LANGUAGE-NAME: A DSL for Safe Blockchain Assets

[Authors/affiliations?]

1 INTRODUCTION

[Blockchain intro.]

Commonly proposed and implemented applications for the blockchain often revolve around the management of "digital assets." **[cite]** One of the most forms of these are a variety of smart contracts that managed assets called "tokens." There are many token standards, especially on the Ethereum blockchain **[cite]**, including **[ERC-20, ERC-721, ERC-777, ERC-1155]**, with others in various stages of the standardization process. Other applications or proposed applications for smart contracts include voting, supply chain management, auctions, lotteries, and other applications which require careful management of their respective assets **[cite/find examples]**.

LANGUAGE-NAME provides features to mark certain values as *assets* as well as a special construct called a *flow*, an abstraction representing an atomic transfer operation, which is widely applicable to smart contracts managing a variety of digital assets. These features combine to prevent common issues in smart contracts. Solidity **[cite]** and Vyper **[cite]**, the most common languages used to write smart contracts on the Ethereum blockchain **[cite]**, being general purpose languages, do not provide analgous support for managing assets.

Contributions. We make the following contributions with LANGUAGE-NAME.

- **Safety guarantees**: LANGUAGE-NAME ensures that assets are properly managed, eliminating reuse, asset-loss, and duplication bugs through the use of the flow abstraction.
- **Flow abstraction**: LANGUAGE-NAME uses a new abstraction called a *flow* to encode semantic information about the flow of assets into the code.
- **Conciseness**: LANGUAGE-NAME makes writing typical smart contract programs more concise by handling common patterns and pitfalls automatically.

2 LANGUAGE DESCRIPTION

A LANGUAGE-NAME program is made of many *contracts*, each containing *declarations*, such as *fields*, *types*, *transactions*, and *views*.

A contract is the high-level grouping of functionality, just as in Solidity; each contract instance in LANGUAGE-NAME represents a contract on the blockchain. Fields function as the peristent storage of the contract, whose data is kept on the blockchain. Type declarations are the primary way to use the type system of LANGUAGE-NAME, providing a way of annotating values as assets (as well as other modifiers, discussed in Section 2.2). In place of Solidity's **function**, we have `transaction` and `view`. In LANGUAGE-NAME, a transaction can change the state of the contract, and a view cannot; it can, however, read the current state of the contract, just like the **view** modifier in Solidity. **[Explain how things map a little bit more. Also, this explanation basically is mostly useful if you already know Solidity. Probably should put more effort into explanation for people who don't know Solidity (at least, if we submit to a non-blockchain venue).]**

Figure 1 shows a simple contract declaring a type, a field, and a transaction, which implements the core functionality of ERC-20's `transfer` function (see Section **??** for more details on ERC-20).

```
1  contract EIP20 {
2      type Token is fungible asset  uint256
3      balances : map address => Token
4      transaction transfer(dst : address, amount : uint256):
5          balances[msg.sender] --[ amount ]-> balances[dst]
6  }
```

Figure 1: A contract with a simple `transfer` function in LANGUAGE-NAME, which transfers `amount` tokens from the first to the second account. It is implemented with a single flow, which automatically checks all the preconditions to ensure the transfer is valid.

2.1 Syntax

Figure 2 shows a fragment of the syntax of the core calculus of LANGUAGE-NAME, which uses A-normal form and makes several other simplifications to the surface LANGUAGE-NAME language. These simplifications are performed automatically by the compiler. **[TODO: We have formalized this core calculus (in K???).]**

2.2 Modifiers

Modifiers can be used to place constraints on how values are managed: `asset`, `fungible`, `unique`, `immutable`, and `consumable`. A `asset` is a value that must not be reused or accidentally lost. A `fungible` value represents a quantity which can be **merged**, and it is **not** `unique`. A `unique` value can only exist in at most one storage; it must be `immutable`. A `immutable` value cannot be changed; in particular, it cannot be source or destination of a flow, the only state-changing construct in LANGUAGE-NAME**[this is the goal, anyway, which I think is true, but need to verify]**. A `consumable` value is an `asset` that it is sometimes appropriate to dispose of; however, this disposal must be done via the `consume` construct, a

q	::= ! | **any** | **nonempty**	(selector quantifiers)
Q	::= q | **empty** | **every**	(type quantities)
T	::= **bool** | **nat** | **map** $\tau \Rightarrow \sigma$ | t | ...	(base types)
τ	::= $Q\ T$	(types)
$\mathcal{V}$	::= n | **true** | **false** | **emptyval** | ...	(values)
$\mathcal{L}$	::= x | $x.x$	(locations)
E	::= $\mathcal{V}$ | $\mathcal{L}$ | **total** t | ...	(expressions)
s	::= $\mathcal{L}$ | **everything** | q $x : \tau$ **s.t.** E	(selector)
S	::= $\mathcal{L}$ | **new** t	(sources)
$\mathcal{D}$	::= $\mathcal{L}$ | **consume**	(destinations)
F	::= $S \xrightarrow{s} \mathcal{D}$	(flows)
Stmt	::= F | Stmt; Stmt | ...	(statements)
M	::= **fungible** | **immutable** | **unique**	
	| **consumable** | **asset**	(type modifiers)
Decl	::= **type** t **is** $\overline{M}\ T$	(type declaration)
	| **transaction** $m(\overline{x : \tau}) \rightarrow x : \tau$ **do** Stmt	(transactions)
	| ...	
Con	::= **contract** C { $\overline{\text{Decl}}$ }	(contracts)

Figure 2: A fragment of the abstract syntax of the core calculus of LANGUAGE-NAME, a simplified form of the surface language.

way of documenting that the disposal is intentional. All of these constraints, except for uniqueness, are enforced statically.

For example, ERC-20 tokens are **fungible**, while ERC-721 tokens are best modeled as being both **unique** and **immutable**. By default, neither is **consumable**, but one of the common extensions of both standards is to add a burn function, which allows tokens to be destroyed by users with the appropriate authentication. In this case, it would be appropriate to add the **consumable** modifier.

