CREXX
Programming Guide

The CREXX team

July 26, 2026

THE REXX LANGUAGE ASSOCIATION
CREXX Programming Series
ISBN 978-94-039116-2-5

Publication Data

The responsible publisher of this edition is identified as *IBizz IT Services and Consultancy*, Amsteldijk 14, 1074 HR Amsterdam, a registered company governed by the laws of the Kingdom of The Netherlands.

This edition is registered under ISBN 978-94-039116-2-5

Content is up to date with version *crexx-1.0.0-beta.3+local.g78c83e4a0b88*

Contents

The CREXX Publication Series

This book is part of a library, the *CREXX Programming Series*, documenting the CREXX programming language and its use and applications. This section lists the other publications in this series, and their roles. These books can be ordered in convenient hardcopy and electronic formats from the Rexx Language Association.

Programming Guide	The Programming Guide explains the working of the tools and has examples for building programs on the platforms it supports.
Language Reference	Referred to as the CRL, this is meant as the formal definition for the language, documenting its syntax and semantics, and prescribing minimal functionality for language implementers.
Library Reference	A complete reference of all included functionality in the CREXX distribution.
VM Specification	The CREXX VM Specification documents the CREXX Assembly Language and its execution by the CREXX Virtual Machine. It also contains low level virtual machine and ABI specifications.

Typographical conventions

In general, the following conventions have been observed in the CREXX publications:

- Body text is in this font
- Examples of language statements are in a `keyword` or **bold** type
- Items that are introduced, or emphasized, are in an *italic* type
- Included program fragments are listed in this fashion:

```
/* salute the reader */
say 'hello reader'
```

Console output generated from programs contained in the text is shown in this form:

```
hello reader
```

1 About This Book

This programming guide explains how to use the CREXX Release 1 beta line toolchain in practice.

It focuses on the tools and workflows around Level B source:

- compiling `.crexx` with `rxc` or `crexx`
- learning practical Level B syntax, modules, classes, interfaces, and tested examples through the Level B tutorial
- assembling `.rxas` with `rxas`
- running `.rxbin` with the crexx virtual machine executables `rxvm` and `rxvme`
- linking modules with `rxlink`
- packaging native executables with `crexx -native` and `rxcpack`
- using standard libraries, classes, interfaces, plugins, compiler exits, and integration APIs

The language itself is defined in the CREXX *Language Reference*. The lower level bytecode, assembly, and runtime model are described in the CREXX *VM Specification*.

1.1 Audience

This guide is for developers who want to build and run CREXX programs, integrate CREXX into tools, or understand how the command-line pieces fit together. Plugin

library and compiler-exit authors should read this guide together with the VM specification and the relevant public headers.

PART I

Guide

2

Level B Tutorial

This tutorial is for experienced Rexx programmers who want to write practical crexx Level B programs in the Release 1 beta 3 documentation line. It assumes you know Rexx ideas such as `say`, `parse`, `do`, `arg`, and built-in functions, but it does not assume you have used cRexx's typed compiler, modules, classes, or interfaces before.

Level B is Rexx-family code, but it is not Classic Rexx compatibility mode. It is the implemented typed crexx language used by the toolchain, libraries, tests, and examples in this repository. When this tutorial says "Level B", it means what the current compiler accepts now, not a planned Level C or Level G feature.

For exact syntax details, keep these reference pages close:

- Data types
- Statements
- Binary memory
- Classes and interfaces
- Namespaces
- Toolchain overview

For more implemented examples in a source download, start with these paths:

- `examples/hello.crexx`,
- `examples/ooBank.crexx`,
- `lib/rxfnsb/rexx/stem.crexx`,

- `compiler/tests/rexx_src/interface_showcase_same_module.crexx`, and
- `compiler/tests/rexx_src/reference_source_iterator.crexx`.

2.1 Why Level B Still Feels Like Rexx

The first pleasant surprise is that much of the surface still reads like Rexx:

- `say` prints a value.
- `arg` receives command-line or procedure arguments.
- `do`, `if`, `select`, `leave`, `iterate`, and `return` shape control flow.
- `parse`, `address`, `signal`, and `trace` are present in the current beta surface.
- The `rxfnsb` library provides many familiar built-in functions.

The biggest difference is that Level B gives the compiler real information: values have types, modules have namespaces, procedures declare arguments and return types, and object contracts are checked.

That trade is worth leaning into. Do not write vague Classic Rexx and hope the compiler guesses your intent. Write a small amount of explicit Level B so the compiler and VM can help you.

2.2 First Program and Workflow

The usual source extension for new crexx code is `.crexx`. For reusable code, start with `options levelb` and import the Level B standard library when you use it.

```
options levelb
import rxfnsb

say "Hello Level B"
say "length=" || length("Rexx")
```

From the repository root:

```
crexx hello.crexx
```

Expected output:

```
Hello Level B
length=4
```

The `crexx` driver wraps the normal compile, assemble, and run path. The underlying stages are still visible:

1. `rxc` compiles source to `.rxas`.
2. `rxas` assembles `.rxas` to `.rxbin`.
3. `rxvm` or `rxvme` runs `.rxbin`.
4. `rxlink` can combine modules into a linked image.

The commands in this tutorial assume CREXX has been installed and `crexx`, `rxc`, `rxas`, and `rxvm` are on your `PATH`. In a source download before installation, use the matching binaries from your build directory instead. For day-to-day scripts, use `crexx`. For multi-module class/interface examples where provider modules must be present at runtime, compile the modules and run or link the resulting `.rxbin` files explicitly. The mini-project below shows that shape.

2.3 Source File Conventions

Use this header shape for committed Level B code. This is a header fragment, not a complete program:

```
options levelb
namespace myapp expose public_symbol
import rxfnsb
```

Keep `options`, `namespace`, and `import` in the leading header block. The compiler pre-scans that header before full parsing, and the rest of the repo uses this layout consistently.

The pieces mean:

- `options levelb`: compile this file as Level B.
- `namespace myapp expose public_symbol`: publish selected symbols from this module.
- `import rxfnsb`: make another namespace visible to this source file.

Headerless top-level scripts run through the `crexx` driver get practical defaults, including Level B and `rxfnsb`, but tutorial and library code should be explicit.

Command-line arguments arrive through `arg`:

```
options levelb

arg words = .string[]

if words[0] = 0 then do
  say "no arguments"
  return
end

say "count=" || words[0]
do i = 1 to words[0]
  say i || ":" || words[i]
end
```

Run with arguments by putting `-args` last:

```
crexx args.crexx -args alpha beta
```

Expected output:

```
count=2
1:alpha
2:beta
```

2.4 Types You Will Use Immediately

Level B values have types. The most common built-in source spellings are:

Type	Use
`.string`	UTF-8 text
`.int`	integer values
`.float`	binary floating-point values
`.decimal`	decimal numeric values
`.boolean`	truth values
`.binary`	arbitrary bytes
`.object`	object-shaped values
`.void`	no return value

Write `.boolean` in current Level B source. Some Rexx or C habits may suggest

`.bool`, but `.boolean` is the documented Level B spelling.

A local variable can be inferred from its first assignment, or you can declare the type first and assign the value next. The latter is useful when the value will be filled later or when you want a specific target type.

```
options levelb
import rxfnsb

title = .string
count = .int
price = .float
ready = .boolean
payload = "4f4b"x as .binary
words = .string[]
people = .stem()

title = "Level B"
count = 2
price = 1.5
ready = 1

call arrayappend words, "alpha"
call arrayappend words, "beta"
words.3 = "gamma"

people.ada = "analytical"
people["grace.hopper"] = "compiler"
tail = "ada"

say title || ":" || count
say typeof(price)
say typeof(ready)
say binlength(payload)
say words[0] || ":" || words.1 || "," || words[2] || "," || words.3
say people.tail
say people["grace.hopper"]
```

Expected output:

```
```

Practical notes:

- Arrays are typed: `.string[]`, `.int[]`, `.object[]`, and so on.
- Default growable arrays are commonly used with one-based element indexes, but Level B also supports explicit bounds such as `.int[0 to 10]` and

`.int[-2 to *]`.

- `array[0]` or `array.0` reads the current high water mark.
- Elements can be read and written with bracket notation (`words[1]`) or Rexx stem-style dot notation (`words.1`).
- Do not assign to index `0`; use ordinary element assignment or helpers such as `arrayappend`, `arrayinsert`, and `arraydelete`.
- The `.stem` class is a string-to-string keyed container for classic Rexx compound-variable style data. It supports dotted tails such as `people.ada` and bracket keys such as `people["grace.hopper"]`.
- `.string` is valid UTF-8 text in normal builds. Use `.binary` for bytes.

2.5 Procedures, Arguments, Returns, and Varargs

A Level B procedure can declare its return type after `=`, and it binds call arguments with `arg`.

```
options levelb

main: procedure
  total = sum(2, 3, 5)
  x = 10
  call bump(x)

  say "total=" || total
  say "x=" || x
  return

sum: procedure = .int
  arg first = .int, ... = .int
  total = first
  do i = 1 to arg[]
    total = total + arg[i]
  end
  return total

bump: procedure = .void
  arg expose value = .int
  value = value + 1
  return
```

Expected output:

```
total=10
x=11
```

The important habits are:

- `arg name = .type` is by value. Changes in the procedure do not update the caller's variable.
- `arg expose name = .type` is by reference. Use it only when the procedure is meant to update caller-owned state.
- A final `... = .type` accepts a variable-length tail.
- `arg[]` is the number of values in that tail, and `arg[i]` reads tail values.

`procedure expose` is different from `arg expose`: it binds module-global storage into a procedure. Prefer explicit arguments and return values until you really need shared module state.

2.6 Expressions, Assignment, Casts, and Comparisons

Level B keeps Rexx operators, but type checking happens before bytecode is emitted.

Use these idioms:

- Assign a variable once with the type you mean; later assignments must be compatible.
- Use `expr as .type` when you need a checked conversion or object cast.
- Use `expr is .type` for object/interface tests.
- Use `typeof(expr)` for diagnostics and runtime type introspection.
- Use = for Rexx-style equality and == when you mean exact string comparison after string promotion.

For binary data, be explicit. This is a fragment:

```
payload = "4f4b"x as .binary
```

For objects, cast at the boundary. This is a fragment:

```
generic = selected as .object
restored = generic as .asset
```

If a cast cannot succeed, Level B raises a conversion signal instead of silently changing the meaning of the value.

2.7 Control Flow

The usual Rexx control-flow tools are present. `select` has both Classic Rexx condition style and a switch-like expression style. `leave` and `iterate` work on loops. Expression-form `do ... end` can return a value with `leave with`.

```
options levelb

arg words = .string[]

if words[0] = 0 then words[1] = "red"

kept = 0
do i = 1 to words[0]
  select
    when words[i] = "skip" then iterate
    when words[i] = "stop" then leave
    otherwise say "word=" || words[i]
  end
  kept = kept + 1
end

summary = do
  if kept > 1 then leave with "many"
  leave with "few"
end

say "summary=" || summary
```

Run:

```
crexx control_flow.crexx -args alpha skip beta stop gamma
```

Expected output:

```
word=alpha
word=beta
summary=many
```

2.8 Rexx-Flavoured Features

Level B keeps several Rexx features that make the language feel direct:

- `say` prints strings.
- `parse` is implemented through the certified `PARSE` exit.
- `address` sends commands to an external environment and can capture output.
- `signal` raises and handles condition objects.
- `trace` enables VM-backed tracing for debugging.

This example keeps the output small:

```
options levelb
import rxfnsb

main: procedure
  parse value "Ada Lovelace,1815" with first last "," year
  say last || ":" || year

  out = .string[]
  err = .string[]
  address crexx "echo 42" output out error err
  say "address=" || out[1]

  handled = ""
  do
    signal other "from block"
  on signal other as problem
    handled = problem.name() || ":" || problem.message()
  end
  say handled
  return
```

Expected output:

```
Lovelace:1815
address=42
OTHER:from block
```

Trace is useful but intentionally noisy. These are reference-only forms, not a runnable tutorial example:

```
trace results
```

```
trace asm
trace llm to file "trace.jsonl"
trace unsuppress namespace rxfnsb
trace off
```

Do not add `TRACE`, `PARSE`, or `ADDRESS` directly to `lib/rxfnsb/rexx/` library sources while debugging the library build. Those files are built with compiler exits disabled. Write a small caller program instead.

2.9 Modules and Libraries

Each source file is a module. Public module identity comes from `namespace`, not from the filename alone. A module exposes selected symbols; another module imports the namespace and calls those symbols.

Library module:

```
options levelb
namespace greetings expose salutation count_items

salutation: procedure = .string
  arg name = .string
  return "Hello, " || name

count_items: procedure = .int
  arg items = .string[]
  return items[0]
```

Caller:

```
options levelb
import greetings

names = .string[]
names[1] = "Ada"
names[2] = "Grace"

say greetings..salutation(names[1])
say "names=" || greetings..count_items(names)
```

Run the caller with the directory containing `greetings.crexx` on the source import path:

```
crexx modules_main.crexx -s/path/to/examples
```

Expected output:

```
Hello, Ada
names=2
```

Use `namespace..symbol` for qualified calls. The older `namespace::symbol` spelling is accepted for compatibility, but new examples should prefer the double-dot form.

2.10 Classes

Classes hold attributes and methods. A class factory is written as `*: factory` or `name: factory`; do not write a return type on a factory. In a class factory, bare `return` returns the object being constructed. This fragment is from the complete interface example in the next section:

```
fileasset: class implements .asset
  _name = .string

  *: factory
    arg name = .string
    _name = name
    return

  kind: method = .string
    return "file"

  name: method = .string
    return _name
```

Ordinary application classes should let the compiler allocate attribute storage. The `with register.N` form exists for low-level VM/native interop, not for day-to-day business objects.

2.11 Interfaces

Interfaces define contracts. Classes implement them. Interface methods can be abstract, or they can provide a default/final body. In current Level B, an interface method with a body is final: a class can rely on it, but cannot override it.

Here is a complete same-file interface and provider example:

```
options levelb
```

```
namespace tutorial_objects

main: procedure
  selected = .asset("log.txt")
  fallback = .asset("memo")

  say selected.describe()
  say fallback.describe()

  if selected is .fileasset then say "file selected"

  generic = .object
  generic = selected as .object
  restored = generic as .asset
  say typeof(restored)
  return

asset: interface
  *: factory
  arg name = .string

  describe: method = .string
    return kind() || ":" || name()

  kind: method = .string
  name: method = .string

fileasset: class implements .asset
  _name = .string

  *: match
    arg name = .string
    if name = "log.txt" then return 100
    return 0

  *: factory
    arg name = .string
    _name = name
    return

  kind: method = .string
    return "file"

  name: method = .string
    return _name

cacheasset: class implements .asset
  _name = .string
```

```
*: factory
  arg name = .string
  _name = name
  return

kind: method = .string
  return "cache"

name: method = .string
  return _name
```

Expected output:

```
file:log.txt
cache:memo
file selected
.tutorial_objects..fileasset
```

The `asset` interface owns the public factory. `fileasset` and `cacheasset` are providers. The `*: match` method is a class-side selector. Positive scores accept the provider, zero or negative scores reject it, and the highest score wins.

2.12 Object Use

Use object values through factories, methods, interfaces, type tests, and checked casts:

- `.asset("log.txt")` calls an interface factory.
- `.fileasset("log.txt")` calls a concrete class factory.
- `value.method()` invokes a method.
- `value is .asset` tests whether the concrete object implements an interface.
- `value as .asset` casts after a runtime check.
- `typeof(value)` reports the concrete runtime type.

Level B does not implicitly call a `toString()` method for arbitrary objects. If you want text, expose a method such as `format`, `describe`, or `name`, then call it explicitly.

Level B also does not currently implement interface inheritance, interface attributes, overloads, singleton declarations, or destructor/finalizer syntax.

2.13 Collections and Iteration

Current Level B collection patterns are explicit:

- Use typed arrays such as `.string[]`, `.int[]`, and `.object[]`.
- Default growable arrays are usually shown with one-based element indexes, but array bounds can be explicit.
- Read `array[0]` for the current count.
- Use bracket notation (`items[1]`) or stem-style dot notation (`items.1`) for array elements.
- Use `arrayappend`, `arrayinsert`, `arraydelete`, and related `rxfnsb` helpers for mutating array shape.
- Use `objectarrayappend`, `objectarrayinsert`, and related helpers for `.object[]` when you need object-array mutation helpers.
- Store concrete objects in `.object[]` with an explicit `as .object` upcast, then cast back to the expected interface or class at the read boundary.
- Use `.stem` when the natural model is a REXX compound variable or keyed string map rather than an ordered typed array.

The mini-project below uses this pattern. This is a fragment:

```
items = .object[]
items[1] = .ledger..entry("coffee", -3) as .object

item = items[1] as .ledger..entry
say item.format()
```

For iterator-like classes, current Level B uses explicit references when the iterator should see live container state. Use snapshots when the iterator should see a stable copy.

2.14 References

References are explicit weak aliases to storage. They do not keep a target alive. If the target storage goes away, `refvalid(ref)` returns false and using the reference raises `REFERENCE_INVALID`.

Use the four source forms deliberately:

- `reference target`: create a weak reference to aliasable storage.
- `local = dereference ref`: link a current-scope local to the target.
- `snapshot ref`: make a deep copy of the current target.
- `refvalid(ref)`: test whether the target is still valid.

This example contrasts a live reference with a snapshot:

```
options levelb
import rxfnsb
namespace tutorial_references

main: procedure
  list = .Bag()
  call list.add("red")
  call list.add("blue")

  live = list.liveCounter()
  snap = list.snapshotCounter()

  call list.add("green")

  say "live=" || live.count()
  say "snapshot=" || snap.count()
  return

Bag: class
  values = .string[]
  item_count = .int

  *: factory
    initial = .string[]
    values = initial
    item_count = 0
    return

  add: method = .void
    arg value = .string
    call arrayappend values, value
    item_count = item_count + 1
    return

  size: method = .int
    return item_count

  liveCounter: method = .LiveCounter
    return .LiveCounter(reference self)

  snapshotCounter: method = .SnapshotCounter
    return .SnapshotCounter(reference self)
```

```
LiveCounter: class
  bag_ref = reference .Bag

  *: factory
    arg source = reference .Bag
    bag_ref = source
    return

  count: method = .int
    if \refvalid(bag_ref) then return -1
    bag = dereference bag_ref
    return bag.size()

SnapshotCounter: class
  bag_copy = .Bag

  *: factory
    arg source = reference .Bag
    bag_copy = snapshot source
    return

  count: method = .int
    return bag_copy.size()
```

Expected output:

```
```

Keep reference boundaries visible. Level B does not let you write convenience forms such as `list_ref.add(...)` or `items_ref[i]` directly through the reference.

2.15 Mini-Project: A Typed Ledger Module

This mini-project uses two source files: a reusable module with an interface and class, and a small application that imports it.

`ledger.crexx`:

```
options levelb
namespace ledger expose entry ledger_total

entry: interface
  *: factory
  arg label = .string, amount = .int
```

```
  label: method = .string
  amount: method = .int

  format: method = .string
    return label() || "=" || amount()

ledgerentry: class implements .entry
  _label = .string
  _amount = .int

  *: factory
    arg label = .string, amount = .int
    _label = label
    _amount = amount
    return

  label: method = .string
    return _label

  amount: method = .int
    return _amount

ledger_total: procedure = .int
  arg entries = .object[]
  total = 0
  do i = 1 to entries[0]
    item = entries[i] as .entry
    total = total + item.amount()
  end
  return total
```

ledger_app.crexx:

```
options levelb
import ledger

items = .object[]
items[1] = .ledger..entry("coffee", -3) as .object
items[2] = .ledger..entry("book", -12) as .object
items[3] = .ledger..entry("gift", 20) as .object

do i = 1 to items[0]
  item = items[i] as .ledger..entry
  say item.format()
end

say "total=" || ledger..ledger_total(items)
```

Compile both modules, then run with the standard library, provider module, and

application module loaded:

```
CREXX_BIN=$(dirname "$(command -v crexx)")
rxc -i "$CREXX_BIN" -o main ledger_app.crexx
rxas -o main.rxbin main
rxc -i "$CREXX_BIN" -o ledger ledger.crexx
rxas -o ledger.rxbin ledger
rxvm "$CREXX_BIN/library.rxbin" ledger.rxbin main.rxbin
```

Expected output:

```
coffee=-3
book=-12
gift=20
total=5
```

The explicit `rxvm` command matters here because interface factory providers are discovered from the modules present in the runtime image. For deployable programs, `rxlink` is the usual way to combine those modules into one image.

2.16 Migration Guide For Classic Rexx Habits

Classic Rexx habit: Rely on untyped variables.

Level B habit: Let obvious literals infer types, or declare variables with `.string`, `.int`, `.float`, `.boolean`, `.binary`, arrays, or object contracts.

Classic Rexx habit: Assume built-in functions are always visible.

Level B habit: `import rxfnsb` in reusable source, unless you are deliberately writing a headerless `crexx` driver script.

Classic Rexx habit: Use stems as flexible bags of values.

Level B habit: Use typed arrays and array helper functions for ordered typed data. Use `.stem` for classic compound-variable style keyed strings. Treat `array[0]` as a count, not as a writable stem slot.

Classic Rexx habit: Share state freely through globals.

Level B habit: Prefer arguments and returns. Use `namespace ... expose`, `procedure expose`, or `arg expose` only for intentional sharing.

Classic Rexx habit: Treat text and bytes as the same thing.

Level B habit: Use `.string` for valid UTF-8 text and `.binary` for arbitrary bytes. Convert explicitly.

Classic Rexx habit: Expect dynamic object behavior.

Level B habit: Define an interface, implement it with classes, and use checked casts and type tests at boundaries.

2.17 Known Beta 3 Boundaries

Level B is the main implemented Release 1 beta language, but this is still a beta line. Current boundaries that matter to tutorial code:

- Level B is not Classic Rexx compatibility mode. Level C is where Classic compatibility work belongs, and normal Level C compilation is not a beta 3 contract unless the beta 3 release notes explicitly say otherwise.
- Level G has real early library work, including `rxfnsg`, but it is not the baseline user language for this release line.
- Interface default methods are final.
- Interface inheritance, interface attributes/state, overloads, singleton declarations, external-object adoption syntax, and destructor/finalizer syntax are not current Level B features.
- Objects do not implicitly convert to strings.
- References are explicit and weak. They are powerful, but not a general replacement for ordinary arguments, returns, or object methods.
- Multi-module interface-provider programs need all provider modules loaded or linked at runtime.

3 Packed Binary Memory

Packed binary memory is for code that needs to treat a `.binary` value as a byte-addressed record space. Typical examples are sorted lookup tables, parser keyword tables, indexes, protocol records, and compact tree nodes.

Use this feature when a packed layout avoids many ordinary Rexx variables, temporary substrings, or repeated binary slice copies. Do not use it just to make simple code look lower-level; normal variables are clearer when the data is not layout-sensitive.

The exact language rules are in the Language Reference chapter Binary Memory.

3.1 Choosing Intrinsics Or Helpers

Use binary-memory intrinsics for hot direct access:

```
hash = <at..u32>(node + NODE_HASH) arena
if <compare..binary>(arena, key_offset, wanted_key) = 0 then say "hit"
```

Use ordinary helper functions for buffer management and mutation:

```
call binresize arena, NODE_SIZE * 128
call binmemmove arena, dst_offset, src_offset, count
call bincopy target, 0, source, source_offset, length
```

The compiler may direct-lower selected helpers, but the source still reads as an ordinary action. That is deliberate: moving, resizing, and opening gaps are operations, not field references.

3.2 Performance Shape

Direct binary-memory field access is designed for hot packed-data loops. The Release 1 VM keeps the strict bounds checks but uses direct fixed-width load/store helpers for common scalar widths. In microbenchmarks, repeated `<at..u8>`, `<at..u32>`, and `<at..int>` access is faster than indexed attribute-array access when the loop is mostly scalar field reads or writes.

That does not mean every data structure should be packed. A binary layout wins when it removes copying, stores dense records, or lets the program scan bytes without materializing substrings. A normal Rexx array or set of registers can still win when the algorithm naturally works with scalar values and the packed layout forces repeated manual offset arithmetic.

Good binary-memory candidates:

- byte scanners and generated lexers;
- fixed-width token or index tables;
- lookup records where key bytes can be compared without slicing;
- large binary constants used directly with `<at>` and `<compare>`;
- compact external record formats whose layout is already byte-based.

