Towards a Proof-Theoretic Semantics of Programming

Noam Zeilberger

Université Paris 7 PPS / $\pi\, r^2$

21 October 2009

Part I

Some History

Proof-Theoretic Justification

Hilbert's program: develop a proof-theoretic justification of the laws of mathematics

- Two answers by two G's...

Two Ideas by Gentzen[1]

1. **Natural deduction**: a more "natural" system of reasoning
2. **Sequent calculus**: encodes ND and implies consistency, etc.

Yet G did not view SC as "semantics" for ND, quite the contrary!...

[1] Untersuchungen über das Logische Schließen

Ur-quote of Proof-Theoretic Semantics

The introductions are so to say the 'definitions' of the symbols concerned, and the eliminations are ultimately only consequences thereof, what can approximately be expressed as follows: In eliminating a symbol, the formula concerned—of which the outermost symbol is in question—may only 'be used as that what it means on the ground of the introduction of that symbol'.†

Die Einführungen stellen sozusagen die „Definitionen" der betreffenden Zeichen dar, und die Beseitigungen sind letzten Endes nur Konsequenzen hiervon, was sich etwa so ausdrücken läßt: Bei der Beseitigung eines Zeichens darf die betreffende Formel, um deren äußerstes Zeichen es sich hier handelt, nur „als das benutzt werden, was sie aufgrund der Einführung dieses Zeichens bedeutet". [Untersuchungen, §5.13]

†Translation by Wagner de Campos Sanz and Thomas Piecha

Two Ideas by Prawitz[2]

1. Direct normalization of ND, without trip through SC
2. Inversion Principle: assume "canonical" proofs to justify elims

For Prawitz: canonical proof = ends in intro

Inversion Principle elaborates ur-quote (and Lorenzen '55)

[2] ND: A Proof Theoretical Study; various papers in the 1970s

Prawitz's Inversion Principle

$$\frac{A \quad B}{A \wedge B}$$

from above intro rule, one can justify the two elim rules,

$$\frac{A \wedge B}{A} \qquad \frac{A \wedge B}{B}$$

by the following reductions:

$$\dfrac{\dfrac{\begin{array}{cc}\vdots & \vdots \\ A & B\end{array}}{A \wedge B}}{A} \rightsquigarrow \begin{array}{c}\vdots \\ A\end{array} \qquad \dfrac{\dfrac{\begin{array}{cc}\vdots & \vdots \\ A & B\end{array}}{A \wedge B}}{B} \rightsquigarrow \begin{array}{c}\vdots \\ B\end{array}$$

Two Ideas by Dummett[3]

1. More restricted canonical pfs justify stronger inv principles
2. Can also consider canonical consequences...

For Dummett: canonical = sequence of intros/elims

Contrasts verificationist and pragmatist meaning-theories:

- V: props defined by intro (i.e., by how you verify them)
- P: props defined by elim (i.e., by how you use them)

[3] 1976 Lectures (The Logical Basis of Metaphysics)

Dummett's Inversion Principles

$$\frac{A \wedge (B \vee C) \qquad \begin{matrix}[A][B] \\ D\end{matrix} \qquad \begin{matrix}[A][C] \\ D\end{matrix}}{A}$$

More restricted canonical proofs justify stronger inversion principles

Dummett's Inversion Principles

$$\dfrac{\dfrac{\vdots}{A} \quad \dfrac{\vdots}{B \vee C} \qquad \begin{matrix}[A][B] \\ D\end{matrix} \qquad \begin{matrix}[A][C] \\ D\end{matrix}}{A}$$

More restricted canonical proofs justify stronger inversion principles

Dummett's Inversion Principles

$$\dfrac{\dfrac{\vdots}{A} \quad \dfrac{\dfrac{\vdots}{B}}{B \vee C} \qquad \begin{matrix}[A][B] \\ D\end{matrix} \qquad \begin{matrix}[A][C] \\ D\end{matrix}}{A} \quad \rightsquigarrow \quad \begin{matrix}\vdots \\ D\end{matrix} \qquad \begin{matrix}\vdots \\ A\end{matrix} \quad \begin{matrix}\vdots \\ B\end{matrix}$$

