Introduction

Integrated Development Environments (IDEs) and other more lightweight code editors are by far the most used tool of software developers. Yet, improvements of language intelligence, i.e. code completion, debugging as well as static code analysis and enrichment, have traditionally been subject to both the language and the editor used. Language support is thereby brought to IDEs by the means of platform dependent extensions that require repeated efforts for each platform and hence varied a lot in performance, feature-richness and availability. Recent years have seen different works [refs?] towards editor independent code intelligence implementations and unified language independent protocols one of which being put forward by Microsoft - the Language Server Protocol [ref] which is discussed in greater detail in sec. **??**. These approaches reduced the effort of implementing language intelligence from $\mathcal{O}(E \times L)$ to $\mathcal{O}(1 \times L)$ where E stands for the number of *editors* and L for *languages*. As a side effect this also allows for developers to stay in their preferred developing environment instead of needing to resort to e.g. Vim or Emacs emulation or loosing access to other plugins.

Being independent of the editors, the choice of language to implement language servers in lies with the developer. In effect, it is possible for language developers to integrate essential parts of the existing language implementation for a language server. By now the LSP has become the most popular choice for cross-platform language tooling with implementations [langservers and microsoft] for all major and many smaller languages.

Speaking of smaller languages is significant, as both research communities and industry continuously develop and experiment with new languages for which tooling is unsurprisingly scarce. Additionally, previous research [ref], that shows the importance of language tools for the selection of a language, highlights the importance of tooling for new languages to be adopted by a wider community. While previously implementing language tools that integrate with the developer's environment was practically unfeasible for small projects due to the incompatibility between different extension systems, leveraging the LSP reduces the amount of work required considerably.

Problem Definition

Yet, while many of the implementations are freely available as Open Source Software [ref?], the methodology behind these servers is often poorly documented,

especially for smaller languages. There are some experience reports [ref: merlin, and others] and a detailed video series on the Rust Analyzer[ref or footnote], project, but implementations remain very opinionated and poorly guided through. The result is that new implementations keep repeating to develop existing solutions.

Moreover, most projects do not formally evaluate the Language Server on even basic requirements. Naïvely, that is, the server should be *performant* enough not to slow down the developer, it should offer *useful* information and capabilities and of course be *correct* as well as *complete.*

To guide future implementations of language servers for primarily small scale languages the research presented in this thesis aims to answer the following research questions at the example of the Nickel Project[1]:

RQ.1 How to develop a language server for a new language with the abovementioned requirements in mind?

RQ.2 How can we assess the implementation both quantitatively based on performance measures and qualitatively based on user satisfaction?

RQ.3 Do the methods used to answer RQ.1 meet the expected requirements under the assessment developed in RQ.2?

Goals

The goal of this research is to describe a reusable approach for representing programs that can be used to query data to answer requests on the Language Server Protocol efficiently. The research is conducted on an implementation of the open source language Nickel[^https://nickel-lang.org] which provides the *Diagnostics*, *Jump to ** and *Hover* features as well as limited *Auto-Completion* and *Symbol resolution.* Although implemented for and with integration of the Nickel runtime, the goal is to keep the internal format largely language independent. Similarly, the Rust based implementation should be described abstractly enough to be implemented in other languages. To support the chosen approach, a user study will show whether the implementation is able to meet the expectations of its users and maintain its performance in real-world scenarios.

Non-Goals

The reference solution portrayed in this work is specific for the Nickel language. Greatest care is given to present the concepts as generically and transferable as possible. However, it is not a goal to explicitly cover a problem space larger than the Nickel language, which is a pure functional language based on lambda calculus featuring recursive record types and optional typing.

Research Methodology

What are the scientific methods

---

[1] https://nickel-lang.org

Structure of the thesis

Background

This thesis illustrates an approach of implementing a language server for the Nickel language which communicates with its clients, i.e. editors, over the open Language Server Protocol (in the following abbreviated as *LSP*). The current chapter provides the background on the technological details of the project. As the work presented aims to be transferable to other languages using the same methods, this chapter will provide the means to distinguish the nickel specific implementation details.