It supports data structures that make working with assets easier, such as *linkings*, a bidirectional map between keys and collections of values, with special operations to support modeling of "token accounts" (i.e., addresses which have a balance consisting of a set of tokens). A linking is essentially a map K =>C V with extra operations, where K is the key type, V is the value type, and C is some collection type constructor, such as list or set. These extra operations include querying which key owns which value, removing the need to maintain "parallel" data structures which separately keep track of various information about an asset; this operation is often useful (see Figure 4b for some examples). Though one can partially implemeny linkings as a library in Solidity, using the **unique** feature of LANGUAGE-NAME, we can guarantee that looking up the owner of a value which has a unique type is a **well-defined** operation.

2.3 Flows

LANGUAGE-NAME is built around the concept of a *flow*, an atomic, state-changing operation describing the transfer of a asset. Each flow has a *source* and a *destination*; they may optionally have a *selector* or a *transformer*, which default to **everything** and the identity transformer, respectively. The source of a flow *provides* values, the destination of the flow *accepts* these values, and the selector describes which subpart of the value(s) in the source should be transferred to the destination. All flows fail if the selected assets

are not present in the source, or if the selected assets cannot be added to the destination. For example, a flow of fungible assets fails if there is not enough of the asset in the source, or if there is too much in the destination; for example, the latter may occur in the case of overflow. Flows can also fail for other reasons: a developer may specify that a certain flow must send all assets matching a predicate, but in addition specify an expected *quantity* that must be selected: any number, exactly one, or at least one.

[Discuss transformers]

2.4 Error Handling

Computation on blockchains like the Ethereum blockchain are grouped into units called *transactions*. [which makes the use of the transaction keyword a little awkward, because multiple transaction calls can happen in a single transaction on the blockchain...] Transactions either succeed or they fail and revert all changes. LANGUAGE-NAME also has transactional semantics: a sequence of flows will either all succeed, or, if a single flow fails, the rest will fail as well. If a sequence of flows fails, it "bubbles up" like an exception, until it either: a) reaches the top level, at which point the entire transaction fails; or b) reaches a **catch**, in which case only the changes made inside the corresponding **try** block will be reverted, and the code inside the **catch** block will be executed.

3 CASE STUDIES

3.1 ERC-20

ERC-20 [cite] is a standard for smart contracts that manage **fungible** tokens, and provides a bare-bones interface for this purpose. Each ERC-20 contract manages the "bank accounts" for its own tokens, keeping track of which users, identified by addresses, have some number of tokens. We focus on one core function from ERC-20, the transfer function. ERC-20 is one of the commonly implemented standards on the Ethereum blockchain [cite]. [Cite all Solidity code] Figure 3 shows a Solidity implementation of the ERC-20

```
1  contract EIP20 {
2      mapping (address => uint256) private _balances;
3      function transfer (address to, uint256 value)
4          public returns (bool) {
5          require(value <= _balances[msg.sender]);
6          _balances[msg.sender] = _balances[msg.sender].sub(
                value);
7          _balances[to] = _balances[to].add(value);
8          return true;
9      }
10 }
```

Figure 3: An implementation of ERC-20's transfer function in Solidity. All preconditions are checked manually. Note that we must include the SafeMath library (not shown), which checks for underflow/overflow, to use the add and sub functions.

function transfer (cf. Figure 1). Note that event code has been omitted, because LANGUAGE-NAME handles events in the same way

as Solidity. This example shows several advantages of the flow abstraction:

- **Precondition checking**: For a flow to succeed, the source must have enough assets and the destination must be capable of receiving the assets flowed. In this case, the balance of the sender must be greater than the amount sent, and the balance of the destination must not overflow when it receives the tokens. Code checking these two conditions is automatically inserted, ensuring that the checks cannot be forgotten.
- **Data-flow tracking**: It is clear where the resources are flowing from the code itself, which may not be apparent in more complicated implementations, such as those involving transfer fees. Furthermore, developers must explicitly mark all times that assets are *consumed*, and only assets marked as consumable may be consumed. This restriction prevents, in this example, tokens from being consumed, and can also be used to ensure that other assets, like ether, are not consumed.
- **Error messages**: When a flow fails, LANGUAGE-NAME provides automatic, descriptive error messages, such as " Cannot flow '<amount>' Token from account[<src>] to account[< dst >]: source only has <balance> Token.". The default implementation provides no error message forcing developers to write their own. Flows enable the generation of the messages by encoding the semantic information of a transfer into the program, instead of using low-level operations like increment and decrement or insert and delete.

3.2 ERC-721

ERC-721 **[cite]**, like ERC-20, is a token standard for the Ethereum blockchain. ERC-721 is a standard for tokens managing *nonfungible* tokens; that is, tokens that are unique. The ERC-721 standard requires many invariants hold, including: the tokens must be unique, the tokens must be owned by at most one account, at most one non-owning account can have "approval" for a token, and only if that token has been minted, we must be able to support "operators" who can manage all of the tokens of a user, among others. Because LANGUAGE-NAME is designed to manage assets, it has features to help developers ensure that these correctness properties hold.

Figures 4a and 4b show implementations in Solidity and LANGUAGE-NAME, respectively, of the ERC-721 function transferFrom. The Solidity implementation is extracted from a reference implementation of ERC-721 given on the official Ethereum EIP page.

A LANGUAGE-NAME implementation has several benefits: because of the asset abstraction, we can be sure that token references will not be duplicated or lost; because Token has been declared as **unique**, we can be sure that we will not mint two of the same token. In addition to the invariants required by the specification, there are also internal invariants which the contract must maintain, such as the connection between idToOwner and ownerToNFTokenCount, which are handled automatically in the LANGUAGE-NAME version. This example demonstrates the benefits of having **unique** assets and the linking data structure built into the language itself.

3.3 Voting

One proposed use for the blockchain is smart contracts for managing voting **[cite]**. Figures 5a and 5b show the core of an implementation of a simple voting contract in Solidity and LANGUAGE-NAME, respectively. This example shows that LANGUAGE-NAME is suited for a range of applications, as we can use the **unique** and **immutable** modifiers to remove certain incorrect behaviors, shown in Figure 5.