More restricted canonical proofs justify stronger inversion principles

Dummett's Inversion Principles

$$\dfrac{\dfrac{\vdots}{A} \quad \dfrac{\dfrac{\vdots}{C}}{B \vee C} \qquad \begin{matrix}[A][B] \\ D\end{matrix} \qquad \begin{matrix}[A][C] \\ D\end{matrix}}{A \wedge (B \vee C)} \qquad A$$

More restricted canonical proofs justify stronger inversion principles

Dummett's Inversion Principles

$$\dfrac{\vdots\ A \quad \dfrac{\vdots\ C}{B \vee C}}{\dfrac{A \wedge (B \vee C) \quad [A][B]\ D \quad [A][C]\ D}{A}} \rightsquigarrow \begin{array}{c} \vdots\ A \quad \vdots\ C \\ \vdots\ D \end{array}$$

More restricted canonical proofs justify stronger inversion principles

Dummett's Inversion Principles (pragmatist)

$$\frac{A \wedge B}{A} \qquad \frac{A \wedge B}{B}$$

from above elim rules, one can justify the intro rule,

$$\frac{A \quad B}{A \wedge B}$$

by the following reductions:

$$\frac{\dfrac{A \quad B}{A \wedge B}}{A} \quad \vdots \quad \rightsquigarrow \quad A \quad \vdots \qquad \frac{\dfrac{A \quad B}{A \wedge B}}{B} \quad \vdots \quad \rightsquigarrow \quad B \quad \vdots$$

Harmony

How to relate V and P? D gives a perceptive but mixed analysis. . .

- Notes some connectives difficult to interpret as both V and P
- Nonetheless, insists on this harmony for V and P to be able to understand eachother

1.5 Ideas by (Andreoli + Girard)

1. Proof search alternating between inversion and focalization
2. Two classes of connectives (A: (a)synchronous, G: $\pm$ polarity)
3. Can explicitly polarize the connectives of IL and CL

Reveals new light on Gentzen/Prawitz/Dummett-style PTS:

- Inversion Principle $\leftrightarrow$ Inversion Phase
- **Strong** (D's) canonical proofs/consequences $\leftrightarrow$ Focalization
- V/P distinction $\leftrightarrow$ +/- polarity

And suggests new approach to harmony: simply accept diversity!

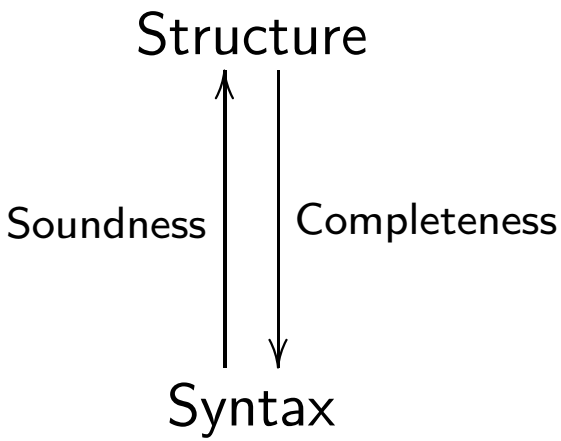
Part II

PTS: The Big Picture

Question

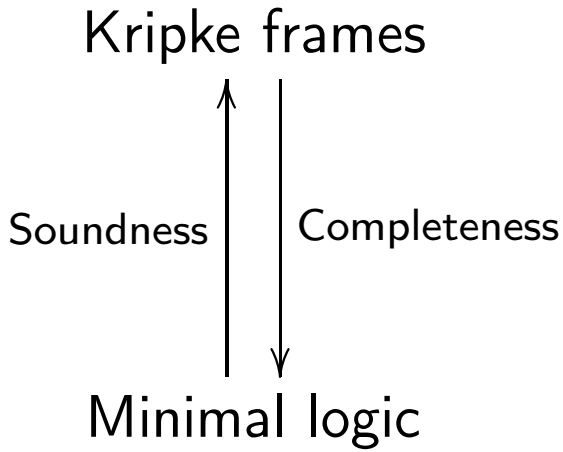
In what sense is PTS really semantics?

