

Going for a miniKanren implementation.

JONATHAN FRECH, Universität Tübingen, Germany

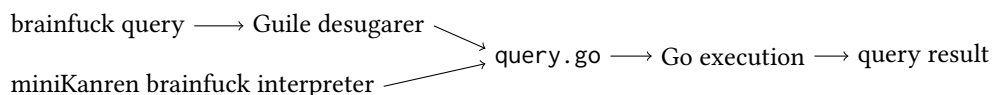
As a final project for the seminar “Relational Programming with miniKanren” held in the winter term of 2020/21, I wrote a miniKanren[2] interpreter in Go[5] and used it to explore the space of all brainfuck[4] programs. Jonathan Frech <info@jfrech.com>, 2021-03-31

Additional Key Words and Phrases: relational programming, miniKanren, Go, concurrency, brainfuck

1 INTRODUCTION

In contrast to classical imperative thought and the uniform data flow associated with it, logic programming explores the space of all possible values, rejecting those that do not meet the desired properties. Amazingly, what at first glance might sound like a purely existential axiomatically guaranteed comprehension is in certain settings computable on physically present hardware.

One particularly minimal specification of a logic programming language is miniKanren, as detailed in *The Reasoned Schemer*[2]. Following this book, I implemented a Go interpreter for the miniKanren language. Furthermore, I kneaded brainfuck programs through a brainfuck interpreter I wrote in miniKanren to solve various problems in \mathbb{Z}_{2^8} . Since the homoiconic beauty of Scheme is not present in Go, I wrote a GNU Guile desugarer to translate miniKanren programs represented as S-expressions to Go source code:



1.1 Installation

In the vein of embedding, the above mentioned desugarer produces a Go snippet; miniKanren programs are meant to be desugared, embedded and run as standalone Go programs:

```
% mkdir -p ~/src && cd ~/src
% curl -fsSL https://papers.jfrech.com/\
> 2021-03-31_jonathan-frech_going-for-a-minikanren-implementation/gomik.tar -o gomik.tar
% sha512sum gomik.tar | fold -w64 | sed 2q
ec40966528c616a8a2a47029ce98815cc973dc38b606a5479a5628af55956898
abedb08b50332ea4988eabaacb8e247a6ebed5690501d48b83568d3ffe4bf6b0
% tar -vxf gomik.tar && cd gomik
% ./test.sh
```

Instead of downloading the tar ball `gomik.tar` off the world-wide web and with access to this paper’s digital form, one may also choose to *locally download* `gomik.tar` *through this very base64-encoded data URI*.

2 IMPLEMENTATION

My miniKanren implementation is a direct port of Friedman et al.’s Scheme implementation, with at times a more imperative approach. The major deviation is the choice of channels to implement goals and mutable maps to implement substitutions:

```
type Substitution = map[Variable]Term
type Stream = chan Substitution
type Goal = func(Substitution) Stream
```

Utilizing Go’s higher-order capabilities, goals are combined as points:

```
var disj2 func(Goal, Goal) Goal
var conj2 func(Goal, Goal) Goal
```

This implementation neither supplies an embedding framework nor is meant to be executed as a standalone binary. By using a desugarer (`desugar.sh`, in turn running `desugarer.scm`), input miniKanren programs are transformed into Go snippets which are compiled together with the implementation to form an executable query. As a corollary, each query requires an invocation of the Go compiler. Source splicing is implemented in `run.sh`, which allows for a seamless invocation of miniKanren queries.

2.1 Symbolic purity

I chose not to support non-symbolic atoms found in many Scheme dialects (i. e. Booleans, integers, characters, strings, etc.) as I feel they distract from the profound and precisely meaningless beauty of S-expressions. Exacerbated by miniKanren’s inability to reason about these other atom’s structure – making them essentially equivalent in opaqueness to pure symbols – the only fundamental types I chose to implement are `Null`, `Symbol`, `Pair` and `Variable`.

2.2 Reification

Friedman et al. specify miniKanren’s semantics only loosely, stating “[...] the order of the values does *not* matter.”[2, p. 14] yet later reasoning about precisely how ordered query answers came to be. Due to the concurrent nature of channels and goroutines, I cannot guarantee for any order necessitating to sort the output for testing.

Reification semantics, however, are kept the same – with their implementation being imperative. All internal S-expressions are pretty-printed without shortening them.

