A simple class of Kripke-style models in which logic and computation have equal standing*

Michael Gabbay[1]** and Murdoch J. Gabbay[2]

[1] http://www.kcl.ac.uk/philosophy/people/fellows/mgabbay.html
[2] www.gabbay.org.uk

Abstract. We present a sound and complete model of lambda-calculus reductions based on structures inspired by modal logic (closely related to Kripke structures). Accordingly we can construct a logic which is sound and complete for the same models, and we identify lambda-terms with certain simple sentences (predicates) in this logic, by direct compositional translation. Reduction then becomes identified with logical entailment. Thus, the models suggest a new way to identify logic and computation. Both have elementary and concrete representations in our models; where these representations overlap, they coincide.
In a concluding speculation, we note a certain subclass of the models which seems to play a role analogous to that played by the cumulative hierarchy models in axiomatic set theory and the natural numbers in formal arithmetic — there are many models of the respective theories, but only some, characterised by a fully second order interpretation, are the 'intended' ones.

1 Introduction

We try to unify logic and computation by using a class of structures which are (very nearly) Kripke structures. It turns out that these structures allow sound and complete interpretations of both a logic (an extension of second-order propositional logic), and computation (the untyped λ-calculus). Furthermore, we are able to compositionally translate λ-terms and formulae into our logic, and when we do so, the 'computational' reduction $\twoheadrightarrow$ maps to logical entailment, and λ maps to a kind of logical quantifier.

Combining logic and computation is of course not a new idea. The two notions are clearly related and intertwined, and there are good theoretical and practical reasons to be interested in these questions.

A naive combination of logic and computation can lead to some famous contradictions. Consider untyped λ-calculus quotiented by computational equivalence, e.g. β-equivalence. Suppose also the naive addition of some basic logical

* The authors are grateful to the anonymous referees for their comments
** Supported by British Academy grant PDF/2006/509, IFCoLog Hilbert-Bernays Project, and grant RYC-2006-002131 at the Polytechnic University of Madrid.

operations, e.g. negation $\neg$. Then we can form a term encoding the Liar paradox $L = \lambda p.\neg(p{\cdot}p){\cdot}\lambda p.\neg(p{\cdot}p)$. Then $L = \neg L$ and from this a contradiction quickly follows. This is, in a nutshell, the argument in [15].[3]

We can persist with adding logic to λ-terms. This motivated the development of types, as in higher-order logic [21], where the paradox is avoided by restricting which terms can receive an interpretation. Also, with the same intuitions but different design decisions, illative combinatory logics [3,2] can admit an untyped system but restrict which λ-terms count as 'logical'.

Conversely, we can view, e.g. the negation of a λ-term as a category error, and think of computation as an object or vehicle of logical systems. So for example *rewriting logic* has a logic with an oriented equality $\twoheadrightarrow$ representing reduction; whereas *deduction modulo* [4,7] has a background 'computational engine' which may be triggered to efficiently compute equalities between terms, but the logic is modulo this and reasons up to computational equivalence. In both cases the interpretations of the logical and computational parts of the languages are separated, like sentences are separated from terms in first-order languages and semantics.

The model of logic and computation of this paper is like rewriting logic in that we explicitly represent computations (reductions) by an arrow, $\rightarrow$. However, in line with the original intuition equivocating logic and computation, $\rightarrow$ is the same arrow as is used to represent logical implication.

In our models, λ-terms are interpreted as sets on a domain and $\rightarrow$ is interpreted as 'complement union'. Entailment and reduction are both therefore represented in the models as subset inclusion (see Definition 2.13). That is, this arrow $\rightarrow$ *really is* standard implication, just as the reader is used to. We discuss in Section 5.1 how this relates to the paradox above.

The kernel of the ideas in this paper is a class of models, presented in Section 2.1; the rest of the paper can be considered to arise just by considering their structure. It turns out that it is possible to consider λ-abstraction as a kind of quantifier and to consider reduction as subset inclusion (Section 2.2). The models are sets-based: The usual logical connectives such as conjunction, negation, and quantification are interpreted in the models as operations on sets; and logical entailment is interpreted as subset inclusion. We obtain an extension of classical second-order propositional logic with quantifiers (Section 2.3). We make our logic rich enough that it captures the structure we used to interpret λ-abstraction and reduction; because reduction is interpreted as subset inclusion, it maps directly to logical entailment.

The idea of modelling λ-reduction is not new. Models of reduction where terms are interpreted as points in an ordering are discussed by Selinger in [19]; and models, based on *graph models*, where terms are interpreted as sets on a domain of functions, are given in [16]. The models in Section 2 have similarities with these. One significant difference is that our models have a *Boolean structure*, in

[3] Historically, Church wanted to base maths on the notion of function as opposed to a notion of set. The λ-calculus was invented as a foundation for logic, not a foundation for computation. [15] proved that this foundation was logically inconsistent.

particular every denotation has a complement and a universal quantification. So the denotational domains for λ-terms also display the structure of denotational domains of propositions of classical logic.

So: logic and computation are not identical in this story (we do not claim that logic *is* computation, or vice versa) — but the two notions overlap in the models, and in this overlap, they coincide.

One possible use for our models is that they might provide a relatively elementary and systematic framework for extensions of the pure λ-calculus with 'logical' constructs, and indeed the design of other logics. We do not consider that there is anything sacred about the particular languages we use in this paper. However, they do have the virtues of being simple and well-behaved. In particular we can give tight soundness and completeness results for both the logic and the λ-calculus (Theorems 2.27 and 4.13).

The structure and main results of this paper are as follows: We develop a simple model theory and the syntax it motivates in Section 2; in Section 3 we verify that we can faithfully translate a system of λ-reduction into this syntax (and thus obtain a new model theory for λ-reduction); in Section 4 we prove completeness for an axiomatisation of the model theory. We conclude with some comments on the significance of the models presented here both for current and further research.

2 The models, computation, logic

2.1 Frames

Definition 2.1 If W is a set, write $\mathcal{P}(W)$ for the set of subsets of W.

A **frame** F is a 3-tuple $(W, \bullet, H)$ of

- W a set of **worlds**,
- $\bullet$ an **application function** from $W \times W$ to $\mathcal{P}(W)$, and
- $H \subseteq \mathcal{P}(W)$.

Remark 2.2 Frames are not monoids or any other form of applicative structure — an applicative structure would map $W \times W$ to W and not, as we do here, to $\mathcal{P}(W)$. One reasonable way to think of $\bullet$ is as a non-deterministic 'application' operation, although we suggest a better way in the concluding section 5.2 (where we also discuss in more detail the differences between frames and other known structures that 'look' like frames).

Subsets of W will serve as denotations of sentences (Definitions 2.7 and 2.13). We can interpret both computational and logical connectives as elementary operations on sets of worlds (e.g. we interpret logical conjunction as sets intersection).

Remark 2.3 $H \subseteq \mathcal{P}(W)$ ('H' for 'Henkin') plays a similar role to the structure of Henkin models for higher-order logic [1,12,20]. This makes our completeness results possible and is a famous issue for second- and higher-order logics. Powersets are too large; for completeness results to be possible we must cut them down

— at least when we quantify. This is why in Definitions 2.7 and 2.13, the binders restrict quantification from $\mathcal{P}(W)$ down to H. More on this in the concluding section 5.2.

The reader familiar with modal logic can think of $\bullet$ as a ternary 'accessibility relation' R such that $Rw_1w_2w_3$ if and only if $w_3 \in w_1 \bullet w_2$. We can also think of $\bullet$ is as a non-deterministic 'application' operation, but note that frames are not applicative structures — an applicative structure would map $W \times W$ to W, whereas in the case of frames, $W \times W$ maps to $\mathcal{P}(W)$. However, $\bullet$ does induce an applicative structure on $\mathcal{P}(W)$:

Definition 2.4 Suppose $F = (W, \bullet, H)$ is a frame. Suppose $S_1, S_2 \subseteq W$ and $w \in W$.

The function $\bullet$ induces functions from $W \times \mathcal{P}(W)$ and $\mathcal{P}(W) \times \mathcal{P}(W)$ to $\mathcal{P}(W)$ by:

$$w \bullet S_1 = \bigcup\{w \bullet w' \mid w' \in S_1\} \qquad S_1 \bullet S_2 = \bigcup\{w_1 \bullet w_2 \mid w_1 \in S_1,\ w_2 \in S_2\}$$

2.2 λ-terms

Definition 2.5 Fix a countably infinite set of **variables**. p, q, r will range over distinct variables (we call this a *permutative* convention)

Define a language $\mathcal{L}_\lambda$ of **λ-terms** by:

$$\mathtt{t} ::= p \mid \lambda p.\mathtt{t} \mid \mathtt{t}\cdot\mathtt{t}$$

λp binds in $\lambda p.\mathtt{t}$. For example, p is bound (not free) in $\lambda p.p\cdot q$. We identify terms up to α-equivalence.

We write $\mathtt{t}[p::=\mathtt{s}]$ for the usual capture-avoiding substitution. For example, $(\lambda p'.q)[q::=p] = \lambda p'.p$, and $(\lambda p.q)[q::=p] = \lambda p'.p$ where p' is a fixed but arbitrary choice of fresh variable.

Definition 2.6 Suppose $F = (W, \bullet, H)$ is a frame. A **valuation** (to F) is a map from variables to sets of worlds (elements of $\mathcal{P}(W)$). v will range over valuations.

If p is a variable, $S \subseteq W$, and v is a valuation, then write $v[p::=S]$ for the valuation mapping q to $v(q)$ and mapping p to S.

Definition 2.7 Define an **denotation** of $\mathtt{t}$ inductively by:

$$[\![p]\!]^v = v(p) \qquad [\![\mathtt{t}\cdot\mathtt{s}]\!]^v = [\![\mathtt{t}]\!]^v \bullet [\![\mathtt{s}]\!]^v$$

$$[\![\lambda p.\mathtt{t}]\!]^v = \{w \mid w \bullet h \subseteq [\![\mathtt{t}]\!]^{v[p::=h]} \text{ for all } h \in H\}$$

Reduction on terms is defined in Figure 1 on page 9.

Remark 2.8 We will be particularly interested in models where the denotation of every λ-term is a member of H. This is because Definition 2.7 interprets λ as a kind of quantifier over all members of H. β-reduction then becomes a form of universal instantiation and so requires that all possible instantiations (i.e. the denotation of any term) is a member of H. More on this in Section 5.2.

Lemma 2.9 *β-reduction and η-expansion are sound, if we interpret reduction as subset inclusion:*

$$\begin{array}{lrl} \beta\text{-reduction} & [\![\lambda p.\mathtt{t}]\!]^v \bullet [\![\mathtt{s}]\!]^v \subseteq [\![\mathtt{t}[p{::=}\mathtt{s}]]\!]^v & (\text{if } [\![\mathtt{s}]\!]^v \in H) \\ \eta\text{-expansion} & [\![\mathtt{t}]\!]^v \subseteq [\![\lambda p.(\mathtt{t}\cdot p)]\!]^v & (p \text{ not free in } \mathtt{t}) \end{array}$$

Proof. By routine calculations from the definitions. We prove a more general result in Theorem 2.23.

Remark 2.10 It may help to give some indication of what the *canonical* frame used in the completeness proof for $\mathcal{L}_\lambda$ (Definition 3.5) looks like: worlds are β-reduction-η-expansion closures of λ-terms $\mathtt{t}$, and for each $h \in H$ there exists some $\mathtt{t}$ such that h is the set of worlds that contain $\mathtt{t}$.

As we emphasised in Remark 2.2, our frames are not applicative structures, and the denotations of λ-terms are not worlds, but sets of worlds. Thus, in the canonical frame, the denotation of a λ-term $\mathtt{t}$ is *not* the set of its reducts (i.e. not some world in the canonical frame). Rather, the denotation of $\mathtt{t}$ is the set of all worlds that *reduce* to $\mathtt{t}$.

We can identify a world with its 'top' term, so roughly speaking, in the canonical model a world $w \in W$ is a term $\mathtt{t}$, and an $h \in H$ (or any denotation) is a set of all terms which reduce to some particular term $\mathtt{s}$.