Weaker candidates:

- small scalar metadata that is read one field at a time;
- code that needs many separate offset calculations for every logical action;
- ordinary string processing where Unicode-aware `.string` operations dominate;
- structures where a register or `.int[]` representation already avoids copies.

Keep offsets in local variables inside hot loops. For repeated access to fields in one record, compute the row or node base once, then add field offsets:

```
row = HEADER_ROWS + i * ROW_SIZE
hash = <at..u32>(row + ROW_HASH) table
kind = <at..u8>(row + ROW_KIND) table
```

Use smaller fixed-width storage only when it matches the format. `.u8` and `.u32` reduce the binary footprint and can help dense structures, but they also add range semantics. Use `.int` when the field is genuinely a Rexx integer and the extra bytes do not matter.

3.3 Layout Constants

Write layout constants in bytes. This makes offset arithmetic visible and keeps the source independent of the storage type used for each field.

```
options levelb
namespace packedindex expose find_node

packedindex_layout: procedure expose NODE_LEFT NODE_RIGHT NODE_HASH
    NODE_KEY NODE_SIZE
  constant NODE_LEFT = 0
  constant NODE_RIGHT = NODE_LEFT + <sizeof..u32>
  constant NODE_HASH = NODE_RIGHT + <sizeof..u32>
  constant NODE_KEY_LEN = NODE_HASH + <sizeof..u32>
  constant NODE_KEY = NODE_KEY_LEN + <sizeof..u16>
  constant NODE_SIZE = 32
  return
```

For reusable modules, expose public procedures. Keep Release 1 layout constants inside the source module and declare them in an explicit procedure scope. Top-level executable code in a library-shaped file can synthesize an implicit `main()`, which is not usually what a packed-layout module wants.

For script-style examples that need a separate layout-constant declaration procedure, add an explicit `main: procedure`. Otherwise the statements after the declaration procedure are part of that procedure body until the next callable boundary.

3.4 Header Check

Binary constants and binary variables use the same direct-access syntax for reads and compares.

```
options levelb

check_header: procedure = .int
  arg page = .binary
  constant MAGIC = "4352584944583031"x as .binary  /* CRXIDX01 */

  if <compare..binary>(page, 0, MAGIC) = 0 then return 1
  return 0
```

Use = only when comparing already-materialized values:

```
magic_copy = <at..binary>(0, <blen>(MAGIC)) page
if magic_copy = MAGIC then say "copied compare"
```

For lookup loops, prefer `<compare..binary>` so the compiler can emit a direct binary compare.

3.5 Sorted Table Lookup

This example shows the intended pattern for a packed sorted table. The table header stores the row count. Each row has a hash and a variable-size key area. The example omits the full binary-search loop to focus on field access.

```
options levelb

find_row: procedure = .int
  arg table = .binary, wanted_hash = .int, wanted_key = .binary

  constant HEADER_COUNT = 0
  constant HEADER_ROWS = 8
  constant ROW_HASH = 0
  constant ROW_KEY_LEN = 4
  constant ROW_KEY = 6
  constant ROW_SIZE = 40

  count = <at..u32>(HEADER_COUNT) table
  do i = 0 to count - 1
    row = HEADER_ROWS + i * ROW_SIZE
    hash = <at..u32>(row + ROW_HASH) table
    if hash = wanted_hash then do
      key_len = <at..u16>(row + ROW_KEY_LEN) table
      if key_len = <blen>(wanted_key) then
        if <compare..binary>(table, row + ROW_KEY, wanted_key) = 0 then
          return row
    end
  end

  return -1
```

The hot loop reads fixed-width fields directly, checks the stored key length, and compares the key without making a binary slice. The compiler test fixture for this shape inspects emitted RXAS and fails if `bslice`, string extraction, or helper calls appear in the lookup path.

3.6 Insert And Delete Paths

Insertion and deletion usually need memory movement. Keep that as helper-style source. It may still be direct-lowered, but it does not need an angle-bracket intrinsic unless profiling proves the helper spelling is a problem.

```
options levelb

insert_gap: procedure = .void
  arg expose table = .binary, row = .int, count = .int
  constant ROW_SIZE = 40

  old_len = <blen>(table)
  new_len = old_len + ROW_SIZE
  call binresize table, new_len

  move_from = row
  move_to = row + ROW_SIZE
  move_len = old_len - row
  if move_len > 0 then call binmemmove table, move_to, move_from,
      move_len

  call binfillat table, row, ROW_SIZE, 0
```

`binfillat(table, offset, length, byte)` is the Release 1 span-fill helper. It currently uses a small checked helper loop; whole-buffer `binfill(table, byte)` is backed by the RXAS `bfill` instruction. `binmemmove` is safe for overlapping ranges and currently uses conservative library code; direct lowering remains a future optimization if profiling shows helper overhead in hot mutation paths.

3.7 Text And Decimal Fields

There are two text layouts, but new packed structures should prefer zero-terminated UTF-8 fields:

- zero-terminated UTF-8 fields, read with `<at..string>(offset)` and compared
- fixed-codepoint UTF-8 source spans, read with `<at..string>(offset, codepoints)` only when the program already knows the source span in codepoints.

For zero-terminated fields, the program must know that the binary format stores a zero byte after the text. Omit the length to read the field, or use `<compare..string>`

to compare it without allocating a string:

```
keyword = <at..string>(keyword_offset) record
if <compare..string>(record, keyword_offset, "select") = 0 then say "
    keyword"
```

Invalid UTF-8 signals `UNICODE_ERROR`.

Decimal fields use zero-terminated decimal text. This keeps the packed format simple for variable-width decimal values; the program only needs the starting byte offset:

```
amount = <at..decimal>(amount_offset) record
<at..decimal>(amount_offset) record = amount + 1d
```

The read form scans to the zero byte, validates the selected text, and parses it through the active decimal plugin. The write form converts the decimal value to decimal text, writes the bytes, and writes one zero byte. The binary buffer must already have room for the complete text plus terminator.

3.8 When Not To Pack

Prefer ordinary Rexx values when:

- the data is small and not in a lookup hot path;
- field layout is not externally visible;
- text processing is mostly Unicode-aware string work;
- the code would need many manual offsets but little measurable speedup;
- a class or array would express the model more safely.

Packed binary memory is a performance tool. Use it when it removes real copying or lets a data structure stay cache-friendly.

4 Overview of the Toolchain

CREXX is built as a small set of separate tools with visible boundaries. That is deliberate: users can inspect the generated assembly, assemble it directly, link modules, disassemble bytecode, or package a program for native deployment.

The common flow is:

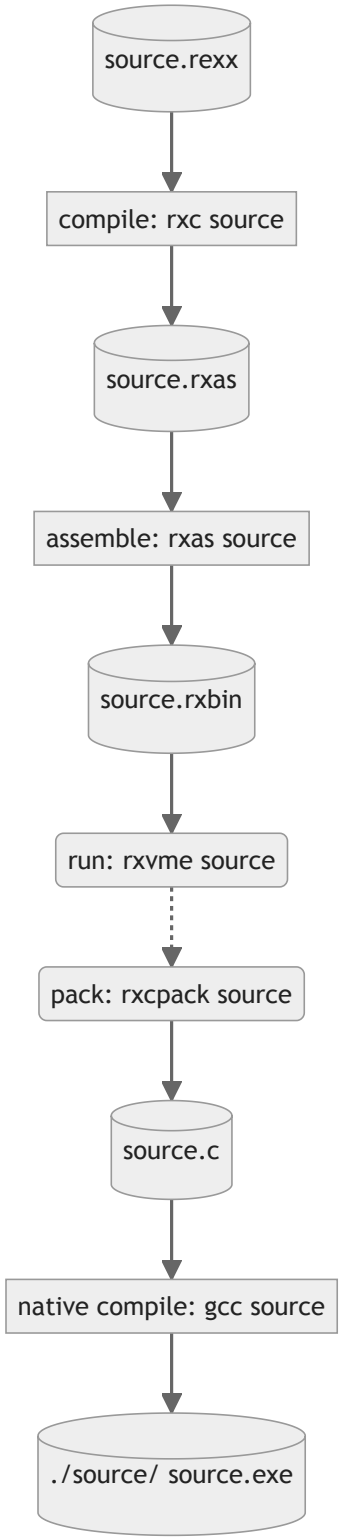
1. `rxc` compiles CREXX source to `.rxas` assembly.
2. `rxas` assembles `.rxas` assembly to `.rxbin` bytecode.
3. `rxvm`, `rxvme`, `rxbvm`, or `rxbvme` executes one or more `.rxbin` files.
4. `rxlink` optionally combines modules into a linked image with a shared constant pool.
5. `rxcpack` can package a linked image as C data for native executable builds.

The `crexx` driver (see the crexx tool on page 63 wraps the usual compile, assemble, run, link, and package steps for day-to-day use. For more elaborate application systems, with separately compiled source modules and library use, a build system like Ninja or CMake[1] is recommended.

4.1 Source, Assembly, and Bytecode

Source files are UTF-8 text. CREXX source normally uses `.crexx`, with `.crx` as a short alias. `.rexx` remains the compatibility/classic extension. The compiler also

[1] The native executables in the CREXX distribution are built with CMake for all Operating System / Instruction Set Archi-



tecture combinations.

accepts an arbitrary non-reserved initial source extension for that compile. Toolchain artifacts are less flexible: *rxas* assembly uses the fixed `.rxas` suffix and *rxbin* bytecode uses the fixed `.rxbin` suffix. The compiler output is RXAS assembly:

```
rxc hello.crexx
```

The assembler output is `.rxbin` bytecode:

```
rxas hello.rxas
```

`.rxbin` is the module format consumed by the VM, linker, disassembler, and packager.

For `rxc`, `rxas`, and `rxlink`, `-o` names an output stem/path unless it already ends in the tool's fixed artifact suffix. For example, `rxas -o app` writes `app.rxbin`, `rxas -o app.rxbin` writes `app.rxbin`, and `rxas -o app.debug` writes `app.debug.rxbin`.

4.2 Module Discovery

The compiler separates source discovery from binary discovery:

- source roots contain `.crexx`, `.crx`, `.rexx`, and any arbitrary extension used by the initial source file for this compile
- binary roots contain `.rxbin`, optional `.rxas`, and `.rxplugin` files

Use `rxc -s path` to add a source import root and `rxc -i path` to add a binary import root. The `crexx` driver exposes the same distinction for compilation. The directory of the source being compiled is searched for source imports, but it is not searched for `.rxbin` imports unless it is also passed with `-i`. That avoids stale generated bytecode shadowing the source you are editing.

When a source file has no `options level...` clause, `.crexx`, `.crx`, and arbitrary initial extensions default to Level G; `.rexx` defaults to Level C.

The runtime library path is separate. Use `crexx -l...` or the VM's own library loading rules when a compiled program must load bytecode or plugins at runtime.

4.3 Linking

`rxlink` combines one or more `.rxbin` modules into a linked image with one shared constant pool. This reduces duplicate constants and resolves module and interface-provider relationships up front while preserving the module records needed by the VM loader.

The use of `rxlink` is a good choice for release-style deployable images and native packaging.

4.4 Native Packaging

Native packaging uses `rxlink` before `rxcpack`. The linked image is serialized to C data and linked with the VM runtime library by the platform C toolchain. The generated executable contains the program image and the runtime pieces selected by the build.

Plugin libraries (which are native libraries) in the CREXX distributions have static and dynamic versions. Using a static or dynamic versions of a plugin is a build choice. Although the intention is to keep distributions equal over operating systems and hardware architecures, it is not realistic to expect that every distribution contains every optional native plugin. What is delivered, however, is documented in the *Library Reference* publication.

5 Global Procedures

5.1 Exposing Global Procedures

A procedure defined in a module file can be made available to other modules (provided it is in an imported namespace) by including the procedure in the namespace instruction.

```
namespace mynamespace expose myfunc1 myfunc2
```

5.2 Importing Global Procedures

When compiling a program, the compiler looks for procedures that are called but not defined in the current module. It searches in various locations in the following order:

1. The primary source root: source files in the directory of the source being compiled, or in the working location selected with `-l` when the source name has no directory.
2. Additional source roots specified with `rxc -s`: source files.
3. Binary roots specified with `rxc -i`: `.rxbin` files, optional `.rxas` files when `--import-rxas` is enabled, and CREXX native function libraries (plugins).
4. The directory where the compiler executable (`rxc`) is located: deployed `.rxbin` files, optional `.rxas` files when `--import-rxas` is enabled, and plugins.

The source root is not automatically searched for `.rxbin` files. Pass `-i .` or `-i build-dir` when a local binary module is an intended compile-time dependency. This separation avoids stale generated `.rxbin` files beside source files being imported by accident.

Source roots search `.crexx`, `.crx`, and `.rexx`. If the initial source file uses another extension, such as `.the`, that extension is searched for this compile too. Files without an explicit `options level...` default to Level G for `.crexx` and `.crx`, and arbitrary initial extensions, and Level C for `.rexx`.

The compiler only uses external procedures that are in the same namespace as the module being compiled or for namespaces listed in the `import` instruction.

```
import rxfnsb mynamespace anothernamespace
```

Metadata encoded with the procedures allows the compiler to determine the type and argument types of imported procedures.

The CREXX level b standard library namespace is `rxfnsb`, to access these procedures programs should import this namespace.

```
import rxfnsb
```

Without the `import` statement for the `rxfnsb` namespace, the level B built-in functions cannot be found. One could say that the built-in functions are not as built-in as they are in Classic Rexx. This is different in CREXX level C, the compatibility level.

6

Global Variables

Global variables are shared among modules and are linked to the namespace of the modules, similar to exposed procedures.

Namespace auto-exposing: Variables exposed through the module `namespace` instruction are automatically mapped into the local scope of all procedures within the module file. Use this when the variable is part of the module's published namespace surface.

```
namespace myproject expose global_var
```

Procedure-level exposing: The `expose` keyword in a `procedure` instruction binds the listed names to module-global storage for that procedure only. Use this for private module state shared by selected local procedures without publishing that variable through the namespace.

```
proc1: procedure = .void expose global_var
  global_var = "Hello World"
  return

proc2: procedure = .void expose global_var
  say "global_var is" global_var
  return
```

A procedure that does not list `global_var` does not see this shared storage unless `global_var` is also in the file-level `namespace ... expose ...` list.

Physical Placement and Hoisting: To ensure robust scoping, global variables physically reside in the **Namespace Scope** of the module. During the compilation process, when a variable is identified as global (either via the `namespace expose`

list or a `procedure expose` list), the compiler physically **hoists** the variable symbol from its local discovery scope up to the Namespace Scope. This ensures that standard tiered resolution (Local -> Namespace -> Universe) correctly identifies the variable across all procedures and blocks within the module. This hoisting is strictly upward; the compiler never demotes a namespace-level variable to a local scope.

Namespace Isolation of Global Variables: Unlike exposed functions, which can be called across imported modules, **exposed global variables are strictly isolated to their originating module's namespace**. The compiler explicitly prevents cross-namespace variable resolution. To read or write a global variable from a module in a different namespace, you must wrap the variable access in an exposed function (a getter/setter) and import that function.

Typing and Shadowing of Global Variables: Global variables are typed based on their first use or declaration. All modules exposing the variable must agree on this type.

- **Top-Level Declarations:** In the top-level scope of a procedure (i.e. directly inside `procedure`), a typed declaration like `x = .int` does *not* create a local shadow variable. Instead, it asserts and binds the type to the global variable `x`.
- **Sub-Scope Declarations (Shadowing):** In sub-scopes (e.g. inside a `do...end` block or an `if ... do` block), a typed declaration like `x = .int` **creates a shadow local variable** that hides the global variable for the duration of that block.
- **Untyped Assignments:** In single-instruction branches (e.g., `if cond then x = 10`), an untyped assignment binds directly to and updates the global variable.
- Inside a `DO` block (e.g., `if cond then do; x = 10; end`), if `x` does not already exist in an outer scope, it creates a block-local variable. If it does exist (including as a global), it updates that variable.

Note: The distinction between single-instruction branches and `DO` blocks is an intentional design choice in Level B.

6.1 Level B System Symbols

In Level B programming, namespaces, procedure names, class names, and variable names beginning with an underscore "_" are private. Level B is specifically designed to offer low-level system capabilities. These variables and procedures are intended for use in low-level capabilities that support other language levels.

7 Intralanguage calls

This chapter discusses calls from one CREXX procedure to another, including the built-in function package. Search order is an intrinsically related concept: how is the called component found. There are some differences with Classic REXX and ooRexx, which can call (and interpret) external procedures in source. In this respect, CREXX behaves like NetREXX, where a called component needs to be compiled, executable code; for NetREXX a `.class` file and in CREXX an `.rxbin` file. Level B introduces a package (module) system where a program can be part of a package and be imported into calling code.

7.1 At compile time

At compile time, a program uses the `CALL` statement, or the function notation with parentheses (also called round brackets). Going forward, and moving into object oriented notations for other REXX variants, the latter is going to gain importance, while the `CALL` statement will be fixed in its current functionality. For that reason, most examples will be in the function notation.

The compiler needs to verify if it is possible to call the called code: it must be present in executable form, and it needs to have the right *signature*[2]. The compiler will not automatically compile a callee of which the source can be located but the executable form is missing; existing systems based on interpreters will happily interrupt their work and tokenize another source file when called; the CREXX `rxc`

[2] with signature we mean the combination of parameter types and return type.

compiler will not.

This implies that there are inherent dependencies to be followed while building an application system that consists of multiple modules; this is not different than in other compiled languages. Building utilities like Make or Ninja can provide these services, and these can be orchestrated by meta-build tools like CMake. The CREXX toolchain itself is built using CMake and from its build specification in CMake most of these patterns can be gleaned.

The `import` statement tells the compiler we want to import functions from a certain package.

7.2 Imported interfaces and classes

Level B extends the same import model to classes and interfaces. An imported namespace may expose:

- procedures
- interfaces
- classes

An interface call is normally written against the contract name:

```
import garage

current = .vehicle("mini")
say current.describe()
```

When two imported namespaces expose the same contract name, qualify the reference with `namespace..`:

```
import qifa
import qifb

left  = .qifa..vehicle("one")
right = .qifb..vehicle.from_name("two")
```

The token to the left of `..` must be an imported namespace. Qualification is a disambiguation mechanism; it does not create a second way to reach symbols that have not been imported. `::` remains accepted as a compatibility alias.

7.3 Casts and type inspection

Object contracts can be checked explicitly:

```
generic = .car("roadster") as .object
vehicle = generic as .garage..vehicle
current = vehicle as .garage..car

say typeof(current)
say current is .garage..vehicle
say current is .garage..car
```

`as` performs a checked object cast, `is` performs a boolean type test, and `typeof` returns the canonical runtime type name.

7.4 Source and binary imports

At compile time, `rxc` distinguishes between:

- source roots, used for CREXX source files
- binary roots, used for `.rxbin`, optional `.rxas`, and plugins

The source file being compiled contributes its own directory as the primary source root. Additional source roots may be passed with `-s`. Binary roots are passed with `-i`. The primary source directory is not automatically a binary root, so a sibling `.rxbin` is visible only when that directory is also passed with `-i`.

Source-root discovery includes `.crexx`, `.crx`, and `.rexx`. If the initial source file has another extension, for example `.the`, that extension is added to source-root discovery for that compile. The default language level is Level G for `.crexx`, `.crx`, and arbitrary initial extensions, and Level C for `.rexx`; explicit `options level...` always wins.

`.rxpp` remains the extension for existing RXPP preprocessor workflows. A future idempotent preprocessor path may reduce the need for a separate preprocessor extension, but that is not part of the source import rule.

For ordinary project work this means:

- keep project source modules in source roots
- keep packaged binary libraries in binary roots

- use namespace qualification only when imported namespaces collide

This split is deliberate: it keeps active source preferred during development and stops stale generated `.rxbin` files beside source files from silently satisfying imports.

8

Interlanguage calls

A CREXX program is able to call programs written in the REXX language, but also programs native to the platform, using a number of calling conventions:

8.1 Current Mechanisms

Address the `address` statement can use the shell and I/O indirection to start native executables and provide input, and retrieve the output.

Plugins the plugin system that is available to the CREXX environment.

Crexxsaa the host integration facade for C applications that want to embed cRexx, register ADDRESS environments, and run `.rxbin` or source files through the CREXX toolchain and runtime.

9

Crexxsaa host integration

`crexxsaa` is the initial CREXX compatibility API for C hosts that want a REXXSAA-shaped entry point without taking on the full historical REXXSAA ABI. It is deliberately small: the API is a stable facade over the current CREXX `rxvml`, `ADDRESS`, sandbox, and exposed-variable contracts.

This enables the use of CREXX as an integrated part of applications or operating environments; it enables CREXX to function as a universal macro language. Supplied examples show how CREXX functions as an edit macro language[3] or to enable the use of embedded SQL[4] in a CREXX program.

The first supported host model is command-environment execution:

- host C code creates a `crexxsaa_context`
- the host registers one or more `ADDRESS` callback environments
- the host configures the CREXX compiler, assembler, import directory, and optional cache directory
- the host runs either an existing `.rxbin` or a source file that `crexxsaa` compiles and caches

This is not a full `RexxStart()` or `RexxVariablePool()` clone. Future adapter entry points may be added where real integrations need them, but the internal runtime model remains the modern CREXX `ADDRESS` model.

[3] In the THE editor, a faithful reproduction of the VM system editor by Mark Hessling.
[4] For sqlite, the universally used in-process SQL engine.

9.1 Writing hosted source

`crexxsaa` compiles source exactly as supplied. It does not add `OPTIONS LEVELB`, an `ADDRESS` clause, imports, or host-specific boilerplate.

Hosted source should therefore declare the language level and command environment it needs:

```
options levelb
address the
'emsg CREXX_PROFILE_HOSTED'
```

The script owns its language mode and its command routing. `crexxsaa` owns only compile, cache, load, and run orchestration.

9.2 Minimal host flow

```
crexxsaa_context *ctx = NULL;

crexxsaa_create(location, library_rxbin, &ctx);
crexxsaa_set_compiler(ctx, rxc_path, rxas_path, import_dir);
crexxsaa_register_address_environment(ctx, "THE", the_callback,
    userdata);
crexxsaa_set_address_environment(ctx, "THE");
crexxsaa_run_source(ctx, profile_path, "THE", 0, argc, argv, &
    program_rc);
crexxsaa_destroy(ctx);
```

`crexxsaa_run_source()` and `crexxsaa_run_rxbin()` return zero when the host API successfully compiled, loaded, and ran the program. The Rexx program status is returned separately through `program_rc`: an integer `main` return value and a top-level `exit n` both set `program_rc` to `n`.

The cache namespace argument, `"THE"` in this example, is part of the cache identity. Different hosts can compile the same source path without sharing a cache bucket accidentally.

9.3 Low-level rxvml calls

Hosts that bypass the `crexxsaa` facade and call `rxvml` directly must use the descriptor invocation APIs:

- `rxvml_call_procedure_descriptor()`

- `rxvml_call_factory_descriptor()`
- `rxvml_call_method_descriptor()`

Descriptors have the form `rxsig1|name|return_type|args`, for example `rxsig1|pkg.main|.int|pat` The runtime checks the descriptor against the callable metadata before invoking, so a name match with the wrong return or argument signature is rejected. This descriptor requirement is part of the `RXVML_ABI_VERSION` 8 break. `CREXXSAA_ABI_VERSION` is 3 so source-cache entries produced by older host/runtime pairs are not reused.

9.4 ADDRESS callbacks

A host registers an environment with `crexxsaa_register_address_environment()`. When CREXX executes a command in that environment, the callback receives a `crexxsaa_address_request` containing the environment name, command text, and active context.

The callback fills a `crexxsaa_address_response`:

- `rc`: command return code
- `condition_name`: optional CREXX condition name
- `diagnostic`: optional diagnostic text

The callback should return zero for a successfully handled dispatch. Non-zero callback return values indicate host-side failure in the bridge itself.

9.5 Variable access during callbacks

`crexxsaa` exposes simple name-based helpers for active `ADDRESS` callbacks:

- `crexxsaa_address_variable_get_alloc()`
- `crexxsaa_address_variable_set()`
- `crexxsaa_free()`

These helpers are valid only while a native `ADDRESS` callback is active. They resolve variables in this order:

1. Direct scalar `ADDRESS ... EXPOSE name` bindings.

2. Direct stem/array `ADDRESS ... EXPOSE name[]` bindings.
3. The active `ADDRESS ... SANDBOX pool` object.
4. The request's standard sandbox fallback.