2.3 A technicality regarding laziness

Querying the countably infinite space of all S-expressions requires special attention: Friedman et al. directly embed laziness into their stream data structure, defining it as a sum type of a finite list and a stream generating function [2, p. 153]. This way, postponing evaluation of the entire stream becomes possibly by piecewise trickling.

In my implementation, I opted to use Go’s native concurrency model and thus chose to implement streams as channels of substitutions. However, since Go is as a C derivative fundamentally imperative, an unmodified goal combinator evaluates both of its arguments before analyzing its first, making implementing recursive relations impossible. To allow the declaration of recursive relations, special versions of `disj2` and `conj2` are crafted, called `disj2Lazy` and `conj2Lazy`: taking as their second argument a value of the type `func() Goal`, thus facilitating right-laziness.

A more fundamental problem is the intrinsically unidirectional nature of streams. Whereas laziness requires a communication between the data’s user and its generator, no one listens when one yells that there is no need for a value at the receiving end of a channel. Thus, when spawning a goroutine to fill a channel, it will eagerly traverse the search space, looking for an answer even if there is no one who asks for such an answer. The net effect of this dissociation between data generation and usage is a buildup of unnecessarily working goroutines, wasting resources.

As a consequence, `conde`’s reordering property is even weaker than in Friedman et al.’s Scheme implementation: *a conde line may only possess one recursive relation which ought to appear as its last line.*

Unfortunately, the above culminates in an inefficient approach to laziness making my implementation inferior to *The Reasoned Schemer*’s Scheme implementation. See 3.4 where I detail how I ran the miniKanren queries presented in this paper.

2.4 Testing

I am no proponent of mindlessly and meticulously testing the smallest of implementation details; a reasonable encapsulation ethic paired with mindful program engineering leaves little room for local incorrect program behavior. A far more powerful approach is that of wholistic testing: when running the testing shell script `test.sh`, a set of fully-fledged miniKanren queries is invoked and compared to a statically reasoned about set of expected answers. Since this testing scenario is not trivialized by locally present assumptions, a high testing coverage is guaranteed, enstilling confidence in the implementation’s correctness.

3 APPLICATION

One of the truly extraordinary possibilities logic programming offers is the ability to inquiry sub-implementations of other languages. Herein, I will use miniKanren to reason about `brainfuck`[4] – Urban Müller’s minimalistic reimagination of a Turing machine.

3.1 Defining brainfuck

`brainfuck` programs consist of a finite string composed of eight instructions `+-<>[] . ,`, modify a writing head on a two-way infinitely long tape of bytes. Imperative in nature, a natural mapping from `brainfuck` source to C source exists, where `static char p[∞]`; `p += ∞/2`, i.e. `p` points somewhere in the middle of the infinitely long¹, initially zeroed tape:

```
+ -----> ++*p;
- -----> --*p;
> -----> ++p;
< -----> --p;
[ -----> while (*p) {
] -----> }
. -----> fputc(*p, stdout);
, -----> { int c = fgetc(stdin);
           *p = c == EOF ? 0 : c; }
```

As an example, `,[->+++<]>` multiplies the inputted byte by three and outputs it.

¹Note that `char[∞]` and `p += ∞/2` are an abuse of notation justified by the theoretic nature of the proposed C pseudo code; any real-world `brainfuck` implementation necessarily has to accept restrictions of finiteness.

Going for a miniKanren implementation.

3.2 Encoding brainfuck

To explore the space of all brainfuck programs, the space of all strings of eight characters $\prod_{n \geq 0} \mathbb{Z}_8^n$ is too large, as a brainfuck program's brackets are always balanced. A more apt space to query is a subspace of the space of all S-expressions isomorphic to the space of all valid brainfuck programs in which `+-<> . ,` are unique symbols $\{\text{INC, SUB, LFT, RGT, PUT, GET}\}$ and any loop of the form `[★]` is a pair `(LOP . ★)`.

A beneficial consequence of this encoding is the nonnecessity to store a program counter: a loop may or may not replicate itself based on the tape state, propagating itself until terminating. See 3.3 for further detail on the interpreter's implementation.

3.3 Interpreting brainfuck

I wrote my brainfuck interpreter in a very direct functional-to-relational translatory way. Since I am writing vanilla miniKanren without any extensions, e. g. support for finite domains, I enumerate the complete complement of `B00` in \mathbb{Z}_2^8 with a 255 cases long `disj` expression to facilitate disjoint branching encountered in brainfuck loops.

