





Code Loading



Note

This chapter covers the technical details of package loading. To install packages, use Pkg, Julia's built-in package manager, to add packages to your active environment. To use packages already in your active environment, write import X or using X, as described in the Modules documentation.

Definitions

Julia has two mechanisms for loading code:

- 1. Code inclusion: e.g. include("source.jl"). Inclusion allows you to split a single program across multiple source files. The expression include ("source.jl") causes the contents of the file source.jl to be evaluated in the global scope of the module where the include call occurs. If include("source.jl") is called multiple times, source.jl is evaluated multiple times. The included path, source.jl, is interpreted relative to the file where the include call occurs. This makes it simple to relocate a subtree of source files. In the REPL, included paths are interpreted relative to the current working directory, pwd().
- 2. Package loading: e.g. import X or using X. The import mechanism allows you to load a package i.e. an independent, reusable collection of Julia code, wrapped in a module—and makes the resulting module available by the name X inside of the importing module. If the same X package is imported multiple times in the same Julia session, it is only loaded the first time—on subsequent imports, the importing module gets a reference to the same module. Note though, that import X can load different packages in different contexts: X can refer to one package named X in the main project but potentially to different packages also named X in each dependency. More on this below.

Code inclusion is quite straightforward and simple: it evaluates the given source file in the context of the caller. Package loading is built on top of code inclusion and serves a different purpose. The rest of this chapter focuses on the behavior and mechanics of package loading.

A package is a source tree with a standard layout providing functionality that can be reused by other Julia projects. A package is loaded by import X or using X statements. These statements also make the module named X—which results from loading the package code—available within the module where

3/20/21, 09:34 1 of 13

the import statement occurs. The meaning of X in import X is context-dependent: which X package is loaded depends on what code the statement occurs in. Thus, handling of import X happens in two stages: first, it determines what package is defined to be X in this context; second, it determines where that particular X package is found.

These questions are answered by searching through the project environments listed in LOAD_PATH for project files (Project.toml or JuliaProject.toml), manifest files (Manifest.toml or JuliaManifest.toml), or folders of source files.

Federation of packages

Most of the time, a package is uniquely identifiable simply from its name. However, sometimes a project might encounter a situation where it needs to use two different packages that share the same name. While you might be able fix this by renaming one of the packages, being forced to do so can be highly disruptive in a large, shared code base. Instead, Julia's code loading mechanism allows the same package name to refer to different packages in different components of an application.

Julia supports federated package management, which means that multiple independent parties can maintain both public and private packages and registries of packages, and that projects can depend on a mix of public and private packages from different registries. Packages from various registries are installed and managed using a common set of tools and workflows. The Pkg package manager that ships with Julia lets you install and manage your projects' dependencies. It assists in creating and manipulating project files (which describe what other projects that your project depends on), and manifest files (which snapshot exact versions of your project's complete dependency graph).

One consequence of federation is that there cannot be a central authority for package naming. Different entities may use the same name to refer to unrelated packages. This possibility is unavoidable since these entities do not coordinate and may not even know about each other. Because of the lack of a central naming authority, a single project may end up depending on different packages that have the same name. Julia's package loading mechanism does not require package names to be globally unique, even within the dependency graph of a single project. Instead, packages are identified by universally unique identifiers (UUIDs), which get assigned when each package is created. Usually you won't have to work directly with these somewhat cumbersome 128-bit identifiers since Pkg will take care of generating and tracking them for you. However, these UUIDs provide the definitive answer to the question of "what package does X refer to?"

Since the decentralized naming problem is somewhat abstract, it may help to walk through a concrete scenario to understand the issue. Suppose you're developing an application called App, which uses two packages: Pub and Priv. Priv is a private package that you created, whereas Pub is a public package that you use but don't control. When you created Priv, there was no public package by the name Priv. Subsequently, however, an unrelated package also named Priv has been published and become popular.