The (Usual) Picture of Denotational Semantics



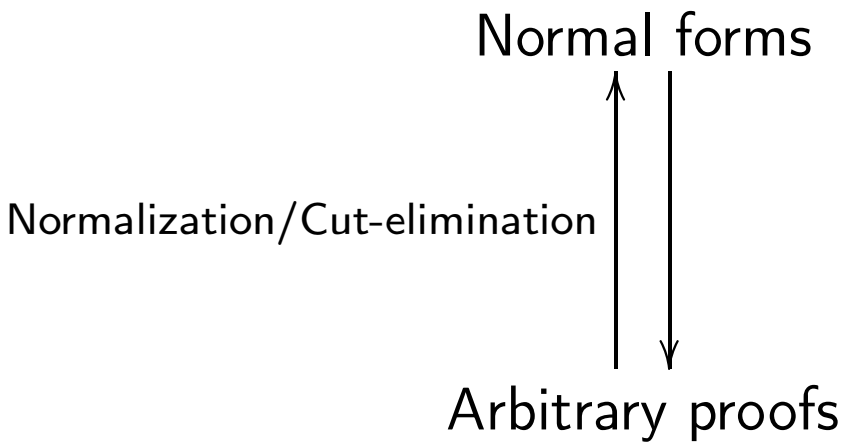
(i.e., compositional interpretation of syntax into something "nice")

(e.g.)



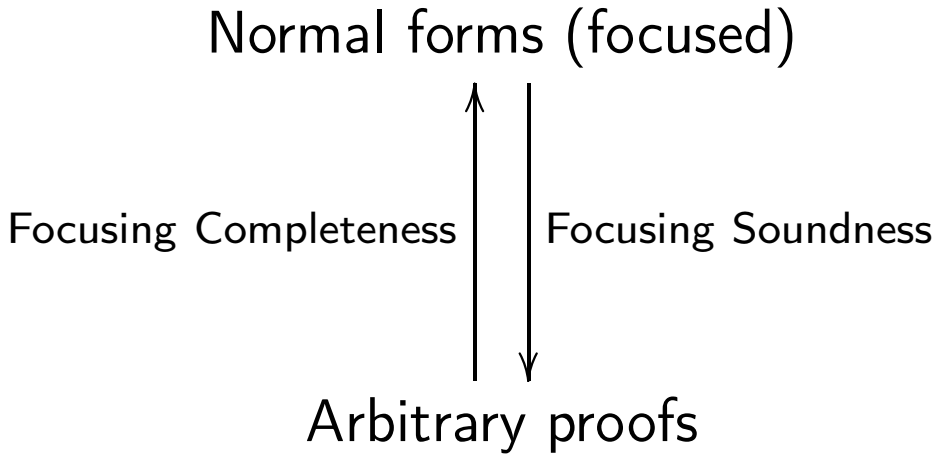
(Completeness $\circ$ Soundness = NBE for simply-typed λ-calculus)

A Blurry Picture of PTS



Problem: "interpretation" not defined compositionally.

Focusing the Picture



Key point: can compose normal forms directly

Definition (Polarization)

Let $|-|$ be the operation that "forgets" polarity, e.g.:

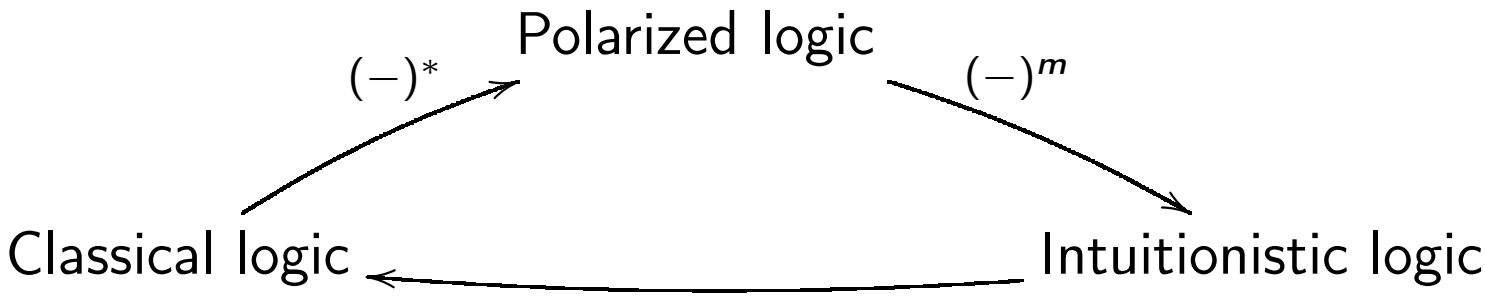
$$|A \otimes B| = |A| \wedge |B| = |A \& B| \qquad |A \oplus B| = |A| \vee |B| = |A \parr B|$$
$$|\stackrel{+}{\neg} A| = \sim |A| = |\stackrel{-}{\neg} A| \qquad |{\downarrow}A| = |A| = |{\uparrow}A|$$

A **polarization** is a mapping $(-)^*$ from ordinary formulas to polarized formulas, such that b and $|b^*|$ are classically equivalent.