For `+` and `-`, a 256 elements large lookup table for `+` is hardcoded and `-` defined as its inverse.

The brainfuck program to be interpreted is represented using one list, with executed instructions getting popped off and loops potentially pushing new instructions to the front. Since the tape head may move in either direction, the tape is represented using two lists, each representing one side. Input and output are each represented using a single list.

See `miks/bf.mik` or equivalently 4.1 for the full implementation.

3.4 Running queries

Although I am confident in my implementation's correctness, albeit not its resourcefulness nor speed, querying benefits from an interactive execution approach; one for which my implementation is outclassed by a Scheme REPL. Thus, the following queries where performed using the implementation provided by the *The Reasoned Schemer*[3], evaluated using GNU Guile 2.2.6[1], called through `mik` which can be installed using the provided install script which fetches the miniKanren implementation from GitHub and installs `/usr/bin/mik` with auxiliary files placed at `/usr/bin/.mik`:

```
# miks/install-the-the-reasoned-schemer-\
> second-edition-minikanren-implementation.sh
```

One may run the queries shown below using `./run.sh mik/bf.mik.q`, where one appends the query to the brainfuck interpreter implementation. However, as previously described in 2.3, some goroutines unfortunately may overflow the stack, rendering querying of large problems nearly impossible.

3.5 Querying brainfuck

3.5.1 *brainfuck interpreter correctness.* A basic query is forward-facing evaluation. The brainfuck program `++++[>+++<-]>` computes $4 \cdot 3 = 12 = 0x0c$ which is represented using the symbol `B0c`:

```
% mik miks/bf.mik
> (run* (o) (bf '() '(INC INC INC INC (LOP . (RGT INC INC INC LFT DEC)) RGT PUT) o))
$1 = (((B0c)))
```

Since an exhaustive `run*` has finished, this output is unique – as expected.

3.5.2 *Six sixes.* Unlike a classical interpreter, however, one is not restricted to only ask for the *result* of a computation and can query what inputs *may lead to* a given output. Since `bf` is a ternary relation with its center argument being brainfuck programs, one can generate brainfuck programs exhibiting a specified behavior by underspecifying the list of instructions, e. g. searching for six brainfuck programs which evaluate to 6:

```
% mik miks/bf.mik
> (run 6 (p) (bf '() p '(B06)))
$1 = (((INC INC INC INC INC INC PUT)) ((INC INC INC INC INC INC PUT INC))
((INC INC INC INC INC INC PUT DEC)) ((INC INC INC INC INC INC INC DEC PUT))
((INC INC INC INC INC INC DEC INC PUT)) ((INC INC INC INC INC DEC INC INC PUT)))
```

With an infinite number of brainfuck programs which evaluate to 6, `run*` is too ambitious a goal. Annotated with standard brainfuck notation, the above six programs are

```
(((INC INC INC INC INC INC PUT)) ; +++++.
((INC INC INC INC INC INC PUT INC)) ; +++++.+
((INC INC INC INC INC INC PUT DEC)) ; +++++.-
((INC INC INC INC INC INC INC DEC PUT)) ; ++++++-.
((INC INC INC INC INC INC DEC INC PUT)) ; ++++++-.
((INC INC INC INC INC DEC INC INC PUT)) ; +++++-++.
```

and indeed all evaluate to 6.

3.5.3 *Kneading subtraction.* Querying for unrestricted programs is in most cases a vast task. More manageably, one can query for holes in a carefully crafted program specifying a problem. As an example, to calculate $8 - 5$, one may ask how to fill the brainfuck program `+++++++*` to output 5. Of course, replacing `*` with `>+++++` does not answer the original question, yet more immediate solutions do:

```
% mik miks/bf.mik
> (run 4 (p) (bf `() `(INC INC INC INC INC INC INC INC . ,p) '(B05)))
$1 = (((DEC DEC DEC PUT)) ((DEC DEC DEC PUT INC))
((RGT INC INC INC INC INC PUT)) ((DEC DEC DEC PUT DEC)))
```

From all but the third possible hole replacements one can count the leading DEC instructions and deduce the answer to the original question to be 3.

