

Shoggoth: A Formalised Logic for Strategic Rewriting

Xueying Qin¹

Rob van Glabbeek¹, Peter Höfner², Liam O'Connor¹, Michel Steuwer¹

¹ The University of Edinburgh ² Australian National University

July 17, 2023

0. Introduction

1. Semantics of Our Strategic Rewriting Language

2. Location Based Weakest Precondition Calculus

3. Conclusion and Future Work

Introduction

Strategic Rewriting Languages

The core calculus of strategic rewriting languages like ELEVATE[1], Stratego[4] and Strafunski[2] has a simple construct: atomic strategies and strategy combinators.

Atomic strategy

An atomic strategy is a *rewrite rule*:

 $add_{com}: a + b \rightsquigarrow b + a$

 $mapFusion : map f (map g xs) \rightsquigarrow map (f \circ g) xs$

Composed strategy

add_{com} ; add_{id} add_{com} <+ mult_{com} repeat(mapFusion)

Strategy combinator

Strategy combinators compose strategies together and controls the application of atomic strategies:

s; t Sequential composition, apply s then t

 $s \prec t$ Left choice, if fail to apply s then t

repeat(s) Keep applying s until inapplicable

- Strategic rewriting languages provide programmers with combinators and generic traversals that allow them to: 1) control the application of rewrite rules; 2) reuse rewrite rules
- Many application areas: for writing program optimisation (ELEVATE), writing interpreter/compiler for DSLs (Spoofax/Stratego) etc.

Strategies can go wrong

- **Result in error** an atomic strategy is not defined for certain expressions or strategies are not well composed
- Do not terminate for example : repeat(SKIP)
- Do not rewrite an expression into desired form

Therefore, we would like a formal understanding of these strategies and a framework that allows us to formally reason about these strategies.

- We design, formalise and mechanise using Isabelle **denotational semantics** of a strategic rewriting language, we prove computational soundness and adequacy between the denotational semantics and an existing big-step operational semantics.
- We design, formalise and mechanise using Isabelle the **weakest precondition calculus** for the strategic rewriting language. We prove its soundness w.r.t. the denotational semantics.
- We demonstrate how to use the weakest precondition calculus to prove properties of stratetigic rewriting.

Semantics of Our Strategic Rewriting Language

System S

System S [5] is the core calculus of the strategic rewriting languages we study that contains atomic strategies (rewrite rules) and strategy combinators which compose strategies and perform expression traversals.

Expression

The expressions being rewritten by strategies are in the form of:

 $Expressions(\mathbb{E}) \quad e := Id \mid Fun e e \mid App e e$

System S - Syntax (1)

```
Strategy
          Strategy(S) s := SKIP (Always success) | ABORT (Always fail)
                            atomic (Atomic strategy)
                            X (Variable)
                            | s_1; s_2 (Sequential composition)
                            | s_1 < + s_2 (Left choice)
                            | s_1 <+> t_2 (Nondeterministic choice)
                            | one(s) (Apply s to one child)
                            some(s) (Apply s to as many children as possible)
                            all(s) (Apply s to all children)
                            \mid \mu X.s (Fixed point operator)
```

We define a denotational semantics. Since our strategies include a fixed point operator, we need to ensure the denotational semantics of strategies is monotone. To do so, we provide a *Plotkin powerdomain* with *Egli-Milner ordering* where divergence (*div*) is the bottom element, and point-wise lift it to the domain (which is a CPO) for defining our denotational semantics.

The Plotkin powerdomain

$$\mathfrak{D}_{p} = \mathcal{P}_{\neg \emptyset}(\mathbb{E} \cup \{err\} \cup \{div\})$$

Egli-Milner ordering

$$a \leq b \iff (\forall x \in a. \exists y \in b. x \leq y) \land (\forall y \in b. \exists x \in a. x \leq y)$$

$$\mathfrak{D} = \mathbb{E} \to \mathfrak{D}_p$$

Denotational Semantics by Examples - Skip, Abort and Atomic



Denotational Semantics by Examples - Sequential Composition



Denotational Semantics by Examples - Left Choice



Denotational Semantics by Examples - Nondeterministic Choice

Example $add_{id}: 0 + a \rightsquigarrow a$ add_{com} : $a + b \rightsquigarrow b + a$ $0 + 3 \xrightarrow{add_{id} <+> add_{com}} 3$ $0 + 3 \xrightarrow{add_{id} <+> add_{com}} 3 + 0$ **Semantics** $(\langle + \rangle_c) : \mathfrak{D} \to \mathfrak{D} \to \mathfrak{D}$ $(s < +>_{s} t)(e) = \{x \mid x \in s(e) \cap \mathbb{E}\} \cup \{div \mid div \in s(e)\}$ $\cup \{ v \mid v \in t(e) \cap \mathbb{E} \} \cup \{ div \mid div \in t(e) \}$ \cup {*err* | *err* \in *s*(*e*) \cap *t*(*e*)} $[[s <+> t]]\xi = [[s]]\xi <+>_{s} [[t]]\xi$

