Decorated semantics for an imperative language with exceptions

Dominique Duval, with Jean-Guillaume Dumas, Burak Ekici, Damien Pous and Jean-Claude Reynaud

Work in progress

GdT Plume, ENS Lyon, 21 mars 2016

The language IMP-EX

Syntax

Arithmetic expressions:

$$a ::= 0 \mid 1 \mid -1 \mid 2 \mid -2 \mid ... \mid \ell_1 \mid \ell_2 \mid \cdots \mid a+a \mid a-a \mid a \times a$$

Boolean expressions:

$$b ::= true \mid false \mid \neg b \mid b \land b \mid b \lor b \mid a = a \mid a > a$$

Commands:

$$c ::= \begin{cases} \text{skip} \mid c; c \mid \ell_i := a \mid \\ \text{if}(b) \text{then}(c) \text{else}(c) \mid \text{while}(b) \text{do}(c) \mid \\ \text{throw}(exn_i) \mid \text{try}(c) \text{catch}(exn_i \Rightarrow c) \end{cases}$$

Programs:

$$pg ::= c; return(a) | c; return(b)$$

Semantics

Denotational: in the category of sets and partial functions

Operational: small-step, big-step

Predicate transformer semantics, ...

Theorem "All semantics for IMP-EX coincide".

Aims and tools

Aims.

- ▶ Design a "kind of" equational logic L, close to the syntax, for reasoning about imperative programs with exceptions.
- ► Translate the syntax of IMP-EX into the logic £.
- **Prove** properties of programs of IMP-EX in the logic \mathcal{L} .
- Implement this proof system in Coq.

Tools.

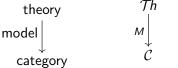
- [Moggi 1989] "effects as monads".
 Terms of type B with a parameter of type A are not interpreted by morphisms from A to B but by morphisms from A to T(B) for some monad T.
- Here, more generally, "effects as functors".
 Terms of type B with a parameter of type A are not interpreted by morphisms from A to B but by morphisms from H(A) to H(B) for some functor H.

Outline

Logic and categories

- Syntax and equational semantics: a theory Th (a category with a congruence ≡) generated by a signature and equations.
- ▶ Denotational semantics: a model $M: \mathcal{T}h \to \mathcal{C}$ (a functor mapping \equiv to =) where \mathcal{C} is "given by mathematics" (e.g., $\mathcal{C} = \mathcal{S}et$ or $\mathcal{C} = \mathcal{P}art$).

$$c = c c c c c c = c a c c$$





Soundness: granted

Remark: usually structured categories and functors

Decorated logic: theories and models

Simply "enlarge" the previous diagram

$$\begin{array}{cccc}
\mathcal{T}h_0 & \xrightarrow{\subseteq} \mathcal{T}h_1 & \xrightarrow{\subseteq} \cdots & \xrightarrow{\subseteq} \mathcal{T}h_n \\
\mathcal{C}_0 & & & \downarrow \\
\mathcal{C}_0 & & & \downarrow \\
H_1 & & & \downarrow \\
H_2 & & & \downarrow \\
H_n & & & \downarrow \\
M_n & &$$

where the functor $\mathcal{T}h_{i-1} \subseteq \mathcal{T}h_i$

- ▶ is the identity on objects
- ▶ preserves \equiv and is " \equiv -faithful": for all $f, g: X \to Y$ in $\mathcal{T}h_{i-1}$ $f \equiv g$ in $\mathcal{T}h_{i-1} \iff f \equiv g$ in $\mathcal{T}h_i$

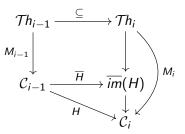
Decoration of terms (notation): $f^{(d)}$ iff $f \in \mathcal{T}h_d$ conversions: $f^{(d)} \implies f^{(d+1)}$

Soundness: if each H_i is faithful

Full image

The full image of a functor $H: \mathcal{C}_{i-1} \to \mathcal{C}_i$ is the category $\overline{im}(H)$ with:

- ▶ the same objects as C_{i-1}
- ▶ an arrow $f: X \to Y$ for each $f: H(X) \to H(Y)$ in C_i .



