The Theory of Relations, Complex Terms, and a Connection Between λ and ϵ Calculi

Edward N. Zalta*
Center for the Study of Language and Information
Stanford University
zalta@stanford.edu

Abstract

This paper motivates and introduces a new method of interpreting complex relation terms in a second-order quantified modal language. The new method of interpreting these terms establishes an interesting connection between λ and ϵ calculi, and the resulting semantics provides a precise understanding of the theory of relations. In addition to motivating the new method generally, several research problems in connection with previous, algebraic methods for interpreting complex relation terms are discussed and solved.

Relations are not sets and predication is not set membership. To assert that John loves Mary or that $1 < 2$ is to assert that John bears a certain (two-place) relation, *loves*, to Mary and that 1 bears a certain (two-place) relation, *less than*, to 2. It is not to assert that $\langle \text{John}, \text{Mary} \rangle$ or $\langle 1, 2 \rangle$ is an element of some set. Similarly, to assert that John is happy or that 2 is prime is to assert that John or 2 *has* (or *exemplifies* or *instantiates*) a certain property (i.e., one-place relation), namely, *being happy* or *being prime*. It is not to assert that John or 2 is an element of some set. Yet atomic predications of the form $F^n x_1 \ldots x_n$ (e.g., Rxy or Px) in predicate calculi are standardly modeled and interpreted as claims solely about set membership: n-place predicates of the predicate calculus are standardly interpreted as denoting or signifying sets of n-tuples and the n-place predicates of the modal predicate calculus are standardly interpreted as denoting or signifying functions that map each possible world to a set of n-tuples. Although this standard interpretation allows us to investigate the metatheoretical properties of these calculi in set-theoretic terms, such an interpretation is nevertheless philosophically incorrect. In a philosophically proper interpretation, predicates denote or signify *relations*, not sets or functions from worlds to sets, and if we want to use set theory to represent or model the truth conditions of exemplification claims, relations should play some role in those truth conditions.

In Section 1, I rehearse the *prima facie* case for this last claim and thereby provide general motivation for developing an intensional interpretation of the modal predicate calculus in which the predicates denote relations. Although such intensional interpretations have been proposed before, they give rise to a number of research problems. These are described in Section 2 and the discussion there motivates specific features of the system presented in the main sections of the paper, namely, Sections 3 – 5. This system achieves the research goals implicitly defined by the discussion in Sections 1 and 2, and one of its distinguishing features is that an ϵ-calculus in the metalanguage is used to interpret the λ-calculus in the object language. The system not only provides a better conception of the predications expressed by primitive atomic formulas of the (modal) predicate calculus, but also provides us with a formalism for asserting a precise theory of relations conceived as genuine entities in their own right and not some other thing.

1 General Motivation

To focus our attention, consider a second-order modal language with definite descriptions (i.e., complex individual terms, interpreted rigidly for simplicity) and λ-expressions interpreted relationally rather than functionally (thereby construing them as complex predicates or n-place relation terms). In the traditional interpretation of this language, relations are assumed to be functions from possible worlds to sets of indi-

*I'd like to thank Allen Hazen, Bernard Linsky, Christopher Menzel, Seyed Mousavian, Uri Nodelman, Paul Oppenheimer, and Susanne Riehemann for the insightful comments and observations they made about this paper. I'd like to thank an anonymous referee of this journal for suggesting that I be more explicit about the motivation for the ideas in this paper. I'd also like to thank the members of the logic group at the University of Alberta/Edmonton for reading and discussing the first draft of the paper.

viduals and so the latter are assigned as the semantic values of the simple predicates and λ-expressions. Though this traditional interpretation suffices for the study of the metatheoretical properties of this language, it fails to offer a philosophically proper understanding of the language as a whole, for the following reasons:

- The interpretation represents relations and predication purely in terms of sets and set membership, and so doesn't acknowledge the fact that relations aren't sets and that predication involves those relations.
- The interpretation turns the second-order comprehension principle into a comprehension principle for sets or functions rather than for relations.
- The interpretation doesn't allow us to assert that there are necessarily equivalent but distinct relations.
- The interpretation allows one to mistakenly suppose that the truth value of a sentence changes from world-to-world because the meaning of the sentence changes from world-to-world. This would be a misconception of modal language.
- Relations and predication have 0-place cases (a 0-place relation is a proposition and the 0-place case of predication is just truth). But sets and set membership, and functional application, don't have a 0-place case.

We discuss these in turn.

1.1 Relations are not Sets

The reason that relations are not sets and predication is neither set membership nor functional application is that a relation *characterizes* its arguments, whereas a set merely *collects* its members and a function merely *correlates* its arguments and values. There is a important difference between *characterization*, *collection*, and *correlation*.[1] When one asserts:

[1] In the following remarks, one might substitute the notions of *classify* and *classification* for *collect* and *collection*. However, in some literatures, classification is based on one or more shared common charactertistics and so already presupposes the notion of characterization. The notions of *collect* and *collection* don't carry this presupposition and so better capture the essential difference between a set and its members.

- Bill is human.
- 3 is prime.

one is predicating properties (i.e., 1-place relations) of objects. That is why the logical form of both sentences is Px. One is characterizing Bill and the number 3 as being a certain way, and the expressions “being human” and “being prime” are used to indicate the characterizations. We're not merely collecting Bill and the number 3 into certain sets because, as a mere collection, no set *characterizes* its members in the way that properties characterize the objects of which they're truly predicated. Of course, one can specify a set with the help of predication. When one specifies the set of *humans*, one is appealing to the property of being human and *relying* on predications/characterizations of the form “x is human” (‘Hx’) to identify the members of the set. Without that predication, the set of humans is just a container whose members could be specified with a list. Similarly, the core idea of a *function* is that of a mapping. But to map one thing to another is not to predicate anything of the first thing. In the above, bulleted predications, one is not merely using the property of being human to map Bill to The True, nor merely using the property of being prime to map the number 3 to the True. Such a mapping is a mere correlation. In mapping (associating) a to b, there is no predication going on, only a connection. By contrast, a property characterizes an object in a way that a function does not.

Similar considerations about characterization, collection and correlation apply when we move from one-place atomic predications of the form Px to two-place atomic predications of the form Rxy. Indeed, predicational statements of the form $F^n x_1 \ldots x_n$ (of which Rxy is a 2-place instance) are more fundamental than functional application statements of the form $f(x_1,\ldots,x_{n-1}) = x_n$. The relational form Rxy correctly analyzes predications of the form:

- John loves Mary.
- Russell thought about the number 1.
- $3 < \pi$.

Frege analyzed these sentences in terms of functional application, which he took as basic. Set theorists have, in turn, analyzed functional application and, subsequently, relational predications, in terms of membership and sets of ordered pairs or sets of n-tuples. But functional application and set membership are only mathematical *models* of predication and they *lose* information precisely because neither correlation or collection fully represent characterization, as suggested above. Thus, representing predication by functional application or set membership doesn't capture everything that is asserted when we predicate properties and relations of things.

Indeed, predication is so fundamental that it cannot be analyzed in terms of any other notion; instead, we have to take it as primitive and develop a *theory* of relations and exemplification, just as in set theory we develop a theory of sets and set membership. Unfortunately, philosophers haven't been as quick as set theorists to supply precise, axiomatic theories of relations. But such an axiomatic theory will be used in what follows to *ground* the intensional interpretation of the predicate calculus described in subsequent sections, in the same way that axiomatic set theory grounds extensional interpretations of the calculus, indeed, in the same way that axiomatic set theory grounds the interpretation of the language of set theory.

Note that it is precisely because the predicate calculus is so fundamental in character that we use it to formalize mathematics. In the canonical formulation of set theory, a membership relation is denoted by relation term, so that $x \in y$ is infix notation for a statement of the form Rxy. And though it is often thought that we can either (a) take functions (or sets) as basic and define relations (as Frege did) or (b) take relations as basic and define functions (as Russell did), evidence has emerged recently that suggests relations are more fundamental than functions; relational type theory, at least, can capture some forms of reasoning that functional type theory can't capture (Oppenheimer & Zalta 2011). Thus, time and energy should be spent investigating an interpretation of the predicate calculus in which relations are taken as primitive semantic elements.

The above considerations don't require us to eliminate sets and set membership altogether from the semantics. Rather, it means that we should, at the very least, (a) assign the predicate 'R' a denotation, $\boldsymbol{d}(R)$, in a primitive domain of relations, (b) assign each relation $\boldsymbol{r}$ in the domain of relations an *extension* $\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r})$, that varies from world to world, and (c) assign truth conditions to the formula 'Rab' with respect to a world $\boldsymbol{w}$ as follows: $\langle \boldsymbol{d}(a), \boldsymbol{d}(b)\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{d}(R))$.

1.2 Comprehension Conditions

The standard comprehension schema for relations expressible in a second-order modal language is formulated as:

$$\exists F^n \Box \forall x_1 \ldots \forall x_n (F^n x_1 \ldots x_n \equiv \varphi), \text{ where } \varphi \text{ has no free } F^n\text{s.}$$

We can accept this schema as asserting existence conditions comprehending the domain of relations, the instances of which constitute true statements asserting the existence of relations; e.g., where G and S are free variables:

- $\exists F \Box \forall x (Fx \equiv \neg Gx)$
- $\exists R \Box \forall x \forall y (Rxy \equiv Syx)$
- etc.

Note that under the traditional interpretation of the modal predicate calculus, the schema asserts existence conditions comprehending a domain consisting of sets or functions. But no one would accept the above comprehension schema as a correct theory of sets or functions; we have much better theories about the existence conditions of those entities, such as the existence axioms of Zermelo-Fraenkel set theory.

Thus, in a proper interpretation of the language in which the comprehension schema for relations is couched, the quantifier $\exists F^n$ in the schema should range over a domain of *relations*, conceived as primitive entities in their own right. But that is only a start. Interpretations with a domain of primitive relations don't yet offer any obvious means of interpreting the complex predicates (i.e., the λ-expressions) except by way of the set-theoretic satisfaction conditions of the matrix φ occurring in the expression $[\lambda x_1 \ldots x_n\, \varphi]$. For how are we to relate the primitive property denoted by $[\lambda x\, Px \,\&\, Qx]$ to the primitive properties denoted by the predicates P and Q if not by using set theory to represent the extension of the complex predicate in terms of the extension of the simple ones? This question brings us to the next reason why the traditional interpretation isn't a philosophically correct interpretation.

1.3 Identity Conditions

A philosophically proper theory of relations should allow us to assert that certain necessarily equivalent relations are distinct. To demonstrate this in terms of a canonical example, let's formally represent the following expressions from natural language in the following manner:[2]

- being red and not red: $[\lambda x\, Rx \,\&\, \neg Rx]$
- being a barber who shaves all and only those who don't shave themselves: $[\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)]$

Necessarily, nothing exemplifes either property and so it follows that:

$$\Box\forall z([\lambda x\, Rx \,\&\, \neg Rx]z \equiv [\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)]z)$$

Moreover, we know that the following equation is true:

$$\{x \,|\, Rx \,\&\, \neg Rx\} = \{x \,|\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)\}$$

since both sets are the empty set. Hence, the function that maps every possible world to $\{x \,|\, Rx \,\&\, \neg Rx\}$ is identical to the function that maps every possible world to $\{x \,|\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)\}$.

Clearly, then, we can't use sets or functions from worlds to sets to represent properties and consistently assert the following inequality:

$$[\lambda x\, Rx \,\&\, \neg Rx] \neq [\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)]$$

But it is reasonable to assert the above inequality, for the following reasons:

- Someone (e.g., someone who hasn't thought through the logical implications) could believe that there is a barber who shaves all and only those who don't shave themselves, i.e., believe

  $$\exists z([\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)]z)$$

  without believing that there is a something that both exemplifies and fails to exemplify being red, i.e., without believing:

  $$\exists z([\lambda x\, Rx \,\&\, \neg Rx]z)$$

- The *argument* to the conclusion that some object, say a, fails to exemplify $[\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)]$ is very different from the argument to the conclusion that a fails to exemplify $[\lambda x\, Rx \,\&\, \neg Rx]$; the former argument involves claims about the B property and S relation, while the latter involves claims about the R property and doesn't appeal to any claims about relations. The radical difference in logical roles can be considered evidence that the properties are different.

- One can tell a story about a barber who shaves all and only those who don't shave themselves without telling a story about something which both is and isn't red.

[2]Note that in what follows, λ-expressions will be completely contained within square brackets, so that, where φ is any formula, $[\lambda x\, \varphi]$ is a 1-place relation term and $[\lambda x_1 \ldots x_n\, \varphi]$ is an n-place relation term.

No doubt, one may find all of these reasons controversial, either on the grounds that the data are controversial or on the grounds that the data don't inevitably lead one to the conclusion. The point is only they these reasons provide a *prima facie* case for putting time and energy into developing interpretations of the predicate calculus and relational λ-calculus which don't collapse equivalent relations.

1.4 Avoiding a Misconception

It is a truism that the truth of the sentence 'Snow is white' depends both on the fact that it means that snow is white and on the fact that the substance snow has the property of being white, for the sentence 'Snow is white' would have been false: (a) if our language had been different (e.g., if 'snow is white' had meant that grass is purple), or (b) if our world had been different (e.g., if the substance snow had had the property of being green). When we express ourselves using modal language, we are concerned with the *second* rather than the first alternative. For example, when we ask, "Might snow have failed to be white?", we are not asking about whether the sentence 'Snow is white' could have expressed a falsehood, but rather whether snow might have been some other color. But the traditional interpretations of modal propositional and predicate calculi do not distinguish these two alternatives. The **V** (valuation) function of traditional interpretations maps the atomic sentences of the language to a truth value at each world, and maps each predicate to some set at each world. The **V** function doesn't exclude the suggestion that the reason that $\mathbf{V}(p, \boldsymbol{w}_1) = \mathbf{T}$ and $\mathbf{V}(p, \boldsymbol{w}_2) = \mathbf{F}$ is that the proposition letter 'p' means one thing in $\boldsymbol{w}_1$ and means something else in $\boldsymbol{w}_2$, and moreover, what it means in $\boldsymbol{w}_1$ is true at $\boldsymbol{w}_1$ and what it means in $\boldsymbol{w}_2$ is false at $\boldsymbol{w}_2$. But what an interpretation *should* do is make it clear that 'p' has the same meaning in both w_1 and w_2 and that what it means is true at w_1 and false at w_2.

This is what our **V** function will do in the interpretation developed in Sections 3 – 5, for the 0-place relation term 'p' will simply denote a proposition (i.e., a 0-place relation). The denotation of p is not relative to a world; 'p' simply has a denotation, where the denotation function, relative to an interpretation and assignment to the variables, is not a binary function (with the first argument being a term and the second a possible world) but rather a unary function on terms. In the first instance, relations, not relation terms, will be assigned extensions at each world. Thus, the *extension* of a proposition (i.e., 0-place relation) at each world w is a truth value. A sentence or other 0-place relation term acquires a truth value at each world in virtue of denoting a proposition that has a truth value at each world. So in the interpretations developed below, a change of meaning is not the reason why the truth value of 'p' can vary from world to world: the sentence letter 'p' has only one meaning, namely, the proposition it denotes.

Similarly, traditional interpretations don't exclude the suggestion that the reason why $w_1 \models Pa$ and $w_2 \not\models Pa$ is that both (a) the predicate 'P' means one thing in w_1 and means something else in w_2, and (b) given what 'P' means in w_1, the formula 'Pa' is true at w_1 and, given what 'P' means in w_2, the formula 'Pa' fails to be true at w_2. But what an interpretation *should* do is make it clear that: (a) 'P' has the same meaning in both w_1 and w_2, (b) that its meaning is a certain property, say r^1, and (c) the reason 'Pa' has different truth values at w_1 and w_2 is that a is in the extension of the property r^1 at w_1 but fails to be in the extension of r^1 at w_2. This will be a feature of the interpretation developed in Sections 3 – 5, for the 1-place relation term 'P' will simply denote a 1-place relation r^1 and the *extension* of this property will vary from world to world. A change of meaning cannot be the reason why 'Pa' changes its truth-value from world to world.[3]

1.5 Propositions and Truth

The final problem with traditional interpretations is that they don't naturally accommodate the fact that the 0-place case of the predicational form '$F^n x_1 \ldots x_n$' is simply the 0-place relation term 'F^0' and the 0-place case of exemplification is therefore truth. Hereafter, we use 'p' as shorthand for 'F^0'. A 0-place relation term denotes a 0-place relation, i.e., a proposition. The 0-place term 'p' is therefore both a term and a formula. It is possible to have both expressions of the form '$p = p$', in which 'p' is functioning as a term, as well as expressions of the form '$p \equiv p$', in which in which 'p' is functioning as a formula. When 'p' functions as a formula, we read it as "p is true", and so we read '$p \equiv p$' as: p is true if and only if p is true. That is, just as we use 'exemplification' to read '$F^n x_1 \ldots x_n$' as $x_1 \ldots x_n$ *exemplifies* F^n (when $n \geq 1$), we use *truth* (i.e., the 0-place case of exemplification) to read the formula 'p'.

Now these facts show the inadequacy of the traditional interpretations of the (modal) second-order predicate calculus, for the traditional interpretation of '$F^n x_1 \ldots x_n$' ($n \geq 1$) doesn't generalize in a simple way to the interpretation of 'p' (i.e., F^0). Neither set membership nor functional application has a 0-place case. In the traditional interpretation, to generalize the predicate calculus so as to include 0-place relation terms, one has to interpret these terms by introducing a *different* element into the semantics, namely, a truth-value. For, as we've seen, the predicate 'F' in the formula 'Fx' is typically interpreted as a function from worlds to sets. But how should we interpret 'p'? There is no 0-place case of the membership relation $x \in y$ corresponding to the 0-place case of exemplification. Nor is there a 0-place case of functional application; a function must have an argument and a value; otherwise, there is no mapping. So one has to interpret 'p' as a function from worlds to *truth values*. This introduces a new semantic value into the semantics. It does no good to suggest that the 0-place case of a function is a constant. A constant is a symbol and is the wrong kind of thing to serve as the denotation of the proposition symbol 'p'. Our goal is to interpret p in a domain of appropriate entities, such as sets, functions or propositions. The logician's trick of modeling 0-place functions as constants doesn't achieve this goal; a constant doesn't represent a proposition.

Contrast the above with an interpretation in which the predicates F^n are interpreted as denoting relations. For one simply includes a domain of 0-place relations along with all the other domains of n-place relations for $n \geq 1$. We don't need a different kind of entity to serve as values for the proposition terms. We don't have to add truth values as entities to be denoted or signified by any term of the language.

These facts are reinforced by considerations from type theory: a re-

[3]For a fuller discussion of the issue discussed in this subsection, see Zalta 1993.

lational type theory can get by with just a single type i (individuals) and a complex type $\langle t_1, \ldots, t_n \rangle$. (In the modal case, one often sees an additional primitive type for possible worlds, but we may put this aside for now.) Propositions can then be represented in relational type theory by the empty type $\langle\, \rangle$, in which $n = 0$. But in functional type theory, to represent propositions type-theoretically, one needs to add a new primitive type for propositions or truth-values![4] For example, Church (1940, 56) uses ι (individuals) and o (truth values) as primitive types and $(\alpha\beta)$ as a complex functional type. Montague (1974, 256) uses e (individuals) and t (truth-values) as primitive types, and both $\langle \alpha, \beta \rangle$ and $\langle s, \alpha \rangle$ as complex functional types (here 's' is the type of a possible world). In relational type theory, however, one doesn't need Church's type o or Montague's type t, since the type for propositions comes for free as the 0-place case of the complex type!

This offers another reason why relations are in some sense more fundamental than sets or functions and, hence, why an intensional interpretation of the modal predicate calculus, in which n-place relations ($n \geq 0$) are taken as primitive, deserves investigation.

2 Problems to Solve

Now that we have reasons for developing an interpretation that doesn't misconceive the predications expressed by the atomic formulas of the predicate calculus, it would serve well to examine the problems with relatively recent attempts to build such intensional interpretations. The principal method that has been developed for building intensional interpretations that also interpret the λ-calculus relationally is that of algebraic semantics. After explaining this method (Section 2.1), we go on to outline its the research problems it presents (Section 2.2), and these serve to motivate the new method for building intensional interpretations that we develop in subsequent sections.

2.1 Previous Method: Algebraic Semantics

Philosophical logicians have previously developed intensional, relational interpretations for the predicate calculus with λ-expressions by employing algebraic semantics. Algebraic logical functions, modeled on the

[4] For a full discussion of this issue, see Oppenheimer & Zalta 2011.

predicate functors in Quine 1960 but transferred to the semantics, harness simple properties, relations and propositions into complex ones. Quine (1960) focused only on a simple predicate calculus without complex terms, and showed how the variables might be eliminated. In that paper, for example, Quine introduces (1960, 344) a *derelativization* operator 'Der' on predicates. Quoting Quine: if 'B' is the 2-place predicate 'bites', then 'Der B' is a 1-place predicate of 'biting something'. Generally, Quine defines, (Der $P)x_1 \ldots x_{n-1}$ iff there is something x_n such that $Px_1 \ldots x_n$. So the sentence 'Der Der B' asserts that something bites something, and this eliminates the variables and variable-binding operators in the formula $\exists x \exists y Bxy$. Quine goes on to introduce a variety of other unary predicate-operators: Inv (major inversion), inv (minor inversion), Ref (reflection), and Neg (negation), as well as the binary predicate-operator × (Cartesian multiplication — to handle conjunctions).

