LANGUAGE-NAME: A DSL for the Safe Blockchain Assets

Reed Oei
reedoei2@illinois.edu
University of Illinois at Urbana-Champaign
Urbana, USA

ACM Reference Format:
Reed Oei. 2020. LANGUAGE-NAME: A DSL for the Safe Blockchain Assets. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

[Authors/affiliations?]

1 INTRODUCTION

[Blockchain intro]

One of the most common applications **[cite]** of smart contracts is managing "digital 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. However, despite their widespread adoption, Solidity **[cite]**, the most common language used to write smart contracts on the Ethereum blockchain **[cite]**, provides no special support for writing such smart contracts **[cite?]**. To this end, we have developed LANGUAGE-NAME, a DSL for implementing programs which manage assets, targeted at writing smart contracts. LANGUAGE-NAME provides a special construct called a *flow* an abstraction representing an atomic transfer operation, which is widely applicable to smart contracts managing many kinds of assets, and in particular, those managing tokens.

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

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

[Potential benefits of the language. Some of these are already discussed in the paper.

- **Good expression of financial assets: fungible, nonfungible/general uniqueness constraints, consumable vs. nonconsumable. NOTE: These are things that Obsidian doesn't express automatically. Emphasize the uniqueness stuff is actually not any more difficult/inefficient than existing solutions.**
- **Atomic flow construct encodes semantic intent**
- **Maybe the language is efficient, but would need an implementation to evaluate this.**
- **Flows are interesting?**

]

2 LANGUAGE DESCRIPTION

A LANGUAGE-NAME program is made of many *contracts*, each containing *declarations*, such as *transactions*, *views*, *types*, and *fields*. A contract is a high-level unit of functionality, which behaves **["very similarly"]** to a contract in Solidity. In LANGUAGE-NAME, we distinguish between two kinds of function: transactions, which can change the state of the contract, and views, which cannot.

Figure 1 shows a simple contract declaraing a type, a field, and a transaction, which implements the core functionality of ERC-20's **transfer** function. The transaction **transfer** contains a single flow, transferring the desired amount of tokens from the transaction's sender to the destination account.

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

Figure 1: An implement of ERC-20's `transfer` function in LANGUAGE-NAME.

2.1 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 at least a *source* and a *destination*; they may optionally have a *selector* or a *transformer*, which default to **`everything`** and the identity transformer, respectively.

There are two special kinds of assets: *fungible* and *nonfungible*. **[I think there's also assets which are neither fungible nor nonfungible.] [Not sure about these definitions.]** A *fungible* assets are those whose values are not unique and can be combined: for example, ERC-20 tokens are fungible, because two accounts may have the same number of tokens—the number isn't the token, but instead describes **how many** tokens there are. A *nonfungible* asset is an asset that is unique and immutable, and can be held in at

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

$\mathfrak{q}$	::=	! | **any** | **nonempty**	(selector quantifiers)
Q	::=	$\mathfrak{q}$ | **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** | $\mathfrak{q}\ x : \tau$ **s.t.** E	(selector)
$\mathcal{S}$	::=	$\mathcal{L}$ | **new** t	(sources)
$\mathcal{D}$	::=	$\mathcal{L}$ | **consume**	(destinations)
F	::=	$\mathcal{S} \xrightarrow{s} \mathcal{D}$	(flows)
Stmt	::=	F | **Stmt**; **Stmt** | ...	(statements)
M	::=	**fungible** | **nonfungible**	
	|	**consumable** | **asset**	(type modifiers)
Decl	::=	**type** t **is** $\overline{M}\ T$	(type declaration)
	|	**transaction** $m(\overline{x : \tau})$ **returns** $x : \tau$ **do Stmt**	(transactions)
	|	...	
Con	::=	**contract** C { $\overline{\textbf{Decl}}$ }	(contracts)

Figure 2: A fragment of the abstract syntax of the core calculus of LANGUAGE-NAME.

most one location. For example, ERC-721 **[cite]** (discussed in more depth in Section 3.2) tokens are nonfungible—each token is unique and can be held by at most one account at a time. LANGUAGE-NAME dynamically ensures that all newly created nonfungible assets are unique, and statically ensures that assets are not duplicated, reused, lost, or modified (if immutable). Furthermore, it supports data structures that make working with assets easier, such as *linkings*, a bidirectional mapping between keys and a collection of values, with special operations to support modeling of "token accounts" (i.e., addresses which have a balance consisting of a set of tokens.