The host gets one API, while the CREXX script chooses whether a command uses direct exposure or sandbox storage.

Compound names such as `FILENAME.1` are mapped to exposed stems by splitting at the first dot and using the remaining text as the stem key. Names are matched case-insensitively at this facade layer to fit legacy host expectations. The standard CREXX sandbox already normalises keys internally.

For compatibility with command processors that treat a scalar result as a one-item stem, an exposed scalar also has a limited compound view:

- `name.0` reads as `"1"`
- `name.1` reads the scalar value
- writes to `name.0` are accepted and ignored
- writes to any other `name.tail` update the scalar value

This rule is applied only when no exposed stem with the same base name exists. Real `EXPOSE name[]` stems keep their normal stem semantics.

The variable facade is not a general variable-pool enumerator and does not provide arbitrary access to unexposed CREXX locals.

9.6 Source cache

`crexxsaa_run_source()` compiles source through `rxc` and `rxas`, then stores the resulting `.rxbin` in a disposable cache. Normal source edits, CREXX rebuilds, compiler path changes, and library rebuilds cause a recompile without manual cache clearing.

Cache location rules:

- `CREXXSAA_CACHE_DIR` overrides the platform default
- a host may call `crexxsaa_set_cache_dir()`
- if both `CREXXSAA_CACHE_DIR` and a host-provided cache directory are present, `CREXXSAA_CACHE_DIR` wins

- platform defaults:
  - macOS: `$HOME/Library/Caches/crexx/crexxsaa`
  - Windows: `%LOCALAPPDATA%\crexx\crexxsaa`, falling back under `%USERPROFILE%\AppData\`
  - other Unix-like systems: `$XDG_CACHE_HOME/crexx/crexxsaa`, falling back to `$HOME/.cache/crexx/crexxsaa`

Runtime cache controls:

- `CREXXSAA_CACHE_DISABLE=1`: compile source through temporary files only
- `CREXXSAA_CACHE_REFRESH=1`: ignore a valid cache hit and replace it
- `CREXXSAA_CACHE_TRACE=1`: write cache decisions to `stderr`

The trace currently emits events such as `miss`, `hit`, `stale`, `refresh`, and `disabled`.

Compiler path controls:

- hosts normally call `crexxsaa_set_compiler(ctx, rxc, rxas, import_dir)`
- `CREXXSAA_RXC`, `CREXXSAA_RXAS`, and `CREXXSAA_IMPORT_DIR` override the configured compiler values

9.7 Cache maintenance tool

The CREXX build/install includes a `crexxsaa` maintenance binary in the normal `bin` directory. It is a troubleshooting tool for the compiled-script cache, not a script runner.

Common commands:

```
crexxsaa --location
crexxsaa --list
crexxsaa --clear
crexxsaa --cache-dir /tmp/crexxsaa-cache --list
crexxsaa --cache-dir /tmp/crexxsaa-cache --clear --list
```

Default invocation with no arguments prints the cache location and lists cache entries. `--location` alone prints only the resolved cache directory.

Example list output:

```
cache: /Users/adrian/Library/Caches/crexx/crexxsaa
source: /path/to/profile.the
  bucket: 0379ad70148bf7ca
```

```
rxbin: /Users/adrian/Library/Caches/crexx/crexxsaa/v1/0379
    ad70148bf7ca/b0ec3f1ffcc51e4c.rxbin
rxbin_size: 1216
source_hash: 9100aa6921615883
config_hash: 5af0757a39c4eba1
```

9.8 Technical cache layout

The cache schema is versioned. Current entries live under `v1`:

```
<cache-root>/v1/<source-key>/
  manifest
  <object-hash>.rxbin
```

`source-key` is an FNV-1a 64-bit hash rendered as 16 hex characters. It includes the host namespace and canonical source path when available.

`object-hash` is also an FNV-1a 64-bit hash rendered as 16 hex characters. It includes the source content hash and compiler/library configuration hash.

Manifest fields:

```
version=1
source_path=/absolute/or/supplied/source/path
source_hash=<16-hex-content-hash>
source_size=<bytes>
source_mtime=<mtime>
config_hash=<16-hex-config-hash>
rxbin=<absolute/cache/path/to/object.rxbin>
```

The content hash, not only the timestamp, decides whether source changed. Size and mtime are recorded for diagnostics and human inspection.

The configuration hash includes:

- cache schema and `CREXXSAA_ABI_VERSION`
- configured or environment-overridden `rxc` path
- configured or environment-overridden `rxas` path
- configured or environment-overridden import directory
- loaded `library.rxbin` path
- file size and mtime for compiler, assembler, import directory, and library path where the platform can stat them

Compile and update sequence:

- compute source hash and configuration hash
- read the source bucket manifest if it exists
- on a valid hit, load the cached `.rxbin`
- on miss, stale, or refresh:
  - run `rxc -i <import-dir> -o <temp-base> <source>`
  - run `rxas -o <temp-base> <temp-base>.rxas`
  - rename the temporary `.rxbin` into the source bucket
  - write a new manifest through a temporary file and rename it into place
  - remove the previous cached `.rxbin` for the bucket when it has been superseded

Invalidation:

- `crexxsaa_invalidate_source(ctx, source_path, namespace)`
- `crexxsaa_invalidate_all(ctx)`
- `crexxsaa_clear_cache(cache_dir_override)`
- `crexxsaa --clear`
- `crexxsaa --cache-dir DIR --clear`

The cache is disposable. Deleting it should affect performance only, not program semantics. The current implementation uses atomic file replacement for compiled objects and manifests, but it is not a full cross-process locking protocol. For troubleshooting, clear the cache while the host application is idle.

10 Compiler exits

This chapter introduces CREXX compiler exits: what they are, how to enable them, and how to use them effectively. It also shows concrete examples, including how variable shadowing works inside exit-generated code.

10.1 What are compiler exits?

A compiler exit is a compile-time hook written in REXX that can inspect tokens and inject replacement code into the current compilation unit. Exits are packaged as a module that the compiler (`rxc`) loads on demand.

This chapter primarily describes the current general-exit model. The ongoing `ADDRESS` / REXXSAA work has also introduced a certified/system-exit model for core-team-curated exits that may own reserved keywords and use a richer compiler-facing planning contract. The working design record is address_rexxsaa_-working.md.

Typical uses:

- Lightweight source transforms for diagnostics and logging (e.g., a `dump` helper that expands into runtime `say` statements).
- Compile-time queries over the token stream (`query` exit).
- Project-specific code generation patterns.

Injected code is grafted into the AST and then compiled like normal source.

10.2 Enabling exits

The compiler looks for an exit bundle (a packed `.rxbin`) in the import path.

- Provide an import path that contains your bundle (e.g., the build `bin` directory):

```
rxc -i /path/to/bin -o out in.rexx
```

- Select the exit module to load using the `RXCP_EXIT_MODULE` environment variable. For the built-in bundle this is `rxcexits`:

```
RXCP_EXIT_MODULE=rxcexits rxc -i /path/to/bin -o out in.rexx
```

Notes:

- The exit module is discovered dynamically at compile time.
- All exits within the bundle are registered (e.g., `dump`, `query`).

10.3 Built-in demo exits

The project ships a small bundle of exits primarily for demonstration and testing:

- `dump` — emits `say` statements to print variables (including arrays) with type information.
- `query` — queries token properties.

You can find their sources under `compiler/exits/` in the repository.

The more elaborate `pprint` exit is a standalone packaged example under `examples/exits/pprint`. It lowers `pprint array,...` to the companion `arrayformatdemo` module under `examples/functions/array-formatting`. Neither the exit nor those formatter procedures are part of the default exit bundle or the Level B/Level G library contract. The example README gives the explicit bundle and import paths.

10.4 Certified/System Exits

Current user exits are intended for non-reserved keywords. The compiler now also has certified/system exits for core-team-policed language features.

Current certified exits:

- `ADDRESS`
- `PARSE`

Key distinctions in the current model:

- Certified/system exits may own reserved keywords.
- General user exits should not be able to override those reserved bindings.
- Certified/system exits may use a richer planning response than the legacy string-only `pre_process()` contract.

Treat the working note as the design authority for the evolving model: address_-rexxsaa_working.md.

10.5 Example: Dumping variables and arrays

Given

```
options levelb
import rxfnsb
main: procedure
  a = .int[5]
  a[1] = 42
  dump a
  return 0
```

bundle:

```
RXCP_EXIT_MODULE=rxcexits rxc -i ./cmake-build-debug/bin -o test ./path
    /to/file.rexx
```

At compile time, `dump a` is replaced with a small loop that prints every element:

```
a (.int[5])[1] = 42
a (.int[5])[2] = 0
a (.int[5])[3] = 0
a (.int[5])[4] = 0
a (.int[5])[5] = 0
```

10.6 Shadowing and scoping inside exits

Exit replacements are grafted back into the main AST and then normalized like ordinary source. Structured replacements are supported, including nested `DO ... END`, `IF ... THEN/ELSE`, and nested instruction blocks. Local scopes for those generated blocks are materialized during the normal symbol-building passes, so identifiers created by injected code do not collide with or overwrite variables in the host program.

Practical consequences:

- A typed declaration inside an exit fragment (e.g., `i = .int`) creates an exit-local variable that is not visible outside the replacement block.
- If the host program already has a variable named `i`, it is not affected by the exit's internal `i`.

Illustration (conceptual)

```
options levelb
main: procedure
  i = .int; i = 7
  dump i   -- exit may internally use a loop variable named i
  say i    -- still prints 7; host i is untouched
  return 0
```

Behind the scenes, any generated grouped blocks are compiled with their own local scopes so that temporary loop counters and helper variables remain confined to the exit-generated block structure.

10.7 Writing your own exit (overview)

An exit is a Rexx module that exports a `process` procedure taking a token descriptor and returning a replacement string or an empty string.

High-level shape

```
namespace myexits expose process

process: procedure
  arg ti = .token
  -- Inspect ti.get_type(), ti.get_text(), ti.get_value_type(), ...
  -- Optionally build and return source replacement text
  return ""
```

Package your exit module into a bundle (`.rxbin`) and place it in the compiler's import path, then set `RXCP_EXIT_MODULE` to the bundle name.

10.8 Planning Hooks

Current behaviour:

- `pre_process()` is used for early structural planning such as hoisting implicit variables.
- General exits may still communicate that planning back to the compiler as a space-separated list of 1-based token indices.
- Certified exits can now return a structured `rxcp.exitplan` object for binding hoists, contextual keyword claims, and planning diagnostics.

Planned direction:

- the Stage 1 `ADDRESS` work continues to extend the structured planning model with richer typed planning data and broader address-environment contracts.

See the working note for the proposal details: address_rexxsaa_working.md.

Tips

- Prefer uncommon temporary names and make typed declarations for any locals you introduce in the replacement.
- For detailed compiler diagnostics, use `rxc -d2` or `rxc -d3`, but redirect both stdout and stderr to a temp log and inspect the log with `tail`, `sed`, or `grep` rather than streaming the raw trace to the terminal.

10.9 Troubleshooting

- `EXIT_BRIDGE_DISPATCH_FAILED` — The compiler could not find a handler for the exit keyword. Ensure your bundle is on the import path and `RXCP_EXIT_MODULE` names it correctly.
- Replacement compiles but behaves oddly — verify your replacement string is valid Level B Rexx and, when using loops, that all loop variables are de-

clared locally inside the replacement (typed) so they do not depend on host context.

PART II

Tools Reference

11

The crexx tool

The `crexx` tool is the convenience driver for common CREXX workflows. It can compile, assemble, execute, link, and package programs without requiring the user to call each toolchain binary by hand.

It is also useful in larger builds because it keeps the release defaults in one place: source-level defaults are delegated to `rxc` by file type, native packaging runs through `rxlink` before `rxcpack`, and source/binary import paths are passed to the compiler phase consistently.

11.1 Use cases

- Executing a simple script with REXX statements and built-in functions, without having to run the tools in the chain individually and having to specify files and options on each tool
- Using plugins and class libraries with minimal overhead in programs
- Combining multiple source files containing functions and classes and executing them as a unit
- Building a larger application using separate compilation and linking
- Developing Class libraries[5] together with their consumers
- See which code is produced for the CREXX virtual machine and tools using verbosity levels

[5] or function libraries

11.2 Options

Options are used to differentiate between the choices that can be made while building a program. All options have defaults so that they can be left out in standard cases.

11.3 Options description

The following options are available (single and double dashes work for all options):

`-help` Display help on the usage of this tool.

`-version` Display the version of this tool. This is the same as the compiler and interpretr version.

`-exec` Execute the compiled `.rxbin` under `rxvme` (default).

`-args` Stop `crexx` option parsing; all remaining command-line arguments are passed to the executed program as separate argv entries. Use this as the final driver option when program arguments contain spaces or characters that a shell would normally interpret.

`-noexec` Compile only; do not execute the resulting `.rxbin`.

`-compile` Compile all CREXX program files on the command line to `.rxbin` files (default).

`-nocompile` Skip the `rxc` and `rxas` phases and reuse an existing `<stem>.rxbin`.

`-native` Compile to a native executable; default `--nonative`. This produces an executable file for the current operating system and instruction set architecture. The native route now links the compiled program with `rxlink` before `rxcpack` generates C source.

`-nonative` Disable native packaging.

`-verbose[0-4]` Report on progress; default verbose0, which only issues error messages when the compile fails. Verbose 2 shows the rendered argv used for the toolchain utilities rxc, rxas and rxvme. Verbose 3 includes options and source listings, while verbose 4 includes the contents of the generated assembly code.

`-[no]colo[u]r` Enable or disable colourized progress output.

`-[no]optimize` Enable or disable optimization.

`-keep` Keep compile/link intermediates (default).

`-nokeep` Delete compile/link intermediates after the run.

`-decimal` Use decimal arithmetic where the driver has to choose arithmetic mode.

`-l[library path]` Load a binary/runtime library. A value containing a directory component is used as the exact file path; a bare packaged name remains relative to `CREXX_HOME/bin`. This permits applications to name an approved library file without searching runtime paths. Runtime/native library loading is separate from the compiler's `-s` and `-i` import-discovery paths.

For native packaging, `crexx` now separates `-l` inputs into two groups:

- packaged CREXX libraries (`.rxbin` inputs such as `classlib`) are passed into `rxlink`
- plugin/static-native libraries are linked natively when a matching platform static library is available

This keeps the direct interpreter path fast while still producing compact native executables.

`-s[path]` or `--source path` Add an extra source import root for the `rxc` phase. This is for off-directory `.rexx` modules that should be visible to source import discovery.

`-i[path]` Add an extra raw binary import root for the `rxc` phase. This is for `.rxbin` imports discovered during compilation.

`--import-rxas` Allow the `rxc` phase to auto-import `.rxas` files from binary roots. This is off by default.

`--linkmap path` When using `-native`, ask `rxlink` to write a link map.

`--link-keep-source` When using `-native`, keep source/TRACE debug metadata in the linked intermediate instead of using the default stripped output.

`--link-keep-inline` When using `-native`, keep inline-body metadata in the linked intermediate. The native link strips this metadata by default because it is only needed by later compiler imports and debugging/tooling checks.

`-s`, `-i`, and `--import-rxas` are compile-time controls only. They do not automatically add user runtime modules to `rxvme` or to native links. For runtime/native library loading, continue to use `-l`.

Source files without an `OPTIONS` clause use the `rxc` file-type defaults: `.rexx` defaults to Level C Classic REXX, while `.crexx` and `.crx` default to Level G. Explicit

`OPTIONS LEVELC` scripts use the normal source header and the driver's standard runtime module set, including the Classic compatibility runtime `rxfnsc`. Level C compilation is incremental: supported Classic Rexx shapes lower and run, while constructs outside the implemented slice are rejected with an unsupported-shape diagnostic.

The driver invokes its toolchain phases through the CREXX ADDRESS command environment using direct argv dispatch. That avoids platform shell parsing for normal compile, assemble, link, pack, native-compile, and execute steps while keeping verbose output readable.

11.4 Examples

11.4.1 Just run it

The simplest way to run a crexx program is to just specify its source file as input to the `crexx` program. It will execute the compiler, the assembler and start it with the standard threaded runtime interpreter. All included libraries and plugins are linked automatically.

```
options levelb
import rxfnsb

say 'hello crexx!'
say 'today''s date is:' date()
```

```
hello crexx!
today's date is: 26 Jul 2026
```

11.4.2 Make an executable module

crexx can make a native executable from a Rexx program. For this, the `rxlink` utility will be called, to arrange the modules into an `.rxbin` library, while insuring that there is no duplication of resources. This `.rxbin` will be processed by the `rxcpack` program, after which the `gcc` or `clang` compilers and the os linkage editor will do their work.

The interesting part here is that all this is accomplished by the `--native` option on the crexx utility. When we take the previous program and compile it with:

```
crexx hello --native
```

then we will have a `hello` program (`hello.exe` on windows) which can be executed by specifying its name in the shell, or even by double clicking it in the File Explorer, Finder or equivalent OS GUI interface. This program can be executed on machines without a CREXX installation.

11.4.3 Compile and run a program with commandline arguments

When using the CREXX compiler driver, commandline arguments as input to the program can be specified with the `--args` option; *this option needs to be the last option on the CREXX command*, because everything that follows will be sent to the program as a string array (`.string[]`).

11.5 Verbosity

With the default verbosity level the tools behaves in the standard unix way where a lack of messages indicates success. This level can be increased gradually to a full explanation of everything that is done. The default is `--verbose0`, which gives no inidcation of what happened unless something went wrong. This is the way to run known-good programs without any overhead.

All verbosity level examples run the following short script:

```
options levelb
import rxfnsb
say 'hello rexx!'
say 'today it''s' date('w')
'echo 42'
```

11.6 `--verbose1`

With `--verbose1`, the driver tells in a very condensed way what it did and how it went. When the return codes from the `rxc` and `rxas` are 0, these are displayed with an 'OK' between square brackets.

```
[ OK ] rxc - Compiled hello.crexx
[ OK ] rxas - Assembled hello.rxas
hello rexx!
today it's Sunday
42
```

It issues some reassuring messages about the compiler and the assembler running successfully and skips the starting of the runtime engine, because the output of the program follows these messages.

11.7 `--verbose2`

With the `--verbose2` setting, there is more tourist information.

```
CREXX compiler driver crexx-1.0.0-beta.3+local.g78c83e4a0b88 (macOS 64
20260726)
cREXX License (MIT)
Copyright (c) 2020-2026 Adrian Sutherland, Peter Jacob, René Jansen
See https://github.com/adesutherland/CREXX for project details
using relpath : /Users/rvjansen/apps/crexx_release/
using lpath : bin
using s roots :
using i roots :
rxc command : "/Users/rvjansen/apps/crexx_release/bin/rxc" "-i" "/User-
s/rvjansen/apps/crexx_release/bin" "-o" "hello" "hello.crexx"
[ OK ] rxc - Compiled hello.crexx
rxas command : "/Users/rvjansen/apps/crexx_release/bin/rxas" "-o" "hello"
"hello"
[ OK ] rxas - Assembled hello.rxas
crexx executes: "/Users/rvjansen/apps/crexx_release/bin/rxvme" "hello"
"/Users/rvjansen/apps/crexx_release/bin/classlib" "/Users/rvjansen/app-
s/crexx_release/bin/classlib_native" "/Users/rvjansen/apps/crexx_release/bin/rxfnsc"
"/Users/rvjansen/apps/crexx_release/bin/rexxscript" "/Users/rvjansen/app-
s/crexx_release/bin/rx_treemap" "/Users/rvjansen/apps/crexx_release/bin/rx_-
```

```
system" "/Users/rvjansen/apps/crexx_release/bin/rx_llist" "/Users/rvjansen/app-
s/crexx_release/bin/rx_keyaccess" "/Users/rvjansen/apps/crexx_release/bin/rx_-
id" "-a"
hello rexx!
today it's Sunday
42
```

It starts by identifying the exact release of the crexx version. After this, a number of paths are shown:

Name	Meaning
relpath	The path where the CREXX system is installed
lpath	The path from which executables and libraries are found
s roots	additional sourcefile lookup locations
i roots	additional binary (.rxbin) lookup locations

Source-level defaults are owned by `rxc`, so the driver does not insert language or import flags for headerless files. The verbose output shows the exact `rxc` command and the source extension passed to the compiler.

With this verbosity level, the exact invocations of the tools in the toolchain are documented, including the complete paths and arguments. Also, the invocation of the runtime engine, with all libraries explicity named - and this includes the libraries that are not used. With native compiles, the `rxlink` tool will extract the used code from these libraries into the executable. It also shows the `-a` flag as a last option, even if this is unused.

11.8 --verbose3

In verbosity level 3, the output level is on par with traditional IBM compiler output, with a complete summary of used and unused compiler options.

```
CREXX compiler driver crexx-1.0.0-beta.3+local.g78c83e4a0b88 (macOS 64
20260726)
cREXX License (MIT)
Copyright (c) 2020-2026 Adrian Sutherland, Peter Jacob, René Jansen
```

```
See https://github.com/adesutherland/CREXX for project details
======> RexxLA cREXX compiler crexx-1.0.0-beta.3+local.g78c83e4a0b88 26
Jul 2026 15:31:36
INVOCATION OPTIONS
Options in effect:
NOCOLOR
DECIMAL
EXEC
NONATIVE
KEEP
OPTIMIZE
NOIMPORT-RXAS
VERBOSE 3
SOURCE ROOTS
BINARY ROOTS
using relpath : /Users/rvjansen/apps/crexx_release/
using lpath : bin
using s roots :
using i roots :
rxc command : "/Users/rvjansen/apps/crexx_release/bin/rxc" "-i" "/User-
s/rvjansen/apps/crexx_release/bin" "-o" "hello" "hello.crexx"
[ OK ] rxc - Compiled hello.crexx
rxas command : "/Users/rvjansen/apps/crexx_release/bin/rxas" "-o" "hello"
"hello"
[ OK ] rxas - Assembled hello.rxas
======> RexxLA cREXX compiler crexx-1.0.0-beta.3+local.g78c83e4a0b88 26
Jul 2026 15:31:36
crexx executes: "/Users/rvjansen/apps/crexx_release/bin/rxvme" "hello"
"/Users/rvjansen/apps/crexx_release/bin/classlib" "/Users/rvjansen/app-
s/crexx_release/bin/classlib_native" "/Users/rvjansen/apps/crexx_release/bin/rxfnsc"
"/Users/rvjansen/apps/crexx_release/bin/rexxscript" "/Users/rvjansen/app-
s/crexx_release/bin/rx_treemap" "/Users/rvjansen/apps/crexx_release/bin/rx_-
system" "/Users/rvjansen/apps/crexx_release/bin/rx_llist" "/Users/rvjansen/app-
s/crexx_release/bin/rx_keyaccess" "/Users/rvjansen/apps/crexx_release/bin/rx_-
id" "-a"
```

```
hello rexx!
today it's Sunday
42
```