[Solidity impl. comes from "Solidity by Example" page]

3.4 The DAO attack

One of the most financially impactful bugs in a smart contract on the Ethereum blockchain was the bug in the DAO contract which allowed a large quantity of ether, worth about $50 million dollars at the time, to be stolen **[cite, verifying dollar amount]**. The bug was caused by a reentrancy-unsafe function in the contract, illustrated below. **[The below is from https://consensys.github. io/smart-contract-best-practices/known_attacks/]**

```
1  function withdrawBalance() public {
2      uint  amountToWithdraw = userBalances[msg.sender];
3      // At this point, the caller's code is executed, and
4      // can call  withdrawBalance again
5      require(msg.sender.call.value(amountToWithdraw)(""));
6      userBalances[msg.sender] = 0;
7  }
```

In LANGUAGE-NAME, this attack could not have occurred for several reasons. Consider the following implementation of the same function in LANGUAGE-NAME given below.

```
1  transaction withdrawBalance():
2      userBalances[msg.sender] --> msg.sender.balance
```

Because of the additional information encoded in the flow construct, the compiler can output the safe version of the above code—reducing the balance before performing the external call—without any developer intervention. Additionally, LANGUAGE-NAME forbids any reentrant call from an external source, a similar approach to the Obsidian language **[cite]**, which would also prevent more complicated reentrancy attacks.

4 DISCUSSION

5 RELATED WORK

[Obsidian, Scilla, Move, etc.?] [TODO: The safety guarantees provided by LANGUAGE-NAME differ from those provided by other languages because...Obsidian has similar concept of assets, but doesn't allow expressing immutability or uniqueness (could do fungibility via an interface, I suppose); this is probably going to be similar to other languages built around linear type systems. Obsidian's reentrancy scheme is slightly different, Nomos has a similar global lock, it seems.]

6 FUTURE WORK

7 CONCLUSION

A FORMALIZATION

A.1 Syntax

```
1  contract NFToken {
2    mapping ( uint256 => address ) idToOwner;
3    mapping ( uint256 => address ) idToApproval;
4    mapping ( address => uint256 ) ownerToNFTokenCount;
5    mapping ( address => mapping ( address => bool ) )
          ownerToOperators;
6
7    modifier canTransfer( uint256 tokId ) {
8      address tokenOwner = idToOwner[tokId];
9      require(tokenOwner == msg.sender ||
10             idToApproval[tokId] == msg.sender ||
11             ownerToOperators[tokenOwner][msg.sender]);
12     _;
13   }
14   function transferFrom( address _from, address _to, uint256
          tokId )
15     external canTransfer(tokId) {
16     require(idToOwner[tokId] != address(0));
17     require(idToOwner[tokId] == _from);
18     require(_to != address(0));
19     if (idToApproval[tokId] != address(0)) {
20       delete idToApproval[tokId];
21     }
22     require(idToOwner[tokId] == _from);
23     ownerToNFTokenCount[_from] = ownerToNFTokenCount[_from]
            - 1;
24     delete idToOwner[tokId];
25     require(idToOwner[tokId] == address(0));
26     idToOwner[tokId] = _to;
27     ownerToNFTokenCount[_to] = ownerToNFTokenCount[_to].add(
            1);
28   }
29 }
```

(a) A Solidity implementation of ERC-721's transferFrom.

```
1  contract NFToken {
2    type Token is unique immutable asset uint256
3    type Approval is unique immutable consumable asset uint256
4    balances : linking address ⇔ set Token
5    approval : linking address ⇔ set Approval
6    ownerToOperators : linking address ⇔ set address
7
8    view canTransfer(tokId : uint256) returns bool :=
9    // `A in B` is true iff we can select `A` from `B`.
10   // It can be implemented efficiently if the LHS is hashable.
11     tokId in balances[msg.sender] or
12     tokId in approval[msg.sender] or
13     msg.sender in ownerToOperators[balances.ownerOf(tokId)]
14
15   transaction transferFrom(_from : address, _to : address, tokId
            : uint256):
16     only when _to != 0x0 and canTransfer(tokId)
17     if approval.hasOwner(tokId) {
18       approval[approval.ownerOf(tokId)] --[ tokId ]-> consume
19     }
20     balances[_from] --[ tokId ]-> balances[_to]
21 }
```

(b) LANGUAGE-NAME implementation of ERC-721's transferFrom.

[We have public and private transactions...we could also have a public/private type? The difference being that public types can be transferred between contracts.] [Random note: We can optimize merge operations on unique types by dropping some checks.]

In the surface language, "collection types" (i.e., $Q \, C \, \tau$ or a transformer) are by default **any**, but all other types, like **nat**, are !.

[Some simplification ideas] [Could get rid of selecting by locations and only allowed selecting with quantifiers, and just optimize things like ! $x : \tau$ s.t. $x = y$ into a lookup. Also, allowing any type quantity in a selector lets us do away with everything. Would actually be even nicer if we allowed any type quantity to appear in the quantifier, because then we wouldn't even need a special rule for everything.] [We could also get rid of "if" and instead do something like any $x : \tau$ s.t. if b then $x = y$ else $false$]

[Contract types should be consumable assets by default (consuming a contract is a self-destruct?)]

A.2 Statics

DEFINITION 1. *Define* **Quant** = {**empty**, **any**, **!**, **nonempty**, **every**}, *and call any* $Q \in$ **Quant** *a type quantity. Define* **empty** < **any** < ! < **nonempty** < **every**.

$\boxed{\tau \; \textbf{asset}}$ Asset Types
[The syntax for record "fields" and type environments is the same...could just use it]

$(Q \, T) \, \textbf{asset} \Leftrightarrow Q \neq \textbf{empty}$ and ($\textbf{asset} \in \textbf{modifiers}(T)$ or
$\qquad\qquad (T = \tau \leadsto \sigma$ and $\sigma \, \textbf{asset})$ or
$\qquad\qquad (T = C \, \tau$ and $\tau \, \textbf{asset})$ or
$\qquad\qquad (T = \{\overline{y : \sigma}\}$ and $\exists x : \tau \in \overline{y : \sigma}.(\tau \, \textbf{asset})))$

$\boxed{\tau \; \textbf{consumable}}$ Consumable Types

$(Q \, T) \, \textbf{consumable} \Leftrightarrow \textbf{consumable} \in \textbf{modifiers}(T)$ or $\neg((Q \, T) \, \textbf{asset})$

