Formal Specification of the Plutus Language

Darryl McAdams

Email: darryl.mcadams@iohk.io Slack: @darryl.mcadams

I. Introduction

In this document, we outline the Plutus Language's overall design and give its big-step semantics. We also discuss high-level design considerations of the language, including the distinction between the Core language and the user-facing language. The latter language is accompanied by a proof theoretic definition of the elaboration and bidirectional typing process.

II. PLUTUS

The Plutus language is a variant of the λ calculus with functions and user-defined recursive data types, as well as a variety of build-in functions for cryptographic calculations. The language is split into two main parts, similar to Haskell: Plutus proper, which is the user-facing language that has convenient ways of declaring data and values, and a core language (Plutus Core, or just or Core for short), which is the actual language that defines the programs to run after Plutus is elaborated into it. The reason for this distinction will be discussed in detail later.

A. Plutus Core

Plutus Core is an untyped λ calculus. The grammar of the language is given in Figure 1, using the conventions of Harper [1]. As is standard in λ calculi, we have variables, λ abstractions, and application. In addition to this, there are also local let bindings, data constructors which have arguments, case terms which match on lists of terms, declared names, success and failure terms, and built-in functions. Terms live within a top-level program which consists of some declarations.

We do not bother to give an example program in Plutus Core here, as the abstract binding tree notation is not reader-friendly, nor is it intended to be read by humans.

The execution of a program in Plutus Core does not in itself result in any evaluation. Instead, the declarations are bound to their appropriate names. Then, designated names can be chosen to be evaluated in the declaration environment generated by the declarations. For instance, we might designate the name main to be the name who's definition we evaluated, as is done in Haskell. The generation of this declaration environment is given by the relation $D^* \gg \delta$, defined in Figure 2.

To give the computation rules for Plutus Core, we must define what the values are of the language, as given in Figure 3. These then let us define a parameterized binary relation $M \downarrow_{\delta} M'$ which means M evaluates to M' using declarations δ , in Figure 4.

Tm	m	::=	x	variable
			decname[n]	declared name
			$\mathtt{let}(m;x.m)$	local declaration
			$\lambda(x.m)$	λ abstraction
			app(m;m)	function application
			$\mathtt{con}[\mathrm{n}](m^*)$	constructed data
			$\mathtt{case}(m^*;c^*)$	case
			$\mathtt{success}(m)$	success
			failure	failure
			$\mathtt{bind}(m;x.m)$	computation bind
			$\mathtt{builtin}[\mathrm{n}](m^*)$	built-in function
Cl	c	::=	$\mathtt{cl}(p^*;m)$	case clause
Pat	p	::=	x	variable pattern
			$\mathtt{con}[\mathrm{n}](p^*)$	constructor pattern
Prg	g	::=	$\mathtt{program}(d^*)$	program
Dec	d	::=	$\mathtt{dec}(n;m)$	name declaration

Fig. 1. Grammar of Plutus Core

$$D \gg \delta$$

Declarations D generate environment δ

$$\begin{array}{c|c} \hline \epsilon \gg \epsilon \\ \hline D \gg \delta & M \Downarrow_{\delta} M' \\ \hline D, \deg(n;M) \gg \delta, n \mapsto M' \end{array}$$

Fig. 2. Declaration Environment Generation

Evaluation of case terms is defined in terms of pattern matching, given in Figure 5. Note that in the definition of evaluation for case terms, we assume that $\operatorname{cl}(\vec{P};N)$ is the first clause of the case term for which \vec{M}' matches the pattern.

Note also that the success and failure terms are not effectful. That is to say, failure does not throw an exception of any sort. They are merely primitive values that represent computational success and failure. They are analogous to Haskell Maybe values, except that they cannot be inspected, and all computational control is done via the *bind* construct.

B. Plutus

Plutus proper, the user-facing language, is a strictly-typed, parametrically polymorphic language similar to Haskell and ML. It provides a convenient, user-friendly way to program, that can be elaborated down into Plutus Core. Plutus proper does not actually factor into any computation; it can be seen

Fig. 3. Values of Plutus Core

$$M \Downarrow_{\delta} M'$$

M evaluates to M' in environment δ

Fig. 4. Evaluation of Plutus Core

entirely as a tool for writing Plutus Core programs. We will discuss this topic more in the next section.