By contrast, Bealer 1979, McMichael and Zalta 1980, Bealer 1982, Zalta 1983, and Menzel 1986 recast Quine's predicate functors into the semantics of a language containing complex terms, and in particular, containing λ-expressions interpreted relationally. Instead of Quine's derelativization operator Der on predicates, these authors would introduce into the semantics a *projection* operator on relations, $\mathbf{PROJ}_i$ ($1 \leq i \leq n$), that maps an n-place relation to an $n-1$-place relation. For example, $\mathbf{PROJ}_2$ maps the 2-place relation of *biting* to the property (1-place relation) of *biting something*, while $\mathbf{PROJ}_1$ maps *biting* to the 1-place relation of *being bitten by something*. Thus the numerical index on $\mathbf{PROJ}_i$ doesn't indicate arity but rather indicates the place of the relation being projected onto some object in the domain. Then, the two λ-expressions:

- $[\lambda x\, \exists y Bxy]$: being an x such that x bites something
- $[\lambda x\, \exists y Byx]$: being an x such that something bites x

can be assigned, respectively, a denotation on the basis of the denotation of the 2-place predicate B, as follows. Where $\boldsymbol{d}$ is a denotation function with suppressed relativizations to some interpretation $\mathcal{I}$ of the language and assignment f to the variables:

$$\boldsymbol{d}([\lambda x\, \exists y Bxy]) = \mathbf{PROJ}_2(\boldsymbol{d}(B))$$

$$\boldsymbol{d}([\lambda x\, \exists y Byx]) = \mathbf{PROJ}_1(\boldsymbol{d}(B))$$

A constraint is then placed on the extension of the complex relation at every possible world by introducing a semantic function, $\mathbf{ex}_{\boldsymbol{w}}$, that maps an n-place relation and world $\boldsymbol{w}$ to a set of n-tuples of objects that stand in the relation at that world. In the above examples, the constraint would require:

$$\mathbf{ex}_{\boldsymbol{w}}(\mathbf{PROJ}_2(\boldsymbol{d}(B))) = \{\boldsymbol{o} \mid \exists \boldsymbol{o}'(\langle \boldsymbol{o}, \boldsymbol{o}' \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{d}(B)))\}$$

$$\mathbf{ex}_{\boldsymbol{w}}(\mathbf{PROJ}_1(\boldsymbol{d}(B))) = \{\boldsymbol{o} \mid \exists \boldsymbol{o}'(\langle \boldsymbol{o}', \boldsymbol{o} \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{d}(B)))\}$$

The first of these stipulates that the extension at world $\boldsymbol{w}$ of the property denoted by $[\lambda x\, \exists y Bxy]$ is the set of all objects $\boldsymbol{o}$ such that for some object $\boldsymbol{o}'$, the ordered pair $\langle \boldsymbol{o}, \boldsymbol{o}' \rangle$ is in the extension at $\boldsymbol{w}$ of the property denoted by B. The second stipulates something analogous for the case where the first place of the relation is projected.

The constraint can be formulated generally for $1 \le i \le n$:

$$\mathbf{ex}_{\boldsymbol{w}}(\mathbf{PROJ}_i(\boldsymbol{d}(R^n))) = \{\langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_{i-1}, \boldsymbol{o}_{i+1}, \ldots \boldsymbol{o}_n \rangle \mid \exists \boldsymbol{o}_i(\langle \boldsymbol{o}_1, \ldots \boldsymbol{o}_{i-1}, \boldsymbol{o}_i, \boldsymbol{o}_{i+1}, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{d}(R)))\}$$

Clearly, when an n-place relation is projected in its i-th place, the result is an $n-1$-place relation.

As the previously cited works show, a group of algebraic operations along these lines can be defined; which ones are defined depend on the primitives of the language, of course. In what follows, we'll assume that our language uses negation, conditionals (rather than conjunctions), and the universal quantifier (instead of the existential quantifier) as basic. So, an algebraic interpretation of such a language would include such algebraic logical functions as: **NEG** (negation), **COND** (conditionalization), $\mathbf{REFL}_{i,j}$ (reflection or reflexivization), $\mathbf{CONV}_{i,j}$ (conversion), and $\mathbf{UNIV}_i$ (universalization). These are, roughly, the semantic counterparts to the Quinean syntactic operators. Moreover, such an algebraic interpretation would include:

- $\mathbf{PLUG}_i$, to handle λ-expressions with singular terms such as constants or free variables. Thus, $\boldsymbol{d}([\lambda x\, Rxa])$ is the 1-place property $\mathbf{PLUG}_2(\boldsymbol{d}(R), \boldsymbol{d}(a))$, with the constraint, when $\boldsymbol{r}$ is a 2-place relation:

  $$\mathbf{ex}_{\boldsymbol{w}}(\mathbf{PLUG}_2(\boldsymbol{r}, \boldsymbol{o})) = \{\boldsymbol{o}' \mid \langle \boldsymbol{o}', \boldsymbol{o} \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r})\}.$$

  Note that Quine had no need of a **PLUG** operation on predicates since the point of his paper (1960) was to eliminate singular terms. Quine supposed that a constant could be replaced by a definite description, and that definite descriptions could, in turn, be eliminated in favor of existence and uniqueness claims.

- **NEC**, to handle modality. For example, $\boldsymbol{d}([\lambda x\, \Box Px])$ is the 1-place property $\mathbf{NEC}(\boldsymbol{d}(P))$, with the constraint, when $\boldsymbol{r}$ is a 1-place relation, that

  $$\mathbf{ex}_{\boldsymbol{w}}(\mathbf{NEC}(\boldsymbol{r})) = \{\boldsymbol{o} \mid \forall \boldsymbol{w}'(\boldsymbol{o} \in \mathbf{ex}_{\boldsymbol{w}'}(\boldsymbol{r}))\}.$$

  Quine had no need of a **NEC** operator for predicates either, since he was skeptical of the meaningfulness of modal language and thought it had no place in a proper science.

- $\mathbf{VAC}_i$, to handle variables vacuously bound by the λ. For example, $\boldsymbol{d}([\lambda x\, Pa])$ is the 1-place property $\mathbf{VAC}_1(\mathbf{PLUG}_1(\boldsymbol{d}(P), \boldsymbol{d}(a)))$, with the constraint, when $\boldsymbol{r}$ is a 0-place relation (i.e., a proposition), that

  $$\mathbf{ex}_{\boldsymbol{w}}(\mathbf{VAC}_1(\boldsymbol{r})) = \{\boldsymbol{o} \mid \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}) = \mathbf{True}\}.$$

  Notice how the constraint placed on the world-relative extension of the property $\mathbf{VAC}_1(\boldsymbol{r})$ is vacuous. Quine didn't introduce a similar predicate functor since he wasn't attempting to interpret λ-expressions (which might have vacuously bound variables), but only the sentences of first-order logic without complex terms and without vacuously bound variables.

With this sketch then, we shall henceforth presume familiarity with this previous strand of research, namely, the algebraic technique for giving an intensional interpretation to the (complex) relation terms in a second-order, quantified modal language. It should be clear that such an algebraic interpretation distinguishes the relations from their extensions at a world, and so allows one to assert that necessarily equivalent relations are distinct. Readers interested in more detail can consult the works cited above.

2.2 The Issues

A group of interesting issues arise in connection with the algebraic interpretations just sketched. In what follows, let $\nu_1, \ldots, \nu_n$ be any distinct object variables. and call the formula φ in the λ-expression $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ the *matrix* formula.

> Issue 1: Is there a way to interpret an impredicative λ-expression, i.e., a λ-expression whose matrix formula contains quantifiers binding relation variables?

For example, there is no obvious way to use the algebraic functions described so far to interpret such expressions as $[\lambda x\, \exists F\, Fx]$, $[\lambda xy\, \forall F(Fx \equiv Fy)]$, etc. The reason is that the algebraic functions for producing complex relations all operate on the *argument* places of the original relations, and the constraints on the new relations govern how their extensions among the objects match up with the extensions of the original relations. It is not obvious how to define algebraic functions that allow for quantification over the relations themselves. If we go back to Quine's original syntactic operations, there has been work on eliminating the relation variables in second-order logic (e.g., Došen 1988), and such work might be adaptable to the interpretation of impredicative λ-expressions in a second-order quantified modal logic. But the technique below offers an alternative that is a much simpler way of doing this, and eliminates the algebraic functions altogether.

A second issue is:

> Issue 2: What should one do about the fact that the algebraic logical operations overgenerate relations? How does one define a denotation for a given λ-expression when there are many appropriate relations to choose from? Indeed, what does the algebra consist of: do the algebraic logical operations generate the relations themselves or generate formal objects that represent relations?

To take a simple example, the 0-place term $[\lambda\, \exists x \exists y Bxy]$ could denote either $\mathbf{PROJ}_1(\mathbf{PROJ}_2(\boldsymbol{d}(B)))$ or $\mathbf{PROJ}_1(\mathbf{PROJ}_1(\boldsymbol{d}(B)))$, i.e., either (a) the relation produced by first projecting the 2nd-place of the B relation and then the 1st-place of the resulting relation, or (b) the relation produced by first projecting the 1st-place of the B relation and then the 1st-place of the resulting relation. Indeed, given that we have two perfectly good semantic descriptions of the relation denoted by $[\lambda \exists x \exists y Bxy]$, should we consider these semantic descriptions as directly describing two different relations or as describing two formal objects that represent the same relation? The answer is not obvious.

To see how the possibilities ramify, consider the following example. Suppose the language contains the constant a, the 1-place predicate P, and a 2-place predicate S. Suppose further that the denotation function $\boldsymbol{d}$, relative to some interpretation $\mathcal{I}$ and variable assignment f, assigns to these primitive expressions the values $\boldsymbol{d}(a)$, $\boldsymbol{d}(P)$, and $\boldsymbol{d}(S)$, respectively (again, suppressing the indices to $\mathcal{I}$ and f). Then the algebraic logical functions described or referenced in the previous subsection would generate the following relations in the domain of relations:

$$\mathbf{REFL}_{1,2}(\mathbf{COND}(\boldsymbol{d}(P), \mathbf{PLUG}_1(\boldsymbol{d}(S), \boldsymbol{d}(a))))$$
$$\mathbf{PLUG}_2(\mathbf{REFL}_{1,3}(\mathbf{COND}(\boldsymbol{d}(P), \boldsymbol{d}(S))), \boldsymbol{d}(a))$$
$$\mathbf{REFL}_{1,2}(\mathbf{PLUG}_2(\mathbf{COND}(\boldsymbol{d}(P), \boldsymbol{d}(S)), \boldsymbol{d}(a)))$$

Now consider the expression $[\lambda y\, Py \rightarrow Say]$. The question arises: which of the above relations should be assigned as its denotation? The constraints on the world-relative extensions of the algebraic functions $\mathbf{PLUG}_i$, $\mathbf{REFL}_{i,j}$, and $\mathbf{COND}$ combine to ensure that each of the above relations can serve as the denotation of $[\lambda y\, Py \rightarrow Say]$. That is, the constraints on their world-relative extension functions will ensure each of the above relations has an extension at each possible world that guarantees the truth of λ-conversion:

$$\Box \forall x([\lambda y\, Py \rightarrow Say]x \equiv Px \rightarrow Sax).$$

So how does one assign a denotation to the λ-expression when there are several algebraically-described relations to choose from? Indeed, are we to consider the algebraic descriptions displayed above as directly describing different, but equivalent, relations or as describing several different formal objects that represent the same relation?

A final issue is:

> Issue 3: If definite descriptions are included as primitive terms in the language, then since a λ-expression may contain descriptions that fail to denote, what should the denotation of such a λ-expression be?

Suppose, for example, nothing at the distinguished actual world of an interpretation is uniquely P, so that the description $\imath zPz$ denotes nothing (recall we are interpreting descriptions rigidly—if they fail to denote at the distinguished actual world, they simply have no denotation). Then the denotation of the expression $[\lambda y\, Ry\imath zPz]$ could be assigned in one of two ways: it can be undefined (i.e., denotationless), on the grounds that no appropriate relation can be contructed algebraically if one of the parts of the λ-expression fails to denote, or it could be assigned a necessarily empty property, on the basis of the fact that at each possible world, no object satisfies the formula $Ry\imath zPz$.

In Zalta 1983, 1988a, and elswhere, these three issues were addressed within an algebraic framework as follows. With regard to Issue 1, impredicative complex terms were banished from the language; only predicative complex terms were allowed. With regard to Issue 2, the λ-expressions were partitioned into exhaustive and mutually exclusive syntactic classes.[5] Consider how this partition works in connection with the more complicated example we discussed in connection with this issue. The expression $[\lambda y\, Py \rightarrow Say]$ is syntactically categorized as the 1,2-reflection of the expression $[\lambda yz\, Py \rightarrow Saz]$, which in turn is categorized as the conditionalization of the two expressions $[\lambda y\, Py]$ and $[\lambda z Saz]$. The first of these latter is *elementary* and so denotes $\boldsymbol{d}(P)$, and the second is categorized as the 1st-plugging of the elementary expression $[\lambda yz\, Syz]$ (which denotes $\boldsymbol{d}(S)$) by the term a (which denotes $\boldsymbol{d}(a)$). Consequently, the denotation function assigns

$$\mathbf{REFL}_{1,2}(\mathbf{COND}(\boldsymbol{d}(P), \mathbf{PLUG}_1(\boldsymbol{d}(S), \boldsymbol{d}(a))))$$

to $[\lambda y\, Py \rightarrow Say]$. In this way, the denotation function can, for each complex relation term, single out a unique element of the domain for it to denote. The question, whether the various candidates for the denotation of the λ-expression algebraic description directly describe several different relations or describe several different formal objects that represent the same relation, was left unanswered.

With regard to Issue 3, the assumption was made that if a description doesn't denote, then neither does any complex term containing the description. The logic was appropriately adjusted so that a free logic applied to any formula containing either a description or a λ-expression containing a description, since both kinds of terms might fail to denote.

In the present paper, however, we develop a different approach to all three issues. (1) Impredicative λ-expressions will be allowed and a new approach to the interpretation of both predicative and impredicative λ-expressions.[6] (2) We will *not* use algebraic logical functions to build complex relations, and so no question arises as to whether the algebraic descriptions of complex relations directly describe different relations or describe different formal objects that represent a single relation.

[5]Bealer developed a similar partition in his 1979 and 1982. The partition in Zalta 1983 was developed with the help of Alan McMichael. See also the partition in Menzel 1986.

[6]Impredicative λ-expressions have proven extremely useful in object theory; e.g., Zalta 1999.

Moreover, we don't partition the λ-expressions into classes of expressions having similar logical forms but instead into classes of expressions that differ in logically insignificant ways. (3) We won't assume that λ-expressions containing non-denoting descriptions fail to denote.

To present our new approach to the three issues, we formulate a second-order language with complex terms and then formulate a semantics. The semantics has the interesting feature that it draws a connection between the λ-calculus (Church 1932), interpreted relationally, and the ϵ-calculus (Hilbert 1922, Ackermann 1924). The key differences between the algebraic approach and the new approach developed here are these:

- On the algebraic approach, the algebraic logical functions basically introduce complex *formal* objects that help us to represent complex relations. But on the new approach presented below, no such formal objects are used; the domain of relations consists simply of primitive relations, without any special formal structural representations of those relations. Instead, the relations will be described semantically using ϵ-terms, and a precise theory of relations, statable in the object language, is used to axiomatize the primitive relations.

- On the algebraic method, a unique member of the domain of relations is singled out for a λ-expression to denote. However, on the method presented below, it suffices to identify some appropriate element of the domain of relations to serve as the denotation of the λ-expression. In other words, in order to give a precise technical account of the meaning of the λ-expression, we shall suppose that it suffices to assign it an appropriate meaning.

- On the algebraic approach, during the simultaneous definition of denotation and truth, the denotation of the description $\imath\nu\varphi$ (where ν is any object variable) is assigned on the basis of the truth conditions of φ, while the denotation of the λ-expression $[\lambda\nu_1 \ldots \nu_n\, \varphi]$ $(n \geq 0)$ is assigned on the basis of the *logical form* of the matrix φ, not on the basis of φ's truth conditions. Thus, on the algebraic approach, although $\lambda\nu_1 \ldots \nu_n$ and $\imath\nu$ are both are term-forming operators that combine with formulas φ to produce terms, they are assigned denotations in very different ways. However, on the new approach presented here, both the denotation of $\imath\nu\varphi$ and the denotation of $[\lambda\nu_1 \ldots \nu_n\ \varphi]$ are assigned on the basis of the (simultaneously defined) truth conditions of the matrix formula φ.

- Our new technique has the added advantage that if a λ-expression contains a non-denoting description, it still denotes a relation. Since the truth conditions for the matrix formula φ will be well-defined even if φ contains a non-denoting description, the denotation function will successfully identify a denotation, given certain constraints on interpretations. So one need not deploy a free logic to govern the complex terms containing descriptions; a free logic is needed only for the descriptions themselves.

To summarize, then, the new method (1) will assign denotations even to impredicative λ-expressions, (2) will require no algebraic logical operators but rather a partition of the λ-expressions into classes the members of which differ in logically insignificant ways from each other, and (3) will assign denotations to λ-expressions containing non-denoting descriptions.

We emphasize, finally, that our goal in what follows is simply to give a new formal interpretation of the complex terms of a second-order quantified modal language. It is not an attempt to build a model of any axiom set. It is already known that second-order logic is consistent, that it can be consistently extended by the λ-calculus, and that primitive descriptions can be consistently incorporated into such a logic. So the goal is not to build a model of second-order logic extended by the axioms for the relational λ-calculus and Russell's axiom for descriptions. The goal is to provide a precise intensional interpretation of a second-order quantified modal language with complex terms that both addresses issues (1) – (3) and achieves the implicit goals defined by the discussion in Section 1.

3 The Grammar: A BNF Definition

We may precisely define our target language using Backus-Naur-Form (BNF). In the BNF definition, we use the following metavariables.

δ	individual constants
ν	individual variables
Σ^n	n-place relation constants ($n \geq 0$)
Ω^n	n-place relation variables ($n \geq 0$)
α	variables (i.e., individual or n-place relation variables)
κ	individual terms
Π^n	n-place relation terms ($n \geq 0$)
φ	formulas
τ	terms

The BNF grammar for our language is:

$$
\begin{array}{lrcl}
(n \geq 0) & & & \\
& \delta & ::= & a_0, a_1, \ldots \\
& \nu & ::= & x_0, x_1, \ldots \\
& \Sigma^n & ::= & P^n_0, P^n_1, \ldots \\
& \Omega^n & ::= & F^n_0, F^n_1, \ldots \\
& \alpha & ::= & \nu \mid \Omega^n \\
& \kappa & ::= & \delta \mid \nu \mid \imath\nu\varphi \\
& \Pi^n & ::= & \Sigma^n \mid \Omega^n \mid [\lambda\nu_1 \ldots \nu_n\ \varphi] \\
& \varphi & ::= & \Pi^n\kappa_1 \ldots \kappa_n \mid \neg\varphi \mid (\varphi \rightarrow \varphi) \mid \forall\alpha\varphi \mid \Box\varphi \\
& \tau & ::= & \kappa \mid \Pi^n
\end{array}
$$

Thus, if one defines an instance of our language by listing a finite vocabulary of simple terms and giving a limiting value to n, the sentences of the resulting grammar can be parsed by any appropriately-configured off-the-shelf parsing engine. Note that terms of the form $[\lambda\ \varphi]$ (i.e., 0-place λ-expressions) also become defined as formulas, since they are 0-place relation terms and the base clause for formulas φ stipulates that 0-place relation terms are formulas (i.e., when $n = 0$, the formula $\Pi^n\kappa_1 \ldots \kappa_n$ consists solely of the 0-place relation term Π^0).[7]

In what follows, we sometimes, for readability, substitute simpler-style variables. For example, we'll write the atomic formulas $F^2_1x_1x_2$ and $P^2_1a_1a_2$ as Fxy and Rab, respectively. Also, instead of $F^0_1, F^0_2, \ldots$, we'll use the variables $p, q, \ldots$.

[7] Note that our language is typed: no relation term can appear in argument position. Thus, the λ-expressions can't be predicated of other λ-expressions. These facts, coupled with the fact that there are no mechanisms of self-reference, ensure that no Russell- or Curry-style paradox can be formulated.

4 The Semantics

To keep the semantics as simple as possible, we both: (a) use a fixed-domain interpretation of modal logic without an accessibility relation (thereby assuming S5), and (b) interpret the definite description $\imath\nu\varphi$ as rigidly denoting the object, if there is one, that is uniquely φ at the distinguished actual world, and nothing otherwise.[8] With (b), we avoid the need for a world-relative denotation function. The key idea in what follows is to use an ϵ-calculus in the metalanguage to interpret the λ-calculus in our object language.

4.1 Interpretations

A formal semantic interpretation $\mathcal{I}$ for our second-order quantified modal language with complex terms is a 6-tuple: $\langle \mathbf{D}, \mathbf{R}, \mathbf{W}, \boldsymbol{w}_0, \mathbf{ex}, \mathbf{V}\rangle$, defined as follows:

- A nonempty domain of individuals $\mathbf{D}$.
- A nonempty domain of relations $\mathbf{R}$, where $\mathbf{R} = \bigcup_{n\geq 0} \mathbf{R}_n$, i.e., $\mathbf{R}$ is the general union of a sequence of nonempty subdomains $\mathbf{R}_n$ for $n \geq 0$.
- A nonempty domain of possible worlds, $\mathbf{W}$.
- A distinguished element of $\mathbf{W}$, $\boldsymbol{w}_0$, known as the actual world.
- A binary function, $\mathbf{ex}$, that assigns each n-place relation $\boldsymbol{r}^n$ in $\mathbf{R}$ ($n \geq 0$) an *exemplification* extension at each possible world $\boldsymbol{w}$ as follows:
  - for $n \geq 1$, $\mathbf{ex}$ assigns to each pair consisting of a relation $\boldsymbol{r}^n$ in $\mathbf{R}$ and possible world $\boldsymbol{w}$ in $\mathbf{W}$, a set of n-tuples whose members are in $\mathbf{D}$; i.e., $\mathbf{ex} : \mathbf{R}_n \times \mathbf{W} \to \wp(\mathbf{D}^n)$.
  - for $n = 0$, $\mathbf{ex}$ assigns, to each pair consisting of a relation $\boldsymbol{r}^0$ in $\mathbf{R}$ and possible world $\boldsymbol{w}$ in $\mathbf{W}$, one of the two truth-values **True** or **False**; i.e., $\mathbf{ex} : \mathbf{R}_0 \times \mathbf{W} \to \{\mathbf{True}, \mathbf{False}\}$.

  (We henceforth index $\mathbf{ex}$ to its second argument.) The idea here is that (a) for $n \geq 1$, the n-tuples in the set $\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n)$ represents the various ordered sets of individuals that exemplify $\boldsymbol{r}^n$ at $\boldsymbol{w}$, and (b) for $n = 0$, $\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0)$ is the truth-value of $\boldsymbol{r}^0$ at $\boldsymbol{w}$.
- An interpretation function $\mathbf{V}$ that assigns a meaning to the (primitive) constants of our language:
  - where τ is any individual constant, $\mathbf{V}(\tau)$ is an element of $\mathbf{D}$,
  - where τ is any n-place relation constant ($n \geq 0$), then $\mathbf{V}(\tau)$ is an element of $\mathbf{R}_n$, and so an element of $\mathbf{R}$.