In fact, the Pub package has started to use it. Therefore, when you next upgrade Pub to get the latest bug fixes and features, App will end up depending on two different packages named Priv—through no action of yours other than upgrading. App has a direct dependency on your private Priv package, and an indirect dependency, through Pub, on the new public Priv package. Since these two Priv packages are different but are both required for App to continue working correctly, the expression import Priv must refer to different Priv packages depending on whether it occurs in App's code or in Pub's code. To handle this, Julia's package loading mechanism distinguishes the two Priv packages by their UUID and picks the correct one based on its context (the module that called import). How this distinction works is determined by environments, as explained in the following sections.

Environments

An *environment* determines what import X and using X mean in various code contexts and what files these statements cause to be loaded. Julia understands two kinds of environments:

- 1. A project environment is a directory with a project file and an optional manifest file, and forms an *explicit environment*. The project file determines what the names and identities of the direct dependencies of a project are. The manifest file, if present, gives a complete dependency graph, including all direct and indirect dependencies, exact versions of each dependency, and sufficient information to locate and load the correct version.
- 2. A package directory is a directory containing the source trees of a set of packages as subdirectories, and forms an *implicit environment*. If X is a subdirectory of a package directory and X/src/X.jl exists, then the package X is available in the package directory environment and X/src/X.jl is the source file by which it is loaded.

These can be intermixed to create a stacked environment: an ordered set of project environments and package directories, overlaid to make a single composite environment. The precedence and visibility rules then combine to determine which packages are available and where they get loaded from. Julia's load path forms a stacked environment, for example.

These environment each serve a different purpose:

- Project environments provide reproducibility. By checking a project environment into version control—e.g. a git repository—along with the rest of the project's source code, you can reproduce the exact state of the project and all of its dependencies. The manifest file, in particular, captures the exact version of every dependency, identified by a cryptographic hash of its source tree, which makes it possible for Pkg to retrieve the correct versions and be sure that you are running the exact code that was recorded for all dependencies.
- Package directories provide convenience when a full carefully-tracked project environment is unnecessary. They are useful when you want to put a set of packages somewhere and be able to

directly use them, without needing to create a project environment for them.

• Stacked environments allow for adding tools to the primary environment. You can push an environment of development tools onto the end of the stack to make them available from the REPL and scripts, but not from inside packages.

At a high-level, each environment conceptually defines three maps: roots, graph and paths. When resolving the meaning of import X, the roots and graph maps are used to determine the identity of X, while the paths map is used to locate the source code of X. The specific roles of the three maps are:

roots: name::Symbol → uuid::UUID

An environment's roots map assigns package names to UUIDs for all the top-level dependencies that the environment makes available to the main project (i.e. the ones that can be loaded in Main). When Julia encounters import X in the main project, it looks up the identity of X as roots[:X].

graph: context::UUID → name::Symbol → uuid::UUID

An environment's graph is a multilevel map which assigns, for each context UUID, a map from names to UUIDs, similar to the roots map but specific to that context. When Julia sees import X in the code of the package whose UUID is context, it looks up the identity of X as graph[context][:X]. In particular, this means that import X can refer to different packages depending on context.

paths: uuid::UUID × name::Symbol → path::String

The paths map assigns to each package UUID-name pair, the location of that package's entry-point source file. After the identity of X in import X has been resolved to a UUID via roots or graph (depending on whether it is loaded from the main project or a dependency), Julia determines what file to load to acquire X by looking up paths[uuid, :X] in the environment. Including this file should define a module named X. Once this package is loaded, any subsequent import resolving to the same uuid will create a new binding to the already-loaded package module.

Each kind of environment defines these three maps differently, as detailed in the following sections.



For ease of understanding, the examples throughout this chapter show full data structures for roots, graph and paths. However, Julia's package loading code does not explicitly create these. Instead, it lazily computes only as much of each structure as it needs to load a given package.