11.9 --verbose4

Verbosity level 4 is for the moment the most verbose level of tool output. In this level, the complete Rexx input source is expanded (when it is produced by the rxpp preprocessor), and all generated assembler output is visible and available for inspection and debugging, as well as the set of generated import statements for runtime linking. It is advisable to page this output through a less-like processor.

```
CREXX compiler driver crexx-1.0.0-beta.3+local.g78c83e4a0b88 (macOS 64
20260726)
cREXX License (MIT)
Copyright (c) 2020-2026 Adrian Sutherland, Peter Jacob, René Jansen
See https://github.com/adesutherland/CREXX for project details
======> RexxLA cREXX compiler crexx-1.0.0-beta.3+local.g78c83e4a0b88 26
Jul 2026 15:31:36
INVOCATION OPTIONS
Options in effect:
NOCOLOR
DECIMAL
EXEC
NONATIVE
KEEP
OPTIMIZE
NOIMPORT-RXAS
VERBOSE 4
SOURCE ROOTS
BINARY ROOTS
using relpath : /Users/rvjansen/apps/crexx_release/
using lpath : bin
using s roots :
using i roots :
```

```
======> RexxLA cREXX compiler crexx-1.0.0-beta.3+local.g78c83e4a0b88 26
Jul 2026 15:31:36
LINENUM ----+----+----+----+----+----+----+----+----+----+----+----+----
+----+----+
000001 options levelb
000002 import rxfnsb
000003 say 'hello rexx!'
000004 say 'today it''s' date('w')
000005 'echo 42'
rxc command : "/Users/rvjansen/apps/crexx_release/bin/rxc" "-i" "/User-
s/rvjansen/apps/crexx_release/bin" "-o" "hello" "hello.crexx"
[ OK ] rxc - Compiled hello.crexx
======> RexxLA cREXX compiler crexx-1.0.0-beta.3+local.g78c83e4a0b88 26
Jul 2026 15:31:37
LINENUM ----+----+----+----+----+----+----+----+----+----+----+----+----
+----+----+
000001 /*
000002 * SOURCE : hello.crexx
000003 */
000004
000005 .globals=0
000006
000007 main() .locals=21
000008 .meta "hello.main"="b" ".void" main() ""
000009 numsci 18,1,1
000010 .meta "hello.main.rc"="b" ".int" r0
000011 .srcstep 1 1 6 "hello.crexx" 5 1 10 "'echo 42'"
000012 null r0
000013 .srcstep 2 2 17 "hello.crexx" 3 1 18 "say 'hello rexx!'"
000014 say "hello rexx!"
000015 .srcstep 37 37 17 "hello.crexx" 4 1 28 "say 'today it''s' date('w')"
000016 .srcstep 36 36 17 "hello.crexx" 4 19 23 "say 'today it''s' date('w')"
000017 .meta "hello.main.input_separator_length"="b" ".int" r1
000018 .srcstep 3 3 25 "date.crexx" 47 3 32 " input_separator_length =
.int"
```

```
000019 null r1
000020 .meta "hello.main.jdn"="b" ".int" r3
000021 .srcstep 4 4 25 "date.crexx" 48 3 13 " jdn = .int"
000022 null r3
000023 .meta "hello.main.output_separator_length"="b" ".int" r5
000024 .srcstep 5 5 25 "date.crexx" 46 3 33 " output_separator_length =
.int"
000025 null r5
000026 .srcstep 6 6 33 "hello.crexx" 4 23 23 "say 'today it''s' date('w')"
000027 load r4,""
000028 .traceevent "L" 4 "R" "S" "r" 4 6 6 0 "" ""
000029 .traceevent "A" 6 "R" "S" "r" 4 6 6 0 "osep" ""
000030 .srcstep 7 7 33 "hello.crexx" 4 23 23 "say 'today it''s' date('w')"
000031 load r2,""
000032 .traceevent "L" 4 "R" "S" "r" 2 7 7 0 "" ""
000033 .traceevent "A" 6 "R" "S" "r" 2 7 7 0 "isep" ""
000034 .srcstep 8 8 25 "date.crexx" 50 13 48 " assembler strlen output_-
separator_length,osep"
000035 .traceevent "V" 6 "R" "I" "r" 5 37 37 0 "output_separator_length"
""
000036 .traceevent "V" 6 "R" "S" "r" 4 37 37 0 "osep" ""
000037 strlen r5,r4
000038 .srcstep 9 9 25 "date.crexx" 51 13 47 " assembler strlen input_-
separator_length,isep"
000039 .traceevent "V" 6 "R" "I" "r" 1 37 37 0 "input_separator_length" ""
000040 .traceevent "V" 6 "R" "S" "r" 2 37 37 0 "isep" ""
000041 strlen r1,r2
000042 .traceevent "V" 6 "R" "I" "r" 5 37 37 0 "output_separator_length"
""
000043 igt r6,r5,1
000044 .traceevent "O" 4 "R" "B" "r" 6 37 37 0 "" ""
000045 brt l361orend,r6
000046 .traceevent "V" 6 "R" "I" "r" 1 37 37 0 "input_separator_length" ""
000047 igt r6,r1,1
000048 .traceevent "O" 4 "R" "B" "r" 6 37 37 0 "" ""
```

```
000049 l361orend:
000050 .traceevent "O" 4 "R" "B" "r" 6 37 37 0 "" ""
000051 brf l360iffalse,r6
000052 .srcstep 10 10 25 "date.crexx" 53 15 98 " assembler signal ÏNVALID_-
ARGUMENTS,ÐATE separators must be empty or one Unicode codepoint"¨
000053 signal "INVALID_ARGUMENTS","DATE separators must be empty or one
Unicode codepoint"
000054 l360iffalse:
000055 .meta "hello.main.adjusted_second"="b" ".int" r7
000056 .srcstep 11 11 25 "date.crexx" 72 3 25 " adjusted_second = .int"
000057 null r7
000058 .meta "hello.main.adjustment"="b" ".int" r8
000059 .srcstep 12 12 25 "date.crexx" 71 3 20 " adjustment = .int"
000060 null r8
000061 .meta "hello.main.after_microseconds"="b" ".int" r9
000062 .srcstep 13 13 25 "date.crexx" 66 3 28 " after_microseconds = .int"
000063 null r9
000064 .meta "hello.main.before_microseconds"="b" ".int" r10
000065 .srcstep 14 14 25 "date.crexx" 65 3 29 " before_microseconds = .int"
000066 null r10
000067 .meta "hello.main.local_second_of_day"="b" ".int" r11
000068 .srcstep 15 15 25 "date.crexx" 70 3 29 " local_second_of_day = .int"
000069 null r11
000070 .meta "hello.main.standard_day"="b" ".int" r12
000071 .srcstep 16 16 25 "date.crexx" 68 3 22 " standard_day = .int"
000072 null r12
000073 .meta "hello.main.standard_second_of_day"="b" ".int" r13
000074 .srcstep 17 17 25 "date.crexx" 69 3 32 " standard_second_of_day =
.int"
000075 null r13
000076 .meta "hello.main.standard_seconds"="b" ".int" r14
000077 .srcstep 18 18 25 "date.crexx" 67 3 26 " standard_seconds = .int"
000078 null r14
000079 .srcstep 19 19 25 "date.crexx" 74 13 38 " assembler mtime before_-
microseconds"
```

```
000080 .traceevent "V" 6 "R" "I" "r" 10 37 37 0 "before_microseconds" ""
000081 mtime r10
000082 .srcstep 20 20 25 "date.crexx" 75 13 34 " assembler time standard_-
seconds"
000083 .traceevent "V" 6 "R" "I" "r" 14 37 37 0 "standard_seconds" ""
000084 time r14
000085 .srcstep 21 21 25 "date.crexx" 76 13 37 " assembler mtime after_-
microseconds"
000086 .traceevent "V" 6 "R" "I" "r" 9 37 37 0 "after_microseconds" ""
000087 mtime r9
000088 .traceevent "V" 6 "R" "I" "r" 9 37 37 0 "after_microseconds" ""
000089 .traceevent "V" 6 "R" "I" "r" 10 37 37 0 "before_microseconds" ""
000090 ilt r6,r9,r10
000091 .traceevent "O" 4 "R" "B" "r" 6 37 37 0 "" ""
000092 brf l406iffalse,r6
000093 .srcstep 22 22 25 "date.crexx" 77 62 83 " if after_microseconds <
before_microseconds then assembler time standard_seconds"
000094 .traceevent "V" 6 "R" "I" "r" 14 37 37 0 "standard_seconds" ""
000095 time r14
000096 l406iffalse:
000097 .srcstep 23 23 25 "date.crexx" 79 3 42 " standard_day = standard_-
seconds 000098 .traceevent "V" 6 "R" "I" "r" 14 23 23 0 "standard_seconds"
""
000099 idiv r12,r14,86400
000100 .traceevent "O" 4 "R" "I" "r" 12 23 23 0 "" ""
000101 .traceevent "A" 6 "R" "I" "r" 12 23 23 0 "standard_day" ""
000102 .srcstep 24 24 25 "date.crexx" 80 3 53 " standard_second_of_day =
standard_seconds // 86400"
000103 .traceevent "V" 6 "R" "I" "r" 14 24 24 0 "standard_seconds" ""
000104 imod r13,r14,86400
000105 .traceevent "O" 4 "R" "I" "r" 13 24 24 0 "" ""
000106 .traceevent "A" 6 "R" "I" "r" 13 24 24 0 "standard_second_of_day"
""
000107 .traceevent "V" 6 "R" "I" "r" 13 37 37 0 "standard_second_of_day"
""
```

```
000108 ilt r6,r13,0
000109 .traceevent "O" 4 "R" "B" "r" 6 37 37 0 "" ""
000110 brf l422iffalse,r6
000111 .srcstep 25 25 25 "date.crexx" 82 5 60 " standard_second_of_day =
standard_second_of_day + 86400"
000112 .traceevent "V" 6 "R" "I" "r" 13 25 25 0 "standard_second_of_day"
""
000113 iadd r13,r13,86400
000114 .traceevent "O" 4 "R" "I" "r" 13 25 25 0 "" ""
000115 .traceevent "A" 6 "R" "I" "r" 13 25 25 0 "standard_second_of_day"
""
000116 .srcstep 26 26 25 "date.crexx" 83 5 36 " standard_day = standard_day
- 1"
000117 .traceevent "V" 6 "R" "I" "r" 12 26 26 0 "standard_day" ""
000118 isub r12,r12,1
000119 .traceevent "O" 4 "R" "I" "r" 12 26 26 0 "" ""
000120 .traceevent "A" 6 "R" "I" "r" 12 26 26 0 "standard_day" ""
000121 l422iffalse:
000122 .srcstep 27 27 25 "date.crexx" 85 3 53 " local_second_of_day =
after_microseconds 000123 .traceevent "V" 6 "R" "I" "r" 9 27 27 0 "after_-
microseconds" ""
000124 idiv r11,r9,1000000
000125 .traceevent "O" 4 "R" "I" "r" 11 27 27 0 "" ""
000126 .traceevent "A" 6 "R" "I" "r" 11 27 27 0 "local_second_of_day" ""
000127 .srcstep 28 28 25 "date.crexx" 86 3 60 " adjustment = local_second_-
of_day - standard_second_of_day"
000128 .traceevent "V" 6 "R" "I" "r" 11 28 28 0 "local_second_of_day" ""
000129 .traceevent "V" 6 "R" "I" "r" 13 28 28 0 "standard_second_of_day"
""
000130 isub r8,r11,r13
000131 .traceevent "O" 4 "R" "I" "r" 8 28 28 0 "" ""
000132 .traceevent "A" 6 "R" "I" "r" 8 28 28 0 "adjustment" ""
000133 .traceevent "V" 6 "R" "I" "r" 8 37 37 0 "adjustment" ""
000134 igt r6,r8,43200
000135 .traceevent "O" 4 "R" "B" "r" 6 37 37 0 "" ""
```

```
000136 brf l447iffalse,r6
000137 .srcstep 29 29 25 "date.crexx" 87 30 61 " if adjustment > 43200
then adjustment = adjustment - 86400"
000138 .traceevent "V" 6 "R" "I" "r" 8 29 29 0 "adjustment" ""
000139 isub r8,r8,86400
000140 .traceevent "O" 4 "R" "I" "r" 8 29 29 0 "" ""
000141 .traceevent "A" 6 "R" "I" "r" 8 29 29 0 "adjustment" ""
000142 br l447ifend
000143 l447iffalse:
000144 .traceevent "V" 6 "R" "I" "r" 8 37 37 0 "adjustment" ""
000145 ilt r15,r8,-43200
000146 .traceevent "O" 4 "R" "B" "r" 15 37 37 0 "" ""
000147 brf l456iffalse,r15
000148 .srcstep 30 30 25 "date.crexx" 88 36 67 " else if adjustment <
-43200 then adjustment = adjustment + 86400"
000149 .traceevent "V" 6 "R" "I" "r" 8 30 30 0 "adjustment" ""
000150 iadd r8,r8,86400
000151 .traceevent "O" 4 "R" "I" "r" 8 30 30 0 "" ""
000152 .traceevent "A" 6 "R" "I" "r" 8 30 30 0 "adjustment" ""
000153 l456iffalse:
000154 l447ifend:
000155 .srcstep 31 31 25 "date.crexx" 90 3 56 " adjusted_second = standard_-
second_of_day + adjustment"
000156 .traceevent "V" 6 "R" "I" "r" 13 31 31 0 "standard_second_of_day"
""
000157 .traceevent "V" 6 "R" "I" "r" 8 31 31 0 "adjustment" ""
000158 iadd r7,r13,r8
000159 .traceevent "O" 4 "R" "I" "r" 7 31 31 0 "" ""
000160 .traceevent "A" 6 "R" "I" "r" 7 31 31 0 "adjusted_second" ""
000161 .traceevent "V" 6 "R" "I" "r" 7 37 37 0 "adjusted_second" ""
000162 ilt r15,r7,0
000163 .traceevent "O" 4 "R" "B" "r" 15 37 37 0 "" ""
000164 brf l470iffalse,r15
000165 .srcstep 32 32 25 "date.crexx" 91 31 62 " if adjusted_second < 0
then standard_day = standard_day - 1"
```

```
000166 .traceevent "V" 6 "R" "I" "r" 12 32 32 0 "standard_day" ""
000167 isub r12,r12,1
000168 .traceevent "O" 4 "R" "I" "r" 12 32 32 0 "" ""
000169 .traceevent "A" 6 "R" "I" "r" 12 32 32 0 "standard_day" ""
000170 br l470ifend
000171 l470iffalse:
000172 .traceevent "V" 6 "R" "I" "r" 7 37 37 0 "adjusted_second" ""
000173 igte r6,r7,86400
000174 .traceevent "O" 4 "R" "B" "r" 6 37 37 0 "" ""
000175 brf l479iffalse,r6
000176 .srcstep 33 33 25 "date.crexx" 92 41 72 " else if adjusted_second
>= 86400 then standard_day = standard_day + 1"
000177 .traceevent "V" 6 "R" "I" "r" 12 33 33 0 "standard_day" ""
000178 iadd r12,r12,1
000179 .traceevent "O" 4 "R" "I" "r" 12 33 33 0 "" ""
000180 .traceevent "A" 6 "R" "I" "r" 12 33 33 0 "standard_day" ""
000181 l479iffalse:
000182 l470ifend:
000183 .srcstep 34 34 41 "date.crexx" 93 3 9 " return standard_day +
2440588"
000184 .traceevent "V" 6 "R" "I" "r" 12 34 34 0 "standard_day" ""
000185 iadd r3,r12,2440588
000186 .traceevent "O" 4 "R" "I" "r" 3 34 34 0 "" ""
000187 .traceevent "A" 6 "R" "I" "r" 3 34 34 0 "jdn" ""
000188 .srcstep 35 35 25 "date.crexx" 57 3 34 " return _dateo(jdn,oformat,osep)"
000189 .traceevent "V" 6 "R" "I" "r" 3 37 37 0 "jdn" ""
000190 load r17,"w"
000191 .traceevent "L" 4 "R" "S" "r" 17 37 37 0 "" ""
000192 .traceevent "V" 6 "R" "S" "r" 4 37 37 0 "osep" ""
000193 load r15,3
000194 swap r16,r3
000195 settp r17,256
000196 settpswapcall r6,_rxsysb._dateo(),r15,r4,256,r18
000197 swap r3,r16
000198 swap r4,r18
```

```
000199 .traceevent "F" 4 "R" "S" "r" 6 37 37 0 "_dateo" ""
000200 scopy r18,r6
000201 br l328bexprend
000202 l328bexprend:
000203 .meta "hello.main.input_separator_length"
000204 .meta "hello.main.jdn"
000205 .meta "hello.main.output_separator_length"
000206 sconcat r18,"today itś",r18
000207 .traceevent "O" 4 "R" "S" "r" 18 37 37 0 "" ""
000208 say r18
000209 .srcstep 47 47 33 "exit_fragment" 2 1 56 "rc=_address('','echo
42',_noredir(),_noredir(),_noredir"
000210 load r16,""
000211 .traceevent "L" 4 "R" "S" "r" 16 47 47 0 "" ""
000212 load r17,"echo 42"
000213 .traceevent "L" 4 "R" "S" "r" 17 47 47 0 "" ""
000214 .srcstep 40 40 33 "exit_fragment" 2 1 9 "_noredir"
000215 .meta "hello.main.handle"="b" ".binary" r6
000216 .srcstep 38 38 25 "_address.crexx" 1682 5 21 " handle = .binary"
000217 null r6
000218 .srcstep 39 39 25 "_address.crexx" 1683 5 18 " return handle"
000219 .traceevent "V" 6 "R" "X" "r" 6 47 47 0 "handle" ""
000220 bcopy r5,r6
000221 br l504bexprend
000222 l504bexprend:
000223 endlife r6
000224 .meta "hello.main.handle"
000225 .srcstep 43 43 33 "exit_fragment" 2 1 9 "_noredir"
000226 .meta "hello.main.handle"="b" ".binary" r6
000227 .srcstep 41 41 25 "_address.crexx" 1682 5 21 " handle = .binary"
000228 null r6
000229 .srcstep 42 42 25 "_address.crexx" 1683 5 18 " return handle"
000230 .traceevent "V" 6 "R" "X" "r" 6 47 47 0 "handle" ""
000231 bcopy r4,r6
000232 br l511bexprend
```

```
000233 l511bexprend:
000234 endlife r6
000235 .meta "hello.main.handle"
000236 .srcstep 46 46 33 "exit_fragment" 2 1 9 "_noredir"
000237 .meta "hello.main.handle"="b" ".binary" r6
000238 .srcstep 44 44 25 "_address.crexx" 1682 5 21 " handle = .binary"
000239 null r6
000240 .srcstep 45 45 25 "_address.crexx" 1683 5 18 " return handle"
000241 .traceevent "V" 6 "R" "X" "r" 6 47 47 0 "handle" ""
000242 bcopy r3,r6
000243 br l518bexprend
000244 l518bexprend:
000245 endlife r6
000246 .meta "hello.main.handle"
000247 load r15,5
000248 settp r16,768
000249 settp r17,768
000250 settp r5,768
000251 swap r18,r5
000252 settp r4,768
000253 swap r19,r4
000254 settpswapcall r0,_rxsysb._address(),r15,r3,768,r20
000255 swap r5,r18
000256 swap r4,r19
000257 swap r3,r20
000258 .traceevent "F" 4 "R" "I" "r" 0 47 47 0 "_address" ""
000259 .traceevent "A" 6 "R" "I" "r" 0 47 47 0 "rc" ""
000260 .srcstep 48 48 1 "hello.crexx" 5 10 10 "'echo 42'"
000261 ret
000262 .meta "hello.main.rc"
000263
000264 /* Imported Declaration from file: _datei@library.rxbin */
000265
000266 _rxsysb._datei() .expose=_rxsysb._datei
000267 .meta "_rxsysb._datei"="b" ".int" _rxsysb._datei() "idate=.string,format=.string,?isep
```

```
000268
000269 /* Imported Declaration from file: _dateo@library.rxbin */
000270
000271 _rxsysb._dateo() .expose=_rxsysb._dateo
000272 .meta "_rxsysb._dateo"="b" ".string" _rxsysb._dateo() "jdn=.int,format=.string,?osep=
000273
000274 /* Imported Declaration from file: _address@library.rxbin */
000275
000276 _rxsysb._address() .expose=_rxsysb._address
000277 .meta "_rxsysb._address"="b" ".int" _rxsysb._address() "?env=.string,?command=.string
...=.string"
rxas command : "/Users/rvjansen/apps/crexx_release/bin/rxas" "-o" "hello"
"hello"
[ OK ] rxas - Assembled hello.rxas
======> RexxLA cREXX compiler crexx-1.0.0-beta.3+local.g78c83e4a0b88 26
Jul 2026 15:31:37
crexx executes: "/Users/rvjansen/apps/crexx_release/bin/rxvme" "hello"
"/Users/rvjansen/apps/crexx_release/bin/classlib" "/Users/rvjansen/app-
s/crexx_release/bin/classlib_native" "/Users/rvjansen/apps/crexx_release/bin/rxfnsc"
"/Users/rvjansen/apps/crexx_release/bin/rexxscript" "/Users/rvjansen/app-
s/crexx_release/bin/rx_treemap" "/Users/rvjansen/apps/crexx_release/bin/rx_-
system" "/Users/rvjansen/apps/crexx_release/bin/rx_llist" "/Users/rvjansen/app-
s/crexx_release/bin/rx_keyaccess" "/Users/rvjansen/apps/crexx_release/bin/rx_-
id" "-a"
hello rexx!
today it's Sunday
42
```

11.10 --verbose[2-4] with native compilation

When the program is compiled and linked to a native executable, no execution of this program follows. Instead the arguments to the `rxpack` object packager and the linkage editor `rxlink` and their results are included.

Also, the complete command line to the c-compiler and its linkage editor step is documented here, and can be copied into private building tools or scripts. Here

the use of static import libraries for the native executables can be seen.

In the following chapters, these tools are documented in detail.

12 rxc - the cREXX Compiler

12.1 Command Line Options

```
cREXX Compiler
Version : crexx-1.0.0-beta.3+local.g78c83e4a0b88
Usage : rxc [options] source_file
Options :
-h Prints help message
-c Prints Copyright & License Details
-v Prints Version
-d[level] Debug Mode (Optional Level)
-dp Stop after parsing and validation
-l location Working Location (directory)
-s source Source import locations - ";" delimited list
-i import Binary import locations - ";" delimited list
--level level Default source level when OPTIONS omits one
--import ns Inject a file-level IMPORT namespace (repeatable)
--import-rxas Enable auto-import scanning of .rxas in binary roots
--diagnostics mode Diagnostic rendering: localized or raw
--diagnostic-locale locale Diagnostic locale such as en_GB or en_US
--no-localisation Use raw diagnostic code/parameter rendering
```

```
-o output_stem RXAS output stem or .rxas file
-n No Optimising
-x Disable compiler exits
--syntaxhighlight Run as a DSLSH syntax-highlighting server (stdio
mode)
--port port Run as a DSLSH syntax-highlighting server (socket mode
on port)
```

12.2 Import Search Roots

`rxc` separates automatic import discovery into two root classes:

- source roots, searched for CREXX source files
- binary roots, searched for `.rxbin`, optional `.rxas`, and `.rxplugin`

The recommended source extensions are:

- `.crexx`: canonical CREXX source, defaulting to Level G when no `options level...` clause is present
- `.crx`: short CREXX source alias, also defaulting to Level G
- `.rexx`: compatibility/classic source, defaulting to incremental Level C lowering

An initial source file may also use another extension, such as `.the` for an editor macro integration. That extension is added to source-root discovery for that compile and defaults to Level G unless the source states another level explicitly. This supports host-specific source names without making every possible extension a recommended CREXX extension.

`.rxpp` remains the preprocessor-oriented source extension for existing RXPP workflows. Longer term, preprocessing may become idempotent and better integrated so ordinary `.crexx` and `.crx` sources can pass through it safely, but that is separate from the source import extension rules.

The primary source root is the directory of the source file being compiled. If the source file name does not include a directory, the compiler uses the working location from `-l`. Additional source roots can be supplied with `-s`, using a semicolon-delimited list.

Binary roots are controlled separately with `-i`. The compiler always includes the executable directory in the binary-root set so deployed libraries remain visible without adding them explicitly. `-i` can be repeated, and repeated `-s` options are accumulated in the same way. The source file's directory is not automatically treated as a binary root; pass `-i .` or `-i build-dir` when a sibling or build-output `.rxbin` is an intended compile-time dependency.

Search order is:

1. primary source root
2. extra `-s` source roots
3. binary roots for `.rxbin`
4. binary roots for `.rxas` when `--import-rxas` is enabled
5. binary roots for `.rxplugin`

This keeps same-project source imports preferred over deployed binary artifacts, while stopping the compiler from treating every binary directory as a source tree. It also prevents stale generated `.rxbin` files left beside edited source files from silently satisfying imports.

12.3 Import Discovery Notes

Automatic `.rxas` import scanning is disabled by default. It can be re-enabled with `--import-rxas` for workflows that intentionally use assembler modules as import sources.

For source imports, `rxc` performs a cheap header scan before doing a full parse. That scan reads the leading `options`, `namespace`, and `import` clauses so the compiler can reject namespace-mismatched source files early and avoid unnecessary full validation work.

Within each source root, the compiler looks for `.crexx`, `.crx`, `.rexx`, and the initial file's extension when it is different from those standard names. Source files with the same module stem are de-duplicated within that source stage so one source module is imported for a stem.

Within a single binary root, when multiple artifacts share the same module stem, the compiler keeps only the freshest candidate. If timestamps tie, `.rxbin` wins over `.rxas`.