$$(X \to Y) \longmapsto (X \to Y)$$

$$\downarrow \qquad \qquad \downarrow$$

$$(X \to Y) \longmapsto (X \to Y)$$

$$\downarrow \qquad \qquad \downarrow$$

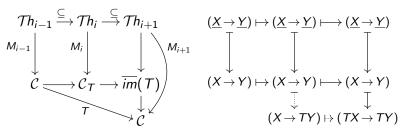
$$(HX \to HY)$$

Soundness: if \overline{H} is faithful

Kleisli category

The Kleisli category of a monad $T: \mathcal{C} \to \mathcal{C}$ is the category \mathcal{C}_T with:

- ightharpoonup the same objects as ${\cal C}$
- ▶ an arrow $f: X \to Y$ for each $f: X \to T(Y)$ in C.



Soundness: if each component of the unit $\eta: Id \Rightarrow T$ is mono

Decorated logic: decorated equations

Notation:
$$\boxed{f \bullet g = g \circ f}$$
 when $\bullet \xrightarrow{f} \bullet \xrightarrow{g} \bullet$

In each theory:

- ▶ a congruence ≡:
 - equivalence relation between parallel terms
 - compatible with composition

$$g_1 \equiv g_2 \implies f \bullet g_1 \bullet h \equiv f \bullet g_2 \bullet h$$

- ► a weak congruence (or several):
 - extends \equiv
 - preorder relation between parallel terms
 - "sometimes" symmetry
 - "sometimes" substitution

$$g_1 \equiv g_2 \implies f \bullet g_1 \equiv f \bullet g_2$$

- "sometimes" replacement

$$g_1 \equiv g_2 \implies g_1 \bullet h \equiv g_2 \bullet h$$

Outline

The language XS-IMP

Syntax

Expressions:

```
a ::= 0 | 1 | -1 | 2 | -2 | ... | \ell | s(a) | p(a)

b ::= true | false | \neg b | a = 0 | a > 0

e ::= a | b
```

Commands:

$$c ::= skip \mid c; c \mid \ell := a$$

Programs:

$$pg ::= c; return(e)$$

Restrictions (easy to remove):

- ightharpoonup only one location ℓ
- no binary operation on expressions

Later:

exceptions, conditionals, loops

Decorated logic for states

 \sim_{st} satisfies substitution and pure replacement:

$$g_1 \sim_{st} g_2 \implies f \bullet g_1 \bullet h^{(0)} \sim_{st} f \bullet g_2 \bullet h^{(0)}$$

Pure operations and equations

The pure theory Th_0 contains:

- ▶ sorts 1, A, B
- ▶ operations $0, 1, -1, ... : 1 \rightarrow A$, $s, p : A \rightarrow A$, true, false : $1 \rightarrow B$, not : $B \rightarrow B$, null?, pos? : $A \rightarrow B$
- ▶ equations $s(0) \equiv 1$, $p(0) \equiv -1$,..., $s \bullet p \equiv id_A$, $p \bullet s \equiv id_A$, true \bullet not \equiv false, ...

 $M_0: \mathcal{T}h_0 \to \mathcal{S}et$ interprets A as the set A of integers, B as the set B of truth values, etc

Operations and equations for states

In Set: a set of states S with (here) $S \cong A$, denoted $x \leftrightarrow x$ Then Th_1 and Th_2 are generated from Th_0 by two operations:

$\texttt{lookup}^{(1)}: \mathbb{1} \to A$	$\texttt{update}^{(2)}:A\to\mathbb{1}$
$\texttt{lookup}: \mathcal{S} \to \mathcal{A}$	$\mathtt{update}: S {\times} A \to S$
$lookup: x \mapsto x$	$\mathtt{update}: (x, y) \mapsto y$

one weak equation:

$$\begin{array}{c} \text{update} \bullet \text{lookup} \sim_{st} id_A \\ \\ \text{update} \bullet \text{lookup} = \varepsilon_A \\ \\ ([x], y) \mapsto [y] \mapsto y \end{array}$$

and decorated rules...