Denotational Semantics by Examples - One (0)



Denotational Semantics by Examples - One (1)



Semantics

$$(one_{s}): \mathfrak{D} \to \mathfrak{D}$$

$$one_{s}(s)(e) = \{ \begin{array}{l} n \\ x \\ e_{2} \end{array} | e = \begin{array}{l} n \\ e_{1} \\ e_{2} \end{array} \land x \in s(e_{1}) \cap \mathbb{E} \}$$

$$\cup \{ \begin{array}{l} n \\ e_{1} \\ x \end{array} | e = \begin{array}{l} n \\ e_{1} \\ e_{2} \end{array} \land x \in s(e_{2}) \cap \mathbb{E} \}$$

$$\cup \{ div | e = \begin{array}{l} n \\ e_{1} \\ e_{2} \end{array} \land div \in s(e_{1}) \cup s(e_{2}) \}$$

$$\cup \{ err | e = ld \lor (e = \begin{array}{l} n \\ e_{1} \\ e_{2} \end{array} \land err \in s(e_{1}) \cap s(e_{2})) \}$$

$$[[one(s)]]\xi = one_{s}([[s]]]\xi)$$

Denotational Semantics by Examples - Fixed Point Operator(0)





Denotational Semantics by Examples - Fixed Point Operator(2)

Semantics
$$\llbracket X \rrbracket \xi = \xi X$$
 $\llbracket \mu X.s \rrbracket \xi = \mu X.\llbracket s \rrbracket (\xi [X \mapsto X])$

Big-Step operational semantics

The existing big-step operational semantics does not model divergence. In general, it takes the form of:

$$e \xrightarrow{s} e'$$

where e' can be either an expression or an error.

Correspondence to Existing Big-Step Operational Semantics(1)





Equivalence?

We believe if we add divergence to the big-step operational semantics, we would be able to conclude that these two semantics are equivalent.

Location Based Weakest Precondition Calculus

Strategies Can Go Wrong



Weakest precondition

Given a program S and a postcondition Q, a weakest precondition is a predicate P_w such that for any precondition P:

 $\{P\}S\{Q\} \Leftrightarrow (P \Longrightarrow P_w)$

The challenge in design

We have strategies that can traverse the syntax tree and control at what location in the syntax tree to apply a strategy — we need a notion of "location" in our formulas.

Our solution

We introduce the location as a path in the syntax tree into our formulas.





Helper partial functions

$$\begin{aligned} update : \mathbb{L} \to \mathbb{E} \to \mathcal{P}_{\neg \emptyset}(\mathbb{E} \cup \{err\} \cup \{div\}) \to \mathcal{P}_{\neg \emptyset}(\mathbb{E} \cup \{err\} \cup \{div\}) \\ update(\epsilon, e, xs) &= xs \end{aligned}$$

$$\begin{aligned} update(\ell \triangleleft l, \bigwedge_{e_1 e_2}^n, xs) &= \{\bigwedge_{x e_2}^n | x \in update(l, e_1, xs) \cap \mathbb{E}\} \cup (xs \cap \{err, div\}) \\ update(\ell \triangleleft l, \bigwedge_{e_1 e_2}^n, xs) &= \{\bigwedge_{e_1 x}^n | x \in update(l, e_2, xs) \cap \mathbb{E}\} \cup (xs \cap \{err, div\}) \\ lookup : \mathbb{L} \to \mathbb{E} \to \mathbb{E} \\ lookup(\epsilon, e) &= e \\ lookup(\ell \triangleleft l, \bigwedge_{e_1 e_2}^n) &= lookup(l, e_1) \\ lookup(\ell \triangleleft l, \bigwedge_{e_1 e_2}^n) &= lookup(l, e_2) \end{aligned}$$

$$\begin{array}{c} \hline \textbf{Definition} \\ wp_{s@l}\zeta(P) = Q & (wp_{s} : \mathbb{L} \to \Gamma_{L} \to \mathcal{P}(\mathbb{E}) \to \mathcal{P}(\mathbb{E})) \end{array}$$

If for a set of expressions described by the postcondition P, a non-empty set of expressions described by the precondition Q can be found,

then each expression in the precondition set will be successfully transformed into an expression in the postcondition set, by applying strategy s at location I in the logic environment ζ .