Theorem (Focusing Completeness)

If $\Gamma \vdash \Delta$ *has a classical proof then* $\Gamma^* \vdash \Delta^*$ *has a focusing proof.*

Polarization as (decomposing) $\neg\neg$ translation



- $(-)^*$ is <u>any</u> polarization
- $(-)^m$ faithful, into fragment w/$\supset$ restricted to minimal $\neg$
- Different $(-)^*$ yield different $\neg\neg$ translations $(-)^m \circ (-)^*$

Towards a PTS of Programming

Type-theoretic interpretation of the previous slides:

- Focusing proofs are β-normal, η-long, CPS terms
- Soundness $\circ$ Completeness = Normalization-by-Focusing
- Different polarizations $\Rightarrow$ different CPS semantics

But there's more...

- Deep connection between dual inversion principles, pattern-matching, and generic/untyped computation
- More realistic opsem: cut-elimination as environment machine
- Theory of extrinsic typing ($\cap$, $\cup$, subtyping, etc.)

Part III

From Patterns to Proofs

Judgments

Convention

Every proposition has a definite **polarity**, *positive or negative. We indicate the polarity of a proposition A with a superscript A^+ or A^-, unless it is unimportant or already clear from context.*

Convention

We write "A true" for the **affirmation** *of A, abbreviated "A", and "A false" for the* **denial** *of A, abbreviated "$\bullet A$".*

Frames

Definition (Frames)

A **frame** Δ is a tree of affirmations and denials. We write $\cdot$ for the empty frame, and Δ_1, Δ_2 for the join of two frames. We call $B \in \Delta$ and $\bullet B \in \Delta$ the **holes** of Δ. A frame without any non-atomic holes B^+ or $\bullet B^-$ is called **simple**.

Dictionaries

Definition (Dictionaries, definition ordering)

A **dictionary** Σ is an inductively defined relation $\Delta \Vdash \Delta'$, where Δ is a simple frame and Δ' is either A^+ or $\bullet A^-$.

Given a dictionary, we describe the **definition ordering** $\prec$ as the least transitive relation satisfying the following:

- If $\Delta \Vdash A^+$ then $\Delta \prec A^+$, and if $\Delta \Vdash \bullet A^-$ then $\Delta \prec A^-$
- If $\bullet B^+ \in \Delta$ then $B^+ \prec \Delta$, and if $B^- \in \Delta$ then $B^- \prec \Delta$

For arbitrary propositions A, we write $\prec_A$ for the restriction of $\prec$ below A, and similarly $\prec_\Delta$ for the restriction of $\prec$ below Δ.

(Example 1)

$$\frac{}{\cdot \Vdash 1} \qquad \frac{\Delta_1 \Vdash A^+ \quad \Delta_2 \Vdash B^+}{\Delta_1, \Delta_2 \Vdash A^+ \otimes B^+} \qquad \frac{\Delta \Vdash A^+}{\Delta \Vdash A^+ \oplus B^+} \qquad \frac{\Delta \Vdash B^+}{\Delta \Vdash A^+ \oplus B^+}$$

$$\frac{\Delta \Vdash \bullet A^-}{\Delta \Vdash \bullet A^- \,\&\, B^-} \qquad \frac{\Delta \Vdash \bullet B^-}{\Delta \Vdash \bullet A^- \,\&\, B^-} \qquad \frac{}{\cdot \Vdash \bullet \bot} \qquad \frac{\Delta_1 \Vdash \bullet A^- \quad \Delta_2 \Vdash \bullet B^-}{\Delta_1, \Delta_2 \Vdash \bullet A^- \parr B^-}$$

$$\frac{\Delta_1 \Vdash A^+ \quad \Delta_2 \Vdash \bullet B^-}{\Delta_1, \Delta_2 \Vdash \bullet A^+ \to B^-} \qquad \frac{}{A^- \Vdash \downarrow A^-} \qquad \frac{}{\bullet A^+ \Vdash \bullet \uparrow A^+}$$