12.4 Inline Assembler

On page **??** the inline assembler function of the CREXX compiler is discussed. This enables the incorporation of `rxas` assembler instructions into a REXX source file.

12.5 Optimizer

The compiler can do a number of optimizations that can make the execution of a program much faster; the next example shows how an operation can be done at compile time, to avoid instruction scheduling and execution at runtime:

```
options levelb
say 0.5**2 'should be 0.25'
```

```
 0.25 should be 0.25
```

```
/*
 * SOURCE                 : fpowtest.crexx
 */

.globals=0

main() .locals=0
   .meta "fpowtest.main"="b" ".void" main() ""
   numsci 18,1,1
   .srcstep 1 1 17 "fpowtest.crexx" 2 1 28 "say 0.5**2 'should be 0.25'
       "
   say "0.25 should be 0.25"
   .srcstep 2 2 1 "fpowtest.crexx" 2 28 28 "say 0.5**2 'should be 0.25'
       "
   ret
```

This works because, for a large number of operations, the `rxc` compiler can assume the result is never going to be different, and will determine that result during compile time. In the same vein, results from operations that are not displayed or handled further in the program, will lead to the operation being skipped entirely.

12.6 Debugging and Validation

For compiler developers and advanced users, `rxc` provides debug modes that enable internal validation of the AST and Symbol table and stress-test the fixpoint loop.

- `-d2` --- Structural Validator
  - Runs the built-in AST/Symbol validator between passes.
  - Catches parent/child inconsistencies, broken bi-directional symbol links, and illegal scope parentage.
- `-d3` --- Validator + Idempotency Stress Test
  - Everything in `-d2`, plus: the fixpoint loop is forced to run extra iterations and selected walkers are invoked multiple times per iteration.
  - Proves walker idempotency and overall convergence.

Example:

```
# Validate (structure only)
rxc -d2 -o out input.rexx

# Stress idempotency and convergence
rxc -d3 -o out input.rexx
```

Notes:

- These modes are for diagnostics and do not change user-visible semantics.
- All walkers inside the fixpoint loop are designed to be idempotent.
- See also: `compiler/docs/fixpoint_idempotency_and_scope.md` for design details.
- See also: `compiler/docs/testing.md` for information on regression testing and how to update golden files.

13 rxas - the CREXX Assembler

The purpose of the CREXX assembler `rxas` is to translate a text file with `rxvm` assembler instructions to a file with binary contents containing these instructions in their binary, executable form. Its main use is to translate an `rxas` file produced by the CREXX compiler `rxc` to a binary `.rxbin` object module.

13.1 Program Structure

The assembler processing goes through a number of steps in a single pass: first, the Lexer / Scanner tokenises the `.rxas` code. After that, the Parser parses the structure into a series of instructions. The binary writer generates the binary code and constant pool for the program at hand. The backpatcher runs last and handles forward references.

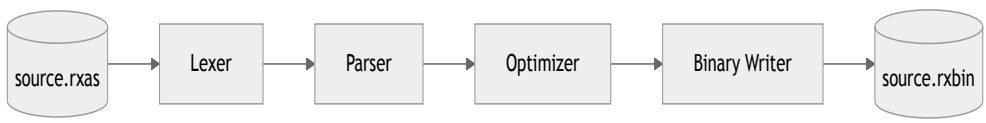


FIGURE 1: Program Structure

13.2 Input/Output

The `rxas` assembler has a `.rxas` file as input and produces an `rxbin` file as output, which can be considered an *object module*, as it has unresolved addresses, which can be resolved by the linkage editor component of the `rxvm` virtual machine interpreter. It also produces a report to stdout (in case of errors only) and can produce a trace file in Debug/verbose mode (option `-d`).

13.3 Character sets

The input file is assumed to be valid UTF8. The assembler, like the compiler, operates using two character sets. The first is for symbols in the assembler language statements. These are all composed of the ASCII subset of Unicode. The second character set is used for data; the contents of variables. Here the whole of Unicode can be used.

13.4 Command Line Arguments

When the command line argument -h is specified the options are shown:

```
cREXX Assembler
Version : crexx-1.0.0-beta.3+local.g78c83e4a0b88
Usage : rxas [options] source_file
Options :
-h Help Message
-c Copyright & License Details
-v Version
-a Architecture Details
-i Print Instructions
-d Debug/Verbose Mode
-l location Working Location (directory)
-o output_stem RXBIN output stem or .rxbin file
-n No Optimising
```

```
Notes :
- source_file : The source file to be assembled; filetype (rxas) is
added to the name.
```

13.5 Optimizer

The assembler contains an optimizer; this is different from the optimizer which is part of the compiler. This phase of the assembler is running always, except when switched off by the `-n` option. When there is any doubt whether any encountered problem is caused by the optimizer, switching it off can help diagnosing the problem.

13.6 Assembler Directives

For machine instructions, see the *cRexx VM Specification* publication. This section discusses instructions to the assembler, which are called *directives* to clearly distinguish them from virtual machine instructions. These are necessary to pass information into the compiled `.rxbin` binary file, to enable execution by the `rxvm` virtual machine. In the following itemized list, *italic* descriptors are categories, while items in roman type are literal directives.

comments A block comment can be made by surrounding the text block with `/*` and `*/` indicators. The `*` (asterisk) can be used as a line comment. The remainder of the line after a line comment is ignored.

labels A label (a string ending with a colon, indicated in the machine instructions documentation as an `ID`, is a target for branching-type instructions.

Example:

```
loop:
    fndnblnk r3,r1,r3   /* find first/next non blank offset         */
    ilt r5,r3,0         /* if <0, nothing found, end search         */
    brt break,r5        /* branch to break                          */
    inc r6              /* else increase word count                 */
                        /* offset of word is in R3                  */
    copy r8,r3          /* save offset of word                      */
    fndblnk r3,r1,r3    /* from offset find next blank offset       */
    ieq r7,r6,r2        /* is this the word we are looking for?     */
    brt wordf,r7        /* go fetch it                              */
```

```
ilt r5,r3,0         /* if <0, nothing found, end search         */
brt break,r5        /* no word found                            */
bct loop,r4,r3      /* continue to look for next non blank char */
```

registers `r0 ...*n` are names of registers, indicated in the machine instructions documentation as `REG`.

.globals=*INT* Defines *int* global variable *g0 ... g*n. These can be used within any procedure in the file.

.locals=*INT* The number of local registers (local to the source program). This number needs to be 1 greater than the highest used register number.

.expose=*ID* Any global register marked as exposed is available to any file which also has the corresponding exposed index/name.

.srcstep Used to document source-step metadata. This optional directive records a self-contained source anchor: module-local step id, clause id, provenance flags, source file, source line number, active column range, and the whole source line. It is used by TRACE, SOURCELINE, panic reporting, and debuggers.

.traceevent Used to document semantic TRACE events. This optional directive records compact event/value codes, mode visibility, value source, value type, source-step id, clause id, and optional symbol or resolved compound names. TRACE handlers use these records instead of guessing values from source text.

.proc A procedure is a scope delimiting mechanism for the access of registers. The registers of a procedure are independent of the caller's registers. The VM maps its registers to the registers in the caller. Each time a procedure/-function is called a new *stack frame* is provided. This means that the called function has its own set of registers.

The function header defines how many registers (called locals) the function can access - for practical purposes one can consider that any number of registers can be assigned to a function. In addition, each file defines a number of global registers that can be shared between procedures.

13.7 Examples

Here are several examples of how to use `rxas` to assemble a program into an object module.

13.7.1 In-line assembly

The CREXX compiler `rxc` enables[6] inline assembly through the `assembler` statement. When used in this way, a lot of the complications of an assembly language program can be handled by the CREXX compiler, like assigning registers to variables, and the conversion of datatypes like `integer` for display as `string`.

```
/* crexx ipow test - inline assembler */
options levelb

number = 10
power = 2
result = 0

say 'test IPOW'
assembler do
   ipow result,number,power
end
say "result =" result /* The compiler converts integer to string */
```

```
test IPOW
result = 100
```

This is a simple and straightforward way to complement the low level assembler instructions with the power of the REXX language. The following example intends to explain how this is implemented; it can be skipped without consequences.

In this example, the compiler generates the following assembler source:

```
/*
 *SOURCE : pow.crexx
 */

.globals=0

main() .locals=4
```

[6] When used with `options level b`

```
.meta "pow.main"="b" ".void" main() ""
numsci 18,1,1
.srcstep 1 1 17 "pow.crexx" 4 1 12 "number = 10"
.meta "pow.main.number"="b" ".int" r0
load r0,10
.traceevent "L" 4 "R" "I" "r" 0 1 1 0 "" ""
.traceevent "A" 6 "R" "I" "r" 0 1 1 0 "number" ""
.srcstep 2 2 17 "pow.crexx" 5 1 10 "power = 2"
.meta "pow.main.power"="b" ".int" r1
load r1,2
.traceevent "L" 4 "R" "I" "r" 1 2 2 0 "" ""
.traceevent "A" 6 "R" "I" "r" 1 2 2 0 "power" ""
.srcstep 3 3 17 "pow.crexx" 6 1 11 "result = 0"
.meta "pow.main.result"="b" ".int" r2
load r2,0
.traceevent "L" 4 "R" "I" "r" 2 3 3 0 "" ""
.traceevent "A" 6 "R" "I" "r" 2 3 3 0 "result" ""
.srcstep 4 4 17 "pow.crexx" 8 1 16 "say 'test IPOW'"
say "test IPOW"
.srcstep 5 5 17 "pow.crexx" 10 4 28 " ipow result,number,power"
.traceevent "V" 6 "R" "I" "r" 2 0 0 0 "result" ""
.traceevent "V" 6 "R" "I" "r" 0 0 0 0 "number" ""
.traceevent "V" 6 "R" "I" "r" 1 0 0 0 "power" ""
ipow r2,r0,r1
.srcstep 6 6 17 "pow.crexx" 12 1 22 "say \"result =\" result /*The compiler converts
     integer to string */"
itos r2
.traceevent "V" 6 "R" "S" "r" 2 6 6 0 "result" ""
sconcat r3,"result =",r2
.traceevent "O" 4 "R" "S" "r" 3 6 6 0 "" ""
say r3
.srcstep 7 7 1 "pow.crexx" 12 22 22 "say \"result =\" result /*The compiler converts
     integer to string */"
ret
.meta "pow.main.number"
.meta "pow.main.power"
.meta "pow.main.result"
```

The `.srcstep` directives (intended for trace, sourceline, panic reports, and debuggers) indicate where the work is done. Related `.traceevent` directives identify values that are actually available for `trace` display. The variables are assigned, as integers, to the registers `r1` and `r2`. The line `ipow, number, power` becomes `ipow r3,r1,r2`, and the display on the terminal is handled by the `itos`,`sconcat` and `say` instructions.

This is an example, with the remark that in this case, the microcode for `ipow` is always executed, the example in `crexx` on page **??** shows that the CREXX optimizer of the compiler can eliminate this code entirely.

The use of Assembler directives is not allowed in inline assembly, so (as an example) is it not possible to define procedures in an inline assembler block.

13.8 Troubleshooting

The assembler will give messages when there are problems in a source file. These are hopefully of enough clarity to resolve the immediate problem with syntactic issues or typos. When a program assembles correctly but its behaviour is unexpected, or its output is incorrect, a number of different strategies can be followed.

13.8.1 Adding say statements

It is easy to add `say` statements to your program. Unlike Rexx, there is no trace statement for assembler programs. However, it is easy to disassemble (see rxdas on page **??**) an `.rxbin` module, and reassemble it with added statements.

13.8.2 Using the debugger

The `rxdb` debugger has a mode for assembler. This can be used to set breakpoints and/or step through the code; here the registers can be traced so variables in your program can be followed and the comparisons and branches can be checked. For more information about the debugger, see rxdb on page 115).

13.8.3 Using the Trace statement

The `trace asm` statement in CREXX is designed to trace `.rxas` code as it is executed. This works only from a REXX program.

As an example, when we want to know which VM instructions are executed:

```
# trace the world
trace asm
say 'hello trace'
say 5**2
```

```
Error: compilation of extensionless files is not supported.
(and no hellotrace.crexx was found)
```

and we see that the optimiser did its work by deciding the answer will never be something else than 25. If we want to see how it is calculated, we need to compile without optimisation:

```
crexx hellotrace --nooptimise
```

and the resulting trace will show

```
Error: compilation of extensionless files is not supported.
(and no hellotrace.crexx was found)
```

14 rxlink - the cREXX Linker

`rxlink` combines one or more `.rxbin` modules into a linked image with a shared constant pool. The result is still a normal `006` format record stream and can be loaded by `rxvm`, `rxbvm`, or passed on to `rxcpack`.

14.1 Command Line Options

```
cREXX Linker
Usage: rxlink [options] input_file [input_file ...]
Options:
-o output_stem Linked output stem or .rxbin file
-c control_file Control file with INPUT/ROOT/INCLUDE/OMIT/OUTPUT/MAP/STRIP
-r root_member Root module selector (may be repeated)
-m map_file Write a simple link map
-l location Working location for input/output resolution
-s Strip source/TRACE debug metadata from linked output
-i Preserve inline-body metadata in linked output
-d Debug mode
-h Help
```

14.2 What It Produces

The linker writes:

1. one shared-pool record
2. one shared-backed module record for each selected module

This reduces duplication across modules while preserving the module boundaries that the virtual machine still uses for load and runtime link work.

14.3 When To Use It

The simplest workflows can stop at `rxas` and run one or more `.rxbin` files directly under `rxvm`. `rxlink` becomes useful when you want a single deployable linked image:

- to package an application root together with the modules it needs
- to reduce duplicated constant-pool data before `rxcpack`
- to produce a tidier deployable artifact than a loose set of `.rxbin` files
- to remove source/TRACE debug metadata from a deployment image

The linker does not try to replace all runtime loading. It links the modules you give it, resolves what it can within that set, and leaves other imports available for later runtime resolution if that is how the program is designed.

14.4 Module Selection

The linker can be driven from the command line or from a control file.

- `-r` selects one or more root modules directly.
- `INCLUDE` forces a module into the linked output even if it is not reached from the current roots.
- `OMIT` excludes a module from consideration.
- If no root is specified, `rxlink` selects modules containing `main`.
- If no selected module contains `main`, it falls back to the modules from the first input file.

After roots are chosen, `rxlink` follows exposed imports, `srcfprocsel` interface-factory references, and interface relationships to pull in the modules needed by those roots.

In normal application linking, explicit roots are usually enough. `INCLUDE` and `OMIT` are mainly for advanced scenarios:

- preparing a linked image that is still expected to do some dynamic or late loading
- carrying an optional provider that is not reached by the normal root graph
- excluding an alternative provider while testing or packaging a chosen variant

14.5 Selectors

Where a directive expects a module selector, `rxlink` accepts:

- the full module name
- the basename
- the source stem
- `input_path::member` for a specific member inside a multi-record input

This is especially useful when one input file contains multiple linked or concatenated modules.

14.6 Control Files

Control files are line-oriented and are useful when the link set is too large or too deliberate to keep on one command line. Blank lines are ignored, and lines beginning with * or # are treated as comments.

The supported directives are:

TABLE 1: Linker Directives.

Parm	Arg
INPUT	path
ROOT	selector
INCLUDE	selector

Parm	Arg
OMIT	selector
OUTPUT	path
MAP	path
STRIP	SOURCE

Selectors can use a module name, basename, stem, or `input_path::member` form.

The most common pattern is:

1. list one or more `INPUT` files
2. choose a `ROOT`
3. optionally ask for `STRIP SOURCE`
4. choose an `OUTPUT`

14.6.1 Minimal Application Control File

```
* app.ctl
INPUT build/app.rxbin
INPUT build/library.rxbin
ROOT app
STRIP SOURCE
OUTPUT build/app_linked
```

This tells `rxlink` to start from module `app`, pull in the providers it needs from the given inputs, strip source/TRACE debug metadata, and write `build/app_linked.rxbin`.

Run it with:

```
rxlink -c app.ctl
```

14.6.2 Advanced Include Example

```
# app_with_optional_provider.ctl
INPUT build/app.rxbin
INPUT build/providers.rxbin
ROOT app
INCLUDE providers::fallback_logger
OUTPUT build/app_with_optional_provider
```

Here `fallback_logger` is forced into the linked image even if the normal root/import walk would not select it. This is an advanced packaging tool: useful when you

know a module will be located indirectly or later than the normal static reachability analysis can see.

14.6.3 Omit One Provider Variant

```
* app_choose_provider.ctl
INPUT build/app.rxbin
INPUT build/provider_a.rxbin
INPUT build/provider_b.rxbin
ROOT app
OMIT provider_b
OUTPUT build/app_provider_a
```

This form is useful when you have alternative providers and want to make the choice explicitly at packaging time.

14.6.4 Add A Map File

```
INPUT build/app.rxbin
INPUT build/library.rxbin
ROOT app
MAP build/app_linked.map
OUTPUT build/app_linked
```

The map file records which modules were selected and which imports, if any, remain unresolved inside the current link set.

14.7 Stripping Source Metadata

`rxlink` strips inline-body metadata from linked output by default. Inline metadata is intended for library artifacts consumed by `rxc`; it is not needed after a final linked image has been built. Use `rxlink -i` to preserve it for diagnostic or tooling builds. The equivalent control-file form is:

```
PRESERVE INLINE
```

`rxlink -s` strips source/TRACE debug metadata from the linked output. The equivalent control-file form is:

```
STRIP SOURCE
```

This removes the serialized source-step records and semantic TRACE event records while preserving runtime metadata such as exposed symbols and interface/class contract data. Keep the linked image unstripped when classic TRACE, `TRACE LLM`, or RXDB source-level debugging is needed.

This option is intended for deployable `.rxbin` artifacts and for downstream `rxcpack` / wrapped images where source breadcrumbs are not needed.

14.8 Example

```
rxlink –o app linked_root.rxbin helpers.rxbin
rxvm app
```

`rxlink -o` names a linked-image stem unless the value already ends in `.rxbin`, so the first command writes `app.rxbin`.

To produce a smaller deployable linked image:

```
rxlink –s –o app_small linked_root.rxbin helpers.rxbin
```

The equivalent control-file driven form is:

```
INPUT linked_root.rxbin
INPUT helpers.rxbin
ROOT linked_root
STRIP SOURCE
OUTPUT app_small
```

15 Assembler Programming Guide

This chapter has general information about writing programs in assembly language using the `rxas` assembler programs. It discusses the assembler file format, linkage conventions, and other useful programming information, including the CREXX level B feature which allows including assembly language statements in REXX programs.

15.1 The `rxas` command

The assembler is started by invoking the `rxas` executable and takes options, which are specified before the name of the source file. For the options, see page 89. The input file has a fixed filename extension of `.rxas`, the output file replaces the extension with `rxbin`. The input is UTF-8[7] and the output is `rxbm` bytecode, which is platform independent.

15.2 The format of an assembler source file

The format of the source lines follows the general layout of assembly source, without a fixed format - there are no mandatory start columns. As most `.rxas` files will

[7] except for older mainframe operating system versions, where it is EBDCIC.

be generated by the rxc compiler from Rexx source, let's have a quick look at what it looks like.

```
options levelb
import rxfnsb

say "hello, rxas!"
say "Today is" date()
```

```
hello, rxas!
Today is 26 Jul 2026
```

This is the generated assembler source:

```
/*
 * SOURCE                 : hellox.crexx
 */

.globals=0

main() .locals=7
   .meta "hellox.main"="b" ".void" main() ""
   numsci 18,1,1
   .srcstep 1 1 17 "hellox.crexx" 4 1 19 "say \"hello, rxas!\""
   say "hello, rxas!"
   .srcstep 2 2 17 "hellox.crexx" 5 1 20 "say \"Today is\" date()"
   load r0,5
   settp r1,512
   settp r2,512
   settp r3,512
   settp r4,512
   settpcall r6,rxfnsb.date(),r0,r5,512
   .traceevent "F" 4 "R" "S" "r" 6 2 2 0 "date" ""
   sconcat r6,"Today is",r6
   .traceevent "O" 4 "R" "S" "r" 6 2 2 0 "" ""
   say r6
   .srcstep 3 3 1 "hellox.crexx" 5 20 20 "say \"Today is\" date()"
   ret

/* Imported Declaration from file: date@library.rxbin */

rxfnsb.date() .expose=rxfnsb.date
   .meta "rxfnsb.date"="b" ".string" rxfnsb.date() "?oformat=.string,?
      idate=.string,?iformat=.string,?osep=.string,?isep=.string"
```

Because this assembler source is generated by the compiler, it contains more than strictly necessary for a successful assembly; to be precise, support for the source-

line() function, tracing and numeric settings. We see a `main() .locals=7` in line 8. Lines 9-14 set defaults for numeric handling, while line 15 includes the sourceline from the CREXX program. Line 16 has the `say hello, rxas`; `say` is the assembler statement which assembles into an instruction for the runtime to put out this line.

When hand-written, this program would look more like:

```
main() .locals=7
  say "hello, rxas!"
  load r1,5
  call r6,date(),r1
  sconcat r6,"Today is",r6
  say r6
  ret

date() .expose=rxfnsb.date
```

We can see that the first component of the lines are the procedure names or the instruction mnemonics ; this program has no labels. After the instructions, their parameters are placed.

15.3 Register names

The CREXX Virtual Machine has a virtually unlimited number of registers. Registers are called r1..r*n. A procedure needs to specify the number of local registers it uses.

15.4 Labels

Labels are used, as in other assembly languages, as targets for branches and comparisons in the code: The label is a string that is ended by a colon (:'). Its purpose is to be a symbolic location to branch into when decisions are taken. It is recommended that a label starts in the first column of the line.[8] Branches need to have a

[8] it will not hurt, however, if it does not start in column one. It must be the first word of the line, though.

target within a module; they cannot jump into other procedures, this is reserved for the `call` operation.

15.5 Procedure names

Procedure names are identifiers for routines in the program. Procedure `main()` carries a special role, as that is the one the VM runtime gives control to when a program is started. Procedures return either nothing (`.void`) or data of a specified type (like `.int` or `.string`). Procedures can be called and will return a value of the specified type.

15.6 Linkage conventions

15.7 Optimizing actions of the `rxas` assembler

`rxas` is an optimising assembler[9] which can rearrange instructions or eliminate them entirely. This optimisation can be switched off when in doubt. This will cause loss of performance, but will make debugging of the code more straightforward, because there is a more direct correspondence between the sourcelines and the generated assembly code.

15.8 Embedded Rexx Assembler

The `assembler` command enables the inclusion of arbitrary assembler instructions within a Rexx program. The compiler validates the instruction mnemonic and the complete variable-length argument list against the opcode signature, ensuring that variables are converted into the appropriate register number. Inline assembly is therefore not limited to three operands.

This example uses the `linkarg` assembler instruction with two variables and an integer constant.

```
assembler linkarg e,i,5
```

[9] optimization is default but can be switched off, for example when debugging.

16 rxcpack - the cREXX C Packer

The C Packer program converts the `.rxbin` files into a C language structure which links together all needed modules, and a large part of the Virtual Machine infrastructure, which file then can be compiled and link edited by the C compiler. `gcc` and `clang` are the targeted compiler toolchains for Linux, macOS and Windows.

16.1 Command Line Options

```
cREXX rxcpack. Tool to convert rxbin files (or any binary file) into
a c file
that can be linked to a C exe
Version : crexx-1.0.0-beta.3+local.g78c83e4a0b88
Usage : rxcpack [options] input_file_1 input_file_2 ... input_file_-
n
(.rxbin is appended to input file names)
Options :
-h Help Message
-c Copyright & License Details
-v Version
-o output_file Binary Output File (.c is appended - default is input_-
file_1.c)
```

16.2 Source Code

Note that the files that are produced by `rxcpack` are valid C source code but only offer the minimal structures to link dumps of binary objects together. These are already platform dependent and can be combined with the statically linked plugin images for the target platform. The flags used for the compiler and linkage editor can be seen in the verbose output of the `crexx` compiler driver; this syntax is slightly different for `gcc` and `clang`.

16.3 The rxlink step

Though not strictly required, it is advisable to use the the `rxlink` tool in combination with the `rxpack`ed modules. It is automatically called by the `crexx` compiler driver to de-duplicate variable occurrences and select only referenced modules from libraries, shrinking the resulting native executable considerably.

