Technical Report: A General Approach to Define Binders Using Matching Logic*

Xiaohong Chen and Grigore Roşu
University of Illinois at Urbana-Champaign
{xc3,grosu}@illinois.edu

June 9, 2020

Abstract

We propose a novel definition of binders using matching logic, where the binding behavior of object-level binders is directly inherited from the built-in $\exists$ binder of matching logic. We show that the behavior of binders in various logical systems such as λ-calculus, System F, π-calculus, pure type systems, can be axiomatically defined in matching logic as notations and logical theories. We show the correctness of our definitions by proving conservative extension theorems, which state that a sequent/judgment is provable in the original system if and only if it is provable in matching logic, in the corresponding theory. Our matching logic definition of binders also yields *models* to all binders, which are deductively complete with respect to formal reasoning in the original systems. For λ-calculus, we further show that the yielded models are representationally complete, a desired property that is not enjoyed by many existing λ-calculus semantics. This work is part of a larger effort to develop a logical foundation for the programming language semantics framework $\mathbb{K}$ (`http://kframework.org`).

1 Introduction

In this paper, we propose a novel definition of binders using *matching logic* [81, 21], where the binding behavior of object-level binders is directly inherited from the built-in $\exists$ binder of matching logic. An appealing aspect of our definition is that it automatically yields *models* to all binders. Therefore, it is interesting and motivating to define a logical system that features binding in matching logic, because it allows us to study the resulting model theory and properties, in addition to the proof theory. We define λ-calculus [28], System F [43, 79], pure type systems [7], and π-calculus [66] in matching logic as *theories* and prove the correctness of definitions as *conservative extension theorems* (Theorems 36 and 49). We also show that the models that matching logic yields for these theories are *deductively complete* with respect to formal reasoning in each of the respective systems (Sections 7 and 9.2). For λ-calculus, we show that the corresponding matching logic models are also *representationally complete* for all λ-theories, a desired property that is not known to hold for many existing λ-calculus semantics [85, 12, 44, 17, 87, 86, 33, 75, 83, 57] (see discussion in Section 8.2.2).

We use λ-calculus as an example to illustrate our definition of binders in matching logic. We define λ-abstraction, $\lambda x.\, e$, as the following matching logic formula (called *pattern*; see Definition 2):

$$\lambda x.\, e \equiv \mathsf{lambda}\,(\mathsf{intension}\,\exists x\colon \mathit{Var}.\,\langle x, e\rangle) \tag{1}$$

Intuitively, $\langle x, e\rangle$ builds an argument-value pair; $\exists$ is the built-in binder in matching logic that thus creates the binding of x to e; $\exists x\colon \mathit{Var}.\,\langle x, e\rangle$ builds the set-theoretic union of all argument-value pairs $\langle x, e\rangle$, as x

*This technical report completes the ICFP'20 conference paper [23] with all the proofs that were not possible to include there due to space restrictions.

ranging over all variables of sort *Var*; this union set is called the *graph* of the function $x \mapsto e$, which is then "packed" by the operator $\mathsf{intension}$ into an object and passed to lambda. Finally, lambda decodes/retracts the packed object and returns the intended interpretation of $\lambda x.\, e$. Binders in the other systems may require different retracts other than lambda, but all take the same packed object as argument, which for convenience we write $[x\colon \mathit{Var}]\, e \equiv \mathsf{intension}\, \exists x\colon \mathit{Var}.\, \langle x, e\rangle$.

The main goal of this paper is to show that the matching logic definition of binders as illustrated in Eq. (1), is mathematically interesting and can serve as a foundation of binders in language frameworks. In Section 2, we start with a discussion on the major existing approaches to dealing with binders and we compare them with our approach. Then we make the following contributions:

- We propose a novel functional variant of matching logic that is more suitable to capture binders, and we comprehensively study its model theory (Section 3); we demonstrate the expressiveness of this functional variant of matching logic by defining several important mathematical instruments (such as equality and sorts) as theories and notations (Section 4);
- We define λ-calculus (Section 5) as a theory in matching logic (Section 6), as an illustrative case study. Then we prove the conservative extension theorem for λ-calculus and show that matching logic yields complete models, in terms of deduction, for λ-calculus (Sections 7-8). We also discuss the representability problem in λ-calculus and show that matching logic yields models that are representationally complete, in Section 8.2.2;
- We generalize our method to arbitrary binders (Section 9).

Finally, we conclude the paper with future work in Sections 10-11.

This paper marks an important step towards formalizing the logical foundation of the $\mathbb{K}$ semantic framework (`http://kframework.org`), which has been used to define complete formal semantics of several real-world languages [15, 68, 50, 51, 30]. Prior attempts have been made to propose a logical foundation of $\mathbb{K}$ using formalisms like rewriting logic [82, 65] and graph rewriting [88], but none of them were satisfactory. Recently, matching logic has been proposed as an alternative [81, 21]. The main idea is that arbitrarily complex programming languages and calculi defined in $\mathbb{K}$ become theories in matching logic, and all the tools offered by $\mathbb{K}$, such as execution engines, symbolic reasoning, and even full functional correctness verification of program or language properties, become proof search heuristics in matching logic, which admits a small proof system and thus a small trust base. Several important logical systems have been defined in matching logic, but none where binders play a major role, like λ-calculus or type systems. On the other hand, the current $\mathbb{K}$ implementations already provide built-in support for user-defined binders of certain restricted forms (Remark 43). Thus, this paper fills this gap by giving the theoretical results about how to define logical systems that feature binders in matching logic and thus in $\mathbb{K}$, without any foundational compromise.

> This technical report accompanies the conference paper [23].
> All proof details can be found in the appendix.

2 Related Work: Existing Approaches to Defining Binders

We discuss some existing approaches to defining binders and compare them with our approach using matching logic. These approaches include: (1) *de Bruijn* techniques [31], which give α-equivalent terms identical encodings; (2) *combinators* [28], which translate terms with binders to binder-free combinator terms; (3) *nominal logic* [72], which uses first-order logic (FOL) to axiomatize *name-swapping* and *freshness*, and uses them to axiomatize object-level binding; (4) *higher-order abstract syntax* [70] (abbreviated HOAS), which uses fixed binders in the meta-language, often a variant of typed λ-calculus, to define arbitrary binders in the object-level systems; (5) *explicit substitution* [1], which uses customized calculi where the meta-level operation of capture-free substitution is incarnated in an object-level operation as part of the calculi; (6) *term-generic logic* [77] (abbreviated TGL), which is a FOL variant parametric in a generic term set, defined

axiomatically and not constructively, which can be *instantiated* by a concrete binder syntax. We discuss how these approaches handle binders and binding behavior using the following λ-expression as an example (a closed expression with distinct bound variables, which requires α-renaming during reduction to avoid variable-capture):

$$(\lambda z.\,(zz))(\lambda x.\,\lambda y.\,(xy)) \tag{$\dagger$}$$

De Bruijn encodings eliminate bound variables by replacing them with indexes that denote the number of (nested) binders that are in scope between them and their corresponding binders.[1] For example, the de Bruijn encoding of ($\dagger$) is $(\lambda(11))(\lambda\lambda(21))$, where 1 means that it is bound by the closest binder and 2 means that it is bound by the second closest binder. Bound variables are eliminated so α-equivalent expressions have the same de Bruijn encoding. However, substitution requires index shifting, to adjust the indexes. De Bruijn techniques are used as the internal representations of terms in several theorem provers, but the encoding is not human readable, implementations are often tricky to get right, and efficiency problems can still appear on large terms.

Combinators translate binders to binder-free terms, which are built with constants like k and s, and application. This translation is called *abstraction elimination*, and can be implemented using term rewriting [55]. It may cause exponential growth in the translated term size. Reduction of combinatory terms is done using equations like $kxy = x$ and $sxyz = (xz)(yz)$ regarded as rewrite rules. Combinatory terms are not human readable; for example, (one of) the equivalent combinator term of ($\dagger$) is $s(skk)(skk)s(s(ks)(s(kk)(skk)))(k(skk))$. Using combinators, the binding behavior of λ is captured *implicitly* through abstraction elimination.

Nominal logic refers to a family of FOL theories whose signatures contain a *name-swapping* operation $(x\,y)\cdot e$ that swaps all (free and bound) occurrences of x and y in e, and a *freshness* predicate $x \# e$ stating that x has no free occurrences in e. The notions of α-equivalence and capture-free substitution are then axiomatized using additional FOL axioms on top of the axioms of name-swapping and freshness. As an example, the following is an axiom in [72, Appendix A.3] that states that swapping two fresh names that do not occur free in a term has not effect:

$$\text{(F1)}\quad \forall x\colon \mathit{Var}.\,\forall y\colon \mathit{Var}.\,\forall e\colon \mathit{Exp}.\,x \# e \wedge y \# e \to (x\,y)\cdot e = e$$

where *Var* and *Exp* are the sorts of variables (also called atoms) and expressions, respectively. Nominal logic also defines a new sort $[\mathit{Var}]\mathit{Exp}$ and a FOL binary function $\_.\_\colon \mathit{Var} \times \mathit{Exp} \to [\mathit{Var}]\mathit{Exp}$ for binding, whose properties such as α-equivalence are axiomatized. Then, β-reduction in λ-calculus, e.g., can be defined in the following way [74, pp. 251, Eq. (12.17)]:

$$(\beta \text{ IN NOMINAL LOGIC})\qquad \forall x\colon \mathit{Var}.\,\forall e\colon \mathit{Exp}.\,\forall e'\colon \mathit{Exp}.\,\mathit{app}(\mathit{lam}(x.e), e') = \mathit{subst}((x.e), e')$$

where $\mathit{subst}(\_,\_)$ is a binary function defined by four axioms (see [72, pp. 8]), in accordance to the four possible forms that e can take (i.e., the variable x; a variable distinct from x; application; or abstraction). E.g., the following is the substitution axiom for abstraction [74, Eq. (12.20)]:

$$\forall x\colon \mathit{Var}.\,\forall y\colon \mathit{Var}.\,\forall e\colon \mathit{Exp}.\,\forall e'\colon \mathit{Exp}.\,y \# e' \to \mathit{subst}(x.\,\mathit{lam}(y.\,e), e') = \mathit{lam}(y.\,\mathit{subst}(x.\,e, e'))$$

Besides nominal logic and its metatheory [24, 25, 39], there is a wider range of research on *nominal techniques* in general, including studies on using Fraenkel-Mostowski sets [38], nominal sets [73] or similar set-theoretic structures [90] as well as category-theoretic notions [41] to formalize and reason about binders and operations on them, and have resulted in practical implementations that support complex recursive and inductive reasoning over terms with bindings as well as algorithms for unification [3] and narrowing [4]. These nominal approaches deal with variable names and bindings *directly*, treat variable names as normal data that can be manipulated, quantified, and reasoned about, and give explicit definitions to operations such as free variables and capture-free substitution (via name-swapping and freshness).

Nominal approaches can be directly exploited in matching logic, because FOL is a methodological fragment of matching logic. Indeed, [81, Section 7] shows how matching logic *symbols* (see Definition 2) can be

[1] Other de Bruijn encodings count the binders from the top of the terms.

used to uniformly represent both FOL predicates and FOL functions (Sections 3.2.1 and 3.2.2), in a way where FOL theories become matching logic theories as are, without any translations. Therefore, nominal logic variants can be defined as theories in matching logic straightforwardly, via the FOL capability of matching logic. Future research shall reveal more direct methods that capture the essence of nominal techniques (e.g., nominal sets) within matching logic, without going through FOL. In this paper, however, we explore a different, more HOAS-style treatment of binders using matching logic, where the built-in $\exists$ binder is used to define binders in object-languages (explained below and revisited in Remark 1).

Higher-order abstract syntax (HOAS) is a design pattern where some expressive higher-order calculus, usually one of the variants of typed λ-calculus [70, 48, 63, 69, 34, 42] or second-order equational logic [37, 34], is used as a foundation to define object-level binders. As an example, we show (part of) the HOAS-style definition of (untyped) λ-calculus in the Twelf system [71]:

```
exp : type.                          // the type for λ-expressions
app : exp -> exp -> exp.             // application is defined as a constant of a function type
lam : (exp -> exp) -> exp.           // lambda is defined as a constant of a function type whose
                                     //    argument also has a function type; e.g., the encoding of (†)
                                     //    is app (lam ([z] (app z z))) (lam ([x] lam ([y] (app x y))))
red : exp -> exp -> type.            // reduction relation (its type result makes it a binary predicate)
red-beta : red (app (lam ([x] (F x))) E) (F E).   // β-reduction, discussed below
```

where [x] _ is the built-in binder of (the HOAS variant underlying) Twelf; E is a variable of type exp; F is a variable of the function type exp -> exp; and (F x) is the (metalevel) application of F to x. Higher-order matching is needed when red-beta is applied, and the internal substitution mechanism of Twelf is triggered when F is applied to E. The binding behavior of λ is obtained from the binding behavior of the built-in binder [x] _, via a constant lam; specifically, $\lambda x.e$ is encoded as lam ([x] e). Object-level substitution is avoided, but clearly this is not how β-reduction is usually defined (for the usual definition, see (β, Reduction) below). Application in λ-calculus is defined by a simple desugaring to the builtin application, using a different constant app; that is, $e_1\, e_2$ is defined as $\mathsf{app}\, e_1\, e_2$ (rather than $e_1\, e_2$). Thus, the definition needs to be justified by proving *adequacy theorems* that establish a bijection between the expressions and formal proofs of λ-calculus, and the HOAS terms and type derivations, which is a tedious and nontrivial task [26].

Explicit substitution turns the implicit meta-level substitution operation into more explicit and atomic steps, in order to provide a better understanding of the operational semantics and execution models of higher-order calculi (see [54, pp. 1–2]; see also [14, pp. 4] for historical remarks). By doing so, it bridges the gap between higher-order formalisms and their implementations, and has resulted in several practical tools. For example, [89] proposes a calculus for explicit substitution whose implementation allows us to define executable formal representations of many logical systems featuring binders with a close-to-zero representational distance.

Term-generic logic (TGL) is a FOL variant, where the set of terms T is generic and given as a *parameter* that exports two operations—free variables and capture-free substitution—satisfying certain properties [77, Definition 2.1]. TGL formulas are then defined constructively as in FOL, from predicates $\pi(e_1, \ldots, e_n)$ and equations $e_1 = e_2$, to compound formulas built using $\wedge$, $\neg$, and $\exists$, with the important exception that $e_1, \ldots, e_n$ are not constructive terms like in FOL, but generic terms in T. In the case of λ-calculus, the set of λ-expressions Λ can be proved to satisfy the definition of a generic term set in TGL, so we can *instantiate* TGL by Λ. The binding behavior of λ is inherited automatically, through the T instance. The metalevel of λ-calculus can be defined by TGL axioms. For example, β-reduction is captured either as an equation or as a relation:

$$(\beta,\ \textsc{Equation}) \quad (\lambda x.\, e)\, e' = e'[e/x] \qquad\qquad (\beta,\ \textsc{Reduction}) \quad \mathit{reduces}\big((\lambda x.\, e)\, e', e'[e/x]\big)$$

where *reduces* is a binary predicate; $(\lambda x.\, e)\, e', e'[e/x] \in \Lambda$ are generic terms (schemas) that represent all the concrete instances. TGL has been used to define various systems featuring bindings. In this paper, we use TGL as an intermediate to capture other systems with binders within matching logic.

Our Approach Using Matching Logic Our matching logic encoding of binders is inspired by the key observation that the meaning of a term with binders, say $\lambda x.\, e$, can be given on top of the function that maps x to e, which can be encoded as its *graph*: the set of argument-value pairs $\bigcup_x \{(x, e)\}$. This set is then packed as an object and passed to a retraction function lambda that retracts/decodes the intended meaning of the term. We recall the encoding of $\lambda x.\, e$ in Eq. (1) below:

$$\lambda x.\, e \equiv \mathsf{lambda}\,(\mathsf{intension}\,\exists x\colon \mathit{Var}.\,\langle x, e\rangle)$$

Note that by introducing the following notation

$$[x\colon \mathit{Var}]\,e \equiv \mathsf{intension}\,\exists x\colon \mathit{Var}.\,\langle x, e\rangle$$

the encoding of $\lambda x.e$ becomes $\mathsf{lambda}\,([x\colon \mathit{Var}]\,e)$, where Var is the sort for λ-calculus variables and thus a subsort of Exp for expressions (see Section 6). Note that our matching logic encoding of binders is reminiscent of both the nominal encoding $\mathit{lam}(x.e)$ and the HOAS encoding $\mathsf{lam}\ (\mathsf{[x]}\ \mathsf{E})$.

An important aspect of our approach is that it yields *models*. We will give a comprehensive study on the *model theory* of matching logic, by which every theory is associated with default models that can be used to give *semantic interpretations* of all matching logic formulas (called *patterns*) of that theory. In particular, the matching logic theory of λ-calculus will also yield a precise and insightful description of how $\lambda x.\, e$ is interpreted (semantically) in matching logic models.

Models are insightful. They help us understand a logical system better, from a different angle. It is not unusual that more than one notion or class of models are proposed for one logic, because each has its unique merit in helping us understand the logic from a certain perspective. Since matching logic has a built-in notion of models, by defining a logical system as a matching logic theory we can immediately study its resulting model theory and properties. For example, in Section 8.2.2, we show how by defining λ-calculus in matching logic, we obtain a new semantics of λ-calculus that is representationally complete for all λ-theories.

The importance of models has also been recognized by several HOAS approaches. For example, [35] proposes presheaf models of variable binding in a second-order syntax of binding terms, where the initial model is used to define recursive/inductive operations; this work also yields an explicit connection to the scope-safe variant of De Bruijn approaches. [36, 37] propose for the same binding syntax yet another category of models, called second-order universal algebras, together with completeness and conservative extension results w.r.t. first-order universal algebras; however, the conservative extension w.r.t. the original logical systems that feature binding and their formal reasoning is not investigated at our knowledge, and not known if it holds. In our work using matching logic, we shall prove the conservative extension for all logical systems that feature binders considered in the paper, but will not cover the topics of inductive reasoning and/or initial models (although a special initial algebra will be discussed in Section 8 for λ-calculus); this topic is left as future work (see Section 10).

As a logic that features binding, we expect matching logic to be definable within HOAS. Such a definition will likely work fine in capturing the syntax and binding behavior of matching logic formulas/patterns as well as its proof theory, but it will not capture the semantics or models of matching logic; see related discussion in Remark 52. In this paper, we will discuss the other direction, that is to capture HOAS by matching logic. We will do that indirectly, by firstly capturing term-generic logic (TGL) and then re-using the existing TGL definitions of HOAS (see [76]). This indirect approach has the advantage that we will be able to examine how the very general TGL models are translated and preserved when defined in matching logic.

Remark 1. Dealing with binders has been and still is an active research topic. The variety of proposals and approaches has occasionally caused heated arguments. We conclude this section by reminding the reader that matching logic was designed to serve as a unified logical foundation for the $\mathbb{K}$ framework, which is intended to support all languages and all definitional styles as logical theories. That is, when looked at through the matching logic lenses, the various approaches to binders above become different methodologies for how to define matching logic theories.

3 Functional Variant of Matching Logic

Matching logic has been recently proposed in its full generality in [81, 21]. In this paper, we will use a variant of matching logic that has a more similar representation to functional programming languages, where the main constructs are function application and constants. Since matching logic is relatively new, we will not assume the reader familiar with it. Therefore, this section has a dual goal: to introduce the reader to the basic intuitions and notations of matching logic, and to propose and present in detail a functional variant of it. Section 3.1 defines its syntax and Section 3.2 its models and semantics. We define matching logic theories in Section 3.3.

3.1 Matching Logic Syntax

Matching logic is parametric in a *signature* that includes *variables* and *constant symbols*:

Definition 2. A *signature* is a tuple $\mathbb{\Sigma} = (EV, SV, \Sigma)$, where $EV \cap SV = \emptyset$ and

1. EV is a countably infinite set of *element variables* denoted $x, y, \ldots$;
2. SV is a countably infinite set of *set variables* denoted $X, Y, \ldots$;
3. Σ is an at most countable set of *(constant) symbols*, or just *symbols*, denoted $\sigma, \sigma_1, \sigma_2, \ldots$.

Matching logic formulas, called $\mathbb{\Sigma}$*-patterns* or simply *patterns*, are inductively defined as follows:

$$\varphi ::= x \mid X \mid \sigma \mid \varphi_1\, \varphi_2 \mid \bot \mid \varphi_1 \to \varphi_2 \mid \exists x.\, \varphi \tag{2}$$

where $\varphi_1\, \varphi_2$ is called an *application* and is assumed associative to the left; $\exists x.\, \varphi$ is the built-in *binder* in matching logic that binds x within φ. Note that $\exists$ only binds element variables and not set variables. We use $\textsc{Pattern}(\mathbb{\Sigma})$, or simply $\textsc{Pattern}$, to denote the set of all $\mathbb{\Sigma}$-patterns.

Remark 3. The syntax of the original matching logic has sorts and multiary many-sorted operations [81, 21]. Our functional variant syntax in Definition 2 is much simpler: it has no sorts and contains only one binary operation, the application, and constants. Yet, as seen in this paper, this simpler variant has the same expressiveness and reasoning capability.

As a convention, we assume the scope of $\exists$ goes as far as possible to the right, so for example, $\exists x.\, y \to x$ should be understood as $\exists x.\, (y \to x)$. In addition, we assume the standard notions of *free variables* $\mathrm{FV}(\varphi) \subseteq EV \cup SV$, α*-equivalence* $\varphi_1 \equiv_\alpha \varphi_2$, and *capture-free substitution* $\varphi[\psi/x]$, which are all summarized in Fig. 1. We regard α-equivalent patterns as *syntactically identical* patterns; in other words, $\varphi_1 \equiv_\alpha \varphi_2$ implies that $\varphi_1 \equiv \varphi_2$. A set of common derived constructs are also included in Fig. 1 in the usual way as syntactic sugar, and we assume the standard precedence among them.

The matching logic syntax of patterns given in Eq. (2) is similar to the FOL syntax of terms and formulas, except that we drop the distinction between terms and formulas, and unify them as patterns.[2] Also, we drop the multiary functions/predicates in FOL, and replace them with a set of constant symbols that can be *applied* to other patterns using the built-in application $\varphi_1\, \varphi_2$. This simpler syntax of matching logic makes it easier to develop its metatheory, and yet, as we will show in Section 4, we do not lose any specification or reasoning power, and can still define important and necessary mathematical instruments as theories and notations in matching logic.

By unifying the syntax of terms and formulas, we can *bind variables in terms*, using the built-in matching logic binder $\exists$. A minimal example is $\exists x.\, x$, where x is bound by $\exists x$, so $\mathrm{FV}(\exists x.\, x) = \emptyset$. While $\exists x.\, x$ is a well-formed matching logic pattern, it is neither a well-formed term nor a well-formed formula in FOL. As we will see in Section 6, being able to build terms *and* create bindings over them is what makes our encoding of various binders in matching logic possible, and novel.

[2] The syntax of a logic should be in harmony with its semantics. FOL distinguishes terms and formulas because their interpretations are different: terms are interpreted as elements and formulas are interpreted as truth values. As we will see in Section 3.2, the matching logic semantics interprets patterns uniformly to the sets of elements that match them, so there is no need to distinguish terms and formulas. Other such examples include modal logic [13] (which abandons terms entirely) and separation logic [80] (which merges the syntax for memory heaps with formulas).

free variables:

$\mathrm{FV}(x) = \{x\}$ $\quad \mathrm{FV}(X) = \{X\}$ $\quad \mathrm{FV}(\sigma) = \emptyset$ $\quad \mathrm{FV}(\varphi_1\, \varphi_2) = \mathrm{FV}(\varphi_1) \cup \mathrm{FV}(\varphi_2)$

$\mathrm{FV}(\bot) = \emptyset$ $\quad \mathrm{FV}(\varphi_1 \to \varphi_2) = \mathrm{FV}(\varphi_1) \cup \mathrm{FV}(\varphi_2)$ $\quad \mathrm{FV}(\exists x.\, \varphi) = \mathrm{FV}(\varphi) \setminus \{x\}$ α-renaming: $\exists x.\, \varphi \equiv \exists y.\, \varphi[y/x]$, for $y \notin \mathrm{FV}(\varphi)$

capture-free substitution (where y distinct from x and z is fresh):

$(\exists x.\, \varphi)[\psi/x] \equiv \exists x.\, \varphi$ $\quad\quad$ $(\exists x.\, \varphi)[\psi/y] \equiv \exists z.\, \varphi[z/x][\psi/y]$

---

derived constructs defined as syntactic sugar:

$\neg\varphi \equiv \varphi \to \bot$ $\quad \varphi_1 \vee \varphi_2 \equiv \neg\varphi_1 \to \varphi_2$ $\quad \varphi_1 \wedge \varphi_2 \equiv \neg(\neg\varphi_1 \vee \neg\varphi_2)$

$\top \equiv \neg\bot$ $\quad \forall x.\, \varphi \equiv \neg\exists x.\, \neg\varphi$ $\quad \varphi_1 \leftrightarrow \varphi_2 \equiv (\varphi_1 \to \varphi_2) \wedge (\varphi_2 \to \varphi_1)$

Figure 1: Above line: standard notions of free variables, α-equivalence, and capture-free substitution for $\exists$ in matching logic. Below line: usual derived constructs defined as syntactic sugar. Standard precedence assumed.

3.2 Matching Logic Semantics

Matching logic patterns are interpreted on an underlying carrier set of elements, and each pattern is then interpreted as a *set of elements*, which are those that *match* the pattern. This is called the *pattern matching semantics* of matching logic, and is what inspired the name "matching logic".