$Q \oplus \mathcal{R}$ represents the quantity present when flowing $\mathcal{R}$ of something to a storage already containing Q. $Q \ominus \mathcal{R}$ represents the quantity left over after flowing $\mathcal{R}$ from a storage containing Q.

```
1   contract Ballot {
2       struct Voter {
3           uint weight;
4           bool voted;
5           uint vote;
6       }
7       struct Proposal {
8           bytes32 name;
9           uint voteCount;
10      }
11
12      address public chairperson;
13      mapping(address => Voter) public voters;
14      Proposal[] public proposals;
15
16
17      function giveRightToVote(address voter) public {
18          require(msg.sender == chairperson,
19              "Only chairperson can give right to vote.");
20          require(!voters[voter].voted,
21              "The voter already voted.");
22          voters[voter].weight = 1;
23      }
24      function vote(uint proposal) public {
25          Voter storage sender = voters[msg.sender];
26          require(sender.weight != 0, "Has no right to vote");
27          require(!sender.voted, "Already voted.");
28          sender.voted = true;
29          sender.vote = proposal;
30          proposals[proposal].voteCount += sender.weight;
31      }
32      function winningProposal() public view
33              returns (bytes32 winningProposal_) {
34          uint winningVoteCount = 0;
35          for (uint p = 0; p < proposals.length; p++) {
36              if (proposals[p].voteCount > winningVoteCount) {
37                  winningVoteCount = proposals[p].voteCount;
38                  winningProposal_ = proposals[p].name;
39              }
40          }
41      }
42  }
```

(a) Solidity

```
1   contract Ballot {
2       type Voter is nonfungible asset address
3
4
5       type ProposalName is nonfungible asset string
6
7
8
9
10
11
12      chairperson : address
13      voters : set Voter
14      proposals : linking ProposalName ⇔ set Voter
15      winningProposalName : string
16
17      transaction giveRightToVote(voter : address):
18          only when msg.sender = chairperson
19          new Voter(voter) --> voters
20
21
22
23
24      transaction vote(proposal : string):
25          voters[msg.sender] --> proposals[proposal][msg.sender]
26          if total proposals[proposal] > total proposals[
27              winningProposalName] {
28              winningProposalName := proposal
29          }
30
31      view winningProposal() returns string :=
32          winningProposalName
```

(b) LANGUAGE-NAME

Figure 5: A voting contract with a set of proposals, for which each user must first be given persmission to vote by the chairperson, assigned in the constructor of the contract (not shown)> Each user can vote exactly once for exactly one proposal. The proposal with the most votes wins.

DEFINITION 2. *Let $Q, \mathcal{R} \in \mathbf{Quant}$. Define the commutative operator $\oplus$, called* combine*, as the unique function $\mathbf{Quant}^2 \to \mathbf{Quant}$* *such that*

$$
\begin{aligned}
Q \oplus \texttt{empty} &= Q \\
Q \oplus \texttt{every} &= \texttt{every} \\
\texttt{nonempty} \oplus \mathcal{R} &= \texttt{nonempty} &\textit{if } \texttt{empty} < \mathcal{R} < \texttt{every} \\
! \oplus \mathcal{R} &= \texttt{nonempty} &\textit{if } \texttt{empty} < \mathcal{R} < \texttt{every} \\
\texttt{any} \oplus \texttt{any} &= \texttt{any}
\end{aligned}
$$

$C \in \textsc{ContractNames}$ $m \in \textsc{TransactionNames}$

$t \in \textsc{TypeNames}$ $x, y, z \in \textsc{Identifiers}$

$n \in \mathbb{Z}$

q	::=	$!$ | **any** | **nonempty**	(selector quantifiers)
$Q, \mathcal{R}, \mathcal{S}$	::=	q | **empty** | **every**	(type quantities)
C	::=	**option** | **set** | **list**	(collection type constructors)
T	::=	**bool** | **nat** | $C\,\tau$	
		| **map** $\tau \Rightarrow \tau$ | **mapitem** $\tau \Rightarrow \tau$	
		| **linking** $\tau \Leftrightarrow \tau$ | **link** $\tau \Leftrightarrow \tau$	
		| $\tau \rightsquigarrow \tau$ | $\{\overline{x : \tau}\}$ | t	(base types)
τ, σ, π	::=	$Q\,T$	(types)
$\mathcal{V}$	::=	n | **true** | **false** | **emptyval** | $\lambda x : \tau.E$	(values)
$\mathcal{L}$	::=	x | $x.x$	(locations)
E	::=	$\mathcal{V}$ | $\mathcal{L}$ | $x.m(\overline{x})$ | **single**(x) | s **in** x | $\{\overline{x : \tau \mapsto x}\}$	
		| **let** $x : \tau := E$ **in** E | **if** x **then** E **else** E	
		| $x = x$ | $x \neq x$ | **total** x | **total** t	(expressions)
s	::=	$\mathcal{L}$ | **everything** | q $x : \tau$ **s.t.** E	(selector)
$\mathcal{S}$	::=	$\mathcal{L}$ | **new** t	(sources)
$\mathcal{D}$	::=	$\mathcal{L}$ | **consume**	(destinations)
F	::=	$\mathcal{S} \xrightarrow{s} x \to \mathcal{D}$	(flows)
Stmt	::=	F | E | **revert**(E) | **pack** | **unpack**(x) | **emit** $E(\overline{x})$	
		| **try** Stmt **catch**$(x : \tau)$ Stmt | **if** x **then** Stmt **else** Stmt	
		| **var** $x : \tau := E$ **in** Stmt | Stmt; Stmt	(statements)
M	::=	**fungible** | **unique** | **immutable** | **consumable** | **asset**	(type declaration modifiers)
Decl	::=	$x : \tau$	(field)
		| **event** $E(\overline{x : \tau})$	(event declaration)
		| **type** t **is** $\overline{M}\,T$	(type declaration)
		| [**private**] **transaction** $m(\overline{x : \tau}) \to x : \tau$ **do** Stmt	(transactions)
		| **view** $m(\overline{x : \tau}) \to \tau := E$	(views)
		| **on create**$(\overline{x : \tau})$ **do** Stmt	(constructor)
Con	::=	**contract** $C\,\{\,\overline{\textbf{Decl}}\,\}$	(contracts)
Prog	::=	$\overline{\textbf{Con}}$; S	(programs)

Figure 6: Abstract syntax of the core calculus of LANGUAGE-NAME.