The source and destination of a flow are two storages which *provide* and *accept* assets, and the selector describes which subpart of the value of the asset(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, and a flow of a nonfungible asset fails if the selected values don't exist in the source location. 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.

2.2 Syntax

Figure 6 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???).]**

3 CASE STUDIES

3.1 ERC-20

[Cite all Solidity code properly]

Figure 3 shows implementations of the ERC-20 **[cite]** standard in both Solidity and LANGUAGE-NAME, one of the most commonly implemented standards on the Ethereum blockchain **[cite]**. Only the core functions of **transfer**, transferFrom, and approve are shown, with the exception of totalSupply in the LANGUAGE-NAME implementation (included because to show the use of the **total** operator). All event code has been omitted, because LANGUAGE-NAME handles events in the same way as Solidity. This contract 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 **[TODO: **will provide**]** automatic, descriptive error messages, such as "Cannot flow '<amount>'Token from account[<src>] to account[<dst>]: source only has <amount>Token .". **[Not sure exactly what the error message should be.]** 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 incrementing and decrementing.

3.2 ERC-721

The ERC-721 standard **[cite]** requires many invariants hold: the tokens must be unique, at most one non-owning account can have "approval" for a token, 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 handle assets, it has features to help developers ensure that these correctness properties hold. 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 **nonfungible**, we can be sure that we will not mint two of the same token.

Figure 4 shows an implementation ERC-721's transferFrom function in both Solidity and LANGUAGE-NAME. The Solidity implementation is extracted from one of the reference implementations of ERC-721 given on its official Ethereum EIP page. In addition to the invariant required by the specification, there are also internal invariant which the contract must maintain, such as the connection between idToOwner and ownerToNFTokenCount, which are handled by LANGUAGE-NAME. This example demonstrates the benefits

```
contract EIP20 {
    mapping (address => uint256) balances;
    mapping (address => mapping (address => uint256)) allowed;
    function transferFrom(address from, address to, uint256 value)
        public returns (bool success) {
        require(balances[from] >= value &&
                allowed[from][msg.sender] >= value);
        balances[to] += value;
        balances[from] -= value;
        allowed[from][msg.sender] -= value;
        return true;
    }
    function approve(address spender, uint256 value)
        public returns (bool success) {
        allowed[msg.sender][spender] = value;
        return true;
    }
}
```

```
contract EIP20 {
    type Token is fungible asset uint256
    type Approval is fungible consumable asset uint256
    accounts : map address => Token
    allowances : map address => map address => Approval
    transaction transferFrom(src : address, dst : address, amount :
        uint256):
        allowances[src][dst] --[ amount ]-> consume
        account[src] --[ amount ]-> account[dst]
    transaction approve(dst : address, amount : uint256):
        allowances[msg.sender][dst] --> consume
        new Approval --[ amount ]-> allowances[msg.sender][dst]
}
```

Figure 3: A Solidity and a LANGUAGE-NAME implementation of the core functions of the ERC-20 standard.

of having nonfungible assets and linkings built into the language itself.

3.3 Voting

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

[Can include this section if we don't only want to talk about tokens...] The **`nonfungible`** modifier is useful in many programs, even those not dealing with financial assets, like ERC-721 contracts. We can also use **`nonfungible`** to remove certain incorrect behaviors from a voting contract, shown in Figure 5.