Intuitively, the pattern $\bot$ (called *bottom*) is matched by no elements, while $\top$ (called *top*, defined in Fig. 1) is matched by all elements. Conjunction $\varphi_1 \wedge \varphi_2$ is matched by the elements that match both φ_1 and φ_2, disjunction $\varphi_1 \vee \varphi_2$ by the elements that match φ_1 or φ_2, negation $\neg\varphi$ by the elements that do not match φ, and implication $\varphi_1 \to \varphi_2$ by all elements a such that if a matches φ_1 then a matches φ_2. Element variable x is matched by the element to which x evaluates (see Definition 7). Set variable X is matched by the set of elements to which X evaluates; this set can be empty, or total, or any subset of the carrier set. Quantification $\exists x.\, \varphi$ is matched by the elements that match φ for *some valuation* of x; that is, it *abstracts away* the irrelevant part x from the matched part φ.

Definition 4. Given $\mathbb{\Sigma} = (EV, SV, \Sigma)$, a $\mathbb{\Sigma}$-*model* (or just *model*) is a tuple $(M, \_\bullet\_, \{\sigma_M\}_{\sigma\in\Sigma})$, where

1. M is an underlying carrier set, required to be non-empty ($M \neq \emptyset$);
2. $\_\bullet\_ : M \times M \to \mathcal{P}(M)$ is called the *interpretation of application*, where $\mathcal{P}(M)$ is the powerset;
3. $\sigma_M \subseteq M$ is a subset, called the *interpretation of* σ, defined for every $\sigma \in \Sigma$.

We often use the same letter M to denote the above model and refer to $\mathbb{\Sigma}$ as the *signature of* M.

Let us compare matching logic and FOL, w.r.t. models. Both logics require their models to have nonempty carriers, so they agree on (1). For (3), however, FOL models interpret constants to elements, while matching logic models interpret constants to any carrier subsets. Similarly, for (2), FOL models interpret application (regarded as a binary function) as a function of $M \times M \to M$ that returns one element, while matching logic models interpret application to a function that returns a set. We use the terminology *functional interpretation* to refer to how FOL interprets functions and terms. Functional interpretation is in harmony with the syntax of FOL terms, which represent elements. Similarly, the *set-theoretic interpretation* of matching logic application and symbols is in harmony with its syntax of patterns, which represent sets of elements.

Note that the FOL functional interpretation can be seen as a special instance of the matching logic set-theoretic interpretation, due to the bijection between an element a and the singleton $\{a\}$: for any set M, the set of all singletons of M is isomorphic to M itself. This justifies our abuse of notation (used often in this paper) in which $\{a\}$ is written as a when there is no confusion. We will use two examples to illustrate how the functional interpretation is a special instance of the set-theoretic interpretation. These examples are also related to the model theory of λ-calculus, so we will re-visit them later; for now, we only use them as examples of matching logic models.

(CURRY.1) $k = s(s(ks)(s(kk)k))(k(skk))$

(CURRY.2) $s = s(s(ks)(s(k(s(ks)))(s(k(s(kk)))s)))(k(k(skk)))$

(CURRY.3) $s(s(ks)(s(kk)(s(ks)k)))(kk) = s(kk)$

(CURRY.4) $s(ks)(s(kk)) = s(kk)(s(s(ks)(s(kk)(skk)))(k(skk)))$

(CURRY.5) $s(k(s(ks)))(s(ks)(s(ks))) = s(s(ks)(s(kk)(s(ks)(s(k(s(ks)))s))))(ks)$

(MEYER-SCOTT) $\forall x.\, \forall y.\, (\forall z.\, xz = yz) \to s(k(skk))x = s(k(skk))y$

Figure 2: Five axioms of Curry and the Meyer-Scott axiom for λ-models [6, pp. 94] ($\bullet_A$ is omitted).

Example 5. Let $(A, \_\bullet_A\_)$ be an *applicative structure* [6, Definition 5.1.1], where A is a nonempty carrier set and $\_\bullet_A\_ : A \times A \to A$ is an application function. Let matching logic signature $\mathbb{\Sigma}^{\emptyset}$ contain no symbols. We define a $\mathbb{\Sigma}^{\emptyset}$-model $(M, \_\bullet\_, \{\})$, where $M = A$ and $a \bullet b = \{a \bullet_A b\}$ for all $a, b \in A$. Then, M is isomorphic to A under the bijection between elements and singletons.

Example 6. Let $(A, \_\bullet_A\_, k, s)$ be a *combinatory algebra* [6, Definition 5.1.7], where $(A, \_\bullet_A\_)$ is an applicative structure and $k, s \in A$ are distinguished elements such that $k \bullet_A a \bullet_A b = a$ and $s \bullet_A a \bullet_A b \bullet_A c = (a \bullet_A c) \bullet_A (b \bullet_A c)$, for all $a, b, c \in A$. A is called a *λ-model* [6], if it additionally satisfies the five axioms of Curry [6, Theorem 5.2.5] and the Meyer-Scott axiom [6, Definition 5.2.7], shown in Fig. 2. Let $\mathbb{\Sigma}^{ks}$ be the matching logic signature $\mathbb{\Sigma}^{ks} = \{\mathsf{k}, \mathsf{s}\}$ and define a $\mathbb{\Sigma}^{ks}$-model $(M, \_\bullet\_, \{\mathsf{k}_M, \mathsf{s}_M\})$, where $M = A$, $\mathsf{k}_M = \{k\}$, $\mathsf{s}_M = \{s\}$, and $a \bullet b = \{a \bullet_A b\}$ for all $a, b \in A$. Then M is isomorphic to A under the element-singleton bijection.

Examples 5 and 6 show that the functional interpretation (of application and constants) is a special instance of the set-theoretic interpretation of matching logic, and that applicative structures, combinatory algebras, and λ-models are special instances of matching logic models. In Section 4, we will show how to enforce functional interpretation in matching logic models, *axiomatically*.

We continue with the semantics of matching logic and define the interpretation of patterns.

Definition 7. Let M be a matching logic model like in Definition 4. We extend the interpretation of application $\_\bullet\_$ *pointwisely*, from over elements to over sets, as $A \bullet B = \bigcup_{a \in A, b \in B} a \bullet b$ for any $A, B \subseteq M$. An *M-valuation* (or simply *valuation*), written $\rho\colon (EV \cup SV) \to M \cup \mathcal{P}(M)$, is a function that maps element variables to elements and set variables to sets, i.e., $\rho(x) \in M$ for $x \in EV$ and $\rho(X) \subseteq M$ for $X \in SV$. It yields a *pattern valuation*, written $|\_|_\rho : \text{PATTERN} \to \mathcal{P}(M)$, defined as:

1. $|x|_\rho = \{\rho(x)\}$ for $x \in EV$;
2. $|X|_\rho = \rho(X)$ for $X \in SV$;
3. $|\sigma|_\rho = \sigma_M$ for $\sigma \in \Sigma$;
4. $|\varphi_1\, \varphi_2|_\rho = |\varphi_1|_\rho \bullet |\varphi_2|_\rho$, where $\_\bullet\_$ is pointwisely extended to sets;
5. $|\bot|_\rho = \emptyset$;
6. $|\varphi_1 \to \varphi_2|_\rho = M \setminus (|\varphi_1|_\rho \setminus |\varphi_2|_\rho)$, where "$\setminus$" denotes set difference;
7. $|\exists x.\, \varphi|_\rho = \bigcup_{a \in M} |\varphi|_{\rho[a/x]}$, where $\rho[a/x]$ is the valuation ρ' such that $\rho'(x) = a$, $\rho'(y) = \rho(y)$ for all $y \in EV$ distinct from x, and $\rho'(X) = \rho(X)$ for all $X \in SV$.

Remark 8. The above semantic rules should not be unexpected. Rules (1) and (2) interpret variables according to ρ. Rules (3) and (4) interpret symbols and application according to M. For rules (5)-(7), if we regard $\emptyset$ as "false" and M as "true", then these rules become precisely the FOL semantic rules of bottom, implication, and $\exists$-quantification, respectively.

We can prove that the derived constructs in Fig. 1 have the expected semantics:

Proposition 9. The following propositions hold:

1. $|\neg\varphi|_\rho = M \setminus |\varphi|_\rho$;
2. $|\varphi_1 \vee \varphi_2|_\rho = |\varphi_1|_\rho \cup |\varphi_2|_\rho$;
3. $|\varphi_1 \wedge \varphi_2|_\rho = |\varphi_1|_\rho \cap |\varphi_2|_\rho$;
4. $|\top|_\rho = M$;
5. $|\varphi_1 \leftrightarrow \varphi_2|_\rho = M \setminus (|\varphi_1|_\rho \triangle |\varphi_2|_\rho)$, where "$\triangle$" denotes set symmetric difference;
6. $|\forall x.\, \varphi|_\rho = \bigcap_{a \in M} |\varphi|_{\rho[a/x]}$.

Proof. We only prove (1) and (6). The others are in Appendix A. For (1), we have $|\neg\varphi|_\rho = |\varphi \to \bot|_\rho = M \setminus (|\varphi|_\rho \setminus |\bot|_\rho) = M \setminus (|\varphi|_\rho \setminus \emptyset) = M \setminus |\varphi|_\rho$. For (6), we have $|\forall x.\, \varphi|_\rho = |\neg\exists x.\, \neg\varphi|_\rho = M \setminus |\exists x.\, \neg\varphi|_\rho = M \setminus \bigcup_{a \in M} |\neg\varphi|_{\rho[a/x]} = M \setminus \bigcup_{a \in M} (M \setminus |\varphi|_{\rho[a/x]}) = M \setminus (M \setminus \bigcap_{a \in M} |\varphi|_{\rho[a/x]}) = \bigcap_{a \in M} |\varphi|_{\rho[a/x]}$. □

Remark 10. Definition 7 and Proposition 9 show that there is a close connection between the matching logic pattern constructs and the set operations in set theory: conjunction corresponds to intersection of two sets; disjunction corresponds to union of two sets; negation corresponds to set complement; top ($\top$) corresponds to the total set; bottom ($\bot$) corresponds to the empty set; $\exists$-quantification corresponds to the (big) union of a collection of sets; and $\forall$-quantification corresponds to the (big) intersection of a collection of sets. This connection to the set-theoretic operations can be useful to understand the intuitive meaning of complex matching logic patterns.

3.2.1 Predicate Patterns

A difference between FOL formulas and matching logic patterns is that the former can only be interpreted as either true or false, while the latter can be interpreted as any subsets of the carrier set. Following up on Remark 8, we identify two special sets, M and $\emptyset$, and use them to represent (logical) true and false, respectively. Obviously, not all patterns are interpreted as M or $\emptyset$. Given a model M, we call φ an *M-predicate*, if $|\varphi|_\rho \in \{\emptyset, M\}$ for all ρ. We call φ a *predicate* (or *predicate pattern*), if it is an M-predicate in all M. Predicate patterns can be built from $\bot$, $\top$, and matching logic logical constructs, e.g., $\forall x.\, (\sigma\, x) \wedge \neg(\sigma\, x)$. More interesting patterns can be built from symbols and application. For example, $\sigma\, x_1 \cdots x_n$ is a predicate pattern if the underlying matching logic theory (discussed in Section 3.3) enforces the models to interpret σ as a predicate (i.e., either $\emptyset$ or M). We will see more predicate patterns in Section 4 and throughout the paper. Roughly speaking, predicate patterns are the matching logic counterparts of FOL formulas. They make "statements", and can take only two possible values: M if the statements are facts, and $\emptyset$ if the statements are not facts. Note that except the application, all matching logic constructs (primitive or derived) preserve the predicate-ness of patterns. We can then use application to build FOL-style predicates, and this way regard predicate logic as a methodological fragment of matching logic.

3.2.2 Functional Patterns

Examples 5 and 6 emphasized that any set M is isomorphic to the set of singletons of M, and that functional interpretation is a special instance of set-theoretic interpretation. Formally, given M, we call φ an *M-functional pattern* if $|\varphi|_\rho$ is a singleton for all ρ. We call φ a *functional pattern*, if it is an M-functional pattern for all M. Roughly speaking, functional patterns are the matching logic counterparts of FOL terms. A functional pattern denotes exactly one element; e.g., x is the simplest functional pattern. More interesting functional patterns can be built by symbols and application; e.g., $\sigma\, x_1 \cdots x_n$ is a function pattern if the underlying matching logic theory (discussed in Section 3.3) enforces the models to interpret σ as a function. We will show many examples of functional patterns in Section 4 and throughout the paper.

3.3 Matching Logic Theories

Examples 5 and 6 show that we sometimes want to consider only a subclass of matching logic models, those that satisfy certain properties. This can be achieved by defining a matching logic *theory*—a set of patterns regarded as *axioms*—and considering only the satisfying models. Formally:

Definition 11. For M and φ, we say M *validates* φ, or φ *holds* in M, written $M \models \varphi$, iff $|\varphi|_\rho = M$ for all ρ. For a pattern set Γ, we say M *validates* Γ, written $M \models \Gamma$, iff $M \models \psi$ for all $\psi \in \Gamma$. We write $\Gamma \models \varphi$, iff $M \models \Gamma$ implies $M \models \varphi$ for all M. A matching logic *theory* $(\mathbb{\Sigma}, \Gamma)$ is a pair, where $\mathbb{\Sigma}$ is a signature and Γ is a set of $\mathbb{\Sigma}$-patterns. We often abbreviate $(\mathbb{\Sigma}, \Gamma)$ as Γ, if $\mathbb{\Sigma}$ is understood.

Note that φ holds in M if it represents a "logical truth", i.e., its interpretation is the total set M.

Remark 12. The axiom set Γ may contain patterns that have free variables. By Definition 11, free (element and set) variables are effectively *universally quantified*, as we need to check the validity of each axiom on all possible valuations. Free element variables in an axiom can be eliminated using $\forall$-quantification, defined in Fig. 1, as in FOL. However, free set variables in an axiom *cannot be eliminated*, because $\forall$-quantification is not applicable to set variables. Allowing free set variables in axioms to be effectively universally quantified, makes matching logic more expressive (in terms of capturing models) than FOL (see Section 4.4), and comparable to the fragment of *monadic second-order logic* [29, 91] where all quantifiers over sets are universal quantifiers and only appear at the top.

We will define various matching logic theories in the rest of the paper. To define a theory, we need to define its sets of element variables, set variables, symbols, and axioms. We often omit explicit definitions of the variable sets and only specify the symbol and axiom sets. For readability, we mix the definitions of the symbol and axiom sets in our narrative texts. For example, when we say "we consider/define a symbol $\sigma \in \Sigma$", we mean to add σ to the symbol set of the theory we are defining. Similarly, when we say that "we define/assume an axiom ψ", we mean to add ψ to the axiom set of the theory we are defining. We will often define a theory Γ' by building it upon another more basic theory Γ. In that case, Γ' is assumed to include all components of Γ.

4 Important Mathematical Instruments

In this section, we (axiomatically) define several important mathematical instruments, like functions and equality, which are required in order to define binders as *theories* within matching logic (as opposed to extensions of the logic). We also propose appropriate notations for them. In Section 4.1, we define the *definedness symbol* and use it to define equality, membership, set-theoretic inclusion, and functional constants. In Section 4.2, we define the *inhabitant symbol* and use it to define sorts, subsorting, and many-sorted functions and partial functions. This allows us to reason about sorts and to capture logical systems with sorts, in the unsorted matching logic. In Sections 4.3 and 4.4, we define matching logic theories that completely capture the models of *product sets* and *powersets*.

4.1 Definedness Symbol and Related Instruments

Recall the pattern matching semantics of matching logic: the interpretation of pattern φ is the set of elements that match it. When φ is matched by at least one element, we say that φ is *defined*. The definedness symbol (Definition 13) takes any pattern φ, and builds a new *definedness pattern* $\lceil \varphi \rceil$, which is a predicate pattern stating that φ is defined. Many important mathematical instruments such as equality and membership, can be derived from the definedness symbol as syntactic sugar.

Definition 13. Let us consider a (constant) symbol written $\lceil \_ \rceil \in \Sigma$, which we call the *definedness symbol*. We write $\lceil \varphi \rceil$ to mean $\lceil \_ \rceil \, \varphi$, obtained by applying $\lceil \_ \rceil$ to φ. We define the following axiom:

$$\text{(Definedness)} \quad \lceil x \rceil \qquad // \text{ or, equivalently, } \forall x \,.\, (\lceil \_ \rceil \, x)$$

We define totality $\lfloor \_ \rfloor$, equality $\_ = \_$, membership $\_ \in \_$, and set inclusion $\_ \subseteq \_$ as derived constructs:

$$\lfloor \varphi \rfloor \equiv \neg\lceil\neg\varphi\rceil \qquad \varphi_1 = \varphi_2 \equiv \lfloor \varphi_1 \leftrightarrow \varphi_2 \rfloor \qquad x \in \varphi \equiv \lceil x \wedge \varphi \rceil \qquad \varphi_1 \subseteq \varphi_2 \equiv \lfloor \varphi_1 \rightarrow \varphi_2 \rfloor$$

Intuitively, (Definedness) states that every individual element x is defined. This is clearly true with our intended meaning of $\lceil \_ \rceil$, because x is matched by *exactly one element* to which it evaluates; this intended meaning is precisely what the (Definedness) axiom captures. Specifically, in any model that validates (Definedness), $\lceil x \rceil$ is interpreted as the total set, according to matching logic validity (Definition 11). Now, consider any pattern φ that is defined, and that φ is matched by one element, say x. By *pointwise extension* (Definition 7), the interpretation of $\lceil \varphi \rceil$ must include the interpretation of $\lceil x \rceil$, which we know is the total set. Therefore, $\lceil \varphi \rceil$ is also interpreted as the total set, as intended. On the other hand, if φ is *undefined*, its interpretation is the empty set, and by pointwise extension, $\lceil \varphi \rceil$ is also interpreted as the empty set. This intuition is formalized below.

Proposition 14. For any model M, patterns $\varphi, \varphi_1, \varphi_2$, element variable x, and valuation ρ, we have

1. $\lceil a \rceil_M = M$ for any $a \in M$, where $\lceil a \rceil_M$ means $\lceil \_ \rceil_M \cdot a$ and $\lceil \_ \rceil_M$ is the interpretation of $\lceil \_ \rceil$;
2. $|\lceil \varphi \rceil|_\rho = M$ if $|\varphi|_\rho \neq \emptyset$; otherwise, $|\lceil \varphi \rceil|_\rho = \emptyset$;
3. $|\lfloor \varphi \rfloor|_\rho = M$ if $|\varphi|_\rho = M$; otherwise, $|\lfloor \varphi \rfloor|_\rho = \emptyset$;
4. $|\varphi_1 = \varphi_2|_\rho = M$ if $|\varphi_1|_\rho = |\varphi_2|_\rho$; otherwise, $|\varphi_1 = \varphi_2|_\rho = \emptyset$;
5. $|x \in \varphi|_\rho = M$ if $\rho(x) \in |\varphi|_\rho$; otherwise, $|x \in \varphi|_\rho = \emptyset$;
6. $|\varphi_1 \subseteq \varphi_2|_\rho = M$ if $|\varphi_1|_\rho \subseteq |\varphi_2|_\rho$; otherwise, $|\varphi_1 \subseteq \varphi_2|_\rho = \emptyset$; note that $|x \subseteq \varphi|_\rho = |x \in \varphi|_\rho$;

Note that all the above patterns in (2)-(6) are predicate patterns (Section 3.2.1).

Not all models validate (Definedness). Indeed, as said in Section 3.3, the purpose of axioms and theories is to restrict models under consideration. As an example, a model whose interpretation of application is a function that always returns the empty set does not validate (Definedness), as it fails to satisfy Proposition 14(1). On the other hand, models that satisfy (Definedness) are also easy to come by. A canonical example is a model M with one distinguished element $\#\mathsf{def}$ such that $\#\mathsf{def} \cdot a = M$ for all $a \in M$, and let $\lceil \_ \rceil_M$, the interpretation of $\lceil \_ \rceil$, to be $\{\#\mathsf{def}\}$. Then we have $|\lceil x \rceil|_\rho = \lceil \_ \rceil_M \cdot \rho(x) = \{\#\mathsf{def}\} \cdot \{\rho(x)\} = \#\mathsf{def} \cdot \rho(x) = M$, and thus M validates (Definedness). In fact, any model can be extended into one that validates (Definedness) by adding an element like $\#\mathsf{def}$ above to it and letting $\lceil \_ \rceil_M$ be $\{\#\mathsf{def}\}$. Since definedness is so useful, we assume it in all subsequent theories defined in this paper, and hereby we do not consider the models that do not satisfy the axiom (Definedness).

Remark 15. We explain why defining equality needs the definedness symbol, when there is already the logical biconditional construct $\varphi_1 \leftrightarrow \varphi_2$, given in Fig. 3. It is *not always the case* that $|\varphi_1 = \varphi_2|_\rho = |\varphi_1 \leftrightarrow \varphi_2|_\rho$ for all ρ. By Proposition 14, $\varphi_1 = \varphi_2$ is a *predicate* stating that φ_1 and φ_2 are matched by the same set of elements, while by Proposition 9, $\varphi_1 \leftrightarrow \varphi_2$ is a pattern (not necessarily a predicate) that is matched by the elements a, such that a matches φ_1 iff a matches φ_2. If $|\varphi_1|_\rho = |\varphi_2|_\rho$, then both $\varphi_1 \leftrightarrow \varphi_2$ and $\varphi_1 = \varphi_2$ are interpreted as the total set, but if otherwise, $\varphi_1 = \varphi_2$ is interpreted as the empty set, while $\varphi_1 \leftrightarrow \varphi_2$ is the complement of set difference. The fact that we can define equality axiomatically, i.e. without extending the logic, to mean *precise* identity in models is particularly useful in our subsequent developments, albeit surprising. Indeed, it is well-known that equality cannot be defined in FOL (which justifies the extension of FOL *with equality*), while in second-order logic it requires quantification over sets.

As a simple example, we can use the definedness symbol (and derived constructs) to axiomatize *functional constants*, which are matching logic symbols whose interpretations are singletons.

Example 16. Let $\sigma \in \Sigma$ be a matching logic symbol. Let us consider the following axiom

$$\text{(Functional Constant)} \quad \exists x.\, \sigma = x$$

Then for any model M that validates this axiom, we have $|\exists x.\, \sigma = x|_\rho = \bigcup_{a \in M} |\sigma = x|_{\rho[a/x]} = M$. By Proposition 14, $|\sigma = x|_{\rho[a/x]}$ is either $\emptyset$ or M, so there exists $a \in M$ such that $|\sigma = x|_{\rho[a/x]} = M$, which implies that $\sigma_M = |x|_{\rho[a/x]} = \{a\}$, i.e., σ is interpreted as a singleton in M.

4.2 Inhabitant Symbol and Related Instruments

Matching logic is an unsorted logic, but we can capture sorts by defining a set of functional constants (Example 16) that represent the *names* of the sorts, and define a special symbol, which we call the *inhabitant symbol*, to get the actual *inhabitant set* of each sort. This intuition is made formal below. From now on, we will always assume the definedness symbol and the (Definedness) axiom.

Definition 17. A *sort constant* (or simply *sort*) is a symbol $s \in \Sigma$, which is a functional constant, as defined in Example 16. Let us consider another symbol $[\![\_]\!] \in \Sigma$, which we call the *inhabitant symbol*. We write $[\![s]\!]$ to mean $[\![\_]\!]\, s$, obtained by applying $[\![\_]\!]$ to s, and call it the *inhabitant of* s.

In other words, the pattern s is matched by the sort name s itself, while $[\![s]\!]$ is matched by the actual elements of sort s. For example, for two sorts *Nat* and *Int* of natural and integer numbers, *Nat* is matched by one element—the sort name *Nat*; *Int* is matched by one element—the sort name *Int*; $[\![Nat]\!]$ is matched by all natural numbers; and $[\![Int]\!]$ is matched by all integer numbers. Note that Definition 17 does not enforce any particular axioms about sorts or the inhabitant symbol. Their interpretations are determined by the models and can be constrained by axioms. For example, subsorting $s_1 \leq s_2$ is a partial ordering on sorts that enforces the subset relation between the inhabitants of s_1 and s_2. In matching logic, subsorting can be axiomatically captured:

$$\text{(Subsorting)} \quad [\![s_1]\!] \subseteq [\![s_2]\!]$$

which states that the inhabitant of s_1 is included in the inhabitant of s_2. In this paper we use subsorting to define the syntax of λ-calculus and other logical systems that feature bindings. In Section 6 we define a sort *Var* for λ-calculus variables and a sort *Exp* for λ-expressions, and we define the *subsorting axiom* $[\![Var]\!] \subseteq [\![Exp]\!]$ to specify that λ-calculus variables are also λ-expressions.