Plutus proper is defined by the grammar given in Figure 6. This includes definitions for the types of Plutus, as well as terms and declarations (of both types and terms). The grammar is intentionally Haskell-like. Blocks that should be parsed via indentation are specified with braces. Additionally, we assume that let, λ and do terms, and \forall types, are sugarable as in Haskell to have iterated bindings. Finally, while the grammar permits pattern matching definitions of functions, we treat term declarations in the elaborator as if they only permit defining a name to be a single term. We do this because the former can be transformed into the latter in conventional ways, thus simplifying things greatly.

An example program in Plutus is the canonical factorial function over Peano numerals. First we must define the relevant

$$M \sim P \triangleright \sigma$$

M matches pattern P yielding substitutions σ

Fig. 5. Pattern Matching in Plutus Core

Fig. 6. Grammar of Plutus

types, and basic functions, and then we can define the factorial function. This is shown in Figure 7.

The statics for Plutus can be given in the form of a collection of elaboration judgments. Whole program elaboration uses the judgment Σ ; $\Delta \vdash P \dashv \Sigma'$; Δ' , which says that a program P elaborates by transforming the signature Σ and declarations Δ into a new signature Σ' and new declarations Δ' .

A signature consists of a number of declarations, for both type constructor arity (of the form $n:\star^k$), where k is the number of type arguments that the type constructor n takes, and term constructor arity (of the form $n:[\vec{\alpha}](\vec{A})B$, where $\vec{\alpha}$ is a collection of type variables that the constructor's arguments and return type can refer to, \vec{A} are the types of the arguments to the constructor n, and B is the type of the value that n constructors. So for example, the type constructor for natural numbers would be in a signature as Nat: \star^0 , with 0 type arguments because it's a nullary type constructor, while list

data Nat = Zero | Suc Nat plus : Nat \rightarrow Nat \rightarrow Nat plus Zero n = nplus (Suc m) n = Suc (plus m n) times : Nat \rightarrow Nat \rightarrow Nat times Zero n = Zero times (Suc m) n = plus n (times m n) fac : Nat \rightarrow Nat fac Zero = Suc Zero fac (Suc n) = times (Suc n) (fac n)

Fig. 7. Fibonacci in Plutus

$$\Sigma; \Delta \vdash P \dashv \Sigma'; \Delta'$$

Program P elaborates by transforming signature Σ and declarations Δ into new signature Σ' and new declarations Δ'

Fig. 8. Elaboration of Programs in Plutus

would be List: \star^1 , with one type argument (the parameter for its element values). Their respective constructors would be in a signature as zero: []()Nat, suc: []()Nat, nil: $[\alpha]()$ (List α), and cons: $[\alpha](\alpha, \text{List }\alpha)$ (List α).

Declaration Δ consist of entries of the form $n:A\mapsto M'$, where n is the name of a declared term, A is its type, and M' is its definition, which is a term in Plutus Core.

The overall elaboration process uses the judgments to turn a set of type and term declarations into a signature and set of declared values that define a program. These judgments are defined in Figure 8. The essence of program elaboration is that we just elaborate each declaration, in sequence, threading the signature and declarations through each sub-elaboration to accumulate the results.

Naturally, we now need to explain how type declarations and term declarations are elaborated. Type declarations will simply elaborate each alternative for the type, adding them to the signature, along with the arity of the whole type constructor. This is shown in Figure 9.

Elaboration of term declarations is simpler, requiring only that we check that the type is valid and that the definition has that type, as shown in Figure 10. This results in the extension of the declarations with the new term.

$$\Sigma \vdash \mathit{type} \ \mathrm{n} \ \vec{\alpha} \ \mathit{alts} \ \vec{C} \ \dashv \ \Sigma'$$

Type declaration for n with parameters $\vec{\alpha}$ and alternatives \vec{C} elaborates by transforming signature Σ into new signature Σ'

$$\Sigma, \mathbf{n} : \star^{k} \vdash C_{0} \text{ alt } \mathbf{n} \ \vec{\alpha} \dashv \Sigma'_{0}$$

$$\Sigma'_{0} \vdash C_{1} \text{ alt } \mathbf{n} \ \vec{\alpha} \dashv \Sigma'_{1}$$

$$\dots$$

$$\Sigma'_{j-1} \vdash C_{j} \text{ alt } \mathbf{n} \ \vec{\alpha} \dashv \Sigma'_{j}$$

$$\Sigma \vdash \text{type } \mathbf{n} \ \vec{\alpha} \text{ alts } C_{0} \mid \dots \mid C_{j} \dashv \Sigma'_{j}$$

$$\Sigma \vdash c \vec{A} \ alt \ n \ \vec{\alpha} \dashv \Sigma'$$

Constructor c with arg types \vec{A} elaborates for type constructor n with parameters $\vec{\alpha}$ by transforming signature Σ into new signature Σ'

$$\frac{\Sigma; \vec{\alpha} \ type \ \vdash \ A_i \ type}{\Sigma \ \vdash \ c \ A_0 \ \dots \ A_k \ alt \ n \ \vec{\alpha} \ \dashv \ \Sigma, c : [\vec{\alpha}](A_0, \dots, A_k)(n \ \vec{\alpha})}$$

