such as Gentzen, Dummett, Prawitz, Tennant, Read.

[35] See Field [18], pp. 281-283. Let me emphasize that I do not consider substructural notions of logical consequence, nor theories that accept inconsistencies.

[36] Field [18], p. 269. My formulation is slightly different from Field's. See also Field [15], pp. 152-153.

[37] Also, a proof-theoretic version of this rule holds in the theory PKF formulated in natural deduction. For PKF, see Halbach and Horsten [24], for its natural deduction version see Horsten [26], p. 188.

Definition 14 (Genuine conditions for conditional-introduction)
I start with a preliminary notion. Recall that every sentence is a subsentence of itself. Define the following sets by transfinite recursion, for Δ a set of $\mathscr{L}_T^{\rightarrow}$-sentences.

$$\varphi \in \mathcal{S}(\Delta) \text{ iff } \begin{cases} \varphi \text{ is a subsentence of some } \psi \in \Delta, \text{ or} \\ \varphi \text{ results from a subsentence of some } \psi \in \Delta \text{ replacing} \\ \quad \text{one occurrence of a subsentence } T\ulcorner\chi\urcorner \text{ of } \psi \text{ with } \chi \end{cases}$$

$$\mathcal{S}^{\alpha}(\Delta) := \mathcal{S}(\bigcup_{\beta<\alpha} \mathcal{S}^{\beta}(\Delta)); \qquad \mathcal{S}^{\infty}(\Delta) := \bigcup_{\alpha \in Ord} \mathcal{S}^{\alpha}(\Delta).$$

By a small notational abuse, if $\Delta = \{\varphi\}$, I will write $\mathcal{S}^{\infty}(\varphi)$. Let's say that a schematic $\mathscr{L}_T^{\rightarrow}$-sentence $\mathcal{C}(\varphi,\psi)$ is a *genuine condition for conditional-introduction* for the logical consequence $\models^*$ if, for all $\varphi,\psi \in \mathscr{L}_T^{\rightarrow}$ and $\Gamma \subseteq SENT_{\mathscr{L}_T^{\rightarrow}}$ the following holds:

$$\Gamma,\varphi \models^* \psi \text{ and } \Gamma \models^* \mathcal{C}(\varphi,\psi) \quad \text{if and only if} \quad \Gamma \models^* \varphi \rightarrow \psi, \qquad \text{(G-C-Intro)}$$

where:

1. $\varphi \in \mathcal{S}^{\infty}(\mathcal{C}(\varphi,\psi))$, or $\psi \in \mathcal{S}^{\infty}(\mathcal{C}(\varphi,\psi))$, or both, and
2. there is no $\chi \in \mathcal{S}^{\infty}(\mathcal{C}(\varphi,\psi))$ s.t. for every $\varphi_0,\psi_0 \in \mathscr{L}_T^{\rightarrow}$:

$$\Gamma \models^* \varphi_0 \rightarrow \psi_0 \text{ if and only if } \Gamma \models^* \chi_0$$

where χ_0 obtains from $\mathcal{S}^{\infty}(\mathcal{C}(\varphi,\psi))$ instead of χ, putting φ_0 and ψ_0 for φ and ψ.

In short, $\mathcal{C}(\varphi,\psi)$ is a genuine condition for conditional-introduction if it is a condition on φ, ψ, or both (item 1), and neither $\mathcal{C}(\varphi,\psi)$ nor any of its proper subschemata (possibly buried into instances of T) are equivalent to the $\models^*$-validity of all the corresponding conditionals, making the condition $\mathcal{C}(\varphi,\psi)$ uninteresting (item 2).

This Definition filters out undesirable rules, e.g. (for Field's $\models_F$):

$$\Gamma,\varphi \models_F \psi \text{ and } \Gamma \models_F (\varphi \vee \neg\varphi) \vee T\ulcorner\psi \vee (\varphi \rightarrow \psi)\urcorner \text{ if and only if } \Gamma \models_F \varphi \rightarrow \psi.$$

This rule is clearly uninformative, since it tells us that, whenever $\varphi \models_F \psi$, we can introduce the corresponding conditional if φ has a classical value, or ψ has value **1**, or ... the whole conditional has value **1**! This example makes it also clear why we need to consider subschemata, possibly buried into the scope of instances of T.

A genuine condition for conditional-introduction seems fundamental to understand the relation between what Field calls "conditional assertion" ($\varphi \models \psi$) and the "assertion of a conditional" ($\models \varphi \rightarrow \psi$). Since conditional assertions are weaker than assertions of a conditional (the former follow always from the latter, but not *vice versa*), a genuine condition for conditional-introduction is central to know in what the difference between these two notions consists, and when we can turn a reasoning under assumption into a conditional expressible in the object-language $\mathscr{L}_T^{\rightarrow}$.

However, it seems difficult to find a genuine condition for conditional-introduction for Field's theory. To my knowledge, Field never claims that a rule stronger than (4) is available. Moreover, the fact that there is no sub-space of always comparable truth-values within **E** (item (b) of Subsection 3.3) seems to doom the prospect to find a genuine condition. This has serious philosophical consequences: if no genuine condition for conditional-introduction can be found for Field's theory, the nature of the conceptually crucial relation between conditional assertion and assertion of a conditional in Field's semantics is bound to remain obscure.

3.5 Axiomatizing Field's conditional?

Even if in general the truth-value of $\varphi \to \psi$ is not the result of a comparison of the truth-values of the components, for each φ and ψ whose logical forms are related in certain ways such comparison is always possible. For example, if ψ is φ, their revision sequences are identical and sentences of the form $\varphi \to \varphi$ have always value **1**. Similar considerations hold for $\varphi \wedge \psi \to \varphi$, $\varphi \to \varphi \vee \psi$, and all the laws that Field's semantics validates. This suggests that we could adopt the validity of some schemata as a key to understand Field's conditional: in other words, one could aim for an *axiomatization* of Field's construction. It seems, however, that this is also a difficult path for the Fieldian theorist. Since obvious complexity reasons show that there is no complete axiomatization of any theory including the truth-set of arithmetic (this is the case for virtually every semantic theory of truth), a widely adopted criterion to establish that a theory axiomatizes a semantic theory of truth is *adequacy*:

Definition 15 (Adequacy (Fischer, Halbach, Speck, Stern [21])[38])
Let Th be a recursively enumerable (r.e.) set of sentences of $\mathscr{L}_T^{\to}$ and $\mathscr{A}$ be a class of models for $\mathscr{L}_T^{\to}$. Th is an adequate axiomatization of $\mathscr{A}$ if and only if for all $S \subset \omega$:

$$(\mathbb{N}, S) \models \mathsf{Th} \ \text{ if and only if } \ (\mathbb{N}, S) \in \mathscr{A}.$$

Field's theory does not admit an adequate first-order axiomatization. The crucial reason is the high computational complexity of the set of sentences having value **1** in Field's theory (henceforth: "Field's truth-set"), that was determined by Welch [41].

Proposition 16
There is no first-order, r.e. theory that axiomatizes adequately Field's truth-set.

Proof (Sketch)
Welch [41] proves that the computational complexity of Field's truth-set exceeds the complexity of Π^1_1-complete subsets of ω. Reasoning as in Fischer, Halbach, Speck, Stern [21], suppose that there is a first-order, r.e. theory F s.t.:

$$(\mathbb{N}, P) \models \mathsf{F} \ \text{ if and only if } \ P \text{ is Field's truth-set.}$$

So, we would have a Δ^1_1 definition of Field's truth-set, which is absurd. □

One could hope to syntactically characterize Field's construction by using infinitary theories. Unfortunately, Field's semantics is too complex even for ω-logics.

Definition 17 (ω_{L}-logic (see McGee [34], p. 150))
Let L be a recursively enumerable set of first-order principles that consists of some logical principles and the two following rules for truth:

$$(T\text{-Intro}) \ \frac{\varphi}{T\ulcorner\varphi\urcorner} \qquad\qquad (T\text{-Elim}) \ \frac{T\ulcorner\varphi\urcorner}{\varphi}$$

such that $\mathsf{PA}_T^{\to}$ formulated in the logic L plus (T-Intro) and (T-Elim) is consistent (at

[38] Fischer, Halbach, Speck, Stern [21] contains a detailed analysis of adequacy criteria. As they show, the criterion adopted here has a preeminent role between other proposed characterizations of adequacy.

least one such L exists).[39] Call the resulting theory $\mathsf{L}(\mathsf{PA}_T^{\rightarrow})$. Fix a non-repeating enumeration of all the closed terms of the language, which I indicate informally with $t_1, t_2, \ldots$. Consider the following rule:

$$\frac{\varphi(t_0) \qquad \varphi(t_1) \qquad \ldots \qquad \varphi(t_n) \qquad \ldots}{\forall n \varphi(t_n)} \ (\omega\text{-Rule})$$

An ω_{L}-*derivation* is a well-ordered sequence of $\mathscr{L}_T^{\rightarrow}$-sentences $\langle \psi_1, \ldots, \psi_\alpha, \ldots \rangle$ s.t. each element ψ_β of the sequence is:

- an atomic or negated atomic arithmetical sentence true in $\mathbb{N}$, or
- an instance of a logical schema of L, or
- obtained from some elements ψ_γ of the sequence, for some $\gamma < \beta$, via the rules of L, or the rule (T-Intro), or the rule (T-Elim), or the (ω-Rule).

The set $\{\varphi \in \mathscr{L}_T^{\rightarrow} \mid$ there is a ω_{L}-derivation of $\varphi\}$ will be called ω_{L}-*logic*.

Proposition 18
No ω_{L}-logic yields Field's truth-set.

Proof (Sketch)
Let L be s.t. the resulting ω_{L}-logic is consistent, call such theory F. There is a positive inductive definition whose least fixed point is identical to F (folklore; see, e.g., McGee [34]). So, the computational complexity of F is $\leq \Pi^1_1$, and thus it cannot be Field's truth-set, by the previously mentioned result of Welch [41]. □

Some consider it to be a nice feature of a semantic construction that it can be captured by a suitable infinitary theory: the least Kripke fixed point is such a construction, as it can be defined by a simple ω_{L}-logic.[40] This is not the case for Field's theory. Propositions 16 and 18 show how the high complexity of Field's theory that many were critical of has a significant impact on the possibility to characterize, and thus understand, the conditional connective it features.

3.6 Primitivism about logical principles, and the search for a new theory

The problems raised in Subsections 3.3-3.5 are worsened by the fact that, as mentioned above, Field never specifies a philosophical interpretation for his conditional, and his theory does not suggest one such interpretation itself.

To address these difficulties, the Fieldian theorist could adopt a minimalist reply, namely something along the following lines: "it is OK to specify no interpretation for the conditional since, independently of that, several principles considered to be worth having can be consistently validated". I am not sure of how this reply can work. If the reply means something like "adopt any reading you like, still certain principles are recovered, and that's all that matters", we have a problem with the "adopt any reading you like" part, since many common and interesting understandings of $\rightarrow$ are

[39] One surely wants these rules, at the very least, if she is to attempt a characterization of Field's theory.
[40] See Martin [32].

barred by Field's theory, or not supported at best. Then, we cannot adopt them, and the Fieldian theorist is challenged to come up with a new reading. If, on the other hand, the reply means something like "it is OK to adopt no interpretation for the conditional *per se*, still certain principles are recovered, and that's all that matters", things are different. Here the Fieldian theorist expresses a sort of *primitivism* about conditional principles: they embody irreducible aspects of the conditional and should be accepted, period. In the remainder of this Subsection, I will refer to "principles" meaning "classically valid conditional principles", unless otherwise specified.

There are two main ways to interpret the primitivist position I just described. One possibility is to argue that there is one specific set of principles that must be recovered. Let's call this position *specific primitivism*. In absence of a philosophically significant interpretation for the conditional, specific primitivism has unpleasant consequences. Nearly every theorist agrees that a theory of naïve truth should provide a classical interpretation of the truth-free part of the language. It is from *classical* semantics that naïve truth forces a deviation – hence the effort to recover as many classical principles as possible, or to approximate them at best. Suppose that a specific primitivist advocates the principles recovered by Field's theory, call them $\mathcal{P}rinc$: what if another construction is found that does not validate all the principles in $\mathcal{P}rinc$ but validates those in another set $\mathcal{P}rinc^{\star}$, which has many principles in common with $\mathcal{P}rinc$, and whose remaining ones look equally natural and fundamental? As Field's theory does not support the three major options for the meaning of the conditional considered in Subsections 3.3-3.5, and it does not provide a new interpretation for it, our theorist accepts some principles while being unable to use the meaning of the conditional to justify their acceptance. On what grounds, then, is our specific primitivist going to choose between $\mathcal{P}rinc$ and $\mathcal{P}rinc^{\star}$, or to argue for $\mathcal{P}rinc$ over $\mathcal{P}rinc^{\star}$? Choosing $\mathcal{P}rinc$ over $\mathcal{P}rinc^{\star}$ for no reason is clearly unacceptable, and especially so in the area of theories of naïve truth, as recalled in Subsection 3.2. Moreover, since we work in a language such as $\mathscr{L}_T^{\rightarrow}$ and we aim at recovering classical principles, many kinds of considerations to choose some principles over others are ruled out, e.g. those involving appeals to natural languages.

