Dialectica logical principles: Not only rules

DAVIDE TROTTA, *Department of Computer Science, University of Pisa, Piano Secondo, Largo Bruno Pontecorvo, 3, 56127 Pisa PI, Italy.*

MATTEO SPADETTO, *School of Mathematics, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, United Kingdom.*

VALERIA DE PAIVA, *Topos Institute, 2140 Shattuck Ave, Suite 610 Berkeley, CA 94704, USA.*

Abstract

Gödel's Dialectica interpretation was designed to obtain the consistency of Peano arithmetic via a proof of consistency of Heyting arithmetic and double negation. In recent years, proof theoretic transformations (the so-called proof interpretations) based on Gödel's Dialectica interpretation have been used systematically to extract new content from proofs and so the interpretation has found relevant applications in several areas of mathematics and computer science. Following our previous work on 'Gödel fibrations', we present a (hyper)doctrine characterization of the Dialectica, which corresponds exactly to the logical description of the interpretation. To show that, we derive the soundness of the interpretation of the implication connective, as expounded on by Spector and Troelstra, in the categorical model. This requires extra logical principles, going beyond intuitionistic logic, namely Markov Principle and the Independence of Premise principle, as well as some choice. We show how these principles are satisfied in the categorical setting, establishing a tight (internal language) correspondence between the logical system and the categorical framework. We make sure that this tight correspondence extends to the use of the principles above, instead of the weaker rules we had proved earlier on. This tight correspondence should come handy not only when discussing the traditional applications of the Dialectica but also when dealing with newer uses in modelling games or concurrency theory.

1 Introduction

Categorical logic is the branch of mathematics in which tools and concepts from category theory are applied to the study of mathematical logic and its connections to theoretical computer science. In broad terms, categorical logic represents both syntax and semantics by a category, and an interpretation by a functor. The categorical framework provides a rich conceptual background for logical and type-theoretic constructions. In many cases, the categorical semantics of a logic provides a basis for establishing a correspondence between theories in the logic and instances of an appropriate kind of category. A classic example is the correspondence between theories of $\beta\eta$-equational logic over simply typed lambda calculus and Cartesian closed categories. Categories arising from theories via term-model constructions can usually be characterized up to equivalence by a suitable universal property. This has enabled proofs of meta-theoretical properties of logics by means of an appropriate categorical algebra. One defines a suitable internal language naming relevant constituents of a category, and then applies categorical semantics to turn assertions in a logic over the internal language into corresponding categorical statements. The goal is to obtain 'internal language theorems' that allow us to pass freely from the logic/type theory to the categorical universe, in such a way that we can solve issues in whichever framework is more appropriate.

Several kinds of categorical universe are available. Our previous joint work [30] on Gödel's Dialectica Interpretation [6] used the fibrational framework expounded by Jacobs [10]. The

Vol. 32, No. 8,
Advance Access Publication on 8 November 2022 https://doi.org/10.1093/logcom/exac079

identification of syntax-free notions of quantifier-free formulae using categorical concepts is the key insight to our results preliminarily described in [31]. This identification, besides explaining how Gödel's Dialectica interpretation works as a double completion under products and coproducts, is itself of independent interest, as it deepens our ability to think about first-order logic, using categorical notions. There are countless theorems in logic involving quantifier-free elements. Without a categorical notion of quantifier-free element, we would not be able to properly state or write such theorems in categorical terms. For example, we would not be able to write Markov Principle categorically, and then we could not present categorically any theorem involving Markov Principle. Therefore, having an algebraic presentation of these elements (at least in the Dialectica context) is fundamental to present and reason categorically about traditional theorems and results of logic.

Wishing to simplify the treatment of our previous work on generalized Gödel fibrations, we concentrated on poset-based fibrations. This led us to more perspicuous explanations of how the logical interpretation works in the simplified categorical setting, which we wrote in [32]. In that paper, we show that the notions of existential-free and universal-free elements introduced in our categorical setting correspond to well-known (non-intuitionistic but) constructive rules underlying Gödel's Dialectica interpretation. Then, we employ our modelling from [32] to show under which hypotheses it is indeed the case that a Gödel hyperdoctrine is a model of the Principle of Independence of Premises, Markov Principle and the Generalized Markov Principle themselves. While the rule-versions of these principles do not require any additional hypothesis for a given Gödel hyperdoctrine to be satisfied, we show that the logical principles require that quantifier-free elements have to be closed under Heyting operations to be satisfied. This closure condition under Heyting operations for quantifier-free elements appears quite natural from a logical perspective, since, e.g. the conjunction or the disjunction of quantifier-free elements is always a quantifier-free element in logic. These considerations improve our previous results in [32] showing that a Gödel Hyperdoctrine satisfying some additional (but quite natural) hypotheses is a model of the Rule of Independence of Premises and the (Generalized) Markov rule only.

The validity of the Principle of Independence of Premises, (Generalized) Markov Principle and the Principle of Skolemization in every Gödel hyperdoctrine satisfying this closure condition with respect to Heyting operations allows us to conclude that the key equivalence $(\psi \to \phi)^D \leftrightarrow (\psi^D \to \phi^D)$ motivating the translation of the implicational connective in the Dialectica interpretation is satisfied. This result can be considered a strengthening of our previous main theorem in [32], as this can now be obtained as an application of our new result.

2 Logical principles in Dialectica

Gödel's Dialectica interpretation [5, 6] associates to each formula ϕ in the language of arithmetic its *Dialectica interpretation* ϕ^D, i.e. a formula of the form:

$$\phi^D = \exists u.\forall x.\phi_D,$$

where ϕ_D is a quantifier-free formula in the language of system T, trying to be *as constructive as possible*. The associations $(-)^D$ and $(-)_D$ are defined inductively on the structure of the formulae, and we refer to [5, 6] for a complete description. The most complicated clause of the translation (and, in Gödel's words, 'the most important one') is the definition of the translation of the implication connective $(\psi \to \phi)^D$. This involves two logical principles that are usually not acceptable from an intuitionistic point of view, namely a form of the *Principle of Independence of Premise* (IP) and a generalization of *Markov Principle* (MP). The interpretation is given by:

$$(\psi \to \phi)^D = \exists V, X.\forall u, y.(\psi_D(u, X(u,y)) \to \phi_D(V(u), y)).$$

The motivation provided in the collected works of Gödel for this translation is that given a witness u for the hypothesis ψ_D, one should be able to obtain a witness for the conclusion ϕ_D, i.e. there exists a function V assigning a witness $V(u)$ of ϕ_D to every witness u of ψ_D. Moreover, this assignment has to be such that from a counterexample y of the conclusion ϕ_D, we should be able to find a counterexample $X(u,y)$ to the hypothesis ψ_D. This transformation of counterexamples of the conclusion into counterexamples for the hypothesis is what gives Dialectica its essential bidirectional character.

We first recall the technical details behind the translation of $(\psi \to \phi)^D$ ([6]) showing the precise points in which we have to employ the non-intuitionistic principles (MP) and (IP). First notice that $\psi^D \to \phi^D$, i.e.:

$$\exists u.\forall x.\psi_D(u,x) \to \exists v.\forall y.\phi_D(v,y) \tag{1}$$

is equivalent to:

$$\forall u.(\forall x.\psi_D(u,x) \to \exists v.\forall y.\phi_D(v,y)). \tag{2}$$

If we apply a special case of the **Principle of Independence of Premise**, namely:

$$(\forall x.\theta(x) \to \exists v.\forall y.\eta(v,y)) \to \exists v.(\forall x.\theta(x) \to \forall y.\eta(v,y)), \tag{IP*}$$

we obtain that (2) is equivalent to:

$$\forall u.\exists v.(\forall x.\psi_D(u,x) \to \forall y.\phi_D(v,y)). \tag{3}$$

Moreover, we can see that this is equivalent to:

$$\forall u.\exists v.\forall y.(\forall x.\psi_D(u,x) \to \phi_D(v,y)). \tag{4}$$

The next equivalence is motivated by a generalization of **Markov's Principle**, namely:

$$\neg\forall x.\theta(u,x) \to \exists x.\neg\theta(u,x). \tag{MP}$$

By applying (MP), we obtain that (4) is equivalent to:

$$\forall u.\exists v.\forall y.\exists x.(\psi_D(u,x) \to \phi_D(v,y)). \tag{5}$$

To conclude that $\psi^D \to \phi^D = (\psi \to \phi)^D$, we have to apply the **Axiom of Choice** (or **Skolemization**), i.e.:

$$\forall y.\exists x.\theta(y,x) \to \exists V.\forall y.\theta(y, V(y)), \tag{AC}$$

twice, obtaining that (5) is equivalent to:

$$\exists V, X.\forall u, y.(\psi_D(u, X(u,y)) \to \phi_D(V(u), y)).$$

This analysis (from Gödel's Collected Works, page 231) highlights the key role the principles (IP), (MP) and (AC) play in the Dialectica interpretation of implicational formulae. The role of the axiom of choice (AC) has been discussed from a categorical perspective both by Hofstra [8] and in our previous work [31]. We re-examine the two principles (IP) and (MP) in the next subsections, following what we discussed in [31].