Project environments

3/20/21, 09:34 4 of 13

A project environment is determined by a directory containing a project file called Project.toml, and optionally a manifest file called Manifest.toml. These files may also be called JuliaProject.toml and JuliaManifest.toml, in which case Project.toml and Manifest.toml are ignored. This allows for coexistence with other tools that might consider files called Project.toml and Manifest.toml significant. For pure Julia projects, however, the names Project.toml and Manifest.toml are preferred.

The roots, graph and paths maps of a project environment are defined as follows:

The roots map of the environment is determined by the contents of the project file, specifically, its toplevel name and uuid entries and its [deps] section (all optional). Consider the following example project file for the hypothetical application, App, as described earlier:

```
name = "App"
uuid = "8f986787-14fe-4607-ba5d-fbff2944afa9"

[deps]
Priv = "ba13f791-ae1d-465a-978b-69c3ad90f72b"
Pub = "c07ecb7d-0dc9-4db7-8803-fadaaeaf08e1"
```

This project file implies the following roots map, if it was represented by a Julia dictionary:

```
roots = Dict(
   :App => UUID("8f986787-14fe-4607-ba5d-fbff2944afa9"),
   :Priv => UUID("ba13f791-ae1d-465a-978b-69c3ad90f72b"),
   :Pub => UUID("c07ecb7d-0dc9-4db7-8803-fadaaeaf08e1"),
)
```

Given this roots map, in App's code the statement import Priv will cause Julia to look up roots[:Priv], which yields ba13f791-ae1d-465a-978b-69c3ad90f72b, the UUID of the Priv package that is to be loaded in that context. This UUID identifies which Priv package to load and use when the main application evaluates import Priv.

The dependency graph of a project environment is determined by the contents of the manifest file, if present. If there is no manifest file, graph is empty. A manifest file contains a stanza for each of a project's direct or indirect dependencies. For each dependency, the file lists the package's UUID and a source tree hash or an explicit path to the source code. Consider the following example manifest file for App:

```
[[Priv]] # the private one
deps = ["Pub", "Zebra"]
uuid = "ba13f791-ae1d-465a-978b-69c3ad90f72b"
```

```
path = "deps/Priv"
[[Priv]] # the public one
uuid = "2d15fe94-a1f7-436c-a4d8-07a9a496e01c"
git-tree-sha1 = "1bf63d3be994fe83456a03b874b409cfd59a6373"
version = 0.1.5
[[Pub]]
uuid = "c07ecb7d-0dc9-4db7-8803-fadaaeaf08e1"
qit-tree-sha1 = "9ebd50e2b0dd1e110e842df3b433cb5869b0dd38"
version = "2.1.4"
  [Pub.deps]
 Priv = "2d15fe94-a1f7-436c-a4d8-07a9a496e01c"
  Zebra = "f7a24cb4-21fc-4002-ac70-f0e3a0dd3f62"
[[Zebra]]
uuid = "f7a24cb4-21fc-4002-ac70-f0e3a0dd3f62"
git-tree-sha1 = "e808e36a5d7173974b90a15a353b564f3494092f"
version = "3.4.2"
```

This manifest file describes a possible complete dependency graph for the App project:

- There are two different packages named Priv that the application uses. It uses a private package, which is a root dependency, and a public one, which is an indirect dependency through Pub. These are differentiated by their distinct UUIDs, and they have different deps:
 - The private Priv depends on the Pub and Zebra packages.
 - The public Priv has no dependencies.
- The application also depends on the Pub package, which in turn depends on the public Priv and the same Zebra package that the private Priv package depends on.