(Example 2)

$$\frac{}{\cdot \Vdash \mathbb{N}} \qquad \frac{\Delta \Vdash \mathbb{N}}{\Delta \Vdash \mathbb{N}} \qquad \frac{}{\bullet P \Vdash P}$$

Patterns

Definition (Patterns)

A derivation of $\Delta \Vdash A^+$ is called a **proof pattern**. A derivation of $\Delta \Vdash \bullet A^-$ is called a **refutation pattern**.

Principle (Abstraction)

A proposition is literally defined by its set of patterns in Σ, i.e., we can only examine it through the dictionary.

Contexts

Definition (Context)

A (simple) **context** Γ is a list of (simple) frames, and a hole $A \in \Gamma$ or $\bullet A \in \Gamma$ is called a **hypothesis**.

Note: we make no assumption of linear use of hypotheses.

Proofs and Refutations

Given patterns for proofs and refutations, how do we define actual proofs and refutations? We have basically no choice...

Positive Proof

$$\Gamma \vdash A^+$$

Positive Proof

$$\frac{\Delta \Vdash A^+ \quad \Gamma \vdash \Delta}{\Gamma \vdash A^+}$$

Positive Proof

$$\frac{\Delta \Vdash A^+ \quad \Gamma \vdash \Delta}{\Gamma \vdash A^+}$$

i.e., pick a proof pattern and fill in its holes

Positive Refutation

$$\Gamma \vdash \bullet A^{+}$$

Positive Refutation

$$\frac{\Delta \Vdash A^+ \quad \longrightarrow \quad \Gamma, \Delta \vdash \#}{\Gamma \vdash \bullet A^+}$$

Positive Refutation

$$\frac{\Delta \Vdash A^+ \quad \longrightarrow \quad \Gamma, \Delta \vdash \#}{\Gamma \vdash \bullet A^+}$$

i.e., derive contradiction from each proof pattern[4]

[4] Note: this is (very close to) Buchholz's Ω-rule.

Negative Proof

$$\Gamma \vdash A^-$$

Negative Proof

$$\frac{\Delta \Vdash \bullet A^- \quad \longrightarrow \quad \Gamma, \Delta \vdash \#}{\Gamma \vdash A^-}$$

Negative Proof

$$\frac{\Delta \Vdash \bullet A^- \quad \longrightarrow \quad \Gamma, \Delta \vdash \#}{\Gamma \vdash A^-}$$

i.e., derive contradiction from each refutation pattern

Negative Refutation

$$\Gamma \vdash \bullet A^-$$

Negative Refutation

$$\frac{\Delta \Vdash \bullet A^- \quad \Gamma \vdash \Delta}{\Gamma \vdash \bullet A^-}$$

Negative Refutation

$$\frac{\Delta \Vdash \bullet A^- \quad \Gamma \vdash \Delta}{\Gamma \vdash \bullet A^-}$$

i.e., pick a refutation pattern and fill in its holes

Remaining rules

$$\frac{}{\Gamma \vdash \cdot} \qquad \frac{\Gamma \vdash \Delta_1 \quad \Gamma \vdash \Delta_2}{\Gamma \vdash \Delta_1, \Delta_2}$$

$$\frac{\bullet A^+ \in \Gamma \quad \Gamma \vdash A^+}{\Gamma \vdash \#} \qquad \frac{A^- \in \Gamma \quad \Gamma \vdash \bullet A^-}{\Gamma \vdash \#}$$

Canonical Derivations

Definition

A derivation of $\Gamma \vdash J$ using only the foregoing rules is called a **canonical derivation**.

Note: these are really just focusing proofs, in funny notation.