17 rxdas - the disassembler

A disassembler reverses the actions of an assembler; where the assembler turns a text file containing instructions and directives into a binary executable, the disassembler returns this binary file into its text form[10] ;in this case it delivers a disassembly which in itself can be re-assembled - and still works.

17.1 Input/Output

The `rxdas` disassembler has a `.rxbin` file as input and produces a text file as output which goes to `stdout`. In this text file a disassembly has taken place; labels are synthetic and based on the combination of instructions around them. As clearly can be seen in the above, the labels generated by the CREXX compiler are not the same as the ones generated by the disassembler. With option `-p`, the constant pool of the `.rxbin` file is printed first, before the rest of the disassembly.

The current `006` `.rxbin` format is fully supported, including pooled float literals, packed code and constant sections, and the interface/class metadata directives `.class`, `.attr`, `.interface`,`.implements`, and `.member`.

Assembler-built jump tables are reconstructed as procedure-local synthetic `.jtable` and `.jcase` declarations. The disassembly records the packed algorithm actually selected (`linear`, `openhash`, or `acph`) and rewrites `jumps`, `jumpr`, `jumpn`, `jumpb`, `jumpbs`, and `jumpi` operands to the synthetic table name. Numeric tables omit the

[10] as much as possible, given the fact that some information on literals has disappeared

internal NaN compatibility entry, which is recreated by `rxas` when the disassembly is assembled again.

17.2 Command Line Arguments

When the command line argument -h is specified the options are shown:\

```
cREXX Disassembler
Version : crexx-1.0.0-beta.3+local.g78c83e4a0b88
Usage : rxdas [options] binary_file
Options :
-h Help message
-c Copyright & license details
-v Version
-p all constant Pool
-g semantic graph summary
-l location working Location (directory)
-o output_file Output file (default is stdout)
```

17.3 Example

```
/* compute sum of numbers 1 to 100 (5050) */
options levelb
/* compute sum of numbers 1 to 100000  */
sum = 0
do i=1 to 100000
   sum = i+sum
end
say "the sum of the numbers 1 to 100000 is:" sum
return
```

This program, when compiled by the `rxc' compiler, produces the following assembly source:

```
/*
 * SOURCE                 : sumLoop1000.crexx
 */
```

```
.globals=0

main() .locals=4
   .meta "sumloop1000.main"="b" ".void" main() ""
   setnumdgts 18
   setnumfuz 0
   setnumfrm 1
   setnumcas 1
   setnumstd 1
   .srcstep 1 1 17 "sumLoop1000.crexx" 4 1 8 "sum = 0"
   .meta "sumloop1000.main.sum"="b" ".int" r0
   load r0,0
   .traceevent "L" 4 "R" "I" "r" 0 1 1 0 "" ""
   .traceevent "A" 6 "R" "I" "r" 0 1 1 0 "sum" ""
   .srcstep 6 6 17 "sumLoop1000.crexx" 5 1 3 "do i=1 to 100000"
   .srcstep 2 2 17 "sumLoop1000.crexx" 5 4 7 "do i=1 to 100000"
   .meta "sumloop1000.main.i"="b" ".int" r1
   load r1,1
   .traceevent "L" 4 "R" "I" "r" 1 2 2 0 "" ""
   .traceevent "A" 6 "R" "I" "r" 1 2 2 0 "i" ""
   .srcstep 3 3 17 "sumLoop1000.crexx" 5 8 17 "do i=1 to 100000"
   load r2,100000
   .traceevent "L" 4 "R" "I" "r" 2 0 0 0 "" ""
l7dostart:
   .srcstep 3 3 17 "sumLoop1000.crexx" 5 8 17 "do i=1 to 100000"
   igt r3,r1,r2
   brt l7doend,r3
   .srcstep 5 5 17 "sumLoop1000.crexx" 6 4 15 "   sum = i+sum"
   .traceevent "V" 6 "R" "I" "r" 1 5 5 0 "i" ""
   .traceevent "V" 6 "R" "I" "r" 0 5 5 0 "sum" ""
   iadd r0,r1,r0
   .traceevent "O" 4 "R" "I" "r" 0 5 5 0 "" ""
   .traceevent "A" 6 "R" "I" "r" 0 5 5 0 "sum" ""
l7doinc:
   .srcstep 4 4 17 "sumLoop1000.crexx" 5 4 5 "do i=1 to 100000"
   inc r1
   .srcstep 7 7 17 "sumLoop1000.crexx" 7 1 4 "end"
   br l7dostart
l7doend:
   .srcstep 8 8 17 "sumLoop1000.crexx" 8 1 49 "say \"the sum of the
      numbers 1 to 100000 is:\" sum"
   itos r0
   .traceevent "V" 6 "R" "S" "r" 0 8 8 0 "sum" ""
   sconcat r3,"the sum of the numbers 1 to 100000 is:",r0
   .traceevent "O" 4 "R" "S" "r" 3 8 8 0 "" ""
   say r3
   .srcstep 9 9 17 "sumLoop1000.crexx" 9 1 7 "return"
   ret
   .meta "sumloop1000.main.sum"
```

The assembler produces an `rxbin' file from that. What follows is the disassembly from this `rxbin' file.

```
* MODULE - sumLoop1000
* DESCRIPTION - sumLoop1000

* CONSTANT POOL - Size 0x7c0.  Dump of EXPOSED entries only (option -p
    not used):


* CODE SEGMENT - Size 0x1d

.globals=0

main()          .locals=4
                .meta "sumloop1000.main"="b" ".void" main() ""
                setnumdgts 18                                   * 0x000000
                    :00f6 Set Numeric Digits digits=op1 (>4)
                setnumfuz 0                                     * 0x000002
                    :00f9 Set Numeric Fuzz digits=op1 (>=0)
                setnumfrm 1                                     * 0x000004
                    :00fc Set Numeric Form=op1 (1=sci,2=eng)
                setnumcas 1                                     * 0x000006
                    :00ff Set Numeric Case=op1 (1=lower,2=upper)
                setnumstd 1                                     * 0x000008
                    :0102 Set Numeric Standard=op1 (1=common,2=classic)
                .srcstep 1 1 17 "sumLoop1000.crexx" 4 1 8 "sum = 0"
                .meta "sumloop1000.main.sum"="b" ".int" r0
                load r0,0                                       * 0x00000a
                    :0001 Load op1 with op2
                .traceevent "L" 4 "R" "I" "r" 0 1 1 0 "" ""
                .traceevent "A" 6 "R" "I" "r" 0 1 1 0 "sum" ""
                .srcstep 6 6 17 "sumLoop1000.crexx" 5 1 3 "do i=1 to
                    100000"
                .srcstep 2 2 17 "sumLoop1000.crexx" 5 4 7 "do i=1 to
                    100000"
                .meta "sumloop1000.main.i"="b" ".int" r1
                load r1,1                                       * 0x00000d
                    :0001 Load op1 with op2
                .traceevent "L" 4 "R" "I" "r" 1 2 2 0 "" ""
                .traceevent "A" 6 "R" "I" "r" 1 2 2 0 "i" ""
                .srcstep 3 3 17 "sumLoop1000.crexx" 5 8 17 "do i=1 to
                    100000"
                load r2,100000                                  * 0x000010
                    :0001 Load op1 with op2
                .traceevent "L" 4 "R" "I" "r" 2 0 0 0 "" ""
                .srcstep 3 3 17 "sumLoop1000.crexx" 5 8 17 "do i=1 to
                    100000"
lb_13:          igtbr lb_1e,r1,r2                               * 0x000013
    :0157 Int Greater than if (op2>op3) goto op1
```

```
                .srcstep 5 5 17 "sumLoop1000.crexx" 6 4 15 "   sum = i+
                    sum"
                .traceevent "V" 6 "R" "I" "r" 1 5 5 0 "i" ""
                .traceevent "V" 6 "R" "I" "r" 0 5 5 0 "sum" ""
                iadd r0,r1,r0                                   * 0x000017
                    :000f Integer Add (op1=op2+op3)
                .traceevent "O" 4 "R" "I" "r" 0 5 5 0 "" ""
                .traceevent "A" 6 "R" "I" "r" 0 5 5 0 "sum" ""
                .srcstep 4 4 17 "sumLoop1000.crexx" 5 4 5 "do i=1 to
                    100000"
                inc1                                            * 0x00001b
                    :0020 Increment R1++ Int
                .srcstep 7 7 17 "sumLoop1000.crexx" 7 1 4 "end"
                br lb_13                                        * 0x00001c
                    :0024 Branch to op1
                .srcstep 8 8 17 "sumLoop1000.crexx" 8 1 49 "say \"the
                    sum of the numbers 1 to 100000 is:\" sum"
lb_1e:          itos r0                                         * 0x00001e
    :00e5 Set register string value from its int value
                .traceevent "V" 6 "R" "S" "r" 0 8 8 0 "sum" ""
                sconcat r3,"the sum of the numbers 1 to 100000 is:",r0
                    * 0x000020:008b String Concat with space (op1=op2||
                    op3)
                .traceevent "O" 4 "R" "S" "r" 3 8 8 0 "" ""
                say r3                                          * 0x000024
                    :01c2 Say op1
                .srcstep 9 9 17 "sumLoop1000.crexx" 9 1 7 "return"
                ret                                             * 0x000026
                    :0030 Return VOID
                .meta "sumloop1000.main.sum"
```

This file has the output of the disassembler; the procedure name, `main`, is identical, but the first label generated by the compiler, `l7dostart:` is called `lb_9` in the disassembler output. This is of no consequence for a subsequent re-assembly and execution of the program.

Do note that the lines are longer than expected, because every disassembled line has the binary representation and a standard explanation of the generated instruction. This makes it easier to find back instructions in an `.rxbin` load module.

The instructions, however, can be different than expected due to the optimizations the assembler performs. When the compiler has performed optimizations, this is already visible in the `.rxas` file. The assembler optimizations are visible in the disassembled object.

The standard instruction documentation the disassembler affixes is the same as

displayed by the `rxas -i` command, in a line comment after the instructions and their operands.

When stepping through a program using the CREXX Debugger (which is mentioned in the next chapter), the disassembly is the most representative record of what is in the `.rxbin` executable.

18 rxdb - the CREXX Debugger

The debugger is one of the programs in the toolchain delivered with CREXX as its source code; most other programs, at the moment, are compiled from C. It is easily adaptable and can be regarded a *debugger construction set*. By adapting and recompiling the user can implement their own wishes for a debugger. In this sense, it can be seen as an open-ended complement to the REXX `trace` statement. Because it has modes for REXX as well as `rxas` Assembler, it is a useful tool for debugging high-level as well as low-level problems.

18.1 Command Line Options

```
RXDB - REXX Debugger
Usage: rxdb [text|plain|llm]
```

18.2 Runtime Options

After the `rxdb` program is started, a few runtime options appear in the delivered version. This is an example session:

PART III

Plugins

19

cRexx /PA - Plugin Architecture

This facility allows for the compilation, linking, and execution of additional functionality (developed in C) alongside Rexx code. It enables the decoupling of native modules (plugins), which can be developed and packaged either as separate entities or linked statically to the main crexx core solution.

The architecture is designed to completely decouple the plugins from the rest of cRexx, and the plugin client library only consists of a single header file (rexxpa.h)

Plugin developers have the flexibility to structure their code in a way that allows for both dynamic or static linking. While the source code can remain the same, it needs to be built differently based on the desired linking approach, as it relies on macro expansion.

During the Rexx compilation process, the Rexx compiler inspects the plugin to determine the type and argument of any provided functions. This inspection helps ensure type safety. It is recommended for plugin developers to load or initialise dependencies only when an explicitly exposed initialisation Rexx function is called or when the first function is invoked (lazy initialisation) to avoid overhead during build (rather than run). In addition, for the static builds, there is macro definition **DECL_ONLY** which allows definition / implementation code to be excluded from the static library designed to be linked to the compiler only.

For dynamically packaged plugins (with the extension *.rxplugin), the search process for these plugins mirrors how crexx locates Rexx modules (extensions such

as *.rexx, *.rxas, or *.rxbin). In contrast, for static builds, the provided functions are loaded before the execution of the main() function in the core crexx solution, however these functions are placed at the end of the search order, meaning users can override static function definitions with local native or crexx modules.

The Plugin Architecture offers a comprehensive set of resources for developers, including a header file, macros, and a cmake configuration. These tools aim to create a convenient and efficient development environment for plugin creation.

19.1 Dynamic Plugins Recommended

User-provided plugins are recommended to be provided as dynamic `.rxplugin` files. Dynamic plugins offer several advantages: easy site-wide distribution by placing them in the same directory as CREXX binaries, project-specific customization by locating them in the project REXX files directory, and flexible placement in any desired location using `rxc` and `rxvm` options.

In contrast, static plugin packaging is more complex in terms of linking and is intended for core CREXX components that are part of every CREXX release. Static plugins are shipped within the CREXX binaries, ensuring their availability and consistency across distributions.

The choice between dynamic and static plugin packaging depends on the specific requirements and use cases: dynamic plugins provide flexibility and customization for users, while static plugins are designed to provide a robust solution for core CREXX components.

19.2 Example Plugin

The following code demonstrates a decryption plugin.

- The PROCEDURE macro starts the function, and argument access is via macros like ARG().
- Each argument is a CREXX register. This means it holds multiple types or values and these are accesses by macros like GETSTRING() and SETSTRING().
- Errors are handled by defining and returning a SIGNAL via the RETURNSIGNAL macro.

```
/*--------------------------------------------------------------------*/
/*                                                                    */
/* Name: rxdes.C                                                      */
/*                                                                    */
/* Copyright René Jansen june 1st 1993                                */
/*                                                                    */
/* Function: crexx external functions RxDesEncrypt and RxDesDecrypt */
/*                                                                    */
/* dependencies/calls: desbase.c CREXX/rxpa                           */
/*                                                                    */
/* Ported to CREXX by Adrian Sutherland 2024                          */
/*--------------------------------------------------------------------*/

#include <stdio.h>
#include <string.h>
#include "crexxpa.h"    // crexx/pa - Plugin Architecture header file
#include "desbase.h"    // DES encryption/decryption functions

// Encrypt a string using DES - the first argument is the key, the
    second the data
PROCEDURE(RxDesEncrypt)
{
    char   des_out[8];
    char   des_in[8];
    char   key[8];
    char   result[17];

    if( NUM_ARGS != 2)
        RETURNSIGNAL(SIGNAL_INVALID_ARGUMENTS, "2 arguments expected")

    if (strlen(GETSTRING(ARG(0))) != 16 || strlen(GETSTRING(ARG(1))) !=
         16)
        RETURNSIGNAL(SIGNAL_INVALID_ARGUMENTS, "Both arguments must be
            16 hex digits")

    if( hex2bin(GETSTRING(ARG(0)), key) < 0)
        RETURNSIGNAL(SIGNAL_INVALID_ARGUMENTS, "Key is not valid hex")

    if( hex2bin(GETSTRING(ARG(1)), des_in) < 0)
        RETURNSIGNAL(SIGNAL_INVALID_ARGUMENTS, "Data is not valid hex")

    // Encrypt
    desinit(key);
    endes(des_in, des_out);

    // Set return and make sure the signal is reset/ok
    bin2hex(des_out, result);
    SETSTRING(RETURN, result);
    RESETSIGNAL
```

```
}

// Decrypt a string using DES - the first argument is the key, the
    second the data
PROCEDURE(RxDesDecrypt)
{
    char   des_out[8];
    char   des_in[8];
    char   key[8];
    char   result[17];

    if( NUM_ARGS != 2)
        RETURNSIGNAL(SIGNAL_INVALID_ARGUMENTS, "2 arguments expected")

    if (strlen(GETSTRING(ARG(0))) != 16 || strlen(GETSTRING(ARG(1))) !=
         16)
        RETURNSIGNAL(SIGNAL_INVALID_ARGUMENTS, "Both arguments must be
            16 hex digits")

    if( hex2bin(GETSTRING(ARG(0)), key) < 0)
        RETURNSIGNAL(SIGNAL_INVALID_ARGUMENTS, "Key is not valid hex")

    if( hex2bin(GETSTRING(ARG(1)), des_in) < 0)
        RETURNSIGNAL(SIGNAL_INVALID_ARGUMENTS, "Data is not valid hex")

    // Decrypt
    desinit(key);
    dedes(des_in, des_out);

    // Set return and make sure the signal is reset/ok
    bin2hex(des_out, result);
    SETSTRING(RETURN, result);
    RESETSIGNAL
}

// Functions to be provided to rexx - these are loaded either when the
    plugin is loaded (dynamic) or
// before main() is called (static)
LOADFUNCS
//      C Function__, REXX namespace & name, Option_, Return Type_,
    Arguments
ADDPROC(RxDesDecrypt, "rxdes.decrypt",       "b",     ".string",    "
    key=.string,data=.string");
ADDPROC(RxDesEncrypt, "rxdes.encrypt",       "b",     ".string",    "
    key=.string,data=.string");
ENDLOADFUNCS

// End of file
```

The following shows how the defined functions are published to cRexx.

```
// Functions to be provided to rexx \- these are loaded either when the
    *
// plugin is loaded (dynamic) or before main() is called (static)*
LOADFUNCS
//      C Function\_\_, REXX namespace & name, Option\_, Return Type\_,
    *
// Arguments*
ADDPROC (RxDesDecrypt, "rxdes.decrypt",       "b",     ".string",
   "key=.string,data=.string");
ADDPROC (RxDesEncrypt, "rxdes.encrypt",       "b",     ".string",
   "key=.string,data=.string");
ENDLOADFUNCS**
```

The ADDPROC Macro published the function to cRexx. This has the namespace and name, options/level (should be b), the function return type (in level b format) and the arguments (again in level b format). This allows the Compiler and VM to search and ``fix-up'' procedure calls using the same search order and syntax as for procedures developed in Rexx.

19.3 Macros

The following MACROs are provided for plugin developers (defined in rexxpa.h)

Macro	Purpose
LOADFUNCS	Starts and finishes the block of ADDFUNC()'s
ADDFUNC()	Published a function to CREXX
NUM_ARGS	Is the number of arguments passed the the function
ARG()	Returns the nth argument (which is an opaque pointer to the CREXX register)
RETURN	Returns the register used to pass the function's returned value.
GETSTRING()	Gets the String value of a register
SETSTRING()	Sets the String value of a register
GETINT()	Gets the Integer value of a register
SETINT()	Sets the Integer value of a register
GETFLOAT()	Gets the float (double) value of a register
SETFLOAT()	Sets the float (double) value of a register
SIGNAL	Returns the registers used to pass and Signal Information.
RESETSIGNAL	Ensures that the signal register is set no ``no signal''
RETURNSIGNAL()	Sets the signal register and returns from the function. Used for error conditions.
ENDLOADFUNCS	Starts and finishes the block of ADDFUNC()'s

The Signal values are:

Signal	Purpose
SIGNAL_NONE	
SIGNAL_ERROR	Syntax error
SIGNAL_- CONVERSION_- ERROR	conversion error between types
SIGNAL_- UNKNOWN_- INSTRUCTION	unknown RSAS instruction
SIGNAL_- FUNCTION_- NOT_FOUND	attempt to execute an unknown function
SIGNAL_OUT_- OF_RANGE	attempt to access an array element that is out of range
SIGNAL_FAILURE	error in an external function or subroutine
SIGNAL_HALT	an external request to halt execution
SIGNAL_- NOTREADY	IO error, such as a file not being ready
SIGNAL_- INVALID_- ARGUMENTS	invalid arguments are passed to a function
SIGNAL_OTHER	Other (or unknown) error condition

In addition

- **BUILD_DLL** is defined if the dynamic (rather than static) version of the plugin is being built. Developers may check for this macro.

- **DECL_ONLY** is designed for static builds linked to the compiler, enabling plugin developers to package a smaller library with declaration support only. The build process would create a separate static library with full functionality separately to link with rxvm. This facility is not provided for dynamic plugins to prevent confusion between declaration-only and full-function files.

19.4 Build Script

The following example shows a CMake script (CMakeLists.txt) to build a plugin. It uses a provided CMake function **add_plugin_target** from **RXPluginFunc-**

tion.cmake.

```
cmake_minimum_required(VERSION 3.5)
project(rxdes C)

set(CMAKE_C_STANDARD 90)

# Including RXPlugin Build System
include(${CMAKE_SOURCE_DIR}/rxpa/RXPluginFunction.cmake)

# Create module
add_dynamic_plugin_target(des rxdes.c desbase.c desbase.h)
```

This script builds a dynamic version of the plugin, and supports Windows (mingw and VS), macOS and Linux.

19.5 Static Builds

The CREXX Cmake build configuration should be referenced for static build support. Building the static library is simple, but proper linking is crucial to guarantee that the linker recognizes the need to link in the plugin and that the plugin's initialization function is called at program load across various platforms.

19.6 Execution from Rexx

The following is a REXX Level B example using the plugin.

Note that there is no manual loading of the dynamic or static plugin, instead CREXX loads the plugins using the same search rules as it uses for other REXX modules. This means that the REXX program (or programmer) does not need to be concerned about how the external function is provided - Rexx, Native, Dynamic, Static - it all has the same calling syntax. This is designed to meet the objective to simplify programming.

```
options levelb
import rxfnsb
import rxdes    /* Import the rxdes plugin functions */

/* Note that the input and output to the des functions are in hex
    strings */
Plaintext = "0000000000000000"
```

```
key =       "08192A3B4C5D6E7F"

Ciphertext = Encrypt(key,Plaintext)
```

In line 4, the import statement loads the namespace meaning that any Rexx modules or native plugins in the rxdes namespace will be loaded as needed; the function call, encrypt(), follows Rexx syntax, and the compiler can check parameters and return types as normal.

PART IV

Appendices

A

Building the Toolchain

Most users will download and install a binary distribution for their platforms and will not need this information. In some cases when a binary distribution is unavailable, it will be beneficial to be able to build the CREXX system using a standard C toolchain.

This chapter aims to show all you need to know about how to build CREXX from scratch, and then run it. It will show what you need, what to do and when it is build, how to make working programs with it.

A.1 Requirements

Currently CREXX builds on Windows, macOS and Linux[11].

You need one of those and:

TABLE 2: Required tools.

Tool	Function
git	Source code versioning
CMake	Build Tool
gcc[12]	C compiler, install g++
Make	conventional build tool, or
Ninja	fast build tool

[11] There is a separate instruction for VM/370 (and later) mainframe operating systems.
[12] clang is a working equivalent and the default on macOS.

A.2 Platform specific info

On Linux and macOS, this instruction is identical. For macOS, Xcode batch tools need to be installed, which will provide you with git, make and the compiler. Brew will give easy access to Cmake and Ninja-build[13].

On Windows, you will need a compatibility layer like msys - installing this has the additional advantage of easy access to git, gcc, cmake and the rest of the necessary tools. On more modern Windows, the WSL[14] and Ubuntu is not a bad choice.

A.3 Process

Here it is assumed that all tools are installed and working, and available on your PATH environment variable.

Choose or make a suitably named directory on your system to contain the source code. Note that the CREXX source is kept in a different directory on you system than where it is built in, or will run from. Now run this command:

```
git clone https://github.com/adesutherland/CREXX.git
```

This will give you a CREXX subdirectory in the current directory, containing the source of CREXX and its dependencies. This is the `develop' branch, which is the one you would normally want to use. All of these are written in the C99 version of the C programming language, which should be widely supported by C compilers on most platforms.

Make a new subdirectory in the current directory (not in CREXX, but in the one that contains it), like `crexx-build'.

```
mkdir crexx-build
```

and cd into that directory. Now issue the following command (we assume that you installed ninja, otherwise substitute `make' for the two <!-- instances ofninja'): -->

```
cmake -G Ninja -DCMAKE_BUILD_TYPE=Release ../CREXX && ninja && ctest
```

[13] For Linux, you will need to install git (which will be there on most distributions), cmake and gcc or clang.}

[14] WSL: Windows subsystem for Linux.

This will do a lot of things. In fact, if all goes well, you will have a built and tested CREXX system.

A.4 Explaining the build process

Let's zoom in a little on what we did. The first step is to tell CMake to validate your system environment and generate a build script (and a test script) for the chosen build tool. Cmake will read the file CmakeLists.txt and validate that your system can do what it asks it to do. This can yield error messages, for example if the C compiler lacks certain functions or header files. (When that happens, open an issue and someone will have a look at it - or peruse stack overflow which is what we probably also will do).