This dependency graph represented as a dictionary, looks like this:

Given this dependency graph, when Julia sees import Priv in the Pub package—which has UUID c07ecb7d-0dc9-4db7-8803-fadaaeaf08e1—it looks up:

```
graph[UUID("c07ecb7d-0dc9-4db7-8803-fadaaeaf08e1")][:Priv]
```

and gets 2d15fe94-a1f7-436c-a4d8-07a9a496e01c, which indicates that in the context of the Pub package, import Priv refers to the public Priv package, rather than the private one which the app depends on directly. This is how the name Priv can refer to different packages in the main project than it does in one of its package's dependencies, which allows for duplicate names in the package ecosystem.

What happens if import Zebra is evaluated in the main App code base? Since Zebra does not appear in the project file, the import will fail even though Zebra does appear in the manifest file. Moreover, if import Zebra occurs in the public Priv package—the one with UUID 2d15fe94-a1f7-436c-a4d8-07a9a496e01c—then that would also fail since that Priv package has no declared dependencies in the manifest file and therefore cannot load any packages. The Zebra package can only be loaded by packages for which it appear as an explicit dependency in the manifest file: the Pub package and one of the Priv packages.

The paths map of a project environment is extracted from the manifest file. The path of a package uuid named X is determined by these rules (in order):

- 1. If the project file in the directory matches uuid and name X, then either:
- It has a toplevel path entry, then unid will be mapped to that path, interpreted relative to the directory containing the project file.
- Otherwise, uuid is mapped to src/X.jl relative to the directory containing the project file.
- 1. If the above is not the case and the project file has a corresponding manifest file and the manifest contains a stanza matching unid then:
- If it has a path entry, use that path (relative to the directory containing the manifest file).
- If it has a git-tree-sha1 entry, compute a deterministic hash function of uuid and git-tree-sha1—call it slug—and look for a directory named packages/X/\$slug in each directory in the Julia DEPOT_PATH global array. Use the first such directory that exists.

If any of these result in success, the path to the source code entry point will be either that result, the relative path from that result plus src/X.jl; otherwise, there is no path mapping for uuid. When loading X, if no source code path is found, the lookup will fail, and the user may be prompted to install the appropriate package version or to take other corrective action (e.g. declaring X as a dependency).

In the example manifest file above, to find the path of the first Priv package—the one with UUID ba13f791-ae1d-465a-978b-69c3ad90f72b—Julia looks for its stanza in the manifest file, sees that it has a path entry, looks at deps/Priv relative to the App project directory—let's suppose the App code lives in /home/me/projects/App—sees that /home/me/projects/App/deps/Priv exists and therefore loads Priv from there.

If, on the other hand, Julia was loading the *other* Priv package—the one with UUID 2d15fe94-a1f7-436c-a4d8-07a9a496e01c—it finds its stanza in the manifest, see that it does *not* have a path entry, but that it does have a git-tree-sha1 entry. It then computes the slug for this UUID/SHA-1 pair, which is HDkrT (the exact details of this computation aren't important, but it is consistent and deterministic). This means that the path to this Priv package will be packages/Priv/HDkrT/src/Priv.jl in one of the package depots. Suppose the contents of DEPOT_PATH is ["/home/me/.julia", "/usr/local/julia"], then Julia will look at the following paths to see if they exist:

- 1. /home/me/.julia/packages/Priv/HDkrT
- 2. /usr/local/julia/packages/Priv/HDkrT

Julia uses the first of these that exists to try to load the public Priv package from the file packages/Priv/HDKrT/src/Priv.jl in the depot where it was found.

Here is a representation of a possible paths map for our example App project environment, as provided in the Manifest given above for the dependency graph, after searching the local file system:

```
paths = Dict(
    # Priv - the private one:
    (UUID("ba13f791-ae1d-465a-978b-69c3ad90f72b"), :Priv) =>
          # relative entry-point inside `App` repo:
          "/home/me/projects/App/deps/Priv/src/Priv.jl",

# Priv - the public one:
    (UUID("2d15fe94-a1f7-436c-a4d8-07a9a496e01c"), :Priv) =>
          # package installed in the system depot:
          "/usr/local/julia/packages/Priv/HDkr/src/Priv.jl",

# Pub:
    (UUID("c07ecb7d-0dc9-4db7-8803-fadaaeaf08e1"), :Pub) =>
          # package installed in the user depot:
          "/home/me/.julia/packages/Pub/oKpw/src/Pub.jl",

# Zebra:
    (UUID("f7a24cb4-21fc-4002-ac70-f0e3a0dd3f62"), :Zebra) =>
```

```
# package installed in the system depot:
   "/usr/local/julia/packages/Zebra/me9k/src/Zebra.jl",
)
```