Fig. 9. Elaboration of Type Declarations in Plutus

$$\Sigma; \Delta \; \vdash \; \mathit{term} \; \mathsf{n} \; \mathit{type} \; A \; \mathit{def} \; M \; \dashv \; \Delta'$$

Term declaration for decname[n] with type A and definition M elaborates by using signature Σ to transform declarations Δ into new declarations Δ'

$$\frac{\Sigma \vdash A \ type \qquad \Sigma \ ; \ \Delta \ ; \ \mathbf{n} : A \vdash A \ni M \bowtie M'}{\Sigma ; \Delta \vdash \textit{term} \ \mathbf{n} \ \textit{type} \ A \ \textit{def} \ M \ \dashv \ \Delta, \mathbf{n} : A \mapsto M'}$$

Fig. 10. Elaboration of Term Declarations in Plutus

As can be seen in term declaration elaboration, there are also term-level elaboration judgments. In particular, we have two typing judgments, for checking and for synthesizing types: Σ ; Δ ; $\Gamma \vdash T \ni M \rhd M'$, which says that with signature Σ , declarations Δ , and in the context Γ , the type T contains the Plutus term M which elaborates to Core term M', and Σ ; Δ ; $\Gamma \vdash M \rhd M' \in T$, which says that with signature Σ , declarations Δ and in context Γ , the term M elaborates to M', synthesizing the type T. These are given in Figure 11 and Figure 12. Since this gets very verbose, Σ , Δ , and Γ will be suppressed except when they different between premises and conclusion, or when they are relevant to the particular rule at hand.

A number of auxiliary judgments are used in the definition of elaboration. The judgment A type, which is short for $\Sigma; \Gamma \vdash A$ type, is the judgment that A is a type in signature Σ and context Γ . The judgment $\vec{P} \to M \rhd M'$ from \vec{A} to B, which is short for $\Sigma; \Delta; \Gamma \vdash \vec{P} \to M \rhd M'$ from \vec{A} to B, is the judgment that $\vec{P} \to M$ is a clause with patterns for types \vec{A} returning a body with type B which elaborates to M', in signature Σ , declarations Δ , and context Γ . This is in turn defined in terms of a judgment $\Sigma \vdash P$ pattern $A \dashv \Gamma'$, which is the judgment that P is a pattern for A in signature Σ that binds variables Γ' . Finally, the judgment $A \sqsubseteq B$ defines a subtyping judgment used for instantiating quantifiers. These auxiliary judgments are defined in Figure 13. We do not bother

$$\Sigma ; \Delta ; \Gamma \vdash A \ni M \triangleright M'$$

Type A contains term M which elaborates to M' in signature Σ , declarations Δ , and context Γ

$$\Gamma \vdash B \ \forall ppe$$

$$\Gamma \vdash B \ \ni M \ \triangleright M'$$

$$\Gamma, x : B \vdash A \ \ni N \ \triangleright N'$$

$$\Gamma \vdash A \ \ni \ \text{let} \ \{ \ x : B \ \{ \ M \ \} \ \} \ \text{in} \ N \ \triangleright \ \text{let}(M'; x.N')$$

$$\frac{\Gamma, x : A \vdash B \ \ni M \ \triangleright M'}{\Gamma \vdash A \to B \ \ni \lambda x \to M \ \triangleright \lambda(x.M')}$$

$$\Sigma \ \ni \ n \ : \ [\vec{\alpha}](A_0, \dots, A_n)B$$

$$[\sigma]B = B'$$

$$\Sigma \vdash [\sigma]A_i \ \ni M_i \ \triangleright M'_i$$

$$\overline{\Sigma \vdash B' \ \ni n \ M_0 \ \dots \ M_n \ \triangleright \text{con[n]}(M'_0; \dots; M'_n)}$$

$$\frac{A \ \ni M \ \triangleright M'}{\text{Comp} \ A \ \ni \text{success} \ M \ \triangleright \text{success}(M')}$$