Remark 2.11 We suggest an intuition why our models 'have to' satisfy β-reduction and η-expansion. Both β-reduction and η-expansion *lose information*: in the case of β we perform the substitution as is usual; in the case of η-expansion $\lambda p.(\mathtt{t}\cdot p)$ has lost any intensional information that might reside in $\mathtt{t}$. So we consider validating η-expansion as an interesting feature, and not necessarily a bug.

Others have also noted good properties and justification in models for η-expansion [14]. It *is* possible to refine the models to eliminate η-expansion, at some cost in complexity; see the Conclusions.

We will fill in more details of the semantics of λ-terms in Section 2.4, including the role of H, once we have built the logic in Section 2.3.

2.3 The logic

Definition 2.12 Define a language $\mathcal{L}$ with **sentences** ϕ by:

$$\phi ::= p, q, r \ldots \mid \phi \to \phi \mid \forall p.\, \phi \mid \Box\phi \mid \phi\cdot\phi \mid \phi \triangleright \phi \mid \bot$$

$\forall p$ binds in $\forall p.\, \phi$. For example, p is bound in $\forall p.\, (p\cdot q)$. We identify sentences up to α-equivalence.

We now give notions of logical entailment and denotation for $\mathcal{L}$. In Section 2.4 we discuss expressive power and in Sections 3 and 4 we sketch proofs of soundness and completeness.

Definition 2.13 Suppose $F = (W, \bullet, H)$ is a frame and v is a valuation to F. Define $[\![\phi]\!]^v$ the **denotation** of ϕ by:

$$\begin{aligned} [\![p]\!]^v &= v(p) & [\![\bot]\!]^v &= \varnothing \\ [\![\phi{\cdot}\psi]\!]^v &= [\![\phi]\!]^v \bullet [\![\psi]\!]^v & [\![\phi \rhd \psi]\!]^v &= \{w \mid w \bullet [\![\phi]\!]^v \subseteq [\![\psi]\!]^v\} \\ [\![\phi \to \psi]\!]^v &= (W \setminus [\![\phi]\!]^v) \cup [\![\psi]\!]^v & [\![\forall p.\, \phi]\!]^v &= \textstyle\bigcap_{h \in H} [\![\phi]\!]^{v[p \mapsto h]} \end{aligned}$$

$$[\![\Box\phi]\!]^v = \begin{cases} W & [\![\phi]\!]^v = W \\ \varnothing & [\![\phi]\!]^v \neq W \end{cases}$$

Remark 2.14 Intuitions are as follows:

– p, q, r, are variables ranging over subsets of W.

– $\phi \to \psi$ is classical implication.

– $\forall p.\, \phi$ is a quantification over elements of H. Think of $\forall p.\, \phi$ as 'the intersection of the denotation of ϕ for all of a pre-selection of possible denotations of p'. The possible denotation of p are subsets of W and not elements of the W; pre-selection is done by H, which identifies those denotations that 'exist' in the sense of being in the ranges of the quantifiers. More on this later.

– $\Box\phi$ is a notion of *necessity*. $\Box\phi$ is either W or $\varnothing$ depending on whether ϕ is itself W or not.
$\Box$ is the modality of S5 [11].

– $\phi{\cdot}\psi$ is a notion of application; the construction in Definition 2.4 ensures that the interpretation of $\cdot$ it is monotonic with respect to subset inclusion.[4]
The maps $\phi{\cdot}$- and -$\cdot\psi$ behave like the box operator of the modal logic K.

– $\rhd$ is the right adjoint to $\cdot$ with respect to $\to$. It is easily verified from Definition 2.13 that $[\![\phi{\cdot}\psi]\!]^v \subseteq [\![\mu]\!]^v$ exactly when $[\![\phi]\!]^v \subseteq [\![\psi \rhd \mu]\!]^v$. So $\phi \rhd \psi$ is interpreted as the largest subset of W that when applied to ϕ, is included in ψ.

$\mathcal{L}$ is a second-order classical propositional logic enriched with the necessity modality $\Box$ from $S5$, and notions of application $\cdot$ and its right adjoint $\rhd$ (with respect to logical implication $\to$).

When we mix all these ingredients, interesting things become expressible, as we now explore.

2.4 Expressivity

Remark 2.15 We can express truth, negation, conjunction, disjunction, if-and-only-if and existential quantification as below. We also unpack Definition 2.13 to see this denotationally:

$$\begin{aligned} \top &= (\bot \to \bot) & [\![\top]\!]^v &= W \\ \neg\phi &= \phi \to \bot & [\![\neg\phi]\!]^v &= W \setminus [\![\phi]\!]^v \\ \phi \wedge \psi &= \neg(\phi \to \neg\psi) & [\![\phi \wedge \psi]\!]^v &= [\![\phi]\!]^v \cap [\![\psi]\!]^v \\ \phi \vee \psi &= (\neg\phi) \to \psi & [\![\phi \vee \psi]\!]^v &= [\![\phi]\!]^v \cup [\![\psi]\!]^v \\ \exists p.\, \phi &= \neg(\forall p.\, \neg\phi) & [\![\exists p.\, \phi]\!]^v &= \textstyle\bigcup_{h \in H} [\![\phi]\!]^{v[p \mapsto h]}. \end{aligned}$$

[4] That is, $[\![\phi]\!]^v \subseteq [\![\phi']\!]^v$ and $[\![\psi]\!]^v \subseteq [\![\psi']\!]^v$ implies $[\![\phi{\cdot}\psi]\!]^v \subseteq [\![\phi'{\cdot}\psi']\!]^v$.

Note that $\exists.$ quantifies over elements of H. This is all standard, which is the point.

Remark 2.16 For the reader familiar with the expression of product and other types in System F [10], note that $[\![\neg(\phi \to \neg\psi)]\!]^v \neq [\![\forall p.\,(\phi \to \psi \to p) \to p]\!]^v$ in general; H may be too sparse. The equality holds in a frame if $[\![\phi]\!]^v \in H$ for every ϕ. We can specify this condition as an (infinite) theory using an 'existence' predicate E (Definition 2.18).

Definition 2.17 We can express that two predicates *are* equal by: $\phi \approx \psi = \Box(\phi \leftrightarrow \psi)$.

We unpack the denotation of $\phi \approx \psi$ and for comparison also that of $\phi \leftrightarrow \psi$ $((\phi \to \psi) \wedge (\psi \to \phi))$:

$$[\![\phi \approx \psi]\!]^v = \begin{cases} W & [\![\phi]\!]^v = [\![\psi]\!]^v \\ \varnothing & [\![\phi]\!]^v \neq [\![\psi]\!]^v \end{cases}$$

$$[\![\phi \leftrightarrow \psi]\!]^v = \{w \in W \mid (w \in [\![\phi]\!]^v \wedge w \in [\![\psi]\!]^v) \vee (w \notin [\![\phi]\!]^v \wedge w \notin [\![\psi]\!]^v)\}$$

Intuitively, $\phi \leftrightarrow \psi$ holds at the worlds where ϕ and ψ are either both true or both false, whereas $\phi \approx \psi$ represents the statement 'ϕ and ψ are true of the same worlds'.

Definition 2.18 We can express that a predicate is in H by $\mathsf{E}\phi = \exists p.\,(\phi \approx p)$, read '$\phi$ exists'. This is usually called an *existence predicate* [13, Ch.16].

It is not hard to verify that $\mathsf{E}\phi$ has the following denotation:

$$[\![\mathsf{E}\phi]\!]^v = \begin{cases} W & [\![\phi]\!]^v \in H \\ \varnothing & [\![\phi]\!]^v \notin H \end{cases}$$

We are now ready to interpret λ-abstraction in our logic. We also mention a notion of matching, because it comes very naturally out of the logic as a 'dual' to the construction for λ:

Definition 2.19 $\curlywedge p.\,\phi = \forall p.\,(p \rhd \phi) \qquad \mathsf{match}\,p.\,\phi = \forall p.\,(\phi \rhd p)$.

Intuitively, $\curlywedge p.\,\phi$ reads as: 'for any p, if p is an argument (of the function instantiated at this world), then we get ϕ'. As a kind of inverse to this, $\mathsf{match}\,p.\,\phi$ reads as: 'for any p, if ϕ is an argument, then we get p'. So $\mathsf{match}\,p.\,\phi$ is a kind of pattern-matching or inverse-λ.

$\curlywedge$ is a logical quantifier, so we name it reversed by analogy with the reversed A and E of universal and existential quantification.

Theorem 2.20 – $[\![\curlywedge p.\,\phi)]\!]^v = \{w \mid w \bullet h \subseteq [\![\phi]\!]^{v[p\mapsto h]}$ *for all* $h \in H\}$
– $[\![\mathsf{match}\,p.\,\phi]\!]^v = \{w \mid w \bullet [\![\phi]\!]^{v[p\mapsto h]} \subseteq h$ *for all* $h \in H\}$

Proof.

$$\begin{aligned}
[\![\Lambda p.\,\phi]\!]^v &= [\![\forall p.\,(p \rhd \phi)]\!]^v && \text{Definition 2.19}\\
&= \textstyle\bigcap_{h\in H}[\![p \rhd \phi]\!]^{v[p\mapsto h]} && \text{Definition 2.13}\\
&= \textstyle\bigcap_{h\in H}\{w \mid w \bullet [\![p]\!]^{v[p\mapsto h]} \subseteq [\![\phi]\!]^{v[p\mapsto h]}\} && \text{Definition 2.13}\\
&= \{w \mid w \bullet h \subseteq [\![\phi]\!]^{v[p\mapsto h]} \text{ for all } h \in H\} && \text{Definition 2.4}
\end{aligned}$$

The case for $\mathsf{match}\, p.$ is similar.

Lemma 2.21 *If p is not free in ϕ, then for any $h \in H$, $[\![\phi]\!]^v = [\![\phi]\!]^{v[p\mapsto h]}$.*

Proof. An easy induction on ϕ.

Lemma 2.22 (Substitution Lemma) *For any v, $[\![\phi[p::=\psi]]\!]^v = [\![\phi]\!]^{v[p\mapsto[\![\psi]\!]^v]}$.*

Proof. By induction on ϕ, we present the cases for $\forall q$ and $\cdot$:

$$\begin{aligned}
[\![(\forall q.\,\mu)[p::=\psi]]\!]^v &= \textstyle\bigcap_{h\in H}[\![\mu[p::=\psi]]\!]^{v[q\mapsto h]} && \text{Definition 2.13}\\
&= \textstyle\bigcap_{h\in H}[\![\mu]\!]^{v[q\mapsto h, p\mapsto[\![\psi]\!]^{v[q\mapsto h]}]} && \text{Induction Hypothesis}\\
&= \textstyle\bigcap_{h\in H}[\![\mu]\!]^{v[q\mapsto h, p\mapsto[\![\psi]\!]^{v}]} && \text{Lemma 2.21}\\
&= [\![\forall q.\,\mu]\!]^{v[p\mapsto[\![\psi]\!]^v]} && \text{Definition 2.13}
\end{aligned}$$

$$\begin{aligned}
[\![(\mu_1\cdot\mu_2)[p::=\psi]]\!]^v &= [\![\mu_1[p::=\psi]]\!]^v \bullet [\![\mu_2[p::=\psi]]\!]^v && \text{Definition 2.13}\\
&= [\![\mu_1]\!]^{v[p\mapsto[\![\psi]\!]^v]} \bullet [\![\mu_2]\!]^{v[p\mapsto[\![\psi]\!]^v]} && \text{Induction Hypothesis}\\
&= [\![\mu_1\cdot\mu_2]\!]^{v[p\mapsto[\![\psi]\!]^v]} && \text{Definition 2.13}
\end{aligned}$$

Theorem 2.23 *The following hold in any frame:*

$$\begin{array}{lrcll}
\textit{(β-reduction)} & [\![(\Lambda p.\,\phi)\cdot\psi]\!]^v & \subseteq & [\![\phi[p::=\psi]]\!]^v & \textit{(for $[\![\psi]\!]^v \in H$)}\\
\textit{(η-expansion)} & [\![\phi]\!]^v & \subseteq & [\![\Lambda p.\,(\phi\cdot p)]\!]^v & \textit{(for p not free in ϕ)}\\
\textit{(matching)} & [\![(\mathsf{match}\, p.\,\phi)\cdot(\phi[p::=\psi])]\!]^v & \subseteq & [\![\psi]\!]^v & \textit{(for $[\![\psi]\!]^v \in H$)}
\end{array}$$

Proof.

$$\begin{aligned}
&[\![\forall p.\,(p \rhd \phi)\cdot\psi]\!]^v\\
&\quad= [\![\forall p.\,(p \rhd \phi)]\!]^v \bullet [\![\psi]\!]^v && \text{Definition 2.13}\\
&\quad= \textstyle\bigcap_{h\in H}\{w \mid w \bullet h \subseteq [\![\phi]\!]^{v[p\mapsto h]}\} \bullet [\![\psi]\!]^v && \text{Definition 2.13}\\
&\quad\subseteq \{w \mid w \bullet [\![\psi]\!]^v \subseteq [\![\phi]\!]^{v[p\mapsto[\![\psi]\!]^v]}\} \bullet [\![\psi]\!]^v && [\![\psi]\!]^v \in H\\
&\quad\subseteq [\![\phi]\!]^{v[p\mapsto[\![\psi]\!]^v]} && \text{Definition 2.4}\\
&\quad= [\![\phi[p::=\psi]]\!]^v && \text{Lemma 2.22}
\end{aligned}$$

$$\begin{aligned}
[\![\phi]\!]^v &\subseteq \{w \mid w \bullet S \subseteq [\![\phi]\!]^v \bullet S \text{ for any } S \in H\} && \text{Definition 2.4}\\
&= \textstyle\bigcap_{h\in H}\{w \mid w \bullet S \subseteq [\![\phi\cdot p]\!]^{v[p\mapsto h]}\} && p \text{ not free in } [\![\psi]\!]^v\\
&= [\![\Lambda p.\,(\phi\cdot p)]\!]^v && \text{Definition 2.19}
\end{aligned}$$

(matching) follows by a similarly routine calculation.