Γ, Δ, Ξ ::= $\emptyset$ \| $\Gamma, x : \tau$ (type environments)

Define the operator $\ominus$, called *split*, as the unique function $\boldsymbol{Quant}^2 \to \boldsymbol{Quant}$ such that

$$
\begin{aligned}
Q \ominus \textbf{empty} &= Q \\
\textbf{empty} \ominus \mathcal{R} &= \textbf{empty} \\
Q \ominus \textbf{every} &= \textbf{empty} \\
\textbf{every} \ominus \mathcal{R} &= \textbf{every} &&\textit{if } \mathcal{R} < \textbf{every} \\
\textbf{nonempty} - \mathcal{R} &= \textbf{any} &&\textit{if } \textbf{empty} < \mathcal{R} < \textbf{every} \\
! - \mathcal{R} &= \textbf{empty} &&\textit{if } ! \leq \mathcal{R} \\
! - \textbf{any} &= \textbf{any} \\
\textbf{any} - \mathcal{R} &= \textbf{any} &&\textit{if } \textbf{empty} < \mathcal{R} < \textbf{every}
\end{aligned}
$$

Note that we write $(Q\,T) \oplus \mathcal{R}$ to mean $(Q \oplus \mathcal{R})\,T$ and similarly $(Q\,T) \ominus \mathcal{R}$ to mean $(Q \ominus \mathcal{R})\,T$.

DEFINITION 3. *We can consider a type environment* Γ *as a function* IDENTIFIERS $\to$ TYPES $\cup \{\bot\}$ *as follows:*

$$
\Gamma(x) = \begin{cases} \tau & \textit{if } x : \tau \in \Gamma \\ \bot & \textit{otherwise} \end{cases}
$$

We write $\textbf{dom}(\Gamma)$ *to mean* $\{x \in \textsc{Identifiers} \mid \Gamma(x) \neq \bot\}$, *and* $\Gamma|_X$ *to mean the environment* $\{x : \tau \in \Gamma \mid x \in X\}$ *(restricting the domain of* Γ*).*

DEFINITION 4. *Let* Q *and* $\mathcal{R}$ *be type quantities,* T_Q *and* $T_\mathcal{R}$ *base types, and* Γ *and* Δ *type environments. Define the following orderings, which make types and type environments into join-semilattices. For type quantities, define the partial order* $\sqsubseteq$ *as the reflexive closure of the strict partial order* $\sqsubset$ *given by*

$Q \sqsubset \mathcal{R} \Leftrightarrow (Q \neq \textbf{any} \textit{ and } \mathcal{R} = \textbf{any}) \textit{ or } (Q \in \{!, \textbf{every}\} \textit{ and } \mathcal{R} = \textbf{nonempty})$

For types, define the partial order $\leq$ *by*

$$Q\,T_Q \leq \mathcal{R}\,T_\mathcal{R} \Leftrightarrow T_Q = T_\mathcal{R} \textit{ and } Q \sqsubseteq \mathcal{R}$$

For type environments, define the partial order ≤ by

$$\Gamma \le \Delta \Leftrightarrow \forall x. \Gamma(x) \le \Delta(x)$$

Denote the join of Γ and Δ by $\Gamma \sqcup \Delta$.

$\boxed{\Gamma \vdash E : \tau \dashv \Delta}$ **Expression Typing**

These rules are for ensuring that expressions are well-typed, and keeping track of which variables are used throughout the expression. Most rules do **not** change the context, with the notable exceptions of internal calls and record-building operations. We begin with the rules for typing the various literal forms of values.

$$\frac{}{\Gamma \vdash \textbf{emptyval} : \textbf{empty } C \ \tau \dashv \Gamma} \ \text{EMPTY-VAL}$$

$$\frac{\Gamma \vdash x : \tau \dashv \Delta}{\Gamma \vdash \textbf{single}(x) : \ ! \ C \ \tau \dashv \Delta} \ \text{SINGLE}$$

$$\frac{\Gamma, x : \tau \vdash E : \sigma \dashv \Gamma, x : \pi \qquad \neg(\pi \ \textbf{asset})}{\Gamma \vdash (\lambda x : \tau. E) : \ ! \ (\tau \leadsto \sigma) \dashv \Gamma} \ \text{TRANSFORMER}$$

$$\frac{\Gamma \vdash \overline{y : \tau} \dashv \Delta}{\Gamma \vdash \{\overline{x : \tau \mapsto y}\} \dashv \Delta} \ \text{BUILD-REC}$$

Next, the lookup rules. Notably, the DEMOTE-LOOKUP rule allows the use of variables of an asset type in an expression without consuming the variable as LIN-LOOKUP does. However, it is still safe, because it is treated as its demoted type, which is always guaranteed to be a non-asset **[probably should prove this to be sure]**.

$$\frac{}{\Gamma, x : \tau \vdash x : \text{demote}(\tau) \dashv \Gamma, x : \tau} \ \text{DEMOTE-LOOKUP}$$

$$\frac{\text{demote}(\tau) \ne \tau}{\Gamma, x : Q \ T \vdash x : Q \ T \dashv \Gamma, x : \textbf{empty } T} \ \text{LIN-LOOKUP}$$

$$\frac{f : \sigma \in \overline{y : \tau}}{\Gamma, x : \{\overline{y : \tau}\} \vdash x.f : \text{demote}(\sigma) \dashv \Gamma, x : \{\overline{y : \tau}\}} \ \text{RECORD-DEMOTE-LOOKUP}$$

$$\frac{f : Q \ T \in \overline{y : \tau} \qquad \text{demote}(\sigma) \ne \sigma}{\overline{z : \sigma} = (\overline{y : \tau} \setminus \{f : Q \ T\}) \cup \{f : \textbf{empty } T\}}{\Gamma, x : \{\overline{y : \tau}\} \vdash x.f : Q \ T \dashv \Gamma, x : \{\overline{z : \sigma}\}} \ \text{RECORD-LIN-LOOKUP}$$

The expression s **in** x allows checking whether a flow will succeed without the EAFP-style ("Easier to ask for forgiveness than permission"; e.g., Python). A flow $A \xrightarrow{s} B$ is guaranteed to succeed when "s **in** A" is true and "s **in** B" is false.

$$\frac{\Gamma \vdash x \ \textbf{provides}_Q \ \tau \qquad \Gamma \vdash s \ \textbf{selects} \ \text{demote}(\tau)}{\Gamma \vdash (s \ \textbf{in} \ x) : \textbf{bool} \dashv \Gamma} \ \text{CHECK-IN}$$