After CMake has successfully validated the build environment, it will generate a build script (a Makefile in the case of Make and a `build.ninja` file in the case of Ninja). This is specified after the -G flag. The -DCMAKE_BUILD_TYPE=Release flag makes sure we do an optimized build, which means we specify an -O3 flag to the C compiler, which then will spend some time optimizing the executable modules, which makes them run faster (they do!). The alternative is a `debug' build which will yield slower executables, but with more debugging information in them.

The two ampersands (&&) mean we do the next part only if the previous step was successful. This is a `ninja` statement, which will build everything in the `build.ninja` specification file. These are a lot of parts, and the good news is, when they are built once, only the changed source will be built, which will be fast.

After this, the generated test suite is run with the `ctest' command. This knows what to do, as the tests were defined in the Cmake recipes, and will show you successes and failures. If what you checked out if git is not a released version, there is a change that some test cases fail, but generally these should indicate success.

A.5 Useful System Test Subsets

For routine system validation, run the full build and full test suite:

```
cmake --build cmake-build-debug
ctest --test-dir cmake-build-debug --output-on-failure
```

For standard-library and BIF work, useful focused checks are:

```
cmake --build cmake-build-debug --target testbifs
ctest --test-dir cmake-build-debug -R '^ts_.*_(noopt|opt)$' --output-on
    -failure
```

If a BIF source change under `lib/rxfnsb/rexx/` causes compiler RXAS golden tests to fail but the corresponding runtime tests still pass, rebuild the linked standard-library image before judging the golden diff:

```
cmake --build cmake-build-debug --target library
cmake --build cmake-build-debug --target testbifs
```

Then rerun the focused compiler/runtime tests and review the generated RXAS:

```
ctest --test-dir cmake-build-debug/compiler/tests -R '13_stems' --
    output-on-failure
ctest --test-dir cmake-build-debug -R '^ts_stem_(noopt|opt)$' --output-
    on-failure
git diff -- compiler/tests/golden
```

Consumer RXAS import blocks are a snapshot of the callables needed for linking/runtime lookup, not a copy of the full provider API. If the diff only adds or removes unused imported declarations, update goldens only after confirming that the consumer still imports the callables it actually calls. The maintainer testing details are in `compiler/docs/testing.md`.

The `lib/rxfnsb/rexx` BIF build is a bootstrap build. It compiles most BIF source files with compiler exits disabled (`rxc -x`), so explicit certified exits such as `TRACE`, `PARSE`, and `ADDRESS` are not available inside those library source files during the BIF build. To debug a BIF, call it from a normal test fixture or scratch program compiled with exits enabled. Add `TRACE UNSUPPRESS NAMESPACE rxfnsb` when you need library frames in the trace.

The native `system` plugin has its own smoke test:

```
ctest --test-dir cmake-build-debug -R '^test_system$' --output-on-
    failure
```

For TRACE/debug metadata work, combine focused TRACE and linker checks before the full suite:

```
ctest --test-dir cmake-build-debug \
  -R '^(trace_event_metadata|test_trace_|ts_trace_|rxlink_format_check|
      rxlink_rxdas_strip_smoke)' \
  --output-on-failure
```

A.6 Build version and timestamp

The project version is read from the top-level `VERSION` file. To change the CREXX version number, edit that file and use a semantic version such as `1.0.0`, `1.0.0-beta.3`, or `1.0.0-beta.3+build.7`.

Local, non-release builds also include build metadata in the displayed version: the build channel, the short Git commit id when Git is available, and a `dirty` marker when the working tree has local changes. The commit id is the part of the local version string that identifies the source revision. After a `git pull`, the commit id is refreshed the next time CMake configures that build directory. If CMake does not reconfigure automatically, run the same `cmake -S ../CREXX -B .` configure command again before rebuilding.

`CREXX_BUILD_TIMESTAMP` is recorded in `BUILDINFO` for package provenance, but it is not part of the displayed tool version. When it is not supplied, CMake creates a UTC timestamp during the first configure of a build directory and stores it in `CMakeCache.txt`. Release and CI builds should pass an explicit timestamp:

```
cmake -DCREXX_BUILD_TIMESTAMP=20260527T120000Z -S ../CREXX -B .
```

A.7 Use of CREXX to build cRexx

CREXX is used to build the library of built-in functions that are written in REXX (and, for a very small part in REXX Assembler) and need to be compiled (carefully observing the dependencies on other REXX built-in functions) before they are added to the library and the CREXX executables in their binary form.

A.8 What do we have after a successful build

A.8.1 Native executables

When all went well, we have a set of native executables for the platform we built CREXX on. These are

TABLE 3: Delivered products.

Name	Function
CREXX	CREXX compiler driver
rxc	CREXX compiler
rxas	CREXX assembler
rxlink	CREXX linker
rxdas	CREXX disassembler
rxvm	CREXX VM, threaded interpreter
rxpp	CREXX macro preprocessor
rxbvm	CREXX VM, non-threaded conventional interpreter
rxvme	CREXX VM, with linked-in REXX library
rxdb	CREXX debugger
rxcpack	CREXX C-generator for native executables

CREXX can compile the REXX script into an executable file, that can be run standalone, for example, on a computer that has no CREXX and/or C compiler installed.

The CREXX debugging tool `rxdb`, is written in REXX and compiled and packaged into an executable file.

The executables in the above table need to be on the PATH environment variable. These are to be found in the compiler, assembler, disassembler, debugger and cpacker directories of the cRexx-build directory we created earlier. It is up to you to add all these separately to the PATH environment, or to just collect them all into one directory that is already on the PATH.

A.8.2 Production and debug builds

Production builds are optimized, while debug builds are slower in execution time but deliver support for the analysis and debugging of problems in the code. The standard distribution is an optimized production build, while a debug build can be produced with the Debug CMake build option.

TABLE 4: Debug vs Release options.

Name	Function
-DCMAKE_BUILD_TYPE=Release	An Optimized build (default)
-DCMAKE_BUILD_TYPE=Debug	A Debug build

A.8.3 Libraries

We also have a set of libraries. Some are written in CREXX and others can be written in C and other programming languages. All executables and libraries are delivered in the `bin` directory of the distribution package.

A.8.4 Optional libraries - build options

TABLE 5: Optional plugin build options.

Name	Function
-DENABLE_ODBC=ON	Build the ODBC Plugin
-DENABLE_GTK=ON	Build the GTK (GUI) Plugin

Some libraries, with dependencies on installed software products, are only produced when they are opted-in with CMake build options. The defaults for these options are `OFF`.

B

cREXX Assembler Architecture (rxas)

The rxas assembler is responsible for translating human-readable Intermediate Representation (IR) assembly (.rxas) into packed executable bytecode (.rxbin) consumed by the rxvm interpreter.

B.1 1. Assembler Pipeline

The assembler processes source files through a pipelined, pseudo-two-pass architecture:

1. **Lexical Analysis (re2c)**:
   - Source defined in assembler/rxasscan.re.
   - Tokenizes input into REXX Assembly primitives: registers (RREG, GREG, AREG), literal types (STRING, INT, FLOAT, DECIMAL, HEX), symbols (ID, LABEL, FUNC), and assembler directives (e.g., .locals, .globals, .expose, .meta).
   - HEX is the RXAS binary-literal token. It accepts byte-paired 0x.../0X... text, including empty 0x, and is stored as OP_BINARY rather than as an integer literal.
   - The lexer also recognizes the interface metadata keywords used at the end of .meta records: .interface, .implements, and .member.
2. **Parsing (Lemon)**:

- Grammar defined in `assembler/rxasgrmr.y`.
- Enforces the structural integrity of the `.rxas` file (headers, function definitions, variable declarations, and instruction sequences).
- The parser actions invoke Builder API functions directly (e.g., `rxasgen*`, `rxaslabl`, `rxasproc`), translating syntax rules into buffered internal data structures.
- Instruction operands are parsed as a recursive comma-separated list. `rxasquev()` and `rxasgenv()` carry the resulting variable-length token vector; there is no three-operand grammar or queue limit.
- Instruction names are derived from `binutils/include/rxops.h`, with a public-source filter for bytecode/runtime-only entries. `RESERVED_*`, `INULL`, `INTERRUPT`, and `IUNKNOWN` remain opcode-table entries but are deliberately rejected as RXAS mnemonics. Recent core runtime additions include the object/interface instructions and the `sock*` TCP socket instructions.

3. **In-Memory Buffering & Constant Pooling**:
   - Handled in `assembler/rxasassm.c`.
   - Instructions and operand slots are appended sequentially into a dynamic array of `bin_code` elements.
   - Complex and pooled literal types (Strings, Binary literals, Floats, Decimals, Procedure Headers, Metadata) are deduplicated via AVL Trees and injected into a variable-length **Constant Pool**.
4. **Backpatching & Optimization (Second Pass)**:
   - Forward references (branches to undefined labels, calls to undefined procedures) are logged as a linked list of references.
   - At the end of the parse, `backptch()` traverses all trees, resolves symbol addresses, updates the binary stream, and applies peephole jump optimizations.

B.2 2. Core Internal Data Structures

The state of the assembler is held within the `Assembler_Context` struct, specifically within the `context->binary` object, which mirrors the layout of the final `.rxbin` file.

B.2.1 Instruction Stream

Instructions are flattened into an array of unions called `bin_code` (defined in `binutils/include/rxdefs.h`). An instruction is represented by an opcode element, immediately followed by elements representing its operands.

The canonical operand format in `rxops.h` is a NUL-terminated signature: one kind code (`R`, `I`, `S`, `L`, and so on) per operand. Consumers iterate that signature instead of switching over a closed `FMT_*` enum. The in-memory `no_ops` field remains a 32-bit `int`, preserving the eight-byte `bin_code` cell; the practical operand-count bound is therefore the bytecode stream's existing `INT_MAX`, not a small instruction-format ceiling.

```
// Example buffer size handling in gen_instr()
context->binary.binary[context->binary.inst_size].instruction.opcode =
    opcode;
context->binary.binary[context->binary.inst_size++].instruction.no_ops
    = operands;

// Operand elements
context->binary.binary[context->binary.inst_size++].index =
    get_reg_number(...);
context->binary.binary[context->binary.inst_size++].iconst = token->
    integer;
```

B.2.2 Constant Pool

Strings, binary literals, pooled float literals, procedure mappings, debug metadata, and exported symbols are packed into `const_pool`. This is a sequential buffer of dynamically sized records. Every record starts with a `chameleon_constant` header dictating its type and byte size. Types include: `STRING_CONST`, `BINARY_CONST`, `FLOAT_CONST`, `PROC_CONST`, `EXPOSE_REG_CONST`, `EXPOSE_PROC_CONST`, `META_FUNC`, `META_INLINE`, `META_REG`, etc.

The serialized `expose_head` chain includes both `EXPOSE_REG_CONST` and `EXPOSE_PROC_CONST` records. Runtime linking and other module-local walkers now rely on that chain instead of scanning the whole constant pool.

The interface/callable-contract work extends that same metadata path rather than introducing a second binary header mechanism. In addition to `META_CLASS` and `META_ATTR`, the assembler now serializes:

- META_INTERFACE for one interface header
- META_IMPLEMENTS for one concrete-class-to-interface link
- META_MEMBER for one interface method or factory declaration

Cross-file compiler inlining also uses the metadata path. Callable signatures remain in META_FUNC; inline-body templates are carried separately in META_INLINE, emitted in RXAS as:

```
.meta "fully.qualified.callable"=".inline" "I6;c,1,..."
```

The I6 payload is the compiler-owned inline transport described in compiler/docs/inlining_design.md. rxas stores it as META_INLINE, and rxdas must emit it back to the same logical .meta ... ".inline" "I6;..." spelling so source, RXAS, and binary import paths do not drift. Linked final images normally strip META_INLINE; library artifacts preserve it for downstream rxc optimisation. The leading c record is a versioned callable proof summary; rxc reconstructs its facts from the transported body and requires exact agreement before enabling imported inlining.

Source/debug metadata now has two separate identities:

- .srcstep step-id is a module-local id for one concrete source anchor: pooled file name, whole source line, active range, and provenance flags. It is not an instruction address or a source line number.
- .srcstep and .traceevent clause-id is a module-local grouping key for all anchors and semantic events that belong to one logical authored clause. Simple clauses often use the same value for both ids; split clauses, generated helper code, loop pieces, or inlined fragments may use several step ids for one clause id.

0 means there is no source/clause anchor. In a linked image, consumers should treat identity as module plus id because ids are only local to their original module.

The RXAS spellings are:

```
.srcstep <step-id> <clause-id> <flags> "<file>" <line> <start-col> <end
    -col> "<whole-source-line>"
.traceevent "<kind>" <mode-mask> "<value-source>" "<value-type>" "<
    register-type>" <value-ref> <source-step-id> <clause-id> <flags> "<
    symbol>" "<resolved-name>"
```

TRACE event records deliberately store compact codes, not presentation strings. Current event kinds include A assignment, V variable read, C resolved compound

name, `L` literal, `O` binary operation, `P` prefix operation, `F` function result, and `M` message. The TRACE exit handler maps those to classic prefixes such as `>=>`, `>V>`, `>C>`, `>L>`, `>O>`, `>P>`, `>F>`, and `+++`. The mode mask controls visibility in modes such as Results and Intermediates. `value-source` is `R` for a register, `K` for a constant-pool value, or `N` when no value is attached. Register-backed events must name an actual available register at that address; handlers must not infer values from source text or scope metadata.

Metadata-only modules are valid. For example, an interface contract file may compile to `.rxas` containing `.meta` records and no function bodies; `rxas` must still emit a `.rxbin` so import and runtime factory resolution can load the contract metadata.

For interface methods, the member-kind string now distinguishes: `- method` for an abstract interface method `- method final` for a Level B final/default interface method body

B.2.3 Jump Tables

Procedure-local `.jtable` declarations and same-line `.jcase` label decorations are resolved after label optimization. The assembler emits the resulting table as a host-layout-independent `BINARY_CONST`; `jumps`, `jumpr`, `jumpn`, `jumpb`, `jumpbs`, and `jumpi` operands are patched to that pool entry. `jumps` uses exact bytes, `jumpr` canonicalizes trailing-space-padded Rexx text, and `jumpn` canonicalizes numeric strings through the shared parser in `binutils/include/rxnumparse.h`. Release 1 packed algorithms are `linear`, `openhash`, and `acph`. `auto` uses linear for one case. Larger tables use open hashing for average key lengths up to two bytes; tables of at least 256 cases also use it for averages up to four bytes. Remaining tables use ACPH. Shared format constants and hash functions live in `binutils/include/rxjtable.h` and must remain common to the assembler and VM readers.

The complete packed layout, canonicalization, corruption, disassembly, and ownership contract is in `RXBIN_JUMP_TABLES.md`.

B.2.4 Binary Literals

RXAS binary literals are written as byte-paired hex with a `0x` or `0X` prefix. `0x00ff` stores two bytes (`00 ff`) in a `BINARY_CONST`, and `0x` stores an empty binary con-

stant. rxdas emits the canonical lowercase 0x... spelling.

load rDst,0x... uses LOAD_REG_BINARY and loads the VM register's binary slot, not its string slot. Binary-memory VM instructions exposed at RXAS are:

- blen rOut,rBin
- blen rOut,0x...
- bresize rBin,rLength
- bclear rBin
- bfill rBin,rByte
- getbyte rOut,rBin,rIndex
- setbyte rBin,rIndex,rByte
- bcopy rDst,rSrc
- bcopy rDst,rSrc,rOffset
- bcopy rDst,0x...,rOffset
- bgetu8 rOut,rBin,rOffset
- bgeti8 rOut,rBin,rOffset
- bgetu16 rOut,rBin,rOffset
- bgeti16 rOut,rBin,rOffset
- bgetu32 rOut,rBin,rOffset
- bgeti32 rOut,rBin,rOffset
- bgeti64 rOut,rBin,rOffset
- bgetf32 rOut,rBin,rOffset
- bgetf64 rOut,rBin,rOffset
- bgetu8 rOut,0x...,rOffset
- bgeti8 rOut,0x...,rOffset
- bgetu16 rOut,0x...,rOffset
- bgeti16 rOut,0x...,rOffset
- bgetu32 rOut,0x...,rOffset
- bgeti32 rOut,0x...,rOffset
- bgeti64 rOut,0x...,rOffset
- bgetf32 rOut,0x...,rOffset
- bgetf64 rOut,0x...,rOffset
- bsetu8 rBin,rOffset,rValue
- bseti8 rBin,rOffset,rValue

- `bsetu16 rBin,rOffset,rValue`
- `bseti16 rBin,rOffset,rValue`
- `bsetu32 rBin,rOffset,rValue`
- `bseti32 rBin,rOffset,rValue`
- `bseti64 rBin,rOffset,rValue`
- `bsetf32 rBin,rOffset,rFloat`
- `bsetf64 rBin,rOffset,rFloat`
- `bcheckrange rBin,rOffset,rLen`
- `bgets rString,rBin,rOffset`
- `bgets rString,0x...,rOffset`
- `bsets rBin,rOffset,rString`
- `bsets rBin,rOffset,"literal"`
- `sget rString,"literal",rOffset`
- `bmove rDst,rSrc,rLen`
- `bmemmove rBin,rDstOffset,rLen`
- `bcmpb rCmp,rBin,rNeedle`
- `bcmpb rCmp,0x...,rNeedle`
- `bcmpb rCmp,rBin,0x...`
- `bcmpb rCmp,0x...,0x...`
- `bcmps rCmp,rBin,rString`
- `bcmps rCmp,rBin,"literal"`
- `bcmps rCmp,0x...,rString`
- `bcmps rCmp,0x...,"literal"`
- `bconcat rDst,rLeft,rRight`
- `bappend rDst,rRight`
- `setbinpos rBin,rOffset`
- `getbinpos rOut,rBin`
- `bslice rDst,rSrc,rLen`
- `bupdate rDst,rOffset,rSrc`
- `stobin rReg`
- `bintos rReg`

Indexes, offsets, and lengths are bytes and zero-based. The typed bget* and bset* instructions use canonical little-endian storage order; they do not use host-native struct layout, alignment, or padding. bgetu8 is the strict counterpart to current getbyte: getbyte still returns -1 for out-of-range reads, while typed reads raise OUT_OF_RANGE. Typed writes raise OUT_OF_RANGE when the field does not fit or when an integer value is outside the target storage type's range. bgeti64/bseti64 are the Release 1 .int storage form. bgetf32/bsetf32 store IEEE binary32 values and widen/narrow through the VM float register; bgetf64/bsetf64 store IEEE binary64 values.

bresize sets the logical byte length and zero-fills newly exposed bytes. The VM may keep a larger private physical capacity, grown in blocks and reused across later logical resizes; blen reports only the logical length. Negative lengths raise OUT_OF_RANGE, and allocation failure raises FAILURE. bclear sets the logical byte length and cursor to zero. bfill fills the current logical byte range and raises OUT_OF_RANGE for bytes outside 0..255.

bcopy rDst,rSrc,rOffset copies exactly the current logical length of rDst from a binary register or constant source at a byte offset. bcheckrange is a strict side-effect-free range check for rOffset..rOffset + rLen; negative offsets, negative lengths, and ranges past the logical binary length raise OUT_OF_RANGE. bgets/bsets read and write zero-terminated UTF-8 fields in binary memory, including the stored NUL on writes but excluding it from the Rexx string value on reads. sget extracts a codepoint-counted slice from a STRING_CONST, starting at a byte offset that must be a UTF-8 boundary. bcmpb and bcmps compare without materializing a temporary copy; the compare register supplies the input source offset and is overwritten with -1, 0, or 1.

setbinpos, getbinpos, and bslice are legacy cursor instructions retained for compatibility and covered by rxasbinarylegacytests. New Release 1 source lowering uses target-sized bcopy, typed reads/writes, and zero-copy compares instead of cursor-based extraction. setbyte and bupdate are strict and raise OUT_OF_RANGE for invalid indexes, bytes outside 0..255, or overlay writes past the destination length. stobin copies the register's current string bytes into its binary slot. bintos validates the register's current binary bytes as UTF-8 and copies them into its string slot; invalid bytes raise UNICODE_ERROR in UTF builds.

Level B exposes these byte-buffer instructions through the rxfnsb binary helpers

(binlength, binbyte, binsetbyte, binsubstr, binconcat, binoverlay, bininsert, bindelstr, binpos, bincompare, bin2x, x2bin). The source-level || operator also lowers to bconcat when either operand is .binary; blank concat remains a text-only operation.

B.2.5 Attribute Array Helpers

RXAS exposes VM attribute-array operations for array/object backing storage:

- getattrs rOut,rArray
- setattrs rArray,rCount / setattrs rArray,count
- minattrs rArray,rCount / minattrs rArray,count
- linkattr / linkattr1, linktoattr / linktoattr1, and unlinkattr / unlinkattr1
- insattrs rArray,index,count and delattrs rArray,index,count
- insattrs1 rArray,index,count and delattrs1 rArray,index,count

The forms without a 1 suffix use zero-based indexes. The *1 forms use one-based indexes and are the usual fit for Level B array BIFs. Bulk insert accepts an index in 0..num_attributes (1..num_attributes+1 for *1) and inserts before that position. Bulk delete is strict: the requested range must fit inside the current attributes. Passing a count of zero is a no-op, but the index still must be in range. Invalid ranges raise OUT_OF_RANGE.

B.2.6 Reference Helpers

RXAS exposes the VM-first reference surface before any Rexx source syntax is finalized:

- mkref rRef,rSource creates a reference value to the storage currently named by rSource, after existing local links have resolved to their target storage.
- deref rDest,rRef copies the current referenced value into rDest. This is a snapshot copy, including nested attributes, not a live alias.
- linkref rLocal,rRef links a local register to the referenced storage until the existing unlink rLocal restores its base storage.
- setref rRef,rSource copies rSource into the referenced storage.
- refvalid rOut,rRef stores 1 when the reference is still valid and 0 otherwise.

- `unref rRef` clears a reference value and releases its reference-cell retain.

Invalid reference use raises `REFERENCE_INVALID`; `refvalid` is the non-raising probe. A reference becomes invalid when its target storage lifetime ends, such as deleted/shrunk attribute storage or frame-owned local storage after return. Plain overwrite of still-live storage keeps the reference valid. These instructions are optimiser barriers until reference alias effects are modelled more deeply. Assigning a scalar value into the register that holds a reference value clears that register's reference payload; it does not invalidate the referenced target storage.

In UTF builds, RXAS string constants are text, not byte containers. Hand-written string operands are unescaped and validated before entering the constant pool; invalid UTF-8 is an assembly error with guidance to use `0x...` binary literals. Use `load rDst,0x...`, `freadb`, `fwriteb`, `sockrecvb`, or `socksendb` for arbitrary bytes.

B.2.7 Symbol Tracking (AVL Trees)

To deduplicate constants and resolve identifiers in `O(log N)` time, the assembler leverages a custom AVL tree implementation (`avl_tree.h`). Active trees include: `- string_constants_tree - decimal_constants_tree - binary_constants_tree - label_constants_tree - proc_constants_tree - extern_constants_tree`

B.3 3. Two-Pass Resolution (Backpatching)

Assembly requires a two-pass approach because a jump or call can reference a label or procedure defined further down in the file. `rxasassm.c` handles this elegantly by building a `struct backpatching` header for every symbol encountered:

```
struct backpatching {
    int defined;          // 1 if definition encountered, 0 if only
        referenced
    size_t index;         // Final resolved target index in the binary
        or const_pool
    struct backpatching_references *refs; // Linked list of unresolved
        usages
};

struct backpatching_references {
    size_t index;         // The instruction operand index in the
        bin_code array
```

```
    Assembler_Token *token;
    struct backpatching_references *link;
};
```

When the EOF is reached, the backptch(Assembler_Context *context) function orchestrates the final resolution:

1. **optimise_labels()**: Performs peephole optimization. If a label simply targets an unconditional branch (br), the target of the label is recursively collapsed to point directly at the final destination, saving execution cycles.
2. **backpatch_procedures()**: Walks the proc_constants_tree. Raises errors for any procedures referenced but never defined. Resolves valid procedures.
3. **backpatch_labels()**: Walks the label_constants_tree. Updates the placeholder bin_code indices mapped in the refs linked list with the actual instruction offsets.

B.4 4. Binary Emission (.rxbin)

Once parsing and backpatching conclude without errors, the assembler flushes the Assembler_Context state into an RXBIN 007 container.