2.1 Independence of Premise

In logic and proof theory, the **Principle of Independence of Premise** states that:

$$(\theta \to \exists u.\eta(u)) \to \exists u.(\theta \to \eta(u)),$$

where u is not a free variable of θ. While this principle is valid in classical logic (it follows from the law of the excluded middle), it does not hold in intuitionistic logic, and it is not generally accepted constructively [2]. The reason why the principle (IP) is not generally accepted constructively is that, from a constructive perspective, turning any proof of the premise θ into a proof of $\exists u.\eta(u)$ means turning a proof of θ into a proof of $\eta(t)$ where t is a witness for the existential quantifier depending on the proof of θ. In particular, the choice of the witness *depends* on the proof of the premise θ, while the (IP) principle tell us, constructively, that the witness can be chosen independently of any proof of the premise θ.

In the Dialectica translation, we only need a particular version of the (IP) principle:

$$(\forall y.\theta(y) \to \exists u.\forall v.\eta(u,v)) \to \exists u.(\forall y.\theta(y) \to \forall v.\eta(u,v)), \tag{IP*}$$

which means that we are asking (IP) to hold not for every formula, but only for those formulas of the form $\forall y.\theta(y)$ with θ quantifier-free. We recall a useful generalization of the (IP*) principle, namely:

$$(\theta \to \exists u.\eta(u)) \to \exists u.(\theta \to \eta(u)), \tag{IP}$$

where θ is existential-free, i.e. θ does not contain existential quantifiers (of course, it is also assumed that u is not a free variable of θ). Therefore, the condition that (IP) holds for every formula of the form $\forall y.\theta(y)$ with $\theta(y)$ quantifier-free is replaced by asking that it holds for every *existential-free formula*.

A similar formulation of (IP) is introduced in [21] where, starting from the observation that intuitionistic finite-type arithmetic is closed under the independence of premise rule (IPR) for $\exists$-free formula, i.e. formulae that neither contain existential quantifiers nor disjunctions, it is proved that a similar result holds for many set theories including Constructive Zermelo–Fraenkel Set Theory (CZF) and Intuitionistic Zermelo–Fraenkel Set Theory (IZF).

The **Independence of Premise Rule** for existential-free formula (IPR) that we use in this paper states that:

$$\text{if } \vdash \theta \to \exists u.\eta(u)) \text{ then } \vdash \exists u.(\theta \to \eta(u)), \tag{IPR}$$

where θ is existential-free.

2.2 Markov Principle

Markov Principle is a statement that originated in the Russian school of constructive mathematics. Formally, Markov's principle is usually presented as the statement:

$$\neg\neg\exists x.\phi(x) \to \exists x.\phi(x),$$

where ϕ is a quantifier-free formula. Thus, MP in the Dialectica interpretation, namely:

$$\neg\forall x.\phi(x) \to \exists x.\neg\phi(x), \tag{MP}$$

with $\phi(x)$ a quantifier-free formula, can be thought of as a generalization of the Markov Principle above. As remarked in [2], the reason why MP is not generally accepted in constructive mathematics is that in general there is no reasonable way to choose constructively a witness x for $\neg\phi(x)$ from a proof that $\forall x.\phi(x)$ leads to a contradiction. However, in the context of Heyting Arithmetic, i.e. when

x ranges over the natural numbers, one can prove that these two formulations of Markov Principle are equivalent. More details about the computational interpretation of Markov Principle can be found in [19]. We recall the version of **Markov's Rule** (MR) corresponding to Markov Principle:

$$\text{if } \vdash \neg\forall x.\phi(x) \text{ then } \vdash \exists x.\neg\phi(x), \tag{MR}$$

where $\phi(x)$ is a quantifier-free formula.

3 Logical doctrines

One of the most relevant notions of categorical logic that enabled the study of logic from a pure algebraic perspective is that of a *hyperdoctrine*, introduced in a series of seminal papers by F.W. Lawvere to synthesize the structural properties of logical systems [11–13]. Lawvere's crucial intuition was to consider logical languages and theories as fibrations to study their 2-categorical properties, e.g. connectives, quantifiers and equality are determined by structural adjunctions. Recall from [11, 23] that a *hyperdoctrine* is a functor:

$$P \colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Hey}$$

from the opposite of a Cartesian closed category $\mathcal{C}$ to the category of Heyting algebras **Hey** satisfying some further conditions: for every arrow $A \xrightarrow{f} B$ in $\mathcal{C}$, the homomorphism $P_f \colon P(B) \longrightarrow P(A)$ of Heyting algebras, where P_f denotes the action of the functor P on the arrow f, has a left adjoint $\exists_f$ and a right adjoint $\forall_f$ satisfying the Beck–Chevalley conditions. The intuition is that a hyperdoctrine determines an appropriate categorical structure to abstract both notions of first order theory and of interpretation.

Semantically, a hyperdoctrine is essentially a generalization of the contravariant *powerset functor* on the category of sets:

$$\mathcal{P} \colon \mathbf{Set}^{\mathrm{op}} \longrightarrow \mathbf{Hey}$$

sending any set-theoretic arrow $A \xrightarrow{f} B$ to the inverse image functor:

$$\mathcal{P}B \xrightarrow{\mathcal{P}f = f^{-1}} \mathcal{P}A.$$

However, from the syntactic point of view, a hyperdoctrine can be seen as the generalization of the so-called *Lindenbaum–Tarski algebra* of well-formed formulae of a first order theory. In particular, given a first order theory $\mathcal{T}$ in a first order language $\mathcal{L}$, one can consider the functor:

$$\mathcal{LT} \colon \mathcal{V}^{\mathrm{op}} \longrightarrow \mathbf{Hey}$$

whose base category $\mathcal{V}$ is the *syntactic* category of $\mathcal{L}$, i.e. the objects of $\mathcal{V}$ are finite lists $\overrightarrow{x} := (x_1, \dots, x_n)$ of variables and morphisms are lists of substitutions, while the elements of $\mathcal{LT}(\overrightarrow{x})$ are given by equivalence classes (with respect to provable reciprocal consequence $\dashv\vdash$) of well-formed formulae in the context $\overrightarrow{x}$, and order is given by the provable consequences with respect to the fixed theory $\mathcal{T}$. Notice that in this case an existential left adjoint to the weakening functor $\mathcal{LT}_\pi$ is computed by quantifying existentially the variables that are not involved in the substitution given by the projection (by duality the right adjoint is computed by quantifying universally).

Recently, several generalizations of the notion of a Lawvere hyperdoctrine were considered, and we refer e.g. to [15–17] or to [9, 24] for higher-order versions. In this work, we consider a natural generalization of the notion of hyperdoctrine, and we call it simply a *doctrine*. A **doctrine** is just a functor:

$$P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos},$$

where the category $\mathcal{C}$ has finite products and **Pos** is the category of posets.

Depending on the categorical properties enjoyed by P, we get P to model the corresponding fragments of first order logic formally in a way identical to the one for $\mathcal{P}$, which we call a *generalized Tarski semantics* and which continues to be complete. Again, the syntactic intuition behind the notion of doctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ remains the same, one should think of $\mathcal{C}$ as the category of contexts associated to a given type theory. Given such a context A, the elements of the posets $P(A)$ represent the predicates in context A and the order relation of $P(A)$ represents the relation of syntactic provability (with respect to the fragment of first order logic modelled by P). Arrows $B \xrightarrow{f} A$ of $\mathcal{C}$ represent (finite lists of) terms-in-context:

$$b : B \mid f(b) : A$$

in such a way that the functor P_f models the substitution by the (finite list of) term(s) f. For instance, if α element of PA represents a formula in context $a : A \mid \alpha(a)$, then the element $P_f(\alpha)$ of $P(B)$ represents the formula $b : B \mid \alpha(f(b))$ in context B obtained by substituting f into α. Notice that we follow the notation used, e.g. in [18] to denote a term-in-context and a formula-in-context. In particular, we write $b : B \mid f(b) : A$ for a term $f(b)$ of sort A in the context $b : B$, and similarly we write $a : A \mid \alpha(a)$ to denote a formula $\alpha(a)$ in the context $a : A$.