4.2.1 Sorted Quantification

The meaning of $\exists x.\, \varphi$ is the set-theoretic (big) union of the interpretations of φ, with x ranging over all elements in the carrier set (see Remark 10). Now that we have defined sorts, we will want to *restrict* x to range over not all elements, but only those having sort s. For that, we define the following self-explanatory derived constructs, called *sorted quantification*:

$$\exists x{:}s.\, \varphi \equiv \exists x.\, (x \in [\![s]\!] \wedge \varphi) \qquad\qquad \forall x{:}s.\, \varphi \equiv \forall x.\, (x \in [\![s]\!] \to \varphi)$$

4.2.2 Many-Sorted Functions

Given sorts $s, s_1, \ldots, s_n$, we call a (constant) symbol $f \in \Sigma$ a *many-sorted function* from $s_1, \ldots, s_n$ to s, written $f\colon s_1 \times \cdots \times s_n \to s$, if it satisfies the axiom:

$$\text{(Function)} \quad \forall x_1{:}s_1.\, \ldots \forall x_n{:}s_n.\, \exists y{:}s.\, f\, x_1\, \cdots\, x_n = y \tag{3}$$

Application is left-associative (Definition 2), so $f\, x_1\, \cdots\, x_n$ means $(\cdots (f\, x_1)\, \cdots\, x_n)$. Intuitively, (Function) requires that $f\, x_1\, \cdots\, x_n$ consist of exactly one element, y, which is an inhabitant of s, given that $x_1, \ldots, x_n$ are inhabitants of $s_1, \ldots, s_n$, respectively. Note that while f, fx_1, fx_1x_2, ..., $fx_1 \cdots x_{n-1}$ are all well-formed patterns, they are not required to consist of exactly one element.

4.2.3 Many-Sorted Partial Functions

The axiom (Function) above is not unusual; it translates to matching logic a standard encoding of many-sorted functions using an unsorted logic (see [67, pp. 8] for a related discussion). What is a lot harder problem is how to capture *partial functions* that can be undefined in some arguments. Capturing partial functions in a formal system is not just of theoretical interest. It is also a practical concern that has arisen in the formal verification of programs with exceptional expressions, such as division by zero or the head of an empty list, and has resulted in work on partial algebras [18], exception algebras [11], error algebras [46], order-sorted algebras [47], and various logics for partial functions [2, 60].

On the other hand, it is surprisingly easy to capture partial functions in matching logic. We take the axiom (Function) and change the equality $\_ = \_$ to set inclusion $\_ \subseteq \_$:

$$\text{(Partial Function)} \quad \forall x_1{:}s_1.\ \ldots \forall x_n{:}s_n.\, \exists y{:}s.\, f\, x_1\, \cdots\, x_n \subseteq y \tag{4}$$

Intuitively, (Partial Function) requires $f\, x_1\, \cdots\, x_n$ to consist of *at most* one element. The *undefinedness* of f on $x_1, \ldots, x_n$ is captured, by $f\, x_1\, \cdots\, x_n$ returning the empty set $\emptyset$. For notional simplicity, we will write $f\colon s_1 \times \cdots \times s_n \rightharpoonup s$ to mean that f is a partial function from $s_1, \ldots, s_n$ to s.

The reason why partial functions can be directly defined using (Partial Function), without needing to extend or modify matching logic, is due to the pattern matching semantics of matching logic, where patterns are not restricted to a functional interpretation, and are given a more general, set-theoretic interpretation, which unifies (both syntactically and semantically) total functions and FOL terms, predicates and FOL formulas, and partial functions and partial terms.

4.3 Product Sorts

In this and the next sections, we assume the definedness symbol, the inhabitant symbol, and all the related instruments that are given in Sections 4.1 and 4.2. Our goal in this section is to axiomatize the *product sort* $s_1 \otimes s_2$, whose (intended) inhabitant is the (set-theoretic) product of the inhabitants of s_1 and s_2, up to isomorphism. Formally:

Definition 18. Given two sorts s_1, s_2, we consider a functional constant $s_1 \otimes s_2 \in \Sigma$, which we call the *product (sort) of s_1 and s_2*. We define a function $\langle \_, \_ \rangle\colon s_1 \times s_2 \to s_1 \otimes s_2$, called *pairing*, where the function notation was introduced in Section 4.2.2. We write $\langle \varphi_1, \varphi_2 \rangle$ to mean $\langle \_, \_ \rangle\, \varphi_1\, \varphi_2$, obtained by applying $\langle \_, \_ \rangle$ to φ_1, and then to φ_2. We define the following two axioms:

$$\begin{array}{ll} \text{(Product)} & [\![s_1 \otimes s_2]\!] = \exists x_1{:}s_1.\, \exists x_2{:}s_2.\, \langle x_1, x_2 \rangle \\ \text{(Injectivity)} & \forall x_1{:}s_1.\, \forall x_2{:}s_2.\, \forall y_1{:}s_1.\, \forall y_2{:}s_2.\, \langle x_1, x_2 \rangle = \langle y_1, y_2 \rangle \to x_1 = y_1 \wedge x_2 = y_2 \end{array}$$

Intuitively, $\langle x_1, x_2 \rangle$ denotes the pair consisting of x_1 and x_2. (Product) states that the inhabitant of $s_1 \otimes s_2$ is the product of the inhabitants of s_1 and s_2. (Injectivity) states that $\langle \_, \_ \rangle$ is injective.

Proposition 19. For any model M validating the axioms in Definition 18, we have $M_{s_1 \otimes s_2} \cong M_{s_1} \times M_{s_2}$, where we use $M_s = [\![\_]\!]_M \bullet s_M$ to denote the inhabitant of s in M, for any sort s.

4.4 Power Sorts

Our goal in this section is to axiomatize the power sort 2^s, whose (intended) inhabitant is the powerset of the inhabitant of s, up to isomorphism. Formally:

Definition 20. Given a sort s, let us consider a functional constant $2^s \in \Sigma$, which we call the *power (sort) of s*. For clarity, we use the Greek letters $\alpha, \beta, \ldots$ for element variables whose intended range is in sort 2^s. Let us define a (constant) symbol $\mathsf{extension} \in \Sigma$, called the *extension symbol* (explained later), and define the following axioms:

(ARITY) $\forall \alpha{:}2^s.\, (\mathsf{extension}\, \alpha) \subseteq [\![s]\!]$
(POWERSET) $X \subseteq [\![s]\!] \to \exists \alpha{:}2^s.\, (\mathsf{extension}\, \alpha) = X$
(EXTENSIONALITY) $\forall \alpha{:}2^s.\, \forall \beta{:}2^s.\, (\mathsf{extension}\, \alpha) = (\mathsf{extension}\, \beta) \to \alpha = \beta$

Note that set variable X is free in (POWERSET). By Remark 12, it is effectively universally quantified.

Definition 20 needs some explanation. Let us consider an intended model M, where the inhabitant of s is M_s and the inhabitant of 2^s is $M_{2^s} = \mathcal{P}(M_s)$, i.e., the *powerset* of M_s. We use $a, b, \cdots \in M_s$ to denote elements in M_s and $A, B, \cdots \in M_{2^s}$ to denote elements in M_{2^s}, i.e., subsets of M_s. Note that α is an element variable of sort 2^s, so let us assume it evaluates to some $A \in M_{2^s}$. Then, the intended, intuitive meaning of $(\mathsf{extension}\, \alpha)$, is that it is a pattern (of sort s) that is matched by all elements a in A. Please note the difference between α and $(\mathsf{extension}\, \alpha)$. On one hand, α is an element variable of sort 2^s, so it is matched by one "element" A. On the other hand, $(\mathsf{extension}\, \alpha)$ is a pattern of sort s, so it is matched by all elements in the set A. In other words, A is regarded as an individual "element" in sort 2^s but a real "set" in sort s, on which the pointwise extension (Definition 7) can apply. Thus, the matching logic symbol "$\mathsf{extension}$" takes A as an element and returns A itself as a set. This has a similar meaning to the term "extension" in logic and philosophy—an extension of a concept consists of the things to which it applies. Here, we regard the element A of the powerset as an intensional concept and the set A of its elements as its extension.

With the above intuition, the axioms in Definition 20 are self-explanatory. (ARITY) states that $(\mathsf{extension}\, \alpha)$ has sort s whenever α has sort 2^s. (POWERSET) states that any subset of the inhabitant of s, ranged by X, has a corresponding "element" denoted α whose extension is X. Therefore, the inhabitant of 2^s is *at least* as large as the powerset of the inhabitant of s. On the other hand, (EXTENSIONALITY) states that α and β are equal whenever their extensions are equal, so the inhabitant of 2^s is *at most* as large as the powerset of the inhabitant set s. Putting the arguments together, we show that the inhabitant of 2^s is the powerset of the inhabitant of s, up to isomorphism:

Proposition 21. For any model M validating the axioms in Definition 20, we have $M_{2^s} \cong \mathcal{P}(M_s)$.

The reverse of $\mathsf{extension}$, called *intension*, can be defined as the following syntactic sugar:

$$\mathsf{intension}\, \varphi \equiv \exists \alpha{:}2^s.\, \alpha \wedge (\mathsf{extension}\, \alpha = \varphi)$$

Intuitively, φ has sort s; $(\mathsf{intension}\, \varphi)$ has sort 2^s, and is matched by the *unique* element α of sort 2^s such that $\mathsf{extension}\, \alpha = \varphi$; the uniqueness is guaranteed by the axiom (EXTENSIONALITY).

Remark 22. Proposition 21 shows that powersets can be completely, finitely axiomatized in matching logic. This result is known to *not hold* in FOL, because by the Löwenheim-Skolem theorem [59], if a FOL theory has infinite models, then it has a countable model. However, using powersets, we can enforce uncountable models by first enforcing an infinite model and considering its powerset. As an example, we define natural numbers *Nat* using *zero* and *suc*, and define the standard injectivity axioms $zero \neq suc(x)$ and $suc(x) = suc(y) \to x = y$ to enforce *Nat* to be infinite, as it must contain *zero*, $suc(zero)$, $suc(suc(zero))$, etc., which are all distinct. If powersets could have been completely axiomatizable in FOL, then we could define the powerset of natural numbers 2^{Nat} that is uncountable, contradicting the Löwenheim-Skolem theorem.

4.5 Matching Logic Proof System

There is a Hilbert-style proof system for matching logic that defines the provability relation $\Gamma \vdash \varphi$ for matching logic theory Γ and pattern φ. The proof system is not needed in order to understand the technical results discussed in this paper (see Appendix B.3). We only review some meta-theorems about the proof system, which are needed in order to prove the subsequent results, mentioning that any (sound) proof system that has these properties would be equally suitable:[3]

Proposition 23. If Γ contains the definedness symbol and the axiom (DEFINEDNESS), then

[3]Note that Γ is different from typing contexts in type systems (see, e.g., [19]) that share variables with judgment φ. Here, Γ has variables independent from φ and its axioms are implicitly universally quantified; see also Remark 12.

free variables:
$\mathrm{FV}(x) = \{x\}$ $\quad \mathrm{FV}(e_1\, e_2) = \mathrm{FV}(e_1) \cup \mathrm{FV}(e_2)$ $\quad \mathrm{FV}(\lambda x.\, \varphi) \equiv \mathrm{FV}(\varphi) \setminus \{x\}$
α-renaming:
$\lambda x.\, \varphi \equiv \lambda y.\, \varphi[y/x]$, for $y \notin \mathrm{FV}(\varphi)$
capture-free substitution (where y distinct from x and z is fresh):
$(\lambda x.\, \varphi)[\psi/x] \equiv \lambda x.\, \varphi$ $\quad (\lambda x.\, \varphi)[\psi/y] \equiv \lambda z.\, \varphi[z/x][\psi/y]$

Figure 3: Meta-properties about binder λ, similar to those for the binder $\exists$ in matching logic (Fig. 1).

1. $\Gamma \vdash \varphi$, if φ is a propositional tautology over patterns;
2. $\Gamma \vdash \varphi_1$ and $\Gamma \vdash \varphi_1 \to \varphi_2$ imply $\Gamma \vdash \varphi_2$;
3. $\Gamma \vdash \varphi[y/x] \to \exists x.\, \varphi$;
4. $\Gamma \vdash \varphi_1 \to \varphi_2$ and $y \notin \mathrm{FV}(\varphi_2)$ imply $\Gamma \vdash (\exists y.\, \varphi_1) \to \varphi_2$;
5. $\Gamma \vdash \varphi = \varphi$;
6. $\Gamma \vdash \varphi_1 = \varphi_2$ and $\Gamma \vdash \varphi_2 = \varphi_3$ imply $\Gamma \vdash \varphi_1 = \varphi_3$;
7. $\Gamma \vdash \varphi_1 = \varphi_2$ implies $\Gamma \vdash \varphi_2 = \varphi_1$;
8. $\Gamma \vdash \varphi_1 = \varphi_2$ implies $\Gamma \vdash \psi[\varphi_1/x] = \psi[\varphi_2/x]$, known as the Leibniz characterization of equality.

Proposition 23 essentially states that FOL with equality reasoning is supported by the proof system of matching logic, where patterns are conveniently regarded as either "predicates" or "terms", depending on the context. We require Γ to contain the definedness symbol and axiom, because they are needed to define equality $\varphi_1 = \varphi_2$, as discussed in Definition 13.

We review the following *soundness theorem* of the matching logic proof system:

Theorem 24 (Soundness Theorem)**.** $\Gamma \vdash \varphi$ *implies* $\Gamma \vDash \varphi$.

While several *(deductive) completeness results* (i.e., $\Gamma \vDash \varphi$ implies $\Gamma \vdash \varphi$) have been proved for some theories Γ in [81, 21], it is incomplete in general for all Γ and φ. Fortunately, it does not affect this paper. Instead, we prove a *new completeness result* as a corollary of the conservative extension theorem of λ-calculus (Theorem 36), where Γ is the matching logic theory that captures λ-calculus and φ is an equation between λ-expressions; see Section 5.

5 λ-Calculus Preliminaries

The syntax of λ-calculus [28] is parametric in a set of variables V^λ, whose elements are written $x, y, \ldots$. The set Λ of λ-*expressions* is inductively defined by the following grammar:

$$e ::= x \mid e_1\, e_2 \mid \lambda x.\, e$$

Free variables $\mathrm{FV}(e)$, α-equivalence $e_1 \equiv e_2$, and capture-free substitution $e[e'/x]$ are defined as usual, shown in Fig. 3. We regard α-equivalent λ-expressions as identical expressions.

In λ-calculus, we are interested in proving equations of the form $e_1 = e_2$, for $e_1, e_2 \in \Lambda$. Equational reasoning in λ-calculus includes the standard reflexivity, symmetry, transitivity, and congruence proof rules, and the distinguished (β) *axiom schema* that specifies the result of function application:

$$(\beta) \quad (\lambda x.\, e)\, e' = e[e'/x] \quad \text{for all } x \in V^\lambda \text{ and } e, e' \in \Lambda$$

We write $\vdash_\lambda e_1 = e_2$ to mean that $e_1 = e_2$ is provable in λ-calculus.

5.1 Our Goal and the Main Challenges

Our first goal is to define a matching logic theory Γ^λ that faithfully captures λ-calculus, in the sense that λ-expressions are well-formed matching logic patterns and λ-reasoning is captured by matching logic reasoning. Formally, our goal is to prove the *conservative extension* theorem:

$$\Gamma^\lambda \vdash e_1 = e_2 \quad \underset{\text{extensiveness}}{\overset{\text{conservativeness}}{\rightleftharpoons}} \quad \vdash_\lambda e_1 = e_2 \text{ for all } e_1, e_2 \in \Lambda \tag{5}$$

which says that we can safely reduce λ-calculus reasoning to matching logic reasoning, without proving fewer or more equations between λ-expressions. Specifically, the extensiveness direction means that all provable equations between λ-expressions can also be proved in Γ^λ, which is thus an extension of λ-calculus, while the conservativeness direction says that no additional equations between λ-expressions can be proved. Note that we are only concerned with equations *between λ-expressions*. Since matching logic has a richer syntax than λ-calculus, of course there are equations, e.g. $\bot = \bot$, which are provable in matching logic but do not even exist in λ-calculus.

Main Challenges There are two main challenges. The first challenge is to capture the binding behavior of λ, that is, to define $\lambda x.\, e$ as *syntactic sugar* in matching logic such that it satisfies the properties about free variables, α-equivalence, and capture-free substitution in Fig. 3. The key observation is that λ plays two important roles: (i) it builds a *term* $\lambda x.\, e$, and (ii) it builds a *binding* of x into e. Matching logic allows us to separate these two roles, where we define terms using symbols and application as shown in Section 4 and bindings using matching logic's built-in binder $\exists$.

The other challenge is to prove the conservative extension theorem shown as Eq. (5). The extensiveness direction is easy, because equational reasoning is supported in matching logic (Proposition 23). We only need to include all instances of (β) in Γ^λ. The conservativeness direction is more involved and is a major technical contribution of this paper. Indeed, matching logic has a richer syntax and a more complex proof system than λ-calculus; we need to show that this more complex infrastructure cannot be used to prove more equations between λ-expressions.

5.2 Our Plan

We will give two different proofs for the conservativeness of Γ^λ, each providing a unique insight about the construction of Γ^λ. The first is based on a model theory of λ-calculus, discussed in Section 7. It considers a special class of λ-calculus models, called concrete Cartesian closed category models, or simply concrete ccc models, which are known to be complete with respect to λ-calculus reasoning (Lemma 26). This model-based proof is easier to understand due to its close connection to the models, and is what inspired our encoding of the λ binder in matching logic (see Eq. (1)). However, it does not generalize to other logical systems with binders that do not have well-established models. Hence, in Section 8 we give an alternative conservativeness proof, based on the syntax and proof derivations of λ-calculus, and not on models. The syntax-based proof does not depend on the existence of a complete class of models, and is thus easier to generalize to other logical systems.

5.3 Concrete ccc Models of λ-Calculus

We review the concrete Cartesian closed category (ccc) models of λ-calculus [6, Definition 5.5.9]. They will be used in the model-based proof of the conservativeness of Γ^λ.

Definition 25 ([9, Definition 57])**.** Given an applicative structure $(A, \_\bullet_A\_)$, its set of *representable functions* is $R(A) = \{f\colon A \to A \mid \text{there is a } b \in A \text{ such that } f(a) = b \bullet_A a \text{ for all } a \in A\}$. A *pre-model* is a triple $(A, \_\bullet_A\ \_, \mathbb{L})$, where $\mathbb{L}\colon R(A) \to A$ is a *retraction function* such that $\mathbb{A} \circ \mathbb{L}$ is the identity on $R(A)$, where $\mathbb{A}\colon A \to R(A)$ is defined as $\mathbb{A}(b)(a) = b \bullet_A a$ for all $b, a \in A$. A pre-model A is called a *concrete ccc model*, if the following definition of $|e|^\lambda_\rho$ is well-defined for every $\rho\colon V^\lambda \to A$:

$$\begin{array}{lllll} \Gamma^\lambda \vdash e_1 = e_2 & \Longrightarrow_1 & \Gamma^\lambda \vDash e_1 = e_2 & \Longrightarrow_2 & M \vDash e_1 = e_2 \text{ for all matching logic models } M \vDash \Gamma^\lambda \\ & & & & \Downarrow_3 \\ \vdash_\lambda e_1 = e_2 & \Longleftarrow_5 & \vDash_\lambda e_1 = e_2 & \Longleftarrow_4 & A \vDash_\lambda e_1 = e_2 \text{ for all concrete ccc models } A \end{array}$$

Figure 4: The main proof steps of the model-based conservativeness proof of Γ^λ.

1. $|x|^\lambda_\rho = \rho(x)$;
2. $|e_1 e_2|^\lambda_\rho = |e_1|^\lambda_\rho \bullet_A |e_2|^\lambda_\rho$;
3. $|\lambda x.\, e|^\lambda_\rho = \mathbb{L}(f^\rho_{e,x})$ where $f^\rho_{e,x}(a) = |e|^\lambda_{\rho[a/x]}$ for $a \in A$, and that $f^\rho_{e,x} \in R(A)$.

Given a concrete ccc model A, we write $A \vDash_\lambda e_1 = e_2$ iff $|e_1|^\lambda_\rho = |e_2|^\lambda_\rho$ for all ρ. We write $\vDash_\lambda e_1 = e_2$ iff $A \vDash_\lambda e_1 = e_2$ for all concrete ccc models A. In the latter, we say $e_1 = e_2$ is valid in λ-calculus.

We review two important results about concrete ccc models in the model-based conservativeness proof, whose main proof steps are shown in Fig. 4. The first result is that concrete ccc models are a special instance of matching logic models. In other words, Γ^λ includes all concrete ccc models as its validating models. This result will be used in Step 3, from matching logic validity to λ-calculus validity. The second result is that concrete ccc models are *complete* with respect to λ-calculus reasoning, i.e., all valid λ-calculus equations can be proved.[4] This known completeness result is restated in Lemma 26. It will be used in Step 5 in Fig. 4, from λ-calculus validity to provability.

Lemma 26 ([56])**.** $\vDash_\lambda e_1 = e_2$ *implies* $\vdash_\lambda e_1 = e_2$ *for any* $e_1, e_2 \in \Lambda$.

Other λ-Calculus Models

We discuss the other relevant notions of λ-calculus models and discuss why we choose the concrete ccc models in our conservativeness proof (given in Section 7).

There are three main notions of models in λ-calculus; see [61] for a survey. Firstly, there are λ-models [6, Section 5.2], which are combinatory algebras that provide coherent interpretations to all λ-expressions. Secondly, there are categorical models [6, Section 5.5], which are given as thereflexive objects of a Cartesian closed category (ccc), where λ-expressions are interpreted as morphisms. Thirdly, there are Hindley-Longo models [52], which form an alternative presentation of λ-models and interpret λ-expressions directly, without translating them to combinatory terms. The concrete ccc models (Definition 25) in this paper belong to the categorical models, where the underlying categories are strictly concrete categories (see, e.g., [6, Definition 5.5.8]).

We choose concrete ccc models because they have a non-categorical set-theoretical presentation (Definition 25) that fits well with the pattern matching semantics of matching logic. In concrete ccc models, the interpretation of a λ-expression is inductively defined from the interpretation of its sub-expressions, so it is more natural to turn concrete ccc models into matching logic models, needed for the conservativeness proof. In contrast, λ-models and Hindley-Longo models interpret *all* λ-expressions *at the same time*. For example, in Hindley-Longo models, $|\lambda x.\, e|^\lambda_\rho$ is defined as some *unspecified element* that satisfies that $|\lambda x.\, e|^\lambda_\rho \bullet_A a = |e|^\lambda_{\rho[a/x]}$ for all a. In concrete ccc models, instead, $|\lambda x.\, e|^\lambda_\rho$ is interpreted explicitly by $|\lambda x.\, e|^\lambda_\rho = \mathbb{L}(f^\rho_{e,x})$, using a given (by the model) retraction function to encode functions into elements. Therefore, it is more convenient in our context to consider concrete ccc models, as they provide an explicit, constructive interpretation of $\lambda x.\, e$.

[4] Here we use the term "completeness" to mean deductive completeness, as given in Lemma 26. In the literature on λ-calculus, *representability completeness* (of λ-calculus models) is also considered; see related discussion in Section 8.2.2.

6 Defining λ-Calculus in Matching Logic

In this section we define the matching logic theory Γ^λ that captures λ-calculus. Our definition is inspired by the concrete ccc models of λ-calculus discussed in Section 5.3. The key ingredient is the retraction function $\mathbb{L}$ that encodes representable functions into elements. Therefore, we first define representable functions and the retraction function.

Recall that $f^\rho_{e,x}$ is the representable function as defined in Definition 25, which corresponds to the interpretation of $\lambda x.\, e$ under ρ in the concrete ccc model. We can capture $f^\rho_{e,x}$ by defining its *graph*:

$$\mathsf{graph}(f^\rho_{e,x}) = \left\{ \left(a, |e|^\lambda_{\rho[a/x]} \right) \mid \text{for all elements } a \text{ in the concrete ccc model } A \right\} \tag{6}$$

which contains all the argument-value pairs of $f^\rho_{e,x}$. Note that this graph is an element in $\mathcal{P}(A \times A)$, the powerset of $A \times A$, but not every element in $\mathcal{P}(A \times A)$ is the graph of a representable function. Therefore, the retraction function $\mathbb{L}$ is captured as a partial function from $\mathcal{P}(A \times A)$ to A (see Remark 27) which is defined only on the graphs of representable functions, and undefined elsewhere.

Now we start to define Γ^λ following the above intuition. Firstly, we include all λ-calculus variables in V^λ as element (and not set) variables in Γ^λ. Then, we define four sorts: *Var* as the sort of λ-calculus variables; *Exp* as the sort of λ-expressions; $\mathit{Var} \times \mathit{Exp}$ as the product sort of *Var* and *Exp* (Definition 18); and $2^{\mathit{Var} \times \mathit{Exp}}$ as its power sort (Definition 20). Intuitively, $2^{\mathit{Var} \times \mathit{Exp}}$ is the sort of all binary relations, including non-functions, over *Var* and *Exp*, because the inhabitant of $2^{\mathit{Var} \times \mathit{Exp}}$ is the powerset of the Cartesian product of the inhabitants of *Var* and *Exp*, by Propositions 19 and 21.

Next, we define the subsorting axiom (Section 4.2), $[\![\mathit{Var}]\!] \subseteq [\![\mathit{Exp}]\!]$, to specify that all variables are well-formed λ-expressions. We define a partial function (Section 4.2.3), $\mathsf{lambda} \colon 2^{\mathit{Exp}^2} \rightharpoonup \mathit{Exp}$, to represent the retraction function $\mathbb{L}$ in Definition 25, although the partial function requirement is included only for clarity and is technically unnecessary, because it will be automatically validated by the intended canonical models that we construct in Sections 7 and 8.

Remark 27. We include both sorts *Var* and *Exp* in theory Γ^λ so as to be completely faithful w.r.t. the λ-calculus syntax defined in Section 5, which has two syntactic categories: V^λ for variables and Λ for expressions. As a result, lambda is a partial function with the power domain 2^{Exp^2}. A valid alternative is to use $2^{\mathit{Exp} \otimes \mathit{Exp}}$ as the domain. The conservative extension theorem (Theorem 36) still holds, and its model-based proofs shown in Section 7 are still valid, because the models we will construct there interpret both *Var* and *Exp* to the same inhabitant set.