3.4 The DAO attack

One of the most financially impactful bugs in 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 relied on a reentrancy-unsafe function in the contract, illustrated below. **[The below is from https://consensys.github.io/smart-contract-best-practices/known_attacks/]**

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

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.

```
transaction withdrawBalance():
    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.?]

6 CONCLUSION

A FORMALIZATION

A.1 Syntax

[We have public and private transactons...we could also have a public/private type?]

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?)]

```
contract NFToken {
  mapping(uint256 => address) idToOwner;
  mapping(uint256 => address) idToApproval;
  mapping(address => uint256) ownerToNFTokenCount;
  mapping(address => mapping(address => bool)) ownerToOperators;
  modifier canTransfer(uint256 _tokenId) {
    address tokenOwner = idToOwner[_tokenId];
    require(tokenOwner == msg.sender ||
            idToApproval[_tokenId] == msg.sender ||
            ownerToOperators[tokenOwner][msg.sender],
            NOT_OWNER_APPROWED_OR_OPERATOR);
    _;
  }
  modifier validNFToken(uint256 _tokenId) {
    require(idToOwner[_tokenId] != address(0), NOT_VALID_NFT);
    _;
  }
  function transferFrom(address _from, address _to, uint256 _tokenId)
    external override canTransfer(_tokenId) validNFToken(_tokenId) {
    require(idToOwner[_tokenId] == _from, NOT_OWNER);
    require(_to != address(0), ZERO_ADDRESS);
    address from = idToOwner[_tokenId];
    if (idToApproval[_tokenId] != address(0)) {
      delete idToApproval[_tokenId];
    }
    require(idToOwner[_tokenId] == _from, NOT_OWNER);
    ownerToNFTokenCount[_from] = ownerToNFTokenCount[_from] - 1;
    delete idToOwner[_tokenId];
    require(idToOwner[_tokenId] == address(0), NFT_ALREADY_EXISTS)
        ;
    idToOwner[_tokenId] = _to;
    ownerToNFTokenCount[_to] = ownerToNFTokenCount[_to].add(1);
}
```

```
contract NFToken {
    type Token is nonfungible asset nat
    type TokenApproval is nonfungible consumable asset nat
    balances : linking address ⇔ set Token
    approval : linking address ⇔ set TokenApproval
    ownerToOperators : linking address ⇔ set address
    view canTransfer(_tokenId : nat) returns bool :=
        _tokenId in balances[msg.sender] or
        _tokenId in approval[msg.sender] or
        msg.sender in ownerToOperators[balances.ownerOf(_tokenId)]
    view validNFToken(_tokenId : nat) returns bool := balances.
        hasOwner(_tokenId)
    transaction transferFrom(_from : address, _to : address, _tokenId
        : nat):
        only when _to != 0x0 and canTransfer(_tokenId)
        if approval.hasOwner(_tokenId) {
            approval[approval.ownerOf(_tokenId)] --[ _tokenId ]->
                consume
        }
        balances[_from] --[ _tokenId ]-> balances[_to]
}
```

Figure 4: A Solidity and a LANGUAGE-NAME implementation of the transferFrom function of the ERC-721 standard.

A.2 Statics

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

$\boxed{\tau\ \texttt{asset}}$ **Asset Types**

[The syntax for record "fields" and type environments is the same...could just use it]

$$(Q\ T)\ \texttt{asset} \iff Q \neq \texttt{empty} \text{ and } (\texttt{asset} \in \texttt{modifiers}(T) \text{ or } (T = \tau \rightsquigarrow \sigma \text{ and } \sigma\ \texttt{asset}) \text{ or } (T = C\ \tau \text{ and } \tau\ \texttt{asset}) \text{ or } (T = \{\overline{y : \sigma}\} \text{ and } \exists x : \tau \in \overline{y : \sigma}.(\tau\ \texttt{asset})))$$

$\boxed{\tau\ \texttt{consumable}}$ **Consumable Types**

$$(Q\ T)\ \texttt{consumable} \iff \texttt{consumable} \in \texttt{modifiers}(T) \text{ or } \neg((Q\ T)\ \texttt{asset})$$

$Q \oplus \mathcal{R}$ 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.

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

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