Corollary 2.24 *– If p is not free in ϕ and $[\![\bot]\!]^v \in H$ then $[\![(\mathsf{match}\, p.\,\phi)\cdot\phi]\!]^v = \varnothing$.*

$$(Eq)\ \mathtt{t} \twoheadrightarrow \mathtt{t} \qquad (\beta)\ (\boldsymbol{\lambda} p.\mathtt{t}){\cdot}\mathtt{s} \twoheadrightarrow \mathtt{t}[p{::=}\mathtt{s}] \qquad (\eta)\ \mathtt{t} \twoheadrightarrow \boldsymbol{\lambda} p.(\mathtt{t}{\cdot}p) \quad (p \text{ not free in } \mathtt{t})$$

$$(\xi)\ \frac{\mathtt{t} \twoheadrightarrow \mathtt{s}}{\boldsymbol{\lambda} p.\mathtt{t} \twoheadrightarrow \boldsymbol{\lambda} p.\mathtt{s}} \qquad (cong)\ \frac{\mathtt{t}_1 \twoheadrightarrow \mathtt{s}_1 \quad \mathtt{t}_2 \twoheadrightarrow \mathtt{s}_2}{\mathtt{t}_1{\cdot}\mathtt{t}_2 \twoheadrightarrow \mathtt{s}_1{\cdot}\mathtt{s}_2} \qquad (trans)\ \frac{\mathtt{t}_1 \twoheadrightarrow \mathtt{t}_2 \quad \mathtt{t}_2 \twoheadrightarrow \mathtt{t}_3}{\mathtt{t}_1 \twoheadrightarrow \mathtt{t}_3}$$

Fig. 1: $\boldsymbol{\lambda}$-reduction

– *If* $[\![\psi]\!]^v \in H$ *then* $[\![(\mathsf{match}\, p.\, \phi){\cdot}((\reflectbox{λ} p.\, \phi){\cdot}\psi)]\!]^v \subseteq [\![\psi]\!]^v$.
– $\mathsf{match}\, p.\, p = \reflectbox{λ} p.\, p$.

Proof. The first two parts follow easily from Theorem 2.23. The third part follows unpacking definitions, since both are equal to $\forall p.\,(p \rhd p)$.

Read $(\mathsf{match}\, p.\, \phi){\cdot}\mu$ as returning the intersection of all ψ such that μ is equivalent to $\phi[p{::=}\psi]$. If there are many such ψ, e.g. when p is not free in ϕ, then $(\mathsf{match}\, p.\, \phi){\cdot}\mu \to \psi$ for all such ψ and so $(\mathsf{match}\, p.\, \phi){\cdot}\mu$ is included in their intersection.

Definition 2.25 Define a translation τ from $\mathcal{L}_\lambda$ (Definition 2.5) to $\mathcal{L}$ (Definition 2.12) by:

$$p^\tau = p \qquad (\mathtt{t}_1{\cdot}\mathtt{t}_2)^\tau = (\mathtt{t}_1^\tau{\cdot}\mathtt{t}_2^\tau) \qquad (\boldsymbol{\lambda} p.\mathtt{t})^\tau = \reflectbox{λ} p.\, \mathtt{t}^\tau$$

Definition 2.26 Write $\mathtt{t} \twoheadrightarrow \mathtt{s}$ if $\mathtt{t} \twoheadrightarrow \mathtt{s}$ is derivable using the axioms of Figure 1.

Our implementation of λ is *sound* in the following sense:

Theorem 2.27 $\mathtt{t} \twoheadrightarrow \mathtt{s}$ *only if* $[\![\mathtt{t}^\tau \to \mathtt{s}^\tau]\!]^v = W$ *for all* v *and* $F = (W, \bullet, H)$ *such that* $[\![\mathtt{u}^\tau]\!]^v \in H$ *for all* $\mathtt{u}$.[5]

Proof. This follows (mostly) by Theorem 2.23.

3 Completeness for $\boldsymbol{\lambda}$-reduction

In this section we show that the axiomatisation of $\boldsymbol{\lambda}$-reduction of Figure 1 is complete for our interpretation of $\boldsymbol{\lambda}$-terms in terms of $\reflectbox{$\lambda$}$. We do this by proving the converse of Theorem 2.27.

To complete Theorem 2.27 we must show that if $\mathtt{t} \not\twoheadrightarrow \mathtt{s}$ then there is a frame and valuation v where $[\![\mathtt{u}^\tau]\!]^v \in H$ for all $\mathtt{u}$ and $[\![\mathtt{t}^\tau \to \mathtt{s}^\tau]\!]^v \neq W$ (where τ is the translation of Definition 2.25).

Definition 3.1 Say a $\boldsymbol{\lambda}$-term is **complex** if it contains term formers, i.e. is not simply a variable. The **size** of a $\boldsymbol{\lambda}$-term is the number of term formers within it.

[5] In other words, if $[\![\mathsf{E}(\mathtt{u}^\tau)]\!]^v = W$ for all terms $\mathtt{u}$.

Now we can begin the construction of the desired frame. First we add infinitely many new variables $r_1, r_2 \dots$ to the language $\mathcal{L}_\lambda$. Since the language is countable we can enumerate its complex terms $\mathtt{t}_1, \mathtt{t}_2 \dots$ and these new variables $r_1, r_2 \dots$. We describe a one-one function f from terms to variables.

> $f(\mathtt{t}_i) = r_j$ where j is the least number such that $j > i$ and r_j is not free in $\mathtt{t}_i$ nor is the value under f of any $\mathtt{t}_k$ for $k < i$.

Thus f is a one-one function that assigns a distinct 'fresh' variable to each complex term of the language. Thus $f(\mathtt{t})$ is a variable that 'names' $\mathtt{t}$. These play the role of witness constants in the construction of the canonical frame in Theorem 3.7. The $f(\mathtt{t})$ also help us carry out inductions on the size of λ-terms, as $\mathtt{t}[p{::=}f(\mathtt{s})]$ is smaller than $\lambda p.\mathtt{t}$ even if $\mathtt{t}[p{::=}\mathtt{s}]$ might not be.

Definition 3.2 Next we add two new axioms of reduction, denote them by (ζ_f):

$$\mathtt{t} \twoheadrightarrow f(\mathtt{t}) \qquad f(\mathtt{t}) \twoheadrightarrow \mathtt{t} \tag{ζ_f}$$

Write $\mathtt{t} \twoheadrightarrow_{\zeta_f} \mathtt{s}$ when $\mathtt{t} \twoheadrightarrow \mathtt{s}$ is derivable using the (ζ_f) rules in addition to the rules of Figure 1.

Lemma 3.3 *If $\mathtt{t} \twoheadrightarrow_{\zeta_f} \mathtt{s}$ and neither $\mathtt{s}$ or $\mathtt{t}$ contain any of the new variables $r_1, r_2 \dots$, then $\mathtt{t} \twoheadrightarrow \mathtt{s}$.*

Proof. By simultaneously substituting each instance of $f(\mathtt{t}_i)$ with $\mathtt{t}_i$ each instance of (ζ_f) becomes an instance of (Eq) without affecting the rest of the derivation.

Definition 3.4 If $\mathtt{t}$ is a term let $w_\mathtt{t} = \{\mathtt{s} \mid \mathtt{t} \twoheadrightarrow_{\zeta_f} \mathtt{s}\}$. Thus $w_\mathtt{t}$ is the closure of $\mathtt{t}$ under $\twoheadrightarrow_{\zeta_f}$.

Definition 3.5 Define the **canonical λ-frame** $F_\lambda = \langle W_\lambda, \bullet_\lambda, H_\lambda \rangle$ as follows:

- $W_\lambda = \{w_\mathtt{t} \mid \mathtt{t} \text{ is a term}\}$
- For any $w_{\mathtt{t}_1}, w_{\mathtt{t}_2} \in W$, $w_{\mathtt{t}_1} \bullet_\lambda w_{\mathtt{t}_2} = \{w \in W_\lambda \mid \mathtt{t}_1 \cdot \mathtt{t}_2 \in w\}$
- $H_\lambda = \big\{\{w \in W_\lambda \mid \mathtt{t} \in w\} \mid \ \mathtt{t} \text{ is a term}\big\}$

Definition 3.6 Given $F_\lambda = \langle W_\lambda, \bullet_\lambda, H_\lambda \rangle$ and a term $\mathtt{t}$. Let $\|\mathtt{t}\| = \{w \in W_\lambda \mid \mathtt{t} \in w\}$.

Theorem 3.7 *Let F_λ be the canonical λ-frame (Definition 3.5). Let τ be the translation from λ-terms $\mathtt{t}$ to sentences ϕ (Definition 2.25). Let $v(p) = \|p\|$ for any variable p. Then for any term $\mathtt{t}$, $[\![\mathtt{t}^\tau]\!]^v = \|\mathtt{t}\|$.*

Proof. By induction on the size of $\mathtt{t}$ we show that $w \in \|\mathtt{t}\|$ (i.e. $\mathtt{t} \in w$) if and only if $w \in [\![\mathtt{t}^\tau]\!]^v$.

- $\mathtt{t}$ is a variable p. Then $\|p\| = v(p) = [\![p]\!]^v$ by the definition of v.

$-$ $\mathtt{t}$ is $\mathtt{t}_1{\cdot}\mathtt{t}_2$. Then $(\mathtt{t}_1{\cdot}\mathtt{t}_2)^\tau = \mathtt{t}_1^\tau{\cdot}\mathtt{t}_2^\tau$.
Suppose $\mathtt{t}_1{\cdot}\mathtt{t}_2 \in w$, and consider the worlds $w_{\mathtt{t}_1}$ and $w_{\mathtt{t}_2}$ in W_λ. If $\mathtt{s}_1 \in w_{\mathtt{t}_1}$ and $\mathtt{s}_2 \in w_{\mathtt{t}_2}$ then by Definition 3.4, $\mathtt{t}_1 \twoheadrightarrow_{\zeta_f} \mathtt{s}_1$ and $\mathtt{t}_2 \twoheadrightarrow_{\zeta_f} \mathtt{s}_2$. Thus $\mathtt{t}_1{\cdot}\mathtt{t}_2 \twoheadrightarrow_{\zeta_f} \mathtt{s}_1{\cdot}\mathtt{s}_2$ and $\mathtt{s}_1{\cdot}\mathtt{s}_2 \in w$. Then by the definition of $\bullet_\lambda$ we have that $w \in w_{\mathtt{t}_1} \bullet_\lambda w_{\mathtt{t}_2}$. Furthermore, $w_{\mathtt{t}_1} \in \|\mathtt{t}_1\|$ and so by the induction hypothesis, $w_{\mathtt{t}_1} \in [\![\mathtt{t}_1^\tau]\!]^v$. Similarly $w_{\mathtt{t}_2} \in [\![\mathtt{t}_2^\tau]\!]^v$. Hence $w \in [\![\mathtt{t}_1^\tau{\cdot}\mathtt{t}_2^\tau]\!]^v$ by Definition 2.13.