Now, we define λ-expressions as syntactic sugar in matching logic. The λ-calculus variables and application are already well-formed matching logic patterns, where $x \in \mathit{Var}$ is represented by the element variables x and $e_1\, e_2$ is represented by the built-in matching logic application $e_1\, e_2$. Abstraction $\lambda x.\, e$ is defined as the following syntactic sugar, where we extract the general *binding notation* $[x\colon \mathit{Var}]\ e$ for clarity and because it can be used to define any other binders, not only λ:

$$[x\colon \mathit{Var}]\ e \equiv \mathsf{intension}\, \exists x\colon \mathit{Var}.\, \langle x, e \rangle \qquad \text{// the binding notation} \tag{7}$$

$$\lambda x.\, e \equiv \mathsf{lambda}\, [x\colon \mathit{Var}]\ e \qquad \text{// } \lambda\text{-abstraction} \tag{8}$$

We assume that $[x\colon \mathit{Var}]\ e$ binds the tightest, so $\mathsf{lambda}\, [x\colon \mathit{Var}]\ e$ is parsed as $\mathsf{lambda}\ ([x\colon \mathit{Var}]\ e)$.

Eq. (8) is a logical incarnation of the semantics of $\lambda x.\, e$ in the concrete ccc models (Definition 25), into matching logic. Recall that in a concrete ccc model, $|\lambda x.\, e|^\lambda_\rho = \mathbb{L}\left(f^\rho_{e,x}\right)$, where $f^\rho_{e,x}(a) = |e|^\lambda_{\rho[a/x]}$. By Remark 10, $\exists x\colon \mathit{Var}.\, \langle x, e \rangle$ denotes the union set $\bigcup_x \{(x, e)\}$, namely the graph of $f^\rho_{e,x}$. (Note that $\forall x\colon \mathit{Var}.\, \langle x, e \rangle$ also yields the correct binding behavior, but it does not have the right semantic meaning of a graph.) The binding notation $[x\colon \mathit{Var}]\ e$ takes this graph as a *set* of pairs and *packs* them into one object in the power sort $2^{\mathit{Var} \times \mathit{Exp}}$. Then, this packed object is passed to lambda, which decodes/retracts it into the intended interpretation of $\lambda x.\, e$. For now, we do not know any property about lambda, except that it is a partial function from $2^{\mathit{Var} \times \mathit{Exp}}$ to *Exp*. Its intended behavior will be axiomatized by the axiom schema (β)—*the* axiom schema that characterizes λ-abstraction and the semantics of λ.

Variables:	
$x, y, \ldots$	element variables, including all λ-calculus variables in V^λ
Symbols:	
Var	a sort constant
Exp	a sort constant
lambda	the retraction symbol, used to capture λ
Axioms:	
(Subsorting)	$[\![\mathit{Var}]\!] \subseteq [\![\mathit{Exp}]\!]$
(β)	$\forall x_1 \colon \mathit{Var}. \cdots \forall x_n \colon \mathit{Var}. \, (\lambda x. \, e) \, e' = e[e'/x]$ where $x_1, \ldots, x_n$ are all the free variables in $\mathrm{FV}((\lambda x. \, e) \, e')$.

Figure 5: Summary of the matching logic theory Γ^λ that captures λ-calculus (infrastructure definitions omitted)

We emphasize that the encoding of $\lambda x. \, e$ in Eqs. (7)-(8) is only possible because matching logic treats terms and formulas uniformly as patterns, and it allows (FOL-style) quantification to be built on terms. A similar definition will immediately fail in FOL, because FOL enforces a clear distinction between terms and formulas at the syntax level and quantification only applies to formulas.

Remark 28. Under the above notations, all λ-expressions are well-formed matching logic patterns. Particularly, the syntactic sugar $\lambda x. \, e$ in Eqs. (7)-(8) satisfies all binding properties about λ in Fig. 3.

Definition 29. Let Γ^λ be the matching logic theory that contains all the axioms and notations that we have defined in this section, and all instances of the (β) axiom schema, as shown in Fig. 5.

Remark 30. Remark 28 holds, not because of the axioms in Γ^λ, but because of the syntactic sugar definition in Eqs. (7)-(8) and the binding behavior of $\exists$. In other words, the binding behavior of λ is directly inherited from from the binding behavior of the built-in binder $\exists$ in matching logic, and is *not* specified by axioms. The axioms specify the semantic behavior of λ, not its binding behavior.

We finish this section by proving the extensiveness theorem for λ-calculus.

Theorem 31. $\vdash_\lambda e_1 = e_2$ *implies* $\Gamma^\lambda \vdash e_1 = e_2$*, for all* $e_1, e_2 \in \Lambda$.

Proof. By Proposition 23, because Γ^λ contains all instances of (β). □

7 Model-Based Conservativeness Proof

Here we prove the conservativeness of Γ^λ, making use of the concrete ccc models of λ-calculus discussed in Section 5.3. The main proof steps have been discussed in Section 5 and summarized in Fig. 4. The only nontrivial one is Step 3, which requires to show that $M \vDash e_1 = e_2$ for all matching logic models $M \vDash \Gamma^\lambda$ implies $A \vDash_\lambda e_1 = e_2$ for all concrete ccc models A. The following is the key lemma establishing the connection between concrete ccc models and matching logic models of Γ^λ:

Lemma 32. *For any concrete ccc model* A *and any valuation* ρ *into* A*, there exists a matching logic model* $M^A \vDash \Gamma^\lambda$ *and a valuation* ρ^A *into* M^A *such that* $|e|_{\rho^A} = \left\{ |e|^\lambda_\rho \right\}$ *for every* $e \in \Lambda$.

Proof. We give the high-level proof idea. The complete proof can be found in Appendix C. Let us fix a concrete ccc model $(A, \_ \bullet_A \_, \mathbb{L})$, where $R(A)$ is its set of representable functions and $\mathbb{L} \colon R(A) \to A$ is its retraction function. Let the carrier set M_A include A. Recall that Γ^λ defines sorts *Var* and *Exp*, and partial function lambda from $2^{\mathit{Var} \times \mathit{Exp}}$ to *Exp* (Fig. 5). Since A is the domain of both variable valuations and expression interpretations in the concrete ccc model, in M_A we let A be the inhabitants of both *Var* and

Exp (see Remark 27), validating axiom (SUBSORTING). We define lambda_{M^A} accordingly to the retraction function $\mathbb{L}$; i.e., $\mathsf{lambda}_{M^A} \cdot P = \{\mathbb{L}(f)\}$ whenever $P = \mathsf{graph}(f)$ and $f \in R(A)$, and $\mathsf{lambda}_{M^A} \cdot P = \emptyset$, otherwise.

We define ρ^A as $\rho^A(x) = \rho(x)$, for every $x \in V^\lambda$, and prove that $|e|_{\rho^A} = \{|e|^\lambda_\rho\}$ for every $e \in \Lambda$. The proof is based on structural induction on e and the only nontrivial case is when e is $\lambda x.\, e_1$. In this case, we have $|\lambda x.\, e_1|_{\rho^A} = |\mathsf{lambda}\,(\mathsf{intension}\,(\exists x\colon Var.\,\langle x, e_1\rangle))|_{\rho^A} = \mathsf{lambda}_{M^A} \cdot |\mathsf{intension}\,(\exists x\colon Var.\,\langle x, e_1\rangle)|_{\rho^A} = \mathsf{lambda}_{M^A} \cdot |\exists x\colon Var.\,\langle x, e_1\rangle)|_{\rho^A} = \mathsf{lambda}_{M^A} \cdot \bigcup_{a\in A}\{(a, |e_1|_{\rho^A[a/x]})\} = \mathsf{lambda}_{M^A} \cdot \bigcup_{a\in A}\{(a, |e_1|^\lambda_{\rho[a/x]})\} = \mathsf{lambda}_{M^A} \cdot \mathsf{graph}(f^\rho_{e_1,x}) = \{\mathbb{L}(f^\rho_{e_1,x})\} = \{|\lambda x.\, e_1|^\lambda_\rho\}$.

Finally, we show that M^A validates (β). Using the above result, for any $x \in V^\lambda$, $e, e' \in \Lambda$, and ρ, we have that $|(\lambda x.\, e)e'|^\lambda_\rho = |e[e'/x]|^\lambda_\rho$ in A implies $|(\lambda x.\, e)e'|_{\rho^A} = |e[e'/x]|_{\rho^A}$ in M^A. Noting that ρ^A is arbitrary (as ρ is arbitrary), M^A validates (β). □

Remark 33. The operations, $\mathsf{intension}$ and lambda, have been crucial in the proof. Without them, the pattern $\exists x\colon Var.\,\langle x, e\rangle$ itself is merely the graph set and is not even a functional pattern (in the sense discussed in Section 4.2.2), and thus cannot be directly used to interpret $\lambda x.\, e$.

Using Lemma 32, we can immediately prove Step 3 in Fig. 4:

Lemma 34. *If $M \vDash e_1 = e_2$ for all models $M \vDash \Gamma^\lambda$, then $A \vDash_\lambda e_1 = e_2$ for all concrete ccc models A.*

Proof. Let A be any concrete ccc and ρ be any valuation. By Lemma 32, there exists a matching logic model $M^A \vDash \Gamma^\lambda$ and a valuation ρ^A such that $|e|_{\rho^A} = \{|e|^\lambda_\rho\}$ for any $e \in \Lambda$. Since $M^A \vDash e_1 = e_2$, we have $|e_1|_{\rho^A} = |e_2|_{\rho^A}$, and thus $|e_1|^\lambda_\rho = |e_2|^\lambda_\rho$. Since ρ is any valuation, we have $A \vDash_\lambda e_1 = e_2$. □

Theorem 35. $\Gamma^\lambda \vdash e_1 = e_2$ *implies* $\vdash_\lambda e_1 = e_2$, *for all* $e_1, e_2 \in \Lambda$.

Proof. See Fig. 4, where Step 1 is by Theorem 24; Step 2 is by Definition 11; Step 3 is by Lemma 34; Step 4 is by Definition 25; and Step 5 is by Lemma 26. □

Theorem 35 together with Theorem 31 show that Γ^λ is a conservative extension of λ-calculus. In fact, we prove the following equivalence theorem (for $e_1, e_2 \in \Lambda$):

Theorem 36. *These are equivalent: (1) $\Gamma^\lambda \vdash e_1 = e_2$; (2) $\Gamma^\lambda \vDash e_1 = e_2$; (3) $\vDash_\lambda e_1 = e_2$; (4) $\vdash_\lambda e_1 = e_2$.*

Proof. (1) $\Longrightarrow$ (2) is by Theorem 24. (2) $\Longrightarrow$ (3) is by Lemma 34. (3) $\Longrightarrow$ (4) is by Lemma 26. (4) $\Longrightarrow$ (1) is by Theorem 35. Note: Conservative extension theorem is the equivalence (1) $\Longleftrightarrow$ (4). □

Remark 37. The equivalence (2) $\Longleftrightarrow$ (4) shows the *(deductive) completeness* of the matching logic models of Γ^λ with respect to λ-calculus. By defining λ-calculus in matching logic, we automatically obtain, from the model theory of matching logic, models that are complete to λ-calculus.

8 Syntax-Based Conservativeness Proof

In this section we show an alternative conservativeness proof of Theorem 35 that is entirely based on the syntactic structure of λ-expressions, and thus is easier to generalize to other logical systems and binders, especially those which do not have well-established models. This syntax-based proof also shows that Γ^λ is *representationally complete* for λ-calculus; see Section 8.2.2.

8.1 Proof Overview: Using the Term Model to Prove the Conservativeness Theorem

We build a special matching logic model $T \vDash \Gamma^\lambda$, which we call the *term model* of λ-calculus,[5] and follow the term algebra technique [49, 78, 8]: T has as elements the equivalence classes of λ-expressions modulo $\alpha\beta$-equivalence, and each $e \in \Lambda$ is interpreted in T as the equivalence class containing itself, $[e]$. Formally, we will prove this:

Theorem 38. *Let* $[e] = \{e' \in \Lambda \mid \vdash_\lambda e = e'\}$ *be the equivalence class of* e *modulo* $\alpha\beta$*-equivalence. Let* $[\Lambda] = \{[e] \mid e \in \Lambda\}$ *be the set of all these classes. Then, there is a matching logic model* $T \vDash \Gamma^\lambda$*, called* term model*, and a valuation* ρ_T*, called* term valuation*, such that* $|e|_{\rho_T} = \{[e]\}$ *for all* $e \in \Lambda$.

Remark 39. For distinct variables $x, y \in V^\lambda$, we have $[x] \neq [y]$ [6, Fact 2.1.37]. Clearly, $x \in [x]$, but $[x]$ also includes infinitely many expressions: $(\lambda y.\, y)x$, $(\lambda y.\, y)((\lambda y.\, y)x)$, etc.

We will construct T in Section 8.2. For now, we show how to prove Theorem 35 using Theorem 38:

Syntax-Based Proof of Theorem 35. We need to prove $\Gamma^\lambda \vdash e_1 = e_2$ implies $\vdash_\lambda e_1 = e_2$:

$\Gamma^\lambda \vdash e_1 = e_2$	implies	$\Gamma^\lambda \vDash e_1 = e_2$	by Theorem 24
	implies	$T \vDash e_1 = e_2$	by Definition 11
	implies	$\lvert e_1 \rvert_{\rho_T} = \lvert e_2 \rvert_{\rho_T}$	by Proposition 14
	implies	$[e_1] = [e_2]$	by Theorem 38
	implies	$\vdash_\lambda e_1 = e_2$	by Definition of $[e]$ in Theorem 38. $\square$

8.2 Construction of the Term Model T and the Term Valuation ρ_T

In this section we construct T and show that $T \vDash \Gamma^\lambda$. Like for the matching logic model of Γ^λ in the proof of Lemma 32, we need to give interpretations to the sorts *Var* and *Exp*, as well as to the retraction function lambda. For *Var* and *Exp*, we define their inhabitants as $T_{Var} = [V^\lambda]$ and $T_{Exp} = [\Lambda]$, where $[V^\lambda]$ and $[\Lambda]$ are the set of equivalence classes of variables and λ-expressions. Clearly, we have $[V^\lambda] \subseteq [\Lambda]$, which validates the axiom (SUBSORTING) $[\![Var]\!] \subseteq [\![Exp]\!]$. We define the interpretation of application on λ-expressions as the application in λ-calculus, i.e., $[e_1] \cdot [e_2] = [e_1\, e_2]$ for any $e_1, e_2 \in \Lambda$. Note that this definition is well-defined, because $\vdash_\lambda e_1\, e_2 = e_1'\, e_2'$ whenever $\vdash_\lambda e_1 = e_1'$ and $\vdash_\lambda e_2 = e_2'$. Finally, we define the interpretation lambda_T such that

$$\mathsf{lambda}_T \cdot \left(\bigcup_{z \in V^\lambda} \big([z], [e[z/x]]\big) \right) = \big\{[\lambda x.\, e]\big\}, \quad \text{for any } x \in V^\lambda \text{ and } e \in \Lambda. \tag{9}$$

and $\mathsf{lambda}_T \cdot P = \emptyset$, if P is not a graph of the above form. The complete construction of T can be found in Appendix D.

The construction of T, especially Eq. (9), is critically depending on the matching logic encoding $\lambda x.\, e \equiv \mathsf{lambda}\,(\mathsf{intension}\, \exists x\colon Var.\, \langle x, e\rangle)$. The α-equivalence $\lambda x.\, e \equiv \lambda z.\, (e[z/x])$ is captured, both syntactically and semantically, by collecting the pairs $\langle z, e[z/x]\rangle$ for all z, using the matching logic pattern $\exists x\colon Var.\, \langle x, e\rangle$ (see Remark 10 for the connection between the $\exists$-patterns and the set-theoretic unions). Therefore, $\exists x\colon Var.\, \langle x, e\rangle$ encapsulates all the information about $[\lambda x.\, e]$, which is *packed* by $\mathsf{intension}$ and passed to lambda, and then *retracted* to restore the original expression $\lambda x.\, e$. The following proposition shows that the condition in Eq. (9) on lambda_T is not inconsistent:

[5] In the literature on λ-calculus, term models have a different meaning. For example, in [6], term models are special λ-calculus models constructed based on the combinatory algebra semantics; see Section 8.2.1 for a comparison.

Proposition 40. $[\lambda x.\, e] = [\lambda x'.\, e']$, whenever

$$\bigcup_{z \in V^\lambda} ([z], [e[z/x]]) = \bigcup_{z \in V^\lambda} ([z], [e'[z/x']]) \tag{10}$$

Proof. Assume the opposite, i.e., $[\lambda x.\, e] \neq [\lambda x'.\, e']$. Let $z^* \in V^\lambda$ be a fresh variable that does not occur in $\lambda x.\, e$ or $\lambda x'.\, e'$. Then we have $\lambda x.\, e \equiv \lambda z^*.\, e[z^*/x]$ and $\lambda x'.\, e' \equiv \lambda z^*.\, e'[z^*/x']$. By the assumption, we have $[\lambda z^*.\, e[z^*/x]] \neq [\lambda z^*.\, e'[z^*/x']]$, and thus $[e[z^*/x]] \neq [e'[z^*/x']]$. Noting that $[z_1] = [z_2]$ iff $z_1 = z_2$, for every $z_1, z_2 \in V^\lambda$ (Remark 39), we have that the pair $([z^*], [e[z^*/x]])$ is in the LHS of Eq. (10) but not its RHS, which is a contradiction. □

So far, we have constructed the term model T. We now define the term valuation ρ_T. Let

$$\mathit{VarVal} = \{\rho \mid \rho(x) \in [V^\lambda] \text{ for all } x \in V^\lambda\}$$

be the set of valuations that map λ-calculus variables (which have been taken as matching logic element variables; see Section 6) to the equivalence classes of λ-calculus *variables*, and not any λ-expressions. We define the term valuation ρ_T, as $\rho_T(x) = [x]$ for every $x \in V^\lambda$. Clearly, $\rho_T \in \mathit{VarVal}$.

Proposition 41. $|e|_{\rho_T} = \{[e]\}$, and $|e|_{\rho[\rho(z)/x]} = |e[z/x]|_\rho$ for all $\rho \in \mathit{VarVal}$.

Proof. We prove both properties simultaneously by induction on the λ-*depth* $d(e)$ of e, the maximum number of nested λ binders in e. If $d(e) = 0$ then e is a variable or is built from only application and has no λ abstraction. In this case, both properties can be proved by another structural induction on e. If $d(e) \geq 1$ then e has either the form $e_1\, e_2$ where $d(e_1), d(e_2) \leq d(e)$, or the form $\lambda x.\, e_1$ where $d(e_1) \leq d(e) - 1$. Then another structural induction on e proves both properties. □

Proposition 42. If $\vdash_\lambda e = e'$, then $|e|_\rho = |e'|_\rho$ for any $\rho \in \mathit{VarVal}$.

Proof. Note that the interpretation of a λ-expression relies on its free variables. Suppose $\mathrm{FV}(e) \cup \mathrm{FV}(e') = \{x_1, \ldots, x_n\}$ and $\rho(x_i) = [y_i]$ for $i \in \{1, \ldots, n\}$. By Remark 39, y_i is the unique variable that is in $[y_i]$. Since ρ equals to $\rho_T[[y_1]/x_1] \cdots [[y_n]/x_n]$ restricted on $x_1, \ldots, x_n$, we have $|e|_\rho = |e|_{\rho_T[[y_1]/x_1]\cdots[[y_n]/x_n]}$. By Proposition 41, $|e|_{\rho_T[[y_1]/x_1]\cdots[[y_n]/x_n]} = |e[y_1/x_1] \cdots [y_n/x_n]|_{\rho_T} = \{[e[y_1/x_1] \cdots [y_n/x_n]]\}$; similarly $|e'|_\rho = \{[e'[y_1/x_1] \cdots [y_n/x_n]]\}$. Then, $\vdash_\lambda e[y_1/x_1] \cdots [y_n/x_n] = e'[y_1/x_1] \cdots [y_n/x_n]$, i.e., $[e[y_1/x_1] \cdots [y_n/x_n]] = [e'[y_1/x_1] \cdots [y_n/x_n]]$. Hence, $|e|_\rho = |e'|_\rho$. □

The only thing left is to prove Theorem 38. We have shown that $|e|_{\rho_T} = \{[e]\}$ for every $e \in \Lambda$, in Proposition 41. It remains to show that T validates (β), i.e., $|(\lambda x.\, e)\, e'|_\rho = |e[e'/x]|_\rho$ for all $\rho \in \mathit{VarVal}$, which follows immediately from Proposition 42. Note that we only need to consider valuations in VarVal because all variables in (β) are quantified over the sort Var.

8.2.1 Comparing Our Term Model T to the Classical Notion of Term Models in λ-Calculus

In the literature on λ-calculus, a *term model* [6, Definition 5.2.11] is a λ-model (Example 6), where the underlying carrier set A is $[\Lambda]$, the application function is the application function over equivalence classes, and the two special constants are $k = [\lambda x.\, \lambda y.\, x]$ and $s = [\lambda x.\, \lambda y.\, \lambda z.\, (xz)(yz)]$; we will denote this λ-model as A and call it a *classical term model*, to not confuse it with our term model T. Clearly, T and A represent different approaches to capture λ-expressions. While A uses the name-free, combinators approach, where λ is handled by *abstraction elimination*, our term model T gives an explicit and constructive interpretation to λ, as shown in Eq. (9).

8.2.2 The Representabiltiy Problem

There has been a long-standing, concerning and open problem in the study of λ-calculus, called the *representability problem* [10, pp. 8], which asks if a given class of λ-calculus models is *representationally complete*, in the sense that there exists a model in the given class such that any two expressions e_1 and e_2 are provably equal if and only if they are interpreted as the same element/value in that model. Representability completeness indicates that a class of λ-calculus models is sufficient in capturing the formal reasoning in λ-calculus, so one may *reduce* the study of formal reasoning in λ-calculus to the study of models, where more mathematical tools and techniques can be applied. Hence, *reduction* is the main motivation.

λ-calculus models are broadly divided into *syntactic models* and *non-syntactic models* [61, pp. 13], depending on whether their construction is based on the syntax and provability of λ-calculus or not. All the classical term models in λ-calculus, as well as our particular matching logic term model in Section 8.2, are syntactic models. Syntactic models are often representationally complete, but studying them tends to be as hard as studying the syntax and formal reasoning directly, and thus the reduction to syntactic models usually does not help simplify the study of λ-calculus. Thus, for decades researchers have been searching for and studying sub-classes of non-syntactic concrete ccc models, hoping they are also representationally complete. So far, three main such sub-classes have been identified, known as the *main semantics* of λ-calculus: Scott's continuous semantics [85], Berry's stable semantics [12, 44], and Bucciarelli-Ehrhard strongly stable semantics [16]. The representability problem for the main semantics (and their sub-classes) has remained largely open as of today, except for some negative results proved for some sub-classes (e.g., graph models [17]).

Theorem 38 shows that the class of matching logic models of Γ^λ is representationally complete, positively answering the representability problem for our matching logic semantics of λ-calculus. Our proof does not rely on any known results about the representational completeness of any existing semantics; instead, it is entirely based on the model theory of matching logic, which is not specific to λ-calculus but which allows for an appropriate axiomatization of λ-calculus as a theory that is hereby endowed with the desired representationally complete models automatically. We can push Theorem 38 even further to any equational extensions of λ-calculus, known as *λ-theories*. Indeed, the definition of the equivalence class $[e]$ as the set of $\alpha\beta$-equivalent expressions of e, has not been critical in the proof of Theorem 38, and the conclusion still holds if we consider any equivalence class $[e]$ that includes the basic $\alpha\beta$-equivalence. Therefore, we conclude that the matching logic definition of λ-calculus is representationally complete for *all λ-theories*.

Although we do not solve any of the existing open problems, our work suggests the matching logic can be a viable alternative to the existing λ-calculus models within the main semantics. The matching logic models are as good as the existing models for λ-calculus in terms of theoretical properties w.r.t. formal reasoning and semantics, yet unlike the existing models, they are general in the sense that they are not crafted specifically for λ-calculus, but are obtained from the matching logic theory Γ^λ. We give a general solution for all the binders, which for λ-calculus is as good as the state of the art, considering *both* the proof-theoretic *and* the model-theoretic aspects.

9 Defining Binders in Other Logical Systems Using Matching Logic

We showed how to capture the binder λ in matching logic as the following notation (Eqs. (7)-(8)):

$$\lambda x.\, e \equiv \mathsf{lambda}\,[x\colon \mathit{Var}]\,e \tag{11}$$

We defined a matching logic theory, Γ^λ (shown in Fig 5), and proved the conservative extension theorem for λ-calculus, Eq. (5). In this section we show that our approach is not specific to λ-calculus. We provide evidence that matching logic can serve as a general approach to dealing with binders. We will show how to use patterns similar to Eq. (11) to define the binders in a variety of logical systems, including System F [43, 79], pure type systems [7], π-calculus [66], and more, and prove a corresponding conservative extension theorem for each of them. To do that, several challenges need to be solved.