4.2 Variable Assignments

Given any interpretation $\mathcal{I}$, we let an *assignment* function to the variables be a function $f_{\mathcal{I}}$ that maps each individual variable to an element of $\mathbf{D}$ and maps each n-place relation variable ($n \geq 0$) to an element of $\mathbf{R}_n$. Henceforth, we shall suppress the subscript on $f_{\mathcal{I}}$, though the reader should remember that all such assignment functions are defined relative to a given interpretation. For any variable α and entity $\boldsymbol{e}$ in the domain over which α ranges, we may define $f[\alpha/\boldsymbol{e}]$ to be the variable assignment just like f except that it assigns to the variable α the entity $\boldsymbol{e}$.[9] Since we have two kinds of variables, we shall see this definition used in two contexts.

> Context 1: α is an individual variable, and the domain over which α ranges is $\mathbf{D}$. If we are discussing an actual formula in the object-language in which the variable x appears, we use $f[x/\boldsymbol{o}]$ to refer to the assignment just like f except that it assigns to x the object $\boldsymbol{o}$; if we are using the metavariable ν, which ranges over individual variables, to discuss a class of formulas involving ν, we use $f[\nu/\boldsymbol{o}]$

[8] Thus, if one were to develop a logic for such a language, the Russell axiom governing primitive descriptions would be a contingent axiom (Zalta 1988b), and the Rule of Necessitation would have to be configured so as to be applicable only to lines of a proof that don't depend on a contingent axiom or theorem.

[9] This can be defined formally in the usual way, where an assignment function f is represented as a set of ordered pairs, α is a variable, and $\boldsymbol{e}$ is an entity from the domain over which α ranges:

$$f[\alpha/\boldsymbol{e}] = (f \sim \langle \alpha, f(\alpha)\rangle) \cup \{\langle \alpha, \boldsymbol{e}\rangle\}$$

Or, we can define $f[\alpha/\boldsymbol{e}]$ functionally, where β is a variable ranging over the same domain as α, as:

$$f[\alpha/\boldsymbol{e}](\beta) = \begin{cases} f(\beta), & \text{if } \beta \neq \alpha \\ \boldsymbol{e}, & \text{if } \beta = \alpha \end{cases}$$

to refer to the assignment just like f except that it assigns to ν the object $\boldsymbol{o}$.

Context 2: α is an n-place relation variable, and the domain over which α ranges is $\mathbf{R}_n$. Here we typically discuss formulas in the object-language that involve the variable F^n, and so we use $f[F^n/\boldsymbol{r}^n]$ to refer to the assignment just like f except that it assigns the n-place relation $\boldsymbol{r}^n$ (in $\mathbf{R}_n$) to F^n.

Note that we sometimes generalize $f[\alpha/e]$ as follows: when discussing a class of formulas in which the variables $\alpha_1,\ldots,\alpha_n$ may or may not be free, then when the entities $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$ are, respectively, in the domain of $\alpha_1,\ldots,\alpha_n$, we use $f[\alpha_1/\boldsymbol{e}_1,\ldots\alpha_n/\boldsymbol{e}_n]$ to refer to the assignment just like f except that it assigns to α_1 the object $\boldsymbol{e}_1$ and ... and assigns to α_n the object $\boldsymbol{e}_n$. We henceforth abbreviate $f[\alpha_1/\boldsymbol{e}_1,\ldots\alpha_n/\boldsymbol{e}_n]$ as $f[\alpha_i/\boldsymbol{e}_i]$. We leave it to the reader to extend the formal definition of $f[\alpha/\boldsymbol{e}]$ given in footnote 9 to this more general notion.

4.3 Simultaneous Definition of Denotation and World-Relative Truth

The simultaneous definition of denotation and world-relative truth in this section has some distinguishing features that require some preliminary discussion.

First, we deploy a special class of ϵ-terms in our semantic metalanguage in the recursive clauses D5 and D6 below. We call these $\bar{\epsilon}$-terms ("epsilon-bar terms"). The first principles of the metalinguistic calculus of $\bar{\epsilon}$-terms are those of the ϵ-calculus:

- $\bar{\epsilon}$-terms have the form $\bar{\epsilon}\boldsymbol{r}^n(\ldots\boldsymbol{r}^n\ldots)$, where $\boldsymbol{r}^n$ is a variable ranging over the domain $\mathbf{R}_n$ of n-place relations ($n \geq 0$) and $\ldots\boldsymbol{r}^n\ldots$ is a semantic condition expressible in our metalanguage. $\bar{\epsilon}$-terms are denoting terms that pick out some chosen member of the domain $\mathbf{R}_n$ that meets the condition $\ldots\boldsymbol{r}^n\ldots$, if there is one. $\bar{\epsilon}$-terms have *no* existential presuppositions: if nothing meets the condition $\ldots\boldsymbol{r}^n\ldots$, then $\bar{\epsilon}\boldsymbol{r}^n(\ldots\boldsymbol{r}^n\ldots)$ denotes nothing.
- $\bar{\epsilon}$-terms obey the two principles:
  - $\bar{\epsilon}$-**Existence**: If $\exists\boldsymbol{r}(\ldots\boldsymbol{r}\ldots)$, then $\exists\boldsymbol{s}(\boldsymbol{s} = \bar{\epsilon}\boldsymbol{r}(\ldots\boldsymbol{r}\ldots))$
  - $\bar{\epsilon}$-**Conversion**: If $\boldsymbol{s} = \bar{\epsilon}\boldsymbol{r}(\ldots\boldsymbol{r}\ldots)$, then $\ldots\boldsymbol{s}\ldots$
- We do *not* assume a principle of extensionality for $\bar{\epsilon}$-terms, i.e., we do not assume that if $\ldots\boldsymbol{r}^n\ldots$ and $\text{---}\boldsymbol{r}^n\text{---}$ are extensionally equivalent conditions on $\boldsymbol{r}^n$, then $\bar{\epsilon}\boldsymbol{r}^n(\ldots\boldsymbol{r}^n\ldots) = \bar{\epsilon}\boldsymbol{r}^n(\text{---}\boldsymbol{r}^n\text{---})$.

We'll place some further, natural constraints on $\bar{\epsilon}$-terms after we've developed our semantic definitions of truth and denotation, but for now, the above principles suffice.

Second, the two recursive clauses D5 and D6 in the definition below are formulated in terms of three other notions, namely, *alphabetically-variant*, *η-variant*, and *η-irreducible* λ-expressions. Alphabetic variance is well-known, and a full definition need not be developed here. Intuitively the idea is that λ-expressions are alphabetic variants if some sequence of uniform replacements of the bound variables transforms one expression to the other without any variables getting captured by any replacement. Thus, the following pairs of λ-expressions are alphabetic variants:

- $[\lambda x\, Fx]$ and $[\lambda y\, Fy]$
- $[\lambda x\, \forall y Myx]$ and $[\lambda y\, \forall z Mzy]$

In the second pair, the replacement sequence $y \to z, x \to y$ transforms the first member into the second. In clauses D5 and D6 below, a λ-expression and its alphabetic variants are assigned the same denotation.

The notion of η-variant λ-expressions may be unfamiliar in the context of a relational λ-calculus. To define η-variant and η-irreducible expressions, we first say that an n-place λ-expression ($n \geq 1$) is *elementary* iff it has the form $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$, where the matrix formula is an atomic formula of the form $\Pi^n \nu_1 \ldots \nu_n$ in which the arguments to the relation term Π^n are the distinct variables $\nu_1,\ldots,\nu_n$ all of which are bound by the λ and none of which are free in Π^n. (Thus, no λ-expressions of the form $[\lambda\, \varphi]$ are considered elementary.) We say further that the elementary λ-expression $[\lambda \nu_1 \ldots \nu_n \Pi^n \nu_1 \ldots \nu_n]$ is an *η-expansion* of the term Π^n ($n \geq 1$). So, for example,

- $[\lambda xyz\, F^3xyz]$ is an η-expansion of F^3
- $[\lambda xy\, [\lambda uv\, \Box\forall F(Fu \equiv Fv)]xy]$ is an η-expansion of $[\lambda uv\, \Box\forall F(Fu \equiv Fv)]$

Now we say, for $n \geq 0$, that $[\lambda \nu_1 \ldots \nu_n\, \varphi']$ *is an immediate η-variant of* $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ just in case φ' is the result of replacing one n-place relation term Π^n in φ by an η-expansion $[\lambda \nu_1' \ldots \nu_n'\, \Pi^n \nu_1' \ldots \nu_n']$. So, for example,

- $[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$ is an immediate η-variant of $[\lambda y\, Py \rightarrow Say]$
- $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$ is an immediate η-variant of $[\lambda y\, Py \rightarrow Say]$
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ is an immediate η-variant of both $[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$ and $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$
- $[\lambda\, [\lambda z\, Pz]y]$ is an immediate η-variant of $[\lambda\, Py]$

and so on. Now, letting ρ range over n-place λ-expressions ($n \geq 0$), we say that ρ' *is an η-variant of* ρ if there is a sequence of λ-expressions $\rho_1, \ldots, \rho_m$ ($m \geq 1$) with $\rho = \rho_1$ and $\rho' = \rho_m$ such that every member of the sequence is an immediate η-variant of the preceding member of the sequence. So, for example, the sequence

$$[\lambda y\, Py \rightarrow Say]$$
$$[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$$
$$[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$$

establishes that the third is an η-variant of the first.

We may now say that the λ-expression $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ ($n \geq 0$) is *η-irreducible* just in case it is not an η-variant of any other λ-expression. The concepts of η-variant and η-irreducible appear in D5 and D6: non-elementary, η-irreducible λ-expressions and their η-variants will be assigned the same denotation. This completes our preliminary discussion of the special features of our definition, which now proceeds as follows.

If given an interpretation $\mathcal{I}$ and an assignment function f, we simultaneously define:

- $\boldsymbol{d}_{\mathcal{I},f}(\tau)$, i.e., the *denotation* of term τ *relative to* $\mathcal{I}$ *and* f, and
- $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$, i.e., φ is *true at* possible world $\boldsymbol{w}$ *under* $\mathcal{I}$ *and* f,

for all of the terms τ and all of the formulas φ of the language.

The definition has six clauses for denotation D1–D6 and six clauses for truth T1–T6, the base clauses being D1, D2, T1, and T2:

D1. If τ is a constant, then $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \mathbf{V}(\tau)$.

D2. If τ is a variable, then $\boldsymbol{d}_{\mathcal{I},f}(\tau) = f(\tau)$.

T1. If φ is a formula of the form $\Pi^n \kappa_1 \ldots \kappa_n$ ($n \geq 1$), then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n (\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_1)\ \&\ \ldots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$.

T2. If φ is a formula of the form Π^0, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\exists \boldsymbol{r}^0 (\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},f}(\Pi^0)\ \&\ \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$.

T3. If φ is a formula of the form $\neg\psi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if it is not the case that $\boldsymbol{w} \models_{\mathcal{I},f} \psi$, i.e., iff $\boldsymbol{w} \not\models_{\mathcal{I},f} \psi$.

T4. If φ is a formula of the form $\psi \rightarrow \chi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only either $\boldsymbol{w} \not\models_{\mathcal{I},f} \psi$ or $\boldsymbol{w} \models_{\mathcal{I},f} \chi$, i.e., iff $\boldsymbol{w} \not\models_{\mathcal{I},f} \psi \vee \boldsymbol{w} \models_{\mathcal{I},f} \chi$.

T5. If φ is a formula of the form $\forall \alpha \psi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if for every entity $\boldsymbol{e}$ in the domain over which α ranges, ψ is true at $\boldsymbol{w}$ under $\mathcal{I}$ and $f[\alpha/\boldsymbol{e}]$, i.e., iff $\forall \boldsymbol{e} \in \mathrm{dom}(\alpha)(\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \psi)$.

T6. If φ is a formula of the form $\Box\psi$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if for every possible world $\boldsymbol{w}'$, $\boldsymbol{w}' \models_{\mathcal{I},f} \psi$, i.e., iff $\forall \boldsymbol{w}' (\boldsymbol{w}' \models_{\mathcal{I},f} \psi)$.

D3. If τ is a description of the form $\imath\nu\varphi$, then

$$\boldsymbol{d}_{\mathcal{I},f}(\tau) = \begin{cases} \boldsymbol{o}, \text{ if } \boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}]} \varphi\ \&\ \forall \boldsymbol{o}' (\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}']} \varphi \rightarrow \boldsymbol{o}' = \boldsymbol{o}) \\ \text{undefined, otherwise} \end{cases}$$

D4. If τ is an *elementary* λ-expression $[\lambda \nu_1 \ldots \nu_n \Pi^n \nu_1 \ldots \nu_n]$, where $n \geq 1$, then:

$$\boldsymbol{d}_{\mathcal{I},f}(\tau) = \begin{cases} \boldsymbol{d}_{\mathcal{I},f}(\Pi^n), \text{ if } \boldsymbol{d}_{\mathcal{I},f}(\Pi^n) \text{ is defined} \\ \text{undefined, otherwise} \end{cases}$$

D5. If τ is a non-elementary, η-irreducible λ-expression $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ ($n \geq 1$) and τ' is any λ-expression that is either an alphabetic variant or an η-variant of τ, then:

$$\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f}(\tau') =$$

$$\begin{cases} \bar{\epsilon} \boldsymbol{r}^n \forall \boldsymbol{w} \forall \boldsymbol{o}_1 \ldots \forall \boldsymbol{o}_n (\langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n) \equiv \boldsymbol{w} \models_{\mathcal{I},f[\nu_i/\boldsymbol{o}_i]} \varphi), \\ \quad \text{if there is one} \\ \text{undefined, otherwise} \end{cases}$$

D6. If τ is an η-irreducible 0-place λ-expression $[\lambda\,\varphi]$ and τ' is any λ-expression that is either an alphabetic variant or an η-variant of τ, then:

$$\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f}(\tau') = \begin{cases} \bar{\epsilon}\boldsymbol{r}^0 \forall \boldsymbol{w}(\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True} \equiv \boldsymbol{w} \models_{\mathcal{I},f} \varphi), \\ \quad \text{if there is one} \\ \text{undefined, otherwise} \end{cases}$$

Several remarks about these clauses are in order.

Given an interpretation $\mathcal{I}$ and assignment f, D1 says that the denotation of a primitive constant is what **V** assigns that constant, and D2 says that the denotation of a variable is what f assigns to that variable. T1 tells us that an exemplification formula is true at world $\boldsymbol{w}$ iff all of the terms in the formula denote and the n-tuple of individuals denoted by the individual terms is an element of the exemplification extension at $\boldsymbol{w}$ of the relation denoted by the relation term. T2 tells us that a formula consisting of a 0-place relation term is true at $\boldsymbol{w}$ iff the term denotes a proposition and the exemplification extension at $\boldsymbol{w}$ of the proposition denoted is the truth-value **True**. T3–T6 are all standard. Speaking loosely, in terms of of an object $\boldsymbol{o}$ satisfying$_{\mathcal{I},f}$ a formula φ at $\boldsymbol{w}$,[10] D3 identifies the denotation$_{\mathcal{I},f}$ of $\imath\nu\varphi$ as an object $\boldsymbol{o}$ if $\boldsymbol{o}$ uniquely satisfies$_{\mathcal{I},f}$ φ at $\boldsymbol{w}_0$, and is undefined otherwise. D4 identifies the denotation$_{\mathcal{I},f}$ of an elementary λ-expression as the relation denoted by the relation term, if there is one, and undefined otherwise. Speaking loosely, in terms of n-tuples of objects satisfying$_{\mathcal{I},f}$ a formula φ at $\boldsymbol{w}$,[11] we may regard D5 as identifying the denotation$_{\mathcal{I},f}$ of a non-elementary, η-irreducible λ-expression $[\lambda\nu_1\ldots\nu_n\,\varphi]$ $(n \geq 1)$ as some chosen relation $\boldsymbol{r}^n$ (if there is one) that, in every world $\boldsymbol{w}$, has an exemplification extension at $\boldsymbol{w}$ that consists of precisely those n-tuples of objects that satisfy$_{\mathcal{I},f}$ φ at $\boldsymbol{w}$. Finally, D6 identifies the denotation$_{\mathcal{I},f}$ of a non-elementary, η-irreducible 0-place λ-expression $[\lambda\varphi]$ as some chosen proposition $\boldsymbol{r}^0$ (if there is one) that has **True** as its exemplification extension at $\boldsymbol{w}$ iff φ is true at $\boldsymbol{w}$.

It is also important to note the following about this definition. First, T1 assigns well-defined, binary truth conditions to atomic formulas even

[10]We may define this precisely for φ in which ν may or may not be free: an object $\boldsymbol{o}$ *satisfies*$_{\mathcal{I},f}$ φ *at* world $\boldsymbol{w}$ just in case $\boldsymbol{w} \models_{\mathcal{I},f[\nu/\boldsymbol{o}]} \varphi$. An object $\boldsymbol{o}$ *uniquely* satisfies$_{\mathcal{I},f}$ φ just in case it satisfies$_{\mathcal{I},f}$ φ at $\boldsymbol{w}$ and any other object that satisfies$_{\mathcal{I},f}$ φ at $\boldsymbol{w}$ is identical to $\boldsymbol{o}$.

[11]We can extend the definition in the previous footnote for φ where $\nu_1,\ldots,\nu_n$ may or may not be free: an n-tuple $\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle$ *satisfies*$_{\mathcal{I},f}$ φ *at* world $\boldsymbol{w}$ just in case $\boldsymbol{w} \models_{\mathcal{I},f[\nu_i/\boldsymbol{o}_i]} \varphi$.

if the formula happens to contain a non-denoting term (in such a case, the atomic formula is false). Given T1–T6, even if an atomic formula with a non-denoting term appears as a subformula of φ in $[\lambda\nu_1\ldots\nu_n\,\varphi]$, the whole formula φ has binary truth conditions. Thus, D5 and D6 may assign to a λ-expression a denotation picked out by a semantic $\bar{\epsilon}$-term even if the matrix formula φ contains a definite description that denotes nothing. To guarantee that there is a relation picked out by the $\bar{\epsilon}$-term, we need to add one further constraint to our definition of an interpretation. We will do this in Section 4.5. Second, D4 governs not only those cases of elementary λ-expression such as $[\lambda x Px]$, where the relation term Π^n is a simple term, but also governs cases of elementary expressions such as $[\lambda x\,[\lambda y\,\varphi]x]$, where the relation term Π^n is $[\lambda y\,\varphi]$ and hence complex. Third, in D5 and D6, we've deployed Hilbert-style $\bar{\epsilon}$-terms to interpret the λ-expressions and this forges an interesting connection between λ and $\bar{\epsilon}$ calculi. Finally, note that in D3, D5, and D6, $\boldsymbol{d}_{\mathcal{I},f}(\tau)$ recursively calls $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$, and so the method of assigning denotations to the two types of complex terms, descriptions and λ-expressions, is uniform.

4.4 Truth and Validity

In the usual way, we say φ is *true under* $\mathcal{I}$ *and* f just in case φ is true at the distinguished actual world $\boldsymbol{w}_0$ under $\mathcal{I}$ and f. Formally:

$$\models_{\mathcal{I},f} \varphi \ =_{df} \ \boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi$$

We say φ is *true under* $\mathcal{I}$ just in case for every f, φ is true under $\mathcal{I}$ and f. Formally:

$$\models_{\mathcal{I}} \varphi \ =_{df} \ \forall f(\models_{\mathcal{I},f} \varphi)$$

We say φ is *satisfiable* if and only if there is some interpretation $\mathcal{I}$ and assignment f such that φ is true$_{\mathcal{I},f}$, i.e., iff $\exists\mathcal{I}\exists f(\models_{\mathcal{I},f} \varphi)$. We say φ is *valid* or *logically true* if and only if φ is true under every interpretation $\mathcal{I}$. Formally:

$$\models \varphi \ =_{df} \ \forall\mathcal{I}(\models_{\mathcal{I}} \varphi)$$

The other semantics definitions, of logical consequence, logical equivalence, etc., like the foregoing ones, are all standard.[12]

[12]Thus, in the usual way, φ *logically implies* ψ $(\varphi \models \psi)$ just in case, for every interpretation $\mathcal{I}$ and assignment f, if φ is true$_{\mathcal{I},f}$, then ψ is true$_{\mathcal{I},f}$. Also, in the usual way, φ and ψ are

4.5 A Special Class of Semantic Structures

Our semantics now consists of the following, defined notions:

- interpretations $\mathcal{I}$
- assignments f
- primitive $\bar{\epsilon}$-terms of the form $\bar{\epsilon}\boldsymbol{r}^n(\ldots\boldsymbol{r}^n\ldots)$
- the denotation of τ with respect to $\mathcal{I}$ and f
- formula φ is true at $\boldsymbol{w}$ under $\mathcal{I}$ and f
- φ is true under $\mathcal{I}$ and f
- φ is true under $\mathcal{I}$
- φ is valid

With these definitions in place, we conclude our model-theoretic semantics for complex terms by placing constraints on interpretations, assignments, $\bar{\epsilon}$-terms, denotation functions and truth conditions. These constraints are, in effect, are semantic axioms that force these semantic notions to have the structure needed to interpret our λ-expressions. The constraints on the $\bar{\epsilon}$-terms requires them to make certain choices in certain well-defined situations, and forces them to behave in nice logical ways that the classical ϵ-terms of the ϵ-calculus wouldn't otherwise behave:

Constraint (1)
Interpretations, Assignments, and Truth.
(.1) For any formula φ and assignment f:

$$\exists \boldsymbol{r}^n \forall \boldsymbol{w} \forall \boldsymbol{o}_1 \ldots \forall \boldsymbol{o}_n (\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n) \equiv \boldsymbol{w} \models_{\mathcal{I},f} \varphi) \quad (n \geq 1)$$

(.2) For any formula φ and assignment f:

$$\exists \boldsymbol{r}^0 \forall \boldsymbol{w} (\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True} \equiv \boldsymbol{w} \models_{\mathcal{I},f} \varphi)$$