Translation

Expressions:
$$e \mapsto e^{(1)} : \mathbb{1} \to Expr$$
 (where $Expr$ is A or B)

- $ightharpoonup 0,1,...\mapsto 0^{(0)},1^{(0)},..., \; \mathtt{true},\mathtt{false}\mapsto \mathtt{true}^{(0)},\mathtt{false}^{(0)}$
- $ightharpoonup s(a)\mapsto aullet s^{(0)},\ p(a)\mapsto aullet p^{(0)},\ \neg b\mapsto bullet {
 m not}^{(0)},\ \dots$
- ▶ $\ell \mapsto lookup^{(1)}$

Commands:
$$c \mapsto c^{(2)} : 1 \to 1$$

- $skip \mapsto id_{1}^{(0)}$
- $ightharpoonup c_1$; $c_2 \mapsto c_1 \bullet c_2$
- $label{eq:lambda} \ell := a \mapsto a \bullet \mathtt{update}^{(2)}$

Programs:
$$pg \mapsto pg^{(2)} : \mathbb{1} \to Expr$$

ightharpoonup c; return(e) $\mapsto c \bullet e$

Forward semantics

Given a program $pg^{(2)}: \mathbb{1} \to Expr$, find a result $rs^{(0)}: \mathbb{1} \to Expr$ such that $pg \sim_{st} rs$. This means that $pg: S \to S \times Expr$ and $rs: \mathbb{1} \to Expr$ satisfy: pg(s) = (s', rs(x)) for some s'

$$egin{aligned} egin{aligned} egin{aligned} egin{aligned} egin{aligned} S &= & (s', rs(x)) & ext{for some } s' \ S &= & & S imes Expr \ & = & & \downarrow arepsilon_{Expr} \ & 1 &= & & Expr \end{aligned}$$

This requires an initialization of the state and the derived strong equation: for each $u^{(0)}: \mathbb{1} \to A$

$$u \bullet \text{update} \bullet \text{lookup} \equiv u \bullet \text{update} \bullet u$$

Method:

- first ≡ is used inductively, by replacement
- until finally \sim_{st} can be used, by pure replacement

This corresponds to an operational semantics.



Forward semantics: an example

```
Initialization: \ell := u^{(0)} for any u^{(0)} : \mathbb{1} \to A
The given program is
        \ell := u; \ \ell := s(\ell); \ \text{return}(p(\ell))
translated as:
pg^{(2)} = u^{(0)} \bullet \text{update}^{(2)} \bullet \text{lookup}^{(1)} \bullet s^{(0)} \bullet \text{update}^{(2)} \bullet \text{lookup}^{(1)} \bullet p^{(0)}
 1 \xrightarrow{u} A \xrightarrow{\text{update}} 1 \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{update}} 1 \xrightarrow{\text{lookup}} A \xrightarrow{p} A
Conclusion: pg^{(2)} \sim_{st} rs^{(0)} where rs^{(0)} = u. The result is u
```

Backward semantics

Given a program pg = c; return $(post) : \mathbb{1} \to Expr$, find an expression $pre : \mathbb{1} \to Expr$ such that $pg \sim_{st} \text{return}(pre)$. This means that c, post and pre satisfy:

$$post(c(s)) = pre(s)$$

$$S \xrightarrow{c^{(2)}} S \xrightarrow{} S$$

$$= \downarrow post^{(1)} \downarrow \qquad = \downarrow post^{(1)}$$

$$Expr \xrightarrow{id^{(0)}} Expr$$

This requires only the weak equation:

$$ext{update} ullet ext{lookup} \sim_{\mathit{st}} \mathit{id}_{A}$$

Method:

- $ightharpoonup \sim_{st}$ is used inductively, by substitution and pure replacement
- until finally ≡ is used for simplifying pure terms

When Expr = B this corresponds to a weakest precondition semantics (here with a restricted language for conditions)