$wp_{s@l}^{\uparrow}\zeta(P) = Q \qquad (wp_s^{\uparrow} : \mathbb{L} \to \Gamma_L \to \mathcal{P}(\mathbb{E}) \to \mathcal{P}(\mathbb{E}))$

If for a set of expressions described by the postcondition P, a non-empty set of expressions described by the precondition Q can be found,

then each expression in the precondition set will be successfully transformed into an expression in the postcondition set or result in error, by applying strategy s at location I in the logic environment ζ .

It is used for defining the weakest precondition for total correctness.

Definition of Unsuccessful Execution



A Strategy Is Not Well-Composed



A Strategy Is Not Well-Composed for Desired Output



Wp of non-deterministic choice

$$wp_{s<+>t@l}\zeta(P) = (wp_{t@l}^{\uparrow}\zeta(P) \cap wp_{s@l}\zeta(P)) \cup (wp_{s@l}^{\uparrow}\zeta(P) \cap wp_{t@l}\zeta(P))$$
$$wp_{s<+>t@l}^{\uparrow}\zeta(P) = wp_{s@l}^{\uparrow}\zeta(P) \cap wp_{t@l}^{\uparrow}\zeta(P)$$

Checking invalid composition for *P*

 $wp_{mult_{com} <+>mult_{zero}; mult_{com}@\epsilon} \zeta(\{e \mid e = App (*) \ 0 \ m\}) = \emptyset$ (Bad!)

Invalid Input

Example $add_{com}: a + b \rightsquigarrow b + a \quad mult_{com}: a * b \rightsquigarrow b * a$ $3 * 6 \xrightarrow{mult_{com} <+ add_{com}; mult_{com}} 3 * 6 3 + 6 \xrightarrow{mult_{com} <+ add_{com}; mult_{com}} err$ Wp of left choice $wp_{s<+t@l}\zeta(P) = wp_{s@l}\zeta(P) \cup (wp_{c@l}^{\uparrow}\zeta(P) \cap wp_{t@l}\zeta(P))$ $wp^{\uparrow}_{\mathsf{r} \sim \mathsf{I} \mathsf{f} \circ \mathsf{f}}(P) = wp_{\mathsf{s} \otimes \mathsf{I}}\zeta(P) \cup (wp^{\uparrow}_{\mathsf{s} \otimes \mathsf{I}}\zeta(P) \cap wp^{\uparrow}_{\mathsf{r} \otimes \mathsf{I}}\zeta(P))$ **Checking Invalid Input** $wp_{mult_{com} < +add_{com}; mult_{com} @ \epsilon \zeta(U)} = \{e \mid e = App(*) \ a \ b\}$ $3 * 6 \in \{e \mid e = App(*) \ a \ b\}$ (Good!) $3 + 6 \notin \{e \mid e = Abb(*) \mid a \mid b\}$ (Bad!)

Invalid Input for A Desired Output

Example

$$mult_{zero}: 0 * a \rightsquigarrow 0$$
 $mult_{com}: a * b \rightsquigarrow b * a$
 $3 * 4$
 $mult_{zero}: mult_{com} < +mult_{zero}; mult_{com} < +mult_{zero}; mult_{com} < +mult_{com} < +m$

Checking invalid input for P

 $wp_{mult_{com} < +mult_{zero}; mult_{com} @ \epsilon \zeta(\{e \mid e = App (*) \ 0 \ m\}) = \{e \mid e = App (*) \ 0 \ m\}$ 3 * 4 $\notin \{e \mid e = App (*) \ 0 \ m\}$ (Bad!)

Detect Divergence (0)

Example

The given strategy will diverge, i.e., will not lead to any successful execution.

repeat(SKIP) Bad?

Wp of fixed point operator

$$wp_{\mu X.s@l}\zeta(P) = [\mathsf{LFP}\,\mathscr{X}:\Delta](l)(P) \qquad wp_{\mu X.s@l}^{\uparrow}\zeta(P) = [\mathsf{LFP}\,\mathscr{Y}:\Delta](l)(P)$$

Where:

$$\Delta = \begin{cases} \mathcal{X}(l)(P) = wp_{s@l}\zeta[(X, \cdot) \mapsto \mathcal{X}, (X, \uparrow) \mapsto \mathcal{Y}](P) \\ \mathcal{Y}(l)(P) = wp_{s@l}^{\uparrow}\zeta[(X, \cdot) \mapsto \mathcal{X}, (X, \uparrow) \mapsto \mathcal{Y}](P) \\ wp_{X@l}\zeta(P) = \zeta(X, \cdot)(l)(P) \text{ (where } \zeta(X, \cdot) \text{ def.)} \\ wp_{X@l}^{\uparrow}\zeta(P) = \zeta(X, \uparrow)(l)(P) \text{ (where } \zeta(X, \uparrow) \text{ def.)} \end{cases}$$



Checking divergence

 $wp_{repeat(SKIP)@\epsilon}\zeta(\mathbb{U}) = \emptyset$ Bad!