A first challenge is that binders can have more complex binding behavior than in λ-calculus; see Fig. 6. For example, $\lambda x{:}e_1.\, e_2$ in System F binds x within e_2, but not in e_1; $\mathsf{Inp}(x, y, e)$ in π-calculus has the binding

Constructs	Binding Behavior	Meaning	Origins
$\lambda x.\, e$	binding x into e	function abstraction	λ-calculus
$\lambda x{:}e_1.\, e_2$	binding x into e_2	function abstraction	System F
$\lambda t.\, e$	binding t into e	type abstraction	System F
$\Pi t.\, e$	binding t into e	Π-type constructor	System F
$\lambda x{:}e_1.\, e_2$	binding x into e_2	function abstraction	Pure type system
$\pi x{:}e_1.\, e_2$	binding x into e_2	type abstraction	Pure type system
$\mathsf{Inp}(x, y, e)$	binding y into e	input process	π-calculus
$\nu y.\, e$	binding y into e	new process name creation	π-calculus
$\mathsf{Bout}(e_1, x, y, e_2)$	binding y into e_2	bound output transition	π-calculus
$\mathsf{Inp}(e_1, x, y, e_2)$	binding y into e_2	input transition	π-calculus

Figure 6: Some example binding constructs and their binding behavior in logical systems.

variable in the second position (i.e., y), and not the first position. We deal with this binding behavior by desugaring to binders whose binding variable is their first argument and is bound within the second argument only; that is, we desugar an arbitrary binder to a binder of the form $b(x, e_1, \ldots, e_n)$, where x is bound in e_1 but not in $e_2, \ldots, e_n$. Clearly, this desugaring process is just a sequence of argument swappings. Then, we further desugar $b(x, e_1, \ldots, e_n)$ to $b'(b''(x, e_1), e_2, \ldots, e_n)$, where b' is a (binding-free) symbol and b'' is a binder that binds x to e_1, just like λ in λ-calculus. Finally, we define $b''(x, e_1)$ as the following syntactic sugar:

$$b''(x, e) \equiv \mathsf{retraction}_b\, [x{:}\, \mathit{Var}]\, e \tag{12}$$

in the same way as in Eq. (11), except that here we use a new retraction symbol $\mathsf{retraction}_b$ that is specific to the binder b. Each binder has its own retraction symbol, but the other infrastructure symbols, such as products, powersets, and the binding notation $[x{:}\, \mathit{Var}]\, e$, are the same. From now on, we will only consider binders $b(x, e)$ that bind x within e, for technical convenience.

A second challenge is that logical systems featuring bindings are very different from each other, in terms of the kinds of *logical reasoning* that is carried out in them. For example, System F derives *typing judgments* $\Gamma \rhd e_1{:}e_2$ to mean that e_1 has type e_2 under typing environment Γ; π-calculus derives *transitions* $e_1 \xrightarrow{act} e_2$ to mean that process e_1 transits by action *act* to process e_2. It is tedious and non-systematic to consider these logical systems *separately*, because we would need to capture their specific logical reasoning and prove the conservative extension theorem for each of them, more or less similarly to the syntax-based proof in Section 8.

Remark 43. The current $\mathbb{K}$ framework implementation provides a "binder" attribute, which allows one to define a language construct that binds all variables occurring in its first argument within its other arguments. The results demonstrated in this paper, particularly this section, will be used to improve $\mathbb{K}$ and let it support binders with more complex binding behaviors. The reader who is interested in seeing examples about the current $\mathbb{K}$ support for binders may look at [53], where the "binder" attribute is used to define the syntax of λ-calculus.

To capture the various logical systems featuring bindings more systematically, we employ a parametric framework for binders, called *term-generic logic* [77] (TGL). TGL is a parametric variant of FOL, whose syntax is parametric on a set of (generic) terms that are not constructed from constants and functions, but defined axiomatically. When we instantiate TGL with the term syntax of a given system (e.g., λ-calculus, System F, π-calculus, etc), it becomes a (first-order) *meta-logic* of that system and can be used to specify and reason about its meta-properties. Using TGL, we give a systematic treatment of binders in the various logical systems. We will capture TGL in matching logic and prove a conservative extension theorem for TGL, from which the conservative extension theorems for the other logical systems follow as corollaries.

Why not use TGL directly then, but instead use matching logic? There are two reasons. Firstly, TGL in its full generality is not implementable, because it does not deal with any concrete syntax of binders.

Its notion of (generic) terms is given axiomatically and needs to be instantiated, which is what we will do in Section 9.1, where we instantiate TGL to bridge matching logic and other logical systems with binders. The second reason is that TGL is a logic specifically designed for binders, while matching logic serves as the unifying logical foundation for the $\mathbb{K}$ framework, as discussed in Section 1 and other places in the paper. Therefore, matching logic supports reasoning in many mathematical domains other than binders, and thus it is more practical than TGL.

We next first introduce TGL in Section 9.1 and then its matching logic definition in Section 9.2.

9.1 Term-Generic Logic (TGL) Preliminaries

TGL [77] is a variant of many-sorted FOL whose syntax is parametric in a (generic) term set that is defined axiomatically. In TGL, any set T exporting two operations—free variables $\mathrm{FV}(e)$ and capture-free substitution $e[e'/x]$—and satisfying the conditions in [77, Definition 2.1], forms a generic term set. TGL formulas are built like in FOL, from predicates $\pi(e_1, \ldots, e_n)$, equations $e_1 = e_2$, and standard connectives $\wedge, \neg, \exists$, except that $e_1, \ldots, e_n$ are generic terms, that is, arbitrary elements in T. The metatheory of TGL, including its semantics and models, terms/formulas interpretation, proof system, and, importantly, a soundness and completeness theorem, have been studied and presented in detail in [77].

For concreteness, we will not introduce TGL in its full generality. Instead, we instantiate TGL with a concrete, constructive term syntax with binders (defined below) and introduce the metatheory of that TGL instance. From the discussion at the beginning of Section 9, this term syntax is sufficient to capture the binders in various logical systems with more complex bindings (Fig. 6).

Definition 44. A *binder syntax* is a tuple (S, V, F, B), where

1. S is a set of *sorts* denoted s, r, possibly with subscripts; we use $\bar{s} \in S^*$ to mean a list of sorts;
2. $V = \{V_s\}_{s \in S}$ is a sort-wise disjoint family of *variables* denoted $x{:}s$, $y{:}s$, etc;
3. $F = \{F_{\bar{s},r}\}_{\bar{s} \in S^*, s \in S}$ is a family of *many-sorted operations* of argument sorts $\bar{s}$ and result sort r;
4. $B = \{B_{s,s',r}\}_{s,s',r \in S}$ is a family of *binders*, where $b(x{:}s, e)$ binds $x{:}s$ to e (of sort s') and returns a term of sort r, for each $b \in B_{s,s',r}$.

We use *TGLTerm* to denote the set of terms generated by the above syntax, where free variables, α-equivalence, and capture-free substitution are defined in the usual way. We omit sorts when they can be inferred. Note that when $B = \emptyset$, rules (1)-(3) generate the standard FOL terms.

Remark 45. *TGLTerm* forms a TGL generic term set in [77, Definition 2.1].

TGL formulas, interpretations, validity, and provability are defined in the standard way, (almost) identical to FOL, except that terms are interpreted simultaneously instead of constructively. Specifically, the interpretation of compound term $f(e)$ is not defined from the interpretation of its sub-term e,[6] but instead we have a Henkin-style definition for term interpretations:

Definition 46 ([77, Section 2])**.** For a given set of *many-sorted predicates* $\Pi = \{\Pi_{\bar{s}}\}_{\bar{s} \in S^*}$, we define the set *TGLForm* of *TGL formulas* by the following grammar:

$$\varphi ::= e_1 = e_2 \mid \varphi_1 \wedge \varphi_2 \mid \neg\varphi \mid \exists x{:}s'.\, \varphi \mid \pi(e_1, \ldots, e_n) \text{ for } \pi \in \Pi_{s_1 \cdots s_n} \text{ and } e_i \text{ has sort } s_i \text{ for all } i$$

Let $A = \{A_s\}_{s \in S}$ be an S-indexed carrier set. A *TGL valuation* $\rho\colon V \to A$ is a function such that $\rho(x{:}s) \in A_s$ for every $s \in S$ and $x{:}s \in V_s$. Let *TGLVal* be the set of all TGL valuations. A *TGL model* $(\{A_s\}_{s \in S}, \{A_e\}_{e \in TGLTerm}, \{A_\pi\}_{\pi \in \Pi})$ has a Henkin-style definition as follows:

1. $A_s \neq \emptyset$ for every $s \in S$.

[6] TGL in its full generality as in [77] does not even have a notion of compound terms or sub-terms.

2. $A_e\colon \mathit{TGLVal} \to A_s$, where s is the sort of e, such that for any $x\colon s, e, e', \rho$:
   (a) $A_{x\colon s}(\rho) = \rho(x\colon s)$.
   (b) $A_{e[e'/x\colon s]}(\rho) = A_e(S_{e',x\colon s}(\rho))$, where $S_{e',x\colon s}(\rho)$ is the TGL valuation such that $S_{e',x\colon s}(\rho)(x\colon s) = A_{e'}(\rho)$ and $S_{e',x\colon s}(\rho)(y\colon s') = A_{y\colon s'}(\rho)$ for any $y\colon s' \not\equiv x\colon s$.
3. $A_\pi \subseteq A_{s_1} \times \cdots \times A_{s_n}$ for every $\pi \in \Pi_{s_1 \ldots s_n}$.

We let $A_\varphi \subseteq \mathit{TGLVal}$ for $\varphi \in \mathit{TGLForm}$ be the set of valuations under which φ holds, defined as:

1. $\rho \in A_{e_1 = e_2}$ iff $A_{e_1}(\rho) = A_{e_2}(\rho)$;
2. $\rho \in A_{\pi(e_1,\ldots,e_1)}$ iff $(A_{e_1}(\rho), \ldots, A_{e_n}(\rho)) \in A_\pi$;
3. $\rho \in A_{\varphi_1 \wedge \varphi_2}$ iff $\rho \in A_{\varphi_1}$ and $\rho \in A_{\varphi_2}$;
4. $\rho \in A_{\neg\varphi}$ iff $\rho \notin A_\varphi$;
5. $\rho \in A_{\forall x\colon s.\, \varphi}$ iff $\rho[a/x\colon s] \in A_\varphi$ for every $a \in A_s$.

TGL has a sound and complete Gentzen proof system [77, Figs. 1-2], which derives sequents of the form $E \vdash_{\mathsf{TGL}} \Delta_1 \triangleright \Delta_2$ for $E, \Delta_1, \Delta_2 \subseteq \mathit{TGLForm}$, which intuitively means that under TGL theory E, the *conjunction* of the formulas in Δ_1 implies the *disjunction* of the formulas in Δ_2. It is required that E contains formulas without free variables, and Δ_1, Δ_2 are finite sets containing formulas with finitely many free variables; these requirements are needed for TGL's completeness theorem and all TGL sequents considered in this paper satisfy these requirements.

Definition 47 ([77, Sections 2-3])**.** For a TGL model A and $\varphi \in \mathit{TGLForm}$, we write $A \vDash_{\mathsf{TGL}} \varphi$ iff $A_\varphi = \mathit{TGLVal}$. We write $A \vDash_{\mathsf{TGL}} E$ iff $A \vDash_{\mathsf{TGL}} \varphi$ for all $\varphi \in E$. *TGL validity* $E \vDash_{\mathsf{TGL}} \Delta_1 \triangleright \Delta_2$ is defined as $\bigcap_{\varphi\in\Delta_1} A_\varphi \subseteq \bigcup_{\varphi\in\Delta_2} A_\varphi$, for all $A \vDash_{\mathsf{TGL}} E$. *TGL provability* $E \vdash_{\mathsf{TGL}} \Delta_1 \triangleright \Delta_2$ is defined by the Gentzen proof system of TGL in the usual way.

Theorem 48 ([77, Theorem 3.1])**.** *Under the above requirements about E, Δ_1, Δ_2, we have $E \vDash_{\mathsf{TGL}} \Delta_1 \triangleright \Delta_2$ if and only if $E \vdash_{\mathsf{TGL}} \Delta_1 \triangleright \Delta_2$.*

9.2 Defining Term Generic Logic in Matching Logic

In this section we define a matching logic theory Γ^{TGL} and introduce notations such that all TGL terms and formulas are well-formed matching logic patterns. We show that Γ^{TGL} is a conservative extension of TGL, by proving the following equivalence theorem.

Theorem 49. *Under the notations in Theorem 48, the following are equivalent: (1) $(\Gamma^{\mathsf{TGL}} \cup E) \vdash \bigwedge \Delta_1 \to \bigvee \Delta_2$. (2) $(\Gamma^{\mathsf{TGL}} \cup E) \vDash \bigwedge \Delta_1 \to \bigvee \Delta_2$; (3) $E \vDash_{\mathsf{TGL}} \Delta_1 \triangleright \Delta_2$; (4) $E \vdash_{\mathsf{TGL}} \Delta_1 \triangleright \Delta_2$; Here, $\bigwedge \Delta_1$ is the conjunction of patterns in Δ_1 and $\bigvee \Delta_2$ is the disjunction of patterns in Δ_2.*

Thanks to the mathematical instruments and notations that we have introduced in Section 4, the definition of Γ^{TGL} is straightforward. The many-sorted binder syntax (Definition 44) and TGL terms are captured by defining sorts and many-sorted functions as in Section 4.2, and defining binders as in Eq. (12). TGL formulas, except $\pi(e_1, \ldots, e_n)$, are captured by matching logic's derived connectives (Fig. 1) and equality (Definition 13). Predicate $\pi(e_1, \ldots, e_n)$ for $\pi \in \Pi_{s_1 \cdots s_n}$, is captured by defining a matching logic symbol π and the following axiom:

$$(\textsc{Predicate}) \quad \forall x_1\colon s_1.\ \ldots \forall x_n\colon s_n.\, (\pi\, x_1 \cdots x_n = \top) \vee (\pi\, x_1 \cdots x_n = \bot) \tag{13}$$

which specifies that π returns either $\top$ or $\bot$, i.e., it indeed builds predicate patterns. Without such axioms, $\pi\, x_1 \cdots x_n$ could be any subset. Let Γ^{TGL} contain all the above definitions and notations.

Remark 50. Under the above notations and axioms, all TGL terms are matching logic functional patterns (Section 3.2.2) and all TGL formulas are matching logic predicate patterns (Section 3.2.1).

Theorem 49 is proved using a model-based approach similar to Fig. 4. The complete proof can be found in Appendix E. Here we explain the only nontrivial proof step, which is (2) $\Longrightarrow$ (3). This is proved by constructing a matching logic model M^A from any given TGL model A, such that all TGL terms and formulas are interpreted the same in M^A and A, i.e., $|e|_\rho = \{A_e(\rho)\}$ for every $e \in \mathit{TGLTerm}$; $|\varphi|_\rho = M^A$ whenever $\rho \in A_\varphi$, and $|\varphi|_\rho = \emptyset$, whenever $\rho \notin A_\varphi$, for every $\varphi \in \mathit{TGLForm}$.

Remark 51. Using TGL and Theorem 49, we obtain a systematic proof of the conservative extension theorems and deductive completeness theorems for all logical systems that have been defined in TGL and studied in [77, Section 4] and [76, Section 4], including System F [43, 79] (both the typing and reduction versions), λ-calculus (including the untyped [28], sub-typed [20], illative [6], and linear versions [45, 58]), pure type systems [7], and π-calculus [66]. The systematic proof works as follows. For each logical system L, its set of terms Term_L can be captured by a binder syntax using the desugaring discussed at the beginning of Section 9. The proof/type system of L that derives sequents of the form $\vdash_L \Phi$ is captured by a set of TGL axioms E^L, where each axiom corresponds to one type/proof rule of L [77]. An *adequacy theorem* is also proved there for each L, stating that $\vdash_L \Phi$ iff $E^L \vdash_{\mathsf{TGL}} \Phi^{\mathsf{TGL}}$, where Φ^{TGL} (of the form $\Delta_1^\Phi \triangleright \Delta_2^\Phi$) is the corresponding TGL encoding of the L-sequent Φ. Let $\Gamma^L = \Gamma^{\mathsf{TGL}} \cup E^L$ be the matching logic theory that captures L, and $\Phi^{\mathsf{ML}} = \bigwedge \Delta_1^\Phi \to \bigvee \Delta_2^\Phi$ be the matching logic encoding of Φ. By Theorem 48, we have that $\vdash_L \Phi$ in L, iff $E^L \vdash_{\mathsf{TGL}} \Phi^{\mathsf{TGL}}$ in TGL, iff $\Gamma^L \vdash \Phi^{\mathsf{ML}}$ in matching logic, iff $\Gamma^L \vDash \Phi^{\mathsf{ML}}$ in matching logic. Hence, Γ^L is a conservative extension of L and the class of matching logic models of Γ^L is complete with respect to L.

Remark 52. Note that the term "consistency" has different meanings in different contexts. In type systems, inconsistency means the ability to prove any typing judgments $t\colon\tau$. Similarly, in λ-calculus or other equational logic theories, inconsistency means the ability to prove any equations $e_1 = e_2$. However, in matching logic (and also FOL), inconsistency means the ability to prove logical false $\bot$. Thus, inconsistency for classical logics such as matching logic is stricter than that for type systems and λ-calculus. For example, if T is a PTS that contains the typing axiom $\mathit{Type}\colon \mathit{Type}$, then T is inconsistent [62], but its matching logic theory Γ^T is still a consistent matching logic theory and has a model that interprets the typing relation $\_\colon\_$ as the total relation on all PTS terms.

10 Future Work

Inductive Reasoning An important direction for future work is to investigate *inductive reasoning* on terms with binders. We use λ-calculus as an example but the discussion applies to all binders.

The set of λ-expressions Λ is an inductive structure. This means that Λ is the smallest set closed under variables, application, and abstraction, and it admits the *principle of inductive reasoning*, which can be intuitively expressed by the following formula (this should be understood informally; in particular, the inductive hypothesis for $\lambda x.\, e$ in (‡) takes various forms in the literature; e.g., [72, pp. 21] uses the M-quantifier on x, meaning that there exists $x\colon \mathit{Var}$ such that x is not free in e, while [5, pp. 5] uses $\forall$-quantifier to quantify all $x\colon \mathit{Var}$ that are not free in e):[7]

$$\begin{aligned}
&\forall P.\, (\forall x\colon \mathit{Var}.\, x \in P)\\
&\quad \wedge (\forall e\colon \mathit{Exp}.\, \forall e'\colon \mathit{Exp}.\, e \in P \wedge e' \in P \to (e\, e') \in P)\\
&\quad \wedge (\forall e\colon \mathit{Exp}.\, e \in P \to \forall x\colon \mathit{Var}.\, \lambda x.\, e \in P)\\
&\quad \to \forall e\colon \mathit{Exp}.\, e \in P
\end{aligned} \tag{‡}$$

where $P \subseteq \Lambda$ is a property of λ-expressions. Inductive reasoning on terms with binders is known to be hard when the binding behavior of λ yields bindings in the meta-language, making it difficult to write pattern-matching style recursive definitions and reasoning (see, e.g., [40]). For example, if we try to parse the above

[7] [5] gives credits to [64] and mentions that it can be used in many other logics.

inductive principle as a matching logic pattern, we will notice that $\forall x\colon \mathit{Var}$ in (‡) binds *nothing*—x is already bound in $\lambda x.\, e$.

There is relevant research on this topic, e.g., [32, 84, 27] for HOAS approaches and [90, 74] for nominal induction and recursion, which we will investigate and reconcile within matching logic. We believe that matching logic is particularly suitable for defining such inductive principles. Indeed, matching logic allows set variables, which are effectively universally quantified in formulas. Therefore, the second-order quantification $\forall P$ in the inductive principle above can be effectively captured in matching logic by simply dropping the $\forall P$ quantifier and letting the set variable P stay free in the formula.

Replacing Axiom Schemas with Axioms The matching logic theory Γ^{λ} for λ-calculus (Section 6) includes *axiom schema* (β) with meta-variables x, e, e', the same as the original λ-calculus. Thus, Γ^{λ} is a faithful definition of λ-calculus that captures it *as is*. This was intended and desired, because we believe that as a unifying logic for semantic frameworks (like $\mathbb{K}$), matching logic should allow us to define logics, calculi and languages as a mirror of the original, without any encodings or translations except for defining the necessary mathematical instruments and convenient notations. For practical reasons, it is also useful to define λ-calculus (and other binders) using axioms (not schemas) and normal variables (not meta-variables), as in nominal logic axiom (β in Nominal Logic) and HOAS (e.g., Twelf definition red-beta), both shown in Section 2. Thus, one way to eliminate schemas and meta-variables is to follow nominal and/or HOAS approaches methodologically, as explained in Remark 1; that is, we define nominal logic or HOAS in matching logic as theories and notations, and then define binders through them. However, matching logic also gives us an opportunity for alternative definitions. Below, we will show at a high level one example. Studying such alternative encodings of calculi is interesting and practical, but will be addressed in other places.

Recall that $\lambda x.\, e \equiv \mathsf{lambda}\,(\mathsf{intension}\ \exists x\colon \mathit{Var}.\,\langle x, e\rangle)$, where $(\mathsf{intension}\ \exists x\colon \mathit{Var}.\,\langle x, e\rangle)$ denotes the graph of $x \mapsto e$ as an element of sort 2^{Exp^2}. As pointed out in Section 6, not all elements of sort 2^{Exp^2} represent a graph, so we may identify and axiomatize a subsort *Graph* of 2^{Exp^2} that includes precisely all graphs. And thus, the schema (β) can be replaced by the following axiom:

$$(\beta,\ \text{Not a Schema}) \quad \forall g\colon \mathit{Graph}.\,\forall e'\colon \mathit{Exp}.\,(\mathsf{lambda}\ g)\ e' = \mathsf{graph\text{-}lookup}\ g\ e'$$

where g and e' are normal variables and graph-lookup is axiomatized as the graph lookup operation.

11 Conclusion

In this paper, we used (a functional variant of) matching logic to define binders in various logical systems. The binding behavior of binders in the object-level systems is directly inherited from the built-in binder $\exists$ in matching logic. We demonstrated our approach directly by defining λ-calculus as a matching logic theory, and indirectly by capturing term-generic logic (TGL); the latter yields matching logic definitions for many logical systems that feature bindings that were previously defined as TGL theories, including System F, pure type systems, π-calculus, etc. We proved the conservative extension theorems for all of these. We illustrated two proof methods: one based on models that is suitable for object-level systems that come equipped with models, and another based on syntax and proof derivations that is more involved but available even when the system lacks models. Our approach also yields *models* for the defined systems. For the systems discussed in the paper, the obtained models are complete w.r.t. logical reasoning, which follows from the conservative extension theorems. For λ-calculus, the models are representationally complete for all λ-theories, suggesting that matching logic is a promising alternative semantics for λ-calculus.

Acknowledgments

We warmly thank the $\mathbb{K}$ Team for invaluable and continuous feedback on matching logic and its role as a foundation of $\mathbb{K}$, as well as for their creative yet hard work on turning theoretical results into practical tools. We also warmly thank James Cheney, Maribel Fernández, Andrei Popescu, and Thomas Tuegel for many

insightful comments and concrete suggestions. We are indebted to the four anonymous reviewers, whose wit and dedication helped us improve the presentation. This work was supported in part by NSF CNS 16-19275. This material is based upon work supported by the United States Air Force and DARPA under Contract No. FA8750-18-C-0092.

References

[1] M. Abadi, L. Cardelli, P.-L. Curien, and J.-J. Lévy. Explicit substitutions. *Journal of Functional Programming*, 1(4):375–416, 1991.

[2] Areski Nait Abdallah. Partial first-order logic. In *The Logic of Partial Information*, Monographs in Theoretical Computer Science. An EATCS Series, chapter 14, pages 425–452. Springer, Berlin, Heidelberg, 1995.

[3] Mauricio Ayala-Rincón, Washington de Carvalho-Segundo, Maribel Fernández, and Daniele Nantes-Sobrinho. Nominal C-unification. In *Proceedings of the 27th International Symposium on Logic-Based Program Synthesis and Transformation (LOPSTR'17)*, volume 10855 of *Lecture Notes in Computer Science*, pages 235–251, Namur, Belgium, 2018. Springer International Publishing.

[4] Mauricio Ayala-Rincón, Maribel Fernández, and Daniele Nantes-Sobrinho. Nominal narrowing. In *Proceedings of the 1st International Conference on Formal Structures for Computation and Deduction (FSCD'16)*, volume 52 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 11:1–11:17, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[5] Brian Aydemir, Arthur Charguéraud, Benjamin C. Pierce, Randy Pollack, and Stephanie Weirich. Engineering formal metatheory. In *Proceedings of the 35th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL'08)*, pages 3–15, New York, NY, USA, 2008. ACM.

[6] Henk Barendregt. *The lambda calculus, its syntax and semantics*. Studies in Logic. College Publications, King's College London, Strand, London WC2R 2LS, UK, 1984.

[7] Henk Barendregt. Lambda calculi with types. In *Handbook of Logic in Computer Science*, volume 2, background: computational structures, chapter 2, pages 117–309. Oxford University Press, UK, 1993.

[8] John Bell and Moshe Machover. *A course in mathematical logic*. North Holland, Amsterdam, Netherlands, 1977.

[9] Chantal Berline. From computation to foundations via functions and application: the λ-calculus and its webbed models. *Theoretical Computer Science*, 249(1):81–161, 2000.

[10] Chantal Berline. Graph models of λ-calculus at work, and variations. *Mathematical Structures in Computer Science*, 16(2):185–221, 2006.

[11] Gilles Bernot, Michel Bidoit, and Christine Choppy. Abstract data types with exception handling: An initial approach based on a distinction between exceptions and errors. *Theoretical Computer Science*, 46:13–45, 1986.

[12] Gérard Berry. Stable models of typed λ-calculi. In *Automata, Languages and Programming*, pages 72–89, Berlin, Heidelberg, 1978. Springer.