$$\overline{\text{Comp} \ A \ \ni \text{success} \ M \ \triangleright \text{success}(M')}$$

$$\overline{\text{Comp} \ A \ \ni \text{failure}} \ \vdash \text{failure}$$

$$\Gamma \vdash M \ \triangleright M' \ \in \text{Comp} \ A$$

$$\Gamma, x : A \vdash \text{Comp} \ B \ \ni N \ \triangleright N'$$

$$\overline{\Gamma \vdash \text{Comp} \ B \ \ni \text{do}} \ \{ x \leftarrow M \ ; N \ \} \ \triangleright \text{bind}(M'; x.N')$$

$$\frac{\Gamma, \alpha \ type \vdash A \ \ni M \ \triangleright M'}{\Gamma \vdash \forall \alpha . \ A \ \ni M \ \triangleright M'}$$

$$\frac{M \ \triangleright M' \in A \ A \ \sqsubseteq B}{B \ \ni M \ \triangleright M'}$$

Fig. 11. Checking Judgment

to define the judgments $\Sigma \ni n : A, \Delta \ni n : A$, and $\Gamma \ni x : A$ as they are relatively obvious and uninteresting.

III. CORE VS. NON-CORE

The choice to have a core language, Plutus Core, and a non-core language, Plutus, provides a number of benefits. Plutus Core is simpler, and so formal verification of behavior is easier. Additionally, it's simplicity means that it can be more easily serialized. But the main benefit of having Core be separate from Plutus proper is that Plutus proper is not actually part of the blockchain protocol. In effect, Plutus is just a tool for editing Plutus Core. If it becomes desirable to have a language with a more powerful collection of tools, such as type classes and GADTs, or even dependent types, they can be layered on top of Core as another distinct way to use Core, without having to make any changes to the protocol's scripting language at all. The syntax and semantics of Core remains unchanged, while these additions are freely swappable, because they all elaborate down to Core.

This also makes it possible to do interesting things with embedded domain specific languages. Because Core has functions and data constructors, it is possible to define DSLs that layer on top of Core as a special mode of use. Consider, for instance,

$$\Sigma ; \Delta ; \Gamma \vdash M \triangleright M' \in A$$

Term M elaborates to M' synthesizing type A in signature Σ , declarations Δ , and context Γ

 $\frac{\Gamma \ni x : A}{\Gamma \vdash x \triangleright x \in A}$

$$\begin{array}{c} \Delta\ni \mathbf{n}:A\\ \hline \Delta\vdash \mathbf{n} \rhd \operatorname{decname}[\mathbf{n}]\in A\\ \hline \underbrace{A\;type\quad A\ni M\rhd M'}_{M\;:A\; \rhd M'\in A}\\ \hline \underbrace{M\; \rhd M'\in A\to B\quad A\ni N\rhd N'}_{M\; N\; \rhd \operatorname{app}(M';N')\in B}\\ \hline \\ M_i \rhd M_i'\in A_i\quad \vec{P_j}\to N_j \rhd N_j'\; \mathit{from}\; A_0,\ldots,A_m\; \mathit{to}\; B\\ \hline \\ \operatorname{case}\; M_0\mid\ldots\mid M_m\; \operatorname{of}\; \{\; \vec{P_0}\to N_0;\ldots;\vec{P_n}\to N_n\;\} \rhd\\ \\ \operatorname{case}\; (M_0',\ldots,M_m';\operatorname{cl}(\vec{P_0};N_0'),\ldots,\operatorname{cl}(\vec{P_n};N_n'))\in B\\ \hline \\ \Sigma\ni \mathbf{n}\; :\; [\vec{\alpha}](A_0,\ldots,A_k)B\\ \\ [\sigma]B=B'\\ \hline \Sigma\vdash [\sigma]A_i\ni M_i\rhd M_i'\\ \hline \Sigma\vdash [n\; M_0\;\ldots\; M_k\; \rhd \operatorname{builtin}[\mathbf{n}](M_0';\ldots;M_k')\in B'\\ \hline \end{array}$$

Fig. 12. Synthesis Judgment

the MAST proposal from Bitcoin. While Core itself is likely impossible to use for MAST because of higher order functions, general recursion, and declared values, it is entirely possible to define a special data type and evaluation function that encodes a MAST computation. In this way, MAST becomes just a special little library in Core. Using the restricted functionality of MAST would be entirely optional, and the full power of Core would still be available for aspects of the program which do not need to be hidden or compressed.