The primary technology built upon in this thesis is the language server protocol. The first part of this chapter introduces the LSP, its rationale and improvements over classical approaches, technical capabilities and protocol details. The second part is dedicated to Nickel, elaborating on the context and use-cases of the language followed by an inspection of the technical features of Nickel.

Language Server Protocol

Language servers are today's standard of integrating support for programming languages into code editors. Initially developed by Microsoft for the use with their polyglot editor Visual Studio Code[2] before being released to the public in 2016 by Microsoft, RedHat and Codeenvy, the LSP decouples language analysis and provision of IDE-like features from the editor. Developed under open source license on GitHub[3], the protocol allows developers of editors and languages to work independently on the support for new languages. If supported by both server and client, the LSP now supports more than 24 language features[4] including code completion, code navigation facilities, contextual information such as types or documentation, formatting, and more

Motivation

Since its release, the LSP has grown to be supported by a multitude of languages and editors(**lsp-website?**), solving a long-standing problem with traditional IDEs.

---

[2]https://code.visualstudio.com/

[3]https://github.com/microsoft/language-server-protocol/

[4]https://microsoft.github.io/language-server-protocol/specifications/specification-current/

Before the inception of language servers, it was the editors' individual responsibility to implement specialized features for any language of interest. Under the constraint of limited resources, editors had to position themselves on a spectrum between specializing on integrated support for a certain subset of languages and being generic over the language providing only limited support. As the former approach offers a greater business value, especially for proprietary products most professional IDEs gravitate towards excellent (and exclusive) support for single major languages, i.e. XCode and Visual Studio for the native languages for Apple and Microsoft Products respectively as well as JetBrains' IntelliJ platform and RedHat's Eclipse. Problematically, this results in less choice for developers and possible lock-in into products subjectively less favored but unique in their features for a certain language. The latter approach was taken by most text editors which in turn offered only limited support for any language.

Popularity statistics[5] shows that except Vim and Sublime Text, both exceptional general text editors, the top 10 most popular IDEs were indeed specialized products. The fact that some IDEs are offering support for more languages through (third-party) extensions, due to the missing standards and incompatible implementing languages/APIs, does not suffice to solve the initial problem that developing any sort of language support requires redundant resources.

This is especially difficult for emerging languages, with possibly limited development resources to be put towards the development of language tooling. Consequently, community efforts of languages any size vary in scope, feature completeness and availability.

The Language Server Protocol aims to solve this issue by specifying a JSON-RPC[^Remote Procedure Call] API that editors (clients) can use to communicate with language servers. Language servers are programs that implement a set of IDE features for one language and exposing access to these features through the LSP, allowing to focus development resources to a single project that is above all unrelated to editor-native APIs for analytics processing code representation and GUI integration. Consequently, now only a single implementation of a language server is required, instead of one for each editor and editor maintainers can concentrate on offering the best possible LSP client support to their product independent of the language.

JSON-RPC

JSON-RPC (v2) (**json-rpc?**) is a JSON based lightweight transport independent remote procedure call protocol used by the LSP to communicate between a language server and a client.

The protocol specifies the general format of messages exchanges as well as different kinds of messages. The following snippet lst. 0.1 shows the schema for request messages.

The main distinction in JSON-RPC are *Requests* and *Notifications*. Messages with an `id` field present are considered *requests*. Servers have to respond to requests with a message referencing the same `id` as well as a result, i.e. data or error. If the client does not require a response, it can omit the `id` field sending a

---

[5]https://web.archive.org/web/20160625140610/https://pypl.github.io/IDE.html

Listing 0.1 JSON-RPC Request

```
// Requests
{
  "jsonrpc": "2.0"
, "method": String
, "params": List | Object
, "id": Number | String | Null
}
```

notification, which servers cannot respond to, with the effect that clients cannot know the effect nor the reception of the message.