We distinguish between three kinds of calls: view, internal, and external. A view call is guaranteed to not change any state in the receiver, while both internal and external calls may do so. The difference between internal and external calls is that we may transfer assets to an internal call, but **not** to an external call, because we cannot be sure any external contract will properly manage the asset

of our contract.

$$\frac{\begin{array}{c}\text{dom}(\textbf{fields}(C)) \cap \text{dom}(\Gamma) = \emptyset \\ \textbf{typeof}(C, m) = \{\overline{a : \tau}\} \leadsto \sigma \qquad \Gamma, x : C \vdash \overline{y : \tau} \dashv \Gamma, x : C\end{array}}{\Gamma, x : C \vdash x.m(\overline{y}) : \sigma \dashv \Gamma, x : C} \ \text{VIEW-CALL}$$

$$\frac{\begin{array}{c}\text{dom}(\textbf{fields}(C)) \cap \text{dom}(\Gamma) = \emptyset \\ \textbf{typeof}(C, m) = \{\overline{a : \tau}\} \leadsto \sigma \\ \Gamma, \textbf{this} : C \vdash \overline{y : \tau} \dashv \Delta, \textbf{this} : C\end{array}}{\Gamma, \textbf{this} : C \vdash \textbf{this}.m(\overline{y}) : \sigma \dashv \Delta, \textbf{this} : C} \ \text{INTERNAL-TX-CALL}$$

$$\frac{\begin{array}{c}\text{dom}(\textbf{fields}(C)) \cap \text{dom}(\Gamma) = \emptyset \\ (\textbf{transaction } m(\overline{a : \tau}) \to \sigma \ \textbf{do} \ S) \in \textbf{decls}(D) \\ \Gamma, \textbf{this} : C, x : D \vdash \overline{y : \tau} \dashv \Gamma, \textbf{this} : C, x : D\end{array}}{\Gamma, \textbf{this} : C, x : D \vdash x.m(\overline{y}) : \sigma \dashv \Gamma, \textbf{this} : C, x : D} \ \text{EXTERNAL-TX-CALL}$$

$$\frac{(\textbf{on create}(\overline{x : \tau}) \ \textbf{do} \ S) \in \textbf{decls}(C) \qquad \Gamma \vdash \overline{y : \tau} \dashv \Gamma}{\Gamma \vdash \textbf{new } C(\overline{y}) : C} \ \text{NEW-CON}$$

Finally, the rules for If and Let expressions. In LET-EXPR, we must ensure that the newly bound variable is either consumed or is not an asset after the body runs.

$$\frac{\Gamma \vdash x : \textbf{bool} \dashv \Gamma \qquad \Gamma \vdash E_1 : \tau \dashv \Delta \qquad \Gamma \vdash E_2 : \tau \dashv \Xi}{\Gamma \vdash (\textbf{if } x \ \textbf{then } E_1 \ \textbf{else } E_2) : \tau \dashv \Delta \sqcup \Xi} \ \text{IF-EXPR}$$

$$\frac{\Gamma \vdash E_1 : \tau \dashv \Delta \qquad \Delta, x : \tau \vdash E_2 : \pi \dashv \Xi, x : \sigma \qquad \neg(\sigma \ \textbf{asset})}{\Gamma \vdash (\textbf{let } x : \tau := E_1 \ \textbf{in} \ E_2) : \pi \dashv \Xi} \ \text{LET-EXPR}$$

$\boxed{\textbf{elemtype}(T) = \sigma}$

$$\textbf{elemtype}(T) = \begin{cases} \sigma & \text{if } T = C \ \sigma \\ ! \ T & \text{otherwise} \end{cases}$$

$\boxed{\Gamma \vdash \mathcal{S} \ \textbf{provides}_Q \ \tau}$ **Source Typing**

$$\frac{}{\Gamma, S : Q \ T \vdash S \ \textbf{provides}_Q \ \textbf{elemtype}(T)} \ \text{PROVIDE-VAR}$$

$$\frac{(\textbf{type } t \ \textbf{is} \ \overline{M} \ T) \in \textbf{decls}(C)}{\Gamma, \textbf{this} : C \vdash (\textbf{new } t) \ \textbf{provides}_{\textbf{every}} \ ! \ t} \ \text{PROVIDE-SOURCE}$$

$\boxed{\Gamma \vdash \mathcal{D} \ \textbf{accepts} \ \tau}$ **Destination Typing** Note that the type quantities in ACCEPT-ONE are different on the left and right of the turnstile. This is because, for example, when I have $D : \textbf{nonempty set nat}$, it is reasonable to flow into it some $S : \textbf{any set nat}$.

$$\frac{}{\Gamma, D : Q \ T \vdash S \ \textbf{accepts} \ \textbf{elemtype}(T)} \ \text{ACCEPT-VAR}$$

$$\frac{\tau \ \textbf{consumable}}{\Gamma \vdash \textbf{consume accepts} \ \tau} \ \text{ACCEPT-CONSUME}$$

$\boxed{\Gamma \vdash s \; \textbf{selects}_Q \; \tau}$ **Selectors**

$$\frac{\Gamma \vdash \mathcal{L} : Q \; T \dashv \Gamma}{\Gamma \vdash \mathcal{L} \; \textbf{selects}_Q \; \textbf{elemtype}(T)} \; \text{Select-Var}$$

$$\frac{}{\Gamma \vdash \textbf{everything} \; \textbf{selects}_{\textbf{every}} \; \tau} \; \text{Select-Everything}$$

$$\frac{\Gamma, x : \tau \vdash p : \textbf{bool} \dashv \Gamma, x : \tau}{\Gamma \vdash (\textbf{q} \; x : \tau \; \textbf{s.t.} \; p) \; \textbf{selects}_{\textbf{q}} \; \tau} \; \text{Select-Quant}$$

$\boxed{\textbf{validSelect}(s, \mathcal{R}, Q)}$ We need to ensure that the resources to be selected are easily computable. In particular, we wish to enforce that we never select **everything** from a source containing **every** of something, nor do we use a selector like q $x : \tau$ s.t. E on a source containing **every** of something. The following definition captures these restrictions.