Identity and Composition

Principle (Identity)

The following are admissible for canonical derivations:

1. *If* $A^- \in \Gamma$ *then* $\Gamma \vdash A^-$
2. *If* $\bullet A^+ \in \Gamma$ *then* $\Gamma \vdash \bullet A^+$
3. $\Gamma(\Delta) \vdash \Delta$

Principle (Composition)

The following are admissible for canonical derivations:

1. *If* $\Gamma \vdash A$ *and* $\Gamma \vdash \bullet A$ *then* $\Gamma \vdash \#$
2. *If* $\Gamma \vdash \Delta$ *and* $\Gamma(\Delta) \vdash J$ *then* $\Gamma \vdash J$

Identity and Composition

Surprising fact: proof of identity and composition is "immediate"!
. . . conditioned on well-foundedness of definition ordering $\prec_A, \prec_\Delta$

Part IV

Extracting a Language

Schema

Equate terms with derivations (a.k.a., "intrinsic" language def)

(So in some sense we're already done. . .)

. . . But we can develop a type-free notation for typeful derivations

- Relies on subformula property for lack of ambiguity
- Composition (computation) revisited in this notation

Slogans

$V^+ = p[\sigma]$

A (pos) value is just a (val) pattern under a substitution

$K^+ = p \mapsto E$

A (pos) continuation is just a map from (val) pats to exps

$V^- = d \mapsto E$

A (neg) value is just a map from (con) pats ("observations") to exps

$K^- = d[\sigma]$

A (neg) value is just a (con) pattern under a substitution

$$\begin{array}{lll} \sigma & ::= & \cdot \mid (\sigma_1, \sigma_2) \mid K^+ \mid V^- \\ E & ::= & \kappa\ V^+ \mid x\ K^- \end{array}$$

Equality

We suffice with "syntactic" equality. . . but for higher-order syntax!

Equality

$$\frac{\sigma_1 = \sigma_2}{p[\sigma_1] = p[\sigma_2]} \qquad \frac{\forall d \;.\; V_1^-(d) = V_2^-(d)}{V_1^- = V_2^-}$$

$$\frac{\forall p \;.\; K_1^+(p) = K_2^+(p)}{K_1^+ = K_2^+} \qquad \frac{\sigma_1 = \sigma_2}{d[\sigma_1] = d[\sigma_2]}$$

$$\frac{}{\cdot = \cdot} \qquad \frac{\sigma_1 = \sigma_1' \quad \sigma_2 = \sigma_2'}{(\sigma_1, \sigma_2) = (\sigma_1', \sigma_2')}$$

$$\frac{V_1^+ = V_2^+}{\kappa\ V_1^+ = \kappa\ V_2^+} \qquad \frac{K_1^- = K_2^-}{x\ K_1^- = x\ K_2^-}$$

Composition and Identity

Composition and identity principles become operations on terms.

Theorem

Composition is associative, identity is a left/right unit.

Evaluation as cut-elimination

Composition (∼cut-admissibility) does not automatically yield a realistic model of computation (recall "NBF", analogous to NBE)

Instead, investigate iterated composition (∼cut-elim):

$$\frac{\cdot \vdash \Delta_1 \quad \Delta_1 \vdash \Delta_2 \quad \ldots \quad \Delta_1, \ldots, \Delta_{n-1} \vdash \Delta_n \quad \Delta_1, \ldots, \Delta_n \vdash \#}{\cdot \vdash \#}$$

Definition (Environments and Programs)

Let $\Gamma = \Delta_1, \ldots, \Delta_n$. An **environment** $\gamma :: \Gamma$ is a list of substitutions $\gamma = (\sigma_1; \ldots; \sigma_n)$, where $\sigma_i :: (\Delta_1, \ldots, \Delta_{i-1} \vdash \Delta_i)$. A **program** is either a pair of an environment $\gamma :: \Gamma$ and expression $E :: (\Gamma \vdash \#)$, written $\langle \gamma \mid E \rangle$, or a triple of γ and a continuation $K :: (\Gamma \vdash \bullet A)$ and a value $V :: (\Gamma \vdash A)$, written $\langle \gamma \mid K \mid V \rangle$.

Operational semantics

$$\langle \gamma \mid \kappa\ V \rangle \quad \rightsquigarrow \quad \langle \gamma \mid \gamma(\kappa) \mid V \rangle$$
$$\langle \gamma \mid K \mid p[\sigma] \rangle \quad \rightsquigarrow \quad \langle (\gamma; \sigma) \mid K(p) \rangle$$

(plus symmetric negative rules)

Observational equivalence

Definition (Observational equivalence)

Let $E_1, E_2 :: (\Gamma \vdash \#)$ be two expressions. We say that $E_1 \cong E_2$ if for all $\gamma :: \Gamma$ and all results R, we have $\langle \gamma \mid E_1 \rangle \Downarrow R$ iff $\langle \gamma \mid E_2 \rangle \Downarrow R$.