[13] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal logic*. Cambridge University Press, One Liberty Plaza, New York, NY, 2001.

[14] C. J. Bloo. *Preservation of termination for explicit substitution*. PhD thesis, Technische Universiteit Eindhoven, 1997.

[15] Denis Bogdănaş and Grigore Roşu. K-Java: A complete semantics of Java. In *Proceedings of the 42nd Symposium on Principles of Programming Languages (POPL'15)*, pages 445–456, Mumbai, India, 2015. ACM.

[16] Antonio Bucciarelli and Thomas Ehrhard. A theory of sequentiality. *Theoretical Computer Science*, 113(2):273–291, 1993.

[17] Antonio Bucciarelli and Antonino Salibra. The sensible graph theories of lambda calculus. In *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science (LICS'04)*, pages 276–285, Turku, Finland, July 2004. IEEE.

[18] Peter Burmeister. Partial algebras—an introductory survey. In *Algebras and orders*, volume 389 of *NATO ASI Series*, pages 1–70. Springer, Dordrecht, Netherlands, 1993.

[19] Luca Cardelli. Type systems. *ACM Computing Surveys (CSUR)*, 28(1):263–264, 1996.

[20] Luca Cardelli, Simone Martini, John C. Mitchell, and Andre Scedrov. An extension of system F with subtyping. *Information and Computation*, 109(1):4–56, 1994.

[21] Xiaohong Chen and Grigore Roşu. Matching μ-logic. In *Proceedings of the 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS'19)*, pages 1–13, Vancouver, Canada, 2019. IEEE.

[22] Xiaohong Chen and Grigore Roşu. Matching μ-logic. Technical report, University of Illinois at Urbana-Champaign, 2019.

[23] Xiaohong Chen and Grigore Roşu. A general approach to define binders using matching logic. In *Proceedings of the 25th ACM SIGPLAN International Conference on Functional Programming (ICFP'20)*, New Jersey, USA, 2020.

[24] James Cheney. Completeness and Herbrand theorems for nominal logic. *Journal of Symbolic Logic*, 71(1):299–320, 2006.

[25] James Cheney. A simple sequent calculus for nominal logic. *Journal of Logic and Computation*, 26(2):699–726, 2014.

[26] James Cheney, Michael Norrish, and René Vestergaard. Formalizing adequacy: a case study for higher-order abstract syntax. *Journal of Automated Reasoning*, 49(2):209–239, 2012.

[27] Adam Chlipala. Parametric higher-order abstract syntax for mechanized semantics. In *Proceedings of the 13th ACM SIGPLAN International Conference on Functional Programming (ICFP'08)*, pages 143–156, British Columbia, Canada, 2008. ACM.

[28] Alonzo Church. *The calculi of lambda-conversion.* Princeton University Press, Princeton, New Jersey, USA, 1941.

[29] Bruno Courcelle and Joost Engelfriet. *Graph structure and monadic second-order logic: a language-theoretic approach*, volume 138. Cambridge University Press, England, UK, 2012.

[30] Sandeep Dasgupta, Daejun Park, Theodoros Kasampalis, Vikram S. Adve, and Grigore Roşu. A complete formal semantics of x86-64 user-level instruction set architecture. In *Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI'19)*, pages 1133–1148, Phoenix, Arizona, USA, 2019. ACM.

[31] Nicolaas Govert de Bruijn. Lambda calculus notation with nameless dummies, a tool for automatic formula manipulation, with application to the Church-Rosser theorem. *Indagationes Mathematicae*, 75(5):381–392, 1972.

[32] Joëlle Despeyroux, Amy Felty, and André Hirschowitz. Higher-order abstract syntax in Coq. In *Typed Lambda Calculi and Applications*, pages 124–138, Berlin, Heidelberg, 1995. Springer.

[33] Erwin Engeler. Algebras and combinators. *Algebra Universalis*, 13(1):389–392, 1981.

[34] Amy Felty and Alberto Momigliano. Hybrid, a definitional two-level approach to reasoning with higher-order abstract syntax. *Journal of Automated Reasoning*, 48(1):43–105, 2012.

[35] M. Fiore, G. Plotkin, and D. Turi. Abstract syntax and variable binding. In *Proceedings. 14^{th} Symposium on Logic in Computer Science (Cat. No. PR00158)*, pages 193–202, Trento, Italy, 1999. IEEE.

[36] Marcelo Fiore and Chung-Kil Hur. Second-order equational logic (extended abstract). In Anuj Dawar and Helmut Veith, editors, *Computer Science Logic*, pages 320–335, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[37] Marcelo Fiore and Ola Mahmoud. Second-order algebraic theories. In Petr Hliněný and Antonín Kučera, editors, *Mathematical Foundations of Computer Science 2010*, pages 368–380, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[38] M. Gabbay and A. Pitts. A new approach to abstract syntax involving binders. In *Proceedings of the 14^{th} Symposium on Logic in Computer Science (LICS'19)*, pages 214–224, Trento, Italy, July 1999. IEEE.

[39] Murdoch Gabbay and James Cheney. A sequent calculus for nominal logic. In *Proceedings of the 19^{th} Annual IEEE Symposium on Logic in Computer Science (LICS'04)*, pages 139–148, Washington, DC, USA, 2004. IEEE.

[40] Murdoch J. Gabbay. *A theory of inductive definitions with α-equivalence: semantics, implementation, programming language.* PhD thesis, DPMMS and Trinity College, Cambridge University, 2000.

[41] Murdoch J. Gabbay and Michael J. Gabbay. Representation and duality of the untyped λ-calculus in nominal lattice and topological semantics, with a proof of topological completeness. *Annals of Pure and Applied Logic Volume*, 168(3):501–621, October 2017.

[42] Andrew Gacek, Dale Miller, and Gopalan Nadathur. A two-level logic approach to reasoning about computations. *Journal of Automated Reasoning*, 49(2):241–273, 2012.

[43] Jean-Yves Girard. *Interprétation fonctionnelle et élimination des coupures de l'arithmétique d'ordre supérieur.* PhD thesis, Paris Diderot University, Paris, France, 1972.

[44] Jean-Yves Girard. The system F of variable types, fifteen years later. *Theoretical Computer Science*, 45:159–192, 1986.

[45] Jean-Yves Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–101, 1987.

[46] M. Gogolla, K. Drosten, U. Lipeck, and H.-D. Ehrich. Algebraic and operational semantics of specifications allowing exceptions and errors. *Theoretical Computer Science*, 34(3):289–313, 1984.

[47] Joseph Goguen and José Meseguer. Order-sorted algebra, part I: equational deduction for multiple inheritance, overloading, exceptions and partial operations. *Theoretical Computer Science*, 105(2):217–273, 1992.

[48] Robert Harper, Furio Honsell, and Gordon Plotkin. A framework for defining logics. *Journal of the ACM*, 40(1):143–184, 1993.

[49] Gisbert Hasenjaeger. Eine bemerkung zu Henkin's beweis für die vollständigkeit des prädikatenkalküls der ersten stufe. *The Journal of Symbolic Logic*, 18(1):42–48, 1953.

[50] Chris Hathhorn, Chucky Ellison, and Grigore Roşu. Defining the undefinedness of C. In *Proceedings of the 36th annual ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI'15)*, pages 336–345, Portland, OR, 2015. ACM.

[51] Everett Hildenbrandt, Manasvi Saxena, Xiaoran Zhu, Nishant Rodrigues, Philip Daian, Dwight Guth, Brandon Moore, Yi Zhang, Daejun Park, Andrei Ştefănescu, and Grigore Roşu. KEVM: A complete semantics of the Ethereum virtual machine. In *Proceedings of the 2018 IEEE Computer Security Foundations Symposium (CSF'18)*, pages 204–217, Oxford, UK, 2018. IEEE. `http://jellopaper.org`.

[52] Roger Hindley and Giuseppe Longo. Lambda-calculus models and extensionality. *Mathematical Logic Quarterly*, 26(4):289–310, 1980.

[53] K Team. K tutorials—λ-calculus. `https://github.com/kframework/k/tree/master/k-distribution/tutorial/1_k/1_lambda/lesson_2`, 2020.

[54] Delia Kesner. A theory of explicit substitutions with safe and full composition. *Logical Methods in Computer Science*, 5(3):1–29, 2009.

[55] Jan Willem Klop. Term rewriting systems. In *Handbook of Logic in Computer Science*, volume 2, Background: computational structures, chapter 1, pages 1–116. Oxford University Press, Inc., USA, 1993.

[56] C. P. J. Koymans. Models of the lambda calculus. *Information and Control*, 52:306–332, 1982.

[57] Jean Louis Krivine. *Lambda-calculus, types and models.* Ellis Horwood, USA, 1993.

[58] Patrick Lincoln and John Mitchell. Operational aspects of linear lambda calculus. In *Proceedings of the 7th Annual IEEE Symposium on Logic in Computer Science (LICS'92)*, pages 235–246, California, USA, June 1992. IEEE.

[59] Leopold Löwenheim. Über möglichkeiten im relativkalkül. *Mathematische Annalen*, 76(4):447–470, 1915.

[60] Francisca Lucio-Carrasco and Antonio Gavilanes-Franco. A first order logic for partial functions. In *Proceedings of the 6th Annual Symposium on Theoretical Aspects of Computer Science (STACS'89)*, pages 47–58, Paderborn, Germany, 1989. Springer.

[61] Giulio Manzonetto. *Models and theories of lambda calculus.* PhD thesis, Università Ca' Foscari di Venezia, 2008.

[62] Per Martin-Löf. *Twenty five years of constructive type theory*, volume 36 of *Oxford Logic Guides Book*, chapter An intuitionistic theory of types, pages 127–172. Oxford University Press, Oxford, UK, 1998.

[63] Raymond C. McDowell and Dale A. Miller. Reasoning with higher-order abstract syntax in a logical framework. *ACM Transactions on Computational Logic*, 3(1):80–136, 2002.

[64] James McKinna and Robert Pollack. Pure type systems formalized. In Marc Bezem and Jan Friso Groote, editors, *Typed Lambda Calculi and Applications*, pages 289–305, Berlin, Heidelberg, 1993. Springer.

[65] José Meseguer and Grigore Roşu. The rewriting logic semantics project: a progress report. *Information and Computation*, 231:38–69, October 2013. Invited paper at FCT 2011.

[66] Robin Milner, Joachim Parrow, and David Walker. A calculus of mobile processes (part 1). *Information and Computation*, 100(1):1–40, 1992.

[67] Timothy Nelson, Daniel Dougherty, Kathi Fisler, and Shriram Krishnamurthi. On the finite model property in order-sorted logic. Technical report, Worcester Polytechnic Institute, Brown University, 2010.

[68] Daejun Park, Andrei Ştefănescu, and Grigore Roşu. KJS: A complete formal semantics of JavaScript. In *Proceedings of the 36th annual ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI'15)*, pages 346–356, Portland, OR, 2015. ACM.

[69] Lawrence C. Paulson. The foundation of a generic theorem prover. *Journal of Automated Reasoning*, 5(3):363–397, 1989.

[70] Frank Pfenning and Conal Elliott. Higher-order abstract syntax. In *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI'88)*, pages 199–208, New York, NY, USA, 1988. ACM.

[71] Frank Pfenning and Carsten Schürmann. System description: Twelf—a meta-logical framework for deductive systems. In *Proceedings of the 16th International Conference on Automated Deduction (CADE 99)*, pages 202–206, Trento, Italy, 1999. Springer.

[72] Andrew M. Pitts. Nominal logic, a first order theory of names and binding. *Information and Computation*, 186(2):165–193, 2003.

[73] Andrew M. Pitts. Alpha-structural recursion and induction. In Joe Hurd and Tom Melham, editors, *Theorem Proving in Higher Order Logics*, pages 17–34, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[74] Andrew M. Pitts. *Nominal sets: names and symmetry in computer science*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, New York, NY, USA, 2013.

[75] Gordon Plotkin. A set-theoretical definition of application. Technical report, University of Edinburgh, 1972.

[76] Andrei Popescu and Grigore Roşu. Term-generic logic (extended technical report). Technical report, Technische Universitat Munchen, University of Illinois at Urbana-Champaign, 2013.

[77] Andrei Popescu and Grigore Roşu. Term-generic logic. *Theoretical Computer Science*, 577:1–24, 2015.

[78] Robert W. Quackenbush. Completeness theorems for universal and implicational logics of algebras via congruences. *Proceedings of the American Mathematical Society*, 103(4):1015–1021, 1988.

[79] John C. Reynolds. Towards a theory of type structure. In *Programming Symposium*, pages 408–425, Berlin, Heidelberg, 1974. Springer.

[80] John C. Reynolds. Separation logic: A logic for shared mutable data structures. In *Proceedings of the 17th Annual IEEE Symposium on Logic in Computer Science (LICS'02)*, pages 55–74, Copenhagen, Denmark, 2002. IEEE.

[81] Grigore Roşu. Matching logic. *Logical Methods in Computer Science*, 13(4):1–61, 2017.

[82] Grigore Roşu and Traian Florin Şerbănuţă. An overview of the K semantic framework. *Journal of Logic and Algebraic Programming*, 79(6):397–434, 2010.

[83] Harold Schellinx. Isomorphisms and nonisomorphisms of graph models. *Journal of Symbolic Logic*, 56(1):227–249, October 1991.

[84] Carsten Schürmann, Joëlle Despeyroux, and Frank Pfenning. Primitive recursion for higher-order abstract syntax. *Theoretical Computer Science*, 266(1):1–57, 2001.

[85] Dana Scott. Continuous lattices. In *Toposes, Algebraic Geometry and Logic*, pages 97–136, Berlin, Heidelberg, 1972. Springer.

[86] Dana Scott. Data types as lattices. *SIAM Journal on Computing*, 5(3):522–587, 1975.

[87] Dana Scott. Some philosophical issues concerning theories of combinators. In *Proceedings of the International Symposium on λ-Calculus and Computer Science Theory*, pages 346–366, Berlin, Heidelberg, 1975. Springer.

[88] Traian Florin Şerbănuţă and Grigore Roşu. A truly concurrent semantics for the K framework based on graph transformations. In *Proceedings of the 6th International Conference on Graph Transformation (ICGT'12)*, pages 294–310, Bremen, Germany, 2012. Springer.

[89] Mark-Oliver Stehr. CINNI—a generic calculus of explicit substitutions and its application to λ- ς- and ϕ-calculi. *Electronic Notes in Theoretical Computer Science*, 36:70–92, 2000.

[90] Christian Urban. Nominal techniques in Isabelle/HOL. *Journal of Automated Reasoning*, 40(4):327–356, May 2008.

[91] Jonni Virtema, Jeremy Meyers, and Antti Kuusisto. Undecidable first-order theories of affine geometries. *Logical Methods in Computer Science*, 9(4):1–23, 2013.

A Technical Details and Proofs for Section 3

A.1 Proof of Proposition 9

We have proved (1) and (6) in the main text. We prove the rest in the following.

Proof. For (2), we have $|\varphi_1 \vee \varphi_2|_\rho = |\neg\varphi_1 \to \varphi|_\rho = M \setminus (|\neg\varphi_1|_\rho \setminus |\varphi_2|_\rho) = M \setminus ((M \setminus |\varphi_1|_\rho) \setminus |\varphi_2|_\rho) = |\varphi_1|_\rho \cup |\varphi_2|_\rho$. For (3), we have $|\varphi_1 \wedge \varphi_2|_\rho = |\neg\varphi_1 \vee \neg\varphi_2|_\rho = |\neg\varphi_1|_\rho \cup |\neg\varphi_2|_\rho = (M \setminus |\varphi_1|_\rho) \cup (M \setminus |\varphi_2|_\rho) = |\varphi_1|_\rho \cap |\varphi_2|_\rho$ For (4), we have $|\top|_\rho = |\neg\bot|_\rho = M \setminus |\bot|_\rho = M \setminus \emptyset = M$. For (5), we have $|\varphi_1 \leftrightarrow \varphi_2|_\rho = |(\varphi_1 \to \varphi_2) \wedge (\varphi_2 \to \varphi_1)|_\rho = |\varphi_1 \to \varphi_2|_\rho \cap |\varphi_2 \to \varphi_1|_\rho = (M \setminus (|\varphi_1|_\rho \setminus |\varphi_2|_\rho)) \cap (M \setminus (|\varphi_2|_\rho \setminus |\varphi_1|_\rho)) = M \setminus ((|\varphi_1|_\rho \setminus |\varphi_2|_\rho) \cup (|\varphi_2|_\rho \setminus |\varphi_1|_\rho)) = M \setminus (|\varphi_1|_\rho \triangle |\varphi_2|_\rho)$. □

A.2 Proof of Proposition 14

Proof. For (1), we remind the reader that $\lceil a \rceil_M$ means $\lceil \_ \rceil_M \cdot a$. Therefore, for a valuation ρ such that $\rho(x) = a$, we have $|\lceil x \rceil|_\rho = |\lceil \_ \rceil x|_\rho = \lceil \_ \rceil_M \cdot |x|_\rho = \lceil \_ \rceil_M \cdot \{a\} = \lceil \_ \rceil_M \cdot a = M$. For (2), let us first suppose $|\varphi|_\rho \neq \emptyset$. Then, there exists $a \in M$ such that $a \in |\varphi|_\rho$. Then, we have $|\lceil \varphi \rceil|_\rho = \lceil \_ \rceil_M \cdot |\varphi|_\rho \supseteq \lceil \_ \rceil_M \cdot \{a\} = M$. Therefore, $|\lceil \varphi \rceil|_\rho = M$. Now, let us suppose $|\varphi|_\rho = \emptyset$. Then, we have $|\lceil \varphi \rceil|_\rho = \lceil \_ \rceil_M \cdot |\varphi|_\rho = \lceil \_ \rceil_M \cdot \emptyset = \emptyset$, where the final step is by pointwise extension 4. For (3), let us first suppose $|\varphi|_\rho = M$, and thus $|\neg\varphi|_\rho = M \setminus |\varphi|_\rho = M \setminus M = \emptyset$. Then, we have $|\lfloor \varphi \rfloor|_\rho = |\neg\lceil\neg\varphi\rceil|_\rho = M \setminus |\lceil\neg\varphi\rceil|_\rho = M \setminus \emptyset = M$. Now, let us suppose $|\varphi|_\rho \neq M$, and thus there exists $a \in M$ such that $a \notin |\varphi|_\rho$, i.e., $a \in |\neg\varphi|_\rho$. Then, we have $|\lfloor \varphi \rfloor|_\rho = |\neg\lceil\neg\varphi\rceil|_\rho = M \setminus |\lceil\neg\varphi\rceil|_\rho = M \setminus M = \emptyset$. For (4), let us first suppose $|\varphi_1|_\rho = |\varphi_2|_\rho$, and thus $|\varphi_1 \leftrightarrow \varphi_2|_\rho = M \setminus (|\varphi_1|_\rho \triangle |\varphi_2|_\rho) = M \setminus \emptyset = M$. Then, we have $|\varphi_1 = \varphi_2|_\rho = |\lfloor \varphi_1 \leftrightarrow \varphi_2 \rfloor|_\rho = M$. Now, let us suppose $|\varphi_1|_\rho \neq |\varphi_2|_\rho$. Then, we have $(|\varphi_1|_\rho \triangle |\varphi_2|_\rho) \neq \emptyset$, and thus $|\varphi_1 \leftrightarrow \varphi_2|_\rho = M \setminus (|\varphi_1|_\rho \triangle |\varphi_2|_\rho) \neq M$. Therefore, we have $|\varphi_1 = \varphi_2|_\rho = |\lfloor \varphi_1 \leftrightarrow \varphi_2 \rfloor|_\rho = \emptyset$. For (5), let us first suppose $\rho(x) \in |\varphi|_\rho$, and thus $|x \wedge \varphi|_\rho = \{\rho(x)\} \cap |\varphi|_\rho = \{\rho(x)\} \neq \emptyset$. Then, we have $|x \in \varphi|_\rho = |\lceil x \wedge \varphi \rceil|_\rho = M$. Now, let us suppose $\rho(x) \notin |\varphi|_\rho$, and thus $|x \wedge \varphi|_\rho = \{\rho(x)\} \cap |\varphi|_\rho = \emptyset$. Then, we have $|x \in \varphi|_\rho = |\lceil x \wedge \varphi \rceil|_\rho = \emptyset$. For (6), let us first suppose $|\varphi_1|_\rho \subseteq |\varphi_2|_\rho$, and thus $|\varphi_1 \to \varphi_2|_\rho = M \setminus (|\varphi_1|_\rho \setminus |\varphi_2|_\rho) = M \setminus \emptyset = M$. Then, we have $|\varphi_1 \subseteq \varphi_2|_\rho = |\lfloor \varphi_1 \to \varphi_2 \rfloor|_\rho = M$. Now, let us suppose $|\varphi_1|_\rho \not\subseteq |\varphi_2|_\rho$, and thus $|\varphi_1 \to \varphi_2|_\rho = M \setminus (|\varphi_1|_\rho \setminus |\varphi_2|_\rho) \neq M$. Then, we have $|\varphi_1 \subseteq \varphi_2|_\rho = |\lfloor \varphi_1 \to \varphi_2 \rfloor|_\rho = \emptyset$. □

B Technical Details and Proofs for Section 4

B.1 Proof of Proposition 19

Proof. Let us fix a model M and an interpretation of the (constant) pairing symbol $\langle \_, \_ \rangle$, which we denote as $\langle \_, \_ \rangle_M \subseteq M$. For any $a \in M_{s_1}$ and $b \in M_{s_2}$, we abbreviate $\langle \_, \_ \rangle_M \cdot a \cdot b$ as $\langle a, b \rangle_M$. Recall that pairing is a function, defined as $\langle \_, \_ \rangle \colon s_1 \times s_2 \to s_1 \otimes s_2$. By the axiom (Function), we have that $\langle a, b \rangle_M$ is a singleton, for any $a, b \in M_s$. By abuse of notation (see the discussion before Example 5), we denote the only element in the singleton $\langle a, b \rangle_M$ also as $\langle a, b \rangle_M$. By the (Product) axiom, we have that $M_{s_1 \otimes s_2} = \bigcup_{a \in M_{s_1}, b \in M_{s_2}} \langle a, b \rangle_M$, so there exists a surjective function $i \colon M_{s_1} \times M_{s_2} \to M_{s_1 \otimes s_2}$, given as $i(a, b) = \langle a, b \rangle_M$ for any $a \in M_{s_1}, b \in M_{s_2}$. By the axiom (Injectivity), we know that i is an injective function, and thus it is a bijection. Therefore, $M_{s_1} \times M_{s_2} \cong M_{s_1 \otimes s_2}$. □

B.2 Proof of Proposition 21

Proposition 53. For any model M validating the axioms in Definition 20, we have $M_{2^s} \cong \mathcal{P}(M_s)$.

Proof. Let us fix a model M and an interpretation of the (constant) extension symbol $\mathsf{extension}$, which we denote as $\mathsf{extension}_M \subseteq M$. Let us define a function $\mathsf{extension}_M(\_) \colon M_{2^A} \to \mathcal{P}(M_s)$ as $\mathsf{extension}_M(A) = \mathsf{extension}_M \cdot A$, for any $A \in M_{2^A}$. Note that the range of $\mathsf{extension}_M(\_)$ is $\mathcal{P}(M_s)$ because of the axiom (Arity) in Definition 20. In the following, we show that $\mathsf{extension}_M(\_)$ is a bijection. For the injectivity,

FOL Reasoning	(Propositional Tautology)	φ if φ is a propositional tautology over patterns
	(Modus Ponens)	$\dfrac{\varphi_1 \quad \varphi_1 \to \varphi_2}{\varphi_2}$
	(∃-Quantifier)	$\varphi[y/x] \to \exists x.\, \varphi$
	(∃-Generalization)	$\dfrac{\varphi_1 \to \varphi_2}{(\exists x.\varphi_1) \to \varphi_2}$ if $x \notin \mathrm{FV}(\varphi_2)$
Frame Reasoning	(Propagation$_\bot$)	$C[\bot] \to \bot$
	(Propagation$_\vee$)	$C[\varphi_1 \vee \varphi_2] \to C[\varphi_1] \vee C[\varphi_2]$
	(Propagation$_\exists$)	$C[\exists x.\, \varphi] \to \exists x.\, C[\varphi]$ if $x \notin \mathrm{FV}(C)$
	(Framing)	$\dfrac{\varphi_1 \to \varphi_2}{C[\varphi_1] \to C[\varphi_2]}$
Technical Rules	(Set Variable Substitution)	$\dfrac{\varphi}{\varphi[\psi/X]}$
	(Existence)	$\exists x.\, x$
	(Singleton)	$\neg\left(C_1[x \wedge \varphi] \wedge C_2[x \wedge \neg\varphi]\right)$

Figure 7: Matching logic proof system (where $C[\varphi]$ denotes an application pattern $\varphi\,\psi$ or $\psi\,\varphi$ for some ψ)

we consider $A, B \in M_{2^s}$ with $A \neq B$. Then by the axiom (Extensionality), we have $\mathsf{extension}_M(A) \neq \mathsf{extension}_M(B)$. For the surjectivity, we consider an arbitrary $C \in \mathcal{P}(M_s)$ and a valuation ρ such that $\rho(X) = C$, where X is the free set variable that occurs in the axiom (Powerset) in Definition 20. Then by the axiom, there exists $A \in M_{2^s}$ such that $\mathsf{extension}_M(A) = C$. Therefore, $\mathsf{extension}(\_)_M$ is a bijection, and we have proved $M_{2^s} \cong \mathcal{P}(M_s)$. □

Remark 54. We define $\mathsf{intension}_M(\_)\colon \mathcal{P}(M_s) \to M_{2^s}$ to be the inverse of $\mathsf{extension}_M(\_)$.