Another benefit to using Core, instead of a lower-level language, is that it leaves the implementation of the language somewhat free. Any implementation can perfectly well choose to use Core directly, by using an evaluator in the way the standard semantics does, or it can use a virtual machine of some sort. Two such machines are given below, one a CEK machine and the other a stack machine.

IV. BASIC VALIDATION PROGRAM STRUCTURE

The basic way that validation is done in Plutus is slightly different than in Bitcoin Script. Whereas in Bitcoin Script, a validation is successful if the validating script successfully executes and leads true on the top of the stack, in Plutus, we have special constructs for this.

In particular, the success e and failure constructs in the Comp type. Any program which validates a transaction must declare a function validator : $A \to \operatorname{Comp} B$ for some A and B, while the corresponding program supplied by the redeemer must declare redeemer : $\operatorname{Comp} A$. The declarations of both are combined into a single set of declarations, and these two declared terms are then composed with a bind. The overall validation, therefore, involves evaluating the term

$$\Sigma ; \Gamma \; \vdash \; A \; type$$

A is a type using signature Σ and context Γ

$$\frac{A \ type \vdash \alpha \ type}{A \to B \ type}$$

$$\frac{A \ type \quad B \ type}{A \to B \ type}$$

$$\frac{\Sigma \ni \mathbf{n} : \star^k \quad \Sigma \vdash A_i \ type}{\Sigma \vdash \mathbf{n} \ A_0 \dots A_k \ type}$$

$$\frac{\Gamma, \alpha \ type \vdash A \ type}{\Gamma \vdash \forall \alpha. \ A \ type}$$

 $A \sqsubseteq B$

A is a subtype of B

$$\begin{array}{c} A \sqsubseteq B \\ \hline A \sqsubseteq \forall \alpha. \ B \\ \hline \\ \frac{[T/\alpha]A \sqsubseteq B}{\forall \alpha. \ A \sqsubseteq B} \\ \hline \\ A' \sqsubseteq A \quad B \sqsubseteq B' \\ \hline \\ A \to B \sqsubseteq A' \to B' \\ \hline \\ \hline \\ A \sqsubseteq A \end{array}$$

$$\Sigma; \Delta; \Gamma \; \vdash \; \vec{P} \to M \; \triangleright \; M' \; \textit{from} \; \vec{A} \; \textit{to} \; B$$

 $\vec{P} \to M$ is a clause whose body elaborates to M' with pattern types \vec{A} and body type B, using signature Σ , declarations Δ , and context Γ

$$\Sigma \vdash A_i \ \textit{pattern} \ P_i \dashv \Gamma_i'$$

$$\Sigma \ ; \ \Delta \ ; \ \Gamma, \Gamma_0', \dots, \Gamma_k' \vdash B \ni M \bowtie M'$$

$$\Sigma; \Delta; \Gamma \vdash P_0 \mid \dots \mid P_k \to M \bowtie M' \ \textit{from} \ A_0, \dots, A_k \ \textit{to} \ B$$

$$\Sigma \vdash A \ pattern \ P \dashv \Gamma'$$

Type A has pattern P using signature Σ yielding context Γ'

$$\Sigma \vdash A \text{ pattern } x \dashv x : A$$

$$\Sigma \ni \mathbf{n} : [\vec{\alpha}](A_0, \dots, A_n)B$$

$$[\sigma]B = B'$$

$$\Sigma \vdash A_i \text{ pattern } P_i \dashv \Gamma_i'$$

$$\Sigma \vdash B' \text{ pattern } \mathbf{n} P_0 \dots P_k \dashv \Gamma_0', \dots, \Gamma_k'$$

Fig. 13. Auxuliary Elaboration Judgments of Plutus

do $\{x \leftarrow \text{redeemer}; \text{validator } x\}$ (or rather, the Plutus Core term that it corresponds to). If this evaluates to success e for some e, then the transaction is valid, analogous to Bitcoin Script successfully executing and leaving true on the top of stack. On the other hand, if it evaluates to failure, then the transaction is invalid, analogous to Bitcoin Script either failing

to execute, or leaving false on the top of stack. The value returned in the success case is irrelevant by validation.

This approach to validation requires explicit control over the success or failure of the transaction. This is distinct from Bitcoin Script, where failure can occur because of some implicitly effectful operations that can immediately kill the entire execution. In Plutus, there are no effects, so any validation failure must be explicitly stated to be a failure using failure, as must any validation success with success e.

REFERENCES

[1] Harper, R. Practical Foundations for Programming Languages.