Conversely, suppose that $w \in [\![\mathtt{t}_1^\tau{\cdot}\mathtt{t}_2^\tau]\!]^v$. Then there are $w_{\mathtt{s}_1}, w_{\mathtt{s}_2}$ such that $w_{\mathtt{s}_1} \in [\![\mathtt{t}_1^\tau]\!]^v$ and $w_{\mathtt{s}_2} \in [\![\mathtt{t}_2^\tau]\!]^v$ and $w \in w_{\mathtt{s}_1} \bullet_\lambda w_{\mathtt{s}_2}$. By the induction hypothesis $w_{\mathtt{s}_1} \in \|\mathtt{t}_1\|$ and $w_{\mathtt{s}_2} \in \|\mathtt{t}_2\|$. Then $\vdash \mathtt{s}_1 \twoheadrightarrow_{\zeta_f} \mathtt{t}_1$ and $\vdash \mathtt{s}_2 \twoheadrightarrow_{\zeta_f} \mathtt{t}_2$. Furthermore, by the construction of $\bullet_\lambda$, $\mathtt{s}_1{\cdot}\mathtt{s}_2 \in w$ and hence by (*cong*) $\mathtt{t}_1{\cdot}\mathtt{t}_2 \in w$.
$-$ $\mathtt{t}$ is $\lambda p.\mathtt{s}$. Then $[\![(\lambda p.\mathtt{s})^\tau]\!]^v = [\![\Lambda p.\,\mathtt{s}^\tau]\!]^v = \{w \mid \forall h \in H_\lambda . w \bullet_\lambda h \subseteq [\![\mathtt{s}^\tau]\!]^{v[p\mapsto h]}\}$.
Suppose $\lambda p.\mathtt{s} \in w_1$. Suppose that $w_3 \in w_1 \bullet_\lambda w_2$, and that $w_2 \in h$ for some $h \in H_\lambda$, then $h = \|\mathtt{u}\|$ for some term $\mathtt{u}$. By (ζ_f) we have that $\mathtt{u} \twoheadrightarrow_{\zeta_f} r$ and $r \twoheadrightarrow_{\zeta_f} \mathtt{u}$ for some r. So $h = \|r\|$ and $r \in w_2$. By the construction of $\bullet_\lambda$, $\lambda p.\mathtt{s}{\cdot}r \in w_3$ and so $\mathtt{s}[p{::=}r] \in w_3$, i.e. $w_3 \in \|\mathtt{s}[p{::=}r]\|$. By the induction hypothesis $\|\mathtt{s}[p{::=}r]\| = [\![\mathtt{s}^\tau[p{::=}r]]\!]^v$. Furthermore by Lemma 2.22 $[\![\mathtt{s}^\tau[p{::=}r]]\!]^v = [\![\mathtt{s}^\tau]\!]^{v[p\mapsto [\![r]\!]^v]}$. But by the definition of v, $[\![r]\!]^v = \|r\|$, and so $w_3 \in [\![\mathtt{s}^\tau]\!]^{v[p\mapsto \|r\|]}$. But $h = \|r\|$ so $w_3 \in [\![\mathtt{s}^\tau]\!]^{v[p\mapsto h]}$. Thus $w_1 \in \{w \mid \forall h \in H_\lambda . w \bullet_\lambda h \subseteq [\![\mathtt{s}^\tau]\!]^{v[p\mapsto h]}\} = [\![(\lambda p.\mathtt{s})^\tau]\!]^v$. Hence, $\|\lambda p.\mathtt{s}\| \subseteq [\![(\lambda p.\mathtt{s})^\tau]\!]^v$

Conversely, suppose that $\lambda p.\mathtt{s} \notin w_\mathtt{u}$ for some $\mathtt{u}$. Let q be a variable not free in $\mathtt{u}$ or $\mathtt{s}$ and consider the worlds w_q and $w_{\mathtt{u}\cdot q}$. If $\mathtt{s}[p{::=}q] \in w_{\mathtt{u}\cdot q}$ then $\mathtt{u}{\cdot}q \twoheadrightarrow_{\zeta_f} \mathtt{s}[p{::=}q]$, so $\lambda q.(\mathtt{u}{\cdot}q) \twoheadrightarrow_{\zeta_f} \lambda q(\mathtt{s}[p{::=}q])$ by (ξ). But by our choice of q, (η) entails that $\mathtt{u} \twoheadrightarrow_{\zeta_f} \lambda q.(\mathtt{u}{\cdot}q)$. So $\mathtt{u} \twoheadrightarrow_{\zeta_f} \lambda q.\mathtt{s}[p{::=}q]$, which contradicts our initial supposition that $\lambda p.\mathtt{s} \notin w_\mathtt{u}$, therefore $\mathtt{s}[p{::=}q] \notin w_{\mathtt{u}\cdot q}$. In other words $w_{\mathtt{u}\cdot q} \notin \|\mathtt{s}[p{::=}q]\|$. Therefore, by the induction hypothesis, $w_{\mathtt{u}\cdot q} \notin [\![\mathtt{s}^\tau[p{::=}q]]\!]^v$. Since $[\![q]\!]^v = \|q\|$, it follows by Lemma 2.22 that $w_{\mathtt{u}\cdot q} \notin [\![\mathtt{s}^\tau]\!]^{v[p\mapsto \|q\|]}$. But clearly $w_{\mathtt{u}\cdot q} \in w_\mathtt{u} \bullet_\lambda w_q$, so it follows that $w_\mathtt{u} \notin \{w \mid \forall h \in H_\lambda . w \bullet_\lambda h \subseteq [\![\mathtt{s}^\tau]\!]^{v[p\mapsto h]}\}$. By the semantics of $(\lambda q.\mathtt{s})^\tau$ (i.e. $\Lambda q.\,\mathtt{s}^\tau$), this means that $w_\mathtt{u} \notin [\![(\lambda q.\mathtt{s})^\tau]\!]^v$. Hence, since every $w \in W_\lambda$ is $w_\mathtt{u}$ for some $\mathtt{u}$, $[\![(\lambda p.\mathtt{s})^\tau]\!]^v \subseteq \|\lambda p.\mathtt{s}\|$.

We can now prove the converse of Theorem 2.27:

Theorem 3.8 $\mathtt{t} \twoheadrightarrow \mathtt{s}$ *if and only if* $[\![\mathtt{t}^\tau \to \mathtt{s}^\tau]\!]^v = W$ *for all* v *and all frames* $F = (W, \bullet, H)$ *such that* $[\![\mathtt{u}^\tau]\!]^v \in H$ *for all* $\mathtt{u}$.[6]

Proof. The left-right direction is Theorem 2.27.

If $\mathtt{t} \not\twoheadrightarrow \mathtt{s}$ then $\mathtt{t} \not\twoheadrightarrow_{\zeta_f} \mathtt{s}$ and so $\mathtt{s} \notin w_\mathtt{t}$ in F_λ. Therefore $\|\mathtt{t}\| \not\subseteq \|\mathtt{s}\|$ and so by Theorem 3.7 there is a valuation v such that $[\![\mathtt{t}^\tau]\!]^v \not\subseteq [\![\mathtt{s}^\tau]\!]^v$. Furthermore, $H_\lambda = \{[\![\mathtt{u}^\tau]\!]^v \mid \mathtt{u} \text{ is a } \lambda\text{-term}\}$.

4 The axioms, Soundness and Completeness

We can axiomatise the interpretation of $\mathcal{L}$ given by Definition 2.13. Axioms are given in Figure 2.

[6] In other words, if $[\![\mathsf{E}(\mathtt{u}^\tau)]\!]^v = W$ for all terms $\mathtt{u}$.

$(\forall R)$	$\phi \to \forall p.\, \phi \qquad (p \notin \phi)$	$(Prop)$	Propositional Tautologies and Modus Ponens
$(\forall L)$	$\forall p.\, \phi \to \mathsf{E}\psi \to \phi[p{::=}\psi]$	$(\rhd L)$	$((\phi \rhd \psi)\cdot\phi) \to \psi$
$(\forall A)$	$\forall p.\, (\phi \to \psi) \to (\forall p.\, \phi \to \forall p.\, \psi)$	$(\rhd R)$	$\phi \to (\psi \rhd \phi\cdot\psi)$
(Gen)	$\dfrac{\mathsf{E}p \to \phi}{\forall p.\, \phi}$	$(\bot)$	$(\phi\cdot\bot) \to \bot$ $(\bot\cdot\phi) \to \bot$
$(\cdot K)$	$\phi\cdot(\psi \vee \mu) \to (\phi\cdot\psi) \vee (\phi\cdot\mu)$ $(\psi \vee \mu)\cdot\phi \to (\psi\cdot\phi) \vee (\mu\cdot\phi)$	(N)	$\dfrac{\phi_1 \to \cdots \to \phi_n \to \psi}{\Box\phi_1 \to \cdots \to \Box\phi_n \to \Box\psi} \quad 0 \leq n$
$(\cdot C)$	$\phi\cdot\exists p.\, \psi \to \exists p.\, (\phi\cdot\psi)$ $(\exists p.\, \psi)\cdot\phi \to \exists p.\, (\psi\cdot\phi) \qquad (p \notin \phi)$	(T) (5)	$\Box\phi \to \phi$ $\neg\Box\phi \to \Box\neg\Box\phi$
$(\rhd K)$	$(\phi \rhd \psi) \wedge (\phi \rhd \mu) \to \phi \rhd (\psi \wedge \mu)$ $(\psi \rhd \phi) \wedge (\mu \rhd \phi) \to (\psi \vee \mu) \rhd \phi$	$(\Box\cdot)$	$\Box(\phi \to \psi) \to (\phi\cdot\mu) \to (\psi\cdot\mu)$ $\Box(\phi \to \psi) \to (\mu\cdot\phi) \to (\mu\cdot\psi)$
$(\rhd C)$	$\forall p.\, (\phi \rhd \psi) \to \phi \rhd \forall p.\, \psi$ $\forall p.\, (\psi \rhd \phi) \to (\exists p.\, \psi \rhd \phi) \qquad (p \notin \phi)$	$(\Box\rhd)$	$\Box(\phi \to \psi) \to (\psi \rhd \mu) \to (\phi \rhd \mu)$ $\Box(\phi \to \psi) \to (\mu \rhd \phi) \to (\mu \rhd \psi)$

Fig. 2: Axioms for $\mathcal{L}$, we write '$p \notin \phi$' as short 'p is not free in ϕ'.

Definition 4.1 Let $\Gamma, \Delta \ldots$ denote sets of sentences. Write $\vdash \phi$ if ϕ is derivable using the rules of Figure 2. Write $\Gamma \vdash A$ when there are $\phi_1 \ldots \phi_n \in \Gamma$ such that $\vdash \phi_1 \to \ldots \phi_n \to \phi$ (associating to the right).

4.1 Theorems and admissible rules

Theorem 4.2 *The converses of* $(\cdot K)$,$(\cdot C)$,$(\rhd K)$, *and* $(\rhd C)$ *are all derivable. Also derivable are:*

$\forall p.\, \phi \leftrightarrow \forall p.\, (\mathsf{E}p \to \phi)$	$(\Box\phi\cdot\psi) \to \Box\phi$	$\Box\phi \to (\psi\cdot\mu) \to (\psi\cdot(\Box\phi \wedge \mu))$
$\forall p.\, \Box\phi \leftrightarrow \Box\forall p.\, \phi$	$(\psi\cdot\Box\phi) \to \Box\phi$	$\Box\phi \to (\psi\cdot\mu) \to ((\Box\phi \wedge \psi)\cdot\mu)$
$\phi \rhd \top$	$\neg\Box\phi \to (\Box\phi \rhd \psi)$	$\Box\phi \to (\psi \rhd \mu) \to (\psi \rhd (\Box\phi \wedge \mu))$
$\bot \rhd \phi$	$\neg\Box\phi \to (\psi \rhd \neg\Box\phi)$	$\Box\phi \to ((\Box\phi \wedge \psi) \rhd \mu) \to (\psi \rhd \mu)$

Notice that the second sentence of the leftmost column is the Barcan formula [13, Ch.13].