Now we recall from [15, 16, 27] the notions of existential and universal doctrines, and we refer to [23] for a detailed introduction to the theory of doctrines and hyperdoctrines. For further insights and applications to higher-order logic or realizability, we refer to [9, 24, 33].

DEFINITION 3.1
A doctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ is **existential** (resp. **universal**) if, for every A_1 and A_2 in $\mathcal{C}$ and every projection $A_1 \times A_2 \xrightarrow{\pi_i} A_i$, $i = 1, 2$, the functor:

$$PA_i \xrightarrow{P_{\pi_i}} P(A_1 \times A_2)$$

has a left adjoint $\exists_{\pi_i}$ (resp. a right adjoint $\forall_{\pi_i}$), and these satisfy the **Beck–Chevalley condition**: for any pullback diagram:

$$\begin{array}{ccc} X' & \xrightarrow{\pi'} & A' \\ {\scriptstyle f'}\big\downarrow & & \big\downarrow{\scriptstyle f} \\ X & \xrightarrow[\pi]{} & A \end{array}$$

with π and π' projections, for any β in $P(X)$ the equality:

$$\exists_{\pi'} P_{f'} \beta = P_f \exists_{\pi} \beta \text{ (resp. } \forall_{\pi'} P_{f'} \beta = P_f \forall_{\pi} \beta \text{)}$$

holds (however, observe that the inequality $\exists_{\pi'} P_{f'} \beta \leq P_f \exists_{\pi} \beta$ (resp. $\forall_{\pi'} P_{f'} \beta \geq P_f \forall_{\pi} \beta$) always holds).

If a doctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ is existential and $\alpha \in P(A \times B)$ is a formula-in-context $a : A, b : B \mid \alpha(a, b)$ and $A \times B \xrightarrow{\pi_A} A$ is the product projection on the component A, then $\exists_{\pi_A} \alpha \in PA$

represents the formula $a : A \mid \exists b : B.\alpha(a, b)$ in context A. Analogously, if the doctrine P is universal, then $\forall_{\pi_A}\alpha \in PA$ represents the formula $a : A \mid \forall b : B.\alpha(a, b)$ in context A. This interpretation is sound and complete for the usual reasons: this is how classic Tarski semantics can be characterized in terms of categorical properties of the powerset functor $\mathcal{P} \colon \mathbf{Set}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$.

One of the most interesting aspects of this categorical approach to logic is that there is categorical equivalence between logical theories and doctrines, via the so-called *internal language* of a doctrine [14, 23]. The internal language of a doctrine P essentially constitutes a syntax endowed with a semantics induced by P itself: there is a way to interpret every sequent in the fragment of first-order logic modelled by P into a categorical statement involving P. This interpretation is sound and complete; this is precisely why we can deduce properties of P through a purely syntactical procedure. We define the following notation for this syntax, taking advantage of these equivalent ways of reasoning about doctrines and logic.

Notation. From now on, we shall employ the logical language provided by the *internal language* of a doctrine and write:

$$a_1 : A_1, \ldots, a_n : A_n \mid \phi(a_1, \ldots, a_n) \vdash \psi(a_1, \ldots, a_n)$$

instead of:

$$\phi \leq \psi$$

in the fibre $P(A_1 \times \cdots \times A_n)$. Similarly, we write:

$$a : A \mid \phi(a) \vdash \exists b : B.\psi(a, b) \text{ and } a : A \mid \phi(a) \vdash \forall b : B.\psi(a, b)$$

in place of:

$$\phi \leq \exists_{\pi_A}\psi \text{ and } \phi \leq \forall_{\pi_A}\psi$$

in the fibre $P(A)$. Also, we write $a : A \mid \phi \dashv\vdash \psi$ to abbreviate $a : A \mid \phi \vdash \psi$ and $a : A \mid \psi \vdash \phi$. Substitutions via given terms (i.e. reindexings and weakenings) are modelled by pulling back along those given terms. Applications of propositional connectives are interpreted by using the corresponding operations in the fibres of the given doctrine. Finally, when the type of a quantified variable is clear from the context, we will omit the type for the sake of readability.

4 Logical principles via universal properties

It is possible to characterize, in terms of weak universal properties, those predicates of a doctrine that are free from a quantifier. In the following definitions, we pursue this idea of defining those elements of an existential doctrine $P \colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ that are *free from left adjoints* $\exists_\pi$. This idea was originally introduced in [28] and, independently, in [4], and then further developed and generalized in the fibrational setting in [31].

DEFINITION 4.1

Let $P \colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ be an existential doctrine and let A be an object of $\mathcal{C}$. A predicate α of the fibre $P(A)$ is said to be an **existential splitting** if it satisfies the following weak universal property: for every projection $A \times B \xrightarrow{\pi_A} A$ of $\mathcal{C}$ and every predicate β of $P(A \times B)$ such that $\alpha \leq \exists_{\pi_A}(\beta)$, there exists an arrow $A \xrightarrow{g} B$ such that:

$$\alpha \leq P_{\langle 1_A, g\rangle}(\beta).$$

Notice that in Definition 4.1 we require the existence of an arrow, but not the uniqueness. This is due to the fact that doctrines are a proof-irrelevant version of fibration, and this is the reason why in fibrational-version of Definition 4.1 presented in [31] we have to require the unicity of the arrow, while in the context of doctrines we do not. Existential splittings stable under re-indexing are called *existential-free elements*. Thus we introduce the following definition:

Definition 4.2
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ be an existential doctrine and let I be an object of $\mathcal{C}$. A predicate α of the fibre $P(I)$ is said to be **existential-free** if $P_f(\alpha)$ is an existential splitting for every morphism $A \xrightarrow{f} I$.

Employing the presentation of doctrines via internal language, we require of the formula $i : I \mid \alpha(i)$ to be free from the existential quantifier that, whenever $a : A \mid \alpha(f(a)) \vdash \exists b : B.\beta(a,b)$ for some term $a : A \mid f(a) : I$, there is a term $a : A \mid g(a) : B$ such that $a : A \mid \alpha(f(a)) \vdash \beta(a, g(a))$.

Observe that in general we always have that $a : A \mid \beta(a, g(a)) \vdash \exists b : B.\beta(a,b)$, in other words $P_{\langle 1_A, g\rangle}\beta \leq \exists_{\pi_A}\beta$. In fact, it is the case that $\beta \leq P_{\pi_A}\exists_{\pi_A}\beta$ (as this arrow of $P(A \times B)$ is nothing but the unit of the adjunction $\exists_{\pi_A} \dashv P_{\pi_A}$), hence a re-indexing by the term $\langle 1_A, g\rangle$ yields the desired inequality. Therefore, the property that we are requiring for $i : I \mid \alpha(i)$ turns out to be the following: whenever there are proofs of $\exists b : B.\beta(a,b)$ from $\alpha(f(a))$, at least one of them factors through the canonical proof of $\exists b : B.\beta(a,b)$ from $\beta(a, g(a))$ for some term $a : A \mid g(a) : B$. This fact implies that, while freely adding the existential quantifiers to a doctrine, we do not add a new sequent $\alpha \vdash \exists b.\beta(b)$ (where α and $\beta(b)$ are predicates in the doctrine we started from) as long as we do not allow a sequent $\alpha \vdash \beta(g)$ as well, for some term g (see [29] for more details). For the proof-relevant versions of this definition, we refer to [31].

Definition 4.3
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ be an existential doctrine. Then we indicate by $P^{\exists\text{-free}}\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ the subdoctrine of P whose elements of the fibres $P^{\exists\text{-free}}(A)$ are existential-free element of $P(A)$.

We dualize the previous Definitions 4.1 and Definition 4.2 to get the corresponding ones for the universal quantifier.

