Tree shaking in Stak Scheme

@raviqqe

July 20, 2025

Contents

- Stak Scheme
- Progress
- Tree shaking
- Future work

Stak Scheme

- A bytecode compiler and virtual machine (VM) for Scheme
- The compiler is written in Scheme.
- The VM is written in Rust.
- It aims to support R7RS-small.

Progress

- Tree shaking
- Unicode (UTF-8) support in I/O
- The `include` syntax

Tree shaking

- Scheme is an **impure** functional programming language.

- Any procedure calls might have side effects in Scheme.

- Even in libraries, we might have top-level procedure calls.

- Variable definitions might also cause side effects.

- No immutable definitions in the R7RS standard.

- Tree shaking is so difficult. 😃

TypeScript

- Many bundlers, such as Webpack and Rolldown, implement tree shaking by default.

- ECMAScript introduced `const` variable definitions relatively recently.

- ESM is much easier to analyze statically than the other module systems like CommonJS.

  `

```typescript
const foo: number = 42;

const bar = (x: number): number => x * foo;

// These results in static analysis or runtime errors.
foo = 0;
bar = (x) => x;
```

Scheme

- Every variable or function definition is *variable*.

```scheme
(define foo 42)

(define (bar x)
  (* x foo))

; Any top-level definitions might be overwritten at runtime.
(set! foo 0)
(set! bar #f)
```

Tree shaking in Stak Scheme

- Tree shaking in Stak Scheme changes the semantics of Scheme a little bit.
- A top-level procedure call requires all symbols inside the expression for the program.
- A variable in a variable definition depends on all symbols inside its value expression.
  - We assume that variable definitions whose symbols (variable names) are not required can be shaken off.

Algorithm

1. Collect all symbols in a main program and top-level procedure calls in libraries.

2. Mark these symbols required.
   - They are the *root* of symbols.

3. Collect all dependencies between symbols in procedure and variable definitions in libraries.

4. Mark the library symbols required transitively.

5. Remove all definitions of symbols not required.

Results

- PR 2549: Tree shaking

Future work

- Debug mode
  - Stack trace
- Unicode in the `(scheme char)` library