Responses as shown in lst. 0.2, have to be sent by servers answering to any request. Any result or error of an operation is explicitly encoded in the response. Errors are represented as objects specifying the error kind using an error `code` and providing a human-readable descriptive `message` as well as optionally any procedure defined `data`.

Listing 0.2 JSON-RPC Response and Error

```
// Responses
{
  "jsonrpc": "2.0"
  "result": any
  "error": Error
, "id": Number | String | Null
}
```

Clients can choose to batch requests and send a list of request or notification objects. The server should respond with a list of results matching each request, yet is free to process requests concurrently.

JSON-RPC only specifies a message protocol, hence the transport method can be freely chosen by the application.

Commands and Notifications

The LSP build on top of the JSON-RPC protocol described in the previous subsection.

File Notification

Diagnostics

Hover

Completion

Go-To-*

Symbols

code lenses

Shortcomings

Configuration programming languages

Nickel (**nickel?**), the language targeted by the language server detailed in this thesis, defines itself as "configuration language" used to automize the generation of static configuration files.

Static configuration languages such as XML(**xml?**), JSON(**json?**), or YAML(**yaml?**) are language specifications defining how to textually represent structural data used to configure parameters of a program[6]. Applications of configuration languages are ubiquitous especially in the vicinity of software development. While XML and JSON are often used by package managers (**composer?**), YAML is a popular choice for complex configurations such as CI/CD pipelines (**gitlab-runner?**) or machine configurations in software defined networks such as Kubernetes and docker compose.

Such static formats are used due to some significant advantages compared to other formats. Most strikingly, the textual representation allows inspection of a configuration without the need of a separate tool but a text editor and be version controlled using VCS software like Git. For software configuration this is well understood as being preferable over databases or other binary formats. Linux service configurations (files in `/etc`) and MacOS `*.plist` files which can be serialized as XML or a JSON-like format, especially exemplify that claim.

Yet, despite these formats being simple to parse and widely supported (**json?**), their static nature rules out any dynamic content such as generated fields, functions and the possibility to factorize and reuse. Moreover, content validation has to be developed separately, which led to the design of complementary schema specification languages like json-schema (**json-schema?**) or XSD (**xsd?**).

These qualities require an evaluated language. In fact, some applications make heavy use of config files written in the native programming language which gives them access to language features and existing analysis tools. Examples include JavaScript frameworks such as webpack (**webpack?**) or Vue (**vue?**) and python package management using `setuptools`(**setuptools?**).

Despite this, not all languages serve as a configuration language, e.g. compiled languages and some domains require language agnostic formats. For particularly complex products, both language independence and advanced features are desirable. Alternatively to generating configurations using high level languages, this demand is addressed by more domain specific languages. Dhall (**dhall?**), Cue (**cue?**) or jsonnet (**jsonnet?**) are such domain specific languages (DSL), that offer varying support for string interpolation, (strict) typing, functions and validation.

---

[6]some of the named languages may have been designed as a data interchange format which is absolutely compatible with also acting as a configuration language

Infrastructure as Code

A prime example for the application of configuration languages are IaaS[7] products. These solutions arise highly complex solutions with regard to resource provision (computing, storage, load balancing, etc.), network setup and scaling. Although the primary interaction with those systems is imperative, maintaining entire applications' or company's environments manually comes with obvious drawbacks.

Changing and undoing changes to existing networks requires intricate knowledge about its topology which in turn has to be meticulously documented as a significant risk for *config drift*. Beyond that, interacting with a system through its imperative interfaces demands qualified skills of specialized engineers.

The concept of "Infrastructure as Code" (*IaC*) serves the DevOps principle of overcoming the need for dedicated teams for *Dev*elvopent and *Op*erations, by allowing to declaratively specify the dependencies, topology and virtual resources. Today various tools with different scopes make it easy to provision complex networks, in a reproducible way. That is setting up the same environment automatically and independently. Optimally, different environments for testing, staging and production can be derived from a common base and changes to configurations are atomic.