Definition 4.4
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ be a universal doctrine and let A be an object of $\mathcal{C}$. A predicate α of the fibre $P(A)$ is said to be a **universal splitting** if it satisfies the following weak universal property: for every projection $A \times B \xrightarrow{\pi_A} A$ of $\mathcal{C}$ and every predicate β of $P(A \times B)$ such that $\forall_{\pi_A}(\beta) \leq \alpha$, there exists an arrow $A \xrightarrow{g} B$ such that:

$$P_{\langle 1_A, g\rangle}(\beta) \leq \alpha.$$

Definition 4.5
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ be a universal doctrine and let I be an object of $\mathcal{C}$. A predicate α of the fibre $P(I)$ is said to be **universal-free** if $P_f(\alpha)$ is a universal splitting for every morphism $A \xrightarrow{f} I$.

The property we require of the formula $i : I \mid \alpha(i)$, so that it is free from the universal quantifiers, is that, whenever $a : A \mid \forall b : B.\beta(a,b) \vdash \alpha(f(a))$ for some term $a : A \mid f(a) : I$, then there is a term $a : A \mid g(a) : B$ such that $a : A \mid \beta(a, g(a)) \vdash \alpha(f(a))$.

We can present the dual notion of Definition 4.3.

DEFINITION 4.6
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos** be an universal doctrine. Then we indicate by $P^{\forall\text{-free}}\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos** the subdoctrine of P whose elements of the fibres $P^{\forall\text{-free}}(A)$ are universal-free element of $P(A)$.

DEFINITION 4.7
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos** be a doctrine. If P is existential, we say that P has **enough existential-free predicates** if, for every object I of $\mathcal{C}$ and every predicate α of PI, there exist an object A and an existential-free object β in $P(I \times A)$ such that $\alpha = \exists_{\pi_I}\beta$.

Analogously, if P is universal, we say that P has **enough universal-free predicates** if, for every object I of $\mathcal{C}$ and every predicate α of PI, there exist an object A and a universal-free object β in $P(I \times A)$ such that $\alpha = \forall_{\pi_I}\beta$.

Now we can introduce a particular kind of doctrine called a *Gödel doctrine*. This definition works as a synthesis of our process of categorification of the logical notions.

DEFINITION 4.8
A doctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos** is called a **Gödel doctrine** if:

1. the category $\mathcal{C}$ is cartesian closed;
2. the doctrine P is existential and universal;
3. the doctrine P has enough existential-free predicates;
4. the existential-free objects of P are stable under universal quantification, i.e. if α is an element of $P(A)$ and it is existential-free, then $\forall_\pi(\alpha)$ is existential-free for every projection π from A;
5. the subdoctrine $P^{\exists\text{-free}}\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos** of the existential-free predicates of P has enough universal-free predicates.

The fourth point of the Definition 4.8 above implies that, given a Gödel doctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos**, the sub-doctrine $P^{\exists\text{-free}}\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos**, such that $P^{\exists\text{-free}}(A)$ is the poset of existential-free predicates contained in $P(A)$ for any object A of $\mathcal{C}$, is a universal doctrine. From a purely logical perspective, requiring existential-free elements to be stable under universal quantification is quite natural since this can be also read as *if $\alpha(x)$ is an existential-free predicate, then $\forall x : X.\alpha(x)$ is again an existential-free predicate*.

Now we have all of the tools needed to introduce the notion of *quantifier-free predicate* in the categorical setting of Gödel doctrines.

DEFINITION 4.9
An element α of a fibre $P(A)$ of a Gödel doctrine P that is both an existential-free predicate of P and a universal-free predicate in the sub-doctrine $P^{\exists\text{-free}}$ of existential-free elements of P is called a **quantifier-free predicate** of P. The sub-doctrine of quantifier-free elements is denoted by $P^{\exists\text{-free}}\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos**.

In order to simplify the notation, but also to make clear the connection with the logical presentation in the Dialectica interpretation, for a given Gödel doctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos**, we will use the notation α_D to indicate an element α of $P^{\exists\forall\text{-free}}$, i.e. a quantifier-free predicate. Applying the definition of a Gödel doctrine, we obtain the following result.

THEOREM 4.10
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos** be a Gödel doctrine, and let α be an element of $P(A)$. Then there exists a

quantifier-free predicate α_D of $P(I \times U \times X)$ such that:

$$i : I \mid \alpha(i) \dashv\vdash \exists u : U.\forall x : X.\alpha_D(i, u, x).$$

This theorem shows that in a Gödel doctrine every formula admits a presentation of the precise form used in the Dialectica translation.

Now we have all the instruments to state and prove the main theorem of this section, i.e. Theorem 4.11. In particular, we show that employing the properties of a Gödel doctrine we can provide a categorical description and presentation of the chain of equivalences involved in the Dialectica interpretation of the implicational formulae. In particular, we show that the crucial steps where (IP) and (MP) are applied are solved categorically employing universal properties of existential-free and universal-free elements.

For sake of readability, we omit the types of quantified variables as we anticipated in the previous section.

THEOREM 4.11
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ be a Gödel doctrine. Then for every ψ_D in $P(I \times U \times X)$ and ϕ_D in $P(I \times V \times Y)$ quantifier-free predicates of P we have that:

$$i : I \mid \exists u.\forall x.\psi_D(i, u, x) \vdash \exists v.\forall y.\phi_D(i, v, y),$$

if and only if there exist $I \times U \xrightarrow{f_0} V$ and $I \times U \times Y \xrightarrow{f_1} X$ such that:

$$i : I, u : U, y : Y \mid \psi_D(i, u, f_1(i, u, y)) \vdash \phi_D(i, f_0(i, u), y).$$

PROOF. Let us consider two quantifier-free predicates ψ_D of $P(I \times U \times X)$ and ϕ_D of $P(I \times V \times Y)$ of the Gödel doctrine P. The following equivalence follows by definition of left adjoint functor:

$$i : I \mid \exists u.\forall x.\psi_D(i, u, x) \vdash \exists v.\forall y.\phi_D(i, v, y) \iff$$
$$i : I, u : U \mid \forall x.\psi_D(i, u, x) \vdash \exists v.\forall y.\phi_D(i, v, y).$$

Now we employ the fact that the predicate $\forall x.\psi_D(i, u, x)$ is existential-free in the Gödel doctrine, obtaining that there exists an arrow $I \times U \xrightarrow{f_0} V$, such that:

$$i : I, u : U \mid \forall x.\psi_D(i, u, x) \vdash \exists v.\forall y.\phi_D(i, v, y) \iff$$
$$i : I, u : U \mid \forall x.\psi_D(i, u, x) \vdash \forall y.\phi_D(i, f_0(i, u), y).$$

Then, since the universal quantifier is right adjoint to the weakening functor, we have that:

$$i : I, u : U \mid \forall x.\psi_D(i, u, x) \vdash \forall y.\phi_D(i, f_0(i, u), y) \iff$$
$$i : I, u : U, y : Y \mid \forall x.\psi_D(i, u, x) \vdash \phi_D(i, f_0(i, u), y).$$

Now we employ the fact that $\phi_D(i, f_0(u), y)$ is universal-free in the subdoctrine of existential-free elements of P. Notice that since $\psi_D(i, u, x)$ is a quantifier-free element of the Gödel doctrine, we have that $\forall x.\psi_D(i, u, x)$ is existential free. Recall that this follows from the fact that in every Gödel doctrine, existential-free elements are stable under universal quantification (this is the last point

Definition 4.8). Therefore we can conclude that there exists an arrow $I \times U \times Y \xrightarrow{f_1} X$ of $\mathcal{C}$ such that:

$$i : I, u : U, y : Y \mid \forall x.\psi_D(i,u,x) \vdash \phi_D(i, f_0(i,u), y) \iff$$
$$i : I, u : U, y : Y \mid \psi_D(i,u,f_1(i,u,y)) \vdash \phi_D(i, f_0(i,u), y).$$

Then, combining the first and the last equivalences, we obtain the following equivalence:

$$i : I \mid \exists u.\forall x.\psi_D(i,u,x) \vdash \exists v.\forall y.\phi_D(i,v,y) \iff$$
$$\text{there exist } (f_0, f_1) \text{ s.t. } i : I, u : U, y : Y \mid \psi_D(i,u,f_1(i,u,y)) \vdash \phi_D(i, f_0(i,u), y).$$ □

Notice that the arrow $U \xrightarrow{f_0} V$ in Theorem 4.11 represents the *witness function*, i.e. it assigns to every witness u of the hypothesis a witness $f_0(u)$ of the thesis, while the arrow $U \times Y \xrightarrow{f_1} X$ represents the *counterexample function*. Observe that while the witness function $f_0(u)$ depends only of the witness u the counterexample function $f_1(u,y)$ depends on a witness of the hypothesis and a counterexample of the thesis. This is a quite natural fact because, considering the constructive point of view, the counterexample has to be relative to a witness validating the thesis.