This example map includes three different kinds of package locations (the first and third are part of the default load path):

- 1. The private Priv package is "vendored" inside the App repository.
- 2. The public Priv and Zebra packages are in the system depot, where packages installed and managed by the system administrator live. These are available to all users on the system.
- 3. The Pub package is in the user depot, where packages installed by the user live. These are only available to the user who installed them.

Package directories

Package directories provide a simpler kind of environment without the ability to handle name collisions. In a package directory, the set of top-level packages is the set of subdirectories that "look like" packages. A package X is exists in a package directory if the directory contains one of the following "entry point" files:

- X.jl
- X/src/X.jl
- X.jl/src/X.jl

Which dependencies a package in a package directory can import depends on whether the package contains a project file:

- If it has a project file, it can only import those packages which are identified in the [deps] section of the project file.
- If it does not have a project file, it can import any top-level package—i.e. the same packages that can be loaded in Main or the REPL.

The roots map is determined by examining the contents of the package directory to generate a list of all packages that exist. Additionally, a UUID will be assigned to each entry as follows: For a given package found inside the folder X...

- 1. If X/Project.toml exists and has a uuid entry, then uuid is that value.
- 2. If X/Project.toml exists and but does *not* have a top-level UUID entry, uuid is a dummy UUID generated by hashing the canonical (real) path to X/Project.toml.
- 3. Otherwise (if Project.toml does not exist), then uuid is the all-zero nil UUID.

The dependency graph of a project directory is determined by the presence and contents of project files in the subdirectory of each package. The rules are:

- If a package subdirectory has no project file, then it is omitted from graph and import statements in its code are treated as top-level, the same as the main project and REPL.
- If a package subdirectory has a project file, then the graph entry for its UUID is the [deps] map of the project file, which is considered to be empty if the section is absent.

As an example, suppose a package directory has the following structure and content:

```
Aardvark/
    src/Aardvark.jl:
        import Bobcat
        import Cobra
Bobcat/
    Project.toml:
        [deps]
        Cobra = "4725e24d-f727-424b-bca0-c4307a3456fa"
        Dingo = "7a7925be-828c-4418-bbeb-bac8dfc843bc"
    src/Bobcat.jl:
        import Cobra
        import Dingo
Cobra/
    Project.toml:
        uuid = "4725e24d-f727-424b-bca0-c4307a3456fa"
        Dingo = "7a7925be-828c-4418-bbeb-bac8dfc843bc"
    src/Cobra.jl:
        import Dingo
Dingo/
    Project.toml:
        uuid = "7a7925be-828c-4418-bbeb-bac8dfc843bc"
    src/Dingo.jl:
        # no imports
```

Here is a corresponding roots structure, represented as a dictionary:

```
roots = Dict(
    :Aardvark => UUID("00000000-0000-0000-000000000000"), # no project file, ni
    :Bobcat => UUID("85ad11c7-31f6-5d08-84db-0a4914d4cadf"), # dummy UUID based or
    :Cobra => UUID("4725e24d-f727-424b-bca0-c4307a3456fa"), # UUID from project f
    :Dingo => UUID("7a7925be-828c-4418-bbeb-bac8dfc843bc"), # UUID from project f
)
```

Here is the corresponding graph structure, represented as a dictionary:

A few general rules to note:

- 1. A package without a project file can depend on any top-level dependency, and since every package in a package directory is available at the top-level, it can import all packages in the environment.
- 2. A package with a project file cannot depend on one without a project file since packages with project files can only load packages in graph and packages without project files do not appear in graph.
- 3. A package with a project file but no explicit UUID can only be depended on by packages without project files since dummy UUIDs assigned to these packages are strictly internal.