```
contract Ballot {
    struct Voter {
        uint weight;
        bool voted;
        uint vote;
    }
    struct Proposal {
        bytes32 name;
        uint voteCount;
    }
    address public chairperson;
    mapping(address => Voter) public voters;
    Proposal[] public proposals;
    function giveRightToVote(address voter) public {
        require(msg.sender == chairperson,
            "Only chairperson can give right to vote.");
        require(!voters[voter].voted,
            "The voter already voted.");
        voters[voter].weight = 1;
    }
    function vote(uint proposal) public {
        Voter storage sender = voters[msg.sender];
        require(sender.weight != 0, "Has no right to vote");
        require(!sender.voted, "Already voted.");
        sender.voted = true;
        sender.vote = proposal;
        proposals[proposal].voteCount += sender.weight;
    }
    function winningProposal() public view
            returns (uint winningProposal_) {
        uint winningVoteCount = 0;
        for (uint p = 0; p < proposals.length; p++) {
            if (proposals[p].voteCount > winningVoteCount) {
                winningVoteCount = proposals[p].voteCount;
                winningProposal_ = p;
            }
        }
    }
    function winnerName() public view
            returns (bytes32 winnerName_) {
        winnerName_ = proposals[winningProposal()].name;
    }
}
```

```
contract Ballot {
    type Voter is nonfungible asset address
    type ProposalName is nonfungible asset string
    chairperson : address
    voters : set Voter
    proposals : linking ProposalName ⇔ set Voter
    winningProposalName : string
    transaction giveRightToVote(voter : address):
        only when msg.sender = chairperson
        new Voter(voter) --> voters
    transaction vote(proposal : string):
        voters[msg.sender] --> proposals[proposal][msg.sender]
        if total proposals[proposal] > total proposals[
            winningProposalName] {
            winningProposalName := proposal
        }
    view winningProposal() returns string := winningProposalName
}
```

Figure 5: A Solidity and a LANGUAGE-NAME implementation of a simple voting contract.

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

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

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

Definition 3. *We can consider a type environment* Γ *as a function* $\textit{Identifiers} \to \textit{Types} \cup \{\bot\}$ *as follows:*

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

$C \in$ ContractNames $\quad m \in$ TransactionNames

$t \in$ TypeNames $\quad x, y, z \in$ Identifiers

$n \in \mathbb{Z}$

$\mathfrak{q}$	::=	! | **any** | **nonempty**
$\mathcal{Q}, \mathcal{R}, \mathcal{S}$	::=	$\mathfrak{q}$ | **empty** | **every**
C	::=	**option** | **set** | **list**
T	::=	**bool** | **nat** | $C\ \tau$ | $\tau \rightsquigarrow \tau$ | $\{\overline{x : \tau}\}$ | t
τ, σ, π	::=	$\mathcal{Q}\ T$
$\mathcal{V}$	::=	n | **true** | **false** | **emptyval** | $\lambda x : \tau.E$
$\mathcal{L}$	::=	x | $x.x$
E	::=	$\mathcal{V}$ | $\mathcal{L}$ | $x.m(\overline{x})$ | **some**(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
s	::=	$\mathcal{L}$ | **everything** | $\mathfrak{q}\ x : \tau$ **s.t.** E
$\mathcal{S}$	::=	$\mathcal{L}$ | **new** t
$\mathcal{D}$	::=	$\mathcal{L}$ | **consume**
F	::=	$\mathcal{S} \xrightarrow{s} x \to \mathcal{D}$
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**
M	::=	**fungible** | **nonfungible** | **consumable** | **asset**
Decl	::=	$x : \tau$
	|	**event** $E(\overline{x : \tau})$
	|	**type** t **is** $\overline{M}\ T$
	|	[**private**] **transaction** $m(\overline{x : \tau})$ **returns** $x : \tau$ **do Stmt**