Alternatively, one could abandon specific primitivism for a weaker form, quantifying existentially over the principles to be recovered: "it is OK to adopt no interpretation for the conditional *per se*, still there exists a collection of principles that are recovered, and that's all that matters". Let's call this view *generic primitivism*. I don't have an argument against this position in general, also because it seems to be a fundamental view about what we should aim for in devising our theories of truth and it is difficult, and perhaps pointless, to argue in favor or against such fundamental views. Two such views can be dialectically articulated here, diverging on whether we want the addition of naïve truth to allow us to recover some principles or to preserve a certain understanding of the elements of our language.[41] Generic primitivism, how-

[41] A position not far from the latter view is expressed in Martin [32] "[we don't know] what theory [Field's] construction yields a [consistency] proof for." (p. 343); "[...] I don't see how [Field's conditional] is *a* generalization; that is, I don't see what generalization it is supposed to be. In the end, all we are given is the model-theoretic construction and a (necessarily very partial) list of the laws and nonlaws. Contrast this with the connectives in the Kripke case. I would go so far as saying that Kripke's disjunction and conjunction *are* the classical disjunction and conjunction." (*ibid.*, p. 345).

ever, seems very difficult to defend when, as in Field's theory, we cannot combine it with a philosophically significant reading of the conditional, as this argument shows.

- Generic primitivism, by itself, only indicates to validate *at least one* set of principles, no matter *which*. This position is tremendously weak: a theory validating only the uninteresting law $(\varphi \wedge \psi) \rightarrow [(\varphi \rightarrow \chi_0) \rightarrow ((\chi_0 \rightarrow \chi_1) \rightarrow \chi_1)]$ satisfies the *desideratum* of generic primitivism. Moreover, a theory validating only the law $(\varphi \wedge \psi) \rightarrow [(\varphi \rightarrow \chi_0) \rightarrow ((\chi_0 \rightarrow \chi_1) \rightarrow \chi_1)]$ meets the requirements of generic primitivism just as a theory validating $\varphi \rightarrow \varphi$ and other apparently more interesting schemata. So, not only generic primitivism alone gives us no reasons to decide between two theories validating different sets of principles (as in the case $\mathcal{P}rinc$ vs. $\mathcal{P}rinc^\star$ from above), but it actually supports the conclusion that $\varphi \rightarrow \varphi$ is not more important than $(\varphi \wedge \psi) \rightarrow [(\varphi \rightarrow \chi_0) \rightarrow ((\chi_0 \rightarrow \chi_1) \rightarrow \chi_1)]$ and that in general no principle is more fundamental than another.
- These undesirable consequences follow from generic primitivism *alone*: a philosophically significant reading for the conditional would give us a criterion to rule them out. Crucially, however, the Fieldian theorist cannot make this move.
- Since the consequences of generic primitivism I just indicated are unacceptable, one must find a way to single out some principles as more important than others. Well, which are those principles and why should one choose exactly them? It is easy to see that we are drifting again toward the search for a kernel of "really fundamental" principles, and we will encounter again problems of the kind $\mathcal{P}rinc$ vs. $\mathcal{P}rinc^\star$ mentioned above: in the absence of a philosophically significant reading for the conditional, a generic primitivist has either to accept the bad consequences mentioned above, or to opt for an equally untenable specific primitivism.

A primitivist about principles seems to have no alternatives beside specific and generic primitivism: either she aims at recovering a specific set of principles, or an unspecified one. As both positions have intolerable consequences when no philosophically significant reading of the conditional is available, neither is an acceptable position for the Fieldian theorist trying to avoid the problems highlighted previously.

Notice how focusing on the preservation of the meaning of connectives and quantifiers is a more neutral common ground: people disagreeing on which are the correct/best logical principles may agree on a more minimal notion of meaning for connectives and quantifiers. Moreover, the presence of different algebraic interpretations of the conditional does not generate a problem of the kind $\mathcal{P}rinc$ vs. $\mathcal{P}rinc^\star$, i.e. of which meaning is more fundamental: we seem to have an intuition according to which $\varphi \rightarrow \varphi$ is more fundamental than $(\varphi \wedge \psi) \rightarrow [(\varphi \rightarrow \chi_0) \rightarrow ((\chi_0 \rightarrow \chi_1) \rightarrow \chi_1)]$, but it seems unproblematic to say that no algebraic interpretation of the conditional is more fundamental than another – they are just different.

In the light of the above observations, a quite natural question comes forward.

The Main Question Can we address Field's program of adding a conditional to Kripke's theory with a *philosophically significant and useful conditional*, namely a conditional that is interesting for the truth-theorist, simple to characterize and to interpret, and that plays a well-defined role in the theory of truth?

4 A new construction

Now I present a theory devised to address the **Main Question**. The new construction generalizes Kripke's theory, retains its nice features, and allows us to formalize and validate conditional sentences that we would like to have in order to overcome the expressive weaknesses of Kripke's approach. The conditional modeled by this construction has a simple and appealing reading, and it enjoys nice semantic introduction/elimination clauses.

As we have seen in Section 2, Kripke's theory suffers from its expressive difficulties also because it is silent about C-gaps. If we had another notion of gappiness, that befits the framework of inductive constructions, we could expand Kripke's theory to include such notion, treating it on a par with Kripke's truth-values 1 and 0. In the construction I give here, an explicit treatment of gappiness is pursued and gaps are not defined via the complement of Kripke fixed points, but *positively*, as a (partly) independent notion, corresponding to the extra truth-value $1/2$. These gaps will be called *P-gaps*. A motivation for exploring P-gaps is that Kripke's trichotomy (truth-sets, falsity-sets, C-gaps) may seem too coarse-grained: sentences that seem very different are conflated together in C-gaps, and one may want to isolate those that can be evaluated in an inductive construction.[42] Once we have positive gaps together with Kripke's truth-values 1 and 0, we use them to interpret a stronger conditional, adopting the Łukasiewicz 3-valued logic (henceforth Ł3) evaluation schema rather than K3.[43] The reason for this choice is clear: Ł3 incorporates K3 and features a conditional $\rightarrow_{Ł3}$ that is "stronger" than the K3 material conditional, since a sentence $\varphi \rightarrow_{Ł3} \psi$ receives value 1 also when φ and ψ have value $1/2$.

Some caution is necessary here, as Łukasiewicz logics are problematic for naïve truth: every finitely valued Łukasiewicz logic is inconsistent with Intersubstitutivity of truth, and the continuum-valued one is ω-inconsistent with it.[44] These results are often used to defend a stronger claim, namely that Łukasiewicz logics (Ł3 in particular) are not compatible with Kripke's framework. McGee puts this idea thus: "[a]n historically important example of a method for handling truth-value gaps which is not amenable to Kripke's techniques is the 3-valued logic of Łukasiewicz".[45] In the construction developed here, I will show that this stronger claim is mistaken: there are ways to handle "gaps" using Ł3 that are perfectly "amenable to Kripke's techniques".

The first task is to turn the lack of characterization of "gappy" sentences we have in Kripke's theory into some positive information. I will now provide a suitable monotone construction that partly does the job: it does not *build* positive gaps, but it *makes room* for a third value $1/2$, preserving and increasing the sentences that are so evaluated. Just as we can isolate the clauses for the preservation of values 1 and 0 in K3, we can isolate the K3-clauses for value $1/2$. Using the Kripke jump, I give an inductive construction that, taking a set S_3 as hypothesis for sentences valued $1/2$ and two sets S_1 and S_2 for values 1 and 0, increases the former using also the latter.

---

42 E.g. the Curry-like sentences in Field [18], pp. 85-86. I explore this idea further in my [37].

43 For Łukasiewicz logics see Malinowski [31] and Gottwald [22].

44 The proof of the first claim is a generalization of Curry's paradox, see for example Field [18], pp. 85-86. The second claim was proven in Restall [36].

45 McGee [34], p. 87, footnote 1.

Definition 19 (Positive gappiness preservation construction for $\mathscr{L}_T^{\rightarrow}$)
For any three sets $S_1, S_2, S_3 \subseteq \omega$, define the set S_3^*, relative to sets S_1 and S_2, so that $n \in S_3^*$ if $n \in S_3$, or

(i) n is $\neg\varphi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\varphi \in S_3$, or

(ii) n is $\varphi \wedge \psi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\psi \in \mathscr{L}_T^{\rightarrow}$ and $\left\{\begin{array}{l}\varphi \in S_3 \text{ and } \psi \in S_3, \text{ or} \\ \varphi \in S_3 \text{ and } \psi \in \mathfrak{E}_\Phi(S_1,S_2), \text{ or} \\ \varphi \in \mathfrak{E}_\Phi(S_1,S_2) \text{ and } \psi \in S_3\end{array}\right\}$ or

(iii) n is $\varphi \vee \psi$ and $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $\psi \in \mathscr{L}_T^{\rightarrow}$ and $\left\{\begin{array}{l}\varphi \in S_3 \text{ and } \psi \in S_3, \text{ or} \\ \varphi \in S_3 \text{ and } \psi \in \mathfrak{A}_\Phi(S_1,S_2), \text{ or} \\ \varphi \in \mathfrak{A}_\Phi(S_1,S_2) \text{ and } \psi \in S_3\end{array}\right\}$ or

(iv) n is $\forall x \chi(x)$ and $\chi(x) \in FOR_{\mathscr{L}_T^{\rightarrow}}$ and, for at least one $s \in CLTER_{\mathscr{L}_T^{\rightarrow}}$, $\chi(s) \in S_3$ and, for all $t \neq s$, either $\chi(t) \in S_3$ or $\chi(t) \in \mathfrak{E}_\Phi(S_1,S_2)$; or

(v) n is Tt and $dec(t) = \ulcorner\chi\urcorner$ and $\chi \in \mathscr{L}_T^{\rightarrow}$ and $\chi \in S_3$.

This is an inductive definition. Put $\eta(n,S_3,S_1,S_2)$ as the abbreviation of the right-hand side of the above definition. $\eta(n,S_3,S_1,S_2)$ is a positive elementary formula, positive in S_1,S_2,S_3. Associate to $\eta(n,S_3,S_1,S_2)$ a monotone operator $\Psi : \mathcal{P}(\omega) \times \mathcal{P}(\omega) \times \mathcal{P}(\omega) \longmapsto \mathcal{P}(\omega) \times \mathcal{P}(\omega) \times \mathcal{P}(\omega)$ s.t.:

$$\Psi(S_3,S_1,S_2) := \langle\{n \in \omega \,|\, \eta(n,S_3,S_1,S_2)\}, S_1, S_2\rangle.$$

Ψ will be referred to as the *P-gappy jump*. Let $\mathfrak{I}_\Psi(R,P,Q)$ denote the fixed point of Ψ obtained from $\langle R,P,Q\rangle$ and $\mathfrak{H}_\Psi(R,P,Q)$ denote the first member of the triple $\mathfrak{I}_\Psi(R,P,Q)$. By construction, $\mathfrak{I}_\Psi(R,P,Q)$ and $\mathfrak{H}_\Psi(R,P,Q)$ are inductive in P,Q,R.

The operator Ψ builds a Kripke fixed point over $\langle S_1,S_2\rangle$, namely $\mathfrak{I}_\Phi(S_1,S_2)$, then it evaluates P-gappy sentences based on the sentences already in S_3 and $\mathfrak{I}_\Phi(S_1,S_2)$, using the clauses of K3, and it goes on until it reaches a fixed point.[46]

Now that we have room to handle $1/2$-valued sentences, we extend Kripke's construction with clauses for $\rightarrow$ patterned after Ł3.[47]

Definition 20 (Łukasiewicz-Kripke construction for $\mathscr{L}_T^{\rightarrow}$)
For any three sets $S_1,S_2,S_3 \subseteq \omega$, define sets S_1^{T}, S_2^{F}, S_3^{G} that satisfy the following:

1. $n \in S_1^{\mathrm{T}}$ if:
   (i) $n \in \mathfrak{E}_\Phi(S_1,S_2)$, or

   (ii) n is $\varphi \rightarrow \psi$, for φ, $\psi \in \mathscr{L}_T^{\rightarrow}$, and $\left\{\begin{array}{l}\varphi \in \mathfrak{A}_\Phi(S_1,S_2), \text{ or} \\ \psi \in \mathfrak{E}_\Phi(S_1,S_2), \text{ or} \\ \varphi,\psi \in \mathfrak{H}_\Psi(S_3,S_1,S_2).\end{array}\right.$

2. $n \in S_2^{\mathrm{F}}$ if:
   (i) $n \in \mathfrak{A}_\Phi(S_1,S_2)$, or

[46] The starting hypotheses for values 1 and 0 cannot be eliminated. One can give all the conditions under which a sentence gets value 1 (0) in a K3 evaluation using only clauses about value 1 (0) in it (see Halbach [25], pp. 202 ff.) but this does not hold for $1/2$: e.g. not all the conditions under which $\varphi \wedge \psi$ has value $1/2$ in a K3 evaluation can be given using only its subsentences or negated subsentences having value $1/2$.

[47] This answers Martin's question "of what is the new conditional a generalization" for this construction.

(ii) n is $\varphi \to \psi$, for φ, $\psi \in \mathscr{L}_T^{\to}$, and $\varphi \in \mathfrak{E}_\Phi(S_1, S_2)$, $\psi \in \mathfrak{A}_\Phi(S_1, S_2)$.