Backward semantics: an example

The given program is
$$\ell := s(\ell); \ \ell := s(\ell); \ \text{return}(p(\ell))$$
 translated as:
$$pg^{(2)} = \text{lookup}^{(1)} \bullet s^{(0)} \bullet \text{update}^{(2)} \bullet \text{lookup}^{(1)} \bullet s^{(0)} \bullet \text{update}^{(2)} \bullet \text{lookup}^{(1)} \bullet p^{(0)}$$

$$1 \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{update}} 1 \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{update}} 1 \xrightarrow{\text{lookup}} A \xrightarrow{p} A$$

$$1 \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{update}} 1 \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{id}} A \xrightarrow{p} A$$

$$1 \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{id}} A \xrightarrow{p} A$$

$$1 \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{id}} A \xrightarrow{p} A$$

$$2 \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{id}} A \xrightarrow{p} A$$

$$2 \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A \xrightarrow{s} A \xrightarrow{\text{id}} A \xrightarrow{p} A$$

$$2 \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A$$

$$3 \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A \xrightarrow{\text{lookup}} A$$

$$3 \xrightarrow{\text{lookup}} A \xrightarrow{\text{$$

Outline

The language XS-IMP-EX

```
Syntax
Expressions:
    as in XS-IMP
Commands:
    c ::= skip | c; c | ℓ := a | throw | try(c)catch(c)
Programs:
    pg ::= c; return(e)

Restriction (easy to remove):
    only one exception name (thus, omitted)
```

Decorated logic for exceptions (only)

Operations and equations for exceptions (only)

In Set: a set of exceptions E with (here) $\mathbb{1} \cong E$, denoted $\star \leftrightarrow \odot$ Then Th_1 and Th_2 are generated from Th_0 by two operations:

$tag^{(1)}:\mathbb{1} o extstyle{0}$	$ ext{untag}^{(2)}: extstyle{0} ightarrow exttt{1}$
$\texttt{tag}: \mathbb{1} \to E$	$\mathtt{untag}: E \to \mathbb{1} + E$
$tag: \star \mapsto \textcircled{*}$	$\mathtt{untag}: \textcircled{\star} \mapsto \star$

one weak equation:

$$\begin{array}{c} \operatorname{tag} \bullet \operatorname{untag} \sim_{\operatorname{ex}} \operatorname{id}_{1} \\ \\ \operatorname{tag} \bullet \operatorname{untag} = \eta_{1} \\ \\ \star \mapsto \textcircled{*} \mapsto \star \operatorname{and} \overset{}{\circlearrowleft} \mapsto \textcircled{*} \mapsto \star \end{array}$$

and decorated rules...



Decorated logic for states and exceptions Duality is broken!

Functor
$$T(D(X)) = S \times X + S \times E$$

$$\mathcal{T}h_0 \xrightarrow{\subseteq} \mathcal{T}h_1 \xrightarrow{\subseteq} \mathcal{T}h_2 \xrightarrow{\subseteq} \mathcal{T}h_3 \xrightarrow{\subseteq} \mathcal{T}h_4$$

$$(X \to Y) \longmapsto (X \to Y) \longmapsto (X \to Y) \qquad \qquad \downarrow$$

$$(X \to Y) \longmapsto (X \to Y) \longmapsto (X \to Y) \qquad \qquad \downarrow$$

$$\downarrow \qquad \qquad \downarrow$$

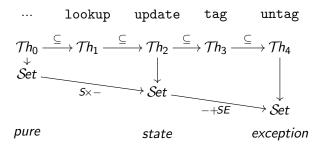
$$(S \times X \to S \times Y) \longmapsto (S \times X \to S \times Y)$$

$$\downarrow \qquad \qquad \downarrow$$

$$(S \times X + S \times E \to S \times Y + S \times E)$$

pure state exception

Operations for states and exceptions: summary



- $f^{(1)}$: may use the state
- $f^{(2)}$: may use and modify the state
- ► f⁽³⁾: may use and modify the state, may raise exceptions and must propagate exceptions
- ► f⁽⁴⁾: may use and modify the state, may raise exceptions and must propagate exceptions, may recover from exceptions