As a notable instance, the Nix(**nix?**) ecosystem even goes as far as enabling declarative system and service configuration using NixOps(**nixops?**).

To get an idea of how this would look like, lst. 0.3 shows the configuration for a deployment of the Git based wiki server Gollum(**gollum?**) behind a nginx reverseproxy on the AWS network. Although targeting AWS, Nix itself is platform-agnostic and NixOps supports different backends through various plugins. Configurations like this are abstractions over many manual steps and the Nix language employed in this example allows for even higher level turing-complete interaction with configurations.

Similarly, tools like Terraform(**terraform?**), or Chef(**chef?**) use their own DSLs and integrate with most major cloud providers. The popularity of these products[8], beyond all, highlights the importance of expressive configuration formats and their industry value.

Finally, descriptive data formats for cloud configurations allow mitigating security risks through static analysis. Yet, as recently as spring 2020 and still more than a year later dossiers of Palo Alto Networks' security department Unit 42 show (**pa2020H1?**) show that a majority of public projects uses insecure configurations. This suggests that techniques(**aws-cloud-formation-security-tests?**) to automatically check templates are not actively employed, and points out the importance of evaluated configuration languages which can implement passive approaches to security analysis.

Nickel

---

[7]Infrastructure as a Service

[8]https://trends.google.com/trends/explore?date=2012-01-01%202022-01-01&q=%2Fg%2F11g6bg27fp,CloudFormation

Gradual typing

Row types

Contracts

Listing 0.3 Example NixOps deployment to AWS

```
{
  network.description = "Gollum server and reverse proxy";
  defaults =
    { config, pkgs, ... }:
    {
      deployment.targetEnv = "ec2";
      deployment.ec2.accessKeyId = "AKIA...";
      deployment.ec2.keyPair = "...";
      deployment.ec2.privateKey = "...";
      deployment.ec2.securityGroups = pkgs.lib.mkDefault [ "default" ];
      deployment.ec2.region = pkgs.lib.mkDefault "eu-west-1";
      deployment.ec2.instanceType = pkgs.lib.mkDefault "t2.large";
    };

  gollum =
    { config, pkgs, ... }:
    {
      services.gollum = {
        enable = true;
        port = 40273;
      };
      networking.firewall.allowedTCPPorts = [ config.services.gollum.port ];
    };

  reverseproxy =
    { config, pkgs, nodes, ... }:
    let
      gollumPort = nodes.gollum.config.services.gollum.port;
    in
    {

      deployment.ec2.instanceType = "t1.medium";

      services.nginx = {
        enable = true;
        virtualHosts."wiki.example.net".locations."/" = {
          proxyPass = "http://gollum:${toString gollumPort}";
        };
      };
      networking.firewall.allowedTCPPorts = [ 80 ];
    };
}
```

Design implementation of NLS

This chapter contains a detailed guide through the various steps and components of the Nickel Language Server (NLS). Being written in the same language (Rust(**rust?**)) as the Nickel interpreter allows NLS to integrate existing components for language analysis. Complementary, NLS is tightly coupled to Nickel's syntax definition. Based on that sec. **??** will introduce the main datastructure underlying all higher level LSP interactions and how the AST described in sec. **??** is transformed into this form. Finally, in sec. **??** the implementation of current LSP features is discussed on the basis of the previously reviewed components.

Illustrative example

The example lst. 0.4 shows an illustrative high level configuration of a server. Throughout this chapter, different sections about the NSL implementation will refer back to this example.