	|	**view** $m(\overline{x : \tau})$ **returns** $\tau := E$
	|	**on create**$(\overline{x : \tau})$ **do Stmt**
Con	::=	**contract** C { $\overline{\textbf{Decl}}$ }
Prog	::=	$\overline{\textbf{Con}}$; S

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

$$\Gamma, \Delta, \Xi \quad ::= \quad \emptyset \mid \Gamma, x : \tau \quad \text{(type environments)}$$

We write $\textbf{dom}(\Gamma)$ *to mean* $\{x \in$ 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* $\mathcal{Q}$ *and* $\mathcal{R}$ *be type quantities,* $T_\mathcal{Q}$ *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*

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

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

$$\mathcal{Q}\ T_\mathcal{Q} \leq \mathcal{R}\ T_\mathcal{R} \iff T_\mathcal{Q} = T_\mathcal{R} \text{ and } \mathcal{Q} \sqsubseteq \mathcal{R}$$

For type environments, define the partial order $\leq$ *by*

$$\Gamma \leq \Delta \iff \forall x.\Gamma(x) \leq \Delta(x)$$

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.

(selector quantifiers)
(type quantities)
(collection type constructors)
(base types)
(types)
(values)
(locations)
(expressions)
(selector)
(sources)
(destinations)
(flows)

$$\frac{}{\Gamma \vdash \textbf{emptyval} : \textbf{empty}\ C\ \tau \dashv \Gamma}\ \textsc{Empty-Val}$$

$$\frac{\Gamma, x : \tau \vdash E : \sigma \dashv \Gamma}{\Gamma \vdash (\lambda x : \tau.E) : \textbf{every list}\ !\ (\tau \rightsquigarrow \sigma) \dashv \Gamma}\ \textsc{Transformer}$$

$$\frac{\Gamma \vdash x : \tau \dashv \Delta}{\Gamma \vdash \textbf{some}(x) : !\ \textbf{option}\ \tau \dashv \Delta}\ \textsc{Some}$$

$$\frac{\Gamma \vdash \overline{y : \tau} \dashv \Delta}{\Gamma \vdash \{\overline{x : \tau \mapsto y}\} \dashv \Delta}\ \textsc{Build-Rec}$$

(statements)
(type declaration modifiers)
(field)
(event declaration)
(type declaration)
(transactions)
(views)
(constructor)
(contracts)
(programs)

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.

$$\frac{}{\Gamma, x : \tau \vdash x : \text{demote}(\tau) \dashv \Gamma, x : \tau}\ \textsc{Demote-Lookup}$$

$$\frac{}{\Gamma, x : \mathcal{Q}\ T \vdash x : \mathcal{Q}\ T \dashv \Gamma, x : \textbf{empty}\ T}\ \textsc{Lin-Lookup}$$

$$\frac{\Gamma \vdash x : \{\overline{y : \tau}\} \vdash \Gamma \qquad f : \sigma \in \overline{y : \tau}}{\Gamma \vdash x.f : \sigma \vdash \Gamma}\ \textsc{Record-Field-Lookup}$$

[Record field lookup rule doesn't take into account that fields can store assets...]

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.

$$\frac{\Gamma \vdash x\ \textbf{provides}_\mathcal{Q}\ \tau \qquad \Gamma \vdash s\ \textbf{selects}\ \text{demote}(\tau)}{\Gamma \vdash (s\ \textbf{in}\ x) : \textbf{bool} \dashv \Gamma}\ \textsc{Check-In}$$

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.

$$\frac{\textbf{typeof}(C, m) = \{\overline{a : \tau}\} \rightsquigarrow \sigma \qquad \Gamma, x : C \vdash \overline{y : \tau} \dashv \Gamma, x : C}{\Gamma, x : C \vdash x.m(\overline{y}) : \sigma \dashv \Gamma, x : C}\ \textsc{View-Call}$$

$$\frac{\textbf{dom}(\textbf{fields}(C)) \cap \textbf{dom}(\Gamma) = \emptyset \qquad \textbf{typeof}(C, m) = \{\overline{a : \tau}\} \rightsquigarrow \sigma \qquad \Gamma, \textbf{this} : C \vdash \overline{y : \tau} \dashv \Delta, \textbf{this} : C}{\Gamma, \textbf{this} : C \vdash \textbf{this}.m(\overline{y}) : \sigma \dashv \Delta, \textbf{this} : C}\ \textsc{Internal-Tx-Call}$$

$$\frac{\textbf{dom}(\textbf{fields}(C)) \cap \textbf{dom}(\Gamma) = \emptyset \qquad (\textbf{transaction}\ m(\overline{a : \tau})\ \textbf{returns}\ \sigma\ \textbf{do}\ S) \in \underline{\textbf{decls}}(D) \qquad \Gamma, \textbf{this} : C, x : D \vdash \overline{y : \tau} \dashv \Gamma, \textbf{this} : C, x : D \qquad \Gamma \vdash \textbf{everything}\ \textbf{selects}_{\textbf{every}}\ \tau}{\Gamma, \textbf{this} : C, x : D \vdash x.m(\overline{y}) : \sigma \dashv \Gamma, \textbf{this} : C, x : D}\ \textsc{External-Tx-Call}\ \textsc{Select-Everything}$$