Constraint (2)
Interpretations, Assignments, $\bar{\epsilon}$-terms, Denotations and Truth:
(.1) For any non-elementary λ-expression $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ $(n \geq 1)$, if assignments f and g agree on all the variables free in $[\lambda \nu_1 \ldots \nu_n\, \varphi]$, then:

$$\bar{\epsilon}\boldsymbol{r}^n \forall \boldsymbol{w} \forall \boldsymbol{o}_1 \ldots \forall \boldsymbol{o}_n (\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n) \equiv \boldsymbol{w} \models_{\mathcal{I},f[\nu_i/\boldsymbol{o}_i]} \varphi) =$$
$$\bar{\epsilon}\boldsymbol{r}^n \forall \boldsymbol{w} \forall \boldsymbol{o}_1 \ldots \forall \boldsymbol{o}_n (\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n) \equiv \boldsymbol{w} \models_{\mathcal{I},g[\nu_i/\boldsymbol{o}_i]} \varphi)$$

and hence, by D5:

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda \nu_1 \ldots \nu_n\, \varphi]) = \boldsymbol{d}_{\mathcal{I},g}([\lambda \nu_1 \ldots \nu_n\, \varphi])$$

(.2) For any λ-expression $[\lambda\, \varphi]$, if assignments f and g agree on all the variables free in $[\lambda\, \varphi]$, then:

$$\bar{\epsilon}\boldsymbol{r}^0 \forall \boldsymbol{w} (\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True} \equiv \boldsymbol{w} \models_{\mathcal{I},f} \varphi) =$$
$$\bar{\epsilon}\boldsymbol{r}^0 \forall \boldsymbol{w} (\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True} \equiv \boldsymbol{w} \models_{\mathcal{I},g} \varphi)$$

and hence, by D6:

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda\, \varphi]) = \boldsymbol{d}_{\mathcal{I},g}([\lambda\, \varphi])$$

Constraint (3)
Interpretations, Assignments, $\bar{\epsilon}$-terms, Denotations and Truth:
(.1) For any non-elementary λ-expression $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ $(n \geq 1)$, if $\tau_1,\ldots,\tau_m$ are, respectively, substitutable for $\alpha_1,\ldots,\alpha_m$ in $[\lambda \nu_1 \ldots \nu_n \varphi]$ and $\boldsymbol{d}_{\mathcal{I},f}(\tau_1) = \boldsymbol{e}_1$ and ... and $\boldsymbol{d}_{\mathcal{I},f}(\tau_m) = \boldsymbol{e}_m$, where $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_m$ are entities in the domain of $\alpha_1,\ldots,\alpha_m$, respectively, then:

$$\bar{\epsilon}\boldsymbol{r}^n \forall \boldsymbol{w} \forall \boldsymbol{o}_1 \ldots \forall \boldsymbol{o}_n (\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n) \equiv \boldsymbol{w} \models_{\mathcal{I},f[\nu_i/\boldsymbol{o}_i]} \varphi^{\tau_1,\ldots\tau_m}_{\alpha_1,\ldots\alpha_m})$$
$$=$$
$$\bar{\epsilon}\boldsymbol{r}^0 \forall \boldsymbol{w} \forall \boldsymbol{o}_1 \ldots \forall \boldsymbol{o}_n (\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n) \equiv \boldsymbol{w} \models_{\mathcal{I},f[\alpha_i/\boldsymbol{e}_i][\nu_i/\boldsymbol{o}_i]} \varphi)$$

and hence, by D5:

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda \nu_1 \ldots \nu_n\, \varphi^{\tau_1,\ldots\tau_m}_{\alpha_1,\ldots\alpha_m}]) = \boldsymbol{d}_{\mathcal{I},f[\alpha_i/\boldsymbol{e}_i]}([\lambda \nu_1 \ldots \nu_n\, \varphi])$$

(.2) For any λ-expression $[\lambda\varphi]$, if $\tau_1,\ldots,\tau_m$ are, respectively, substitutable for $\alpha_1,\ldots,\alpha_m$ in $[\lambda\,\varphi]$ and $\boldsymbol{d}_{\mathcal{I},f}(\tau_1) = \boldsymbol{e}_1$ and ... and $\boldsymbol{d}_{\mathcal{I},f}(\tau_m)$ $= \boldsymbol{e}_m$, where $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_m$ are entities in the domain of $\alpha_1,\ldots,\alpha_m$, respectively, then:

$$\bar{\epsilon}\boldsymbol{r}^0 \forall \boldsymbol{w} (\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True} \equiv \boldsymbol{w} \models_{\mathcal{I},f} \varphi^{\tau_1,\ldots\tau_m}_{\alpha_1,\ldots\alpha_m}) =$$
$$\bar{\epsilon}\boldsymbol{r}^0 \forall \boldsymbol{w} (\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True} \equiv \boldsymbol{w} \models_{\mathcal{I},f[\alpha_i/\boldsymbol{e}_i]} \varphi)$$

and hence, by D6:

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda\, \varphi^{\tau_1,\ldots\tau_m}_{\alpha_1,\ldots\alpha_m}]) = \boldsymbol{d}_{\mathcal{I},f[\alpha_i/\boldsymbol{e}_i]}([\lambda\, \varphi])$$

logically equivalent just in case both $\varphi \models \psi$ and $\psi \models \varphi$. Finally, φ is a *logical consequence* of a set of formulas Γ just in case, for every interpretation $\mathcal{I}$ and assignment f, if every member of Γ is true$_{\mathcal{I},f}$, then φ is true$_{\mathcal{I},f}$.

We explain these constraints in turn.

Clearly, interpretations and assignments that obey Constraints (1.1) and (1.2) *validate* the following Comprehension Schema for Relations ($n \geq 0$):

Comprehension Schema for Relations (RC):
$\exists F^n \Box \forall x_1 \ldots \forall x_n (F^n x_1 \ldots x_n \equiv \varphi)$, where $n \geq 0$ and F^n is not free in φ

So by restricting our attention to those interpretations that satisfy these Constraints (1.1) and (1.2), we are assured that each non-elementary λ-expression has a denotation.[13] For by the properties of $\bar{\epsilon}$-terms and D5, if the term:

$$\bar{\epsilon}\boldsymbol{r}^{\boldsymbol{n}} \forall \boldsymbol{w} \forall \boldsymbol{o}_1 \ldots \forall \boldsymbol{o}_n (\langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^{\boldsymbol{n}}) \equiv \boldsymbol{w} \models_{\mathcal{I}, f[\nu_i/\boldsymbol{o}_i]} \varphi)$$

is to pick out a relation for the non-elementary λ-expression $[\lambda \nu_1 \ldots \nu_n \varphi]$ ($n \geq 1$) to denote, then it must be the case that:

$$(\vartheta)\ \exists \boldsymbol{r}^{\boldsymbol{n}} \forall \boldsymbol{w} \forall \boldsymbol{o}_1 \ldots \forall \boldsymbol{o}_n (\langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^{\boldsymbol{n}}) \equiv \boldsymbol{w} \models_{\mathcal{I}, f[\nu_i/\boldsymbol{o}_i]} \varphi)$$

But (ϑ) is true in interpretations satisfying Constraint (1.1), since we can instantiate f to $f[\nu_i/\boldsymbol{o}_i]$.[14] Similarly, by the properties of $\bar{\epsilon}$-terms and D6, if the term:

$$\bar{\epsilon}\boldsymbol{r}^0 \forall \boldsymbol{w}(\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True} \equiv \boldsymbol{w} \models_{\mathcal{I}, f} \varphi)$$

is to pick out a proposition for the non-elementary λ-expression $[\lambda \varphi]$ to denote, then it must be the case that:

$$(\zeta)\ \exists \boldsymbol{r}^0 \forall \boldsymbol{w}(\mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True} \equiv \boldsymbol{w} \models_{\mathcal{I}, f} \varphi)$$

But (ζ) just is Constraint (1.2).

Constraints (2.1) and (2.2) constitute a *restricted* principle of extensionality for $\bar{\epsilon}$-terms.[15] Semantic structures that obey these constraints have a nice property: when the clauses D5 and D6 pick out a denotation for the expression $[\lambda \nu_1 \ldots \nu_n \varphi]$ ($n \geq 0$) with respect to a variable assignment f, then the same denotation is picked out with respect to any variable assignment g that agrees with f on the free variables in $[\lambda \nu_1 \ldots \nu_n \varphi]$. The $\bar{\epsilon}$-terms in D5 and D6 are thus required to choose the same relation as the denotation of $[\lambda \nu_1 \ldots \nu_n \varphi]$ with respect to both f and g.[16] So, for example, consider the λ-expression $[\lambda x\, Rxy]$ ($= \tau$). If f and g both assign to the variable y the object $\boldsymbol{o}$, then $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$. To give another example in connection these assignments f and g, the denotation of 0-place expression $[\lambda\, Py]$ with respect to f and g is now required to be the same. This is one way in which our $\bar{\epsilon}$-terms behave with less freedom than classical ϵ-terms.

Constraints (3.1) and (3.2) also constitute a restricted principle of extensionality for $\bar{\epsilon}$-terms, and like (2.1) and (2.2), identify the relations picked out by $\bar{\epsilon}$-terms only when logically irrelevant conditions obtain. Semantic structures that obey these two constraints have a further nice property, which we can see by way of example. Consider the two λ-expressions $[\lambda x\, Rxa]$ and $[\lambda x\, Rxy]$. Suppose that $\boldsymbol{d}_{\mathcal{I},f}(a) = \boldsymbol{o}$. Then Constraint (3.1) requires, for the assignments f and $f[y/\boldsymbol{o}]$, that:

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda x\, Rxa]) = \boldsymbol{d}_{\mathcal{I},f[y/\boldsymbol{o}]}([\lambda x\, Rxy])$$

This is a second way in which our $\bar{\epsilon}$-terms behave with less freedom than classical ϵ-terms. Note, however, that our $\bar{\epsilon}$-terms are still free to choose different denotations for logically equivalent expressions such as $[\lambda x\, Rx \,\&\, \neg Rx]$ and $[\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)]$ (that is, in absence of any axioms in the object language that might assert the identity or distinctness of such properties). Semantic structures that obey Constraints (3.1) and (3.2) in addition to Constraints (2.1) and (2.2) are ones for which the classical Substitution Lemma holds not only for formulas but also for all terms.[17]

[13] Elementary λ-expressions will then be guaranteed to have denotations by D4, no matter whether the relation term Π^n in $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ is simple or complex, and if complex, elementary or non-elementary.

[14] Note how Constraints (1.1) and the body of the $\bar{\epsilon}$-term in clause D5 resemble one another. Constraint (1.1) uses an existential quantifier to assert the existence of relations meeting certain conditions, for every assignment f; clause D5 uses a $\bar{\epsilon}$-term to pick out a relation that meets those same conditions relative to an assignment that applies to the variables bound by the λ.

[15] In the classical ϵ-calculus, the principle of extensionality takes the form: $\forall x(\varphi \equiv \psi) \to \epsilon x\varphi = \epsilon x\psi$ (Avigad and Zach 2013). This asserts that if φ and ψ are materially equivalent, then $\epsilon x\varphi = \epsilon x\psi$ (when φ, ψ are empty, the ϵ-terms denote the single junk element). Constraints (2.1) and (2.2) don't quite take this form, but they do assert an $\bar{\epsilon}$-identity whenever a kind of equivalence holds, namely, when $f(\alpha) = g(\alpha)$ for every free α in $[\lambda \nu_1 \ldots \nu_n \varphi]$.

[16] The Assignment Agreement Lemma, which asserts that if f and g agree on all the free variables in formula φ that $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ iff $\boldsymbol{w} \models_{\mathcal{I},g} \varphi$, can be proved without this constraint. However, without the constraint, the Corollary to the Assignment Agreement Lemma, which asserts that if f and g agree on all the free variables in term τ that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$, won't be derivable for λ-expressions. Since this Corollary is needed at various points in the proof of the Substitution Lemma, Constraint (2) is included so as to guarantee that the Corollary holds for all terms. The proof of the Assignment Agreement Lemma and its Corollary appear in the Appendix.

[17] The Substitution Lemma asserts:

4.6 Returning to Our Examples

We can now see how the problematic issues and examples in Section 2.2 fare under our new semantics. Consider $[\lambda x\, \exists F Fx]$, an example used to illustrate Issue 1. D5 yields that:

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda x \exists F Fx]) = \begin{cases} \bar{\epsilon}\boldsymbol{r}^1 \forall \boldsymbol{w} \forall \boldsymbol{o}(\boldsymbol{o} \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^1) \;\equiv\; \boldsymbol{w} \models_{\mathcal{I},f[x/\boldsymbol{o}]} \exists F Fx), \\ \quad \text{if there is one} \\ \text{undefined, otherwise} \end{cases}$$

Since there are no descriptions involved in the formula and our constraints on interpretations guarantee that there is such a relation, we may work through the semantics definitions to conclude: the denotation$_{\mathcal{I},f}$ of $[\lambda x\, \exists F Fx]$ is a 1-place relation $\boldsymbol{r}^1$ whose exemplification extension at a world $\boldsymbol{w}$ consists of all and only those objects $\boldsymbol{o}$ such that $\boldsymbol{o}$ satisfies$_{\mathcal{I},f}$ $\exists F Fx$ at $\boldsymbol{w}$, i.e., such that some 1-place relation $\boldsymbol{s}^1$ is such that $\boldsymbol{o} \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{s}^1)$. Thus, we've interpreted the impredicative λ-expression without having to formulate special algebraic logical functions.

Consider the more complex of the two examples used to illustrate Issue 2. D5 yields that:

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda x Px \rightarrow Sax]) = \begin{cases} \bar{\epsilon}\boldsymbol{r}^1 \forall \boldsymbol{w} \forall \boldsymbol{o}(\boldsymbol{o} \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^1) \;\equiv\; \boldsymbol{w} \models_{\mathcal{I},f[x/\boldsymbol{o}]} Px \rightarrow Sax)), \\ \quad \text{if there is one} \\ \text{undefined, otherwise} \end{cases}$$

Again, since there are no descriptions involved in the formula and our constraints on interpretations guarantee that there is such a relation, we may conclude that the denotation is a relation $\boldsymbol{r}^1$ whose exemplification extension at a world $\boldsymbol{w}$ consists of all and only those objects $\boldsymbol{o}$ such that either $\boldsymbol{o}$ is not in the extension$_{\boldsymbol{w}}$ of $\boldsymbol{d}_{\mathcal{I},f}(P)$ or the pair $\langle \boldsymbol{d}_{\mathcal{I},f}(a), \boldsymbol{o}\rangle$ is in the extension$_{\boldsymbol{w}}$ of $\boldsymbol{d}_{\mathcal{I},f}(S)$. Thus, we've interpreted expressions like $[\lambda x Px \rightarrow Sax]$ without having to partition the λ-expressions into syntactic classes that correspond to the logical components of their syntax. Instead, we partition them only with respect to logically-insignificant variations (i.e., we collapse all the alphabetic- and η-variants).

Consider, finally, the example used to illustrate Issue 3, $[\lambda x\, Rx\imath yPy]$, in an interpretation in which $\imath yPy$ fails to denote. D5 yields that:

If τ is substitutable for α in φ and $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{e}$, where $\boldsymbol{e}$ is an entity in the domain of the variable α, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi^{\tau}_{\alpha}$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$.

The proof of this Lemma also appears in the Appendix.

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda x Rx\imath yPy]) = \begin{cases} \bar{\epsilon}\boldsymbol{r}^1 \forall \boldsymbol{w} \forall \boldsymbol{o}(\boldsymbol{o} \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^1) \;\equiv\; \boldsymbol{w} \models_{\mathcal{I},f[x/\boldsymbol{o}]} Rx\imath yPy), \\ \quad \text{if there is one} \\ \text{undefined, otherwise} \end{cases}$$

Again, our constraints on interpretations guarantee that there is such a relation, so the above picks out a relation $\boldsymbol{r}^1$ whose exemplification extension at a world $\boldsymbol{w}$ consists of all and only those objects $\boldsymbol{o}$ such that (i) there are objects and properties that are denoted by R, x, and $\imath yPy$, and (ii) the ordered pair consisting of $\boldsymbol{o}$ and $\boldsymbol{d}_{\mathcal{I},f[x/\boldsymbol{o}]}(\imath yPy)$ is in the extension$_{\boldsymbol{w}}$ of $\boldsymbol{d}_{\mathcal{I},f[x/\boldsymbol{o}]}(R)$. Since there is no object denoted by $\imath yPy$, no object satisifies condition (i), and hence no objects satisfies (i) and (ii). So we've specified a relation $\boldsymbol{r}^1$ whose exemplification extension is empty at every world $\boldsymbol{w}$. $[\lambda x\, Rx\imath yPy]$, therefore, denotes a necessarily unexemplified relation.

In light of this last example, we leave it as an exercise to show that our definitions do indeed imply the following with respect to two other λ-expressions containing the non-denoting description $\imath yPy$: (a) the denotation of $[\lambda x \neg Rx\imath yPy]$ is a necessarily exemplified relation, and (b) the denotation of $[\lambda x\, Rx\imath yPy \vee Qx]$ is a relation that is necessarily equivalent to the denotation of Q.

5 Observations and Further Developments

Our semantics for complex relation terms using metalinguistic $\bar{\epsilon}$-terms therefore addresses Issues (1) – (3). It gives us an interpretation for impredicative λ-expressions; it requires only a partition of the λ-expressions into classes of expressions that differ in logically-insignificant ways; and it assigns denotations to λ-expressions containing non-denoting descriptions.[18] The price of admission is only that we have to place three constraints on our semantic definitions and primitive $\bar{\epsilon}$-terms.

[18] Notice that none of the issues discussed in Hazen 2012 (60–62), in connection with the σ choice functions used to model counterpart functions, arise in connection with our metalinguistic $\bar{\epsilon}$-terms: we deploy those terms in a highly circumscribed and controlled environment, namely, a flat domain of relations. By contrast, the σ-terms discussed by Hazen are deployed throughout the ZF hierarchy of sets and, as such, have to be interpreted in terms of *class* functions.

5.1 β-Conversion and Descriptions

The principle of β-Conversion in the relational λ-calculus is formulated as follows:[19]

> β-**Conversion**:
> $[\lambda y_1 \ldots y_n\, \varphi]x_1 \ldots x_n \equiv \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$, provided $x_1,\ldots,x_n$ are substitutable, respectively, for $y_1,\ldots,y_n$ in φ

Note that this applies even to λ-expressions in which the matrix formula is impredicative or contains descriptions (we'll discuss the latter below).[20]

One additional advantage of the new semantic method is that we no longer have to add as axioms the special instances of **RC** involving descriptions. To see why, note that under the algebraic interpretation, β-Conversion has to be restricted. The restriction that is needed in the algebraic approach is that φ must be free of descriptions. For otherwise, we would have the following instance:

$$[\lambda y\, \neg Ry\imath zPz]x \equiv \neg Rx\imath zPz$$

This would be false in any interpretation in which $\imath zPz$ fails to denote, since at any world you pick and for any object x you pick, the left side of the biconditional is false at that world (given that it is an atomic formula with a non-denoting λ-expression) while the right side is true at that world (given that the atomic formula $Rx\imath zPz$ is false at that world, making $\neg Rx\imath zPz$ true there). But on the algebraic approach, the following instance of **RC** is valid:

$$\exists F \Box \forall x(Fx \equiv \neg Rx\imath zPz)$$

As one can see from the above considerations, in any interpretation, any necessarily exemplified property F will be a witness to this claim, since for any world you pick, every object x is such that $\neg Rx\imath zPz$ there, given that the description fails to denote.

[19] The use of 'β' in 'β-Conversion' derives from Curry (1963, 117), but this principle first appears in Church 1932 (355, Rule of Procedure II).

[20] For a proof that β-Conversion is valid in those interpretations that satisfy the constraints described in Section 4.5, see the Appendix. The proof is preceded by proofs of some of the other, preliminary lemmas mentioned previously, such as the Assignment Agreement Lemma, Corollary to the Assignment Agreement Lemma, Substitution Lemma, and the Generalized Substitution Lemma.

But, normally, the existence claim would be inferred from the instance of β-Conversion, as follows:

1.	$[\lambda y\, \neg Ry\imath zPz]x \equiv \neg Rx\imath zPz$	Instance
2.	$\forall x([\lambda y\, \neg Ry\imath zPz]x \equiv \neg Rx\imath zPz)$	GEN, 1
3.	$\Box\forall x([\lambda y\, \neg Ry\imath zPz]x \equiv \neg Rx\imath zPz)$	RN, 2
4.	$\exists F\Box\forall x(Fx \equiv \neg Rx\imath zPz)$	$\exists$I, 3

In the algebraic interpretation, this derivation is blocked, given that (1) isn't an instance of β-Conversion, for the reasons stated above. Indeed, generally, one can't derive **RC** from β-Conversion in full generality in the algebraic interpretation (following the above derivation scheme) since no instances of β-Conversion with descriptions are allowed. Instead, one can only derive, in the algebraic interpretation, a version of **RC** that has the same restriction that was placed on β-Conversion. Thus, all of the valid instances of **RC** containing descriptions have to be added as axioms, such as $\exists F\Box\forall x(Fx \equiv \neg Rx\imath zPz)$, $\exists F\Box\forall x(Fx \equiv Qx\imath zPz \rightarrow Q\imath zPz)$, etc. Indeed, it isn't even clear how to formulate axioms that would assert all such validities on the algebraic approach! These claims may have to be asserted piecemeal.

But on the present approach, no restrictions on β-Conversion are necessary. The instances of β-Conversion involving descriptions are valid, since some relation or other is always denoted by every λ-expression, given D5, D6 and our constraints on interpretations. So the existence of relations involving descriptions is guaranteed and **RC** can be derived from β-Conversion in full generality. No special axioms have to be added.

5.2 Identity

There is one element missing from our account. As it stands, identity is *not* a primitive in our target language and so we can't even express standard identities among λ-expressions, such as α- and η-conversion, or other kinds of identities. Of course, this is easily remedied: we are free to add identity as a primitive and start asserting identities among relations. Our BNF definition of the language is easily modified to allow for the formation of identity statements among individual terms and among relation terms. Once modified appropriately, we would be free to assert, for example:

- $[\lambda x\, \neg Fx] = [\lambda y\, \neg Fy]$

- $[\lambda x\, Fx] = F$
- $[\lambda x\, Px \,\&\, Qx] = [\lambda x\, Qx \,\&\, Px]$
- $[\lambda x\, Rx \,\&\, \neg Rx] \neq [\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)]$

Given such assertions, interpretations of our language would have to behave in certain ways if they are to make the above true, and in particular, such assertions would force our $\bar{\epsilon}$-terms to pick relations in certain ways.