3. $n \in S_3^{\mathrm{G}}$ if:
   (i) $n \in \mathfrak{H}_\Psi(S_3, S_1, S_2)$, or
   (ii) n is $\varphi \to \psi$, for φ, $\psi \in \mathscr{L}_T^{\to}$, and $\begin{cases} \varphi \in \mathfrak{E}_\Phi(S_1, S_2) \text{ and } \psi \in \mathfrak{H}_\Psi(S_3, S_1, S_2), \text{ or} \\ \varphi \in \mathfrak{H}_\Psi(S_3, S_1, S_2) \text{ and } \psi \in \mathfrak{A}_\Phi(S_1, S_2). \end{cases}$

The above sets are also inductively defined. Put $\theta_1(n, S_1, S_2, S_3)$ ($\theta_2(n, S_1, S_2, S_3)$, $\theta_3(n, S_1, S_2, S_3)$) as the abbreviation of the right-hand side of item 1 (2, 3, respectively) above. All these formulae are positive elementary, positive in S_1, S_2, and S_3. Associate to them, in a standard way, a monotone operator acting on triples of sets. Define $\Upsilon : \mathcal{P}(\omega) \times \mathcal{P}(\omega) \times \mathcal{P}(\omega) \longmapsto \mathcal{P}(\omega) \times \mathcal{P}(\omega) \times \mathcal{P}(\omega)$ as:

$$\begin{aligned} \Upsilon(S_1, S_2, S_3) := \langle & \{n \in \omega \,|\, \theta_1(n, S_1, S_2, S_3)\}, \{n \in \omega \,|\, \theta_2(n, S_1, S_2, S_3)\}, \\ & \{n \in \omega \,|\, \theta_3(n, S_1, S_2, S_3)\} \rangle. \end{aligned}$$

The whole construction will be referred to as the *ŁK-construction*. $\mathfrak{I}_\Upsilon(P, Q, R)$ denotes the fixed point of Υ obtained from $\langle P, Q, R \rangle$. $\mathfrak{I}_\Upsilon(P, Q, R)$ is inductive in P, Q, R.

The jump Υ takes as argument a triple $\langle P, Q, R \rangle$, reading it as the starting hypothesis for a truth-value distribution of values 1, 0, and $1/2$, in this order. The first stage of the construction, namely $\Upsilon(P, Q, R)$, is the following triple:

$$\Bigg\langle \underbrace{\mathfrak{E}_\Phi(P, Q) \cup \{\varphi \to \psi \in \mathscr{L}_T^{\to} \,|\, \begin{cases} \varphi \in \mathfrak{A}_\Phi(P, Q), \text{ or} \\ \psi \in \mathfrak{E}_\Phi(P, Q), \text{ or} \\ \varphi, \psi \in \mathfrak{H}_\Psi(R, P, Q) \end{cases} \}}_{\text{The first member of the triple } \Upsilon(P,Q,R), \text{ i.e. } P^{\mathrm{T}}},$$

$$\underbrace{\mathfrak{A}_\Phi(P, Q) \cup \{\varphi \to \psi \in \mathscr{L}_T^{\to} \,|\, \varphi \in \mathfrak{E}_\Phi(P, Q) \text{ and } \psi \in \mathfrak{A}_\Phi(P, Q)\}}_{\text{The second member of the triple } \Upsilon(P,Q,R), \text{ i.e. } Q^{\mathrm{F}}},$$

$$\underbrace{\mathfrak{H}_\Psi(R, P, Q) \cup \{\varphi \to \psi \in \mathscr{L}_T^{\to} \,|\, \begin{cases} \varphi \in \mathfrak{E}_\Phi(P, Q) \text{ and } \psi \in \mathfrak{H}_\Psi(R, P, Q), \text{ or} \\ \varphi \in \mathfrak{H}_\Psi(R, P, Q) \text{ and } \psi \in \mathfrak{A}_\Phi(P, Q) \end{cases} \}}_{\text{The third member of the triple } \Upsilon(P,Q,R), \text{ i.e. } R^{\mathrm{G}}} \Bigg\rangle$$

Consider P^{T} (Q^{F} and R^{G} are similar). The first step in defining P^{T} is constructing $\mathfrak{E}_\Phi(P, Q)$, the truth-set of the Kripke fixed point built over our starting hypothesis for values 1 and 0 (value $1/2$ plays no role in Kripke fixed points). Clearly $P \subseteq \mathfrak{E}_\Phi(P, Q)$. In addition to $\mathfrak{E}_\Phi(P, Q)$, P^{T} contains the conditionals $\varphi \to \psi$ such that:

- the antecedent (φ) is in the Kripke falsity-set built over our starting distribution of values 1, 0 and $1/2$; or
- the consequent (ψ) is in the Kripke truth-set built over our distribution; or
- both antecedent (φ) and consequent (ψ) are in the set of P-gappy sentences $\mathfrak{H}_\Psi(R, P, Q)$ built over this truth-value distribution.

Conditionals are interpreted "slowly". At first, $\mathfrak{E}_{\Phi}(P,Q)$, $\mathfrak{A}_{\Phi}(P,Q)$, and $\mathfrak{H}_{\Psi}(R,P,Q)$ are built, and they have no clause for the conditional. Then, we evaluate those conditionals having the antecedent, or the consequent, or both in these fixed points – as shown in the triple above, to the right of the symbol "$\cup$" in each line. After that, the process just described is iterated, until it comes to a halt at the fixed point.

Several differences with Field's construction are visible. First, just as Kripke's theory, the ŁK-construction uses only inductive definitions. This shows that McGee's verdict on Ł3 is mistaken: there are ways to handle "gaps" in Ł3 which are "amenable to Kripke's techniques", namely P-gaps. Indeed, only "Kripke's techniques" were used in this theory: as in Kripke's theory, we have an inductive definition (the ŁK-construction) and a monotone jump (Υ). The clauses to interpret T correspond to Kripke's ones: for every set built in the construction, if φ is in that set, so is $T\ulcorner\varphi\urcorner$. Second, unlike the revision sequence(s) used in Field's theory, the ŁK-construction is monotonic: once a sentence φ goes in a member of the triple built by Υ at some stage, φ remains in it up to the fixed point. In Kripke's theory, sentences are evaluated gradually, starting from the initial hypotheses and progressively interpreting more complex sentences.[48] The ŁK-construction shares this simple picture, extending it to its conditional. Field's revision construction does not seem to provide an equally intuitive picture of how sentences are evaluated.

5 Properties of the ŁK-construction

I will now review the main general properties of the ŁK-construction. By a slight notational abuse, I will use $\langle P_\infty, Q_\infty, R_\infty\rangle$ to indicate $\mathfrak{I}_\Upsilon(P,Q,R)$, and I will also use P_∞ (Q_∞, R_∞) alone to denote the 1st (2nd, 3rd) element of $\mathfrak{I}_\Upsilon(P,Q,R)$.[49] A fixed point $\mathfrak{I}_\Upsilon(P,Q,R)$ is *consistent* if $P_\infty, Q_\infty, R_\infty$ are pairwise disjoint (henceforth *pwd*), *inconsistent* otherwise. I will say that a sentence φ has value 1 (0, $1/2$) in a consistent fixed point $\langle P_\infty, Q_\infty, R_\infty\rangle$ if $\varphi \in P_\infty$ (Q_∞, R_∞, resp.), in symbols $|\varphi|_{(PQR)^\infty} = 1$ (0, $1/2$).

Let's start from the observation that the ŁK-construction, as a generalization of Kripke's theory, inherits its partiality: it is not the case that every sentence has a value in every consistent fixed point of Υ. In fact, as in Kripke's case, consistency requires partiality: if a fixed point of Υ is consistent, some sentence has no value in it.

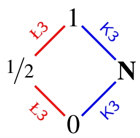
Lemma 21
If $\mathfrak{I}_\Upsilon(P,Q,R)$ is s.t. for every $\varphi \in \mathscr{L}_T^{\rightarrow}$, $\varphi \in (P_\infty \cup Q_\infty \cup R_\infty)$, then $\mathfrak{I}_\Upsilon(P,Q,R)$ is inconsistent.

Proof
Let ι be provably equivalent to $T\ulcorner\iota\urcorner \rightarrow \neg T\ulcorner\iota\urcorner$. If $\iota \in P_\infty$, by the fixed-point property, $T\ulcorner\iota\urcorner \in P_\infty$, so $\neg T\ulcorner\iota\urcorner \in Q_\infty$ and also $T\ulcorner\iota\urcorner \rightarrow \neg T\ulcorner\iota\urcorner \in Q_\infty$, i.e. $\iota \in Q_\infty$, i.e. $\mathfrak{I}_\Upsilon(P,Q,R)$ is inconsistent. The same conclusion follows easily if $\iota \in Q_\infty$ or $\iota \in R_\infty$. □

48 For some epistemic and algorithmic aspects of Kripke's theory, see Cantini [10], p. 69 and following.

49 Writing P_∞ alone is not meaningful, since to know what is in it we must know the other members of the fixed-point triple. However, when writing P_∞ (or Q_∞, or R_∞) alone, the triple will always be clear.

Hartry Field suggested a nice way of presenting the interpretation of the conditional given by the ŁK-construction. Every consistent fixed point $\mathfrak{I}_{\Upsilon}(P,Q,R)$ yields partial, three-valued evaluations for sentences, but since some sentences receive no value, we could add explicitly a fourth value, call it $\mathbf{N}$, indicating a lack of value in $\mathfrak{I}_{\Upsilon}(P,Q,R)$. $\mathbf{N}$ represents the analogue of C-gaps for fixed points of Υ, and it behaves as the intermediate value of K3. The conditional modeled by the ŁK-construction thus shows that it is possible to keep Ł3 and K3 together in the reading of the conditional: in every consistent fixed point of Υ, it is possible to give the stronger Ł3-reading to some conditional sentences, while the remaining ones could be understood as K3 conditionals. The relations between the (now) 4 truth-values are summed up in the following "mixed" Hasse diagram, and they are specified by the labels on its lines:



So, if $f^{\rightarrow}$ is the truth-function associated to $\rightarrow$ by the ŁK-construction, we have that $f^{\rightarrow}(1/2, 1/2) = 1$, but $f^{\rightarrow}(\mathbf{N},\mathbf{N}) = \mathbf{N}$. So, $f^{\rightarrow}$ behaves like the corresponding truth-function of Ł3 for all inputs for which a value other than $\mathbf{N}$ results, and like the corresponding truth-function of K3 in the remaining cases.

Lemma 21 entails that the ŁK-construction has no laws: for every consistent fixed point of Υ, there is no schematic law s.t. all its instances have value 1 in it. In our context, partiality and the loss of laws seem inevitable: this is the price to pay to use, whenever possible, the strong Ł3 evaluation schema. Once this price is paid, however, the ŁK-construction improves on Kripke's semantics for the logical vocabulary and on his treatment of truth-theoretic facts, as the following results indicate.

Proposition 22 (Weak Naïveté)
For every $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $P,Q,R \subseteq \omega$:

- $\varphi \in P_\infty$ (Q_∞, R_∞) *if and only if* $T\ulcorner\varphi\urcorner \in P_\infty$ (Q_∞, R_∞).
- $\varphi \in SENT_{\mathscr{L}_T^{\rightarrow}} \setminus (P_\infty \cup Q_\infty \cup R_\infty)$ *if and only if* $T\ulcorner\varphi\urcorner \in SENT_{\mathscr{L}_T^{\rightarrow}} \setminus (P_\infty \cup Q_\infty \cup R_\infty)$.

Proof

$$T\ulcorner\phi\urcorner \in P_\infty \text{ iff } T\ulcorner\phi\urcorner \in \mathfrak{C}_{\Phi}(P_\infty,Q_\infty) \text{ (df. of } \Upsilon) \text{ iff } \phi \in \mathfrak{C}_{\Phi}(P_\infty,Q_\infty) \text{ (df. of } \Phi). \quad (5)$$

Replacing P_∞ with Q_∞ (R_∞) and $\mathfrak{C}_{\Phi}(P_\infty,Q_\infty)$ with $\mathfrak{A}_{\Phi}(P_\infty,Q_\infty)$ $(\mathfrak{H}_{\Psi}(R_\infty,P_\infty,Q_\infty))$ completes the proof of the first claim. As to the second claim, let $\phi \in SENT_{\mathscr{L}_T^{\rightarrow}} \setminus (P_\infty \cup Q_\infty \cup R_\infty)$. If $T\ulcorner\phi\urcorner \in P_\infty$ (or Q_∞ or R_∞), by (5) also $\phi \in P_\infty$, so $\phi \in P_\infty \cap SENT_{\mathscr{L}_T^{\rightarrow}} \setminus (P_\infty \cup Q_\infty \cup R_\infty)$, which is absurd. The same holds starting with $T\ulcorner\phi\urcorner$. □

Corollary 23 (Strong(er) Naïveté)
For every $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $P,Q,R \subseteq \omega$, if $\varphi \in (P_\infty \cup Q_\infty \cup R_\infty)$, then $\varphi \leftrightarrow T\ulcorner\varphi\urcorner \in P_\infty$.

Proof
Let $\phi \in P_\infty$. By Proposition 22, $T\ulcorner\phi\urcorner \in P_\infty$. Then, both $\phi \rightarrow T\ulcorner\phi\urcorner \in P_\infty^{\mathrm{T}}$ and $T\ulcorner\phi\urcorner \rightarrow \phi \in P_\infty^{\mathrm{T}}$. By the fixed-point property, $P_\infty = P_\infty^{\mathrm{T}}$. So $\phi \leftrightarrow T\ulcorner\phi\urcorner \in \mathfrak{C}_{\Phi}(P_\infty,Q_\infty) \subseteq P_\infty^{\mathrm{T}} = P_\infty$. If $\phi \in Q_\infty$ or $\phi \in R_\infty$, the case is dual. □

Corollary 24 (Field's Intersubstitutivity of truth)
For every $\varphi, \psi, \chi \in \mathscr{L}_T^{\rightarrow}$ and every $P, Q, R \subseteq \omega$, if ψ and χ are alike except that one of them has an occurrence of φ where the other has an occurrence of $T\ulcorner\varphi\urcorner$, then:

$$\psi \in P_\infty(Q_\infty, R_\infty) \text{ if and only if } \chi \in P_\infty(Q_\infty, R_\infty).$$

By Corollary 23, we can strengthen this to: if $\psi \in (P_\infty \cup Q_\infty \cup R_\infty)$, then $\psi \leftrightarrow \chi \in P_\infty$.