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

[Add method typing as transformers]

Finally, the rules for If and Let expressions. In Let-Expr, we ensure that the newly bound variable is either consumed or is not an asset in the body.

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

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

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

$$\frac{}{\Gamma, S : \tau \vdash S\ \textbf{provides}_!\ \tau}\ \textsc{Provide-One}$$

$$\frac{}{\Gamma, S : \mathcal{Q}\ C\ \tau \vdash S\ \textbf{provides}_Q\ \tau}\ \textsc{Provide-Col}$$

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

[Note, it will be too difficult to implement to make every kind of selector work with the sources, because the quantified selector can contain arbitrary expressions. It needs to be restricted somehow; the current rules only ensure you don't flow everything from a source. Could write special Flow-Source rules.]

$\boxed{\Gamma \vdash \mathcal{D}\ \textbf{accepts}\ \tau}$ **Destination Typing** **[Prevent variables that are supposed to store exactly one of something from receiving another?]** 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 : **nonempty set nat**, it is reasonable to flow into it some S : **any set nat**.

$$\frac{}{\Gamma, D : \mathcal{Q}\ T \vdash S\ \textbf{accepts}\ \mathcal{R}\ T}\ \textsc{Accept-One}$$

$$\frac{}{\Gamma, D : \mathcal{Q}\ C\ \tau \vdash S\ \textbf{accepts}\ \tau}\ \textsc{Accept-Col}$$

$$\frac{\tau\ \textbf{consumable}}{\Gamma \vdash \textbf{consume}\ \textbf{accepts}\ \tau}\ \textsc{Accept-Consume}$$

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

$$\frac{\Gamma \vdash \mathcal{L} : \mathcal{Q}\ T \dashv \Gamma}{\Gamma \vdash \mathcal{L}\ \textbf{selects}_Q\ \mathcal{Q}\ T}\ \textsc{Select-Loc}$$

$$\frac{\Gamma \vdash x : \mathcal{Q}\ C\ \tau \dashv \Gamma}{\Gamma \vdash x\ \textbf{selects}_Q\ \tau}\ \textsc{Select-Col}$$

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

$\boxed{\textbf{validSelect}(s, \mathcal{R}, \mathcal{Q})}$ 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 $\mathfrak{q}x : \tau$ s.t. E on a source contianing **every** of something. The following definition captures these restrictions.

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

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

[In the new flow rule, we always use a transformer. However, that just means we desugar something like $A \xrightarrow{s} B$ into $A \xrightarrow{s} (\lambda x : \tau.x) \to B$. In the real compiler, this can be optimized.]

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.]**

(Non)Ambiguity of Flow Rules. Consider the flow $A \xrightarrow{s} f \to B$. **[Actually, the type of f will probably be enough to distinguish the cases, but if we want to desguar the flows into flows containing a transformer always then we would have to infer its type and run into the same issue again.]** The only way that the choice of which Provide, Select, or Accept rules could be ambiguous **[I think...]** is if A and B are both collections containing the same type, and either s is a collection containing the same type or it is **everything**. If A, B, and s are all collections containing the same type, then we could use either version of the rules (the appropriate One rule or the appropriate Col rule). However, regardless of the rule we choose, the outcome will be the same. For example, if we use the Select-Loc rule, A will now store the quantity **any** (unless s is **empty**), which is correct, because we don't know how many values will be transferred by s. Finally, if s is **everything**, the same argument applies—the outcome will be the same regardless of which rule we pick.