Although we certainly want to be able to allow for the assertion of the above equalities and inequalities, the approach to identity just outlined is still not adequate, for a completely general theory of relations should include not just a statement of *existence conditions* for relations but also a statement of *identity conditions* for relations. We can take **RC** as our background comprehension condition for the existence of relations, but if we are to avoid Quine's charge that properties (including relations and propositions) are "creatures of darkness" (1956, 180), we must give mathematically precise identity conditions for our intensional relations. Such identity conditions should be consistent with different intuitions about difficult cases of relation identity: such identity conditions should tell us generally what as to be the case when we assert that relations are or are not identical.

So our plan in what follows is this: instead of taking identity as a primitive and adding axioms for identity, we shall *define* general conditions for the identity of properties, relations and propositions. Moreover, we plan to state these identity conditions for relations entirely in our object language. We won't need the *any* notions from the metalanguage, such as its background language of set theory, to state a theory of identity for relations. This is important given the assumption I made at the outset describing my starting point, namely, that relations are more fundamental than sets. If the first- and second-order (modal) predicate calculus is more fundamental than set theory, as I have suggested, then our question becomes, what identity conditions for relations can be stated using only the resources of predication found in the predicate calculus? Such a statement will then allow us to place final constraints on interpretations, so as to validate the theory of relations.

The statement of identity conditions for relations can be given by extending our target language to the language of *object theory*, as described in Zalta 1983, 1988a, 1999, and elsewhere. The language of object theory differs from the language considered above by having a second kind of atomic formula, namely, formulas of the form xF^1 ('x encodes F^1'). These encoding formulas can be used to form complex formulas and descriptions, but they may *not* appear as subformulas of the matrix φ in λ-expressions.[21] Consequently, object theory singles out a subclass of formulas, namely the *propositional* formulas which are free of encoding subformulas, and requires that λ-expressions be built only from matrix formulas that are propositional. Thus, the class of λ-expressions in the language of object theory is identical to the class of λ-expressions of our target language defined in Section 3.[22]

Indeed, the semantics developed above adapts to the language of object theory with very little additional structure. We need only an encoding extension function, **en**, that maps each 1-place relation in $\mathbf{R}_1$ to a subset of $\mathbf{D}$, so that the truth conditions of 'xF^1' can be stated as $\boldsymbol{d}(x) \in \mathbf{en}(\boldsymbol{d}(F^1))$ (ignoring the relativization to $\mathcal{I}$ and f).[23] Note that the truth conditions of xF^1 make no reference to possible worlds: the encoding extension of a property doesn't vary from world to world. So for any interpretation $\mathcal{I}$, assignment f, and world $\boldsymbol{w}$, if xF^1 is true$_{\mathcal{I},f}$, then so is $\Box xF^1$.

We've now sufficiently developed and described this extension of our target language for our purposes in what follows. Our extended target language allows us to formulate general identity conditions for relations. Readers familiar with object theory already know how, but for those who don't, identity conditions for properties are:[24]

[21] This proscription provides a solution to the paradoxes of encoding. But we shall not discuss these here as this will take us too far afield. See some of the referenced publications on object theory.

[22] Actually, this is not quite accurate, but near enough for our present purposes. If a formula φ contains a description, $\iota x(\ldots xF \ldots)$, in which an encoding formula such as xF appears, then xF is not considered a subformula of φ. So if we say that such formulas contain *description-embedded* encoding formulas, it is possible to have *propositional* formulas that have description-embedded encoding formulas. Thus, in full object theory, we permit λ-expressions whose matrix formula is a propositional formula with description-embedded encoding formulas. But those behave no differently from the λ-expressions described in our original target language in Section 3, and so for our purposes below, they can be ignored.

[23] Strictly speaking, the condition has to allow for descriptions that might fail to denote, and should read: $\exists \boldsymbol{r}^1 \exists \boldsymbol{o}(\boldsymbol{r}^1 = \boldsymbol{d}(F^1) \,\&\, \boldsymbol{o} = \boldsymbol{d}(x) \,\&\, \boldsymbol{o} \in \mathbf{en}(\boldsymbol{r}^1))$.

[24] Given that encoding predications are necessary if true, one might wonder why I include the modal operator in the following definition. The reason is that I take identity to be a modal notion. I want the definition to expressly say that for F and G to be identical, it is necessary that they are encoded by by the same objects.

Property Identity (**PI**):
$F^1 = G^1 =_{df} \Box\forall x(xF^1 \equiv xG^1)$

Now, in terms of this definition, we define both identity for propositions and for n-place relations ($n \geq 2$). Identity for propositions is defined as follows:[25]

Proposition Identity:
$p = q =_{df} [\lambda y\, p] = [\lambda y\, q]$

Identity for n-place relations ($n \geq 2$) is defined as:

Relation Identity:
$$F^n = G^n =_{df} \quad \text{(where } n \geq 2\text{)}$$
$$\forall x_1 \ldots \forall x_{n-1}([\lambda y\, F^n y x_1 \ldots x_{n-1}] = [\lambda y\, G^n y x_1 \ldots x_{n-1}]\ \& \ [\lambda y\, F^n x_1 y x_2 \ldots x_{n-1}] = [\lambda y\, G^n x_1 y x_2 \ldots x_{n-1}]\ \& \ldots \& \ [\lambda y\, F^n x_1 \ldots x_{n-1} y] = [\lambda y\, G^n x_1 \ldots x_{n-1} y])$$

These two definitions reduce the identity of non-monadic relations to the identity of properties. They are central to object theory's *theory* of identity, and they offer *extensional* conditions for the identity of intensional entities (notwithstanding the modal operator in the definition of property identity). Quine's complaint, therefore, is undermined. Our theory tells us, in theoretical terms, what it is we know when we pretheoretically judge, assert, assume, etc., that relations F and G are identical or distinct.

Finally, then, we require that interpretations of the language of object theory obey additional constraints, so as to ensure that when the definientia of these definitions obtain, interpretations *semantically identify* the relations in question. Thus, in the case of properties, we require that an interpretation $\mathcal{I}$ of the language of object theory be such that:[26]

$$\forall \boldsymbol{r}, \boldsymbol{s} \in \mathbf{R}_1 [\forall \boldsymbol{o} \in \mathbf{D}(\boldsymbol{o} \in \mathbf{en}(\boldsymbol{r}) \equiv \boldsymbol{o} \in \mathbf{en}(\boldsymbol{s})) \rightarrow \boldsymbol{r} = \boldsymbol{s}]$$

[25] Cf. Myhill 1963, where a definition like this first appeared.

[26] Note that in order to ensure that **PI** correctly identifies relations no matter what the interpretation, we need *not* preface the antecedent of the following condition with a universal quantifier over possible worlds, since such a quantifier would be vacuous: the encoding extension of a property doesn't vary from world to world, and so if the encoding extensions of two properties are the same, they are the same with respect to every world.

Something similar can be done in the case of relations and propositions.[27] So we shall assume in what follows that the principle of substitution of identicals is valid, even though the antecedents of its instances are defined identity statements. Consequently, **RC** and the above definitions for identity jointly embody a precise and general theory of relations by providing existence and identity conditions for them.

5.3 Asserting Inequalities

Clearly, the semantics we've developed does *not* require that necessarily equivalent relations be identical. We may consistently assert, using the notions of identity just defined, that some necessarily equivalent relations are distinct. Under the intensional conception of relations, there can be open propositional formulas $\varphi(x)$ and $\psi(x)$ such that at every world $\boldsymbol{w}$, exactly the same objects satisfy both $\varphi(x)$ and $\psi(x)$; yet φ and ψ can be used to construct terms that denote different relations. The examples we used in the introduction were:

- x is red and not red:
  $Rx \,\&\, \neg Rx$
- x is a barber who shaves all and only those who don't shave themselves:
  $Bx \,\&\, \forall y(Sxy \equiv \neg Syy)$

The corresponding λ-expressions are:

- $[\lambda x\, Rx \,\&\, \neg Rx]$
- $[\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)]$

Focusing on this particular case, note that the domain $\mathbf{R}_1$ in interpretations $\mathcal{I}$ is guaranteed to have at least one or more properties whose

[27] For n-place relations ($n \geq 2$) and propositions ($n = 0$), one has to add more structure to interpretations $\mathcal{I}$ to ensure that when the definientia obtain, the interpretations identify the relations and propositions. For example, in the case of propositions, we have to ensure that whenever two properties are the vacuous expansions of distinct propositions, the properties have distinct encoding extensions. In the case of relations, we have to ensure that when we plug up distinct n-place relations by $n-1$ objects (plugging the objects into the corresponding argument places of the two relations), the properties that result have distinct encoding extensions. But we leave this for another occasion, as this would take us too far afield of the present paper.

exemplification extensions are empty at every world, since the formulas used in the above λ-expressions can be used to form valid instances of **RC**. To a first approximation, *any* such property is suitable as a denotation for the above two λ-expressions. Clause D5 picks one for each term. All we have to do to ensure that the two λ-expressions don't denote the same property is *assert*, in the object language (using the notion of identity defined above) that the two properties are distinct:

$$[\lambda x\, Rx \,\&\, \neg Rx] \neq [\lambda x\, Bx \,\&\, \forall y(Sxy \equiv \neg Syy)].$$

Such an assertion, when the defined notion of identity is expanded, guarantees that at some possible world, there is an object that encodes the property denoted by one of the two expressions without encoding the property denoted by the other. This couldn't happen if the properties denoted by the λ-expressions were semantically identified.[28] The above inequality is consistent, given our semantics, with the claim that the two properties are necessarily exemplified by the same objects. So our semantic method allows us to assert that some necessarily equivalent relations are distinct. We are free to do this as the occasion and need arises.

5.4 α- and η- Conversion

Note that we can use our system extended with the defined notions of relation identity to enquire into the status of the two other axioms of the classical λ-calculus:

- α-**Conversion**: $[\lambda \nu_1 \ldots \nu_n\, \varphi] = [\lambda \nu'_1 \ldots \nu'_n\, \varphi']$,
  where $[\lambda \nu_1 \ldots \nu_n\, \varphi]$ and $[\lambda \nu'_1 \ldots \nu'_n\, \varphi']$ are alphabetic variants.
- η-**Conversion**: $[\lambda x_1 \ldots x_n\, F^n x_1 \ldots x_n] = F^n$

In these axioms, the identity sign is not a primitive, but defined in our system extended with the addition of encoding formulas. But, as we

[28] In the semantics, we get for free that

$$\forall \boldsymbol{r}, \boldsymbol{s} \in \mathbf{R}_1[\boldsymbol{r} = \boldsymbol{s} \rightarrow \forall \boldsymbol{o} \in \mathbf{D}(\boldsymbol{o} \in \mathbf{en}(\boldsymbol{r}) \equiv \boldsymbol{o} \in \mathbf{en}(\boldsymbol{s}))]$$

This follows by the logic of identity that is operative in the semantics. So, the converse asserts: if the encoding extensions of $\boldsymbol{r}$ and $\boldsymbol{s}$ differ, then $\boldsymbol{r}$ and $\boldsymbol{s}$ are different properties. That is why, the defined claim that $F \neq G$ in the object language, once expanded in terms of the definition of property identity, semantically requires that the encoding extensions of the properties denoted by F and G are different properties.

shall now show, given our semantic stipulations in Section 4, the formulas that result by expanding the defined identity symbol are valid.

Consider α-Conversion first. Given our definitions for identity, we have the following three general classes of instances of α-Conversion:

- $[\lambda \nu_1\, \varphi] = [\lambda \nu'_1\, \varphi']$ $(n = 1)$
- $[\lambda\, \varphi] = [\lambda\, \varphi']$ $(n = 0)$
- $[\lambda \nu_1 \ldots \nu_n\, \varphi] = [\lambda \nu'_1 \ldots \nu'_n\, \varphi']$ $(n \geq 2)$

In each case, when you expand the defined identity sign, the λ-expressions flanking the identity sign become part of formulas flanking a biconditional. D5 and D6 require that alphabetically variant λ-expressions have the same denotation, and so these particular expanded formulas are valid, since the λ-expressions in the formulas flanking a biconditional will denote the same relations.

So, for example, in the 1-place case:

$$[\lambda y\, \neg Fy] = [\lambda z\, \neg Fz]$$

expands to:

$$\Box\forall x(x[\lambda y\, \neg Fy] \equiv x[\lambda z\, \neg Fz])$$

given the definition Property Identity. Since D5 stipulates that $[\lambda y\, \neg Fy]$ and $[\lambda z\, \neg Fz]$ denote the same property in any interpretation, it takes very little reasoning to show that the above formula is valid.

A similar argument applies to η-Conversion, but before we consider it, notice the simple way in which η-Conversion was presented. The left-hand side of the equation in η-Conversion is a specific elementary λ-expression involving the atomic formula $F^n x_1 \ldots x_n$ in which the n-place relation variable F^n has a free occurrence and the distinct variables $x_1, \ldots, x_n$ are all bound by the λ.[29] Given that F^n is free, we can universally generalize on the F^n and then instantiate the generalization to *any* n-place relation term Π^n substitutable for F^n, to get:

$$[\lambda x_1 \ldots x_n\, \Pi^n x_1 \ldots x_n] = \Pi^n, \text{ where } x_1, \ldots, x_n \text{ aren't free in } \Pi^n$$

[29] Compare the statement of η-Conversion with following statement by Curry & Feys 1958 (92) and Hindley & Seldin 1986 (73): $\lambda x M x = M$, where x is not free in M. Here M is clearly a metavariable that ranges over (complex) expressions. By contrast, we've stated η-Conversion using variables of the object language.

And given that alphabetic variants are identical, we can replace x_i by ν_i for any other variable ν we may be informally using. Thus, η-Conversion yields:

η-**Conversion Theorem Schema 1**:
$[\lambda \nu_1 \ldots \nu_n \, \Pi^n \nu_1 \ldots \nu_n] = \Pi^n$, for any elementary λ-expression[30]

Moreover, it doesn't take too much work to show that the following theorem schema is also derivable:

η-**Conversion Theorem Schema 2**
$[\lambda \nu_1 \ldots \nu_n \, \varphi] = [\lambda \nu_1 \ldots \nu_n \, \varphi']$,
where $[\lambda \nu_1 \ldots \nu_n \, \varphi']$ is any η-variant of $[\lambda \nu_1 \ldots \nu_n \, \varphi]$

Although we haven't specified a *deductive system* of axioms and rules, it is easy to see that *instances* of this schema would be derivable from our simple formulation of η-Conversion, given standard formulations of axioms and rules. As an example, we derive:

$$[\lambda y \, Py \rightarrow Say] = [\lambda y \, [\lambda z \, Pz]y \rightarrow [\lambda uv \, Suv]ay]$$

by the following argument, where ηC abbreviates η-Conversion:[31]

1.	$[\lambda z \, Pz] = P$	instance, ηC
2.	$[\lambda uv \, Suv] = S$	instance, ηC
3.	$[\lambda y \, Py \rightarrow Say] = [\lambda y \, Py \rightarrow Say]$	instance, =I
4.	$[\lambda y \, Py \rightarrow Say] = [\lambda y \, [\lambda z \, Pz]y \rightarrow Say]$	=E, 1, 3
5.	$[\lambda y \, Py \rightarrow Say] = [\lambda y \, [\lambda z \, Pz]y \rightarrow [\lambda uv \, Suv]ay]$	=E, 2, 4

With this example in mind, we now appeal to η-Conversion Theorem Schema 1 in an extended argument that generalizes the above argument to *any* η-variants, and thereby show that η-Conversion Theorem Schema 2 is derivable generally. First we establish:

Lemma: η-**Conversion on Immediate Variants**:
If $[\lambda \nu_1 \ldots \nu_n \, \varphi']$ is an *immediate* η-variant of $[\lambda \nu_1 \ldots \nu_n \, \varphi]$, then
$[\lambda \nu_1 \ldots \nu_n \, \varphi] = [\lambda \nu_1 \ldots \nu_n \, \varphi']$

[30]Recall that we've defined elementary λ-expressions so that in the elementary expression $[\lambda \nu_1 \ldots \nu_n \, \Pi^n \nu_1 \ldots \nu_n]$, none of the ν_i are free in Π^n.

[31] Line (3) of the proof that follows in the text cites =I, i.e., Identity Introduction. This abbreviates the following chain of reasoning. Clearly, we can derive $F^1 = F^1$, for by the definition of Property Identity, we have to show $\Box\forall x(xF \equiv xF)$, which is easily derivable as a theorem of logic. So we can universally generalize to establish $\forall F^1(F^1 = F^1)$. Then we can instantiate F^1 to the term $[\lambda y \, Py \rightarrow Say]$, producing line (3) in the text.

Proof: Suppose $[\lambda \nu_1 \ldots \nu_n \varphi']$ is an *immediate* η-variant of $[\lambda \nu_1 \ldots \nu_n \varphi]$. Then, by definition of immediate variant, let Π^n be the relation term in the latter whose η-expansion, $[\lambda \nu_1{}' \ldots \nu_n{}' \, \Pi^n \nu_1{}' \ldots \nu_n{}']$, appears in the former. By the relevant instance of η-Conversion Theorem Schema 1, i.e.,

$$[\lambda \nu_1{}' \ldots \nu_n{}' \, \Pi^n \nu_1{}' \ldots \nu_n{}'] = \Pi^n$$

we may substitute $[\lambda \nu_1{}' \ldots \nu_n{}' \, \Pi^n \nu_1{}' \ldots \nu_n{}']$ for Π^n into the right-side occurrence of φ in the following instance of the reflexivity of identity:[32]

$$[\lambda \nu_1 \ldots \nu_n \, \varphi] = [\lambda \nu_1 \ldots \nu_n \, \varphi]$$

The result, by =E, is:

$$[\lambda \nu_1 \ldots \nu_n \, \varphi] = [\lambda \nu_1 \ldots \nu_n \, \varphi'] \qquad \bowtie$$

The proof of η-Conversion Theorem Schema 2 now is at hand, for whenever ρ' is an η-variant of ρ, there is a finite sequence of λ-expressions such that each member of the sequence is an immediate η-variant of the preceding member of the sequence. So by a finite number of applications of the above Lemma, we can prove $\rho = \rho'$, where ρ' is an η-variant of ρ.

Finally, as to the validity of η-Conversion. Given the three cases for defining the identity symbol, the three cases of η-Conversion are:

- $[\lambda x \, F^1 x] = F^1$
- $[\lambda \, p] = p$
- $[\lambda x_1 \ldots x_n \, F^n x_1 \ldots x_n] = F^n \qquad (n \geq 2)$

Clearly, for each case, if we expand the formula so as to eliminate the defined identity symbol, the result is something easily, if sometimes tediously, established as valid. In each case, the resulting formulas contain a pair of η-variant λ-expressions flanking a biconditional (or flanking the biconditionals in a conjunction of biconditonals). Our semantic

[32]Reflexivity of identity was derived for 1-place relations in footnote 31. Given that n-place relation identity ($n \geq 2$) is defined in terms of 1-place relation identity, the argument easily generalizes to yield $F^n = F^n$ ($n \geq 2$) as a theorem.

rules T1–T6 and D1–D6, and subsequent constraints, combine to guarantee the truth of those formulas in every interpretation.[33]

5.5 η-Conversion and Extensionality

Finally, it may come as a surprise to those who are more familiar with the functional λ-calculus that our extended system validates η-Conversion *without* endorsing any objectionable form of extensionality. It is sometimes thought that η-Conversion is an axiom that imposes extensionality on the λ-calculus (see below). But, clearly, our system doesn't endorse either of the following objectionable forms of extensionality:

- $\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n) \rightarrow F^n = G^n \quad (n \geq 0)$
- $\Box\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n) \rightarrow F^n = G^n \quad (n \geq 0)$

The reason these forms are objectionable was explained in Section 1.3: in many cases, we want to be able to assert that relations are distinct even if materially or necessarily equivalent.

[33] The 1-place case η-Conversion is trivial. The expanded formula is: $\Box\forall y(y[\lambda x Fx] \equiv yF)$, and since $[\lambda x Fx]$ is elementary, D4 guarantees that F and $[\lambda x Fx]$ denote the same property. So the properties denoted by $[\lambda x\, Fx]$ and F have the same encoding extension, and hence $\forall y(y[\lambda x Fx] \equiv yF)$ is true in every interpretation. Since the encoding extension of properties doesn't vary from world to world, the formula $\Box\forall y(y[\lambda x\, Fx] \equiv yF)$ is also true in every interpretation. So η-Conversion is valid for 1-place relations.

For propositions, the definition of Proposition Identity applied to $[\lambda\, p] = p$ yields $[\lambda y\,[\lambda\, p]] = [\lambda y\, p]$. This in turn can be expanded by the definition of Property Identity to: $\Box\forall x(x[\lambda y\,[\lambda\, p]] \equiv x[\lambda y\, p])$. But $[\lambda y\, p]$ is a non-elementary, η-irreducible λ-expression, and $[\lambda y\,[\lambda\, p]]$ is an (immediate!) η-variant of it, So, by D5, they have the same denotation and the reasoning reduces to that of the previous case. D5's requirement that η-variant λ-expressions have the same denotation is crucial, for without it, nothing guarantees that the properties picked out by $\bar{\epsilon}$-terms as the denotations of $[\lambda y\,[\lambda\, p]]$ and $[\lambda y\, p]$ have the same encoding extension.

When the definition of n-place relation identity ($n \geq 2$) is applied to the instance $[\lambda z_1 \ldots z_n\, F^n z_1 \ldots z_n] = F^n$, the result is, for any $x_1, \ldots, x_{n-1}$, a series of conjunctions of biconditionals:

- $[\lambda y\, F^n y x_1 \ldots x_{n-1}] = [\lambda y\,[\lambda z_1 \ldots z_n\, F^n z_1 \ldots z_n] y x_1 \ldots x_{n-1}]$
- …
- $[\lambda y\, F^n x_1 \ldots x_{n-1} y] = [\lambda y\,[\lambda z_1 \ldots z_n\, F^n z_1 \ldots z_n] x_1 \ldots x_{n-1} y]$

To see that each conjunct is logically true, note that each asserts a property identity involving non-elementary, η-irreducible λ-expressions and their (immediate) η-variants. As we've seen previously, each expands to a logical truth, for any assignment to $x_1, \ldots, x_{n-1}$.

But this now raises the question, why do Hindley and Seldin (1986, 74) say:

> … (η) is usually taken as the 'canonical' definition of extensionality in λ-calculus.