If $n = 0$ then (N) becomes a simple necessitation rule stating that $\vdash A$ implies $\vdash \Box A$. From this we get the following group of inference rules, $(Subs)$:

$$(Subs) \qquad \frac{\phi \to \psi}{(\phi\cdot\mu) \to (\psi\cdot\mu)} \qquad \frac{\phi \to \psi}{(\mu\cdot\phi) \to (\mu\cdot\psi)}$$

$$\frac{\phi \to \psi}{(\psi \rhd \mu) \to (\phi \rhd \mu)} \qquad \frac{\phi \to \psi}{(\mu \rhd \phi) \to (\mu \rhd \psi)}$$

4.2 Soundness

Theorem 4.3 *Suppose $F = (W, \bullet, H)$ is a frame. Then $\vdash \phi$ implies $[\![\phi]\!]^v = W$ for any v.*

Proof. By induction on derivations. Assume $\vdash \phi$. We consider each axiom and inference rule in turn.

To show that an axiom of the form $\phi \to \psi$ is sound it is enough to show that $[\![\phi]\!]^v \subseteq [\![\psi]\!]^v$ for any v, for that implies that $(W \setminus [\![\phi]\!]^v) \cup [\![\psi]\!]^v = W$.

– ϕ and is an instance of (*Prop*). The soundness of tautological consequence for the chosen interpretation of $\to$ is well known.

– Instances of ($\forall L$). Given Lemma 2.22 $\bigcap_{h\in H} [\![\phi]\!]^{v[p\mapsto h]} \subseteq \bigcap_{h\in H} [\![\phi]\!]^{v[p\mapsto [\![\psi]\!]^v]}$ provided that $[\![\psi]\!]^v \in H$. But if $[\![\psi]\!]^v \notin H$ then $[\![\mathsf{E}\psi]\!]^v = \varnothing$ and the axiom is validated.

– The cases for ($\forall R$) and ($\forall A$) are equally straightforward.

– Instances of (*Gen*). By induction hypothesis $[\![\mathsf{E}p \to \phi]\!]^v = W$ for any v so $[\![\mathsf{E}p]\!]^{v[p\mapsto h]} \subseteq [\![\phi]\!]^{v[p\mapsto h]}$ for any h. But $[\![\mathsf{E}\phi]\!]^{v[p\mapsto h]} = W$ for all h and so $[\![\forall p.\, A]\!]^v = W$.

– Instances of ($\rhd L$) and ($\rhd R$). We reason using Definitions 2.13 and 2.4:

$$\begin{aligned}
[\![(\phi \rhd \psi)\cdot\phi]\!]^v &= [\![\phi \rhd \psi]\!]^v \bullet [\![\phi]\!]^v && \text{Definition 2.13}\\
&= \{w \mid w \bullet [\![\phi]\!]^v \subseteq [\![\psi]\!]^v\} \bullet [\![\phi]\!]^v && \text{Definition 2.13}\\
&\subseteq [\![\psi]\!]^v && \text{Definition 2.4}
\end{aligned}$$

$$\begin{aligned}
[\![\phi]\!]^v &\subseteq \{w \mid w \bullet [\![\psi]\!]^v \subseteq [\![\phi]\!]^v \bullet [\![\psi]\!]^v\} && \text{Definition 2.4}\\
&= \{w \mid w \bullet [\![\psi]\!]^v \subseteq [\![\phi\cdot\psi]\!]^v\} && \text{Definition 2.13}\\
&= [\![(\psi \rhd \phi\cdot\psi)]\!]^v && \text{Definition 2.13}
\end{aligned}$$

– Instances of ($\cdot K$).

$$\begin{aligned}
&[\![(\phi\cdot\psi) \vee (\phi\cdot\mu)]\!]^v\\
&\quad = ([\![\phi]\!]^v \bullet [\![\psi]\!]^v) \cup ([\![\phi]\!]^v \bullet [\![\mu]\!]^v) && \text{Remark 2.15}\\
&\quad = \begin{array}{l}\bigcup\{w_1 \bullet w_2 \mid w_1 \in [\![\phi]\!]^v \ \&\ w_2 \in [\![\psi]\!]^v\}\\ \cup \bigcup\{w_1 \bullet w_2 \mid w_1 \in [\![\phi]\!]^v \&\ w_2 \in [\![\mu]\!]^v\}\end{array} && \text{Definition 2.4}\\
&\quad = \bigcup\{w_1 \bullet w_2 \mid w_1 \in [\![\phi]\!]^v \ \&\ w_2 \in ([\![\psi]\!]^v \cup [\![\mu]\!]^v)\} && \text{Definition 2.4}\\
&\quad = [\![\phi\cdot(\psi \vee \mu)]\!]^v && \text{Remark 2.15}
\end{aligned}$$

The other case for ($\cdot K$) and the cases for ($\rhd K$) are similar.

– Instances of ($\cdot C$) and ($\rhd C$).

$$\begin{aligned}
[\![\phi\cdot\exists p.\, \psi]\!]^v &= [\![\phi]\!]^v \bullet \textstyle\bigcup_{h\in H} [\![\psi]\!]^{v[p\mapsto h]} && \text{Remark 2.15}\\
&= \textstyle\bigcup_{h\in H}([\![\phi]\!]^v \bullet [\![\psi]\!]^{v[p\mapsto h]}) && \text{Definition 2.4}\\
&= \textstyle\bigcup_{h\in H}([\![\phi]\!]^{v[p\mapsto h]} \bullet [\![\psi]\!]^{v[p\mapsto h]}) && \text{Lemma 2.21}\\
&= [\![\exists p.\, (\phi\cdot\psi)]\!]^v && \text{Remark 2.15}
\end{aligned}$$

The other cases are similar.

– Instances of ($\bot$). $[\![\bot\cdot\phi]\!]^v = [\![\phi\cdot\bot]\!]^v = [\![\phi]\!]^v \bullet [\![\bot]\!]^v = \varnothing$ etc.
– Instances of ($\dot{\Box}$). We must show that $[\![\Box(\phi \to \psi)]\!]^v \cap [\![\phi\cdot\mu]\!]^v \subseteq [\![\psi\cdot\mu]\!]^v$. This is trivial if $[\![\phi \to \psi]\!]^v \neq W$, so we may assume that $[\![\phi]\!]^v \subseteq [\![\psi]\!]^v$. But then $[\![\phi\cdot\mu]\!]^v \subseteq [\![\psi\cdot\mu]\!]^v$. The argument is similar for ($\Box\triangleright$)
– Instances of (N). By induction hypothesis $\bigcap_i [\![\phi_i]\!]^v \subseteq \psi$ for any v. If for some ϕ_i, $[\![\phi_i]\!]^v \neq W$ then $[\![\Box\phi_i]\!]^v = \varnothing$ and so $\bigcap_i [\![\Box\phi_i]\!]^v \subseteq \Box\psi$. On the other hand, if $[\![\phi_i]\!]^v = W$ then $[\![\psi]\!]^v = [\![\Box\psi]\!]^v = W$ and again $\bigcap_i [\![\Box\phi_i]\!]^v \subseteq \Box\psi$.
– Axioms (T) and (5) are easily seen to be sound from semantic conditions for $\Box$.

4.3 Completeness

Definition 4.4 Say that a set of sentences Γ is **consistent** if $\Gamma \nvdash \bot$.

We will show that given a consistent set of sentences Γ we can construct a frame F and a valuation v such that $\bigcap_{\phi\in\Gamma} [\![\phi]\!]^v \neq \varnothing$.

Definition 4.5 A **maximal** set Δ is a consistent set such that
(1) $\phi \in \Delta$ or $\neg\phi \in \Delta$ for any ϕ,
(2) for every sentence ϕ there is some variable p such that $\phi \approx p \in \Delta$ (see Definition 2.17), and
(3) if $\neg\forall p.\, \phi \in \Delta$ then $\neg\phi[p{::=}q], \mathsf{E}q \in \Delta$ for some variable symbol q.

Remark 4.6 The second requirement on maximality ensures that every sentence ϕ is 'named' by some atomic variable. The third requirement is the more familiar condition that every existential quantifier have a 'Henkin witness'.

Lemma 4.7 *If Δ consistent then there exists a maximal set Δ' such that $\Delta \subseteq \Delta'$.*

Proof. Add two infinite collections of new propositional variable symbols $r_1, r_2 \ldots$, $c_1, c_2 \ldots$ to $\mathcal{L}$, then enumerate all sentences $\phi_0, \phi_1 \ldots$ and describe two one-one functions f, g from predicates to variables:

> $f(\phi_i) = r_j$ where j is the least number such that $j > i$ and r_j is not free in ϕ_i nor in Δ nor is the value under f of any $\phi_k < \phi_i$.
>
> $g(\forall p.\, \phi_i) = c_j$ where j is the least number such that $j > i$ and c_j is not free in ϕ_i nor in Δ nor is the value under f of any $\forall p.\, \phi_k < \forall p.\, \phi_i$. We also write $g(\forall p.\, \psi)$ as $g_{\forall p.\, \psi}$.

We now construct Δ_0, Δ_1,... as follows (using the enumeration $\phi_0, \phi_1 \ldots$ above, or a new one):

> $\Delta_0 = \Delta \cup \{\phi \approx f(\phi)\} \cup \{\neg\forall p.\, \phi \to \big(\neg\phi[p{::=}g_{\forall p.\, \phi}] \wedge \mathsf{E}(g_{\forall p.\, \phi})\big)\}$ for all ϕ.
>
> If $\Delta_n \cup \{\phi_n\}$ is inconsistent then $\Delta_{n+1} = \Delta_n \cup \{\neg\phi_n\}$, otherwise $\Delta_{n+1} = \Delta_n \cup \{\phi_n\}$.

Note that $\Delta_i \subseteq \Delta_j$ if $i \leq j$. Let $\Theta = \bigcup_n \Delta_n$. By construction $\Delta \subseteq \Theta$. We must prove Θ is maximal:

– Δ_0 is consistent: Suppose Δ_0 is inconsistent, then there are $\mu_1 \ldots \mu_n \in \Delta_0$ such that $\vdash \mu_1 \to \cdots \to \mu_n \to \bot$. Suppose the μ_i that do not occur in Δ are:

$$\phi_1 \approx f(\phi_1) \ldots \phi_k \approx f(\phi_k)$$

and

$$\neg\forall p.\, \psi_1 \to \neg\psi_1[p{::=}g_{\forall p.\,\psi_1}] \ldots \neg\forall p.\, \psi_l \to \neg\psi_l[p{::=}g_{\forall p.\,\psi_l}]$$

First simultaneously substitute $f(\phi_i)$ with ϕ_i. We get that $\vdash \mu'_1 \to \cdots \to \mu'_n \to \bot$ where each μ'_i is either in Δ or is $\phi \approx \phi$ or is $\neg\forall p.\, \psi' \to \neg\psi'[p{::=}g_{\forall p.\,\psi}]$ (where $\psi' = \psi[f(\phi_i){::=}\phi_i]$).
Let ρ be $\neg\forall p.\, \psi'_j \to \big(\neg\psi'_j[p{::=}g_{\forall p.\,\psi_j}] \wedge \mathsf{E}(g_{\forall p.\,\psi_j})\big)$ where ψ'_j is latest in the enumeration of all sentences. We may assume that $\rho = \mu'_n = \mu'_{m+1}$. Since $\vdash \mu'_1 \to \cdots \to \mu'_m \to \rho \to \bot$ we have by $(Prop)$ that

$$\vdash \mu'_1 \to \cdots \to \mu'_m \to \neg\forall p.\, \psi'_j$$

and

$$\vdash \mu'_1 \to \cdots \to \mu'_m \to \mathsf{E}(g_{\forall p.\,\psi'_j}) \to (\forall p.\, \psi'_j \to \psi'_j[p{::=}g_{\forall p.\,\psi'_j}])$$

but by our choice of $g_{\forall p.\,\psi'_j}$ it follows by the quantifier axioms that $\vdash \mu'_1 \to \cdots \to \mu'_m \to \forall p.\, \psi'_j$. So $\vdash \mu'_1 \to \cdots \to \mu'_m \to \bot$.
We may conclude from this that $\mu'_1 \to \cdots \to \mu'_l \to \bot$ is derivable where each μ'_i $(i < l \leq n)$ is either of the form $\phi \approx \phi$ or is in Δ. But this is impossible by the consistency of Δ.