Observe the following specific instances of these rules in our example:

- Aardvark can import on any of Bobcat, Cobra or Dingo; it does import Bobcat and Cobra.
- Bobcat can and does import both Cobra and Dingo, which both have project files with UUIDs and are declared as dependencies in Bobcat's [deps] section.
- Bobcat cannot depend on Aardvark since Aardvark does not have a project file.
- Cobra can and does import Dingo, which has a project file and UUID, and is declared as a dependency in Cobra's [deps] section.

- Cobra cannot depend on Aardvark or Bobcat since neither have real UUIDs.
- Dingo cannot import anything because it has a project file without a [deps] section.

The paths map in a package directory is simple: it maps subdirectory names to their corresponding entry-point paths. In other words, if the path to our example project directory is /home/me/animals then the paths map could be represented by this dictionary:

Since all packages in a package directory environment are, by definition, subdirectories with the expected entry-point files, their paths map entries always have this form.

Environment stacks

The third and final kind of environment is one that combines other environments by overlaying several of them, making the packages in each available in a single composite environment. These composite environments are called *environment stacks*. The Julia LOAD_PATH global defines an environment stack—the environment in which the Julia process operates. If you want your Julia process to have access only to the packages in one project or package directory, make it the only entry in LOAD_PATH. It is often quite useful, however, to have access to some of your favorite tools—standard libraries, profilers, debuggers, personal utilities, etc.—even if they are not dependencies of the project you're working on. By adding an environment containing these tools to the load path, you immediately have access to them in top-level code without needing to add them to your project.

The mechanism for combining the roots, graph and paths data structures of the components of an environment stack is simple: they are merged as dictionaries, favoring earlier entries over later ones in the case of key collisions. In other words, if we have $stack = [env_1, env_2, ...]$ then we have:

```
roots = reduce(merge, reverse([roots1, roots2, ...]))
graph = reduce(merge, reverse([graph1, graph2, ...]))
paths = reduce(merge, reverse([paths1, paths2, ...]))
```

The subscripted $roots_i$, $graph_i$ and $paths_i$ variables correspond to the subscripted environments, env_i , contained in stack. The reverse is present because merge favors the last argument rather than first when there are collisions between keys in its argument dictionaries. There are a couple of noteworthy features of this design:

- 1. The *primary environment*—i.e. the first environment in a stack—is faithfully embedded in a stacked environment. The full dependency graph of the first environment in a stack is guaranteed to be included intact in the stacked environment including the same versions of all dependencies.
- 2. Packages in non-primary environments can end up using incompatible versions of their dependencies even if their own environments are entirely compatible. This can happen when one of their dependencies is shadowed by a version in an earlier environment in the stack (either by graph or path, or both).

Since the primary environment is typically the environment of a project you're working on, while environments later in the stack contain additional tools, this is the right trade-off: it's better to break your development tools but keep the project working. When such incompatibilities occur, you'll typically want to upgrade your dev tools to versions that are compatible with the main project.

Conclusion

Federated package management and precise software reproducibility are difficult but worthy goals in a package system. In combination, these goals lead to a more complex package loading mechanism than most dynamic languages have, but it also yields scalability and reproducibility that is more commonly associated with static languages. Typically, Julia users should be able to use the built-in package manager to manage their projects without needing a precise understanding of these interactions. A call to Pkg.add("X") will add to the appropriate project and manifest files, selected via Pkg.activate("Y"), so that a future call to import X will load X without further thought.

« Embedding Julia Profiling »

Powered by Documenter.jl and the Julia Programming Language.