The structural output consists of:

1. **Magic Header & Version**: Validates the file format.
2. **Module Directory and Section Sizes**: the module identity, global-register count (context->binary.globals), six fixed section ranges, and exact stored sizes.
3. **Portable Constants and Metadata**: the native in-memory pool is converted to fixed-width, little-endian, ID-linked records.
4. **Bytecode and Semantic Graph**: the resolved instruction slots are variable-integer encoded, constant operands become record IDs, graph-bearing operands become dense graph IDs, and graph facts/indexes occupy their own sections.

As of format version 002 and later, float operands are no longer stored inline as raw double payloads in operand slots. Instead, the operand slot carries an index to a deduplicated FLOAT_CONST record in the constant pool.

rxdas emits float operands and FLOAT_CONST pool entries with enough significant digits to distinguish binary64 values that would otherwise share a six-decimal rendering, while keeping integer-looking values parseable as RXAS float tokens (for example, -0.0 rather than -0).

RXAS still builds the normal in-memory bin_code[] and raw constant pool first. write_module() then emits only 007: signed integers use ZigZag variable integers, portable records replace native pool layouts, and the base sections are uncompressed. There is no legacy-output mode; 006 is rejected and repository artifacts are rebuilt. The 007 container and semantic graph are specified in RXBIN_007_SEMANTIC_GRAPH.md.

The fixed direct-bytecode call forms call1 through call4 name each caller argument register explicitly. RXAS sets RXBIN007_FEATURE_FIXED_CALLS whenever one is present; call with a count register remains the general fallback for higher arity, imported/native and dynamically selected targets. The existing two-operand call result,procedure() remains the zero-argument direct form.

RXAS feeds the common rxbin graph builder from class, interface, implements, member, callable, factory, and signature facts. After ordinary backpatching it resolves local procedure references, finalizes the indexed module graph, and emits the same sectioned 007 container used for linked images, with a one-entry module directory. Unknown external type/member references are complete opaque nodes, not dangling strings.

Parameter and return types remain canonical text on type nodes, including preseeded built-in nodes such as .float; signatures refer to those nodes with module-local dense IDs. Type-, member-, and factory-bearing serialized instruction operands also use module-local graph IDs. RXAS/RXDAS text syntax remains human-readable and descriptor based.

B.5 5. Current Interface-Dispatch Additions

The current interface runtime slice relies on three assembler-visible opcodes and the metadata records above:

- setobjtype rX,"fully.qualified.class" stamps a newly created object value with its concrete runtime class identity and clears any uninitialized-

object marker

- `setobjuninit rX,"fully.qualified.class"` writes a typed object placeholder and marks it uninitialized; the compiler uses this for bare object class defaults such as `.SomeClass`
- `assertinitialized rX` raises `OBJECT_NOT_INITIALIZED` if `rX` is a typed uninitialized object value
- `isinitialized rOut,rX` stores `0` for a typed uninitialized object and `1` otherwise
- `srcmethodsel rProc,rObj,"rxsig1|member|return_type|args"` resolves a concrete procedure from an object value plus a method descriptor
- `srcfprocsel rProc,"rxsig1|fully.qualified.interface|return_type|args",rArgs` resolves the default `*` factory provider for an interface
- `srcfprocsel rProc,"rxsig1|fully.qualified.interface..factory_name|return_type|args",rArgs` resolves a named factory provider for an interface
- `typeof rOut,rObj` returns the canonical source type name of an object value
- `istype rOut,rObj,"type"` checks an object value against an interface, class, or `.object`
- `asserttype rObj,"type"` raises `CONVERSION_ERROR` if the object value does not match the requested interface, class, or `.object`

`istype` and `asserttype` check only object/class/interface compatibility. Initialization is a separate lifecycle check. `assertinitialized` is the raising check; `isinitialized` is the non-raising probe.

`srcfprocsel` supports both the default `*` surface and named factory selectors. Provider selection is a VM concern: the assembler simply emits the opcode and the interface/class metadata needed for runtime lookup. The selector string is a callable descriptor (`rxsig1|name|return_type|args`), which lets the VM and linker reject providers whose metadata has the right name but the wrong signature.

For `call ... , rArgs`, `dcall`, and `srcfprocsel ... , rArgs`, the trailing register operand names the argument-count register. The actual argument values are taken from the contiguous registers immediately after that count register. That hidden contiguous argument block is semantically part of the instruction use set, so optimiser passes must treat those opcodes as barriers unless they fully model the implicit register consumption.

The peephole optimiser also consumes instruction metadata from binutils/include/rxops.h. FLOW_JUMP, FLOW_COND, and FLOW_TERM block NO_HAZARD rule skips automatically. FLG_OPT_BARRIER is for FLOW_NEXT instructions that still must not be skipped, such as calls, signal handler configuration, explicit signal/check instructions, and opaque argument-block operations. FLG_IMPLICIT_REG_USE is for linear instructions whose register effects are not represented as normal operands. For example, inc0, dec0, inc1, and friends are still linear execution, but the optimiser treats them as using the corresponding fixed local register when checking whether an intervening instruction is relevant to a rule.

The opcode-effects inventory is split deliberately without duplicating facts. rxops.h remains authoritative for the dense opcode list, operand format, control flow, opcode flags and source-mnemonic status. binutils/include/rxopeffects.h is the one-to-one semantic sidecar: it has one ordered entry for every opcode slot, including reserved and internal slots. binutils/rxopmeta.c combines both sources through the stable rxop_effects() C API; rxop_effect_count() exposes the mechanically checked inventory size.

RxOpEffects uses the following guarantees:

- state is CLASSIFIED, CONSERVATIVE, RESERVED, or INTERNAL. Conservative, reserved, internal and unknown/out-of-range queries are optimizer barriers and never claim a kill.
- reads and writes are conservative possible-access sets over explicit operand positions. Existing instructions retain the compact three-bit masks; wider instructions use one-bit-per-operand signature strings. Consumers use rxop_effect_reads_operand() and rxop_effect_writes_operand() rather than interpreting either representation directly. A consumer must account for two positions naming the same physical register.
- kills is a proof set: every named operand is definitely overwritten without reading its previous value. It is always a subset of writes and is deliberately narrower than possible writes.
- implicit covers fixed local-register read/modify/write, integer-coded local copy/target registers, argument-register indexes and the runtime-sized local range after operand 3 used by call, dcall, and srcfprocsel.
- branch_targets identifies explicit label operands through the same generic query representation. Packed jump-table instructions instead carry RXOP_-

SEM_INDIRECT_BRANCH.

- semantics covers possible signal/error transfer, direct and dynamic call boundaries, returns, alias creation/release, reference creation/read/write/release, storage-lifetime end, indirect writes, indirect branches and opaque side effects.
- flow and optimizer_barrier are returned in the same object but are derived from the canonical rxops.h flow/flags. A non-classified state also forces the returned barrier bit.

The current inventory has 641 entries: 582 source opcodes (576 classified and six explicitly conservative process/redirect operations), 56 reserved slots and three internal handlers. Coverage is not inferred from an instruction name or format. The audit used the VM handlers between START_OF_INSTRUCTIONS and END_OF_-INSTRUCTIONS, the assembler/compiler behavior described here, and focused semantic tests. The tests mechanically validate table alignment, operand-mask legality, flow/branch/call/implicit consistency, source coverage, reserved/internal treatment and fail-closed unknown queries; they do not claim to parse arbitrary C handler bodies formally.

NR-27 adds a transient whole-procedure machine-flow layer after the unchanged 20-item local peephole and before ordinary RXBIN emission. Stable peephole output is moved, with its token ownership, into a growable procedure stream. The assembler then builds resolved label successors/predecessors, reachability, typed-view liveness and transformation-specific must-availability/may-reach facts over the original queue records. Surviving records still pass through the same assembler emission functions; RXAS syntax, canonical RXBIN and the public ABI do not change.

The NR-27 register universe distinguishes local (r), argument (a) and global (g) register classes and tracks integer, float, string, decimal, binary, attribute and reference views. Explicit and implicit accesses come from rxop_effects(). Register metadata and register-backed TRACE records are observations. Alias/reference/lifetime/indirect/opaque operations taint involved registers and invalidate the relevant facts. Calls and other modeled barriers kill availability rather than excluding an otherwise independent region. Registered branch SIGNAL handlers are conservative asynchronous observation targets for every executable instruction in the procedure. An unresolved label or unknown opcode makes control

flow incomplete and disables all NR-27 rewrites for that procedure. A packed jump-table branch is complete only when its procedure-local declared table, every `.jcase` label and the mandatory miss fallthrough are all present in the same retained procedure stream. The flow graph then includes every case edge plus the miss edge. Undeclared, malformed, cross-procedure, unsupported or off-graph-miss tables remain fail-closed.

The admitted first transformation panel is deliberately narrower than the fact engine:

- `icopy`, `fcopy` and strict-comparison uses of `scopy` may be redirected only when the typed source view is unchanged on every path and every may-reaching destination observation can be redirected atomically; decimal copy remains excluded because it can allocate and signal;
- an exact self `copy` is removed because the VM handler is explicitly a no-op, while nonidentity full-value copies remain excluded pending ownership, payload, attribute, flag and cleanup proof;
- repeated identical integer/bitwise-equal float loads, repeated `null`, and repeated one-register `itof` are removed only under a must-available fact;
- unreachable instructions, TRACE records and source-step records are removed only when the entire procedure CFG is complete; and
- dead-result deletion remains disabled because a nominal numeric destination write may release hidden reference or native-payload state even when its numeric result is dead.

Every admitted rewrite removes at least one executable instruction and inserts none, so fixed-point iteration terminates by a strictly decreasing instruction count. `rxas -d` prints per-candidate accept/reject reasons and a per-procedure summary. `-n` bypasses both the local and whole-procedure optimizers.

NR-18 adds one reverse direction for surviving `icopy`/`fcopy` records. For an immediately adjacent, classified, nonthrowing, single-kill producer, the pass may retarget operand 1 from a disposable local temporary to the copy's final local and delete the copy. The written typed view must be exact; the final view must be dead at the producer boundary; the temporary view must be dead after the copy; the producer must not read the final; and TRACE/source-step, implicit, alias/reference/lifetime, opaque, ownership and asynchronous-handler observations reject the candidate. These conditions make the original and retargeted states equal at

every observable edge while reducing the executable count by one.

Copy rewrites proved from one graph may be applied in the same iteration only when their complete physical source/destination register pairs are disjoint. Their substitutions then commute and cannot invalidate one another's liveness or availability facts. Procedures newly admitted through exact jump-table edges always receive reachability cleanup. Global typed-value analysis is additionally bounded to one million `queue-item x liveness-word` cells for those procedures; above the bound, `rxas -d` reports `scope=reachability-only`. This preserves the linear high-yield cleanup without turning large generated dispatch procedures into quadratic assembly work. Procedures without resolved indirect tables retain the established NR-27 behavior without that bound.

The older compare/branch rule still reads explicit read/kill and implicit register facts through `rxop_effects()` inside the bounded local queue. NR-27 does not change that rule's selection boundary. `test_rxop_metadata` is generated from `op_table` and the complete effects sidecar so instructions cannot be added without a mechanically aligned entry. The NR-27 effects audit also records that decimal literal load and decimal copy may signal, while one- and two-register integer-to-float conversion and float comparison are non-signalling in the current VM handlers.

Attribute/register-view cleanup is also a keyhole concern. A full `copy` already copies the VM value status word, so `copy rA,rB` followed immediately by `acopy rA,rB` is reduced to the full copy for hand-written or legacy RXAS. Compiler-generated RXAS should not emit that pair for typed payload access. A future explicit flag-copy operation would use `acopy`, but there is no current compiler use case. Typed payload copies are `icopy`, `fcopy`, `scopy`, `dcopy`, and `bcopy`; `bcopy` copies only the binary payload and byte cursor, not public/compiler/library status flags. Duplicate `link`/typed-copy/`unlink` reads, and equivalent `linkattr1` reads for the same owner and one-based attribute slot, may reuse the first detached copy when no barrier or mapped-register use intervenes.

Signal support adds action-aware and dynamic-name forms:

- `sigcalla proc(),"NAME"` installs a handler that returns an internal action marker interpreted by the VM as `skip`, `retry`, or `fail`
- `sigpush "NAME"` saves the current handler entry for `NAME` on the current frame's handler stack

- `sigpop "NAME"` restores the most recent saved handler entry for `NAME`
- `sigbrv label,rSignal,"NAME"` installs a branch handler that wraps the raw interrupt object as a Level B `.signal` value in `rSignal` before branching
- `signal rName` raises a signal whose name is read from a string register
- `signal rName,rPayload` raises a dynamic-name signal with a payload object

Unknown dynamic names raise `INVALID_SIGNAL_CODE`. Literal `signal "NAME"` forms are still assembled directly against the static signal table.

Optimiser rule operands use lowercase `r` for a captured register. Uppercase `R`, `G`, and `A` match literal local/global/argument register numbers. The assembler uses this to express rules such as `inc r0 -> inc0` without adding mnemonic-specific C code to the optimiser engine.

Legacy rules retain three inline operand pairs for source compatibility. New superinstruction rules select a variable-length operand-pattern side table; matching and output generation iterate the whole pattern, and captured values are stored in a dynamically grown map rather than ten fixed slots. The wide `cnop` regression exercises nine captured input and output operands through this path.

Opcode arity is likewise defined by the complete format string rather than a three-operand ceiling. RXAS parsing, RXBIN writing and reading, disassembly, link relocation, VM decoding, metadata inspection, compiler emission, and the `rxc ASSEMBLE` path all iterate that declared operand sequence. The NR-09 large-instruction batch uses forms up to eight operands; adding another arity does not require a new fixed-width transport structure.

The NR-09 fused instructions are catalogued in `docs/reference/rxas/instructions/12-large-inst`. They preserve the ordered state changes and signals of their component sequence, including an intermediate write where that write is part of the public instruction contract. Compiler-owned rewrites may instead choose a result-only form when the removed temporary has no source-level observation.

The NR-14 frozen-PARSE instructions are catalogued on the same reference page. `parsewords3`, `parsepos2`, and `parsewords3d` encode complete exact plans in their opcode identities; `parseplan` carries a compiler-generated versioned descriptor in an ordinary string-constant operand. RXBIN 007 writers set `RXBIN007_FEATURE_FROZEN_PARSE` when any of the four opcodes is present. Readers reject an image that uses one without that feature bit, and older readers reject the unknown feature

bit, so a provisional image cannot be silently decoded with different semantics. The descriptor adds no RXBIN section or relocation form.

The NR-15 native-stem family is catalogued under arrays, attributes, references, and objects. `steminit`, `stemget`, `stemset`, `stemreset`, `stemget2`, `stemset2`, `stemsize`, `stemkeyat`, and `stemvalueat` use ordinary register operands and set `RXBIN007_FEATURE_NATIVE_STEM`. The feature declares the instruction contract only: the receiver's private hash-table bytes remain ordinary process-local VM value state and are never serialized as an RXBIN section. Two-segment forms stream `left || "." || right` during lookup and materialize a canonical key only for a proved new insertion.

`typeof`, `istype`, and `asserttype` are object-contract operations. Compiler generated code uses them for object casts/tests/introspection; scalar `typeof`/`is` cases are folded earlier by `rxc`.

Interface default methods use that same path. The assembler does not introduce new opcodes for them; it simply carries `META_MEMBER` kind `method final` and emits the interface method body as an ordinary procedure. The VM method registry then decides whether `srcmethodsel` should bind `class.member` directly or fall back to the interface's emitted default-body procedure.

At the current Level B stage, the runtime selection rule is:

- each provider row may carry an optional resolved class-side `§match` or `§match.member` procedure
- `srcfprocsel` causes the VM to call that effective `match` on every candidate with the same argument list that will later be passed to the selected factory
- if no explicit `match` exists, the candidate behaves as if it had an implicit `match` returning `1`
- candidates with score `<= 0` are rejected
- the highest positive score wins
- tied highest scores break alphabetically by fully qualified concrete class name

B.6 6. Core Socket Instructions

Core TCP sockets are exposed as normal RXAS mnemonics generated from `binutils/include/rxops.h`. They use VM-managed integer handles rather than

OS file descriptors or pointers, so programs cannot accidentally retain a native socket after the VM context is freed. All socket instructions are FLG_OPT_BARRIER because they perform external I/O or mutate context-owned handle state.

The current instruction surface is intentionally raw TCP with optional client TLS connect support:

- socknew rSock
- sockclose rRc,rSock
- sockconnect rSock,rHost,rPort
- sockconnecttls rSock,rHost,rPort
- sockbind rSock,rHost,rPort
- socklisten rRc,rSock,rBacklog
- sockaccept rClient,rServer
- sockshutdown rRc,rSock,rHow
- sockstarttls rRc,rSock,rHost
- socksend rBytes,rSock,rText
- socksendb rBytes,rSock,rBin
- sockrecv rText,rSock,rMax
- sockrecvb rBin,rSock,rMax
- sockpending rBytes,rSock
- socktimeout rRc,rSock,rMillis
- sockblocking rRc,rSock,rFlag
- socknodelay rRc,rSock,rFlag
- sockkeepalive rRc,rSock,rFlag
- sockpeer rText,rSock
- socklocal rText,rSock
- sockstatus rStatus,rSock
- sockerror rText,rSock

sockconnect, sockconnecttls, and sockbind keep the socket handle in their first operand and record the result in the handle's status slot; use sockstatus or sockerror immediately after those operations. The higher-level rxsocket.rexx wrapper turns these into function return codes for ordinary Level B code.

sockconnecttls is the portable client TLS connect primitive. It connects to rHost:rPort, starts TLS before any application bytes are exchanged, and uses rHost for SNI and

certificate name verification. `sockstarttls` remains the lower-level true STARTTLS primitive for protocols that exchange clear-text bytes before TLS; backends that cannot upgrade an existing connection in place return a negative unsupported status instead of reconnecting behind the caller.

Both TLS instructions are present in all builds. When no TLS backend is compiled in, they record a negative socket status; `sockstarttls` also returns that code in `rRc`. Backends are selected at CMake configure time. Fresh builds default to `CREXX_ENABLE_TLS=NETWORK` on Apple platforms, `CREXX_ENABLE_TLS=OPENSSL` on non-Windows Unix-like platforms, and `CREXX_ENABLE_TLS=SCHANNEL` on Windows. The Network backend uses Network.framework, Security.framework, CoreFoundation.framework, and the system trust store for `sockconnecttls`. The OpenSSL backend supports both direct TLS connect and true STARTTLS. The SChannel backend uses Windows SChannel/SSPI and the Windows trust store, and supports both direct TLS connect and true STARTTLS.

C

cREXX Linker Architecture (`rxlink`)

The `rxlink` tool combines one or more assembled `.rxbin` modules into a linked image that keeps module boundaries while deduplicating compatible constant-pool leaves into one shared pool.

C.1 Purpose

Use `rxlink` when you want to:

- bundle a root module with the providers it needs
- turn a loose module set into one deployable linked image
- shrink downstream `rxcpack` / wrapped artifacts by removing duplicated pool entries
- optionally strip source/TRACE debug metadata from deployable images

This is now the normal native-packaging route for the shipped drivers too: `crexx`, `crxc`, `rxpp`, and related wrapped tools link a deployable image first and then pass that linked image to `rxcpack`.

`rxlink` is not a replacement for the VM loader. The output still contains multiple module records, and `rxvm` still performs the final runtime link/load work.

C.2 Output Format

RXLINK reads and writes only RXBIN 007. The complete toolchain moved atomically; 006 inputs are rejected and rebuilt rather than decoded through a compatibility path. The format and semantic graph design are in `RXBIN_007_SEMANTIC_GRAPH.md`.

The 007 output is one fixed-width sectioned container with a module directory, module instruction ranges, shared constant/canonical metadata data, and one image-wide semantic graph. The current shared-pool/module record stream is not carried forward.

Milestone 1 retains metadata-driven module selection. For the selected modules RXLINK rebuilds canonical text-backed type/member/callable/factory nodes, preserves declaration origin, assigns new image-local dense IDs, rewrites graph-bearing instruction references, and rebuilds every adjacency, name, declaration, dispatch, callable, factory, and provider index. It never concatenates module-local indexes.

The common `rxbin` graph library owns structural merge, remap, validation, and fast rule-neutral traversal. Future language-policy adapters may decide inheritance, assignability, override/default conflicts, and provider selection without embedding those language rules in RXBIN itself.

C.3 Selection Model

Inputs are read as a container stream, so one input file may contain a linked multi-module container or concatenated standalone library containers. Module selection then happens in this order:

1. apply `OMIT`
2. apply `INCLUDE`
3. apply explicit `ROOT`
4. if no roots were chosen, select modules containing `main`
5. if there is still no root, select modules from the first input file
6. walk imports, `srcfprocsel` interface references, and interface relationships to pull in required providers

7. reject duplicate selected exports

Selectors match by:

- full module name
- basename
- filename stem with trailing `.rxas` removed
- `input_path::member`

C.4 What The Linker Reads From Metadata

`rxlink` does not need all metadata equally:

- `proc_head` is used to find procedures and detect `main`
- `expose_head` is used to discover imports and exports
- `meta_head` is scanned for `META_INTERFACE` and `META_IMPLEMENTS` so interface definitions and implementations pull each other in
- the instruction stream is scanned for `srcfprocsel` descriptor strings so modules referenced only through runtime interface-factory lookup are still retained
- the instruction stream is scanned for `srcmethodsel` member names. Because the receiver's concrete class can be known only at runtime, modules that expose or declare a matching member name are selected conservatively.
- selected class/interface contracts are then checked against callable descriptors so a provider with the right name but wrong return or argument signature is rejected before output is written.

The linker preserves the metadata chain in output because the VM and tooling still consume it at runtime.

C.5 Constant-Pool Rewriting

Leaf constants are deduplicated across selected modules when their serialized bytes match:

- `STRING_CONST`

- BINARY_CONST
- DECIMAL_CONST
- FLOAT_CONST

Structured constants are rewritten into the shared pool with all referenced offsets updated:

- procedures
- exposed register/procedure entries
- metadata entries
- instruction operands that point into the pool

Instruction rewriting derives its operand count and kind from the canonical variable-length opcode signature. Linker scans and remaps therefore have no three-operand format switch; wide instructions retain every inline operand while constant and graph references are rewritten normally.

C.6 Strip Support

Current conservative strip support has two independent axes:

- CLI: -s
- control file: STRIP SOURCE
- CLI: -i
- control file: PRESERVE INLINE / STRIP INLINE

STRIP SOURCE removes:

- META_SOURCE_STEP
- META_TRACE_EVENT

Trace-event metadata is source-level debugging metadata. Without source-step anchors, classic TRACE value events are not coherent enough to keep in a deployable stripped image, and they may still expose variable names, compound names, constants, or live values. Keep the linked image unstripped when TRACE, RXDB source stepping, or source-level diagnostics are needed.

Inline-body metadata is different from runtime contract metadata. It is useful to libraries consumed by rxc, but it is not needed once a final linked image has been

built. `rxlink` therefore strips `META_INLINE` by default. Use `-i` or `PRESERVE INLINE` only for diagnostic/tooling builds that need to inspect the inline transport after linking.

It intentionally does not remove runtime contract metadata such as:

- `META_CLASS`
- `META_ATTR`
- `META_INTERFACE`
- `META_IMPLEMENTS`
- `META_MEMBER`

That keeps interface/class dispatch and metadata-aware tooling behaviour stable while still removing source text/file path payloads and source-level TRACE value metadata.

C.7 Control Files

Supported directives are:

- `INPUT path`
- `ROOT selector`
- `INCLUDE selector`
- `OMIT selector`
- `OUTPUT path`
- `MAP path`
- `STRIP SOURCE`
- `STRIP INLINE`
- `PRESERVE INLINE`

C.8 Testing Guidance

When changing `rxlink`, keep three layers of coverage in mind:

1. format tests: shared-pool/shared-module record layout
2. behavioural tests: linked images run correctly in both `rxvm` and `rxbvm`

3. toolchain tests: rxdas can still disassemble linked images, including stripped ones

For broader confidence, the runtime _opt path is now wired through linked images in normal ctest coverage. For a focused rerun, use:

- ctest -L linked_opt --output-on-failure
- cmake --build <build-dir> --target linked_opt_sweep

Be conservative with stripping. If a proposed change removes anything beyond the current source-level debug set (META_SOURCE_STEP and META_TRACE_EVENT), verify both runtime contract lookup and metadata introspection in rxvm before assuming it is safe.

List of Tables

Index

ISBN 978-94-039116-2-5