$\textbf{validSelect}(s, \mathcal{R}, Q) \Leftrightarrow \min(Q, \mathcal{R}) < \textbf{every}$ and ($Q = \textbf{every} \implies \exists \mathcal{L}.s = \mathcal{L}$)

$\boxed{\Gamma \vdash S \; \textbf{ok} \dashv \Delta}$ **Statement Well-formedness**

Flows are the main construct for transferring resources. A flow has four parts: a source, a selector, a transformer, and a destination. The selector acts as a function that "chooses" part of the source's resources to flow. These resources then get applied to the transformer, which is an applicative functor applied to a function type. **[Bringing back one would let us do all the collections the same way in all of these flow-related rules, which would be nice.]**

$$\frac{\begin{array}{c} \Gamma \vdash A \; \textbf{provides}_Q \; \tau \qquad \textbf{immutable} \notin \textbf{modifiers}(\Gamma(A)) \\ \Gamma \vdash s \; \textbf{selects}_\mathcal{R} \; \tau \qquad \textbf{validSelect}(s, \mathcal{R}, Q) \\ \Delta = \textbf{update}(\Gamma, A, \Gamma(A) \ominus \mathcal{R}) \qquad \Delta \vdash f : \tau \leadsto \sigma \dashv \Delta \\ \Delta \vdash B \; \textbf{accepts} \; \sigma \qquad \textbf{immutable} \notin \textbf{modifiers}(\Delta(B)) \end{array}}{\Gamma \vdash (A \xrightarrow{s} f \to B) \; \textbf{ok} \dashv \textbf{update}(\Delta, B, \Delta(B) \oplus \min(Q, \mathcal{R}))} \; \text{Ok-Flow}$$

[TODO: Finish handling currying transformers.]

$$\frac{\Gamma \vdash E : \tau \dashv \Delta \qquad \Delta, x : \tau \vdash S \; \textbf{ok} \dashv \Xi, x : \sigma \qquad \neg(\sigma \; \textbf{asset})}{\Gamma \vdash (\textbf{var} \; x : \tau := E \; \textbf{in} \; S) \; \textbf{ok} \dashv \Xi} \; \text{Ok-Var-Def}$$

$$\frac{\Gamma \vdash x : \textbf{bool} \dashv \Gamma \qquad \Gamma \vdash S_1 \; \textbf{ok} \dashv \Delta \qquad \Gamma \vdash S_2 \; \textbf{ok} \dashv \Xi}{\Gamma \vdash (\textbf{if} \; x \; \textbf{then} \; S_1 \; \textbf{else} \; S_2) \; \textbf{ok} \dashv \Delta \sqcup \Xi} \; \text{Ok-If}$$

$$\frac{\Gamma \vdash S_1 \; \textbf{ok} \dashv \Delta \qquad \Gamma, x : \tau \vdash S_2 \; \textbf{ok} \dashv \Xi, x : \sigma \qquad \neg(\sigma \; \textbf{asset})}{\Gamma \vdash (\textbf{try} \; S_1 \; \textbf{catch} \; (x : \tau) \; S_2) \; \textbf{ok} \dashv \Delta \sqcup \Xi} \; \text{Ok-Try}$$

$$\frac{\Gamma \vdash E : \tau \dashv \Gamma \qquad \neg(\tau \; \textbf{asset})}{\Gamma \vdash \textbf{revert}(E) \; \textbf{ok} \dashv \Gamma} \; \text{Ok-Revert}$$

$$\frac{\Gamma \vdash E : \tau \dashv \Delta \qquad \neg(\tau \; \textbf{asset})}{\Gamma \vdash E \; \textbf{ok} \dashv \Delta} \; \text{Ok-Expr}$$

$$\frac{\Gamma \vdash S_1 \; \textbf{ok} \dashv \Delta \qquad \Delta \vdash S_2 \; \textbf{ok} \dashv \Xi}{\Gamma \vdash (S_1; S_2) \; \textbf{ok} \dashv \Xi} \; \text{Ok-Seq}$$

$$\frac{\textbf{this}.f : \tau \in \textbf{fields}(C)}{\Gamma, \textbf{this} : C \vdash \textbf{unpack}(f) \; \textbf{ok} \dashv \Gamma, \textbf{this} : C, \textbf{this}.f : \tau} \; \text{Ok-Unpack}$$

$$\frac{\begin{array}{c} (\Gamma|_{\textbf{dom}(\textbf{fields}(C))}) \le \textbf{fields}(C) \\ \Delta = \{x : \tau \in \Gamma \mid x \notin \textbf{dom}(\textbf{fields}(C))\} \end{array}}{\Gamma, \textbf{this} : C \vdash \textbf{pack} \; \textbf{ok} \dashv \Delta, \textbf{this} : C} \; \text{Ok-Pack}$$

$\boxed{\vdash_C \textbf{Decl ok}}$ **Declaration Well-formedness**

$$\frac{\Gamma = \textbf{this} : C, \textbf{fields}(C), \overline{x : \tau} \qquad \Gamma \vdash E : \sigma \dashv \Gamma}{\vdash_C (\textbf{view} \; m(\overline{x : \tau}) \to \sigma := E) \; \textbf{ok}} \; \text{Ok-View}$$

$$\frac{\begin{array}{c} \textbf{this} : C, \overline{x : \tau}, y : \textbf{empty} \; T \vdash S \; \textbf{ok} \dashv \Delta, \textbf{this} : C, y : Q \; T \\ \textbf{dom}(\textbf{fields}(C)) \cap \textbf{dom}(\Delta) = \emptyset \\ \forall x : \tau \in \Delta.\neg(\tau \; \textbf{asset}) \qquad \neg(Q \; T \; \textbf{asset}) \end{array}}{\vdash_C (\textbf{transaction} \; m(\overline{x : \tau}) \to y : Q \; T \; \textbf{do} \; S) \; \textbf{ok}} \; \text{Ok-Tx-Public}$$

$$\frac{\begin{array}{c} \textbf{this} : C, \overline{x : \tau}, y : \textbf{empty} \; T \vdash S \; \textbf{ok} \dashv \Delta, \textbf{this} : C, y : Q \; T \\ \textbf{dom}(\textbf{fields}(C)) \cap \textbf{dom}(\Delta) = \emptyset \qquad \forall x : \tau \in \Delta.\neg(\tau \; \textbf{asset}) \end{array}}{\vdash_C (\textbf{private transaction} \; m(\overline{x : \tau}) \to y : Q \; T \; \textbf{do} \; S) \; \textbf{ok}} \; \text{Ok-Tx-Private}$$