$$\frac{\Gamma \vdash A\ \texttt{provides}_Q\ \tau \qquad \Gamma \vdash s\ \texttt{selects}_{\mathcal{R}}\ \tau \qquad \texttt{validSelect}(s, \mathcal{R}, Q) \qquad \Delta = \texttt{update}(\Gamma, A, \Gamma(A) \ominus \mathcal{R}) \qquad \Delta \vdash f : \tau \rightsquigarrow \sigma \dashv \Delta \qquad \Delta \vdash B\ \texttt{accepts}\ \sigma}{\Gamma \vdash (A \xrightarrow{s} f \to B)\ \texttt{ok} \dashv \texttt{update}(\Delta, B, \Delta(B) \oplus \min(Q, \mathcal{R}))}\ \textsc{Ok-Flow}$$

[TODO: Finish handling currying transformers.]

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

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

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

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

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

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

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

$$\frac{(\Gamma|_{\texttt{dom}(\texttt{fields}(C))}) \le \texttt{fields}(C) \qquad \Delta = \{x : \tau \in \Gamma \mid x \notin \texttt{dom}(\texttt{fields}(C))\}}{\Gamma, \texttt{this} : C \vdash \texttt{pack}\ \texttt{ok} \dashv \Delta, \texttt{this} : C}\ \textsc{Ok-Pack}$$

$\vdash_C$ `Decl` ok — Declaration Well-formedness

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

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

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

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.]**

$$\frac{Q \neq \texttt{every}}{\vdash_C (x : Q\ T)\ \texttt{ok}}\ \textsc{Ok-Field}$$

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. **[Need to ensure that nonfungible types are immutable.]**

$$\frac{T\ \texttt{asset} \implies \texttt{asset} \in \overline{M}}{\vdash_C (\texttt{type}\ t\ \texttt{is}\ \overline{M}\ T)\ \texttt{ok}}\ \textsc{Ok-Type}$$

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.

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

`Con` ok — Contract Well-formedness

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

`Prog` ok — Program Well-formedness

$$\frac{\forall C \in \overline{\texttt{Con}}.C\ \texttt{ok} \qquad \emptyset \vdash S \dashv \emptyset}{(\overline{\texttt{Con}}; S)\ \texttt{ok}}\ \textsc{Ok-Prog}$$

Other Auxiliary Definitions. **[Eliminate all the locations except for x and then use flows to extract and put stuff back?]**

`modifiers`$(T) = \overline{M}$ — Type Modifiers

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

demote$(\tau) = \sigma$ | 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. **[TODO: Transformer demotion? My current thought is we should split functions and transformers, with the latter being able to "hold" a resource after being partially applied, and therefore being able to be an asset. Alternatively, can just not allow for currying...]**

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

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

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

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

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

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

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

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

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

[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 preconditions 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 `consume` 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.]**