Proof
By (5), the proof is as in Kripke's theory (the strengthening is immediate). □

Corollary 25 (Restricted Compositionality)
For every $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $P, Q, R \subseteq \omega$, if $\varphi \in (P_\infty \cup Q_\infty \cup R_\infty)$, then:

($T\neg$) *If φ is $\neg\psi$, then $(T\ulcorner\neg\psi\urcorner \leftrightarrow \neg T\ulcorner\psi\urcorner) \in P_\infty$.*
($T\wedge$) *If φ is $\psi \wedge \chi$, then $(T\ulcorner\psi \wedge \chi\urcorner \leftrightarrow (T\ulcorner\psi\urcorner \wedge T\ulcorner\chi\urcorner)) \in P_\infty$.*
($T\vee$) *If φ is $\psi \vee \chi$, then $(T\ulcorner\psi \vee \chi\urcorner \leftrightarrow (T\ulcorner\psi\urcorner \vee T\ulcorner\chi\urcorner)) \in P_\infty$.*
($T\rightarrow$) *If φ is $\psi \rightarrow \chi$, then $(T\ulcorner\psi \rightarrow \chi\urcorner \leftrightarrow (T\ulcorner\psi\urcorner \rightarrow T\ulcorner\chi\urcorner)) \in P_\infty$.*
($T\forall$) *If φ is $\forall x\psi(x)$, then $(T\ulcorner\forall x\psi(x)\urcorner \leftrightarrow \forall xT\ulcorner\psi(\dot{x})\urcorner) \in P_\infty$.*

Proof
If $\varphi \in (P_\infty \cup Q_\infty \cup R_\infty)$, then clearly $\varphi \leftrightarrow \varphi \in P_\infty$. The result is then immediate by the stronger form of Corollary 24. □

Lemma 26 (Fieldian Determinateness)
For every $\varphi \in \mathscr{L}_T^{\rightarrow}$ and $P, Q, R \subseteq \omega$, the following holds:

1. *Let $\mathcal{D}(\varphi)$ stand for $\neg(\varphi \rightarrow \neg\varphi)$. Call $\mathcal{D}$ a* Fieldian determinateness operator*, or* determinateness operator *for short. If $\varphi \in Q_\infty \cup R_\infty$, then $\neg\mathcal{D}(\varphi) \in P_\infty$.*
2. *If $\mathfrak{I}_\Upsilon(P, Q, R)$ is consistent and $\varphi \leftrightarrow \neg\varphi \in P_\infty$, then $\varphi, \neg\varphi \in R_\infty$.*

Proof
Ad 1, let $\varphi \in R_\infty$. By the fixed-point property of Υ, the following claims hold:

$$\neg\varphi \in R_\infty, \;\; \varphi \rightarrow \neg\varphi \in P_\infty, \;\; \neg\neg(\varphi \rightarrow \neg\varphi) \in P_\infty, \;\; \text{i.e. } \neg\mathcal{D}(\varphi) \in P_\infty.$$

The same result follows letting $\varphi \in Q_\infty$.

Ad 2, let $\mathfrak{I}_\Upsilon(P, Q, R)$ be consistent and $\varphi \leftrightarrow \neg\varphi \in P_\infty$. By the fixed-point property of P_∞, both the following hold:

(i) either $(\varphi \in Q_\infty)$ or $(\neg\varphi \in P_\infty)$ or $(\varphi, \neg\varphi \in R_\infty)$;
(ii) either $(\neg\varphi \in Q_\infty)$ or $(\varphi \in P_\infty)$ or $(\neg\varphi, \varphi \in R_\infty)$.

If $\varphi \in Q_\infty$, then $\varphi \notin P_\infty$, $\varphi \notin R_\infty$, and $\neg\varphi \notin Q_\infty$, so item (ii) cannot hold if $\varphi \in Q_\infty$ (the same follows if $\neg\varphi \in P_\infty$). Dually, if $\neg\varphi \in Q_\infty$ or $\varphi \in P_\infty$, condition (i) cannot obtain. Since $\mathfrak{I}_\Upsilon(P, Q, R)$ is consistent, $\varphi \in R_\infty$ and $\neg\varphi \in R_\infty$. □

Corollaries 23, 24, 25, and Lemma 26 show how the ŁK-construction improves on Kripke's theory in enabling us to validate claims about the semantics of $\mathscr{L}_T^{\rightarrow}$, formalized in the object-language itself. In Kripke's theory, Tarski $\equiv$-biconditionals hold only for elements of Kripke truth- and falsity-sets and not for their C-gaps. Despite the fact that there are valueless sentences in consistent fixed points of Υ as well (so, an analogue of C-gaps for the ŁK-construction), here we also have P-gaps at our disposal, and P-gappy sentences are quite tractable. In fact, Corollary 23 tells us that we can validate Tarski $\leftrightarrow$-biconditionals for P-gappy sentences as well. Similar remarks go for Corollary 24 and Intersubstitutivity of truth, and for Corollary 25 and the compositional behavior of naïve truth.

Lemma 26 shows that the ŁK-construction enables us to define a determinateness operator *à la* Field, denoted by $\mathcal{D}$, which declares in the object-language that, for every fixed point $\mathfrak{I}_\Upsilon(P,Q,R)$, every sentence in Q_∞ or in R_∞ is "not determinately true" in $\mathfrak{I}_\Upsilon(P,Q,R)$ – by Corollary 24, "not determinate" and "not determinately true" are interchangeable. The operator $\mathcal{D}$ can be applied consistently to standard paradoxes, such as the liar, asserting positively their lack of determinate truth. For example, the fixed point $\mathfrak{I}_\Upsilon(\emptyset,\emptyset,\{\lambda\})$ is consistent, and $|\neg\mathcal{D}(\lambda)|_{(\emptyset\emptyset\{\lambda\})^\infty} = 1$.[50] No device comparable to $\mathcal{D}$ is available in Kripke's theory: since C-gaps are not usable in Kripke fixed points to validate or refute any claim, the only sentences that Kripke's theory can deem "not determinate" are the sentences in the falsity-set of a fixed point, whose lack of truth can already be captured via the K3 negation. Of course, the notion of determinateness given by the ŁK-construction is quite weak, as it cannot treat a revenge-paradox involving it (the sentence λ^* provably equivalent to $\neg\mathcal{D}(T\ulcorner\lambda^*\urcorner)$) – as it is well-known for evaluations based on Ł3.[51] Despite their simplicity, however, fixed points of Υ are expressive enough to make positive claims about the indeterminateness of their P-gappy sentences. Moreover, consistent fixed points of Υ do not give to any sentence an unintended value: if $\langle P_\infty, Q_\infty, R_\infty\rangle$ is consistent and $\varphi \in Q_\infty \cup R_\infty$, then $\neg\mathcal{D}(\varphi) \in P_\infty$ and $\neg\mathcal{D}(\varphi) \notin (Q_\infty \cup R_\infty)$. In consistent fixed points, revenge paradoxes receive no truth-value. Contrast this with Field's theory: that one is a total theory, every sentence has a value in it, and sometimes an unintended one. The ŁK-construction, although less expressive than Field's theory, makes the notion of determinateness more uniform, as we see from Lemma 26.[52]

The ŁK-construction improves on Kripke's theory in recovering features of naïve truth since it extends Kripke's treatment of truth to positive gappy sentences. P-gaps, unlike C-gaps, can be used to interpret (in the simple setting of inductive definitions) a conditional stronger than the K3 material conditional. However, the improvements on Kripke's theory granted by the ŁK-construction are not just a matter of expressive power: they are especially a matter of philosophical significance. A sheer increase in expressive power is obtained in Field's theory; as I argued, however, so many conceptually desirable features of conditionals are lost in Field's account, that the philosophical significance of the new conditional seems doubtful. In the ŁK-construction, instead, the more articulated aim expressed in the **Main Question** is pursued.

[50] Of course, also $|\lambda \leftrightarrow \neg\lambda|_{(\emptyset\emptyset\{\lambda\})^\infty} = 1$, validating a formal version of statement (G) from Section 2.

[51] A similar fact is given for evaluations based on any Łukasiewicz logic.

[52] I thank an anonymous referee for pointing out that consistent fixed points of Υ show that one can use a conditional very similar to Łukasiewicz' one and get ω-models validating Intersubstitutivity of truth.

The conceptual perspective outlined in the statement of the **Main Question** is also the point of view from which I consider the lack of conditional laws of the ŁK-construction. Of course, one can be mainly interested in recovering as many conditional laws as possible – that is perfectly legitimate. However, I tried to argue in Subsection 3.6 that validating conditional laws without having a philosophically significant reading for the conditional, although technically informative, suffers from severe conceptual difficulties. Moreover, I argued in Subsection 3.2 that the questions leading to these difficulties arise naturally even if we see a theory of truth only as showing the consistency of some principles. For these reasons, it seems that losing conditional laws, or schematic laws in general, is acceptable, if one gains a philosophically significant account of connectives and quantifiers. This conclusion is reinforced by the fact that giving up schematic laws does not mean renouncing the goal of accounting for the behavior of connectives and quantifiers: a theory may have good introduction and elimination rules for that. In Kripke's theory, we lose schematic laws but we gain a philosophically significant interpretation and nice semantic rules for $\neg$, $\wedge$, $\vee$, and $\forall$ (consistently with naïveté): the same cannot be said for Kripke's conditional (see Sections 1 and 2). The ŁK-construction addresses the **Main Question** including a new conditional from the Kripkean perspective just sketched.[53]

In the rest of this work, I will argue that the ŁK-construction satisfies the articulated aim given by the **Main Question**. In this Section we have seen general improvements on the expressive weaknesses of Kripke's theory. In Section 6, I will look into the philosophical interpretation of the new conditional, arguing in favor of its significance. I will also show that the ŁK-construction preserves Kripke's semantic rules and add to them a nice conditional-introduction. In Section 7, I will consider some applications of the ŁK-construction, to vindicate its conceptual fruitfulness and to give concrete examples of how it overcomes specific deficiencies of Kripke's theory.

6 Interpreting the ŁK-construction

6.1 Interpreting the conditional in the ŁK-construction

The conditional modeled by Υ can be interpreted as *a tool to compare truth-values*, in the sense of Subsection 3.3: every consistent fixed point $\mathfrak{I}_\Upsilon(P,Q,R)$ gives conditions for $\varphi \to \psi$ to have values 1, 0, *and* 1/2 in terms of simple relations between the values φ and ψ in $\mathfrak{I}_\Upsilon(P,Q,R)$. This Subsection establishes this claim and explores its philosophical significance. Let's only consider consistent fixed points of Υ.

In building $\mathfrak{I}_\Upsilon(P,Q,R)$, the sets P, Q, and R represent a starting distribution of values 1, 0, and 1/2, giving a partial ordering to $\mathscr{L}_T^{\to}$-sentences: call it $\leq_{PQR}$. The first stage of the ŁK-construction evaluates all the conditionals generated by the original truth-value distribution, as explained at the end of Section 4. A conditional $\varphi \to \psi$

[53] Of course, I am not *inciting* to abandon schematic laws: I am arguing that, in addition to the principles that a theory of truth validates, there are also conceptually motivated criteria of philosophical significance that should play a role in deciding whether to accept a theory or not. When those criteria are not met (as I have argued that it is the case for Field's theory), they should outweigh the importance of schematic laws.

gets value 1 if $\varphi \leq_{PQR} \psi$, it gets value 0 if $\varphi >_{PQR} \psi$ and the truth-value difference between them is 1, and it gets value $1/2$ if $\varphi >_{PQR} \psi$ and the truth-value difference between them is $1/2$. We get new sets $P^{\mathrm{T}}, Q^{\mathrm{F}}, R^{\mathrm{G}}$ and a new ordering $\leq_{(PQR)^+}$ that includes the previous one: if $\varphi \leq_{PQR} \psi$, then $\varphi \leq_{(PQR)^+} \psi$. At the second stage, we have truth-value comparisons according to the $\leq_{(PQR)^+}$-ordering, i.e. truth-value comparisons of truth-value comparisons, and so on. Once at the fixed point, we have exhausted all the starting information given by P, Q, and R, determining via $\rightarrow$ all the truth-value relations deriving from them. *All* these truth-value relations are evaluated at a fixed point because of the fixed-point property: any application of Υ to a fixed-point triple yields the same triple, and every conditional that can be evaluated, given the starting distribution, has already a value. Monotonicity ensures that once a truth-value comparison is established it does not change, thus the fact that the stages of the construction perform truth-value comparisons transfers to the fixed point.

This table shows the values of conditionals in consistent fixed points of Υ.

Table 2: $\rightarrow$-comparisons of truth-values at a (consistent) fixed point of Υ

$\rightarrow$	0	$1/2$	1	**N**
0	1	1	1	1
$1/2$	$1/2$	1	1	**N**
1	0	$1/2$	1	**N**
N	**N**	**N**	1	**N**

Let me emphasize that **N** is not a truth-value given by the ŁK-construction, as it indicates a lack of truth-value in a consistent fixed point of Υ. It's status, then, is similar to the status of value $1/2$ when it is associated to the C-gap of a consistent Kripke fixed point.[54] Therefore, in the following, when referring to the truth-values given by the ŁK-construction I will always mean 1, 0, and $1/2$, unless indicated otherwise.

Using Table 2, we can argue that the ŁK-construction confers a conceptually significant reading on the conditional, improving on Kripke's theory and on the problems (a)-(e) raised in conjunction with Field's theory in Subsection 3.3 (see also Table 1).