Therefore, Theorem 4.11 shows that the notion of Gödel doctrine encapsulates in a pure form the basic mathematical feature of the Dialectica interpretation, namely its interpretation of implication, which corresponds to the existence of functionals of types $f_0 : U \to V$ and $f_1 : U \times Y \to X$ as described. One should think of this as saying that a proof of $\exists u.\forall x.\psi_D(i,u,x) \to \exists v.\forall y.\phi_D(i,v,y)$ is obtained by transforming to $\forall u.\exists v.\forall y.\exists x.(\psi_D(i,u,x) \to \phi_D(i,v,y))$, and then Skolemizing along the lines explained in the Section 2 and by Troelstra [6]. So, combining Theorems 4.10 and 4.11, we have strong evidence that the notion of Gödel doctrine really provides a categorical abstraction of the main concepts involved in the Dialectica translation.

Now we show that this kind of doctrine embodies also the *logical principles* involved in the translation. The first principle we consider it the axiom of choice (AC) also sometimes called the principle of Skolemization.

THEOREM 4.12
Every Gödel doctrine $P : \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Pos** validates the following **Skolemization principle**:

$$a : A \mid \forall b : B.\exists c : C.\alpha(a,b,c) \dashv\vdash \exists f : C^B.\forall b : B.\alpha(a,b,\mathrm{ev}(f,b)),$$

where α is any predicate in $P(A \times B \times C)$.

PROOF. Let us assume that $a : A \mid \gamma(a) \vdash \forall b.\exists c.\alpha(a,b,c)$ for some predicate γ of $P(A)$. By point (3) of Definition 4.8, we assume without loss of generality that $\gamma(a)$ is existential-free: otherwise, there is an existential-free predicate γ' covering $\gamma(a)$ and we get back to our hypothesis by using that P is existential.

Since P is universal, it is the case that $a : A, b : B \mid \gamma(a) \vdash \exists c.\alpha(a,b,c)$ and, being $\gamma(a)$ existential-free:

$$a : A, b : B \mid \gamma(a) \vdash \alpha(a,b,g(a,b))$$

for some term in context $a : A, b : B \mid g(a,b) : C$. Being $\mathcal{C}$ cartesian closed, there is a context $f : C^B$ together with a term in context $f : C^B, b : B \mid \mathrm{ev}(f,b) : C$ such that there is a unique term in context

$a : A \mid h(a) : C^B$ satisfying $a : A, a : B \mid g(a,b) = \mathrm{ev}(h(a), b) : C$. Hence:

$$a : A, b : B \mid \gamma(a) \vdash \alpha(a, b, \mathrm{ev}(h(a), b))$$

and being P universal, it is the case that:

$$a : A \mid \gamma(a) \vdash \forall b.\alpha(a, b, \mathrm{ev}(h(a), b)).$$

Finally, since:

$$a : A \mid \forall b.\alpha(a, b, \mathrm{ev}(h(a), b)) \vdash \exists f.\forall b.\alpha(a, b, \mathrm{ev}(f, b)),$$

(this holds for any predicate $\delta(a, -)$ in place of the predicate $\forall b.\alpha(a, b, \mathrm{ev}(-, b))$) we conclude that:

$$a : A \mid \gamma(a) \vdash \exists f.\forall b.\alpha(a, b, \mathrm{ev}(f, b)).$$

We are done by taking $\forall b.\exists c.\alpha(a, b, c)$ as the predicate $\gamma(a)$. □

REMARK 4.13
In the proof of Theorem 4.12 we do not need the property 5. of Definition 4.8. That is why, according to [31], one calls a Skolem doctrine a doctrine satisfying all of the properties satisfied by a Gödel doctrine, except for the property named 5. here.

4.1 Dialectica construction

Recall that the notion of Dialectica category introduced in [3] was generalized to the fibrational setting by Hofstra in [8], and this means that, in particular, we can consider the proof-irrelevant construction associating a doctrine $\mathfrak{Dial}(P)$ to a given doctrine P:

Dialectica construction. Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$ be a doctrine whose base category $\mathcal{C}$ is cartesian closed. The **dialectica doctrine**:

$$\mathfrak{Dial}(P)\colon \mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Pos}$$

is defined as the functor sending an object I into the poset $\mathfrak{Dial}(P)(I)$ defined as follows:

- **objects** are quadruples (I, X, U, α) where I, X and U are objects of the base category $\mathcal{C}$ and $\alpha \in P(I \times X \times U)$;
- **partial order:** we stipulate that $(I, U, X, \alpha) \leq (I, V, Y, \beta)$ if there exists a pair (f_0, f_1), where $I \times U \xrightarrow{f_0} V$ and $I \times U \times Y \xrightarrow{f_1} X$ are morphisms of $\mathcal{C}$ such that:

$$\alpha(i, u, f_1(i, u, y)) \leq \beta(i, f_0(i, u), y).$$

In [31], we proved that a fibration is an instance of the Dialectica construction if and only if it is a Gödel fibration, and to prove this result we employ the decomposition of the Dialectica monad as a free-simple-product completion followed by the free-simple-coproduct completion of fibrations. So we can deduce the same result for the proof-irrelevant version here simply as a particular case.

However, notice that employing Theorems 4.10 and 4.11 we have another simpler and more direct way of proving such correspondence. This is because Theorem 4.11 states that the order defined in the fibres of a Gödel doctrine is exactly the same order defined in a dialectica doctrine. The idea is that if P is a Gödel doctrine and $P^{\exists\forall\text{-free}}$ is the subdoctrine of quantifier-free elements of P it is easy to check that the assignment $P(I) \xrightarrow{(-)^D} \mathfrak{Dial}(P^{\exists\forall\text{-free}})(I)$ sending $\alpha \mapsto (I, X, U, \alpha_D)$ where α_D is the quantifier-free element such that $\alpha(i) \dashv\vdash \exists u \forall x \alpha_D(i, u, x)$ (which exists by Theorem 4.10), provides an isomorphism of posets by Theorem 4.11, and it can be extended to an isomorphism of existential and universal doctrines.

THEOREM 4.14
Every Gödel doctrine P is equivalent to the Dialectica completion $\mathfrak{Dial}(P^{\exists\forall\text{-free}})$ of the subdoctrine $P^{\exists\forall\text{-free}}$ of P consisting of the quantifier-free predicates of P.

Therefore, we have that Theorem 4.14 provides another way of thinking about Dialectica doctrines (or Dialectica categories) since it underlines the logical properties that a doctrine has to satisfy in order to be an instance of the Dialectica construction.

5 Logical principles in Gödel hyperdoctrines

Gödel doctrines provide a categorical framework that generalizes the principal concepts underlying the Dialectica translation, such as the existence of witness and counterexample functions whenever we have an implication $i : I \mid \exists u.\forall x.\psi_D(u,x,i) \vdash \exists v.\forall y.\phi_D(v,y,i)$. The key idea is that, intuitively, the notion of *existential-quantifier-free* objects can be seen as a reformulation of the *independence of premises rule*, while *quantifier-free* objects can be seen as a reformulation of *Markov rule*. Notice that in the proof of Theorem 4.11 existential and universal free elements play the same role that (IP) and (MP) have in the Dialectica interpretation of implicational formulae.

The main goal of this section is to formalize this intuition showing the exact connection between the rules (IPR) and (MR), the principles (IP) and (MP) and Gödel doctrines. So, first of all we have to equip Gödel doctrines with the appropriate Heyting structure in the fibres in order to be able to formally express these principles. Therefore, we have to consider Gödel hyperdoctrines.

DEFINITION 5.1
A hyperdoctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** is said a **Gödel hyperdoctrine** when P is a Gödel doctrine.

From a logical perspective, one might want the quantifier-free predicates to be closed with respect to all of the propositional connectives (or equivalently that P is the dialectica completion of a hyperdoctrine itself—see [29]), since this is what happens in logic. However, for sake of generality, we do not assume such closure condition on quantifier-free elements.

5.1 Logical rules

The main purpose of this section is to show what logical rules are satisfied in a Gödel hyperdoctrine.