Linearization

The focus of the NLS as presented in this work is to implement a working language server with a comprehensive feature set. Prioritizing a sound feature set, NLS takes an eager, non-incremental approach to code analysis, resolving all information at once for each code update (`didChange` and `didOpen` events), assuming that initial Nickel projects remain reasonably small. The analysis result is subsequently stored in a linear data structure with efficient access to elements. This data structure is referred to in the following as *linearization*. The term arises from the fact that the linearization is a transformation of the syntax tree into a linear structure which is presented in more detail in sec. **??**. The implementation distinguishes two separate states of the linearization. During its construction, the linearization will be in a *building* state, and is eventually post-processed yielding a *completed* state. The semantics of these states are defined in sec. **??**, while the post-processing is described separately in sec. **??**. Finally, sec. **??** explains how the linearization is accessed.

States

At its core the linearization in either state is represented by an array of `LinearizationItem`s which are derived from AST nodes during the linearization process as well as state dependent auxiliary structures.

Closely related to nodes, `LinearizationItem`s maintain the position of their AST counterpart, as well as its type. Unlike in the AST, *metadata* is directly associated with the element. Further deviating from the AST representation, the *type* of the node and its *kind* are tracked separately. The latter is used to distinguish between declarations of variables, records, record fields and variable usages as well as a wildcard kind for any other kind of structure, such as terminals control flow elements.

The aforementioned separation of linearization states got special attention. As the linearization process is integrated with the libraries underlying the Nickel interpreter, it had to be designed to cause minimal overhead during normal execution. Hence, the concrete implementation employs type-states(**typestate?**) to separate both states on a type level and defines generic interfaces that allow for context dependent implementations.

At its base the `Linearization` type is a transparent smart pointer(**deref-chapter?**; **smart-pointer-chapter?**) to the particular `LinearizationState` which holds state specific data. On top of that NLS defines a `Building` and `Completed` state.

The `Building` state represents a raw linearization. In particular that is a list of `LinearizationItem`s of unresolved type ordered as they are created through a depth-first iteration of the AST. Note that new items are exclusively appended such that their `id` field is equal to the position at all time during this phase. Additionally, the `Building` state records all items for each scope in a separate mapping.

Once fully built, a `Building` instance is post-processed yielding a `Completed` linearization. While being defined similar to its origin, the structure is optimized for positional access, affecting the order of the `LinearizationItem`s and requiring an auxiliary mapping for efficient access to items by their `id`. Moreover, types of items in the `Completed` linearization will be resolved.

Type definitions of the `Linearization` as well as its type-states `Building` and `Completed` are listed in lsts. 0.5, 0.6, 0.7. Note that only the former is defined as part of the Nickel libraries, the latter are specific implementations for NLS.

Transfer from AST

The NLS project aims to present a transferable architecture that can be adapted for future languages. Consequently, NLS faces the challenge of satisfying multiple goals

1. To keep up with the frequent changes to the Nickel language and ensure compatibility at minimal cost, NLS needs to integrate critical functions of Nickel's runtime
2. Adaptions to Nickel to accommodate the language server should be minimal not to obstruct its development and maintain performance of the runtime.