- Unlike in Kripke's theory, the ŁK-construction can employ a non-classical truth-value and a notion of gappiness in the evaluation of conditionals. While a Kripke fixed point only deems equivalent two sentences if they both have value 1 or both have value 0, a fixed point of Υ also deems equivalent its $1/2$-valued sentences.
- As seen in Subsection 3.3, several problems of Field's conditional are due to its behavior when a conditional sentence has a value different from **1**. In fixed points of Υ, instead, those conditionals that get a value different from 1 get a value that measures how close the difference of the values of antecedent and

54 See the discussion after Proposition 4.

consequent comes to a 1-valued conditional. If $|\varphi|_{(PQR)^\infty} = 1$, $|\psi_1|_{(PQR)^\infty} = 0$, $|\psi_2|_{(PQR)^\infty} = 1/2$, then φ has a greater value than ψ_2 and a still greater value than ψ_1, i.e. $\varphi \to \psi_1 \lneqq_{(PQR)^\infty} \varphi \to \psi_2$. The model recognizes and expresses it:

$$|(\varphi \to \psi_2) \to (\varphi \to \psi_1)|_{(PQR)^\infty} = 1/2. \tag{6}$$

Contrast the failure of comparability expressed by (3). Thus, for every consistent fixed point $\mathfrak{I}_\Upsilon(P,Q,R)$, the Ł3 conditional provides a complete map of all the truth-value relations between sentences getting a value in $\mathfrak{I}_\Upsilon(P,Q,R)$. And this is very relevant in the light of the general interest that truth-value relations have for the semantics of $\mathscr{L}_T^{\to}$, as it was emphasized at the beginning of Subsection 3.3.

- Unlike Field's theory, the ŁK-construction provides a simple understanding of truth-values. Not only, 1, 0, and 1/2 retain their usual numerical meaning, but every truth-value plus **N** may result from evaluating a conditional in the ŁK-construction. We can use the value 1/2 of the ŁK-construction to interpret any kind of sentence, including conditional statements (contrary to what we have in Field's theory). As a result, the value 1/2 of the ŁK-construction allows for a more uniform treatment of paradoxical sentences: a simple but striking example is the Curry sentence. The fixed point $\mathfrak{I}_\Upsilon(\emptyset,\emptyset,\{\kappa\})$ is consistent and $\kappa \leftrightarrow \neg\kappa$ has value 1 in it. Moreover, also $\mathfrak{I}_\Upsilon(\emptyset,\emptyset,\{\kappa,\lambda\})$ is consistent, so both $\lambda \leftrightarrow \neg\lambda$ and $\kappa \leftrightarrow \neg\kappa$ can jointly have value 1, so the ŁK-construction delivers a simple, unified, and numerical treatment of intuitively related paradoxes.[55]
- The interaction of the truth-values with the lack of truth-value (**N**) is also quite intuitive: since **N** is, essentially, the analogue of a Kripkean C-gap for the ŁK-construction, it also interacts with the truth-values in a way similar to C-gaps. Let me explain this point with a specific example, concerning the conditional. If φ lacks a truth-value and ψ has value 1, the conditional $\varphi \to \psi$ gets value 1, while the conditional $\psi \to \varphi$ lacks a truth-value. Now, it might seem somewhat strange that, reading the conditional as a tool to compare truth-values, we can compare the truth-values of φ and ψ also when it is not the case that both φ and ψ have a truth-value, and that, in addition, we can make this comparison only in one direction.[56] This behavior of the conditional, however, is in line with the strong Kleene way of cashing out partiality, that is adopted in Kripke's theory and that I extended to the present framework. In fact, both in consistent Kripke fixed points and in consistent fixed points of Υ, if the consequent of a conditional ($\supset$ and $\to$, respectively) has the designated value, this suffices to assign the designated value to the entire conditional (even if the antecedent lacks a truth-value), but this is not so if the antecedent of a conditional has the designated value and its consequent has no truth-value. One can informally interpret this situation as follows: if the consequent of a conditional has value 1, this suffices to determine that, whatever

[55] Contrast this fact with item (a) in Subsection 3.3. Note that in every consistent fixed point of Υ, for every $\varphi \in \mathscr{L}_T^{\to}$, the sentence $\neg\varphi$ has the same value of the sentence $\varphi \to 0 \neq 0$, while Field's theory does not validate this general equivalence (see footnote 33). There are good reasons to accept this equivalence, such as its classical validity. If one agrees on this equivalence and interprets naïvely the truth predicate, it is natural to see λ and κ as equivalent between them, as λ is equivalent to $\neg\lambda$ and κ is equivalent to $\kappa \to 0 \neq 0$ (*modulo* Intersubstitutivity of truth).

[56] Let me thank an anonymous referee for highlighting the importance of this point to me.

truth-value the antecedent might have, it has to be less than or equal to 1, since 1 is the greatest truth-value. Nothing similar to this reasoning applies to a conditional whose antecedent has value 1 and whose consequent has no truth-value, so the resulting conditional should have no truth-value.[57] Similar remarks go for the other relations between the truth-values and **N** summarized in Table 2. The ŁK-construction, thus, generalizes the strong Kleene treatment of partiality to Ł3 and provides a rationale to identify those truth-value comparisons that can be performed also when some sentence has no truth-value.

- Since (as I have just argued) when $\varphi \rightarrow \psi$ has no truth-value there is no asking about a truth-value comparison between φ and ψ, it is natural to compare **N** to the sub-space **E** of Field's value space **F**. However, in every consistent fixed point of Υ, we can always distinguish between conditionals that are given a value and are treated as in Ł3 semantics, and conditionals that are not. The value zone **E**, instead, has a much more complex and less intuitive behavior, as we have seen in Subsection 3.3, items (a) and (b).
- The Ł3 interpretation, as it is well-known, is too strong to be applied to every conditional sentence, and a residual of unevaluated sentences is bound to be generated by every consistent fixed point of Υ. The ŁK-construction, however, shows that, despite the non-monotonic character of the Ł3 conditional, a partial version of this connective can be used within the framework of inductive definitions *a là* Kripke. So, the general picture given by the ŁK-construction is not conceptually different from what we have in Kripke's theory, but it has been shown to provide a *substantial incremental progress* on that construction.[58]

6.2 Conditional-introduction in the ŁK-construction

The ŁK-construction improves on Kripke's and Field's theories as far as the introduction of the conditional is concerned.[59] Note in the first place that every consistent fixed point $\mathfrak{I}_{\Upsilon}(P,Q,R)$ has the following closure properties: for all $\varphi, \psi \in \mathscr{L}_T^{\rightarrow}$,

- If $\varphi \in P_\infty$ and $\varphi \rightarrow \psi \in P_\infty$, then $\psi \in P_\infty$.
- If $\varphi \in Q_\infty$, or $\psi \in P_\infty$, or both $\varphi, \psi \in R_\infty$, then $\varphi \rightarrow \psi \in P_\infty$.

The second clause can be used to formulate a rule of conditional-introduction that expresses genuinely the difference between *conditional assertion* and *assertion of a conditional* in the context of the ŁK-construction. The next Definition and Proposition, together with the subsequent remarks, establish this point.

[57] Clearly, the situation would be entirely different if we adopted a different interpretation of partiality, e.g. the view embodied by the weak Kleene scheme.

[58] An anonymous referee suggested that the line of thought I follow here could be used to argue in favor of "another inductively defined gap" (and suitable extensions of the treatment of the conditional), "in addition to the P-gaps", along the lines offered by Roy Cook in his [11]. Cook propounds an indefinite extensibility theory in order to address revenge paradoxes. Cook's theory presents a progression of larger and larger languages, and of more and more truth-values. The referee, then, seems to indicate that my P-gaps could be subject to similar extensions. This suggestion is particularly interesting, and I would like to address it in a future work.

[59] I thank an anonymous referee for some useful suggestions on this point.

Definition 27 (Logical consequence generated by consistent fixed points of Υ) Let $\mathcal{U}$ be any set of consistent fixed points of Υ. Define the *ŁK-logical consequence generated by* $\mathcal{U}$, in symbols $\models^{\mathcal{U}}_{ŁK}$, as follows:

$$\Gamma \models^{\mathcal{U}}_{ŁK} \psi \text{ if and only if } \text{ for every } \mathfrak{I}_{\Upsilon}(P,Q,R) \in \mathcal{U}, \text{ if for every } \varphi \in \Gamma, \varphi \in P_{\infty}, \text{ then } \psi \in P_{\infty}.$$

Proposition 28
Let $\mathcal{H}(\varphi,\psi)$ abbreviate the sentence $(\varphi \vee \neg\varphi) \vee \psi \vee [(\varphi \leftrightarrow \neg\varphi) \wedge (\psi \leftrightarrow \neg\psi)]$. For every $\mathcal{U}$ as in Definition 27, and for all $\varphi,\psi \in \mathscr{L}^{\rightarrow}_{T}$ and $\Gamma \subseteq SENT_{\mathscr{L}^{\rightarrow}_{T}}$:

$$\Gamma,\varphi \models^{\mathcal{U}}_{ŁK} \psi \text{ and } \Gamma \models^{\mathcal{U}}_{ŁK} \mathcal{H}(\varphi,\psi) \quad \text{if and only if} \quad \Gamma \models^{\mathcal{U}}_{ŁK} \varphi \rightarrow \psi.$$

Proof
I suppress Γ for simplicity. The right-to-left direction is immediate. From left-to-right, let $\varphi,\psi \in \mathscr{L}^{\rightarrow}_{T}$ be s.t. $\varphi \models^{\mathcal{U}}_{ŁK} \psi$ and $\models^{\mathcal{U}}_{ŁK} \mathcal{H}(\varphi,\psi)$ and consider these cases:

(A) There is a $\mathfrak{I}_{\Upsilon}(P',Q',R') \in \mathcal{U}$ s.t. $\psi \in P'_{\infty}$. In this case, $\varphi \rightarrow \psi \in P'_{\infty}$.
(B) There is a $\mathfrak{I}_{\Upsilon}(P^{\dagger},Q^{\dagger},R^{\dagger}) \in \mathcal{U}$ s.t. $\varphi \vee \neg\varphi \in P^{\dagger}_{\infty}$. In this case, $\varphi \in P^{\dagger}_{\infty}$ or $\neg\varphi \in P^{\dagger}_{\infty}$. If $\neg\varphi \in P^{\dagger}_{\infty}$, then by definition of Υ, $\varphi \rightarrow \psi \in P^{\dagger}_{\infty}$. If $\varphi \in P^{\dagger}_{\infty}$, by our assumption that $\varphi \models^{\mathcal{U}}_{ŁK} \psi$, then also $\psi \in P^{\dagger}_{\infty}$, so $\varphi \rightarrow \psi \in P^{\dagger}_{\infty}$.
(C) There is a $\mathfrak{I}_{\Upsilon}(P^{\ddagger},Q^{\ddagger},R^{\ddagger}) \in \mathcal{U}$ s.t. $[(\varphi \leftrightarrow \neg\varphi) \wedge (\psi \leftrightarrow \neg\psi)] \in P^{\ddagger}_{\infty}$. In this case, $\varphi,\psi \in R^{\ddagger}_{\infty}$ by Lemma 26, so $\varphi \rightarrow \psi \in P^{\ddagger}_{\infty}$.

At least one of (A), (B), and (C) is the case, so $\models^{\mathcal{U}}_{ŁK} \varphi \rightarrow \psi$ holds. □

Let **(ŁK→-Intro)** indicate the conditional-introduction rule given by the left-to-right direction of Proposition 28. This rule is interesting under several respects:

1. **(ŁK→-Intro)** is arguably better than the clauses that allow us to introduce a conditional when LEM holds for the antecedent (e.g. (4) for Field's theory). Unlike them, **(ŁK→-Intro)** can use also information on sentences having value 1/2. In Kripke's theory, we cannot use non-classical values because of the status of C-gaps. In Field's theory, value **1/2** is not very interesting (conditionals cannot receive this value), and we cannot use clause (4) to characterize the conditionals $\varphi \rightarrow \psi$ such that $\models_F \varphi \rightarrow \psi$. These problems are avoided here, thanks to P-gaps.
2. It is easy to see that $\models^{\mathcal{U}}_{ŁK} \mathcal{H}(\varphi,\psi)$ is a *genuine* conditional-introduction condition, in the sense of Definition 14. For every set $\mathcal{U}$ of consistent fixed points of Υ, the schema $\mathcal{H}(\varphi,\psi)$ is not equivalent to the $\models^{\mathcal{U}}_{ŁK}$-validity of all the corresponding conditionals (e.g. $\models^{\mathcal{U}}_{ŁK} \mathcal{H}(0=0,\lambda)$, but $\not\models^{\mathcal{U}}_{ŁK} 0=0 \rightarrow \lambda$), and the same holds of all its subschemata, as the proof of Proposition 28 shows. The formula $\mathcal{H}(\varphi,\psi)$ expresses genuinely the difference between conditional assertion and assertion of a conditional, while this task, as seen in Subsection 3.4, seems difficult for Field's semantics. Moreover, **(ŁK→-Intro)** indicates an interesting link between these two notions: it is consistent with naïveté to express in the object-language a valid semantic inference, via a conditional, not only when the sentences involved have classical values (value 1 in case of the consequent), but also when they have a positive gappy status, i.e. the value 1/2 of a partial version of Ł3.