Decorated equations for states and exceptions

Weak equations

$$SX \longrightarrow \eta_{SX} \rightarrow SX + SE \xrightarrow{f_1} SY + SE \longrightarrow \varphi_{Y} + SE \longrightarrow Y + SE$$

$$\qquad \boxed{f_1 \sim_{\mathsf{ex}} f_2 : X \to Y}$$

interpreted as:
$$\eta_{SX} \bullet f_1 = \eta_{SX} \bullet f_2$$

$$g_1 \sim_{\mathsf{ex}} g_2 \implies f^{(2)} \bullet g_1 \bullet h \sim_{\mathsf{ex}} f^{(2)} \bullet g_2 \bullet h$$

interpreted as:
$$f_1 \bullet (\varepsilon_Y + SE) = f_2 \bullet (\varepsilon_Y + SE)$$

$$g_1 \sim_{st} g_2 \implies f \bullet g_1 \bullet h^{(0)} \sim_{st} f \bullet g_2 \bullet h^{(0)}$$

$$\qquad \qquad |f_1 \sim_{st,ex} f_2: X \to Y |$$

interpreted as:
$$\eta_{SX} \bullet f_1 \bullet (\varepsilon_Y + SE) = \eta_{SX} \bullet f_2 \bullet (\varepsilon_Y + SE)$$

 $g_1 \sim_{st.ex} g_2 \implies f^{(2)} \bullet g_1 \bullet h^{(0)} \sim_{st.ex} f^{(2)} \bullet g_2 \bullet h^{(0)}$

Equations for states and exceptions: summary

$$Th_0 \xrightarrow{\subseteq} Th_1 \xrightarrow{\subseteq} Th_2 \xrightarrow{\subseteq} Th_3 \xrightarrow{\subseteq} Th_4$$

$$\equiv --- \equiv --- \Longrightarrow ---$$

Translation

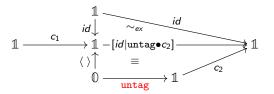
```
Expressions: |e \mapsto e^{(1)} : \mathbb{1} \to Expr
as for XS-IMP
Commands: c \mapsto c^{(3)} : 1 \rightarrow 1 (really (3), not (4))
   ▶ skip, c_1; c_2, \ell := a:
       as for XS-IMP
   ▶ throw \mapsto tag^{(3)} \bullet []_1^{(0)}
       pretends that the exception has type \mathbb{1}, instead of \mathbb{0}
   ▶ \operatorname{try}(c_1)\operatorname{catch}(c_2) \mapsto (\downarrow (c_1 \bullet [id_1|\operatorname{untag}^{(4)} \bullet c_2])^{(3)}
       (next slide)
```

Programs: $pg \mapsto pg^{(3)} : \mathbb{1} \to Expr$ (really (3), not (4)) as for XS-IMP

Translation of try-catch

$$\operatorname{try}(c_1) \operatorname{catch}(c_2) \mapsto (\downarrow (c_1 \bullet [id_1 \mid \operatorname{untag}^{(4)} \bullet c_2])^{(3)}$$

Uses: the decorated coproduct 1 = 1 + 0



and the "downcast" operator ↓

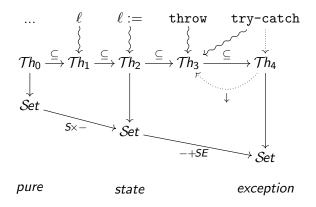
$$(\downarrow (f^{(4)}))^{(3)}$$
 is such that $f \sim_{ex} \downarrow f$

- $ightharpoonup \downarrow f$ is the same as f on non-exceptional arguments
- $ightharpoonup \downarrow f$ propagates exceptions while f may recover from exceptions

Rules for
$$\downarrow$$
 include: \downarrow $(f_1) \equiv \downarrow$ $(f_2) \iff f_1 \sim_{ex} f_2$



Translation of XS-IMP-EX: summary



Backward semantics

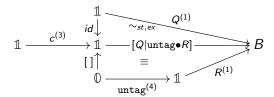
Predicate transformer semantics [Claude Marché, MPRI 2012]

Hoare triples: $\{P\}$ c $\{Q|R\}$ is valid if:

if c is executed in a state satisfying P then:

- if c terminates normally in a state s' then s' satisfies Q
- if c terminates abruptly in a state s' then s' satisfies R

This means that $P \sim_{st,ex} c \bullet [Q \mid \operatorname{untag} \bullet R]$



Outline

The language IMP-EX: syntax, revisited

Expressions:

```
a ::= 0 | 1 | -1 | 2 | -2 | ... | \ell | s(a) | p(a)

b ::= true | false | \neg b | a = 0 | a > 0

e ::= a | b
```