– For every ϕ, either $\phi \in \Theta$ or $\neg\phi \in \Theta$: By the construction, either ϕ or $\neg\phi$ is added to some Δ_i. By the consistency of Θ, it is also deductively closed.

– By the construction of Δ_0, for every sentence ϕ, there is some variable p such that $\phi \leftrightarrow p \in \Delta_0 \subseteq \Theta$.

– If $\neg\forall p.\, \phi \in \Theta$ then for some c, $\neg\forall p.\, \phi \to (\neg\phi[p{::=}c] \wedge \mathsf{E}c) \in \Delta_0 \subseteq \Delta'$ and so $\neg\phi[p{::=}c], \mathsf{E}c \in \Theta$.

Thus Θ is indeed maximal.

Definition 4.8 If Θ is a maximal set then $\mathcal{C}_\Theta = \{\Delta \mid \Delta$ is maximal and $\Box\phi \in \Theta$ implies $\phi \in \Delta\}$.

Definition 4.9 Define the **canonical frame** $F_{\mathcal{C}_\Theta} = \langle \mathcal{C}_\Theta, \bullet_{\mathcal{C}_\Theta}, H_{\mathcal{C}_\Theta}\rangle$:

– For any $w_1, w_2 \in \mathcal{C}_\Theta$, $w_1 \bullet_{\mathcal{C}_\Theta} w_2 = \{w \in \mathcal{C}_\Theta \mid \phi \in w_1 \;\&\; \psi \in w_2$ implies $\phi{\cdot}\psi \in w\}$.
– $H_{\mathcal{C}_\Theta} = \big\{\{w \in \mathcal{C}_\Theta \mid \phi \in w\} \mid \mathsf{E}\phi \in \Theta\big\}$.

It follows by (T) that $\Theta \in \mathcal{C}_\Theta$.

Definition 4.10 Given $F_{\mathcal{C}_\Theta} = (\mathcal{C}_\Theta, \bullet_{\mathcal{C}_\Theta}, H_{\mathcal{C}_\Theta})$ and a sentence ϕ. Let $\|\phi\| = \{w \in \mathcal{C}_\Theta \mid \phi \in w\}$.

Theorem 4.11 *Let $F_{\mathcal{C}_\Theta}$ be the canonical frame, and let $v(p) = \|p\|$ for any (sentential) variable p. Then for any sentence ϕ, $[\![\phi]\!]^v = \|\phi\|$.*

Proof. By induction on ϕ.

– ϕ is p for some variable p. Then $\|p\| = v(p) = [\![p]\!]^v$ by the definition of v.

– ϕ is $\phi_1 \to \phi_2$.

Suppose that $\phi_1 \to \phi_2 \in w$. If $w \in [\![\phi_1]\!]^v$, then by the induction hypothesis $w \in \|\phi_1\|$, i.e. $\phi_1 \in w$. So $\phi_2 \in w$ and $w \in [\![\phi_2]\!]^v$ by the induction hypothesis. Thus $w \in W \setminus [\![\phi_1]\!]^v \cup [\![\phi_2]\!]^v$.

Conversely, suppose that $\phi_1 \to \phi_2 \notin w$. Then by $(Prop)$ $\neg\phi_1 \notin w$ and $\phi_2 \notin w$. By the induction hypothesis, and the maximality of w, we may conclude that $w \notin W \setminus [\![\phi_1]\!]^v$ and $w \notin [\![\phi_2]\!]^v$.

– ϕ is $\bot$. By the consistency of every $w \in \mathcal{C}_\Theta$, $\|\bot\| = \varnothing = [\![\bot]\!]^v$.

– ϕ is $\Box\psi$.

If $\Box\psi \in w$ then by (5) and the construction of $\mathcal{C}_\Theta$, $\neg\Box\psi \notin \Theta$. So $\Box\psi \in \Theta$ and $\psi \in w'$ for all $w' \in \mathcal{C}_\Theta$.

For the converse case suppose that $\psi \in w$ for all $w \in \mathcal{C}_\Theta$. Then since $\mathcal{C}_\Theta$ contains all maximal consistent sets containing $\{\mu \mid \Box\mu \in \Theta\}$ it follows that $\{\mu \mid \Box\mu \in \Theta\} \vdash \psi$. So by (N), $\{\Box\mu \mid \Box\mu \in \Theta\} \vdash \Box\psi$ and so $\Box\psi \in \Theta$. But by (T), (N) and (5), $\vdash \Box\psi \to \Box\Box\psi$, so $\Box\psi \in w'$ for any $w' \in \mathcal{C}_\Theta$.

– ϕ is $\forall p.\, \psi$.

Suppose that $\forall p.\, \psi \in w$ then by $(\forall L)$ $\psi[p{::=}\mu] \in w$ whenever $\mathsf{E}\mu \in w$. By the cases above for $\Box$, $\psi[p{::=}\mu] \in w$ whenever $\mathsf{E}\mu \in \Theta$.[7] By the maximality of Θ, for each μ, there is a variable f_μ such that $\mu \approx f_\mu \in \Theta$. Thus $\|\mu\| = \|f_\mu\|$ and every $h \in H_{\mathcal{C}_\Theta}$ is $\|f_\mu\|$ for some μ such that $\mathsf{E}\mu \in \Theta$. By the induction hypothesis and Lemma 2.22, $w \in [\![\psi[p{::=}f_\mu]]\!]^{v[f_\mu \mapsto h]}$ for all $h \in H_{\mathcal{C}_\Theta}$. Thus $w \in [\![\forall p.\, \mu]\!]^v$.

Conversely, suppose that $\forall p.\, \psi \notin w$. Then by the maximality of w, $\neg\forall p.\, \psi \in w$ and so $\neg\psi'[p{::=}c] \wedge \mathsf{E}c \in w$ for some c. By the induction hypothesis and Lemma 2.22 we have that $w \notin [\![\psi']\!]^{v[p \mapsto [\![c]\!]^v]}$, we also have $\mathsf{E}c \in \Theta$, so $w \notin [\![\forall p.\, \psi]\!]^v$.

– ϕ is $\phi_1 {\cdot} \phi_2$.

Suppose $w_3 \in [\![\phi_1 {\cdot} \phi_2]\!]^v$. Then there are w_1, w_2 such that $w_1 \in [\![\phi_1]\!]^v$ and $w_2 \in [\![\phi_2]\!]^v$ and $w_3 \in w_1 \bullet_{\mathcal{C}_\Theta} w_2$. By the induction hypothesis $w_1 \in \|\phi_1\|$ and $w_2 \in \|\phi_2\|$, so $\phi_1 \in w_1$ and $\phi_2 \in w_2$. This implies $\phi_1 {\cdot} \phi_2 \in w_3$.

For the converse case suppose that $\phi_1 {\cdot} \phi_2 \in w_3$, we must show that there are w_1, w_2 such that $w_1 \in [\![\phi_1]\!]^v$, $w_2 \in [\![\phi_2]\!]^v$ and $w_3 \in w_1 \bullet_{\mathcal{C}_\Theta} w_2$. Given the induction hypothesis, it is enough to construct two maximal sets Δ_1, Δ_2 such that $\phi_1 \in \Delta_1, \phi_2 \in \Delta_2$ and $\psi_1 {\cdot} \psi_2 \in w_3$ for every $\psi_1 \in \Delta_1, \psi_2 \in \Delta_2$. We must then verify that these two sets are in $\mathcal{C}_\Theta$ by showing that $\{\psi \mid \Box\psi \in \Theta\} \subseteq \Delta_1 \cap \Delta_2$. This is done in Lemma 4.12.

[7] $\mathsf{E}\mu$ is short for $\exists p.\, \Box(p \leftrightarrow \mu)$. The axioms for $S5$ (N),(T) and (5) entail that $\vdash \exists p.\, \Box(p \leftrightarrow \mu) \to \Box\exists p.\, \Box(p \leftrightarrow \mu)$. So, by the case for $\Box$, $\mathsf{E}\mu \in w$ iff $\mathsf{E}\mu \in w'$ for any $w, w' \in \mathcal{C}_\Theta$.

– The case where ϕ is $\phi_1 \rhd \phi_2$ is similar to that for $\phi_1{\cdot}\phi_2$ and uses a lemma similar to Lemma 4.12.

Lemma 4.12 *If $\Delta \in \mathcal{C}_\Theta$ and $\phi_1{\cdot}\phi_2 \in \Delta$, then there are two maximal sets $\Delta_1, \Delta_2 \in \mathcal{C}_\Theta$ such that*

(1) $\phi_1 \in \Delta_1, \phi_2 \in \Delta_2$

(2) $\psi_1{\cdot}\psi_2 \in \Delta$ for every $\psi_1 \in \Delta_1, \psi_2 \in \Delta_2$

(3) $\Box\psi \in \Theta$ implies $\psi \in \Delta_1 \cap \Delta_2$ for any ψ.

Proof. Enumerate all sentences $\psi_0, \psi_1 \ldots$ and construct two sequences $\Phi_0, \Phi_1, \ldots$ and $\Psi_0, \Psi_1, \ldots$:

$\Phi_0 = \{\neg\neg\phi_1\}$ and $\Psi_0 = \{\neg\neg\phi_2\}$

If ψ_n is not of the form $\forall p.\, \mu$ then:

$$\Phi_{n+1} = \begin{cases} \Phi_n \cup \{\neg\psi_n\} & \text{if } (\bigwedge\Phi_n \wedge \neg\psi_n){\cdot}(\bigwedge\Psi_n) \in \Delta \\ \Phi_n \cup \{\psi_n\} & \text{otherwise} \end{cases}$$

$$\Psi_{n+1} = \begin{cases} \Psi_n \cup \{\neg\psi_n\} & \text{if} (\bigwedge\Phi_{n+1}){\cdot}(\bigwedge\Psi_n \wedge \neg\psi_n) \in \Delta \\ \Psi_n \cup \{\psi_n\} & \text{otherwise} \end{cases}$$

If ψ_n is of the form $\forall p.\, \mu$ then:

$$\Phi_{n+1} = \begin{cases} \Phi_n \cup \{\neg\forall p.\, \mu, \neg\mu[p{::=}c], \mathsf{E}c\} & \\ \quad \text{if } \begin{array}{l}(\bigwedge\Phi_n \wedge \neg\forall p.\, \mu \wedge \neg\mu[p{::=}c] \wedge \mathsf{E}c){\cdot}(\bigwedge\Psi_n) \in \Delta \\ \text{for some variable } c\end{array} & \\ \Phi_n \cup \{\forall p.\, \mu\} \text{ otherwise} & \end{cases}$$

$$\Psi_{n+1} = \begin{cases} \Psi_n \cup \{\neg\forall p.\, \mu, \neg\mu[x{::=}c], \mathsf{E}c\} & \\ \quad \text{if } \begin{array}{l}(\bigwedge\Phi_{n+1}){\cdot}(\bigwedge\Psi_n \wedge \neg\forall p.\, \mu \wedge \neg\mu[p{::=}c] \wedge \mathsf{E}c) \in \Delta \\ \text{for some variable } c\end{array} & \\ \Psi_n \cup \{\forall p.\, \mu\} \text{ otherwise} & \end{cases}$$

Note that $\Phi_i \subseteq \Phi_j$ and $\Psi_i \subseteq \Psi_j$ if $i \leq j$. Let $\Delta_1 = \bigcup_n \Phi_n$ and $\Delta_2 = \bigcup_n \Psi_n$.