Remark 55. Given M, if $M_{2^s} = \mathcal{P}(M_s)$, then both $\mathsf{extension}_M(\_)$ and $\mathsf{intension}_M(\_)$ become the identity function over $\mathcal{P}(M_s)$, i.e., $\mathsf{extension}_M(C) = \mathsf{intension}_M(C) = C$, for any $C \subseteq M_s$. For such M, it is sometimes confusing whether pointwise extension is triggered because C can mean either an element in M_{2^s} or a subset of M_s, and only the latter requires pointwise extension. To prevent this confusion, we will use $\_\bullet\_\colon M \times M \to \mathcal{P}(M)$ to mean only the interpretation of application, and use a different notation $\_\bullet_{pe}\_\colon \mathcal{P}(M) \times \mathcal{P}(M) \to \mathcal{P}(M)$ to mean its pointwise extension, defined the same as in Definition 4: $A \bullet_{pe} B = \bigcap_{a\in A, b\in B} a \bullet b$ for any $A, B \subseteq M$,

B.3 Matching Logic Proof System, Proof of Proposition 23, and Proof of Theorem 24

The Hilbert-style proof system of ML is shown in Fig. 7. This proof system is obtained by instantiating the Hilbert system given in [21] on the functional variant (Definition 2). Its meta-properties, including Proposition 23 and Theorem 24, need to be proved.

We first prove Proposition 23.

Proof. (1)-(4) can be proved by the FOL reasoning rules (Fig. 7). In the following, we will use standard propositional reasoning without explicitly showing their formal proofs in ML.

For (5), we need to prove $\Gamma \vdash \varphi = \varphi$, i.e., $\Gamma \vdash \lceil\neg(\varphi \leftrightarrow \varphi)\rceil \to \bot$. By propositional reasoning, this can be divided into proving (5a) $\Gamma \vdash \lceil\neg(\varphi \leftrightarrow \varphi)\rceil \to \lceil\bot\rceil$ and (5b) $\Gamma \vdash \lceil\bot\rceil \to \bot$. Note that (5b) is proved by (Propagation$_\bot$). For (5a), we apply (Framing). Then, we need to prove $\Gamma \vdash \neg(\varphi \leftrightarrow \varphi) \to \bot$, which is a propositional tautology.

For (6), we need to prove $\Gamma \vdash \varphi_1 = \varphi_3$. By a similar argument as in (5), we need to prove that $\Gamma \vdash \varphi_1 \leftrightarrow \varphi_3$. In the following, we show that $\Gamma \vdash \varphi_1 = \varphi_2$ implies that $\Gamma \vdash \varphi_1 \leftrightarrow \varphi_2$. Clearly, once we prove that, we can finish the proof of (7) by the standard propositional reasoning. We will prove a more general

result: $\Gamma \vdash \neg\lceil\psi\rceil$ implies $\Gamma \vdash \neg\psi$. By propositional reasoning, we only need to prove that $\Gamma \vdash \neg\lceil\psi\rceil \to \neg\psi$, i.e., $\Gamma \vdash \psi \to \lceil\psi\rceil$, whose proof is given in [22, Corollary 59].[8]

For (7), the proof is trivial by noting $\varphi_1 = \varphi_2 \equiv \lfloor\varphi_1 \leftrightarrow \varphi_2\rfloor$, and by propositional reasoning, $\Gamma \vdash (\varphi_1 \leftrightarrow \varphi_2) \leftrightarrow (\varphi_1 \leftrightarrow \varphi_2)$.

For (8), we can apply a similar argument as in (5) and only need to prove that $\Gamma \vdash \varphi_1 \leftrightarrow \varphi_2$ implies $\Gamma \vdash \psi[\varphi/x] \leftrightarrow \psi[\varphi/x]$. The latter can be proved by structural induction on ψ, and all cases can be proved by standard propositional reasoning. □

Next we prove Theorem 24.

Proof. We prove that all proof rules (and axioms) of the ML proof system in Fig. 7 are sound.

For the first four FOL rules, the proof follows directly by Remark 8.

For (Propagation$_\bot$), we have $|(\bot\,\varphi) \to \bot|_\rho = M \setminus |\bot\,\varphi|_\rho = M \setminus (\emptyset \bullet |\varphi|_\rho) = M \setminus \emptyset = M$. The case for $(\varphi\,\bot)$ is proved similarly.

For (Propagation$_\vee$), we have $|((\varphi_1 \vee \varphi_2)\,\psi) \to \varphi_1\,\psi \vee \varphi_1\,\psi|_\rho = M \setminus (|(\varphi_1 \vee \varphi_2)\,\psi|_\rho \setminus |\varphi_1\,\psi \vee \varphi_1\,\psi|_\rho) = M \setminus ((|\varphi_1 \vee \varphi_2|_\rho \bullet |\psi|_\rho) \setminus (|\varphi_1\,\psi|_\rho \cup |\varphi_1\,\psi|_\rho)) = M \setminus ((|\varphi_1|_\rho \cup |\varphi_2|_\rho \bullet |\psi|_\rho) \setminus (|\varphi_1|_\rho \bullet |\psi|_\rho \cup |\varphi_1|_\rho \bullet |\psi|_\rho))$, which, by pointwise extension, equals to $M \setminus \emptyset = M$. The case for $(\psi\,(\varphi_1 \vee \varphi_2))$ is proved similarly.

For (Propagation$_\exists$), we have $|((\exists x.\,\varphi)\,\psi) \to \exists x.\,(\varphi\,\psi)|_\rho = M \setminus (|(\exists x.\,\varphi)\,\psi|_\rho \setminus |\exists x.\,(\varphi\,\psi)|_\rho) = M \setminus (((\bigcup_a |\varphi|_{\rho[a/x]}) \bullet |\psi|_\rho) \setminus \bigcup_a |\varphi\,\psi|_{\rho[a/x]}) = M \setminus (((\bigcup_a |\varphi|_{\rho[a/x]}) \bullet |\psi|_\rho) \setminus \bigcup_a |\varphi|_{\rho[a/x]} \bullet |\psi|_{\rho[a/x]})$, which, by pointwise extension, equals to $M \setminus \emptyset = M$. The case for $(\psi\,(\exists x.\,\varphi))$ is proved similarly.

For (Framing), we have $|(\varphi_1\,\psi) \to (\varphi_2\,\psi)|_\rho = M \setminus (|\varphi_1\,\psi|_\rho \setminus |\varphi_2\,\psi|_\rho) = M \setminus ((|\varphi_1|_\rho \bullet |\psi|_\rho) \setminus (|\varphi_2|_\rho \bullet |\psi|_\rho))$. Now note that $|\varphi_1 \to \varphi_2|_\rho = M$, which implies that $|\varphi_1|_\rho \subseteq |\varphi_2|_\rho$, and therefore by pointwise extension, $(|\varphi_1|_\rho \bullet |\psi|_\rho) \setminus (|\varphi_2|_\rho \bullet |\psi|_\rho) = \emptyset$, and thus $|(\varphi_1\,\psi) \to (\varphi_2\,\psi)|_\rho = M \setminus ((|\varphi_1|_\rho \bullet |\psi|_\rho) \setminus (|\varphi_2|_\rho \bullet |\psi|_\rho)) = M \setminus \emptyset = M$.

For (Set Variable Substitution), we have $|\varphi[\psi/X]|_\rho = |\varphi|_{\rho[|\psi|_\rho/X]} = M$.

For (Existence), we have $|\exists x.\,x|_\rho = \bigcup_a \{a\} = M$.

For (Singleton), we note that $|x|_\rho$ is a singleton, so exactly one of $|x \wedge \varphi|_\rho$ and $|x \wedge \neg\varphi|_\rho$ is $\emptyset$. Then by pointwise extension, exactly one of $|C_1[x \wedge \varphi]|_\rho$ and $|C_2[x \wedge \neg\varphi]|_\rho$ is $\emptyset$, and thus we have $|C_1[x \wedge \varphi] \wedge C_2[x \wedge \neg\varphi]|_\rho = \emptyset$. Then we have that $|\neg(C_1[x \wedge \varphi] \wedge C_2[x \wedge \neg\varphi])|_\rho = M \setminus |C_1[x \wedge \varphi] \wedge C_2[x \wedge \neg\varphi]|_\rho = M \setminus \emptyset = M$.

In conclusion, all ML proof rules and axioms in Fig. 7 are sound. □

C Technical Details and Proofs for Section 7

In this section, let us fix a concrete ccc model $(A, \_\bullet_A\_, \mathbb{L})$. We recall the following notations (see Definition 25):

1. $R(A) = \{f\colon A \to A \mid \text{there exists a } b \in A \text{ such that } f(a) = b \bullet_A b \text{ for all } a \in A\}$;
2. $\_\bullet_A\_\colon A \times A \to A$;
3. $\mathbb{L}\colon R(A) \to A$.

For any $f\colon A \to A$, we define $\mathsf{graph}(f) = \{(a, f(a)) \mid a \in A\} \subseteq A \times A$.

Remark 56. For two arbitrary elements a, b, we write $a \rightsquigarrow b$ to mean the sequence consisting of a, b. In general, for elements $a_1, \ldots, a_n$, we write $a_1 \rightsquigarrow \cdots \rightsquigarrow a_n$ to mean the sequence consisting of $a_1, \ldots, a_n$. We will use sequences to represent the results of a *partial evaluation*.

[8] The corollary is for the full ML, not for the functional variant, but we can re-use the formal ML proof verbatim, under the notations introduced in Section 4.

C.1 Construction of the ML model M^A

Definition 57. Let Σ^λ be the signature of Γ^λ, which contains:

1. $\lceil \_ \rceil$, the definedness symbol (Definition 13);
2. $[\![\_]\!]$, the inhabitant symbol (Definition 17);
3. $\langle \_, \_ \rangle$, the pairing symbol (Definition 18);
4. $\mathsf{extension}$, the extension symbol (Definition 20);
5. $\mathit{Var}, \mathit{Exp}, \mathit{Var} \times \mathit{Exp}, 2^{\mathit{Var} \times \mathit{Exp}}$, the sort constants (Section 6);
6. lambda, the retraction symbol for λ (Section 6).

Given a concrete ccc model $(A, \_ \bullet_A \_, \mathbb{L})$, we define an ML model M^A of signature Σ^λ in the following way. For notational simplicity, we omit the superscript A and simply write M. Recall that an ML model is a tuple $(M, \_ \bullet \_, \{\sigma_M\}_{\sigma \in \Sigma^\lambda})$; see Definition 4.

Firstly, we define the carrier set M as the *disjoint union* of the following sets:

1. $\{\#\mathsf{def}\}$, where $\#\mathsf{def}$ is a distinguished element, used to interpret the definedness symbol $\lceil \_ \rceil$;
2. $\{\#\mathsf{inh}\}$, where $\#\mathsf{inh}$ is a distinguished element, used to interpret the inhabitant symbol $[\![\_]\!]$;
3. $\{\#\mathsf{Var}, \#\mathsf{Exp}, \#\mathsf{VarExp}, \#\mathsf{2VarExp}\}$, each interpreting the sort names $\mathit{Var}, \mathit{Exp}, \mathit{Var} \times \mathit{Exp}, 2^{\mathit{Var} \times \mathit{Exp}}$;
4. A;
5. $A \times A$;
6. $\mathcal{P}(A \times A)$;
7. $\{\#\mathsf{pair}, \#\mathsf{ext}, \#\mathsf{lam}\}$, each interpreting the (constant) symbols $\langle \_, \_ \rangle, \mathsf{extension}, \mathsf{lambda}$;
8. $\{\#\mathsf{pair} \rightsquigarrow a \mid a \in A\}$, where $\#\mathsf{pair} \rightsquigarrow a$ is the partial evaluation result of applying $\#\mathsf{pair}$ to a.

Secondly, we define the interpretation of application $\_ \bullet \_ : M \times M \to \mathcal{P}(M)$ as follows:

1. $\#\mathsf{def} \bullet a = M$ for every $a \in M$;
2. $\#\mathsf{inh} \bullet \#\mathsf{Var} = A$;
3. $\#\mathsf{inh} \bullet \#\mathsf{Exp} = A$;
4. $\#\mathsf{inh} \bullet \#\mathsf{VarExp} = A \times A$;
5. $\#\mathsf{inh} \bullet \#\mathsf{2VarExp} = \mathcal{P}(A \times A)$;
6. $\#\mathsf{pair} \bullet a = \{\#\mathsf{pair} \rightsquigarrow a\}$ for every $a \in A$;
7. $(\#\mathsf{pair} \rightsquigarrow a) \bullet b = \{(a, b)\}$ for every $a, b \in A$; note that (a, b) is the pair of a and b in $A \times A$;
8. $\#\mathsf{ext} \bullet P = P$ for every $P \in \mathcal{P}(A \times A)$ note that $P \subseteq A \times A \subseteq \mathcal{P}(M)$;
9. $\#\mathsf{lam} \bullet P = \{\mathbb{L}(f_P)\}$, if $P \in \mathcal{P}(A \times A)$, $f_P \in R(A)$, $P = \mathsf{graph}(f_P)$, and $\mathbb{L}(f_P)$ is defined;
10. $a \bullet b = \{a \bullet_A b\}$ for every $a, b \in A$;
11. Otherwise, if none of the above rules applies, $a \bullet b = \emptyset$ for $a, b \in M$.

Thirdly, we give interpretations to all symbols in Σ^λ as follows:

1. $\lceil \_ \rceil_M = \{\#\mathsf{def}\}$;
2. $[\![\_]\!]_M = \{\#\mathsf{inh}\}$;
3. $\langle \_, \_ \rangle_M = \{\#\mathsf{pair}\}$;
4. $\mathsf{extension}_M = \{\#\mathsf{ext}\}$;
5. $Var_M = \{\#\mathsf{Var}\}$;
6. $Exp_M = \{\#\mathsf{Exp}\}$;
7. $(Var \times Exp)_M = \{\#\mathsf{VarExp}\}$;
8. $(2^{Var \times Exp})_M = \{\#\mathsf{2VarExp}\}$;
9. $\mathsf{lambda}_M = \{\#\mathsf{lam}\}$.

And now we finish the construction of the ML model M.

Lemma 58. *For any pattern φ and valuation ρ such that $|\varphi|_\rho \subseteq A \times A$, we have $|\mathsf{intension}\,\varphi|_\rho = \left\{|\varphi|_\rho\right\}$, i.e., the singleton that contains $|\varphi|_\rho$.*

Proof. By the construction of the ML model M and Remark 55. □

C.2 Proof of Lemma 32

Proof. Let us fix a valuation on A, say $\rho\colon V^\lambda \to A$. We define a corresponding ML M-valuation ρ^A as $\rho^A(x) = \rho(x)$ for every $x \in V^\lambda$. We will prove that $|e|_{\rho^A} = \left\{|e|^\lambda_\rho\right\}$ for all $e \in \Lambda$ by structural induction on e.

To prevent confusion, we will use $\_\bullet\_$ to mean only the interpretation of application, and use $\_\bullet_{pe}\_$ to mean its pointwise extension; see Remark 55.

When e is a variable $x \in V^\lambda$, we have $|x|_{\rho^A} = \{\rho^A(x)\} = \{\rho(x)\} = \left\{|x|^\lambda_\rho\right\}$.

When e has the form $e_1\,e_2$, we have $|e_1\,e_2|_{\rho^A} = |e_1|_{\rho^A} \bullet_{pe} |e_2|_{\rho^A} = \left\{|e_1|^\lambda_\rho\right\} \bullet_{pe} \left\{|e_2|^\lambda_\rho\right\} = |e_1|^\lambda_\rho \bullet |e_2|^\lambda_\rho = \left\{|e_1|^\lambda_\rho \bullet_A |e_2|^\lambda_\rho\right\} = \left\{|e_1\,e_2|^\lambda_\rho\right\}$.

When e has the form $\lambda x.\, e_1$, we have $|\lambda x.\, e_1|_{\rho^A} = |\mathsf{lambda}\,(\mathsf{intension}\,\exists x\colon Var.\,\langle x, e_1\rangle)|_{\rho^A} = \{\#\mathsf{lam}\} \bullet_{pe} |\mathsf{intension}\,\exists x\colon Var.\,\langle x, e_1\rangle|_{\rho^A}$. Note that $|\exists x\colon Var.\,\langle x, e_1\rangle|_{\rho^A} = \bigcup_{a\in A} |\langle x, e_1\rangle|_{\rho^A[a/x]} = \bigcup_{a\in A} \{\#\mathsf{pair}\} \bullet_{pe} |x|_{\rho^A[a/x]} \bullet_{pe} |e_1|_{\rho^A[a/x]} = \bigcup_{a\in A} \{\#\mathsf{pair}\} \bullet_{pe} \{a\} \bullet_{pe} \left\{|e_1|^\lambda_{\rho^A[a/x]}\right\} = \bigcup_{a\in A} \left\{\#\mathsf{pair} \bullet a \bullet |e_1|^\lambda_{\rho^A[a/x]}\right\} = \bigcup_{a\in A} \left\{\left(a, |e_1|^\lambda_{\rho^A[a/x]}\right)\right\} = \mathsf{graph}(f^\rho_{e_1,x})$, where $f^\rho_{e_1,x}$ is defined in Definition 25. Then, using Lemma 58, we have that $\{\#\mathsf{lam}\} \bullet_{pe} |\mathsf{intension}\,\exists x\colon Var.\,\langle x, e_1\rangle|_{\rho^A} = \{\#\mathsf{lam}\} \bullet_{pe} \left\{|\exists x\colon Var.\,\langle x, e_1\rangle|_{\rho^A}\right\} = \{\#\mathsf{lam}\} \bullet_{pe} \left\{\mathsf{graph}(f^\rho_{e_1,x})\right\} = \#\mathsf{lam} \bullet \mathsf{graph}(f^\rho_{e_1,x}) = \{\mathbb{L}(f^\rho_{e_1,x})\} = |\lambda x.\, e_1|^\lambda_\rho$. □

D Technical Details and Proofs for Section 8

Some notations used in this section are defined in Appendix C.

D.1 Construction of the Term Model T

Definition 59. Recall that Σ^λ is the signature of Γ^λ that contains:

1. $\lceil \_ \rceil$, the definedness symbol (Definition 13);
2. $[\![\_]\!]$, the inhabitant symbol (Definition 17);

3. $\langle \_, \_ \rangle$, the pairing symbol (Definition 18);
4. $\mathsf{extension}$, the extension symbol (Definition 20);
5. $\mathit{Var}, \mathit{Exp}, \mathit{Var} \times \mathit{Exp}, 2^{\mathit{Var} \times \mathit{Exp}}$, the sort constants (Section 6);
6. lambda, the retraction symbol for λ (Section 6).

We define an ML model T of signature Σ^λ in the following way. Recall that an ML model is a tuple $(T, \_\bullet\_, \{\sigma_T\}_{\sigma \in \Sigma^\lambda})$; see Definition 4.

Firstly, we define the carrier set T as the *disjoint union* of the following sets:

1. $\{\#\mathsf{def}\}$, where $\#\mathsf{def}$ is a distinguished element, used to interpret the definedness symbol $\lceil \_ \rceil$;
2. $\{\#\mathsf{inh}\}$, where $\#\mathsf{inh}$ is a distinguished element, used to interpret the inhabitant symbol $[\![\_]\!]$;
3. $\{\#\mathsf{Var}, \#\mathsf{Exp}, \#\mathsf{VarExp}, \#\mathsf{2VarExp}\}$, each interpreting the sort names $\mathit{Var}, \mathit{Exp}, \mathit{Var} \times \mathit{Exp}, 2^{\mathit{Var} \times \mathit{Exp}}$;
4. $[\Lambda]$; note that this includes $[V^\lambda]$;
5. $[V^\lambda] \times [\Lambda]$;
6. $\mathcal{P}\big([V^\lambda] \times [\Lambda]\big)$;
7. $\{\#\mathsf{pair}, \#\mathsf{ext}, \#\mathsf{lam}\}$, each interpreting the (constant) symbols $\langle \_, \_ \rangle, \mathsf{extension}, \mathsf{lambda}$;
8. $\{\#\mathsf{pair} \rightsquigarrow [e] \mid e \in \Lambda\}$.

Secondly, we define the interpretation of application $\_\bullet\_ : T \times T \to \mathcal{P}(T)$ as follows:

1. $\#\mathsf{def} \bullet a = T$ for every $a \in T$;
2. $\#\mathsf{inh} \bullet \#\mathsf{Var} = [V^\lambda]$;
3. $\#\mathsf{inh} \bullet \#\mathsf{Exp} = [\Lambda]$;
4. $\#\mathsf{inh} \bullet \#\mathsf{VarExp} = [V^\lambda] \times [\Lambda]$;
5. $\#\mathsf{inh} \bullet \#\mathsf{2VarExp} = \mathcal{P}\big([V^\lambda] \times [\Lambda]\big)$;
6. $\#\mathsf{pair} \bullet [e] = \{\#\mathsf{pair} \rightsquigarrow a\}$ for every $e \in \Lambda$;
7. $(\#\mathsf{pair} \rightsquigarrow [e]) \bullet [e'] = \{([e], [e'])\}$ for every $e, e' \in \Lambda$;
8. $\#\mathsf{ext} \bullet P = P$ for every $P \in \mathcal{P}\big([V^\lambda] \times [\Lambda]\big)$ note that $P \subseteq [V^\lambda] \times [\Lambda] \subseteq \mathcal{P}(T)$;
9. $\#\mathsf{lam} \bullet P = \{[\lambda x.\, e]\}$, if $P = \bigcup_{z \in V^\lambda} ([z], [e[z/x]])$; well-definedness is proved in Proposition 40.
10. $[e] \bullet [e'] = \{[e\, e']\}$ for every $e, e' \in [\Lambda]$;
11. Otherwise, if none of the above rules applies, $a \bullet b = \emptyset$ for $a, b \in T$.

Thirdly, we give interpretations to all symbols in Σ^λ as follows:

1. $\lceil \_ \rceil_M = \{\#\mathsf{def}\}$;
2. $[\![\_]\!]_M = \{\#\mathsf{inh}\}$;
3. $\langle \_, \_ \rangle_M = \{\#\mathsf{pair}\}$;
4. $\mathsf{extension}_M = \{\#\mathsf{ext}\}$;

5. $Var_M = \{\#\mathsf{Var}\}$;
6. $Exp_M = \{\#\mathsf{Exp}\}$;
7. $(Var \times Exp)_M = \{\#\mathsf{VarExp}\}$;
8. $(2^{Var \times Exp})_M = \{\#\mathsf{2VarExp}\}$;
9. $\mathsf{lambda}_M = \{\#\mathsf{lam}\}$.

And now we finish the construction of the term model T.

Note that Lemma 58 also holds for T.

D.2 Proof of Proposition 41

Proof. We prove both properties simultaneously by induction on $d(e)$.

When $d(e) = 0$, we have that e is either a variable $x \in V^\lambda$ or an application $e_1\, e_2$ where $d(e_1) = d(e_2) = 0$. We do structure induction on e. When e is variable x, we have $|x|_{\rho_T} = \{\rho_T(x)\} = \{[x]\}$, and $|x|_{\rho[\rho(z)/x]} = \{\rho(z)\} = |z|_\rho = |x[z/x]|_\rho$. When e is variable y distinct from x, we have $|y|_{\rho_T} = \{\rho_T(y)\} = \{[y]\}$, and $|y|_{\rho[\rho(z)/x]} = \{\rho(y)\} = |y|_\rho = |y[z/x]|_\rho$. When e is $e_1\, e_2$, we have $|e_1 e_2|_{\rho_T} = |e_1|_{\rho_T} \bullet_{pe} |e_2|_{\rho_T} = \{[e_1]\} \bullet_{pe} \{[e_2]\} = [e_1] \cdot [e_2] = \{[e_1 e_2]\}$, and $|e_1 e_2|_{\rho[\rho(z)/x]} = |e_1|_{\rho[\rho(z)/x]} \bullet_{pe} |e_2|_{\rho[\rho(z)/x]} = |e_1[z/x]|_\rho \bullet_{pe} |e_2[z/x]|_\rho = |e_1[z/x]\, e_2[z/x]|_\rho = |(e_1 e_2)[z/x]|_\rho$. We have proved the case when $d(e) = 0$ by structural induction on e.