THEOREM 5.2
Every Gödel hyperdoctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** validates the **Rule of Independence of Premise**, i.e. whenever $\beta \in P(A \times B)$ and $\alpha \in P(A)$ is a existential-free predicate, it is the case that:

$$a : A \mid \top \vdash \alpha(a) \rightarrow \exists b.\beta(a,b) \text{ implies that } a : A \mid \top \vdash \exists b.(\alpha(a) \rightarrow \beta(a,b)).$$

PROOF. Let us assume that $a : A \mid \top \vdash \alpha(a) \rightarrow \exists b.\beta(a,b)$. Then it is the case that $a : A \mid \alpha(a) \vdash \exists b.\beta(a,b)$. Since $\alpha(a)$ is free from the existential quantifier, it is the case that there is a term in context $a : A \mid t(a) : B$ such that:

$$a : A \mid \top \vdash \alpha(a) \rightarrow \beta(a,t(a)).$$

Therefore, since:

$$a : A \mid \alpha(a) \to \beta(a, t(a)) \vdash \exists b.(\alpha(a) \to \beta(a, b)),$$

(as this holds for any predicate $\gamma(a, -)$ in place of the predicate $\alpha_D(a) \to \beta(a, -)$) we conclude that:

$$a : A \mid \top \vdash \exists b.(\alpha(a) \to \beta(a, b)).$$

□

Notice that Theorem 5.2 formalizes precisely the intuition that the notion of existential-free element can be seen as a reformulation of the independence of premises rule: in a Gödel hyperdoctrine, we have that existential-free elements are *exactly* elements satisfying the independence of premises rule.

THEOREM 5.3
Every Gödel hyperdoctrine $P \colon \mathcal{C}^{\text{op}} \longrightarrow$ **Hey** satisfies the following **Generalized Markov Rule**, i.e. whenever $\beta_D \in P(A)$ is a quantifier-free predicate and $\alpha \in P(A \times B)$ is an existential-free predicate, it is the case that:

$$a : A \mid \top \vdash (\forall b.\alpha(a, b)) \to \beta_D(a) \text{ implies that } a : A \mid \top \vdash \exists b.(\alpha(a, b) \to \beta_D(a)).$$

PROOF. Let us assume that $a : A \mid \top \vdash (\forall b.\alpha(a, b)) \to \beta_D(a)$. Then it is the case that $a : A \mid (\forall b.\alpha(a, b)) \vdash \beta_D(a)$. Hence, since β_D is quantifier-free and α is existential-free, there exists a term in context $a : A \mid t(a) : B$ such that:

$$a : A \mid \top \vdash \alpha(a, t(a)) \to \beta_D(a),$$

therefore, since:

$$a : A \mid \alpha(a, t(a)) \to \beta(a) \vdash \exists b.(\alpha(a, b) \to \beta_D(a)),$$

we can conclude that:

$$a : A \mid \top \vdash \exists b.(\alpha(a, b) \to \beta_D(a)).$$

□

While for the case of (IPR), we have that existential-free elements of a Gödel hyperdoctrine correspond to formulae satisfying (IPR), we have that the elements of a Gödel doctrine that are quantifier-free, i.e. universal-free in the subdoctrine of existential-free elements, are exactly those satisfying a Generalized Markov Rule by Theorem 5.3. Moreover, notice that this Generalized Markov Rule is exactly the one we need in the equivalence between (4) and (5) in the interpretation of the implication in Section 2. Alternatively, in order to get this equivalence, one requires β_D to satisfy the law of excluded middle and the usual Markov Rule (see Corollary 5.4), as these two assumptions yield the Generalized Markov Rule. In particular, any boolean doctrine (a hyperdoctrine modelling the law of excluded middle) satisfies the Generalized Markov Rule (see Remark 5.5).

To obtain the usual Markov Rule as corollary of Theorem 5.3, we simply have to require the bottom element $\bot$ of a Gödel hyperdoctrine to be *quantifier-free*.

COROLLARY 5.4
Every Gödel hyperdoctrine $P \colon \mathcal{C}^{\text{op}} \longrightarrow$ **Hey** such that $\bot$ is a quantifier-free predicate satisfies **Markov Rule**, i.e. for every quantifier-free element $\alpha_D \in P(A \times B)$ it is the case that:

$$b : B \mid \top \vdash \neg \forall a.\alpha_D(a, b) \text{ implies that } b : B \mid \top \vdash \exists a.\neg \alpha_D(a, b).$$

PROOF. It follows by Theorem 5.3 just by replacing β_D with $\bot$, that is quantifier-free by hypothesis. □

REMARK 5.5
Clearly any boolean doctrine satisfies the Rule of Independence of Premises and the (Modified) Markov Rule, as it models every inference rule of classic first-order logic. In general, these are not satisfied by a usual hyperdoctrine, because they are not satisfied in intuitionistic first-order logic. It turns out that *the logic modelled by a Gödel hyperdoctrine is right in-between intuitionistic first-order and classical first-order logic*: it is powerful enough to guarantee the equivalences in Section 2 that justify the Dialectica interpretation of the implication.

REMARK 5.6
We observe that Theorem 5.3 and Theorem 5.2 deal with the validity of the *rule versions* of (IP) and (MP), and not the usual presentation in form of axioms or *principles*. As pointed out in [21], in general, these are not valid in an arbitrary intuitionistic theory, so it becomes interesting to find out which are the intuitionist theories that validate these rules. The validity of these rules in arbitrary Gödel hyperdoctrines have two main consequences: first, since HA validates these rules, see [26], and Gödel's Dialectica interpretation was originally introduced to provide proofs of the relative consistency of HA, the fact that Gödel doctrines validate these rules too *underscores how faithful the modelling is*. If Gödel doctrines or Dialectica categories didn't validate these rules, it would be hard to say that these categorical constructions abstract the main features of the logical translation, since they could not be employed for giving proofs of relative consistency of HA.

Secondly, we have that validating these rules suggests that the internal logic of Gödel hyperdoctrines could represent an interesting family of theories being intuitionistic, but at the same time validating the rule versions of (IP) and (MP).

We conclude this section on logical rules by presenting two other results about the Rule of Choice and the Counterexample Property previously defined in [29], which follow directly from the definitions of existential-free and universal-free elements.

COROLLARY 5.7
Every Gödel hyperdoctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** such that $\bot$ is a quantifier-free object satisfies the **Counterexample Property**, i.e. whenever:

$$a : A \mid \forall b.\alpha(a,b) \vdash \bot$$

for some predicate $\alpha(a,b) \in P(A \times B)$, then it is the case that:

$$a : A \mid \alpha(a, g(a)) \vdash \bot$$

for some term in context $a : A \mid g(a) : B$.

COROLLARY 5.8
Every Gödel hyperdoctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** such that $\top$ is existential-free satisfies the **Rule of Choice**, i.e. whenever:

$$a : A \mid \top \vdash \exists b.\alpha(a,b)$$

for some existential-free predicate $\alpha \in P(A \times B)$, then it is the case that:

$$a : A \mid \top \vdash \alpha(a, g(a))$$

for some term in context $a : A \mid g(a) : B$.

The rule appearing in Corollary 5.8 is called *Rule of Choice* in [15], while it appears as *explicit definability* in [21].

5.2 Logical principles

In the previous results, we have seen what *rules* hold in Gödel hyperdoctrines. This subsection is devoted to the analysis of the respective logical *principles* in Gödel hyperdoctrines. In detail, if previously we found the right hypotheses that make a Gödel hyperdoctrine into a model of the Rule of Independence of Premise and of (the Generalized) Markov Rule, in this subsection we follow the same process for the corresponding stronger formulation of these as principles. We start by stating the following theorem, involving the Independence of Premise in its formulation as a principle:

THEOREM 5.9
Every Gödel hyperdoctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** such that the existential-free elements are closed with respect to finite conjunction satisfies the **Principle of Independence of Premise**, i.e. whenever $\beta \in P(A \times B)$ and $\alpha \in P(A)$ is an existential-free predicate, it is the case that:

$$a : A \mid \top \vdash (\alpha(a) \to \exists b.\beta(a,b)) \to \exists b.(\alpha(a) \to \beta(a,b)).$$

PROOF. First, since every Gödel doctrine has enough existential-free elements, there exists an existential-free element $\gamma(a,c) \in P(A \times C)$ such that:

$$a : A \mid \exists c.\gamma(a,c) \dashv\vdash \alpha(a) \to \exists b.\beta(a,b).$$