The answer is, our semantics doesn't endorse a principle that corresponds to a principle they presupposed when they asserted the above, namely, the Rule ($\Box$)ξ:[34]

$$\text{From } (\Box)\forall x_1 \ldots \forall x_n(\varphi \equiv \psi), \text{ infer } [\lambda x_1 \ldots x_n\, \varphi] = [\lambda x_1 \ldots x_n\, \psi]$$

(I've put the $\Box$ in parenthesis to emphasize that there is a non-modal and a modal form of Rule ξ.) Clearly, η-Conversion ('ηC') and the Rule ($\Box$)ξ jointly imply that necessarily equivalent relations are identical, by the following argument:[35]

[34] I'm indebted to Allen Hazen for noting this point. I've also benefited from reading Alama 2013 as I was sorting through the issues.

[35] It may not be immediately clear why our version of Rule ($\Box$)ξ corresponds to rule (ξ) found in Curry & Feys 1958 (89) and in Hindley & Seldin 1986 (66). Curry & Feys formulate the rule (ξ) as: from $A = B$, infer $\lambda x A = \lambda x B$; Hindley & Seldin formulate it as: from $M = M'$, infer $\lambda x.M = \lambda x.M'$. Both pairs of authors use (ξ) and (η) to derive extensionality, which is formulated as: $Mx = Nx \rightarrow M = N$, by the following argument (Curry & Feys 1958, 92; Hindley & Seldin 1986, 74):

1.	$Mx = Nx$	Assumption
2.	$\lambda x Mx = \lambda x Nx$	(ξ), 1
3.	$\lambda x Mx = M$	instance of (η)
4.	$\lambda x Nx = N$	instance of (η)
5.	$M = N$	=E, 2, 3, 4
6.	$Mx = Nx \rightarrow M = N$	CP, 1–5

But notice that in this argument, they apply (ξ) to line 1 by taking $Mx = Nx$ as an instance of $M = N$, which is the premise to Rule (ξ). But we can't represent $Mx = Nx$ as $Fx = Gx$, since the latter is not well-defined, no matter what terms Π^1 we substitute for the variables F, G. If we represent $Mx = Nx$ as $(\Box)\forall x(Fx \equiv Gx)$, then we can't regard the latter as an instance of $F = G$ in the same way that $Mx = Nx$ is regarded as an instance of $M = N$ (again, no matter what terms we substitute for the variables F, G).

Consequently, to capture the reasoning that shows (ξ) and (η) imply extensionality, I've formulated the premise of Rule ($\Box$)ξ so that it directly justifies the move from line 1 to line 2 in the proof given by Curry & Feys and Hindley & Seldin. So the premise $Mx = Nx$ has been represented as $(\Box)\forall x_1 \ldots \forall x_n(\varphi \equiv \psi)$, and the conclusion $\lambda x Mx = \lambda x Nx$ has been represented as $[\lambda x_1 \ldots x_n\, \varphi] = [\lambda x_1 \ldots x_n\, \psi]$.

1.	$(\Box)\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n)$	Assumption
2.	$[\lambda x_1 \ldots x_n\, F^n x_1 \ldots x_n] = [\lambda x_1 \ldots x_n\, G^n x_1 \ldots x_n]$	$(\Box)\xi$, 1
3.	$[\lambda x_1 \ldots x_n\, F^n x_1 \ldots x_n] = F^n$	ηC
4.	$[\lambda x_1 \ldots x_n\, G^n x_1 \ldots x_n] = G^n$	ηC
5.	$F^n = G^n$	=E, 2, 3, 4
6.	$(\Box)\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \equiv G^n x_1 \ldots x_n) \rightarrow F^n = G^n$	CP, 1–5

But there is no longer any motivation to endorse either Rule ξ or the objectionable extensionality principles mentioned above—the classical desideratum of having an extensional theory of relations is met by the object-theoretic definitions of relation identity. In light of our semantics, these provide extensional identity conditions for our intensional entities without implying either the objectionable forms of extensionality or endorsing Rule ξ.

6 Appendix: Definitions and Metatheorems

In this Appendix, we state definitions and prove metatheorems for the language and semantics developed in Sections 3 and 4. We leave the corresponding definitions and metatheorems for the extended system discussed in 5 for a another occasion.

6.1 Free Occurrences of Variables

By considering the appropriate clauses in the BNF definition of *term* and *formula*, we recursively define what it is for an occurrence of a variable α to be *free in* term τ or formula φ, as follows:

(.1) If τ is the variable α, then that occurrence of α in τ is free. (If τ is a constant or a variable other than α, then there is no occurrence of α in τ.)

(.2) If φ is $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 0)$, then an occurrence of α in φ is free iff it is an occurrence of α in one of the terms $\kappa_1,\ldots,\kappa_n$, or Π^n that is free.

(.3) If φ is $\neg\psi$ or $\Box\psi$, then an occurrence of α in φ is free iff it is an occurrence of α in ψ that is free. If φ is $(\psi \rightarrow \chi)$, then an occurrence of α in φ is free iff it is an occurrence of α in ψ or χ that is free.

(.4) If φ is $\forall\beta\psi$, then an occurrence of α in φ is free iff (i) it is an occurrence of α in ψ that is free and (ii) it is not an occurrence of β.

(.5) If τ is $\imath\nu\psi$, then an occurrence of α in τ is free iff (i) it is an occurrence of α in ψ that is free and (ii) it is not an occurrence of ν.

(.6) If τ is $[\lambda\nu_1 \ldots \nu_n\, \psi^*]$ $(n \geq 0)$, then an occurrence of α in τ is free iff (i) it is an occurrence of α in ψ^* that is free and (ii) it is not an occurrence of $\nu_1,\ldots,\nu_n$.

(.7) No occurrence of a variable is free in an expression unless it can be so demonstrated by the clauses above.

We henceforth say that a variable α *occurs free* or *is free* in formula φ or term τ if and only if at least one occurrence in of α in φ or τ is free.

6.2 Assignment Agreement Lemma

Lemma. If f and g are any assignment functions that agree on the free variables (if any) in φ, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ iff $\boldsymbol{w} \models_{\mathcal{I},g} \varphi$.

Proof: By induction on the complexity of φ and, in the base case, a secondary induction on the complexity of terms τ occurring in φ. Assume that f and g are arbitrarily chosen assignments that agree on the free variables, if any, in φ.

Formula Induction: Base Case. φ has the form $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 1)$ or the form Π^0.

Term Induction: Base Case. All of the individual and relation terms τ in φ are simple (i.e., constants or variables). If term τ in φ is a constant (i.e., an individual constant or n-place relation constant), then by D1, $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \mathbf{V}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$. If τ is a variable (i.e., an individual variable or an n-place relation variable), then by D2, $\boldsymbol{d}_{\mathcal{I},f}(\tau) = f(\tau)$ and $\boldsymbol{d}_{\mathcal{I},g}(\tau) = g(\tau)$. However, when τ is a variable, then the occurrence of τ is, by definition, free in φ. So, by hypothesis, $f(\tau) = g(\tau)$. So $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$. Hence, no matter whether τ in φ is a constant or a variable, $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$. Now, to show $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ iff $\boldsymbol{w} \models_{\mathcal{I},g} \varphi$, our two cases are:

- *Case A*. $n \geq 1$, φ is $\Pi^n\kappa_1 \ldots \kappa_n$. By T1, we have to show that:

$$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_1)\ \&\ \ldots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$$

if and only if
$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},g}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},g}(\kappa_1)\ \&\ \ldots\ \&$
$\boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},g}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$

- *Case B*. $n = 0$, φ is Π^0, i.e., a 0-place relation constant or variable. By T2, we have to show that:

$\exists \boldsymbol{r}^0(\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},f}(\Pi^0)\ \&\ \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$
if and only if
$\exists \boldsymbol{r}^0(\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},g}(\Pi^0)\ \&\ \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$

But, clearly, in *all* of these cases, if $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$ for every term τ in φ, then any witnesses to the existential claims in the left-side condition are witnesses to the existential claims in the right-side condition, and vice versa.

Term Induction: Inductive Case 1. One or more of the individual terms κ_i in φ is a definite description. So there is only one case: φ is $\Pi^n \kappa_1 \ldots \kappa_n$ ($n \geq 1$) and one or more of the κ_i in $\Pi^n \kappa_1 \ldots \kappa_n$ is a description. (When φ is Π^0, $n = 0$ and there are no κ_i in Π^0.) We may, without loss of generality, assume that exactly one of the κ_i, say κ_1, is a description. So φ is $\Pi^n \imath\nu\psi\kappa_2 \ldots \kappa_n$ and all the other terms in φ are as in the Term Induction Base Case, i.e., all of the other terms τ in φ are such that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$. So to show $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ iff $\boldsymbol{w} \models_{\mathcal{I},g} \varphi$, we have to show:

$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\imath\nu\psi)\ \&\ \boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_2)\ \&$
$\ldots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$
if and only if
$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},g}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},g}(\imath\nu\psi)\ \&\ \boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},g}(\kappa_2)\ \&$
$\ldots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},g}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$

($\rightarrow$) For the left-right direction, assume the left-hand condition, and let $\boldsymbol{R}^n, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ be arbitrarily chosen witnesses to the existential claims. So $\boldsymbol{a}_1 = \boldsymbol{d}_{\mathcal{I},f}(\imath\nu\psi)$. We can reduce this to the previous case if we can show $\boldsymbol{a}_1 = \boldsymbol{d}_{\mathcal{I},g}(\imath\nu\psi)$, since by hypothesis, all the other terms in φ are such that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$. Now by D3, it follows from the fact that $\boldsymbol{a}_1 = \boldsymbol{d}_{\mathcal{I},f}(\imath\nu\psi)$ that:

(ϑ) $\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{a}_1]} \psi\ \&\ \forall \boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}']} \psi \rightarrow \boldsymbol{o}' = \boldsymbol{a}_1)$

Since we want to show $\boldsymbol{a}_1 = \boldsymbol{d}_{\mathcal{I},g}(\imath\nu\psi)$, we need to establish:

(ζ) $\boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{a}_1]} \psi\ \&\ \forall \boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{o}']} \psi \rightarrow \boldsymbol{o}' = \boldsymbol{a}_1)$

Our inductive hypothesis is that any assignments f' and g' that agree on the free variables of ψ are such that $\boldsymbol{w} \models_{\mathcal{I},f'} \psi$ iff $\boldsymbol{w} \models_{\mathcal{I},g'} \psi$. But note that the free variables in ψ include ν and possibly all of the free variables in φ. (The free variables in φ may be free in ψ with the exception of ν.) Since f and g agree on all the free variables of φ, $f[\nu/\boldsymbol{a}_1]$ and $g[\nu/\boldsymbol{a}_1]$ agree on all of the free variables in ψ. So our inductive hypothesis yields:

$$\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{a}_1]} \psi \quad \text{iff} \quad \boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{a}_1]} \psi$$

The first conjunct of (ζ) now follows from this and the first conjunct of (ϑ). To show the second conjunct of (ζ) holds, assume $\boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{b}]} \psi$, where $\boldsymbol{b}$ is arbitrary, to show $\boldsymbol{b} = \boldsymbol{a}_1$. By instantiating $\boldsymbol{b}$ into the second conjunct of (ϑ), we know:

$$\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{b}]} \psi \rightarrow \boldsymbol{b} = \boldsymbol{a}_1$$

But again, since $f[\nu/\boldsymbol{b}]$ and $g[\nu/\boldsymbol{b}]$ agree on all the free variables in ψ, our inductive hypothesis yields:

$$\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{b}]} \psi \quad \text{iff} \quad \boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{b}]} \psi$$

This and our assumption that $\boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{b}]} \psi$ implies $\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{b}]} \psi$, allowing us to conclude $\boldsymbol{b} = \boldsymbol{a}_1$. So we've established ($\zeta$), and hence, that $\boldsymbol{a}_1 = \boldsymbol{d}_{\mathcal{I},g}(\imath\nu\psi)$. ($\leftarrow$) By analogous reasoning.

Term Induction: Inductive Case 2. The n-place relation term in φ is complex and any individual terms κ_i are simple. There are two cases:

Case A. For $n \geq 1$, when φ is $\Pi^n \kappa_1 \ldots \kappa_n$ and the n-place relation term Π^n is a λ-expression of the form $[\lambda \nu_1 \ldots \nu_n\, \psi]$

Case B. For $n = 0$, when φ is Π^0, and Π^0 is a λ-expression of the form $[\lambda\, \psi]$.

So, to show $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ iff $\boldsymbol{w} \models_{\mathcal{I},g} \varphi$, we have to show:

Case A. $\boldsymbol{w} \models_{\mathcal{I},f} [\lambda \nu_1 \ldots \nu_n\, \psi]\kappa_1 \ldots \kappa_n$ iff $\boldsymbol{w} \models_{\mathcal{I},g} [\lambda \nu_1 \ldots \nu_n\, \psi]\kappa_1 \ldots \kappa_n$

Case B. $\boldsymbol{w} \models_{\mathcal{I},f} [\lambda\, \psi]$ iff $\boldsymbol{w} \models_{\mathcal{I},g} [\lambda\, \psi]$

By expanding *Case A* using T1 and by expanding *Case B* using T2, we have to show, letting $\boldsymbol{o}_i = \boldsymbol{d}_{\mathcal{I},f}(\kappa_i)$ abbreviate $\boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_1)\ \&\ \ldots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n)$:

(ϑ1) *Case A*.

$$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}([\lambda \nu_1 \ldots \nu_n \psi]) \ \& \ \boldsymbol{o}_i = \boldsymbol{d}_{\mathcal{I},f}(\kappa_i) \ \& \ \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$$

if and only if

$$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},g}([\lambda \nu_1 \ldots \nu_n \psi]) \ \& \ \boldsymbol{o}_i = \boldsymbol{d}_{\mathcal{I},g}(\kappa_i) \ \& \ \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n \rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$$

(ϑ2) *Case B*.

$$\exists \boldsymbol{r}^0(\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},f}([\lambda\, \psi]) \ \& \ \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$$

if and only if

$$\exists \boldsymbol{r}^0(\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},g}([\lambda\, \psi]) \ \& \ \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$$

So we show that the witnesses to the corresponding existential claims can be identified. In *Case A*, we know by hypothesis that the κ_i are simple, and so by previous reasoning, we know that $\boldsymbol{d}_{\mathcal{I},f}(\kappa_i) = \boldsymbol{d}_{\mathcal{I},g}(\kappa_i)$, for $1 \leq i \leq n$. So it remains to show:

Case A. $\boldsymbol{d}_{\mathcal{I},f}([\lambda \nu_1 \ldots \nu_n\, \psi]) = \boldsymbol{d}_{\mathcal{I},g}([\lambda \nu_1 \ldots \nu_n\, \psi])$

Case B. $\boldsymbol{d}_{\mathcal{I},f}([\lambda\, \psi]) = \boldsymbol{d}_{\mathcal{I},g}([\lambda\, \psi])$

But the above follow from Constraints (2.1) and (2.2), respectively, on our semantics. Thus, both of our cases reduce to the reasoning in the Term Induction Base Case: in each case, every term τ in φ is such that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$, and so, for both biconditionals (ϑ1) and (ϑ2), any witnesses to the quantified claims on one side are witnesses to corresponding quantified claims on the other side.

Formula Induction: Inductive Case 1. φ is $\neg\psi$, $\psi \rightarrow \chi$, or $\Box\psi$. These cases follow straightforwardly from the inductive hypothesis.

Formula Induction: Inductive Case 2. φ is $\forall\alpha\psi$. Then the free variables of φ are those free in ψ with the exception of α. So for any arbitarily chosen entity $\boldsymbol{e}$ in the domain of α, the assignments $f[\alpha/\boldsymbol{e}]$ and $g[\alpha/\boldsymbol{e}]$ agree on all the free variables in ψ. So, by our inductive hypothesis:

$$\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \psi \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},g[\alpha/\boldsymbol{e}]} \psi$$

But since $\boldsymbol{e}$ was arbitrary, we have:

$$\forall \boldsymbol{e} \in \mathrm{dom}(\alpha)[\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \psi \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},g[\alpha/\boldsymbol{e}]} \psi]$$

But this implies:

$$[\forall \boldsymbol{e} \in \mathrm{dom}(\alpha)(\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \psi)] \text{ if and only if } [\forall \boldsymbol{e} \in \mathrm{dom}(\alpha)(\boldsymbol{w} \models_{\mathcal{I},g[\alpha/\boldsymbol{e}]} \psi)]$$

By T5, it follows that:

$$\boldsymbol{w} \models_{\mathcal{I},f} \forall\alpha\psi \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},g} \forall\alpha\psi$$

i.e.,

$$\boldsymbol{w} \models_{\mathcal{I},f} \varphi \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},g} \varphi \qquad \bowtie$$

6.3 Corollary to the Assignment Agreement Lemma

Corollary. If assignments f and g agree on the free variables in term τ, then $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$.

Proof. This is established by appealing to, or repurposing, some of the reasoning in, the Assignment Agreement Lemma. Assume assignments f and g agree on the free variables in τ.

Case A. τ is a constant. This was established in the Term Induction: Base Case of the Assignment Agreement Lemma.

Case B. τ is a variable. This was established in the Term Induction: Base Case of the Assignment Agreement Lemma.

Case C. τ is a definite description of the form $\iota\nu\varphi$. We establish this by adapting the reasoning in Inductive Case 1 of the Term Induction in the Assignment Agreement Lemma. Assume f and g agree on the free variables in the term $\iota\nu\psi$. Now by D3, we know:

$$\boldsymbol{d}_{\mathcal{I},f}(\iota\nu\psi) = \begin{cases} \boldsymbol{o} \text{ if } \boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}]} \psi \ \& \ \forall \boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}']} \psi \rightarrow \boldsymbol{o}' = \boldsymbol{o}) \\ \text{undefined, otherwise} \end{cases}$$

But the free variables in $\iota\nu\psi$ are those in ψ with the exception of ν. So, for any object $\boldsymbol{o}''$, the assignments $f[\nu/\boldsymbol{o}'']$ and $g[\nu/\boldsymbol{o}'']$ agree on all the free variables in ψ. Consequently, by the Assignment Agreement Lemma, it follows that:

$$\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}]} \psi \text{ if and only if } \boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{o}]} \psi$$

It also follows that:

$$\forall \boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}']} \psi \rightarrow \boldsymbol{o}' = \boldsymbol{o}) \text{ iff } \forall \boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{o}']} \psi \rightarrow \boldsymbol{o}' = \boldsymbol{o})$$

So, we have established:

$$\boldsymbol{d}_{\mathcal{I},f}(\iota\nu\psi) = \begin{cases} \boldsymbol{o} \text{ if } \boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{o}]} \psi \ \& \ \forall \boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},g[\nu/\boldsymbol{o}']} \psi \rightarrow \boldsymbol{o}' = \boldsymbol{o}) \\ \text{undefined, otherwise} \end{cases}$$

i.e., by D3, that:

$\boldsymbol{d}_{\mathcal{I},f}(\imath\nu\psi) = \boldsymbol{d}_{\mathcal{I},g}(\imath\nu\psi)$

Case D. τ is a complex n-place relation term, i.e., a λ-expression of the form $[\lambda\nu_1 \ldots \nu_n\, \varphi]$, for $n \geq 0$. Then our corollary holds in virtue of Constraints (2.1) and (2.2) on our semantics, from which it follows that when f and g agree on the free variables in complex n-place relation term τ, $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},g}(\tau)$. ⋈

6.4 Substitutions and Substitutable For

Substitutions

Where τ is any term and α any variable, we use the notation φ^τ_α and ρ^τ_α, respectively, to stand for the result of substituting the term τ for every free occurrence of the variable α in formula φ and in term ρ. This notion may be defined recursively based on the complexity of ρ and φ as follows:

- If ρ is a constant or variable other than α, $\rho^\tau_\alpha = \rho$.
  If ρ is α, $\rho^\tau_\alpha = \tau$
- If φ is $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 0)$, then $\varphi^\tau_\alpha = \Pi^{n\,\tau}_{\ \alpha}\kappa_1{}^\tau_\alpha \ldots \kappa_n{}^\tau_\alpha$.
- If φ is $\neg\psi$ or $\Box\psi$, then $\varphi^\tau_\alpha = \neg(\psi^\tau_\alpha)$ or $\Box(\psi^\tau_\alpha)$, respectively.
  If φ is $\psi \rightarrow \chi$, then $\varphi^\tau_\alpha = \psi^\tau_\alpha \rightarrow \chi^\tau_\alpha$.
- If φ is $\forall\beta\psi$, then $\varphi^\tau_\alpha = \begin{cases} \forall\beta\psi, \text{ if } \alpha = \beta \\ \forall\beta(\psi^\tau_\alpha), \text{ if } \alpha \neq \beta \end{cases}$
- If ρ is $\imath\nu\psi$, then $\rho^\tau_\alpha = \begin{cases} \imath\nu\psi, \text{ if } \alpha = \nu \\ \imath\nu(\psi^\tau_\alpha), \text{ if } \alpha \neq \nu \end{cases}$
- If ρ is $[\lambda\nu_1 \ldots \nu_n\, \psi]$, then
  $\rho^\tau_\alpha = \begin{cases} [\lambda\nu_1 \ldots \nu_n\, \psi], \text{ if } \alpha \text{ is one of } \nu_1, \ldots, \nu_n \\ [\lambda\nu_1 \ldots \nu_n\, \psi^\tau_\alpha], \text{ if } \alpha \text{ is none of } \nu_1, \ldots, \nu_n \end{cases}$

Substitutable For

We recursively define term τ *is substitutable for* the variable α *in* formula φ or *in* term ρ as follows:

- If ρ is a constant or a variable, then τ is substitutable for α in ρ iff (i) ρ is α and (ii) τ and α are terms of the same type.

- If φ is $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 0)$, then τ is substitutable for α in φ iff τ is substitutable for α in one of the terms Π^n, or $\kappa_1, \ldots,$ or κ_n.
- If φ is $\neg\psi$ or $\Box\psi$, then τ is substitutable for α in φ iff τ is substitutable for α in ψ; if φ is $\psi \rightarrow \chi$, then τ is substitutable for α in φ iff τ is substitutable for α in ψ and χ.
- If φ is $\forall\beta\psi$, then τ is substitutable for α in φ iff either α does not occur free in φ or both (i) β does not free occur in τ and (ii) τ is substitutable for α in ψ.
- If ρ is $\imath\nu\psi$, then τ is substitutable for α in ρ iff either α does not occur free in ρ or both (i) ν does not occur free in τ and (ii) τ is substitutable for α in ψ.
- If ρ is $[\lambda\nu_1 \ldots \nu_n\psi]$ $(n \geq 0)$, then τ is substitutable for α in ρ iff either α does not occur free in ρ or both (i) $\nu_1, \ldots, \nu_n$ do not occur free in τ and (ii) τ is substitutable for α in ψ.