– Each $(\bigwedge\Phi_n){\cdot}(\bigwedge\Psi_n) \in \Delta$:
By induction on n. If $n = 0$ then since $\phi_1{\cdot}\phi_2 \in \Delta$ it follows by $(Prop)$ and $(Subs)$ that $\neg\neg\phi_1{\cdot}\neg\neg\phi_2 \in \Delta$.
Assume that $(\bigwedge\Phi_n){\cdot}(\bigwedge\Psi_n) \in \Delta$ but $(\bigwedge\Phi_{n+1}){\cdot}(\bigwedge\Psi_n) \notin \Delta$. First we must show that

$$(\bigwedge\Phi_n \wedge \neg\psi_n){\cdot}(\bigwedge\Psi_n) \notin \Delta \quad \text{implies} \quad (\bigwedge\Phi_n \wedge \psi_n){\cdot}(\bigwedge\Psi_n) \in \Delta. \qquad (\dagger)$$

By $({\cdot}K)$, $(Prop)$ and the consistency of Δ if $(\bigwedge\Phi_n \wedge \neg\psi_n){\cdot}(\bigwedge\Psi_n) \notin \Delta$ and $(\bigwedge\Phi_n \wedge \psi_n){\cdot}(\bigwedge\Psi_n) \notin \Delta$ then

$$\big(\bigwedge(\Phi_n \wedge \psi_n) \vee (\bigwedge\Phi_n \wedge \neg\psi_n)\big){\cdot}(\bigwedge\Psi_n) \notin \Delta.$$

So by (Prop) and (Subs) $\big((\bigwedge\Phi_n \wedge (\psi_n \vee \neg\psi_n))\cdot(\bigwedge\Psi_n)\big) \notin \Delta$. But this entails that $(\bigwedge\Phi_n)\cdot(\bigwedge\Psi_n) \notin \Delta$ which is contrary to our assumption.
So the lemma holds if ψ_n is not of the form $\forall p.\, \mu$. Suppose ψ_n is of the form $\forall p.\, \mu$. We must show that:

$$(\bigwedge\Phi_n \wedge \neg\forall p.\, \mu \wedge \neg\mu[p{::=}c] \wedge \mathsf{E}c)\cdot(\bigwedge\Psi_n) \notin \Delta \quad \text{for all } c,$$

implies that

$$(\bigwedge\Phi_n \wedge \forall p.\, \mu)\cdot(\bigwedge\Psi_n) \in \Delta$$

Given †, we need only show that

$$(\bigwedge\Phi_n \wedge \neg\forall p.\, \mu \wedge \neg\mu[p{::=}c] \wedge \mathsf{E}c)\cdot(\bigwedge\Psi_n) \notin \Delta \quad \text{for all } c,$$

implies that

$$(\bigwedge\Phi_n \wedge \neg\forall p.\, \mu)\cdot(\bigwedge\Psi_n) \notin \Delta$$

If $(\bigwedge\Phi_n \wedge \neg\forall p.\, \mu)\cdot(\bigwedge\Psi_n) \in \Delta$ then by $(\cdot C)$, $\neg\forall p.\, \neg\big((\bigwedge\Phi_n \wedge \neg\forall p.\, \mu \wedge \neg\mu)\cdot(\bigwedge\Psi_n)\big) \in \Delta$.[8] But since Δ is maximal and every negated universal quantification has a witness:

$$\neg\big((\bigwedge\Phi_n \wedge \neg\forall p.\, \mu \wedge \neg\mu[p{::=}c])\cdot(\bigwedge\Psi_n)\big) \wedge \mathsf{E}c \in \Delta \quad \text{for some } c$$

But then $(\bigwedge\Phi_n \wedge \neg\forall p.\, \mu \wedge \neg\mu[p{::=}c] \wedge \mathsf{E}c)\cdot(\bigwedge\Psi_n) \notin \Delta$ (for some c).
So we can conclude that $(\bigwedge\Phi_n)\cdot(\bigwedge\Psi_n) \in \Delta$ implies that $(\bigwedge\Phi_{n+1})\cdot(\bigwedge\Psi_n) \in \Delta$. Analogous reasoning shows that this in turn implies that $(\bigwedge\Phi_{n+1})\cdot(\bigwedge\Psi_{n+1}) \in \Delta$

– Δ_1, Δ_2 are consistent:
Suppose Δ_1 is inconsistent, then there are $\mu_1 \dots \mu_n \in \Delta_1$ such that $\vdash \mu_1 \to \dots \to \mu_n \to \bot$. The μ_i must all be in some $\Phi_k \subseteq \Delta_1$, but as $(\bigwedge\Phi_k)\cdot(\bigwedge\Psi_k) \in \Delta$ this implies by $(\bot)$ that $\bot \in \Delta$. This is impossible since Δ is consistent. We may conclude analogously that Δ_2 is not inconsistent.

– For any ϕ, either $\phi \in \Delta_1$ or $\neg\phi \in \Delta_1$:
This follows from the fact that every $\phi \in \Delta_1$ is a ψ_i, and so either it or its negation is added to Φ_i. Similarly, $\phi \in \Delta_2$ or $\neg\phi \in \Delta_2$ for any ϕ.

– $\neg\forall p.\, \mu \in \Delta_1$ implies $\neg\mu[p{::=}c], \mathsf{E}c \in \Delta_1$ for some c:
$\neg\forall p.\, \mu$ is a ψ_{i+1} and so is added to Φ_{i+1}, but $\Phi_{i+2} = \Phi_{i+1} \cup \{\neg\forall p.\, \mu, \neg\mu[p{::=}c], \mathsf{E}c\}$ for some c.[9] Similarly for Δ_2.

So Δ_1 and Δ_2 are maximal, now we verify that they satisfy the conditions of the lemma.

(1) $\phi_1 \in \Delta_1$ and $\phi_2 \in \Delta_2$: By choice of Φ_0, Ψ_0 we have that $\neg\neg\phi_1 \in \Delta_1$ and $\neg\neg\phi_2 \in \Delta_2$, so by (Prop) and the maximality of Δ_1 and Δ_2 it follows that $\phi_1 \in \Delta_1$ and $\phi_2 \in \Delta_2$.

[8] As we may assume that p is not free in $\bigwedge\Phi_n$.
[9] We chose Φ_0 and Ψ_0 to be $\{\neg\neg\phi_1\}$ and $\{\neg\neg\phi_2\}$, to guarantee the that the construction did not begin with a sentence of the form $\neg\forall p.\, \mu$.

(2) $\psi_1{\cdot}\psi_2 \in \Delta$ for every $\psi_1 \in \Delta_1, \psi_2 \in \Delta_2$. Choosing some suitable large i, we have that $\psi_1 \in \Phi_i$ and $\psi_2 \in \Psi_i$ and the result follows by (*Prop*), (*Subs*) and the fact that $\bigwedge\Phi_i{\cdot}\bigwedge\Psi_i \in \Delta$
(3) If $\Box\psi \in \Theta$ then $\Box\Box\psi \in \Theta$ and so, since $\Delta \in \mathcal{C}_\Theta$, $\Box\psi \in \Delta$. Now, if $\neg\Box\psi \in \Delta_1$ or $\neg\Box\psi \in \Delta_2$ then, by (1) and (2), $(\neg\Box\psi){\cdot}\phi_2 \in \Delta$ or $\phi_1{\cdot}(\neg\Box\psi) \in \Delta$. But $\Box\psi \in \Delta$, so by Theorem 4.2 this implies that $(\Box\psi \wedge \neg\Box\psi){\cdot}\phi_2 \in \Delta$ or $\phi_1{\cdot}(\Box\psi \wedge \neg\Box\psi) \in \Delta$. This is impossible since Δ_1 and Δ_2 are consistent. So $\Box\psi \in \Delta_1 \cap \Delta_2$ and by (T) $\psi \in \Delta_1 \cap \Delta_2$.

Theorem 4.13 *If Δ is consistent then $\bigcap_{\phi\in\Delta}[\![\phi]\!]^v \neq \varnothing$ for some frame F and valuation v.*

Proof. We have shown that Δ can be extended to a maximal set Θ which is in the canonical frame $F_{\mathcal{C}_\Theta}$. Then by Theorem 4.11, $\phi \in \Delta$ implies that $\Theta \in [\![\phi]\!]^v \in W_{\mathcal{C}_\Theta} \in F_{\mathcal{C}_\Theta}$, so $\Theta \in \bigcap_{\phi\in\Delta}[\![\phi]\!]^v \neq \varnothing$.

Corollary 4.14 *If $[\![\phi]\!]^v = W$ for all frames F, then $\vdash \phi$.*

Proof. If $\nvdash \phi$ then $\{\neg\phi\}$ is consistent, so there is a frame F such that $[\![\neg\phi]\!]^v \neq \varnothing$. By the semantics of negation it follows that $[\![\phi]\!]^v \neq W$.

We can use this result together with Definition 4.1 to simplify Theorem 3.8.

Corollary 4.15 $\mathtt{t} \twoheadrightarrow \mathtt{s}$ *if and only if* $\{\mathsf{E}\mathtt{u}^\tau \mid \mathtt{u}$ *is a λ-term*$\} \vdash \mathtt{t}^\tau \to \mathtt{s}^\tau$

It is a further issue whether Corollary 4.15 holds if $\{\mathsf{E}\mathtt{t}^\tau \mid \mathtt{t}$ is a λ-term$\}$ is replaced with $\{\mathsf{E}\phi \mid \phi$ is a sentence$\}$, or even $\{\mathsf{E}\phi \mid \phi$ is a closed sentence$\}$. A result equivalent to the fact that the corollory does not hold for the assumptions $\{\mathsf{E}\mathtt{u}^\tau \mid \mathtt{u}$ is a closed λ-term$\}$ was shown by Plotkin in [18].

5 Conclusion

5.1 Negation and the Liar

How does our logic resolve the paradoxes that affected Church's original system? We can extend τ (Definition 2.25) to translate $(\neg\mathtt{t})^\tau$ to either $\neg(\mathtt{t}^\tau)$ or $(\Lambda p.\,\neg p){\cdot}\mathtt{t}^\tau$. The first case corresponds to negation as a term-former in λ-term syntax; the second case corresponds to treating negation as a constant-symbol.

In the first case, let L^τ be short for $\Lambda p.\,\neg(p{\cdot}p){\cdot}\Lambda p.\,\neg(p{\cdot}p)$. Then we may use Theorem 2.27 and Corollary 4.14 to conclude that $\{\mathsf{E}(\mathtt{u}^\tau) \mid \mathtt{u}$ is a λ-term$\} \vdash L^\tau \to \neg L^\tau$. So just as with Church's system we get a sentence L^τ that implies its own negation. Since $[\![\neg L^\tau]\!]^v = W \setminus [\![L^\tau]\!]^v$ there is only one possible interpretation of L^τ: the empty set.

In the second case similar reasoning applies, but more interpretations of L^τ are available. This is because $(\Lambda p.\,\neg p){\cdot}\mathtt{t}^\tau$ may receive a different interpretation from $\neg\mathtt{t}^\tau$, even if we assume $\mathsf{E}(\mathtt{u}^\tau)$ for all terms $\mathtt{u}$. The reason for this is that $[\![\Lambda p.\,\neg p]\!]^v = \bigcap_{h\in H}\{w \mid w \bullet h \subseteq W \setminus h\}$ and so contains only those $w \in W$

that relate, by $\cdot$, members of H to their complements. So although $h \in H$ has a complement in $\mathcal{P}(W)$, there may be no $w \in W$ to serve in the extension of $[\![\Lambda p.\neg p]\!]^v$.

For example, if F is a frame where $[\![\top]\!]^v = W \in H$ then $w \in [\![\Lambda p.\neg p]\!]^v$ implies $w \bullet W \subseteq \varnothing$ and so $w \bullet S \subseteq \varnothing$ for any $S \subseteq W$ (as $\bullet$ is monotonic with respect to $\subseteq$). So $w \in [\![\Lambda p.\neg p]\!]^v$ implies $w \bullet w' = \varnothing$ for all $w' \in W$. So for such a frame F, $[\![(\Lambda p.\neg p)\cdot\phi]\!]^v = \varnothing = [\![\bot]\!]^v$ for any ϕ!

What moral can we draw from this? The negations of λ-terms can be interpreted perfectly naturally in our models. Paradoxes are averted because they may translate to impossible structural properties on the frames. Our models might help design other extensions of the λ-calculus, by considering how these extensions behave when interpreted in the models.

5.2 Related work

Multiplicative conjunction. $\mathcal{L}$ with its connective $\cdot$ (Definition 2.12) looks like a (classical) logic with a *multiplicative conjunction* $\otimes$, as in linear logic or bunched implications [8,17]. Multiplicative conjunction $\otimes$ does not have contraction, so that for example $A \otimes A$ is not equivalent to A.