To accommodate these goals NLS comprises three different parts as shown in fig. 1. The `Linearizer` trait acts as an interface between Nickel and the language server. NLS implements such a `Linearizer` specialized to Nickel which registers nodes and builds a final linearization. As Nickel's type checking implementation was adapted to pass AST nodes to the `Linearizer`. During normal operation the overhead induced by the `Linearizer` is minimized using a stub implementation of the trait.
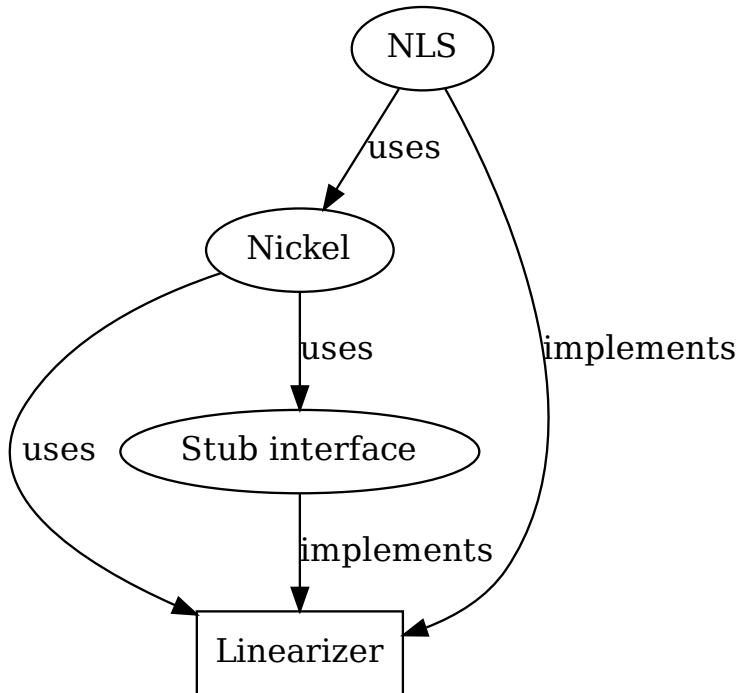

Figure 1: Interaction of Componenets

Usage Graph

At the core the linearization is a simple *linear* structure. Also, in the general case[9] the linearization is reordered in the post-processing step. This makes it impossible to encode relationships of nodes on a structural level. Yet, Nickel's support for name binding of variables, functions and in recursive records implies great a necessity for node-to-node relationships to be represented in a representation that aims to work with these relationships. On a higher level, tracking both definitions and usages of identifiers yields a directed graph.

---

[9]Except single primitive expressions

There are three main kids of vertices in such a graph. **Declarations** are nodes that introduce an identifier, and can be referred to by a set of nods. Referral is represented as **Usage** nodes which can either be bound to a declaration or unbound if no corresponding declaration is known. In practice Nickel distinguishes simple variable bindings from name binding through record fields in recursive records. It also Integrates a **Record** kind to provide deep record destructuring.

During the linearization process this graphical model is recreated on the linear representation of the source. Hence, each `LinearizationItem` is associated with one of the aforementioned kinds, encoding its function in the usage graph. Nodes of the AST that do not fit in a usage graph, a wildcard kind `Structure` is applied.

Scopes

The Nickel language implements lexical scopes with name shadowing.

1. A name can only be referred to after it has been defined
2. A name can be redefined for a local area

An AST can be used to represent this logic. A variable reference always refers to the closest parent node defining the name. Scopes are naturally separated using branching. Each branch of a node represents a sub-scope of its parent, i.e. new declarations made in one branch are not visible in the other.

When eliminating the tree structure, scopes have to be maintained in order to provide auto-completion of identifiers and list symbol names based on their scope as context. Since the bare linear data structure cannot be used to deduce a scope, related metadata has to be tracked separately. The language server maintains a register for identifiers defined in every scope. This register allows NLS to resolve possible completion targets as detailed in sec. **??**.

For simplicity, scopes are represented by a prefix list. Whenever a new lexical scope is entered the prefix list of the outer scope is extended by a unique identifier. With the example in mind lst. 0.4 contains the defintion of a simple record.

Additionally, to keep track of the variables in scope, and iteratively build a usage graph, NLS keeps track of the latest definition of each variable name and which `Declaration` node it refers to.

General Process

From the perspective of the language server, building a linearization is a completely passive process. For each analysis NLS initializes an empty linearization in the `Building` state. This linearization is then passed into Nickel's type-checker along a `Linearizer` instance.

Type checking in Nickel is implemented as a complete recursive depth-first preorder traversal of the AST.

Linearizer

The heart of the linearization the `Linearizer` trait as defined in lst. 0.8. The `Linearizer` lives in parallel to the `Linearization`. Its methods modify on a

shared reference to a `Building Linearization`

`Linearizer::add_term` is used to record a new term, i.e. AST node.