We take divergence seriously — a strategy which can either diverge or successfully execute leads to diverge.



Reasoning About Traversals - One (0)





Reasoning About Traversals - Top Down (0)





Soundness Theorems

 $\forall X | P. \zeta(X, \cdot)(I)(P) = \{e \mid update(I, e, (\xi(X)(lookup(I, e)))) \subseteq P\} \\ \land \zeta(X, \uparrow)(I)(P) = \{e \mid update(I, e, (\xi(X)(lookup(I, e)))) \subseteq P \cup \{err\}\}$

 $wp_{s@l}\zeta(P) = \{e \mid update(l, e, \llbracket s \rrbracket \xi(lookup(l, e))) \subseteq P\}$

(Weakest Precondition for Total Correctness)

 $\forall X \, I \, P. \, \zeta(X, \cdot)(I)(P) = \{e \mid update(I, e, (\xi(X)(lookup(I, e)))) \subseteq P\} \\ \land \zeta(X, \uparrow)(I)(P) = \{e \mid update(I, e, (\xi(X)(lookup(I, e)))) \subseteq P \cup \{err\}\}$

 $wp_{s@l}^{\uparrow}\zeta(P) = \{e \mid update(l, e, \llbracket s \rrbracket\xi(lookup(l, e))) \subseteq P \cup \{err\}\}$

(Weakest May Error Precondition)

Conclusion and Furure Work

Conclusion

- We present the formalised denotational semantics of System S and demonstrate the correspondence (potentially equivalence) between the denotational semantics and big-step operational semantics.
- We present the formalised weakest precondition calculus for System S and demostrated the usage of the weakest precondition calculus for reasoning about the execution of strategies.

Future Work

- Rewriting expressions represented in other forms such as graphs?
- Using weakest precondition calculus for automatic reasoning about the execution of strategies?

We were on the track ahead as the nightmare plastic column of foetid black iridescence oozed tightly onward through its fifteen-foot sinus; gathering unholy speed and driving before it a spiral, re-thickening cloud of the pallid abyss-vapour. It was a terrible, indescribable thing vaster than any subway train—a shapeless congeries of protoplasmic bubbles, faintly self-luminous, and with myriads of temporary eyes forming and unforming as pustles of greenish light all over the tunnel-filling front that bore down upon us, crushing the frantic penguins and slithering over the glistening floor that it and its kind had swept so evilly free of all litter. Still came that eldritch, mocking cry — "Tekeli-li!" And at last we remembered that the daemoniac shoggoths — given life, thought, and plastic organ patterns solely by the Old Ones, and having no language save that which the dot-groups expressed — had likewise no voice save the imitated accents of their bygone masters. —H. P. Lovecraft "From the Mountains of Madness"[3]

Thank you

Xueying Qin [xueying.qin@ed.ac.uk]

Bastian Hagedorn, Johannes Lenfers, Thomas Kœhler, Xueying Qin, Sergei Gorlatch, and Michel Steuwer.

Achieving high-performance the functional way: A functional pearl on expressing high-performance optimizations as rewrite strategies. *Proc. ACM Program. Lang.*, 4(ICFP), aug 2020.

Markus Kaiser and Ralf Lämmel.

An isabelle/hol-based model of stratego-like traversal strategies. In Proceedings of the 11th ACM SIGPLAN Conference on Principles and Practice of Declarative Programming, PPDP '09, page 93–104, New York, NY, USA, 2009. Association for Computing Machinery.

Howard P. Lovecraft.

At the mountains of madness. 1931.

Eelco Visser.

Stratego: A language for program transformation based on rewriting strategies system description of stratego 0.5.

In Aart Middeldorp, editor, *Rewriting Techniques and Applications*, pages 357–361, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.

Eelco Visser and Zine el Abidine Benaissa.

A core language for rewriting.

Electronic Notes in Theoretical Computer Science, 15:422–441, 1998. International Workshop on Rewriting Logic and its Applications.