A field definition is always okay, as long as the type doesn't have the **every** modifier. **[Add this restriction to the rest of the places where we write types.] [Maybe we should always restrict variable definitions so that you can only write named types that appear in the current contract. This isn't strictly necessary, because everything will still work, but you'll simply never be able to get a value of an asset type not created in the current contract.]**

$$\frac{Q \ne \textbf{every}}{\vdash_C (x : Q \; T) \; \textbf{ok}} \; \text{Ok-Field}$$

A type declaration is okay as long as it has the **asset** modifier if its base type is an asset. Note that this restriction isn't necessary, but is intended to help users realize which types are assets without

unfolding the entire type definition.

$$\frac{T \text{ asset} \implies \text{asset} \in \overline{M}}{\vdash_C (\text{type } t \text{ is } \overline{M} \ T) \text{ ok}} \text{ OK-TYPE}$$

Note that we need to have constructors, because only the contract that defines a named type is allowed to create values of that type, and so it is not always possible to externally initialize all contract fields.

$$\frac{\text{fields}(C) = \overline{\text{this}.f : Q \ T} \qquad \text{this} : C, \overline{x : \tau}, \overline{\text{this}.f : \text{empty } T} \vdash S \text{ ok} \dashv \Delta}{\vdash_C (\text{on create}(\overline{x : \tau}) \text{ do } S) \text{ ok}}$$

OK-CONSTRUCTOR

$\boxed{\text{Con ok}}$ **Contract Well-formedness**

$$\frac{\forall d \in \overline{\text{Decl}}.(\vdash_C d \text{ ok}) \qquad \exists! d \in \overline{\text{Decl}}.\exists \overline{x : \tau}, S.d = \text{on create}(\overline{x : \tau}) \text{ do } S}{(\text{contract } C \ \{\overline{\text{Decl}}\}) \text{ ok}}$$

OK-CON

$\boxed{\text{Prog ok}}$ **Program Well-formedness**

$$\frac{\forall C \in \overline{\text{Con}}.C \text{ ok} \qquad \emptyset \vdash S \dashv \emptyset}{(\overline{\text{Con}}; S) \text{ ok}} \text{ OK-PROG}$$

Other Auxiliary Definitions. $\boxed{\text{modifiers}(T) = \overline{M}}$ **Type Modifiers**

$$\text{modifiers}(T) = \begin{cases} \overline{M} & \text{if } (\text{type } T \text{ is } \overline{M} \ T) \\ \emptyset & \text{otherwise} \end{cases}$$

$\boxed{\text{demote}(\tau) = \sigma}$ $\boxed{\text{demote}_*(T_1) = T_2}$ **Type Demotion** demote and demote$_*$ take a type and "strip" all the asset modifiers from it, as well as unfolding named type definitions. This process is useful, because it allows selecting asset types without actually having a value of the desired asset type. Note that demoting a transformer type changes nothing. This is because a transformer is **never** an asset, regardless of the types that it operators on, because it has no storage.

$$\begin{aligned} \text{demote}(Q \ T) &= Q \ \text{demote}_*(T) \\ \text{demote}_*(\text{nat}) &= \text{nat} \\ \text{demote}_*(\text{bool}) &= \text{bool} \\ \text{demote}_*(t) &= \text{demote}_*(T) && \text{where type } t \text{ is } \overline{M} \ T \\ \text{demote}_*(C) &= \text{demote}_*(\{\overline{x : \tau}\}) && \text{where fields}(C) = \{\overline{x : \tau}\} \\ \text{demote}_*(C \ \tau) &= C \ \text{demote}(\tau) \\ \text{demote}_*(\{\overline{x : \tau}\}) &= \left\{\overline{x : \text{demote}(\tau)}\right\} \\ \text{demote}_*(\tau \rightsquigarrow \sigma) &= \tau \rightsquigarrow \sigma \end{aligned}$$

$\boxed{\text{decls}(C) = \overline{\text{Decl}}}$ **Contract Declarations**

$$\text{decls}(C) = \overline{\text{Decl}} \text{ where } (\text{contract } C \ \{\overline{\text{Decl}}\})$$

$\boxed{\text{fields}(C) = \Gamma}$ **Contract Fields**

$$\text{fields}(C) = \{\text{this}.f : \tau \mid f : \tau \in \text{decls}(C)\}$$

$\boxed{\text{typeof}(C, m) = \tau \rightsquigarrow \sigma}$ **Method Type Lookup**

$$\text{typeof}(C, m) = \begin{cases} \{\overline{x : \tau}\} \rightsquigarrow \sigma & \text{if } (\text{private transaction } m(\overline{x : \tau}) \text{ returns } y : \\ \{\overline{x : \tau}\} \rightsquigarrow \sigma & \text{if } (\text{transaction } m(\overline{x : \tau}) \text{ returns } y : \sigma \text{ do } S) \in \\ \{\overline{x : \tau}\} \rightsquigarrow \sigma & \text{if } (\text{view } m(\overline{x : \tau}) \text{ returns } \sigma := E) \in \text{decls}(C) \end{cases}$$

$\boxed{\text{update}(\Gamma, x, \tau)}$ **Type environment modification**

$$\text{update}(\Gamma, x, \tau) = \begin{cases} \Delta, x : \tau & \text{if } \Gamma = \Delta, x : \sigma \\ \Gamma & \text{otherwise} \end{cases}$$

$$\forall y : \sigma \in \Delta.\neg(\sigma \text{ asset})$$

[Asset retention theorem?] [Resource accessiblity?]

[What guarantees should we provide (no errors except for flowing a resource that doesn't exist in the source/already exists in the destination)?]

NOTE: If we wanted to be "super pure", we can implement pre-conditions with just flows by doing something like:

```
1  { contractCreator = msg.sender } --[ true ]-> consume
```

This works because { contractCreator = msg.sender } : set *bool* (specifically, a singleton), so if contractCreator = msg.sender doesn't evaluate to true, then we will fail to consume true from it. **[I don't think actually doing this is a good idea; at least, not in the surface language. Maybe it would simplify the compiler and/or formalization, but it's interesting/entertaining.]**