However $\cdot$ is *not* a species of multiplicative conjunction. This is because multiplicative conjunction is usually taken to be associative and commutative, whereas $\cdot$ is neither; it models application, and we do not usually want $f(x)$ to equal $x(f)$, or $f(g(x))$ to equal $(f(g))(x)$.[10]

Phase spaces. On a related point, a frame F based on a set W with its function $\bullet$ from Definition 2.1 looks like a phase space [9]. Indeed the denotation for λ in Definition 2.7 uses the same idea as the denotation of multiplicative implication $\multimap$ (see for example [9, Section 2.1.1]).

However F is unlike a phase space in one very important respect: $\bullet$ does not make W into a commutative monoid because it maps $W \times W$ to $\mathcal{P}(W)$, and not to W. This is also, as we have mentioned, why W is not an applicative structure.

An interesting interpretation of our models. The 'application operation' $\bullet$ returns not worlds but *sets* of worlds. In Section 2.1 we have already suggested that we can read $\bullet$ as a ternary Kripke accessability relation, or as a non-deterministic application operation. We would now like to suggest another reading, which we have found useful.

Think of worlds $w \in W$ as objects, programs, and/or data. What happens when we apply one object to another (e.g. a rock to a nut; a puncher to some tape; a program to an input; a predicate to a datum)? On the one hand, we obtain an output (a nut that is broken; strip of perforated tape; a return value; a truth-value). Yet, on its own, an output is meaningless. What makes the execution of some action *a computation* is not the raw output, but the concept that this

[10] There is some interest in non-commutative multiplicative conjunction, but this does not change the basic point.

output signifies. That is, we apply w_1 to w_2 not to obtain some w_3, but to obtain some meaning that w_3 signifies. A raw output like '42' tells us nothing; it is only significant relative to the meaning we give it.

The output of a computation, rather than a mere action, is a *concept*.

As is standard, we can interpret concepts extensionally as sets of data. So, when $\bullet$ maps $W \times W$ to $\mathcal{P}(W)$ we can read this as follows: $\bullet$ takes two objects and applies one to the other to obtain a concept.

By this reading, when we write $\lambda p.\mathtt{t}$ or $\forall p.\,\phi$, p quantifies over concepts — not over data. Data is certainly there, and resides in W, but when we compute on W this returns us to the world of concepts. It is certainly possible to envisage frames in which $w_1 \bullet w_2$ is always a singleton $\{w_3\}$ — but this is just a very special case (and our completeness proofs do not generate such frames).

The fact that $\bullet$ maps to $\mathcal{P}(W)$ is a key point of the semantics in this paper. It turns out that this is sufficient to unify logic and computation, as we have seen.

Relevance logic. The notion of a function from $W \times W$ to $\mathcal{P}(W)$ does appear in the form of a ternary relation R on W, when giving denotations to *relevant* implication in *relevance logic* [5], and to implication in other provability and substructural logics such as *independence logic* [22]. For example, the clause for $\triangleright$ in Definition 2.13 is just like the clause for relevant implication $\rightarrow$ in [5, §3.7, p.69].

However, these logics impose extra structure on R; see for example conditions (1) to (4) in [5, §3.7, p.68]. In the notation of this paper, $\bullet$ for relevance logic and other substructural logics models a form of logical conjunction and has structural properties like associativity and commutativity. In our frames $\bullet$ models function application, which does not have these structural properties.

H and Henkin models for higher-order logic. Another feature of frames is the set $H \subseteq \mathcal{P}(W)$. We mentioned in Remark 2.3 that we use H to restrict quantification and 'cut down the size' of powersets so as to obtain completeness. This idea is standard from Henkin semantics for higher-order logics.

Here, two classes of frame are particularly interesting: **full** frames in which $H = \mathcal{P}(W)$ (which we return to below), and **faithful** frames in which the denotation of every possible sentence is in H (see [20, Section 4.3]). Full frames are simple, and may be represented as a pair $F^{\mathrm{full}} = (W, \bullet)$, but they inherit an overwhelming richness of structure from the full powerset. Henkin models are simpler 'first order' [20, Corollary 3.6] — and therefore completely axiomatisable — approximations to the full models.

Henkin semantics for higher-order logic are actually unsound without the assumption of faithfulness. We do not impose a general condition that models must be faithful because we built the models in the general case without committing to one specific logic. Once we fix a logic, conditions analogous to faithfulness begin to appear. See for example Theorems 2.27 and 3.8, and Corollary 4.15.

Investigating the properties of full models is possible further work.

5.3 Summary, and further work

We have presented models in which logic and computation have equal standing. They combine, among other things, the expressive power of untyped λ-calculus and quantificational logic, in a single package. This has allowed us to give interesting answers to two specific questions:

Q. What is the negation of a λ-term?
A. Its sets complement.

Q. What logical connective corresponds to functional (λ) abstraction?
A. $\curlywedge$ from Definition 2.19.

There are many questions to which we do not have answers.

The logic $\mathcal{L}$ is very expressive; interesting things can be expressed other than the λ-calculus, including encodings of first-order logic and simple types, and also less standard constructs. We note in particular matching (Definition 2.19) as a 'dual' to λ. What other programming constructs do the semantics and the logic $\mathcal{L}$ invite?

We can take inspiration from modal logic, and note how different conditions on accessibility in Kripke models match up with different systems of model logic [13]. It is very interesting to imagine that conditions on H and $\bullet$ might match up with systems of λ-reduction and equality.

Can we tweak the frames so that Theorem 2.27 becomes provable for a reduction relation that does not include (η), or perhaps does include its converse (yes, but there is no space here for details). λ-calculus embeds in $\mathcal{L}$, so can a sequent system be given for $\mathcal{L}$ extending the sequent system for λ-reduction of [6]?

Note that logic-programming and the Curry-Howard correspondence both combine logic and computation, where computation resides in proof-search and proof-normalisation respectively.

Where does our combination of logic and computation fit into this picture, if at all?

We can only speculate on applications of all this.

Models can be computationally very useful. For instance, we may falsify a predicate by building a model that does not satisfy it. Our models might have something specific to offer here, because they are fairly elementary to construct and the tie-in to the languages $\mathcal{L}_\lambda$ and $\mathcal{L}$ is very tight, with completeness results for arbitrary theories (see Definition 4.1 and Corollary 4.14).

We have already touched on possible applications to language design; we might use the models and the logic to design new language constructs. We note in passing that the models have a built-in notion of location, given by reading $w \in W$ as a 'world'. It is not entirely implausible that this might be useful to give semantics to languages with subtle constructs reflecting that not all computation takes place on a single thread. To illustrate what we have in mind, consider a simple 'if...then...else' λ-term $\lambda p.\psi\left\{\begin{smallmatrix}\mathtt{t}_1\\ \mathtt{t}_2\end{smallmatrix}\right.$ which is such that $\lambda p.\psi\left\{\begin{smallmatrix}\mathtt{t}_1\\ \mathtt{t}_2\end{smallmatrix}\right.\cdot\mathtt{s}$ reduces to $\mathtt{t}_1[p{::=}\mathtt{s}]$ if ψ, and to $\mathtt{t}_2[p{::=}\mathtt{s}]$ otherwise. But 'where' should ψ be checked? Should it be checked at the world where the function resides, where the argument resides, or where the result resides? Translations into our logic reflect these possibilities:

$$\begin{array}{ll} \text{At the function:} & \forall p.\,\big((\psi \to (p \rhd \phi_1)) \wedge (\neg\psi \to (p \rhd \phi_2))\big) \\ \text{At the argument:} & \forall p.\,\big(((p \wedge \psi) \rhd \phi_1) \wedge ((p \wedge \neg\psi) \rhd \phi_2)\big) \\ \text{At the result:} & \forall p.\,\big(p \rhd ((\psi \to \phi_1) \wedge (\neg\psi \to \phi_2))\big) \end{array}$$

Note that p may be free in ψ.

We conclude the paper with a hypothesis. Consider the full frames where $H = \mathcal{P}(W)$ and using the translation τ of Definition 2.25 consider the relation $\twoheadrightarrow_2$ defined such that $\mathtt{t} \twoheadrightarrow_2 \mathtt{s}$ when $[\![\mathtt{t}^\tau \to \mathtt{s}^\tau]\!]^v = W$ for any valuation v on any full frame. Our hypothesis is this: there are $\mathtt{t}$ and $\mathtt{s}$ such that $\mathtt{t} \twoheadrightarrow_2 \mathtt{s}$ but $\mathtt{t} \not\twoheadrightarrow \mathtt{s}$; furthermore, $\mathtt{t} \twoheadrightarrow_2 \mathtt{s}$ is 'true' in the sense that $\mathtt{t}$ intuitively does compute to $\mathtt{s}$. In other words, we hypothesise that the intuitive concept of computation is captured by the F^{full}, just like our intuitive concept of natural number is captured by the standard model $\mathcal{N}$. We suggest that λ-calculi and axiomatisations of computation are actually first order implementations of $\twoheadrightarrow_2$.

References

1. Christoph Benzmüller, Chad E. Brown, and Michael Kohlhase. Higher-order semantics and extensionality. *Journal of Symbolic Logic*, 69:1027–1088, 2004.
2. Wil Dekkers, Martin Bunder, and Henk Barendregt. Completeness of the propositions-as-types interpretation of intuitionistic logic into illative combinatory logic. *Journal of Symbolic Logic*, 3:869–890, 1998.
3. Wil Dekkers, Martin Bunder, and Henk Barendregt. Completeness of two systems of illative combinatory logic for first-order propositional and predicate calculus. *Archive für Mathematische Logik*, 37:327–341, 1998.
4. Gilles Dowek, Thrse Hardin, and Claude Kirchner. Theorem proving modulo. *Journal of Automated Reasoning*, 31(1):33–72, September 2003.
5. J. Michael Dunn and Greg Restall. Relevance logic. In Dov M. Gabbay, editor, *Handbook of Philosophical Logic*, volume 6, pages 1–136. Kluwer Academic Publishers, second edition, 2002.
6. Michael Gabbay and Murdoch J. Gabbay. Term sequent logic. In *Proceedings of the 17th International Workshop on Functional and (Constraint) Logic Programming (WFLP 2008)*, volume 246 of *Electronic Notes in Theoretical Computer Science*, pages 87–106, August 2009.
7. Benjamin Werner and Gilles Dowek. Arithmetic as a theory modulo. In *Proceedings of RTA*, pages 423–437, 2005.
8. Jean-Yves Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.
9. Jean-Yves Girard. Linear logic: its syntax and semantics. In *Proceedings of the workshop on Advances in linear logic*, pages 1–42, New York, 1995. Cambridge University Press.
10. Jean-Yves Girard, Paul Taylor, and Yves Lafont. *Proofs and types.* Cambridge University Press, 1989.
11. Robert Goldblatt. *Logics of Time and Computation.* Number 7 in CSLI Lecture Notes. Center for the Study of Language and Information, 2. edition, 1992.
12. Leon Henkin. Completeness in the theory of types. *Journal of Symbolic Logic*, 15:81–91, 1950.

13. George E. Hughes and Maxwell J. Cresswell. *A New Introduction to Modal Logic*. Routledge, 1996.
14. C. Barry Jay and Neil Ghani. The virtues of eta-expansion. *Journal of Functional Programming*, 5(2):135–154, April 1995.
15. S. C. Kleene and J. B. Rosser. The inconsistency of certain formal logics. *Annals of Mathematics*, 36(3), 1935.
16. Robert K. Meyer, Martin W. Bunder, and Lawrence Powers. Implementing the 'fool's model' of combinatory logic. *Journal of Automated Reasoning*, 7(4):597–630, 1991.
17. Peter O'Hearn and David Pym. The logic of bunched implications. *Bulletin of Symbolic Logic*, 5(2):215–244, 1999.
18. Gordon D. Plotkin. The lambda-calculus is omega-incomplete. *Journal of Symbolic Logic*, 39(2):313–317, 1974.
19. Peter Selinger. Order-incompleteness and finite lambda reduction models. *Theoretical Computer Science*, 309(1):43–63, 2003.
20. Stewart Shapiro. *Foundations without foundationalism: a case for second-order logic*. Number 17 in Oxford logic guides. Oxford University Press, 2000.
21. Johan van Benthem. Higher-order logic. In *Handbook of Philosophical Logic, 2nd Edition*, volume 1, pages 189–244. Kluwer, 2001.
22. Albert Visser. An overview of interpretability logic. In *Mathematical Logic, Proceedings of the 1988 Heyting Conference*, pages 307–359. Plenum Press, 1997.