3.5.4 *Solving multidimensional linear equations.* Let $f : \mathbb{Z}_{28}^2 \rightarrow \mathbb{Z}_{28}$, $(x, y) \mapsto 2 \cdot x - 3 \cdot y$ and define $V_f := f^{-1}(\{0\})$ as its variety. To calculate V_f , one may implement f in brainfuck via `, [->+<], [->---<]>`. and query for all inputs which evaluate to zero:

```
% mik miks/bf.mik
> (run 5 (x y) (bf `(,x ,y) `(GET (LOP . (DEC RGT INC INC LFT))
... GET (LOP . (DEC RGT DEC DEC DEC LFT)) RGT PUT) '(B00)))
$1 = ((B00 B00) (B03 B02) (B06 B04) (B01 B56) (B09 B06))
```

Where the second to last solution $(0x01, 0x56) = (1, 86)$ uncovers the modular nature of \mathbb{Z}_{28} , as

$$f(1, 86) = 2 \cdot 1 - 3 \cdot 86 = 2 - 258 = -256 \equiv 0.$$

3.5.5 *Kneading division.* In the same vein, one may also invert multiplication. That is, to calculate $12/3$ querying `, [>+<-]>`:

```
% mik miks/bf.mik
> (run* (i) (bf `(,i) '(GET (LOP . (RGT INC INC INC LFT DEC)) RGT PUT) '(B0c)))
$1 = ((B04))
```

Since it is a wholistic query, this *proves* that $3 \cdot x = 12$ has exactly one solution in \mathbb{Z}_{28} .

The same approach can be used to prove $\#\{x \in \mathbb{Z}_{28} : 4 \cdot x = 12\} = 4$:

```
% mik miks/bf.mik
> (run* (i) (bf `(,i) '(GET (LOP . (RGT INC INC INC INC LFT DEC)) RGT PUT) '(B0c)))
$1 = ((B03) (B43) (B83) (Bc3))
```

3.5.6 *Boolean negation.* With a starting fragment `>, [★1]★2` and the specification $\{0 \mapsto 1, 1 \mapsto 0\}$, miniKanren can fill both holes to satisfy the specification, implementing Boolean negation:

```
% mik miks/bf.mik
> (run 1 (p q1 q2) (== p `(INC RGT GET (LOP . ,q2) . ,q1))
... (bf '(B00) p '(B01)) (bf '(B01) p '(B00)))
$1 = (((INC RGT GET (LOP LFT) LFT PUT) (LFT PUT) (LFT)))
```

Replacing both holes $\star_1 := <$. and $\star_2 := <$, one arrives at `>, [<]<`. which indeed implements a not gate.

Requesting four hole substitutions, further solutions `>, [->]<`. as well as `>, [+<]<`. and `>, [>>]<`. are found which also implement not gates yet in a different manner.

```
% mik miks/bf.mik
> (run 4 (p q1 q2) (== p `(INC RGT GET (LOP . ,q2) . ,q1))
... (bf '(B00) p '(B01)) (bf '(B01) p '(B00)))
$1 = (((INC RGT GET (LOP LFT) LFT PUT) (LFT PUT) (LFT))
((INC RGT GET (LOP DEC RGT) LFT PUT) (LFT PUT) (DEC RGT))
((INC RGT GET (LOP INC LFT) LFT PUT) (LFT PUT) (INC LFT))
((INC RGT GET (LOP RGT RGT) LFT PUT) (LFT PUT) (RGT RGT)))
```

REFERENCES

- [1] Free Software Foundation, Inc. *GNU Guile 2.2.6*. See <https://www.gnu.org/software/guile/>; accessed 2021-03-31. 1995–2019.
- [2] Daniel P. Friedman et al. *The Reasoned Schemer*. 2nd ed. The MIT Press, 2018.
- [3] Daniel P. Friedman et al. *trs2-impl.scm*. Hosted on GitHub: <https://raw.githubusercontent.com/TheReasonedSchemer2ndEd/CodeFromTheReasonedSchemer2ndEd/master/trs2-impl.scm>; accessed 2021-03-31. 2018.
- [4] Urban Müller. *brainfuck*. See <https://esolangs.org/wiki/Brainfuck>; accessed 2021-03-31. 1993.
- [5] The Go Project. *Go 1.16*. Accessed 2021-03-31. URL: <https://golang.org/doc/go1.16>.