Finally, it is immediate from the first closure property mentioned at the beginning of this Subsection that *modus ponens* holds for every notion of logical consequence $\models^{\mathcal{U}}_{ŁK}$ given as in Definition 27, and moreover $\neg$, $\wedge$, $\vee$, and $\forall$ are treated as in Kripke's theory by the ŁK-construction (and the same holds for Field's theory). So, if we associate a logical consequence patterned after Definition 27 to Kripke's theory, the resulting clauses to introduce and eliminate $\neg$, $\wedge$, $\vee$, and $\forall$ are inherited by $\models^{\mathcal{U}}_{ŁK}$, for every set $\mathcal{U}$ of consistent fixed points of Υ.[60]

6.3 Recovering Kripke's theory in the ŁK-construction

As I mentioned, a point of interest about the ŁK-construction is its flexibility. The roots of this feature are in the fact that, unlike Φ, the P-gappy jump Ψ does not *build* sets of P-gappy sentences. The Kripke jump Φ assigns values 1 and 0 to $\mathscr{L}^{\rightarrow}_{T}$-sentences based on some starting sets of sentences (possibly empty) *and* facts concerning the base language, in our case arithmetical truths and falsities. The P-gappy jump, instead, does not give value $1/2$ to any arithmetical sentence. I take it for granted that there are no "gappy" arithmetical facts.[61] As a result, the extension of a P-gap determined by Ψ depends only on the starting hypotheses. This feature was designed to allow the ŁK-construction to apply to sets embodying various notions of gappiness. The following lemma shows that the ŁK-construction has room for literally *any* notion of gappiness, as an empty initial choice generates no P-gap.

Lemma 29
For every $P, Q \subseteq \omega$, $\mathfrak{H}_{\Psi}(\emptyset, P, Q) = \emptyset$.

Proof
For all ordinals α, $\Psi^{\alpha}(\emptyset, P, Q) = \langle \emptyset, \mathfrak{E}_{\Phi}(P, Q), \mathfrak{A}_{\Phi}(P, Q)\rangle$ (by an easy induction). □

A corollary is straightforward.

Corollary 30
For $P, Q \subseteq \omega$, let $\langle P^{\emptyset}_{\alpha}, Q^{\emptyset}_{\alpha}, R^{\emptyset}_{\alpha}\rangle := \Upsilon^{\alpha}(P, Q, \emptyset)$. For all ordinals α and all $\varphi, \psi \in \mathscr{L}^{\rightarrow}_{T}$:

1. $\varphi \rightarrow \psi \in P^{\emptyset}_{\alpha}$ *iff either* $\varphi \rightarrow \psi \in P$ *or* $\neg\varphi \vee \psi \in P^{\emptyset}_{\alpha}$.
2. $\varphi \rightarrow \psi \in Q^{\emptyset}_{\alpha}$ *iff either* $\varphi \rightarrow \psi \in Q$ *or* $\varphi \wedge \neg\psi \in P^{\emptyset}_{\alpha}$.
3. $R^{\emptyset}_{\alpha} = \emptyset$.

The proof is routine, I sketch it in a footnote.[62] The latter result shows that the ŁK-construction is a generalization of Kripke's theory, since it can define also all Kripke

[60] In Definition 27, I consider only *consistent* fixed points of Υ, to avoid losing central pieces of semantic reasoning. There are sentences $\varphi, \psi \in \mathscr{L}^{\rightarrow}_{T}$ and sets $\mathcal{V}$ of fixed points of Υ (including inconsistent fixed points) s.t., defining $\models^{\mathcal{V}}_{ŁK}$ as in Definition 27, we have that $\models^{\mathcal{V}}_{ŁK} \varphi \rightarrow \psi$ but $\varphi \not\models^{\mathcal{V}}_{ŁK} \psi$, which upsets a fundamental relation between conditional assertion and assertion of the conditional. It is possible to liberalize Definition 27 to include *some* inconsistent fixed points of Υ without this disastrous consequence, but this does not seem very interesting, as we cannot use arbitrary sets of fixed points of Υ. Identical remarks hold for an analogue of Definition 27 for Kripke's theory and the K3 material conditional $\supset$.

[61] This may be not entirely uncontroversial, but I cannot discuss this point here.

[62] To simplify the notation, let $P = \emptyset = Q$: the general case is immediate from this one (the proof sketch continues on the next page).

fixed points, not adding anything irrelevant to them (i.e. interpreting $\rightarrow$ as a notational variant of $\supset$ and adding the empty set as P-gap). So, the fixed points of Υ range from essentially Kripke fixed points to consistent fixed points validating claims that Kripke's theory cannot account for in any sense, such as the statement (G), or $\lambda \leftrightarrow \kappa$. These sentences do not belong to every fixed point of Υ, as this would impose some fixed features on our notions of gappiness and would make it impossible to recover Kripke fixed points in the above sense. This fact, of course, draws our attention to *non-minimal* fixed points of Υ. Unlike Kripke's theory, whose $\leq_S$-least fixed point is often considered to be the most interesting,[63] the $\leq_S$-least fixed point of Υ is as interesting as Kripke's one. Non-minimal fixed points of the ŁK-construction, as we will see, can be really used to model various intuitions about gappiness, rather than being a mere technical possibility. The existence of "big" non-minimal consistent fixed points of Υ is easy to see, from standard considerations.[64]

Lemma 31 (Maximally consistent fixed points of Υ)
Let $P,Q,R \subseteq \omega$ be s.t. $\mathfrak{I}_\Upsilon(P,Q,R)$ is consistent. There exist sets $P^M, Q^M, R^M \subseteq \omega$ s.t.:

- $P \subseteq P^M$, $Q \subseteq Q^M$, *and* $R \subseteq R^M$.
- $\langle P^M_\infty, Q^M_\infty, R^M_\infty \rangle$ *is maximally consistent, i.e. it is consistent and for all* $\varphi \in \mathscr{L}_T^{\rightarrow}$ *s.t.* $\varphi \notin (P^M_\infty \cup Q^M_\infty \cup R^M_\infty)$, *we have:*

Proof (Sketch)

$$\mathfrak{I}^1_\Upsilon = \begin{array}{l} \Big\langle \mathfrak{E}_\Phi \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \begin{cases} \varphi \in \mathfrak{A}_\Phi \text{ or} \\ \psi \in \mathfrak{E}_\Phi \text{ or} \\ \varphi, \psi \in \mathfrak{H}_\Psi \end{cases} \}; \\ \mathfrak{A}_\Phi \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} | \varphi \in \mathfrak{E}_\Phi, \psi \in \mathfrak{A}_\Phi\}; \\ \mathfrak{H}_\Psi \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \begin{cases} \varphi \in \mathfrak{E}_\Phi, \psi \in \mathfrak{H}_\Psi, \text{ or} \\ \varphi \in \mathfrak{H}_\Psi, \psi \in \mathfrak{A}_\Phi \end{cases} \}\Big\rangle \end{array}$$

$$= \big\langle \mathfrak{E}_\Phi \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \neg\varphi \in \mathfrak{E}_\Phi \text{ or } \psi \in \mathfrak{E}_\Phi\}; \mathfrak{A}_\Phi \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \varphi \in \mathfrak{E}_\Phi, \neg\psi \in \mathfrak{E}_\Phi\}; \emptyset\big\rangle$$

$$= \big\langle \mathfrak{E}_\Phi \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \neg\varphi \vee \psi \in \mathfrak{E}_\Phi\}; \mathfrak{A}_\Phi \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \varphi \wedge \neg\psi \in \mathfrak{E}_\Phi\}; \emptyset\big\rangle.$$

IH: The claim holds for all $\beta < \alpha + 1$

$$\mathfrak{I}^{\alpha+1}_\Upsilon = \begin{array}{l} \Big\langle \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} | \begin{cases} \varphi \in \mathfrak{A}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \text{ or} \\ \psi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \text{ or} \\ \varphi, \psi \in \mathfrak{H}_\Psi(\emptyset, P^\emptyset_\alpha, Q^\emptyset_\alpha) \end{cases} \}; \\ \mathfrak{A}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} | \varphi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha), \psi \in \mathfrak{A}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha)\}; \\ \mathfrak{H}_\Psi(\emptyset, P^\emptyset_\alpha, Q^\emptyset_\alpha) \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} | \begin{cases} \varphi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha), \psi \in \mathfrak{H}_\Psi(\emptyset, P^\emptyset_\alpha, Q^\emptyset_\alpha) \text{ or} \\ \varphi \in \mathfrak{H}_\Psi(\emptyset, P^\emptyset_\alpha, Q^\emptyset_\alpha), \psi \in \mathfrak{A}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \end{cases} \}\Big\rangle \end{array}$$

$$= \begin{array}{l} \big\langle \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \neg\varphi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \text{ or } \psi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha); \\ \mathfrak{A}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \varphi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha), \neg\psi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha)\}; \emptyset\big\rangle \end{array}$$

$$= \begin{array}{l} \big\langle \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \neg\varphi \vee \psi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha); \\ \mathfrak{A}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha) \cup \{\varphi \rightarrow \psi \in \mathscr{L}_T^{\rightarrow} \,|\, \varphi \wedge \neg\psi \in \mathfrak{E}_\Phi(P^\emptyset_\alpha, Q^\emptyset_\alpha)\}; \emptyset\big\rangle. \end{array}$$

□

[63] For a discussion, see Burgess [8] and Martin [32].

[64] The proof is routine, see Cantini [9], Theorem 30.13 and Lemma 31.2.

- $\mathfrak{I}_\Upsilon((P^M \cup \{\varphi\}), Q^M, R^M)$ *is inconsistent.*
- $\mathfrak{I}_\Upsilon(P^M, (Q^M \cup \{\varphi\}), R^M)$ *is inconsistent.*
- $\mathfrak{I}_\Upsilon(P^M, Q^M, (R^M \cup \{\varphi\}))$ *is inconsistent.*

Some interesting non-minimal consistent fixed points of Υ will be discussed in the next Section; before that, however, I present another feature of the ŁK-construction, to clarify its applicability. Non-minimal fixed points may seem unpleasantly arbitrary. However, any fixed point of Υ can be turned into *the* $\leq_S$-least fixed point of a uniformly defined variant of Υ.

Definition 32 (Relativized ŁK-construction)
Let $P, Q, R \subseteq \omega$. Define Υ_{PQR}, the Υ*-construction relative to* $\langle P, Q, R\rangle$ as follows.

- Let $\Phi_{P,Q}$ be the monotone operator obtained by adding the clause "$n \in P$" to the first item in Definition 1, and the clause "$n \in Q$" to its the second item.
- Let $\Psi_{R,P,Q}$ be the monotone operator obtained by:
  - Adding to Definition 19 the clause "$n \in R$".
  - Replacing everywhere in Definition 19 "$\varphi \in \mathfrak{E}_{\Phi_{P,Q}}$" for "$\varphi \in \mathfrak{E}_\Phi(P,Q)$".
  - Replacing everywhere in Definition 19 "$\varphi \in \mathfrak{A}_{\Phi_{P,Q}}$" for "$\varphi \in \mathfrak{A}_\Phi(P,Q)$".
- Define $\Upsilon_{P,Q,R}$ by substituting $\Phi_{P,Q}$ everywhere for Φ, and $\Psi_{R,P,Q}$ everywhere for Ψ, in Definition 20.

Lemma 33
For any $P, Q, R \subseteq \omega$, $\mathfrak{I}_{\Phi_{P,Q}} = \mathfrak{I}_\Phi(P,Q)$, $\mathfrak{I}_{\Psi_{R,P,Q}} = \mathfrak{I}_\Psi(R,P,Q)$, *and* $\mathfrak{I}_{\Upsilon_{P,Q,R}} = \mathfrak{I}_\Upsilon(P,Q,R)$. *Moreover,* $\mathfrak{I}_{\Upsilon_{P,Q,R}}$ *is inductive in* P, Q, R.

6.4 Summing up: general aspects of the ŁK-construction

Similarly to Kripke's theory, the ŁK-construction is a general template to give a fairly rich array of theories of naïve truth, consisting in the fixed points of Υ (or the minimal fixed points of the variants of Υ). Such theories have features that make them good candidates to improve on the weaknesses of Kripke's theory explored in Section 2, without the conceptual problems of Field's theory. Every fixed point of Υ:

(i) preserves Kripke's theory, in the strong sense given by Corollary 30;
(ii) gives a simple and conceptually interesting reading to the conditional, which can be seen as comparing 1, 0, *and the non-classical value* $1/2$ (in the sense of Subsections 3.3 and 6.1), being a partial version of the Ł3 conditional;
(iii) is given using the relatively simple means of inductive definitions, the same means employed by Kripke's construction, being therefore acceptable for anyone accepting Kripke's theory.

In addition, the fixed points of Υ that have a non-empty starting hypothesis for P-gappy sentences improve essentially on Kripke's theory, as they:

(i+) add to Kripke's theory a valuation for P-gappy sentences that makes explicit the notion of gappiness we want to capture – and this improves on Kripke's treatment of truth-theoretic facts (as shown in Section 5);

(ii+) allow us to formalize and validate claims that Kripke's theory cannot give (e.g. statement (G)), yield introduction/elimination clauses that use also non-classical information (and express genuinely the difference between conditional assertion and assertion of the conditional), and define a determinateness operator that is stronger than the K3 negation.

These are general features of the ŁK-construction, but concrete improvements on Kripke's theory can only be given by specific fixed points of Υ. I turn now to an application of the ŁK-construction that arguably constitutes such an improvement, a philosophically motivated collection of fixed points of Υ.