Intuitively, term τ *is substitutable for* the variable α *in* formula φ (term ρ) if and only if (a) τ and α are terms of the same type, and (b) every occurrence of a variable β free in τ remains an occurrence that is free in ρ^τ_α. So when τ is substitutable for the variable α in φ (or in ρ), no occurrence of a variable $\beta, \nu, \nu_1, \ldots, \nu_n$ free in τ becomes bound ('captured') by a variable-binding operator such as $\forall\beta$, $\imath\nu$, or $\lambda\nu_1 \ldots \nu_n$ in φ (or ρ) when τ is substituted for the free occurrences of α in φ (or ρ).

6.5 Substitution Lemma

Lemma. If τ is substitutable for α in φ and $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{e}$, where $\boldsymbol{e}$ is an entity in the domain of the variable α, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi^\tau_\alpha$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$.

Example. If the constant a is substitutable for the variable x in Fx $(= \varphi)$ and $\boldsymbol{d}_{\mathcal{I},f}(a) = \boldsymbol{o}$, where $\boldsymbol{o}$ is an entity in the domain of the variable x, then $\boldsymbol{w} \models_{\mathcal{I},f} Fa$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[x/\boldsymbol{o}]} Fx$.

Proof: By induction on the complexity of φ with the proof of the base case proceeding by a secondary induction on the complexity of terms τ' occurring in φ. Assume τ is substitutable for α in φ and $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{e}$, where $\boldsymbol{e}$ is an entity in the domain of the variable α. Note that by the definition of *substitutable for* and φ^τ_α, τ is substitutable for α in the edge case where there are no free occurrences of α in φ. But in that case, $\varphi^\tau_\alpha = \varphi$, and so the lemma asserts that $\boldsymbol{w} \models_{\mathcal{I},f} \varphi$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$. If α has no free occurrences in φ, then clearly, f and $f[\alpha/\boldsymbol{e}]$ agree on all the free variables in φ, so the lemma follows by the Assignment Agreement Lemma (Section 6.2). So, in what follows, we presume that α is free in φ.

Formula Induction: Base Case. We have two cases:

Case A. φ has the form $\Pi^n\kappa_1\ldots\kappa_n$ $(n \geq 1)$

Case B. φ has the form Π^0

Term Induction: Base Case. All of the individual and relation terms τ' in φ are simple, i.e., constants or variables.

• *Case A*. Suppose φ has the form $\Pi^n\kappa_1\ldots\kappa_n$, and all of the terms are constants or variables. Since we've covered the case where α isn't free in φ, we may suppose at least one of the terms in the present case is α. So either (i) α is one or more of the κ_i and Π^n is simple or (ii) α is Π^n and the κ_i are simple.

• *Case A*(i). α is one or more of the κ_i and Π^n is simple. Without loss of generality, suppose α is κ_1 and not one of the other κ_i. Then τ is an object term and $\kappa_1{}^\tau_\alpha = \tau$. Since α is not any of the other terms in the formula, we know $\kappa_i{}^\tau_\alpha = \kappa_i$, for $2 \leq i \leq n$, and $\Pi^n{}^\tau_\alpha = \Pi^n$. So φ^τ_α is $\Pi^n\tau\kappa_2\ldots\kappa_n$ and φ is $\Pi^n\alpha\kappa_2\ldots\kappa_n$. Consequently, to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \varphi^\tau_\alpha \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$$

we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \Pi^n\tau\kappa_2\ldots\kappa_n \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \Pi^n\alpha\kappa_2\ldots\kappa_n$$

By T1, this means we have to show:

$$(\vartheta)\ \exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\ldots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\tau)\ \&\ \boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_2)\ \&\ \ldots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$$
$$\text{if and only if}$$
$$\exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\ldots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)\ \&\ \boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_2)\ \&\ \ldots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$$

Since Π^n is a constant or a variable other than α, we know (by D1 if Π^n is a constant, or by D2 if Π^n is a variable other than α) that $\boldsymbol{d}_{\mathcal{I},f}(\Pi^n) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\Pi^n)$. Similar reasoning applies to the $\kappa_2,\ldots,\kappa_n$: since they are constants or variables other than α, it follows by D1 (if they are constants) or by D2 (if they are variables other than α) that $\boldsymbol{d}_{\mathcal{I},f}(\kappa_i) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i)$, for $2 \leq i \leq n$. So if we can show $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)$, we're done, for this would establish that any witnesses to any quantifier on one side of (ϑ) would be witnesses to the corresponding quantifier on the other side of (ϑ). But by hypothesis, $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{e}$ and by definition, $\boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha) = \boldsymbol{e}$.

• *Case A*(ii). Π^n is α and $\kappa_1,\ldots,\kappa_n$ are all simple. Then τ is an n-place relation term and $\Pi^n{}^\tau_\alpha = \tau$. Moreover, $\kappa_i{}^\tau_\alpha = \kappa_i$, for $1 \leq i \leq n$. So φ^τ_α is $\tau\kappa_1\ldots\kappa_n$ and φ is $\alpha\kappa_1\ldots\kappa_n$. Consequently, we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \tau\kappa_1\ldots\kappa_n \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \alpha\kappa_1\ldots\kappa_n$$

By T1, this means we have to show:

$$(\zeta)\ \exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\ldots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}(\tau)\ \&\ \boldsymbol{o}_i = \boldsymbol{d}_{\mathcal{I},f}(\kappa_i)\ \&\ \langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$$
$$\text{if and only if}$$
$$\exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\ldots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)\ \&\ \boldsymbol{o}_i = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i)\ \&\ \langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$$

But this follows by reasoning analogous to the previous case: we know both (1) that $\boldsymbol{d}_{\mathcal{I},f}(\tau)$ = (by hypothesis) $\boldsymbol{e}$ = (by definition) $\boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)$, and (2) that since the κ_i are simple and not α, it follows that $\boldsymbol{d}_{\mathcal{I},f}(\kappa_i) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i)$, for $1 \leq i \leq n$. So any witnesses to one side of the biconditional are witnesses to the other side.

• *Case B*. Suppose φ has the form Π^0, and Π^0 is a constant or a variable. But since α is free in φ, Π^0 must be α. So we know all of the following: that α is a 0-place relation variable, that τ is a 0-place relation term, that $\varphi^\tau_\alpha = \tau$, and that $\varphi = \alpha$. Consequently, we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \tau \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \alpha$$

By T2, we have to show:

$$\exists \boldsymbol{r}^0(\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},f}(\tau)\ \&\ \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$$
$$\text{if and only if}$$
$$\exists \boldsymbol{r}^0(\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)\ \&\ \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$$

But by hypothesis, $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{e}$ and by definition, $\boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha) = \boldsymbol{e}$. So any witness to one side of this biconditional is a witness to the other.

Term Induction: Inductive Case 1. There is only one case: φ is $\Pi^n\kappa_1\ldots\kappa_n$ $(n \geq 1)$, where one or more of the κ_i is a description of the form $\iota\nu\psi$ and the remaining terms are all simple. (If φ is of the form Π^0, then $n = 0$ and there are no κ_i.) Without loss of generality, assume κ_1 is $\iota\nu\psi$, and the other κ_i and Π^1 are simple. So φ is $\Pi^n\iota\nu\psi\kappa_2\dots\kappa_n$ and there are two cases: (A) α isn't free in $\iota\nu\psi$, and (B) α is free in $\iota\nu\psi$.

• *Case A*. If α isn't free in $\iota\nu\psi$, then since we know α is free in φ, α must be either Π^n or one or more of $\kappa_2,\dots,\kappa_n$.

• *Case A(i)*. Suppose α is Π^n. Then φ^τ_α is $\tau\iota\nu\psi\kappa_2\dots\kappa_n$ and φ is $\alpha\iota\nu\psi\kappa_2\dots\kappa_n$. Consequently, to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \varphi^\tau_\alpha \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$$

we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \tau\iota\nu\psi\kappa_2\dots\kappa_n \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \alpha\iota\nu\psi\kappa_2\dots\kappa_n$$

i.e., by T1, that:

$$\begin{aligned}&\exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\dots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}(\tau)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\iota\nu\psi)\ \&\\ &\boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_2)\ \&\ \dots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1,\dots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))\\ &\quad\text{if and only if}\\ &\exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\dots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\iota\nu\psi)\ \&\\ &\boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_2)\ \&\ \dots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1,\dots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))\end{aligned}$$

If we can establish that the witnesses to the corresponding existential claims on both sides of the biconditional are identical, we are done. First, by previous reasoning, we established that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)$. Second, since the κ_i, $2 \le i \le n$, are all simple, previous reasoning has also established that $\boldsymbol{d}_{\mathcal{I},f}(\kappa_i) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i)$, for $2 \le i \le n$. So it remains to show: $\boldsymbol{d}_{\mathcal{I},f}(\iota\nu\psi) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\iota\nu\psi)$. However, since α is Π^n and α is not free in $\iota\nu\psi$, we know f and $f[\alpha/\boldsymbol{e}]$ agree on all the free variables in $\iota\nu\psi$. So by the Corollary to the Assignment Agreement Lemma (Section 6.3), it follows that $\boldsymbol{d}_{\mathcal{I},f}(\iota\nu\psi) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\iota\nu\psi)$.

• *Case A(ii)*. Suppose α is one or more of $\kappa_2,\dots,\kappa_n$. Without loss of generality, suppose α is κ_2. Then φ^τ_α is $\Pi^n\iota\nu\psi\tau\kappa_3\dots\kappa_n$ and φ is $\Pi^n\iota\nu\psi\alpha\kappa_3\dots\kappa_n$. Consequently, to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \varphi^\tau_\alpha \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$$

we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \Pi^n\iota\nu\psi\tau\kappa_3\dots\kappa_n \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \Pi^n\iota\nu\psi\alpha\kappa_3\dots\kappa_n$$

i.e., by T1, that:

$$\begin{aligned}&\exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\dots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\iota\nu\psi)\ \&\ \boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f}(\tau)\ \&\\ &\boldsymbol{o}_3 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_3)\ \&\ \dots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1,\dots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))\\ &\quad\text{if and only if}\\ &\exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\dots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\iota\nu\psi)\ \&\\ &\boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)\ \&\ \boldsymbol{o}_3 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_3)\ \&\ \dots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_n)\ \&\\ &\langle \boldsymbol{o}_1,\dots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))\end{aligned}$$

Again, if we can establish that the witnesses to the corresponding existential claims on both sides of the biconditional are identical, we are done. First, since Π^n is simple, then by previous reasoning, we know $\boldsymbol{d}_{\mathcal{I},f}(\Pi^n) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\Pi^n)$. Second, by previous reasoning, we established that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)$. Third, since the κ_i ($3 \le i \le n$) are all simple, previous reasoning has also established that $\boldsymbol{d}_{\mathcal{I},f}(\kappa_i) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i)$, for $2 \le i \le n$. So again it remains to show: $\boldsymbol{d}_{\mathcal{I},f}(\iota\nu\psi) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\iota\nu\psi)$. However, since α is κ_2 and α is not free in $\iota\nu\psi$, we know f and $f[\alpha/\boldsymbol{e}]$ agree on all the free variables in $\iota\nu\psi$. So by the Corollary to the Assignment Agreement Lemma (Section 6.3), it follows that $\boldsymbol{d}_{\mathcal{I},f}(\iota\nu\psi) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\iota\nu\psi)$.

• *Case B*. If α is free in $\iota\nu\psi$, then φ^τ_α is $\Pi^{n}{}^\tau_\alpha[\iota\nu\psi]^\tau_\alpha\kappa_2{}^\tau_\alpha\dots\kappa_n{}^\tau_\alpha$, where we use the square brackets in $[\iota\nu\psi]^\tau_\alpha$ to help indicate the result of substituting τ for every free occurrence of α in $\iota\nu\psi$. Furthermore, since α is free in $\iota\nu\psi$, α can't be ν, so that $[\iota\nu\psi]^\tau_\alpha$ is $\iota\nu(\psi^\tau_\alpha)$. So φ^τ_α is $\Pi^{n}{}^\tau_\alpha\iota\nu(\psi^\tau_\alpha)\kappa_2{}^\tau_\alpha\dots\kappa_n{}^\tau_\alpha$. Consequently, to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \varphi^\tau_\alpha \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$$

we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \Pi^{n}{}^\tau_\alpha\iota\nu(\psi^\tau_\alpha)\kappa_2{}^\tau_\alpha\dots\kappa_n{}^\tau_\alpha \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \Pi^n\iota\nu\psi\kappa_2\dots\kappa_n$$

i.e., by T1, that:

$$\begin{aligned}&\exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\dots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}(\Pi^{1}{}^\tau_\alpha)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\iota\nu(\psi^\tau_\alpha))\ \&\\ &\boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_2{}^\tau_\alpha)\ \&\ \dots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n{}^\tau_\alpha)\ \&\ \langle \boldsymbol{o}_1,\dots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))\\ &\quad\text{if and only if}\\ &\exists \boldsymbol{r}^n\exists \boldsymbol{o}_1\dots\exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\Pi^n)\ \&\ \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\iota\nu\psi)\ \&\\ &\boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_2)\ \&\ \dots\ \&\ \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_n)\ \&\ \langle \boldsymbol{o}_1,\dots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))\end{aligned}$$

But Π^n is simple and so no matter whether Π^n is or isn't another occurrence of α, it follows by D1 (if Π^n is a constant), or D2 (if Π^n is a variable other than α), or previous reasoning (if Π^n is α) that $\boldsymbol{d}_{\mathcal{I},f}(\Pi^{1}{}^\tau_\alpha)$

$= \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\Pi^1)$. Similarly, each κ_i, where $2 \leq i \leq n$, is simple and so no matter whether κ_i is or isn't another occurrence of α, it follows by D1 (if κ_i is a constant), or D2 (if κ_i is a variable other than α), or previous reasoning (if κ_i is α) that $\boldsymbol{d}_{\mathcal{I},f}(\kappa_i{}^\tau_\alpha) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i)$. So to show the above biconditional, it remains to show $\boldsymbol{d}_{\mathcal{I},f}(\imath\nu(\psi^\tau_\alpha)) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\imath\nu\psi)$. Now by D3, we know both:

(ζ1) $\boldsymbol{d}_{\mathcal{I},f}(\imath\nu(\psi^\tau_\alpha)) =$

$$\begin{cases} \boldsymbol{o} \text{ if } \boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}]} \psi^\tau_\alpha \;\&\; \forall \boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}']} \psi^\tau_\alpha \rightarrow \boldsymbol{o}' = \boldsymbol{o}) \\ \text{undefined, otherwise} \end{cases}$$

(ζ2) $\boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\imath\nu\psi) =$

$$\begin{cases} \boldsymbol{o} \text{ if } \boldsymbol{w}_0 \models_{\mathcal{I},f[\alpha/\boldsymbol{e}][\nu/\boldsymbol{o}]} \psi \;\&\; \forall \boldsymbol{o}'(\boldsymbol{w}_0 \models_{\mathcal{I},f[\alpha/\boldsymbol{e}][\nu/\boldsymbol{o}']} \psi \rightarrow \boldsymbol{o}' = \boldsymbol{o}) \\ \text{undefined, otherwise} \end{cases}$$

So, if we can show that:

(ξ) $\forall \boldsymbol{o}''[\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{o}'']} \psi^\tau_\alpha$ if and only if $\boldsymbol{w}_0 \models_{\mathcal{I},f[\alpha/\boldsymbol{e}][\nu/\boldsymbol{o}'']} \psi]$

we are done, for this would allow us to transform the right side of the identity in (ζ1) into the right side of the identity in (ζ2), and vice versa, thereby establishing $\boldsymbol{d}_{\mathcal{I},f}(\imath\nu(\psi^\tau_\alpha)) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\imath\nu\psi)$. Now we can show ($\xi$) by picking an arbitrarily chosen object $\boldsymbol{b}$, and showing:

(ξ') $\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{b}]} \psi^\tau_\alpha$ if and only if $\boldsymbol{w}_0 \models_{\mathcal{I},f[\alpha/\boldsymbol{e}][\nu/\boldsymbol{b}]} \psi$

Now note that our inductive hypothesis is, for any assignment g:

(ϑ) If τ' is substitutable for β in ψ and $\boldsymbol{d}_{\mathcal{I},g}(\tau') = \boldsymbol{e}'$, then $\boldsymbol{w} \models_{\mathcal{I},g} \psi^{\tau'}_\beta$ if and only if $\boldsymbol{w} \models_{\mathcal{I},g[\beta/\boldsymbol{e}']} \psi$

So if we let τ' be τ, let β be α, and let $\boldsymbol{e}'$ be $\boldsymbol{e}$, and let g be $f[\nu/\boldsymbol{b}]$, then our inductive hypothesis yields that:

(ϑ') If τ is substitutable for α in ψ and $\boldsymbol{d}_{\mathcal{I},f[\nu/\boldsymbol{b}]}(\tau) = \boldsymbol{e}$, then $\boldsymbol{w} \models_{\mathcal{I},f[\nu/\boldsymbol{b}]} \psi^\tau_\alpha$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\nu/\boldsymbol{b}][\alpha/\boldsymbol{e}]} \psi$

To detach the biconditional in the consequent of (ϑ'), note *first* that, by hypothesis, τ is substitutable for α in φ. Since, in the present case, α is free in φ and α is free in $\imath\nu\psi$, it follows that τ is substitutable for α in $\imath\nu\psi$. (For otherwise, if a variable free in τ were captured when τ is substituted for α in $\imath\nu\psi$, that variable would be captured when τ is substituted for all the free occurrences of α in φ, contradicting τ's substitutability for α in φ.) So by definition of *substitutable for*, we know:

τ is substitutable for α in ψ

It also follows from the fact that τ is substitutable for α in $\imath\nu\psi$, when α is free in $\imath\nu\psi$, that ν isn't free in τ. So f and $f[\nu/\boldsymbol{b}]$ agree on the free variables in τ, and by the Corollary to the Assignment Agreement Lemma (Section 6.3), it follows that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f[\nu/\boldsymbol{b}]}(\tau)$. But now note *second* that the other conjunct of our very first hypothesis is that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{e}$. So it follows that:

$$\boldsymbol{d}_{\mathcal{I},f[\nu/\boldsymbol{b}]}(\tau) = \boldsymbol{e}.$$

So, we may infer from our last two displayed conclusions and (ϑ') that:

$\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{b}]} \psi^\tau_\alpha$ if and only if $\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{b}][\alpha/\boldsymbol{e}]} \psi$

But, by definition, $f[\nu/\boldsymbol{b}][\alpha/\boldsymbol{e}] = f[\alpha/\boldsymbol{e}][\nu/\boldsymbol{b}]$. So we may transform our last conclusion into (ξ'):

(ξ') $\boldsymbol{w}_0 \models_{\mathcal{I},f[\nu/\boldsymbol{b}]} \psi^\tau_\alpha$ if and only if $\boldsymbol{w}_0 \models_{\mathcal{I},f[\alpha/\boldsymbol{e}][\nu/\boldsymbol{b}]} \psi$

And this is what we had to show to prove $\boldsymbol{d}_{\mathcal{I},f}(\imath\nu(\psi^\tau_\alpha)) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\imath\nu\psi)$ and thereby complete the proof of *Case B*.

Term Induction: Inductive Case 2. There are two cases: (A) φ has the form $\Pi^n\kappa_1\ldots\kappa_n$, where Π^n is $[\lambda\nu_1\ldots\nu_n\,\psi]$ and the κ_i are simple, and (B) φ has the form Π^0, where Π^0 is $[\lambda\,\psi]$.

• *Case A*. φ has the form $\Pi^n\kappa_1\ldots\kappa_n$, where Π^n is $[\lambda\nu_1\ldots\nu_n\,\psi]$ and the κ_i are simple. Then there are two further cases: (i) α isn't free in $[\lambda\nu_1\ldots\nu_n\,\psi]$, and ($ii$) α is free in $[\lambda\nu_1\ldots\nu_n\,\psi]$.

• *Case A*(i). If α isn't free in $[\lambda\nu_1\ldots\nu_n\,\psi]$, then α must be one or more of the κ_i (since we know α is free in φ and the κ_i are all simple). Without loss of generality, suppose α is κ_1 and all the other κ_i are constants or variables other than α. Then φ^τ_α is $[\lambda\nu_1\ldots\nu_n\,\psi]\tau\kappa_2\ldots\kappa_n$ and φ is $[\lambda\nu_1\ldots\nu_n\,\psi]\alpha\kappa_2\ldots\kappa_n$. Consequently, to show:

$\boldsymbol{w} \models_{\mathcal{I},f} \varphi^\tau_\alpha$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$

we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} [\lambda\nu_1 \ldots \nu_n\, \psi]\tau\kappa_2 \ldots \kappa_n$$
if and only if
$$\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} [\lambda\nu_1 \ldots \nu_n\, \psi]\alpha\kappa_2 \ldots \kappa_n$$

i.e., by T1, that:

$$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}([\lambda\nu_1 \ldots \nu_n\, \psi]) \;\&\; \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f}(\tau) \;\&\; \boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f}(\kappa_2) \;\&\; \ldots \;\&\; \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f}(\kappa_n) \;\&\; \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$$
if and only if
$$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}([\lambda\nu_1 \ldots \nu_n\, \psi]) \;\&\; \boldsymbol{o}_1 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha) \;\&\; \boldsymbol{o}_2 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_2) \;\&\; \ldots \;\&\; \boldsymbol{o}_n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_n) \;\&\; \langle \boldsymbol{o}, \ldots \boldsymbol{o}_n\rangle \in \mathbf{en}(\boldsymbol{r}^n))$$

Now by previous reasoning, we know both that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\alpha)$ and that, for $2 \leq i \leq n$, $\boldsymbol{d}_{\mathcal{I},f}(\kappa_i) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i)$. So our biconditional is proved if we can show: $\boldsymbol{d}_{\mathcal{I},f}([\lambda\nu_1 \ldots \nu_n\, \psi]) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}([\lambda\nu_1 \ldots \nu_n\, \psi])$. However, since α is not free in $[\lambda\nu_1 \ldots \nu_n\, \psi]$, we know f and $f[\alpha/\boldsymbol{e}]$ agree on all the free variables in $[\lambda\nu_1 \ldots \nu_n\, \psi]$. So by the Corollary to the Assignment Agreement Lemma (Section 6.3), it follows that $\boldsymbol{d}_{\mathcal{I},f}([\lambda\nu_1 \ldots \nu_n\, \psi]) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}([\lambda\nu_1 \ldots \nu_n\, \psi])$.