In particular, we have that $a : A, c : C \mid \gamma(a,c) \vdash \alpha(a) \to \exists b.\beta(a,b)$. Hence we have that:

$$a : A, c : C \mid \gamma(a,c) \wedge \alpha(a) \vdash \exists b.\beta(a,b).$$

Notice that $\gamma(a,c) \wedge \alpha(a)$ is an existential-free element because both $\gamma(a,c)$ and $\alpha(a)$ are existential-free elements and existential-free elements are closed with respect finite conjunction by hypothesis. Therefore, we can conclude that there exists a term $a : A, c : C \mid t(a,c) : B$ such that:

$$a : A, c : C \mid \gamma(a,c) \wedge \alpha(a) \vdash \beta(a, t(a,c)),$$

and then we have:

$$a : A, c : C \mid \gamma(a,c) \vdash \alpha(a) \to \beta(a, t(a,c)).$$

Now, since $\alpha(a) \to \beta(a, t(a,c))$ is exactly $(\alpha(a) \to \beta(a,b))[t(a,c)/b]$ and it always holds that:

$$a : A, c : C \mid (\alpha(a) \to \beta(a,b))[t(a,c)/b] \vdash \exists b.(\alpha(a) \to \beta(a,b))$$

we get that:

$$a : A, c : C \mid \gamma(a,c) \vdash \exists b.(\alpha(a) \to \beta(a,b)).$$

Therefore we can conclude that:

$$a : A \mid \exists c.\gamma(a,c) \vdash \exists b.(\alpha(a) \to \beta(a,b)).$$

Since $a : A \mid \exists c.\gamma(a,c) \dashv\vdash \alpha(a) \to \exists b.\beta(a,b)$, it is the case that:

$$a : A \mid \top \vdash (\alpha(a) \to \exists b.\beta(a,b)) \to \exists b.(\alpha(a) \to \beta(a,b)).$$

□

As a corollary of the previous result, we obtain the following presentation of the principle (IP*) introduced in Section 2.1 in terms of Gödel hyperdoctrine. We recall that (IP*) is precisely the form of the Principle of Independence of Premise we need in the Dialectica interpretation.

COROLLARY 5.10
Every Gödel hyperdoctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** such that the existential-free elements are closed with respect to finite conjunction satisfies (IP*), i.e. whenever $\beta \in P(C \times B)$ and $\alpha_D \in P(A)$ is an quantifier-free predicate, it is the case that:

$$- \mid \top \vdash (\forall a.\alpha_D(a) \to \exists b.\forall c.\beta(c,b)) \to \exists b.(\forall a.\alpha_D(a) \to \forall c.\beta(c,b)).$$

PROOF. It follows from Theorem 5.9 and from the fact that existential-free if α_D is quantifier-free then $\forall a.\alpha_D$ is existential-free. □

Similarly, we can prove the following result about the Generalized Markov Principle:

THEOREM 5.11
Every Gödel hyperdoctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** such that the existential-free elements are closed with respect to implication satisfies the following **Generalized Markov Principle**, i.e. whenever $\beta_D \in P(A)$ is a quantifier-free predicate and $\alpha \in P(A \times B)$ is an existential-free predicate, it is the case that:

$$a : A \mid \top \vdash (\forall b.\alpha(a,b) \to \beta_D(a)) \to \exists b.(\alpha(a,b) \to \beta_D(a)).$$

PROOF. First notice that since α is an existential-free predicate and β_D is a quantifier-free predicate we have that $\forall b.\alpha(a,b) \to \beta_D(a)$ is an element of $P^{\exists\text{-free}}(A)$ (because by hypothesis existential-free elements are closed with respect to implication). Thus, since $P^{\exists\text{-free}}$ has enough quantifier-free elements by definition of Gödel doctrine, there exists a universal-free predicate of $P^{\exists\text{-free}}$, i.e. a quantifier-free predicate $\sigma_D \in P^{\exists\text{-free}}(\times C)$ such that $a : A \mid \forall c.\sigma_D(a,c) \dashv\vdash \forall b.\alpha(a,b) \to \beta_D(a)$. In particular, we have that $a : A \mid \forall c.\sigma_D(a,c) \wedge \forall b.\alpha(a,b) \vdash \beta_D(a)$, and hence $a : A \mid \forall c.\forall b.(\sigma_D(a,c) \wedge \alpha(a,b)) \vdash \beta_D(a)$. Now, since β_D is quantifier-free, i.e. it is universal-free in $P^{\exists\text{-free}}$, there exist two terms $a : A \mid t(a) : B$ and $a : A \mid t'(a) : C$ such that:

$$a : A \mid \sigma_D(a,t'(a)) \wedge \alpha(a,t(a)) \vdash \beta_D(a).$$

Therefore, we have that $a : A \mid \sigma_D(a,t'(a)) \vdash (\alpha(a,b) \to \beta_D(a))[t(a)/b]$. Now, since we always have that $a : A \mid \forall c.\sigma_D(a,c) \vdash \sigma_D(a,t'(a))$ and $a : A \mid (\alpha(a,b) \to \beta_D(a))[t(a)/b] \vdash \exists b.(\alpha(a,b) \to \beta_D(a))$, we can conclude that:

$$a : A \mid \top \vdash (\forall b.\alpha(a,b) \to \beta_D(a)) \to \exists b.(\alpha(a,b) \to \beta_D(a)).$$

□

In order to obtain the usual presentation of Markov Principle as a corollary of Theorem 5.11, we simply have to require the bottom element $\bot$ of a Gödel hyperdoctrine to be *quantifier-free*:

COROLLARY 5.12
Every Gödel hyperdoctrine $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** such that the existential-free elements are closed with respect to implication and falsehood $\bot$ is a quantifier-free predicate, satisfies **Markov Principle**, i.e. for every quantifier-free element α_D in $P(A \times B)$ it is the case that:

$$b : B \mid \top \vdash \neg\forall a.\alpha_D(a,b) \to \exists a.\neg\alpha_D(a,b).$$

PROOF. It follows by Theorem 5.11 just by replacing β_D with $\bot$, that is quantifier-free by hypothesis. □

We have proved that under suitable hypotheses, a Gödel hyperdoctrine satisfies (IP), (MP), (GMP) and the principle of Skolemization.

Therefore, combining Theorem 5.9, Theorem 5.11 (and Corollary 5.12), and Theorem 4.12, we can repeat the chain of equivalences we provided in Section 2, and obtain the following main result.

THEOREM 5.13
Let $P\colon \mathcal{C}^{\mathrm{op}} \longrightarrow$ **Hey** be a Gödel hyperdoctrine such that:

- existential-free elements are closed with respect to implication and finite conjunction;
- falsehood $\bot$ is a quantifier-free predicate.

Then for every ψ_D in $P(I \times U \times X)$ and ϕ_D in $P(I \times V \times Y)$ quantifier-free predicates of P we have that the formula:

$$i : I \mid \exists u.\forall x.\psi_D(i,u,x) \to \exists v.\forall y.\phi_D(i,v,y)$$

is provably equivalent to:

$$i : I \mid \exists f_0, f_1.\forall u,y.(\psi_D(i,u,f_1(i,u,y)) \to \phi_D(i,f_0(i,u),y)).$$

Theorem 5.13 fully represents a categorical version of the translation of the implication connective in the Dialectica interpretation. In particular, it shows that the equivalence $(\psi \to \phi)^D \leftrightarrow (\psi^D \to \phi^D)$ presented in Section 2 is perfectly modelled by a Gödel hyperdoctrine satisfying the natural additional closure properties of Theorem 5.13.

REMARK 5.14
Observe that Theorem 5.13 can be considered a stronger version of Theorem 4.11. Hence, once more, it converts the rule stated in the latter theorem into an actual principle.

In detail, by the thesis of Theorem 5.13, it is enough to observe that the first sequent of the statement of Theorem 4.11 is equivalent to the sequent:

$$i : I \mid \top \vdash \exists u.\forall x.\psi_D(i,u,x) \to \exists v.\forall y.\phi_D(i,v,y)$$

by the elimination and introduction rules for the implication, and that the second one is equivalent to the following:

$$i : I \mid \top \vdash \exists f_0, f_1.\forall u,y.(\psi_D(i,u,f_1(i,u,y)) \to \phi_D(i,f_0(i,u),y)).$$

For this second equivalence, one applies the implicational elimination and introduction to convert the second sequent of 4.11 into:

$$i : I, u : U, y : Y \mid \top \vdash \psi_D(i,u,f_1(i,u,y)) \to \phi_D(i,f_0(i,u),y),$$

which is actually equivalent to $i : I \mid \top \vdash \exists f_0, f_1.\forall u,y.(\psi_D(i,u,f_1(i,u,y)) \to \phi_D(i,f_0(i,u),y))$, by Corollary 5.8 and being the formula:

$$i : I \mid \forall u,y.(\psi_D(i,u,f_1(i,u,y)) \to \phi_D(i,f_0(i,u),y))$$

existential-free.