Theorem (Congruence)

If $E_1 = E_2$ then $E_1 \cong E_2$.

Theorem (Separation)

In the presence of two distinct observable results (Ω and $\mho$) and integer state, for any pure E_1, E_2, if $E_1 \neq E_2$ then $E_1 \ncong E_2$.

Part V

Extrinsic types

Refinement types

On top of this intrinsic language definition, we can build an extrinsic **refinement type system** to capture more precise semantic properties (e.g., rule out $\lightning$).

Refinement relation between extrinisic and intrinsic: $S \sqsubseteq A$

Given a value with intrinsic type A (or continuation accepting A), we can hope to show it also has (accepts) extrinsic type S, e.g.

Note: typical Curry-style typing is the special case where there is only one (or rather two) intrinsic type(s).

Refinement types

Extrinsic type system on same principles as intrinsic (details in phd).

Identity Coercion Interpretation

Definition (Subtyping)

Let $S, T \sqsubseteq A$. We say that $S \leq_A T$ (or simply $S \leq T$) if $x : S \vdash Id_x : T$.

Many advantages of this definition over direct axiomatization!

. . . but it doesn't explain why certain principles, such as $\neg\neg S \cap \neg\neg T \leq \neg\neg(S \cap T)$, are unsound i.t.p.o. effects

No-Counterexamples Interpretation

Definition (Safety)

Let $\gamma :: \Gamma$, $V :: (\Gamma \vdash A)$, and $K :: (\Gamma \vdash \bullet A)$. We write $\gamma \models V \perp K$ if $\langle \gamma \mid V \mid K \rangle \not\Downarrow \mho$. Similarly, for $\sigma :: (\Gamma \vdash \Delta)$ and $E :: (\Gamma, \Delta \vdash \#)$, we write $\gamma \models \sigma \perp E$ if $\langle (\gamma; \sigma) \mid E \rangle \not\Downarrow \mho$. We write $V \perp K$ and $\sigma \perp E$ if these hold in all environments.

Definition (NCE Subtyping)

Let $S, T \sqsubseteq A$. We say that $S \leqslant_A T$ if for all closed values $V :: A$ and continuations $K :: \bullet A$, if $V : S$ and $K : \bullet T$ then $V \perp K$.

Soundness

Theorem (Soundness)

If $S \leq T$ then $S \leqslant T$.

Proof.

Almost immediate by type safety. □

Conditional Completeness

Definition (Choice operators)

A frame Δ is said to have (binary) **demonic choice** if for every pair of expressions $E_1, E_2 :: (\Delta \vdash \#)$, there is an expression $E_1 \curlywedge E_2 :: (\Delta \vdash \#)$, such that $(E_1 \curlywedge E_2)^\perp = E_1^\perp \cap E_2^\perp$ [and that admits an appropriate refinement typing rule]. It is has **angelic choice** if there is an expression $E_1 \curlyvee E_2$ such that $(E_1 \curlyvee E_2)^\perp = E_1^\perp \cup E_2^\perp$ [with an appropriate typing rule].

Theorem (Conditional Completeness)

If A has "ancestral" demonic and angelic choice, then $S \not\leq_A T$ implies $S \not\leq_A T$.

Where are we?

Focusing and polarity as a fresh take on PTS:

- Enables more interesting dialogue between "verificationist" and "pragmatist"
- Very strong, compositional notion of canonical form

and polarized type theory as a fresh take on Curry-Howard:

- Evaluation order encoded in types
- Pattern-matching as a primitive, which enables a uniform account of untyped computation, intrinsic types, and extrinsic refinement types.

Where can we go?

Some important unresolved questions:

- Type theory: How to accomodate dependent/higher-order quantifiers? Note: dependent quantifiers needed for "meta-circular interpreter"!
- Programming languages: If polarized logic is the logic of CPS, what is the logic of delimited CPS? An answer seems essential for encapsulating effects.
- Logic: What is the precise relationship between proofs and countermodels? (The ludics program.)
- Topology: How can we explain away the "infinities" of canonical derivations as *façons de parler*? More precisely, can we compile them away? Note: this may already be answered by infinitary PT.