• *Case A(ii)*. If α is free in $[\lambda\nu_1 \ldots \nu_n\, \psi]$, then φ^τ_α is:

$$[\lambda\nu_1 \ldots \nu_n\, \psi]^\tau_\alpha \kappa_1{}^\tau_\alpha \ldots \kappa_n{}^\tau_\alpha.$$

Furthermore, since α is free in $[\lambda\nu_1 \ldots \nu_n\, \psi]$, α can't be any of the ν_i, so that $[\lambda\nu_1 \ldots \nu_n\, \psi]^\tau_\alpha$ is $[\lambda\nu_1 \ldots \nu_n\, \psi^\tau_\alpha]$. So φ^τ_α is:

$$[\lambda\nu_1 \ldots \nu_n\, \psi^\tau_\alpha]\kappa_1{}^\tau_\alpha \ldots \kappa_n{}^\tau_\alpha$$

Consequently, to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} \varphi^\tau_\alpha \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \varphi$$

we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} [\lambda\nu_1 \ldots \nu_n\, \psi^\tau_\alpha]\kappa_1{}^\tau_\alpha \ldots \kappa_n{}^\tau_\alpha$$
if and only if
$$\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} [\lambda\nu_1 \ldots \nu_n\, \psi]\kappa_1 \ldots \kappa_n$$

i.e., by T1, that:

(ϑ) $\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f}([\lambda\nu_1 \ldots \nu_n\, \psi^\tau_\alpha]) \;\&\; \boldsymbol{o}_i = \boldsymbol{d}_{\mathcal{I},f}(\kappa_i{}^\tau_\alpha) \;\&\; \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$
if and only if
$\exists \boldsymbol{r}^n \exists \boldsymbol{o}_1 \ldots \exists \boldsymbol{o}_n(\boldsymbol{r}^n = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}([\lambda\nu_1 \ldots \nu_n\, \psi]) \;\&\; \boldsymbol{o}_i = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i) \;\&\; \langle \boldsymbol{o}_1, \ldots, \boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n))$

Since the κ_i are all simple, then no matter whether κ_i is a constant, a variable other than α, or α itself, we know from previous reasoning that $\boldsymbol{d}_{\mathcal{I},f}(\kappa_i{}^\tau_\alpha) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}(\kappa_i)$, for $1 \leq i \leq n$. So it remains to show:

(ζ) $\boldsymbol{d}_{\mathcal{I},f}([\lambda\nu_1 \ldots \nu_n\, \psi^\tau_\alpha]) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}([\lambda\nu_1 \ldots \nu_n\, \psi])$

But (ζ) follows from Constraint (3.1) of our semantics (Section 4.5) when $m = 1$.

• *Case B*. φ has the form Π^0, where Π^0 is $[\lambda\, \psi]$. Since α is free in φ, we know α is free in $[\lambda\, \psi]$. So φ^τ_α is $[\lambda\, \psi]^\tau_\alpha$, which in turn, by definition, is $[\lambda\, \psi^\tau_\alpha]$. So, we have to show:

$$\boldsymbol{w} \models_{\mathcal{I},f} [\lambda\, \psi^\tau_\alpha] \text{ if and only if } \boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} [\lambda\, \psi]$$

i.e., by T2, that:

(ϑ) $\exists \boldsymbol{r}^0(\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},f}([\lambda\, \psi^\tau_\alpha]) \;\&\; \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$
if and only if
$\exists \boldsymbol{r}^0(\boldsymbol{r}^0 = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}([\lambda\, \psi]) \;\&\; \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^0) = \mathbf{True})$

To establish (ϑ), it remains to show:

(ζ) $\boldsymbol{d}_{\mathcal{I},f}([\lambda\, \psi^\tau_\alpha]) = \boldsymbol{d}_{\mathcal{I},f[\alpha/\boldsymbol{e}]}([\lambda\, \psi])$

But (ζ) follows from Constraint (3.2) of our semantics (Section 4.5) when $m = 1$.

Formula Induction: Inductive Case 1. φ is $\neg\psi$, $\psi \rightarrow \chi$, or $\Box\psi$. These cases follow straightforwardly from the inductive hypothesis.

Formula Induction: Inductive Case 2. φ is $\forall\beta\psi$. There is only one case, namely, α is free in φ. (The case where α is not free in φ was covered at the very outset.) So $\alpha \neq \beta$, and hence φ^τ_α is $\forall\beta(\psi^\tau_\alpha)$. So we have to show:

(ϑ) $\boldsymbol{w} \models_{\mathcal{I},f} \forall\beta(\psi^\tau_\alpha)$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \forall\beta\psi$

By T5, we know the following about the left condition of (ϑ):

($\zeta1$) $\boldsymbol{w} \models_{\mathcal{I},f} \forall\beta(\psi^\tau_\alpha)$ if and only if $\forall \boldsymbol{e}' \in \mathrm{dom}(\beta)(\boldsymbol{w} \models_{\mathcal{I},f[\beta/\boldsymbol{e}']} \psi^\tau_\alpha)$

Moreover, by T5, we know the following about the right condition of (ϑ):

($\zeta2$) $\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}]} \forall\beta\psi$ if and only if $\forall \boldsymbol{e}' \in \mathrm{dom}(\beta)(\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}][\beta/\boldsymbol{e}']} \psi)$

So if we can show, for an arbitrary entity $\boldsymbol{b}$ in the domain of β that:

(ξ) $\boldsymbol{w} \models_{\mathcal{I},f[\beta/\boldsymbol{b}]} \psi^{\tau}_{\alpha}$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\alpha/\boldsymbol{e}][\beta/\boldsymbol{b}]} \psi$

we are done, for that would allow us to transform the right condition of (ζ1) into the right condition of (ζ2), and vice versa, thereby establishing (ϑ). Note that where $\boldsymbol{e}''$ is any entity in the domain of the variable γ and g any assignment, our inductive hypothesis is:

(H) if τ' is substitutable for γ in ψ and $\boldsymbol{d}_{\mathcal{I},g}(\tau') = \boldsymbol{e}''$, then $\boldsymbol{w} \models_{\mathcal{I},g} \psi^{\tau'}_{\gamma}$ if and only if $\boldsymbol{w} \models_{\mathcal{I},g[\gamma/\boldsymbol{e}'']} \psi$.

So if we let τ' be τ, g be $f[\beta/\boldsymbol{b}]$, γ be α, and $\boldsymbol{e}''$ be $\boldsymbol{e}$, our inductive hypothesis implies:

(H′) if τ is substitutable for α in ψ and $\boldsymbol{d}_{\mathcal{I},f[\beta/\boldsymbol{b}]}(\tau) = \boldsymbol{e}$, then $\boldsymbol{w} \models_{\mathcal{I},f[\beta/\boldsymbol{b}]} \psi^{\tau}_{\alpha}$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\beta/\boldsymbol{b}][\alpha/\boldsymbol{e}]} \psi$.

If we can detach the biconditional in the consequent of (H′), we'll be a step away from proving (ξ). To this end, note *first* that, by hypothesis, τ is substitutable for α in φ. Since α is free in φ, it follows from the definition of *substitutable for* that:

τ is substitutable for α in ψ.

It also follows from the fact that τ is substitutable for α in φ, when α is free in φ, that β doesn't occur free in τ. So f and $f[\beta/\boldsymbol{b}]$ agree on all the free variables in τ. So by the Corollary to the Assignment Agreement Lemma, it follows that $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{d}_{\mathcal{I},f[\beta/\boldsymbol{b}]}(\tau)$. But note *second* that, by hypothesis, $\boldsymbol{d}_{\mathcal{I},f}(\tau) = \boldsymbol{e}$. So it follows that:

$\boldsymbol{d}_{\mathcal{I},f[\beta/\boldsymbol{b}]}(\tau) = \boldsymbol{e}$.

From the two most recently displayed conclusions and (H′), we may derive the consequent of (H′):

$\boldsymbol{w} \models_{\mathcal{I},f[\beta/\boldsymbol{b}]} \psi^{\tau}_{\alpha}$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\beta/\boldsymbol{b}][\alpha/\boldsymbol{e}]} \psi$.

But since $f[\beta/\boldsymbol{b}][\alpha/\boldsymbol{e}]$ is identical to $f[\alpha/\boldsymbol{e}][\beta/\boldsymbol{b}]$, this last displayed conclusion is equivalent to (ξ), which is what we had to show to complete the proof of (ϑ). ⋈

6.6 Generalized Substitution Lemma

Where $\tau_1,\ldots,\tau_m$ are any terms and $\alpha_1,\ldots,\alpha_m$ are any variables, let $\varphi^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ stand for the result of simultaneously substituting the term τ_i for each free occurrence of the corresponding variable α_i in φ, for each i such that $1 \leq i \leq m$. In other words, $\varphi^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ is the result of making all of the following substitutions simultaneously: (a) substituting τ_1 for every free occurrence of α_1 in φ, (b) substituting τ_2 for every free occurrence of α_2 in φ, etc. Similarly, where $\tau_1,\ldots,\tau_m$ are any terms and $\alpha_1,\ldots,\alpha_m$ are any variables, we let $\rho^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ stand for the result of simultaneously substituting the term τ_i for each free occurrence of the corresponding variable α_i in ρ, for each i such that $1 \leq i \leq m$.[36] Then we have:

Lemma. If $\tau_1,\ldots,\tau_n$ are substitutable, respectively, for $\alpha_1,\ldots,\alpha_n$ in φ and $\boldsymbol{d}_{\mathcal{I},f}(\tau_1) = \boldsymbol{e}_1$ and … and $\boldsymbol{d}_{\mathcal{I},f}(\tau_n) = \boldsymbol{e}_n$, where the $\boldsymbol{e}_i$, respectively, are entities in the domain of the variable α_i, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n}$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[\alpha_i/\boldsymbol{e}_i]} \varphi$.

Example. If $x_1,\ldots,x_n$ are substitutable, respectively, for $y_1,\ldots,y_n$ in φ, and $\boldsymbol{d}_{\mathcal{I},f}(x_1) = \boldsymbol{o}_1$ and … and $\boldsymbol{d}_{\mathcal{I},f}(x_n) = \boldsymbol{o}_n$, then $\boldsymbol{w} \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots y_n}$ if and only if $\boldsymbol{w} \models_{\mathcal{I},f[y_i/\boldsymbol{o}_i]} \varphi$.

Proof: By generalizing the proof of the Substitution Lemma in Section 6.5. ⋈

6.7 β-Conversion is Valid

To prove the validity of β-Conversion in interpretations that satisfy Constraints (1) – (3), we show:

$\models [\lambda y_1 \ldots y_n \varphi]x_1 \ldots x_n \equiv \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$, provided $x_1,\ldots,x_n$ are substitutable, respectively, for $y_1,\ldots,y_n$ in φ

By working through the definitions of $\models \psi$, $\models_{\mathcal{I}} \psi$, $\models_{\mathcal{I},f} \psi$, and $\boldsymbol{w} \models_{\mathcal{I},f} \psi$ (and especially T4 of the latter), we have to show, for arbitrary $\mathcal{I}$ and f:

[36] The notions $\varphi^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ and $\rho^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m}$ can be defined more precisely, but the definition is somewhat tedious. For example, the base case for terms ρ is:

- If ρ is a constant or variable other than $\alpha_1,\ldots,\alpha_m$, $\rho^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m} = \rho$.
  If ρ is α_1, $\rho^{\tau_1,\ldots,\tau_n}_{\alpha_1,\ldots,\alpha_n} = \tau_1$

  ⋮

  If ρ is α_m, $\rho^{\tau_1,\ldots,\tau_m}_{\alpha_1,\ldots,\alpha_m} = \tau_m$

The base case for formulas φ is even more tedious. Since the various clauses would be extremely difficult to read but relatively straightforward to understand, we omit the remainder of the definition.

$$\boldsymbol{w}_0 \models_{\mathcal{I},f} [\lambda y_1 \ldots y_n\, \varphi]x_1 \ldots x_n \text{ iff } \boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$$

Since all of the terms in the left condition have denotations, we can expand the left condition by T1 and then simplify it, so that we have to prove:

$$\langle \boldsymbol{d}_{\mathcal{I},f}(x_1),\ldots,\boldsymbol{d}_{\mathcal{I},f}(x_n)\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n \varphi])) \text{ iff } \boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$$

For the remainder of the proof, we use the following arbitrary names for the denotations of the x_i:

$$\begin{aligned} &\boldsymbol{d}_{\mathcal{I},f}(x_1) = f(x_1) = \boldsymbol{o}_1 \\ &\quad\vdots \\ &\boldsymbol{d}_{\mathcal{I},f}(x_n) = f(x_n) = \boldsymbol{o}_n \end{aligned}$$

So we have to show:

$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi])) \text{ iff } \boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$$

In light of the fact that alphabetic- and η-variants of $[\lambda y_1 \ldots y_n\, \varphi]$ receive the same denotation, there are only two cases to consider.

Case 1. $[\lambda y_1 \ldots y_n\, \varphi]$ is elementary, in which case it has the form $[\lambda y_1 \ldots y_n\, \Pi^n y_1 \ldots y_n]$. So we have to show:

$$(\vartheta)\quad \langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \Pi^n y_1 \ldots y_n])) \text{ iff } \boldsymbol{w}_0 \models \Pi^n x_1 \ldots x_n$$

But the left-hand side of (ϑ) reduces, by D4, to

$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}(\Pi))$$

And the right-hand side of (ϑ), by T1, expands to:

$$\langle \boldsymbol{d}_{\mathcal{I},f}(x_1),\ldots,\boldsymbol{d}_{\mathcal{I},f}(x_n)\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}(\Pi^n)),$$

i.e.,

$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}(\Pi^n)).$$

So the two sides of the biconditional we had to show are identical conditions.

Case 2. $[\lambda y_1 \ldots y_n\, \varphi]$ is non-elementary and η-irreducible. Again, we have to show:

$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi])) \text{ iff } \boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$$

By D5 and Constraint (1.1) on Interpretations, we know:

$$\begin{aligned} &\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi]) = \\ &\quad \bar{\epsilon}\boldsymbol{r}^n \forall \boldsymbol{w} \forall \boldsymbol{o}_1{}' \ldots \forall \boldsymbol{o}_n{}'(\langle \boldsymbol{o}_1{}',\ldots,\boldsymbol{o}_n{}'\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{r}^n) \;\equiv\; \boldsymbol{w} \models_{\mathcal{I},f[y_i/\boldsymbol{o}_i{}']} \varphi) \end{aligned}$$

Now the $\bar{\epsilon}$-Conversion principle (Section 4.3) that governs the logic of $\bar{\epsilon}$-terms is: if $\boldsymbol{s} = \bar{\epsilon}\boldsymbol{r}(\ldots \boldsymbol{r} \ldots)$ then $(\ldots \boldsymbol{s} \ldots)$. So we can substitute $\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n \varphi])$ for $\boldsymbol{r}^n$ to infer:

$$\forall \boldsymbol{w} \forall \boldsymbol{o}_1{}' \ldots \forall \boldsymbol{o}_n{}'(\langle \boldsymbol{o}_1{}',\ldots,\boldsymbol{o}_n{}'\rangle \in \mathbf{ex}_{\boldsymbol{w}}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi])) \;\equiv\; \boldsymbol{w} \models_{\mathcal{I},f[y_i/\boldsymbol{o}_i{}']} \varphi)$$

If we instantiate $\forall \boldsymbol{w}$ to $\boldsymbol{w}_0$, and instantiate $\forall \boldsymbol{o}_1{}' \ldots \forall \boldsymbol{o}_n{}'$ to $\boldsymbol{o}_1,\ldots,\boldsymbol{o}_n$, respectively, we derive the following fact:

Fact 1:
$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi])) \;\equiv\; \boldsymbol{w}_0 \models_{\mathcal{I},f[y_i/\boldsymbol{o}_i]} \varphi$$

Now recall that we have to show:

$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi])) \text{ iff } \boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$$

$(\rightarrow)$ For the left-right direction, assume:

$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi])).$$

By Fact 1, it follows that:

$$\boldsymbol{w}_0 \models_{\mathcal{I},f[y_i/\boldsymbol{o}_i]} \varphi$$

But notice that it is a condition on β-Conversion that the x_i are substitutable, respectively, for the y_i in φ. Moreover, by hypothesis, $\boldsymbol{d}_{\mathcal{I},f}(x_i) = \boldsymbol{o}_i$. So by the Generalized Substitution Lemma (Section 6.6):[37]

$$\boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n} \text{ if and only if } \boldsymbol{w}_0 \models_{\mathcal{I},f[y_i/\boldsymbol{o}_i]} \varphi$$

So:

$$\boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$$

$(\leftarrow)$ Now for the right-left direction, assume:

$$\boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$$

Now by hypothesis we know $f(x_1) = \boldsymbol{o}_1$ & … & $f(x_n) = \boldsymbol{o}_n$, i.e.,

[37] Here is where Constraints (2.1), (2.2) (3.1), and (3.2) are needed, since they are used to prove the Substitution Lemma, and hence, the Generalized Substitution Lemma.

$$f = f[x_i/\boldsymbol{o}_i]$$

So we can use the same reasoning that led us to Fact 1 to obtain Fact 2, which governs the expression $[\lambda x_1 \ldots x_n\, \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}]$:

Fact 2:
$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda x_1 \ldots x_n\, \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}])) \equiv \boldsymbol{w}_0 \models_{\mathcal{I},f} \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$$

Given our assumption, we may conclude:

$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda x_1 \ldots x_n\, \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}]))$$

But we also know that:

$[\lambda x_1 \ldots x_n\, \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}]$ and $[\lambda y_1 \ldots y_n\, \varphi]$ are alphabetic variants.

So, by D5, it follows that:

$$\boldsymbol{d}_{\mathcal{I},f}([\lambda x_1 \ldots x_n\, \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}]) = \boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi])$$

Hence,

$$\langle \boldsymbol{o}_1,\ldots,\boldsymbol{o}_n\rangle \in \mathbf{ex}_{\boldsymbol{w}_0}(\boldsymbol{d}_{\mathcal{I},f}([\lambda y_1 \ldots y_n\, \varphi])),$$

which is what we had to show. ⋈

Bibliography

Ackerman, W., 1924, "Begründung des 'tertium non datur' mittels der Hilbertschen Theorie der Widerspruchsfreiheit," *Mathematische Annalen*, 93: 1–36.

Alama, J., 2013, "The Lambda Calculus," in the *Stanford Encyclopedia of Philosophy* (Summer 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2013/entries/lambda-calculus/>.

Avigad, J., and R. Zach, 2013, "The Epsilon Calculus," in the *Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2013/entries/lambda-calculus/>.

Bealer, G., 1979, "Theories of Properties, Relations, and Propositions," *The Journal of Philosophy*, 76(11): 634–648.

———, 1982, *Quality and Concept*, Oxford: Clarendon.

Church, A., 1932, "A Set of Postulates for the Foundation of Logic," *The Annals of Mathematics* (2nd Series), 33(2): 346–366.

———, 1940, "A Formulation of the Simple Theory of Types," *The Journal of Symbolic Logic*, 5(2): 56–68.

———, 1941, *The Calculi of Lambda-Conversion*, Princeton: Princeton University Press.

Curry, H., 1963, *Foundations of Mathematical Logic*, New York: McGraw-Hill; page reference is to the reprint, New York: Dover, 1977.

Curry, H., and R. Feys, 1958, *Combinatory Logic* (Volume 1), Amsterdam: North-Holland.

Došen, K., 1988, "Second-order Logic Without Variables," in W. Buszkowski, W. Marciszewski, and J. van Benthem (eds.), *Categorial Grammar*, Amsterdam: John Benjamins.

Ewald, W. B. (ed.), 1996, *From Kant to Hilbert. A Source Book in the Foundations of Mathematics* (Volume 2), Oxford: Oxford University Press.

Hazen, A., 2012, "Reflections on Counterpart Theory," in K. Bennett and D. Zimmerman (eds.), *Oxford Studies in Metaphysics* (Volume 7), Oxford: Oxford University Press, 41–64.

Hilbert, D., 1922, "Neubegründung der Mathematik: Erste Mitteilung," *Abhandlungen aus dem Seminar der Hamburgischen Universität*, 1: 157–177; English translations in Mancosu 1998, 198–214 and Ewald 1996, 1115–1134.

Hindley, J., and J. Seldin, 1986, *Introduction to Combinators and λ-Calculus*, Cambridge: Cambridge University Press.

Mancosu, P., 1998, *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920s*, Oxford: Oxford University Press.

McMichael, A., and E. Zalta, 1980, "An Alternative Theory of Nonexistent Objects," *Journal of Philosophical Logic*, 9: 297–313.

Menzel, C., 1986, "A Complete, Type-Free 'Second-order' Logic and Its Philosophical Foundations," Technical Report #csli-86-40, Stanford: Center for the Study of Language and Information Publications.

Montague, R., 1974, "The Proper Treatment of Quantification in Ordinary English," in R. Montague, *Formal Philosophy*, R. Thomason (ed.), New Haven: Yale University Press.

Oppenheimer, P., and E. Zalta, 2011, "Relations Versus Functions at the Foundations of Logic: Type-Theoretic Considerations," *Journal of Logic and Computation*, 21: 351–374.

Quine, W.V.O., 1956, "Quantifiers and Propositional Attitudes," *Journal of Philosophy*, 53 (5): 177–187; reprinted in *The Ways of Paradox and Other Essays*, revised and enlarged edition, Cambridge, MA: Harvard, 1976, 185–196.

———, 1960, "Variables Explained Away," in *Selected Logical Papers*, New York: Random House, 227–235.

Zalta, E., 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: D. Reidel.

———, 1988a, *Intensional Logic and the Metaphysics of Intentionality*, Cambridge, MA: MIT/Bradford.

———, 1988b, "Logical and Analytic Truths That Are Not Necessary," *The Journal of Philosophy*, 85(2): 57–74.

———, 1993, "A Philosophical Conception of Propositional Modal Logic," *Philosophical Topics*, 21(2): 263-281.

———, 1999, "Natural Numbers and Natural Cardinals as Abstract Objects: A Partial Reconstruction of Frege's *Grundgesetze* in Object Theory," *Journal of Philosophical Logic*, 28(6): 619–660.