When $d(e) \geq 1$, we have that e is either $e_1 e_2$ with $d(e_1), d(e_2) \leq d(e)$, or $\lambda x.\, e_1$ with $d(e_1) \leq d(e) - 1$. We do structural induction on e. When e is $e_1 e_2$, the proof is the same as above for the case $d(e) = 0$, and thus we omit it here. When e is $\lambda y.\, e_1$ for fresh y, we have that $|(\lambda y.\, e_1)[z/x]|_\rho = |\lambda y.\, (e_1[z/x])|_\rho = |\mathsf{lambda}\ (\mathsf{intension}\ \exists y\colon Var.\ \langle y, e_1[z/x]\rangle)|_\rho = \{\#\mathsf{lam}\} \bullet_{pe} |\mathsf{intension}\ \exists y\colon Var.\ \langle y, e_1[z/x]\rangle|_\rho$

$$
\begin{aligned}
&= \{\#\mathsf{lam}\} \bullet_{pe} \left\{ |\exists y\colon Var.\ \langle y, e_1[z/x]\rangle|_\rho \right\} \\
&= \{\#\mathsf{lam}\} \bullet_{pe} \left\{ \textstyle\bigcup_{[w]\in[V^\lambda]} |\langle y, e_1[z/x]\rangle|_{\rho[[w]/y]} \right\} \\
&= \{\#\mathsf{lam}\} \bullet_{pe} \left\{ \textstyle\bigcup_{[w]\in[V^\lambda]} \{\#\mathsf{pair}\} \bullet_{pe} \{[w]\} \bullet_{pe} |e_1[z/x]|_{\rho[[w]/y]} \right\} \\
&= \{\#\mathsf{lam}\} \bullet_{pe} \left\{ \textstyle\bigcup_{[w]\in[V^\lambda]} \{\#\mathsf{pair}\} \bullet_{pe} \{[w]\} \bullet_{pe} |e_1|_{\rho[[w]/y][\rho(z)/x]} \right\} \\
&= \{\#\mathsf{lam}\} \bullet_{pe} \left\{ \textstyle\bigcup_{[w]\in[V^\lambda]} \{\#\mathsf{pair}\} \bullet_{pe} \{[w]\} \bullet_{pe} |e_1|_{\rho[\rho(z)/x][[w]/y]} \right\} \\
&= \{\#\mathsf{lam}\} \bullet_{pe} \left\{ |\exists y\colon Var.\ \langle y, e_1\rangle|_{\rho[\rho(z)/x]} \right\} = |\mathsf{lambda}\ (\mathsf{intension}\ \exists y\colon Var.\ \langle y, e_1\rangle)|_{\rho[\rho(z)/x]} = |\lambda y.\, e_1|_{\rho[\rho(z)/x]}\,.
\end{aligned}
$$

□

E Technical Details and Proofs for Section 9

We remind the reader that in this paper, we do not consider TGL in its most general form. We are considering TGL instances, where the term syntax of the binder syntax as defined in Definition 44. For language simplicity, we will call these TGL instances simply TGL.

E.1 The Gentzen-Style Proof System of TGL and the Proof of Theorem 48

We show in Fig. 8 the Gentzen-style proof system of TGL [77, Figs. 1-2]. We add additionally the last rule (Binder) to specify that equality is a congruence relation on binders. Note that the substitution rule (Sbs) implies that equality is a congruence relation on functions.

Now we prove Theorem 48.

Proof. The proof is the same to the proof of [77, Theorem 3.2], by noting that the new rule (Binder) can be simulated by adding axioms of the form:

$$\forall x.\, \forall\, \mathrm{FV}(t, t').\, t = t' \to b(x, t) = b(x, t').$$

□

E.2 Construction of the ML Theory Γ^{TGL}

Let us fix a binder syntax (S, V, F, B) like the one in Definition 44.

Definition 60. Let the ML signature Σ^{TGL} contain the following symbols:

1. $\lceil \_ \rceil$, the definedness symbol (Definition 13);
2. $[\![\_]\!]$, the inhabitant symbol (Definition 17);
3. $\langle \_, \_ \rangle$, the pairing symbol (Definition 18);
4. $\mathsf{extension}$, the extension symbol (Definition 20);
5. s, a sort constant for every $s \in S$ (Section 4.2);
6. $s^2, 2^{s^2}$, the square and power sorts of s, for every $s \in S$ (Section 4.2);
7. f for every $f \in F$;
8. π for every $\pi \in \Pi$;
9. $\mathsf{restraction}_b$, the retraction symbol for every binder $b \in B$ (Definition 44).

Let the ML theory Γ^{TGL} contain the infrastructure axioms about sorts, product sorts, power sorts, the functional axioms of $f \in F$, and predicate axioms of $\pi \in \Pi$, the partial function axioms of $\mathsf{restraction}_b$ for each $b \in B$, and the following *functional binder axioms*:

$$\text{(Functional Binder)} \quad \forall x{:}s.\, \forall \mathrm{FV}(e).\, \exists y.\, b(x, e) = y$$

The purpose of (Functional Binder) is to enforce that in any model $M \vDash \Gamma^{\mathsf{TGL}}$, the interpretations of binders $b(x, e)$ is a singleton. Therefore, all TGL terms are functional ML patterns.

E.3 Construction of the ML model M^A from a TGL model A

Let us fix a TGL model $(\{A_s\}_{s\in S}, \{A_e\}_{e\in TGLTerm}, \{A_\pi\}_{\pi\in\Pi})$. We show how to construct a corresponding ML model M^A of Σ^{TGL} below. For notational simplicity, we will omit the superscript and write M for M^A throughout this section.

We first prove a lemma about the TGL model A.

Lemma 61. *Let $x_1, \ldots, x_n, x'_1, \ldots, x'_n$ be variables and ρ, ρ' be valuations. If $\rho(x_1) = a_1$, …, $\rho(x_n) = a_n$, $\rho'(x'_1) = a_1$, …, $\rho'(x'_n) = a_n$, then we have $A_{f(x_1,\ldots,x_n)}(\rho) = A_{f(x'_1,\ldots,x'_n)}(\rho')$.*

Proof. Let $y_1, \ldots, y_n$ be n fresh variables. Since $f(x_1, \ldots, x_n) \equiv f(y_1, \ldots, y_n)[x_1/y_1] \cdots [x_n/y_n]$, we have that $A_{f(x_1,\ldots,x_n)}(\rho) = A_{f(y_1,\ldots,y_n)[x_1/y_1]\cdots[x_n/y_n]}(\rho) = A_{f(y_1,\ldots,y_n)}(\delta)$, where $\delta = \rho[\rho(x_1)/y_1] \cdots [\rho(x_n)/y_n] = \rho[a_1/y_1] \cdots [a_n/y_n]$, by [77, Definition 2.4]. Similarly, we have $A_{f(x'_1,\ldots,x'_n)}(\rho') = A_{f(y_1,\ldots,y_n)}(\delta')$ where $\delta' = \rho'[a_1/y_1] \cdots [a_n/y_n]$. Note that $\mathrm{FV}(f(y_1, \ldots, y_n)) = \{y_1, \ldots, y_n\}$, and $\delta\big|_{y_1,\ldots,y_n} = \delta'\big|_{y_1,\ldots,y_n}$. By [77, Lemma 2.5], we have $A_{f(y_1,\ldots,y_n)}(\delta) = A_{f(y_1,\ldots,y_n)}(\delta')$, and thus $A_{f(x_1,\ldots,x_n)}(\rho) = A_{f(x'_1,\ldots,x'_n)}(\rho')$. □

Definition 62. We define the ML model $(M, \_\bullet\_, \{\sigma_M\}_{\sigma\in\Sigma^{\mathsf{TGL}}})$ as follows. Firstly, we define M to be the *disjoint union* of the following sets:

1. $\{\#\mathsf{def}\}$, where $\#\mathsf{def}$ is a distinguished element, used to interpret the definedness symbol $\lceil \_ \rceil$;
2. $\{\#\mathsf{inh}\}$, where $\#\mathsf{inh}$ is a distinguished element, used to interpret the inhabitant symbol $[\![\_]\!]$;
3. $\{\#\mathsf{s}, \#\mathsf{sos}', \#\mathsf{2sos}'\}$, each interpreting the sorts $s, s \otimes s', 2^{s\otimes s'}$, for every $s, s' \in S$;

4. A_s, for every $s \in S$;
5. $A_s \times A_{s'}$, for every $s, s' \in S$;
6. $\mathcal{P}(A_s \times A_{s'})$, for every $s, s' \in S$;
7. $\{\#\mathsf{pair}, \#\mathsf{ext}, \#\mathsf{f}, \#\mathsf{pi}, \#\mathsf{resb}\}$, each interpreting $\langle \_, \_ \rangle, \mathsf{extension}, f, \pi, \mathsf{restraction}_b$;
8. $\{\#\mathsf{pair} \rightsquigarrow a \mid a \in \bigcup_s A_s\}$;
9. $\{\#\mathsf{f} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_k \mid f \in F_{s_1 \cdots s_k \cdots s_n, s}, 1 \leq k < n, a_1 \in A_{s_1}, \ldots, a_k \in A_{s_k}\}$;
10. $\{\#\mathsf{pi} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_k \mid \pi \in \Pi_{s_1 \cdots s_k \cdots s_n}, 1 \leq k < n, a_1 \in A_{s_1}, \ldots, a_k \in A_{s_k}\}$.

As we have seen in Definitions 57 and 59, Items 8-10 include the results of partial evaluations.

Secondly, we define the interpretation of application $\_ \bullet \_ : M \times M \to \mathcal{P}(M)$ as follows:

1. $\#\mathsf{def} \bullet a = T$ for every $a \in T$;
2. $\#\mathsf{inh} \bullet \#\mathsf{s} = A_s$, for every $s \in S$;
3. $\#\mathsf{inh} \bullet \#\mathsf{sos}' = A_s \times A_{s'}$, for every $s, s' \in S$;
4. $\#\mathsf{inh} \bullet \#\mathsf{2sos}' = \mathcal{P}(A_s \times A_{s'})$, for every $s, s' \in S$;
5. $\#\mathsf{pair} \bullet a = \{\#\mathsf{pair} \rightsquigarrow a\}$ for every $a \in \bigcup_s A_s$;
6. $(\#\mathsf{pair} \rightsquigarrow a) \bullet a' = \{(a, a')\}$ for every $a, a' \in \bigcup_s A_s$;
7. $\#\mathsf{ext} \bullet P = P$ for every $P \in \mathcal{P}(A_s \times A_{s'})$;
8. $(\#\mathsf{f} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_{k-1}) \bullet a_k = \{\#\mathsf{f} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_{k-1} \rightsquigarrow a_k\}$ for every $f \in F_{s_1 \cdots s_n, s}$, $1 \leq k < n$, $a_1 \in A_{s_1}, \ldots, a_k \in A_{s_k}$;
9. $(\#\mathsf{f} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_{n-1}) \bullet a_n = \{A_{f(x_1,\ldots,x_n)}(\rho)\}$ where $\rho(x_1) = a_1, \ldots, \rho(x_n) = a_n$, for every $f \in F_{s_1 \cdots s_n, s}$, $a_1 \in A_{s_1}, \ldots, a_n \in A_{s_n}$; we should verify that the choices of $x_1, \ldots, x_n$ and ρ do not matter; see Lemma 61;
10. $(\#\mathsf{pi} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_{k-1}) \bullet a_k = \{\#\mathsf{pi} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_{k-1} \rightsquigarrow a_k\}$ for every $\pi \in \Pi_{s_1 \cdots s_n}$, $1 \leq k < n$, $a_1 \in A_{s_1}, \ldots, a_k \in A_{s_k}$;
11. $(\#\mathsf{pi} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_{n-1}) \bullet a_n = M$, for every $\pi \in \Pi_{s_1 \cdots s_n}$, $a_1 \in A_{s_1}, \ldots, a_n \in A_{s_n}$, such that $(a_1, \ldots, a_n) \in A_\pi$;
12. $(\#\mathsf{pi} \rightsquigarrow a_1 \rightsquigarrow \cdots \rightsquigarrow a_{n-1}) \bullet a_n = \emptyset$, for every $\pi \in \Pi_{s_1 \cdots s_n}$, $a_1 \in A_{s_1}, \ldots, a_n \in A_{s_n}$, such that $(a_1, \ldots, a_n) \notin A_\pi$;
13. $\#\mathsf{resb} \bullet P = \{A_{b(x,t)}(\rho)\}$, if $P \subseteq A_s \times A_{s'}$, $b \in B_{s,s',r}$, and there exists $\mathcal{F} \colon M_s \to M_{s'}$ defined as $\mathcal{F}(a) = A_t(\rho[a/x])$ for every $a \in M_s$ such that $P = \mathsf{graph}(\mathcal{F})$; the well-definedness is proved in Lemma 63;
14. Otherwise, if none of the above rules applies, $a \bullet b = \emptyset$ for $a, b \in T$.

Thirdly, we define symbol interpretations as follows:

1. $\lceil \_ \rceil_M = \{\#\mathsf{def}\}$;
2. $[\![\_]\!]_M = \{\#\mathsf{inh}\}$;
3. $\langle \_, \_ \rangle_M = \{\#\mathsf{pair}\}$;

4. $\mathsf{extension}_M = \{\#\mathsf{ext}\}$;
5. $s_M = \{\#\mathsf{s}\}$;
6. $(s \otimes s')_M = \{\#\mathsf{sos}'\}$;
7. $(2^{s \otimes s'})_M = \{\#\mathsf{2sos}'\}$;
8. $f_M = \{\#\mathsf{f}\}$;
9. $\pi_M = \{\#\mathsf{pi}\}$;
10. $(\mathsf{restraction}_b)_M = \{\#\mathsf{resb}\}$.

Now we finish the construction of M.

Lemma 63. *The application interpretation $\#\mathsf{resb} \bullet P$ given in Definition 62 is well-defined.*

Proof. We need to show that the choice of $b(x,t)$ and ρ does not matter. Therefore, let us assume there are x, t, ρ and x', t', ρ' such that they yield the same function $\mathcal{F}$, i.e.:

$$A_t(\rho[a/x]) = A_{t'}(\rho'[a/x']), \text{ for all } a \in A_s.$$

Our goal is to prove that $A_{b(x,t)}(\rho) = A_{b(x',t')}(\rho')$.

Let y be a fresh variable. We enumerate all the free variables in $b(x,t)$ as $\mathrm{FV}(b(x,t)) = \{z_1, \ldots, z_m\}$. Let $z_1'', \ldots, z_m''$ be fresh variables, and we define $t'' = t[y/x][z_1''/z_1] \cdots [z_m''/z_m]$. Clearly, We have $\mathrm{FV}(t'') \subseteq \{y, z_1'', \ldots, z_m''\}$. Similarly, we enumerate $\mathrm{FV}(b(x',t')) = \{z_1', \ldots, z_m'\}$. Let $z_1''', \ldots, z_{m'}'''$ be fresh variables, and define $t''' = t'[y/x][z_1'''/z_1'] \cdots [z_{m'}'''/z_{m'}']$. We have $\mathrm{FV}(t''') \subseteq \{y, z_1''', \ldots, z_{m'}'''\}$.

Let us consider valuation ρ^*, such that $\rho^*(z_1'') = \rho(z_1)$, $\ldots$, $\rho^*(z_1'') = \rho(z_1)$, $\rho^*(z_1''') = \rho'(z_1')$, $\ldots$, $\rho^*(z_{m'}''') = \rho'(z_{m'}')$. By [77, Lemma 2.5], we have $A_{t_i}(\rho[a/x]) = A_{t_i''}(\rho^*[a/y])$ and $A_{t_i'}(\rho[a/x]) = A_{t_i'''}(\rho^*[a/y])$ for every $a \in A_s$. By [77, Definition 2.4], we have $\rho^* \in A_{\forall y:\, s.\, t''=t'''}$. Recall that our goal is to prove $A_{b(x,t)}(\rho) = A_{b(x',t')}(\rho')$. By [77, Lemma 2.5], we need to prove $A_{b(y,t'')}(\rho^*) = A_{b(y,t''')}(\rho^*)$, i.e., to prove $\rho^* \in A_{b(y,t'')=b(y,t''')}$, which holds by the proof rule (Binder). □

E.4 Proof of Theorem 49

Let us first prove the following lemma:

Lemma 64. *For any $\rho \in TGLVal$, $t \in T$, and $\varphi \in TGLForm$, we have that $|t|_\rho = \{A_t(\rho)\}$, and*

$$|\varphi|_\rho = \begin{cases} M & \text{if } \rho \in A_\varphi \\ \emptyset & \text{if } \rho \notin A_\varphi \end{cases}$$

Proof. The proof is by structural induction on t and φ. Let us first prove that $|t|_\rho = \{A_t(\rho)\}$.

When t is a variable x, we have $|x|_\rho = \{\rho(x)\} = \{A_x(\rho)\}$.

When t has the form $f(t_1, \ldots, t_n)$ for $f \in F_{s_1 \cdots s_n, s}$, we have $|f(t_1, \ldots, t_n)|_\rho = \{\#\mathsf{f}\} \bullet_{pe} |t_1|_\rho \bullet_{pe} \cdots \bullet_{pe} |t_n|_\rho = \{\#\mathsf{f}\} \bullet_{pe} \{A_{t_1}(\rho)\} \bullet_{pe} \cdots \bullet_{pe} \{A_{t_n}(\rho)\} = \#\mathsf{f} \bullet A_{t_1}(\rho) \bullet \cdots \bullet A_{t_n}(\rho) = \{A_{f(x_1,\ldots,x_n)}(\rho')\}$, where $\rho'(x_1) = A_{t_1}(\rho)$, $\ldots$, $\rho'(x_n) = A_{t_n}(\rho)$. By [77, Lemma 2.6], we have $\{A_{f(x_1,\ldots,x_n)}(\rho')\} = \{A_{f(t_1,\ldots,t_n)}(\rho)\}$.

When t has the form $b(x, t_1)$ for $b \in B_{s,s',r}$, we have that $|b(x,t_1)|_\rho = |\mathsf{restraction}_b\, (\mathsf{intension}\, \exists x{:}\, s.\, \langle x, t_1\rangle)|_\rho = \{\#\mathsf{resb}\} \bullet_{pe} |\mathsf{intension}\, \exists x{:}\, s.\, \langle x, t_1\rangle|_\rho = \{\#\mathsf{resb}\} \bullet_{pe} \{|\exists x{:}\, s.\, \langle x, t_1\rangle|_\rho\} = \#\mathsf{resb} \bullet |\exists x{:}\, s.\, \langle x, t_1\rangle|_\rho = \#\mathsf{resb} \bullet \bigcup_{a \in A_s} \{\#\mathsf{pair}\} \bullet_{pe} \{a\} \bullet_{pe} |t_1|_{\rho[a/x]} = \#\mathsf{resb} \bullet \bigcup_{a \in A_s} \{\#\mathsf{pair}\} \bullet_{pe} \{a\} \bullet_{pe} \{A_{t_1}(\rho[a/x])\} = \#\mathsf{resb} \bullet \bigcup_{a \in A_s} \{(a, A_{t_1}(\rho[a/x]))\} = \{A_{b(x,t_1)}(\rho)\}$, by Definition 62.

Now we prove the conclusion about $|\varphi|_\rho$ by structural induction. For notational simplicity, let us define the "indicator operator" $\mathbb{I}$ such that $\mathbb{I}(S) = M$ if S is a valid mathematical statement and $\mathbb{I}(S) = \emptyset$, otherwise. Then, our goal is to prove that $|\varphi|_\rho = \mathbb{I}(\rho \in A_\varphi)$.

When φ is $\pi(t_1, \ldots, t_n)$, we have $|\pi(t_1, \ldots, t_n)|_\rho = \{\#\mathsf{pi}\} \bullet_{pe} |t_1|_\rho \bullet_{pe} \cdots \bullet_{pe} |t_n|_\rho = \{\#\mathsf{pi}\} \bullet_{pe} \{A_{t_1}(\rho)\} \bullet_{pe} \cdots \bullet_{pe} \{A_{t_n}(\rho)\} = \#\mathsf{pi} \bullet A_{t_1}(\rho) \bullet \cdots \bullet A_{t_n}(\rho) = \mathbb{I}((A_{t_1}(\rho), \ldots, A_{t_n}(\rho)) \in A_\pi)$, by Definition 62. Also note that $(A_{t_1}(\rho), \ldots, A_{t_n}(\rho)) \in A_\pi$ iff $\rho \in A_{\pi(t_1,\ldots,t_n)}(\rho)$.

When φ is $t = t'$, we have $|t = t'|_\rho = \mathbb{I}(|t|_\rho = |t'|_\rho)$, by Proposition 14. Then, we have $|t|_\rho = |t'|_\rho$, iff $A_t(\rho) = A_{t'}(\rho)$, iff $\rho \in A_{t=t'}(\rho)$.

When φ is $\varphi_1 \wedge \varphi_2$, we have $|\varphi_1 \wedge \varphi_2|_\rho = |\varphi_1|_\rho \cap |\varphi_2|_\rho = \mathbb{I}(\rho \in A_{\varphi_1}) \cap \mathbb{I}(\rho \in A_{\varphi_2}) = \mathbb{I}(\rho \in A_{\varphi_1}$ and $\rho \in A_{\varphi_2}) = \mathbb{I}(\rho \in A_{\varphi_1 \wedge \varphi_2})$.

When φ is $\neg\varphi_1$, we have $|\neg\varphi_1|_\rho = M \setminus |\varphi_1|_\rho = M \setminus \mathbb{I}(\rho \in A_{\varphi_1}) = \mathbb{I}(\rho \notin A_{\varphi_1}) = \mathbb{I}(\rho \in A_{\neg\varphi_1})$.

When φ is $\forall x.\, \varphi_1$ where x has sort s, we have $|\forall x.\, \varphi_1|_\rho = \bigcap_{a \in A_s} |\varphi_1|_{\rho[a/x]} = \bigcap_{a \in A_s} \mathbb{I}(\rho[a/x] \in A_{\varphi_1}) = \mathbb{I}($for all $a \in A_s$, $\rho[a/x] \in A_{\varphi_1}) = \mathbb{I}(\rho \in A_{\forall x.\, \varphi_1})$. □

Now we prove Theorem 49.

Proof. We will prove that (3) $\Longrightarrow$ (4) $\Longrightarrow$ (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3). The only nontrivial case is (2) $\Longrightarrow$ (3). Indeed, (3) $\Longrightarrow$ (4) is by Theorem 48. (4) $\Longrightarrow$ (1) is by Proposition 23 and noting that the TGL proof system is identical to the FOL proof system. (1) $\Longrightarrow$ (2) is by Theorem 24.

To prove (2) $\Longrightarrow$ (3), we assume the opposite. Then, there exists a TGL model A such that $A \vDash E$, but $\bigcap_{\varphi \in \Delta_1} A_\varphi \not\subseteq \bigcup_{\varphi \in \Delta_2} A_\varphi$. By Definition 46, we have $\bigcap_{\varphi \in \Delta_1} A_\varphi = A_{\bigwedge \Delta_1}$ and $\bigcup_{\varphi \in \Delta_2} A_\varphi = A_{\bigvee \Delta_2}$, and thus $\rho \notin A_{\bigwedge \Delta_1 \to \bigvee \Delta_2}$. By Lemma 64, we know that for the ML model $M \vDash \Gamma^{\mathsf{TGL}}$ as defined in Definition 62, we have $M \vDash E$ and $M \nvDash \bigwedge \Delta_1 \to \bigvee \Delta_2$, which is a contradiction. Therefore, we prove that (2) $\Longrightarrow$ (3). □

(Ax)	$\dfrac{\cdot}{\Delta_1 \rhd \Delta_2}$ if $\Delta_1 \cap \Delta_2 \neq \emptyset$
(Left$\to$)	$\dfrac{\Delta_1 \rhd \Delta_2, \varphi_1 \quad \Delta_1, \varphi_2 \rhd \Delta_2}{\Delta_1, (\varphi_1 \to \varphi_2) \rhd \Delta_2}$
(Right$\to$)	$\dfrac{\Delta_1, \varphi \rhd \Delta_2, \varphi_2}{\Delta_1 \rhd \Delta_2, (\varphi_1 \to \varphi_2)}$
(Left$\wedge$)	$\dfrac{\Delta_1, \varphi_1, \varphi_2 \rhd \Delta_2}{\Delta_1, (\varphi_1 \wedge \varphi_2) \rhd \Delta_2}$
(Right$\wedge$)	$\dfrac{\Delta_1 \rhd \Delta_2, \varphi_1 \quad \Delta_1 \rhd \Delta_2, \varphi_2}{\Delta_1 \rhd \Delta_2, (\varphi_1 \wedge \varphi_2)}$
(Left$\forall$)	$\dfrac{\Delta_1, \forall x.\, \varphi, \varphi[t/x] \rhd \Delta_2}{\Delta_1, \forall x.\, \varphi \rhd \Delta_2}$
(Right$\forall$)	$\dfrac{\Delta_1 \rhd \Delta_2, \varphi[y/x]}{\Delta_1 \rhd \Delta_2, \forall x.\, \varphi}$ if y fresh
(Reflexivity)	$\dfrac{\Delta_1, t = t \rhd \Delta_2}{\Delta_1 \rhd \Delta_2}$
(Symmetry)	$\dfrac{\Delta_1 \rhd \Delta_2, t_1 = t_2 \quad \Delta_1, t_2 = t_1 \rhd \Delta_2}{\Delta_1 \rhd \Delta_2}$
(Transitivity)	$\dfrac{\Delta_1 \rhd \Delta_2, t_1 = t_2 \quad \Delta_1 \rhd \Delta_2, t_2 = t_3 \quad \Delta_1, t_1 = t_3 \rhd \Delta_2}{\Delta_1 \rhd \Delta_2}$
(Cmp$_\pi$)	$\dfrac{\Delta_1 \rhd \Delta_2, t_i = t_i' \quad \Delta_1 \rhd \Delta_2, \pi(t_1, \ldots, t_n) \quad \Delta_1, \pi(t_1', \ldots, t_n') \rhd \Delta_2}{\Delta_1 \rhd \Delta_2}$ for every $i \in \{1, \ldots, n\}$
(Sbs)	$\dfrac{\Delta_1 \rhd \Delta_2, t_1 = t_2 \quad \Delta_1, t[t_1/x] = t[t_2/x] \rhd \Delta_2}{\Delta_1 \rhd \Delta_2}$
(Binder)	$\dfrac{\Delta_1 \rhd \Delta_2, t = t' \quad \Delta_1, b(x,t) = b(x,t') \rhd \Delta_2}{\Delta_1 \rhd \Delta_2}$

Figure 8: The Gentzen-style proof system of TGL [77, Figs. 1-2], plus one congruence rule (Binder) for binders