> It's responsibility is to combine context information stored in the `Linearizer` and concrete information about a node to extend the `Linearization` by appropriate items.

`Linearizer::retype_ident` is used to update the type information for a current identifier.

> The reason this method exists is that not all variable definitions have a corresponding AST node but may be part of another node. This is especially apparent with records where the field names part of the record node and as such are linearized with the record but have to be assigned there actual type separately.

`Linearizer::complete` implements the post-processing necessary to turn a final `Building` linearization into a `Completed` one.

> Note that the post-processing might depend on additional data

`Linearizer::scope` returns a new `Linearizer` to be used for a sub-scope of the current one.

> Multiple calls to this method yield unique instances, each with their own scope.
>
> It is the caller's responsibility to call this method whenever a new scope is entered traversing the AST.
>
> Notably, the recursive traversal of an AST ensures that scopes are correctly backtracked.

While data stored in the `Linearizer::Building` state will be accessible at any point in the linearization process, the `Linearizer` is considered to be only *scope safe*. No instance data is propagated back to the outer scopes `Linearizer`. Neither have `Linearizers` of sibling scopes access to each other's data.

Metadata

Records

Static access

Integration with Nickel

Scope

Retyping

Post-Processing

Resolving Elements

Resolving by position

As part of the post-processing step discussed in sec. **??**, the `LinearizationItems` in the `Completed` linearization are reorderd by their occurence of the corresponding AST node in the source file. To find items in this list three preconditions have to hold:

1. Each element has a corresponding span in the source

2. Items of different files appear ordered by `FileId`
3. Two spans are either within the bounds of the other or disjoint.

$$\text{Item}^2_{\text{start}} \geq \text{Item}^1_{\text{start}} \wedge \text{Item}^2_{\text{end}} \leq \text{Item}^1_{\text{end}}$$

4. Items referring to the spans starting at the same position have to occur in the same order before and after the post-processing. Concretely, this ensures that the tree-induced hierarchy is maintained, more precise elements follow broader ones

This first two properties are an implication of the preceding processes. All elements are derived from AST nodes, which are parsed from files retaining their position. Nodes that are generated by the runtime before being passed to the language server are either ignored or annotated with synthetic positions that are known to be in the bounds of the file and meet the second requirement. For all other nodes the second requirement is automatically fulfilled by the grammar of the Nickel language. The last requirement is achieved by using a stable sort during the post-processing.

The algorithm used is listed in lst. 0.9. Given a concrete position, that is a `FileId` and `ByteIndex` in that file, a binary search is used to find the *last* element that *starts* at the given position. According to the aforementioned preconditions an element found there is equivalent to being the most specific element starting at this position. In the more frequent case that no element starting at the provided position is found, the search instead yields an index which can be used as a starting point to iterate the linearization *backwards* to find an item with the shortest span containing the queried position. Due to the third requirement, this reverse iteration can be aborted once an item's span ends before the query. If the search has to be aborted, the query does not have a corresponding `LinearizationItem`.

Resolving by ID

During the building process item IDs are equal to their index in the underlying List which allows for efficient access by ID. To allow similarly efficient access to nodes with using IDs a `Completed` linearization maintains a mapping of IDs to their corresponding index in the reordered array. A queried ID is first looked up in this mapping which yields an index from which the actual item is read.

Resolving by scope

During the construction from the AST, the syntactic scope of each element is eventually known. This allows to map scopes to a list of elements defined in this scope. Definitions from higher scopes are not repeated, instead they are calculated on request. As scopes are lists of scope fragments, for any given scope the set of referable nodes is determined by unifying IDs of all prefixes of the given scope, then resolving the IDs to elements. The Rust implementation is given in lst. 0.10 below.