Observe that Theorem 5.13 follows as a consequence of the fragment of first-order logic under which the internal language of a Gödel hyperdoctrine is closed. So far we can reinforce what stated in Remark 5.5 and say that this fragment contains at least the whole intuitionistic first-order logic together with the Principle of Independence of Premise, the Generalized Markov Principle and the

Principle of Skolemization. These principles, together with the rules of intuitionistic first-order logic, are precisely what is needed to get the equivalence $(\psi \to \phi)^D \leftrightarrow (\psi^D \to \phi^D)$ in a Gödel hyperdoctrine.

6 Conclusion

We have recast our previous fibrational based modelling of Gödel's interpretation [31] in terms of categorical (hyper)doctrines. We show that the notions we considered in our previous work (existential-free and universal-free objects) really provide a categorical explanation of the traditional syntactic notions as described in [6]. This means that we are able to mimic completely the purely logical explanation of the interpretation, given by Spector and expounded on by Troelstra [6], using categorical notions. We show how to interpret logical implications using the Dialectica transformation. Through this process we explain how we go beyond intuitionistic principles, adopting both the Independence of Premise (IP) principle and Markov Principle (MP) as well as the axiom of choice in the logic.

Our main results show the perfect correspondence between the logical and the categorical tools, in the cases of Markov principle (MP) and the independence of premise (IP) principle. Thus, we improve our results discussed in [32] showing that, under certain hypotheses, they hold even in the stronger form of principles.

This shows that the categorical modelling really captures all the essential features of the interpretation. But it also opens new possibilities for modelling of constructive set theories (in the style of Nemoto and Rathjen [21]) and of categorical modelling of intermediate logics (intuitionistic propositional logic plus (IP) or (MK), see [1, 7]). This also leads into applications such as the study of foundations of functional abstract machines [22], of dependent type theory [20], of reverse mathematics [21], of concurrency theory [34] and of quantified modal logic [25], among others. In future work, we want to investigate some of these applications as well as to pursue further abstract characterizations of categorical logic.

Acknowledgements

We would like to thank Maria Emilia Maietti for ideas and discussions that inspired this work and the anonymous referees for extremely useful and precise comments.

Funding

Research supported by the project MIUR PRIN 2017FTXR IT-MaTTerS (Trotta) and by a School of Mathematics EPSRC Doctoral Studentship (Spadetto).

References

[1] F. Aschieri and M. Manighetti. On natural deduction for Herbrand constructive logics II: Curry–Howard correspondence for Markov's Principle in first-order logic and arithmetic. CoRR, abs/1612.05457, 2016.

[2] J. Avigad and S. Feferman. Gödel's functional (Dialectica) interpretation. *Handbook of Proof Theory*, **137**, 02, 1970.

[3] V. de Paiva. The Dialectica categories. *Categories in Computer Science and Logic*, **92**, 47–62, 1989.

[4] J. Frey. Categories of partial equivalence relations as localizations. arXiv preprint https://arxiv.org/abs/1912.06726, 2020.

[5] K. Gödel. Über eine bisher noch nicht benützte erweiterung des finiten standpunktes. *Dialectica*, **12**, 280–287, 1958.

[6] K. Gödel and S. Feferman., et al. *Kurt Gödel: Collected Works: Volume II: Publications 1938–1974*, vol. **2**. Oxford University Press, 1986.

[7] H. Herbelin. An intuitionistic logic that proves Markov's principle. In *2010 25th Annual IEEE Symposium on Logic in Computer Science*, pp. 50–56, 2010.

[8] P. Hofstra. The Dialectica monad and its cousins. *Models, Logics, and Higher Dimensional Categories: A tribute to the Work of Mihály Makkai*, **53**, 107–139, 2011.

[9] J. M. E. Hyland, P. T. Johnstone and A. M. Pitts. Tripos theory. *Mathematical Proceedings of the Cambridge Philosophical Society*, **88**, 205–232, 1980.

[10] B. Jacobs. *Categorical Logic and Type Theory*. Studies in Logic and the Foundations of Mathematics, vol. **141**. North Holland Publishing Company, 1999.

[11] F. W. Lawvere. Adjointness in foundations. *Dialectica*, **23**, 281–296, 1969.

[12] F. W. Lawvere. Diagonal arguments and cartesian closed categories. In *Category Theory, Homology Theory and Their Applications*, vol. **2**, pp. 134–145. Springer, 1969.

[13] F. W. Lawvere. Equality in hyperdoctrines and comprehension schema as an adjoint functor. In *New York Symposium on Application of Categorical Algebra*, vol. **2**, A. Heller., ed., pp. 1–14. American Mathematical Society, 1970.

[14] M. E. Maietti. Modular correspondence between dependent type theories and categories including pretopoi and topoi. *Mathematical Structures in Computer Science*, **15**, 1089–1149, 2005.

[15] M. E. Maietti, F. Pasquali and G. Rosolini. Triposes, exact completions, and Hilbert's ε-operator. *Tbilisi Mathematical Journal*, **10**, 11, 2017.

[16] M. E. Maietti and G. Rosolini. Quotient completion for the foundation of constructive mathematics. *Logica Universalis*, **7**, 371–402, 2013.

[17] M. E. Maietti and G. Rosolini. Unifying exact completions. *Applied Categorical Structures*, **23**, 43–52, 2013.

[18] M. E. Maietti and G. Rosolini. *Relating Quotient Completions via Categorical Logic*, pp. 229–250. De Gruyter, 2016.

[19] M. Manighetti. Computational interpretations of Markov's principle. arXiv preprint https://arxiv.org/abs/1611.03714, 2016.

[20] S. K. Moss and T. von Glehn. Dialectica models of type theory. In *33rd Annual ACM/IEEE Symposium on Logic in Computer Science*, pp. 739–748. Association for Computing Machinery, New York, NY, USA, 2018.

[21] T. Nemoto and M. Rathjen. The independence of premise rule in intuitionistic set theories. arXiv preprint https://arxiv.org/abs/1911.08027, 2019.

[22] P. M. Pédrot. A functional functional interpretation. In *CSL-LICS 2014 Annual Conference on Computer Science Logic and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science*, 2014.

[23] A. M. Pitts. Categorical logic. In *Handbook of Logic in Computer Science*, vol. **6**, S. Abramsky, D. M. Gabbay and T. S. E. Maibaum., eds, pp. 39–129. Oxford Univ. Press, 1995.

[24] A. M. Pitts. Tripos theory in retrospect. *Mathematical Structures in Computer Science*, **12**, 265–279, 2002.

[25] T. Shimura and R. Kashima. Cut-elimination theorem for the logic of constant domains. *Mathematical Logic Quarterly*, **40**, 153–172, 1994.

[26] A. S. Troelstra and D. van Dalen. *Constructivism in Mathematics, Vol 1*. Constructivism in Mathematics: An Introduction. Elsevier Science, 1988.
[27] D. Trotta. The existential completion. *Theory and Applications of Categories*, **35**, 1576–1607, 2020.
[28] M.E. Maietti and D. Trotta. Generalized existential completions and their regular and exact completions. arXiv preprint https://arxiv.org/abs/2111.03850, 2021.
[29] D. Trotta and M. Spadetto. Quantifier completions, choice principles and applications. arXiv preprint https://arxiv.org/abs/2010.09111, 2020.
[30] D. Trotta, M. Spadetto and V. de Paiva. The Gödel fibration (long version). arXiv preprint https://arxiv.org/abs/2104.14021, 2020.
[31] D. Trotta, M. Spadetto and V. de Paiva. The Gödel Fibration. In *46th International Symposium on Mathematical Foundations of Computer Science (2021)*. LIPIcs, vol. 202, pp. 87:1–87:16, 2021.
[32] D. Trotta, M. Spadetto and V. de Paiva. Dialectica logical principles. In *Logical Foundations of Computer Science*, S. Artemov and A. Nerode., eds, pp. 346–363. Springer International Publishing, Cham, 2022.
[33] J. van Oosten. *Realizability: An Introduction to Its Categorical Side*. ISSN. Elsevier Science, 2008.
[34] G. Winskel. Making concurrency functional. arXiv preprint https://arxiv.org/abs/2202.13910, 2022.

Received 30 March 2022