Commands:

$$c ::= \begin{cases} \text{skip} \mid c; c \mid \ell := a \mid \\ \text{throw} \mid \text{try}(c) \text{catch}(c) \mid \\ \text{if}(b) \text{then}(c) \text{else}(c) \mid \text{repeat}(c) \end{cases}$$

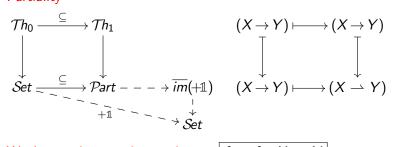
Programs:

$$pg ::= c; return(e)$$

- As before: only one location, no binary operation on expressions, only one exception name
- ▶ In addition: repeat(c) "instead of" while(b)do(c)

Decorated logic for non-termination

Partiality



Weak equations are inequations $f_1 \succcurlyeq f_2 : X \rightarrow Y$

$$f_1 \succcurlyeq f_2 : X \to Y$$

interpreted as: $f_1 \ge f_2 : X \rightharpoonup Y$ (as partial functions)

Part with > is a 2-category

> satisfies replacement and substitution:

$$g_1 \succcurlyeq g_2 \implies f \bullet g_1 \bullet h \succcurlyeq f \bullet g_2 \bullet h$$

Operations and equations for non-termination

 Th_1 is generated from Th_0 by one operation constructor:

$$\mathsf{loop}(c)^{(1)}: X \to X \text{ for each } c^{(1)}: X \to X$$

 $\mathsf{loop}(c): X \rightharpoonup X \text{ is the least fixed point of } f \mapsto c \bullet f$

one strong equation:

$$loop(c) \equiv c \bullet loop(c)$$

and decorated rules, including:

$$f \equiv c \bullet f \implies f \succcurlyeq loop(c)$$

A weak congruence

The "weakest" congruence for states, exceptions and non-termination is $\leq_{st,ex}$. For instance:

$$f \preccurlyeq_{\mathsf{st},\mathsf{ex}} u^{(0)} : X \to Y$$

is a concise way to express the following:

 $f: S \times X + S \times E \longrightarrow S \times Y + S \times E$ and $u: X \to Y$ are such that: if f(s,x) is defined, then it returns (s', u(x)) for some s'.

This is the kind of relation required between a "program" f and its "result" u



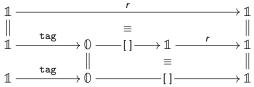
Translation

Translation is obvious:

▶ repeat $(c) \mapsto loop(c)$

Example $repeat(throw) \equiv throw$

repeat(throw) is translated as $r = loop(tag \bullet [\]_{1})$



because r (like all commands) propagates exceptions

Decorated logic for conditions

Weak equations are conditional equations

$$f_1 \sim f_2: X \to Y \text{ if } b$$

where \sim is any of the previous (strong or weak) congruence and b is a boolean expression.

For replacement, conditional \sim has the same properties as \sim .

 $\mathcal{T}h_1$ is generated from $\mathcal{T}h_0$ by two operation constructors ("conditional non-determinism"):

choose
$$(c_1, c_2)^{(1)}: X \to Y$$
 for each $c_1^{(1)}, c_2^{(1)}: X \to Y$

$$\downarrow_b (c)^{(0)}: X \to Y \text{ for each } c^{(1)}: X \to Y$$

$$[c_1|c_2]: X + X \to Y$$

$$\downarrow_b ([c_1|c_2]) = b \bullet [c_1|c_2]: X \to Y$$

Decorated logic for IMP-EX

Combine the decorated logics for:

- states
- exceptions
- non-termination
- and "conditional non-determinism"

by composing the corresponding functors and extending the corresponding weak congruences

Outline

Conclusion

Remark.

Effects as functors, with their weak congruences, can be seen as a kind of generalization of 2-categories, with decorated categorical notions as a generalization of lax categorical notions.

To do...

- "work in progress":
 - which is the best order for composing the effects?
- ▶ Define while loops by:
 while(b)do(c) =
 try(repeat(if(b)then(c)else(throw))) catch(skip)
 Prove that indeed such a while loop is the least fixed point
 of f → if(b)then(c;f)else(skip)
- Complete the implementation in Coq
- ► Towards richer languages (C, C++, Java,...)