LSP Server

Diagnostics and Caching

Capabilities

Hover

Completion

Jump to Definition

Show references

Symbols

Listing 0.4 Nickel example with most features shown

```
let Port | doc "A contract for a port number" = contracts.from_predicate (fun value =
  builtins.is_num value &&
  value % 1 == 0 &&
  value >= 0 &&
  value <= 65535) in

let Container = {
  image | Str,
  ports | List #Port,
} in

let NobernetesConfig = {
  apiVersion | Str,
  metadata.name | Str,
  replicas | #nums.PosNat
           | doc "The number of replicas"
           | default = 1,
  containers | { _ : #Container },

} in

let name_ = "myApp" in

let metadata_ = {
    name = name_,
} in

let webContainer = fun image => {
  image = image,
  ports = [ 80, 443 ],
} in

{
  apiVersion = "1.1.0",
  metadata = metadata_,
  replicas = 3,
  containers = {
    "main container" = webContainer  "k8s.gcr.io/#{name_}"
  }
} | #NobernetesConfig
```

Listing 0.5 Definition of Linearization structure

```
pub trait LinearizationState {}

pub struct Linearization<S: LinearizationState> {
    pub state: S,
}
```

Listing 0.6 Type Definition of Building state

```rust
pub struct Building {
    pub linearization: Vec<LinearizationItem<Unresolved>>,
    pub scope: HashMap<Vec<ScopeId>, Vec<ID>>,
}
impl LinearizationState for Building {}
```

Listing 0.7 Type Definition of Completed state

```rust
pub struct Completed {
    pub linearization: Vec<LinearizationItem<Resolved>>,
    scope: HashMap<Vec<ScopeId>, Vec<ID>>,
    id_to_index: HashMap<ID, usize>,
}
impl LinearizationState for Completed {}
```

Listing 0.8 Interface of linearizer trait

```rust
pub trait Linearizer {
    type Building: LinearizationState + Default;
    type Completed: LinearizationState + Default;
    type CompletionExtra;

    fn add_term(
        &mut self,
        lin: &mut Linearization<Self::Building>,
        term: &Term,
        pos: TermPos,
        ty: TypeWrapper,
    )

    fn retype_ident(
        &mut self,
        lin: &mut Linearization<Self::Building>,
        ident: &Ident,
        new_type: TypeWrapper,
    )

    fn complete(
        self,
        _lin: Linearization<Self::Building>,
        _extra: Self::CompletionExtra,
    ) -> Linearization<Self::Completed>
    where
        Self: Sized,

    fn scope(&mut self) -> Self;
}
```

Listing 0.9 Resolution of item at given position

```rust
impl Completed {
  pub fn item_at(
    &self,
    locator: &(FileId, ByteIndex),
  ) -> Option<&LinearizationItem<Resolved>> {
    let (file_id, start) = locator;
    let linearization = &self.linearization;
    let item = match linearization
      .binary_search_by_key(
        locator,
        |item| (item.pos.src_id, item.pos.start))
    {
      // Found item(s) starting at `locator`
      // search for most precise element
      Ok(index) => linearization[index..]
        .iter()
        .take_while(|item| (item.pos.src_id, item.pos.start) == locator)
        .last(),
      // No perfect match found
      // iterate back finding the first wrapping linearization item
      Err(index) => {
        linearization[..index].iter().rfind(|item| {
          // Return the first (innermost) matching item
          file_id == &item.pos.src_id
          && start > &item.pos.start
          && start < &item.pos.end
        })
      }
    };
    item
  }
}
```

Listing 0.10 Resolution of all items in scope

```rust
impl Completed {
  pub fn get_in_scope(
    &self,
    LinearizationItem { scope, .. }: &LinearizationItem<Resolved>,
  ) -> Vec<&LinearizationItem<Resolved>> {
    let EMPTY = Vec::with_capacity(0);
    // all prefix lengths
    (0..scope.len())
      // concatenate all scopes
      .flat_map(|end| self.scope.get(&scope[..=end])
        .unwrap_or(&EMPTY))
      // resolve items
      .map(|id| self.get_item(*id))
      // ignore unresolved items
      .flatten()
      .collect()
  }
}
```