7 An application of the ŁK-construction: $\mathscr{L}_T$-grounded conditionals

I present now an application that isolates an interesting family of fixed points of Υ and embodies a *meta-theoretical* view of the role of the conditional in $\mathscr{L}_T^{\rightarrow}$. Let's start from the distribution of values obtained (*not explicitly*) in a consistent Kripke fixed point *for the language* $\mathscr{L}_T$. Let $A,B \subseteq SENT_{\mathscr{L}_T}$ be s.t. $\mathfrak{I}_\Phi(A,B)$ is consistent. Put $C^{A,B} := SENT_{\mathscr{L}_T} \setminus (\mathfrak{E}_\Phi(A,B) \cup \mathfrak{A}_\Phi(A,B))$. Let $\langle A_\alpha, B_\alpha, C_\alpha^{A,B}\rangle := \Upsilon^\alpha(A,B,C^{A,B})$ and $\langle A_\infty, B_\infty, C_\infty^{A,B}\rangle := \mathfrak{I}_\Upsilon(A,B,C^{A,B})$. $C^{A,B}$ is an important kind of set in a partial semantics. Kripke's theory was conceived for $\mathscr{L}_T$, the sublanguage of $\mathscr{L}_T^{\rightarrow}$ that does not have $\rightarrow$. Assuming $C^{A,B}$, we make available to the broader construction via Υ *all* the negative information relative to C-Gaps in $\mathscr{L}_T$, integrating it in P-Gaps. Some sentences are not in $(A_\infty \cup B_\infty \cup C_\infty^{A,B})$, such as κ, but this is in line with the rationale of fixed points such as $\langle A_\infty, B_\infty, C_\infty^{A,B}\rangle$: they ignore sentences that are not *grounded*, in the Kripkean sense, *in the conditional-free language* $\mathscr{L}_T$.[65] Fixed points of the form $\langle A_\infty, B_\infty, C_\infty^{A,B}\rangle$ have some nice features: to begin with, they are always consistent.

Proposition 34 (Consistency)
Every fixed point of the form $\langle A_\infty, B_\infty, C_\infty^{A,B}\rangle$ *is consistent.*

Proof
The proof is made of distinct claims, similar between them. I will only do some cases.

1 A_0, B_0 and $C_0^{A,B}$ are pwd. Immediate by definition of $C_0^{A,B}$.
2a For every α, if A_α, B_α, $C_\alpha^{A,B}$ are pwd, then $\mathfrak{E}_\Phi(A_\alpha,B_\alpha) \cap \mathfrak{A}_\Phi(A_\alpha,B_\alpha) = \emptyset$.
If $\alpha = 0$, it's immediate by our assumption. Let $\alpha > 0$ and suppose by contradiction that for a $\varphi \in \mathscr{L}_T^{\rightarrow}$, $\varphi \in \mathfrak{E}_\Phi(A_\alpha,B_\alpha) \cap \mathfrak{A}_\Phi(A_\alpha,B_\alpha)$. Then, there is a unique $\delta \geq 0$ s.t.: $\varphi \notin \mathfrak{E}_\Phi^\delta(A_\alpha,B_\alpha) \cap \mathfrak{A}_\Phi^\delta(A_\alpha,B_\alpha)$ but $\varphi \in \mathfrak{E}_\Phi^{\delta+1}(A_\alpha,B_\alpha) \cap \mathfrak{A}_\Phi^{\delta+1}(A_\alpha,B_\alpha)$. Let's argue by cases according to the logical form of φ.

IH: For all $\gamma \leq \delta$, $\mathfrak{E}_\Phi^\gamma(A_\alpha,B_\alpha)$, $\mathfrak{A}_\Phi^\gamma(A_\alpha,B_\alpha)$, and $\mathfrak{H}_\Psi^\gamma(C_\alpha^{A,B},A_\alpha,B_\alpha)$ are pwd.

φ is $(s = t)$ for $s,t \in TER_{\mathscr{L}_T^{\rightarrow}}$. Then $\varphi \in \mathfrak{E}_\Phi^{\delta+1}(A_\alpha,B_\alpha) \cap \mathfrak{A}_\Phi^{\delta+1}(A_\alpha,B_\alpha)$ iff s and t both coincide and differ in their value, which is absurd.
φ is $\neg\psi$. So $\psi \in \mathfrak{E}_\Phi^\delta(A_\alpha,B_\alpha) \cap \mathfrak{A}_\Phi^\delta(A_\alpha,B_\alpha)$, against the IH.

[65] I am referring primarily to Kripke's informal idea of groundedness (see [29], p. 694 and p. 701), but the present discussion could also be related to the more formal notion that Kripke discusses on p. 706 and following. For an analysis of Kripkean groundedness, see Yablo [44].

φ is $\psi \wedge \chi$. So ($\psi, \chi \in \mathfrak{E}^{\delta}_{\Phi}(A_\alpha, B_\alpha)$) and ($\psi$ or $\chi \in \mathfrak{A}^{\delta}_{\Phi}(A_\alpha, B_\alpha)$), against the IH.
φ is $\psi \vee \chi$. Dual to the above case.
φ is $\psi \to \chi$. There is no clause for $\to$ in Φ. So, $\varphi \in \mathfrak{E}^{\delta+1}_{\Phi}(A_\alpha, B_\alpha) \cap \mathfrak{A}^{\delta+1}_{\Phi}(A_\alpha, B_\alpha)$ iff $\varphi \in \mathfrak{E}^{\delta}_{\Phi}(A_\alpha, B_\alpha) \cap \mathfrak{A}^{\delta}_{\Phi}(A_\alpha, B_\alpha)$, against the IH.
φ is $\forall x \chi(x)$. So ($\chi(t) \in \mathfrak{E}^{\delta}_{\Phi}(A_\alpha, B_\alpha)$ for all $t \in CTER_{\mathscr{L}_T^{\to}}$) and ($\chi(s) \in \mathfrak{A}^{\delta}_{\Phi}(A_\alpha, B_\alpha)$ for some $s \in CTER_{\mathscr{L}_T^{\to}}$), against the IH.
φ is Tt, $dec(t) = \ulcorner \chi \urcorner$ and $\chi \in \mathfrak{E}^{\delta}_{\Phi}(A_\alpha, B_\alpha) \cap \mathfrak{A}^{\delta}_{\Phi}(A_\alpha, B_\alpha)$, against the IH.

2b For every α, if A_α, B_α, $C^{\text{A,B}}_\alpha$ are pwd, then $\mathfrak{E}_{\Phi}(A_\alpha, B_\alpha) \cap \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha) = \emptyset$.
The proof is substantially similar to the proof of claim **2a**.

2c For every α, if A_α, B_α, $C^{\text{A,B}}_\alpha$ are pwd, then $\mathfrak{A}_{\Phi}(A_\alpha, B_\alpha) \cap \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha) = \emptyset$.
The proof is substantially similar to the proof of claim **2a**.

3a For every ordinal α, if A_α, B_α, and $C^{\text{A,B}}_\alpha$ are pwd, then $A_{\alpha+1} \cap B_{\alpha+1} = \emptyset$.
See the proof of claim **3b** below.

3b For every ordinal α, if A_α, B_α, and $C^{\text{A,B}}_\alpha$ are pwd, then $A_{\alpha+1} \cap C^{\text{A,B}}_{\alpha+1} = \emptyset$.
Suppose that $A_{\alpha+1} \cap C^{\text{A,B}}_{\alpha+1} \neq \emptyset$. $\mathfrak{E}_{\Phi}(A_\alpha, B_\alpha)$, $\mathfrak{A}_{\Phi}(A_\alpha, B_\alpha)$, $\mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha)$ are pwd by **2abc**, so there is a $\psi \to \chi$ s.t. $\psi \to \chi \in A_{\alpha+1} \cap C^{\text{A,B}}_{\alpha+1} \setminus A_\alpha \cap C^{\text{A,B}}_\alpha$. So one of the following obtains:

1. $\psi \in \mathfrak{A}_{\Phi}(A_\alpha, B_\alpha)$ and $\begin{cases} \psi \in \mathfrak{E}_{\Phi}(A_\alpha, B_\alpha) \text{ and } \chi \in \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha), \text{ or} \\ \psi \in \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha) \text{ and } \chi \in \mathfrak{A}_{\Phi}(A_\alpha, B_\alpha) \end{cases}$
against our assumption and claims **2abc**.
2. $\chi \in \mathfrak{E}_{\Phi}(A_\alpha, B_\alpha)$ and $\begin{cases} \psi \in \mathfrak{E}_{\Phi}(A_\alpha, B_\alpha) \text{ and } \chi \in \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha), \text{ or} \\ \psi \in \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha) \text{ and } \chi \in \mathfrak{A}_{\Phi}(A_\alpha, B_\alpha) \end{cases}$
against our assumption and claims **2abc**.
3. $\psi, \chi \in \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha)$ and: $\begin{cases} \psi \in \mathfrak{E}_{\Phi}(A_\alpha, B_\alpha), \chi \in \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha), \text{ or} \\ \psi \in \mathfrak{H}_{\Psi}(C^{\text{A,B}}_\alpha, A_\alpha, B_\alpha), \chi \in \mathfrak{A}_{\Phi}(A_\alpha, B_\alpha) \end{cases}$
against our assumption and claims **2abc**.

3c For every ordinal α, if A_α, B_α, and $C^{\text{A,B}}_\alpha$ are pwd, then $B_{\alpha+1} \cap C^{\text{A,B}}_{\alpha+1} = \emptyset$.
The proof is substantially similar to the proof of claim **3b**.

4 For every limit ordinal γ, A_γ, B_γ, and $C^{\text{A,B}}_\gamma$ are pwd.
This is straightforward from claims **3abc** since Υ is monotone, so for γ a limit:

$$\langle A_\gamma, B_\gamma, C^{\text{A,B}}_\gamma \rangle = \Upsilon^\gamma(A, B, C^{\text{A,B}}) = \bigcup_{\delta<\gamma} \Upsilon^\delta(A, B, C^{\text{A,B}}) = \bigcup_{\delta<\gamma} \langle A_\delta, B_\delta, C^{\text{A,B}}_\delta \rangle$$

where A_δ, B_δ, and $C^{\text{A,B}}_\delta$ are pwd for all $\delta < \gamma$. □

In addition to being consistent, fixed points such as $\langle A_\infty, B_\infty, C^{\text{A,B}}_\infty \rangle$ confer a real expressive capability to $\to$, as opposed to the fixed points presented in Corollary 30.

Lemma 35 (Non-triviality)
For every $n \in \omega$ there are $\varphi_n, \psi_n \in \mathscr{L}_T^{\to}$ s.t. $\varphi_n \to \psi_n \in A_{n+1} \setminus A_n$ but $\neg\varphi_n \vee \psi_n \in C^{\text{A,B}}_{n+1}$.

Proof
Let $\chi \in C^{\text{A,B}}$. $\mathfrak{I}_{\Phi}(A, B)$ is consistent, so $C^{\text{A,B}} \neq \emptyset$. Define by primitive recursion:

$$\begin{cases} \varphi_0 \text{ is } \chi \\ \varphi_{n+1} \text{ is } 0=0 \to \varphi_n \end{cases}$$

and put $\varphi_n = \psi_n$; the claim follows by a straightforward induction. □

It is easy (but long) to show that non-triviality holds at some infinite stages too. A simple example is the following: define a function t by primitive recursion as follows:

$$\begin{cases} t(0,\ulcorner\chi\urcorner) := \chi \\ t(n+1,\ulcorner\chi\urcorner) := T\ulcorner 0=0 \to \underset{.}{t}(n,\ulcorner\chi\urcorner)\urcorner \end{cases}$$

It's straightforward to verify that:

- $(\forall n)[T\ulcorner \underset{.}{t}(n,\ulcorner\chi\urcorner)\urcorner] \to (\forall n)[T\ulcorner \underset{.}{t}(n,\ulcorner\chi\urcorner)\urcorner] \in A_{\omega+2} \setminus A_{\omega+1}$.
- $\neg(\forall n)[T\ulcorner \underset{.}{t}(n,\ulcorner\chi\urcorner)\urcorner] \vee (\forall n)[T\ulcorner \underset{.}{t}(n,\ulcorner\chi\urcorner)\urcorner] \in C^{\mathrm{A,B}}_{\omega+2} \setminus C^{\mathrm{A,B}}_{\omega+1}$.[66]

The construction of $\langle A_\infty, B_\infty, C^{\mathrm{A,B}}_\infty\rangle$ is non-trivial: new sentences are actually added as the construction grows. At the same time, the new conditional $\to$, interpreted by this kind of fixed points, is essentially more expressive than $\supset$, being able to interpret *all* the P-gappy sentences relative to the starting choice, and determining their truth-value relationships with other sentences in the build-up process.

As for the expression of truth-theoretic facts, Proposition 22, Corollaries 23, 24, 25 and Lemma 26 hold for every fixed point of the form $\langle A_\infty, B_\infty, C^{\mathrm{A,B}}_\infty\rangle$ as well. The divergence between weak and strong naïveté (Proposition 22 and Corollary 23) is now interpreted in the light of the role that this application confers to the new conditional – a sort of meta-theoretical role made explicit – still retaining the reading of $\to$ as a tool to compare truth-values given in general by the ŁK-construction. The fixed points of the form $\langle A_\infty, B_\infty, C^{\mathrm{A,B}}_\infty\rangle$ are designed to express *exactly* the if-then statements and equivalences between $\mathscr{L}_T$-sentences that no application of Kripke's construction can express, plus all the derived conditionals and equivalences between $\mathscr{L}_T^{\to}$-sentences obtained starting from the original Kripkean partitions. From the point of view of this application, the language one is really interested in is $\mathscr{L}_T$: the detour through $\mathscr{L}_T^{\to}$ is instrumental to express better the semantics of $\mathscr{L}_T$.

One could claim that fixed points of the form $\langle A_\infty, B_\infty, C^{\mathrm{A,B}}_\infty\rangle$ embody the addition of a new conditional to the Kripkean construction for its intended language: such fixed points preserve the original construction by Kripke for $\mathscr{L}_T$ and add to it exactly what was missing to make it expressive. The C-gaps of consistent Kripke fixed points for the language $\mathscr{L}_T$, which cannot be used to extract any information in the inductive framework of Kripke's theory, are now converted into positive information thanks to P-gaps, and used in a more extended inductive framework to overcome the limitations (K2), (K3), and (K4) highlighted in the Introduction.[67] Most importantly,

[66] t is a variant of the so-called McGee function (see Halbach [25], p. 157). This process holds for larger ordinals and closes at some ordinal $\leq \omega_1^{\mathrm{CK}}$, however saying something more precise on this point would require to deal with some complicated and ultimately not relevant issues on ordinal notations.

[67] Admittedly, the ŁK-construction does not overcome limitation (K1), namely the absence of schematic laws: I argued that this is an acceptable price to pay at the end of Section 5.

these limitations are addressed via a conditional that has a conceptually interesting reading, following the aim set forth in the **Main Question**.

As to (K2), note that whenever φ is a $\mathscr{L}_T$-sentence or is a $\mathscr{L}_T^{\rightarrow}$-sentence derived from a Kripkean partition given by some A, B, and $C^{\mathrm{A,B}}$, the identity of value between φ and $T\ulcorner\varphi\urcorner$ is preserved in $\langle A_\infty, B_\infty, C_\infty^{\mathrm{A,B}}\rangle$ and internalized via $\rightarrow$, as the corresponding Tarski biconditional $\varphi \leftrightarrow T\ulcorner\varphi\urcorner$ is in A_∞, i.e. it receives value 1. So, the equivalence between φ and $T\ulcorner\varphi\urcorner$ is validated by the theory for every sentence targeted by Kripke's theory plus all the results of truth-value comparisons between them. As to (K3) and (K4), the paradoxes that are relevant for the original theory by Kripke for the language $\mathscr{L}_T$ are also accommodated nicely by fixed points of the form $\langle A_\infty, B_\infty, C_\infty^{\mathrm{A,B}}\rangle$. Not only it is possible to classify them explicitly as P-gappy in the fixed-point triples of Υ, we can also classify explicitly all the sentences in the resulting set $C_\infty^{\mathrm{A,B}}$ as not determinately true. Finally, the treatment of the liar paradox does justice to the intuition that it is equivalent to its own negation, since every fixed point of the form $\langle A_\infty, B_\infty, C_\infty^{\mathrm{A,B}}\rangle$ validates the sentence $\lambda \leftrightarrow \neg\lambda$, which formalizes the statement (G) mentioned in Section 2.[68]

An important element of this class of fixed points is $\mathfrak{I}_\Upsilon(\emptyset,\emptyset,C^{\emptyset,\emptyset})$, i.e. the fixed point of Υ applied to the $\leq_S$-least Kripke fixed point and its valueless sentences of $\mathscr{L}_T$. For $A,B \subseteq SENT_{\mathscr{L}_T}$, if $\mathfrak{I}_\Phi(A,B)$ is consistent, $C^{\emptyset,\emptyset} \supseteq C^{\mathrm{A,B}}$, so this fixed point exploits in full a very rich assumption about sentences given value $1/2$.

A derivative case of this kind of application is obtained replacing $C^{\mathrm{A,B}}$ with

$$C^{\mathrm{PKF}} := \{\varphi \in \mathscr{L}_T \,|\, \mathsf{PKF} \vdash \varphi \Leftrightarrow \neg\varphi\}.^{69}$$

Note that $\mathfrak{I}_\Upsilon(\emptyset,\emptyset,C^{\mathrm{PKF}})$ is consistent, and that the results holding for fixed points of the form $\langle A_\infty, B_\infty, C_\infty^{\mathrm{A,B}}\rangle$ hold for the above one too. Moreover, $\mathfrak{I}_\Upsilon(\emptyset,\emptyset,C^{\mathrm{PKF}})$ is purely inductive, since C^{PKF} is recursively enumerable. This shows that it is possible to have a paracomplete theory of naïve truth (featuring a conceptually interesting conditional) that improves on Kripke's theory *and* that is quite simple. At the same time, the use of sets such as C^{PKF} prompts us to develop *calculi* to determine positive assertions of gappiness, that could be interpreted by models such as the ŁK-construction.[70]

[68] Feferman [13] notes that the consistency of the biconditionals $\varphi \leftrightarrow T\ulcorner\varphi\urcorner$ for the Łukasiewicz 3-valued conditional *restricted to* $\varphi \in \mathscr{L}_T$ can be proven along the lines of the consistency proof of similarly restricted instances of naïve comprehension given by Brady [6]. This result, however, is immediate from Proposition 34, which gives a very simple proof of it. Feferman urges also for some improvement on Brady's unsatisfactory restriction. He proposes to supplement classical logic with an *intensional* conditional $\rightarrow^\star$ and its standardly defined biconditional $\leftrightarrow^\star$, s.t. all $\leftrightarrow^\star$-Tarski biconditionals hold (see Aczel and Feferman [1]). The semantics of $\rightarrow^\star$, however, is quite disappointing: e.g., not even *modus ponens* is validated for it (Feferman [13], §11). Fixed points of the form $\langle A_\infty, B_\infty, C_\infty^{\mathrm{A,B}}\rangle$ may be seen as an answer to Feferman's instigation to recover more Tarski biconditionals, with a conceptually motivated restriction of the excluded instances (also less draconian than Brady's one), and a better-behaved conditional.

[69] For the theory PKF, see Halbach and Horsten [24] (see also footnote 37).

[70] This would be interesting in order to associate a logic to some collection of fixed points of Υ (by Corollary 30, presumably we do not want to consider *all* the fixed points of Υ). In the case of consistent fixed points extending $\mathfrak{I}_\Upsilon(\emptyset,\emptyset,C^{\mathrm{PKF}})$, for example, it is clear how to describe such calculus. The arithmetical and truth-theoretic parts of the theory would be given by the axioms and rules of PKF, and its logic would consist of all the axioms and rules of PKF for $\neg$, $\wedge$, $\vee$, and $\forall$ plus, for the conditional, *modus ponens* and the following rule for introducing the conditional: the theory proves $\Rightarrow \varphi \rightarrow \psi$ whenever the theory proves $\varphi \Rightarrow \psi$ and one of the following is the case: (i) the theory proves $\Rightarrow \varphi \vee \neg\varphi$, (ii) the theory proves $\Rightarrow \psi$, (iii) the theory proves both $\varphi \Leftrightarrow \neg\varphi$ and $\psi \Leftrightarrow \neg\psi$.

References

1. Aczel, Peter and Solomon Feferman, Consistency of the unrestricted abstraction principles using an intensional equivalence operator, in J. P. Seldin and J. R. Hindley (Eds.), *To H. B. Curry: Essays on combinatory logic, lambda calculus and formalism*, 67-98, Academic Press, New York (1980).
2. Bacon, Andrew, A New Conditional for Naive Truth Theory, *Notre Dame Journal of Formal Logic*, 54, 1, 87-104 (2013).
3. Beall, JC, *Spandrels of Truth*, Oxford University Press, Oxford (2009).
4. Belnap, Nuel D. Jr., A useful four-valued logic, in J. Michael Dunn and George Epstein (Eds.), *Modern Uses of Multiple-Valued Logic*, 8-37, Reidel, Dordrecht (1977).
5. Blamey, Stephen, Partial Logic, in Dov M. Gabbay and Franz Guenthner (Eds.), *Handbook of Philosophical Logic*, second edition, volume 5, 261-353, Kluwer Academic Publishers, Dordrecht, Boston (2002).
6. Brady, Ross, The consistency of the axioms of abstraction and extensionality in three-valued logic, *Notre Dame Journal of Formal Logic*, 12, 4, 447-453 (1971).
7. Brady, Ross, The Non-Triviality of Dialectical Set Theory, in Graham Priest, Richard Routley and Jean Norman (Eds.), *Paraconsistent Logic: Essays on the Inconsistent*, 437-470, Philosophia Verlag, München (1989).
8. Burgess, John P., Friedman and the axiomatization of Kripke's theory of truth, in Neil Tennant (Ed.), *Foundational Adventures. Essays in Honour of Harvey Friedman*, College Publications, London (2009).
9. Cantini, Andrea, *Logical Frameworks for Truth and Abstraction. An Axiomatic Study*, volume 135 of *Studies in Logic and the Foundations of Mathematics*, Elsevier, Amsterdam (1996).
10. Cantini, Andrea, Osservazioni su Autoriferimento e Verità, *Annali del Dipartimento di Filosofia dell'Università di Firenze*, VIII, 53-76 (2002).
11. Cook, Roy T., Embracing Revenge: On the Indefinite Extensibility of Language, in JC Beall (Ed.), *The Revenge of the Liar. New Essays on the Paradox*, 31-52, Oxford University Press, Oxford (2007).
12. Dummett, Michael E.A., Truth, *Proceedings of the Aristotelian Society*, New Series, 59 (1958-1959),141-162 (1958).
13. Feferman, Solomon, Toward Useful Type-Free Theories. I, *The Journal of Symbolic Logic*, 49, 1, 75-111 (1984).
14. Field, Hartry, Saving the Truth Schema from Paradox, *Journal of Philosophical Logic*, 31, 1, 1-27 (2002).
15. Field, Hartry, A Revenge-Immune Solution to the Semantic Paradoxes, *Journal of Philosophical Logic*, 32, 2, 139-177 (2003).
16. Field, Hartry, The Consistency of the Naïve Theory of Properties, *The Philosophical Quarterly*, 54, 214, 78-104 (2004).
17. Field, Hartry, Solving the Paradoxes, Escaping Revenge, in JC Beall (Ed.), *The Revenge of the Liar. New Essays on the Paradox*, cit., 78-144 (2007).
18. Field, Hartry, *Saving Truth from Paradox*, Oxford University Press, Oxford (2008).
19. Field, Hartry, Comments on Martin's and Welch's Comments, *The Review of Symbolic Logic*, 4, 3, 360-366 (2011).
20. Field, Hartry, Naive Truth And Restricted Quantification: Saving Truth A Whole Lot Better, *Review of Symbolic Logic*, 7, 147-91 (2014).
21. Fischer, Martin, Volker Halbach, Jönne Speck and Johannes Stern, Axiomatizing Semantic Theories of Truth?, *The Review of Symbolic Logic*, 8, 257-278 (2015).
22. Gottwald, Siegfried, *A Treatise on Many-Valued Logics*, Studies in Logic and Computation, volume 9, Research Studies Press, Baldock (2001).
23. Gupta, Anil and Nuel D. Belnap Jr., *The Revision Theory of Truth*, MIT Press, Cambridge, Massachusetts (1993).
24. Halbach, Volker and Leon Horsten, Axiomatizing Kripke's Theory of Truth, *The Journal of Symbolic Logic*, 71, 2, 677-712 (2006).
25. Halbach, Volker, *Axiomatic Theories of Truth*, Cambridge University Press, Cambridge (2011).
26. Horsten, Leon, *The Tarskian Turn*, MIT Press, Cambridge, Massachusetts (2011).
27. Kaye, Richard, *Models of Peano Arithmetic*, Oxford University Press, Oxford (1991).
28. Kremer, Michael, Kripke and the Logic of Truth, *Journal of Philosophical Logic*, 17, 3, 327-332 (1988).
29. Kripke, Saul, Outline of a Theory of Truth, *The Journal of Philosophy*, 72, 19, 690-716 (1975).

30. Leitgeb, Hannes, On the Metatheory of Field's "Solving the Paradoxes, Escaping Revenge", in JC Beall (Ed.), *The Revenge of the Liar. New Essays on the Paradox*, cit., 159-183, Oxford University Press, Oxford (2007).
31. Malinowski, Grzegorz, *Many-Valued Logics*, Oxford Logic Guides, 25, Clarendon Press, Oxford (1993).
32. Martin, Donald A., Field's *Saving Truth from Paradox*: Some Things It Doesn't Do, *The Review of Symbolic Logic*, 4, 3, 339-347 (2011).
33. McGee, Vann, Applying Kripke's Theory of Truth, *The Journal of Philosophy*, 86, 10, 530-539 (1989).
34. McGee, Vann, *Truth, Vagueness, and Paradox: An Essay on the Logic of Truth*, Hackett Publishing Company, Indianapolis and Cambridge (Massachusetts) (1991).
35. Moschovakis, Yiannis N., *Elementary Induction on Abstract Structures*, North-Holland and Elsevier, Amsterdam, New York (1974), unabridged republication by Dover Books, New York (2008).
36. Restall, Greg, Arithmetic and Truth in Łukasiewicz's Infinitely Valued Logic, *Logique Et Analyse*, 139 (140), 303-312 (1992).
37. Rossi, Lorenzo, Graphs, Truth, and Conditional(s), under review.
38. Soames, Scott, *Understanding Truth*, Oxford University Press, New York and Oxford (1999).
39. Tarski, Alfred, The Concept of Truth in Formalized Languages, Polish original version 1933, German translation 1935, English translation of the German version in Alfred Tarski, *Logic, Semantics, Metamathematics*, second edition, John Corcoran (Ed.), Hackett Publishing Company, Indianapolis (1983).
40. van Dalen, Dirk, *Logic and Structure*, Springer, London Heidelberg New York Dordrecht (2012).
41. Welch, Philip, Ultimate truth *vis à vis* stable truth, in *The Review of Symbolic Logic*, 1, 1, 126-142 (2008).
42. Welch, Philip, Truth, Logical Validity and Determinateness: A Commentary on Field's *Saving Truth from Paradox*, *The Review of Symbolic Logic* 4, 3, 248-259 (2011).
43. Welch, Philip, Some Observations on Truth Hierarchies, *The Review of Symbolic Logic*, 7, 1, 1-30 (2014).
44. Yablo, Stephen, Grounding, Dependence, and Paradox, *Journal of Philosophical Logic*, 11, 1, 117-137 (1982).