

Implementacja protokołu LSP dla wybranego środowiska zintegrowanego

(Implementation of the LSP protocol
for selected IDE)

Wiktor Adamski

Praca inżynierska

Promotor: dr Wiktor Zychla

Uniwersytet Wrocławski
Wydział Matematyki i Informatyki
Instytut Informatyki

30 stycznia 2018 r.

Wiktor Adamski

.....

.....

(adres zameldowania)

.....

.....

(adres korespondencyjny)

PESEL:

e-mail:

Wydział Matematyki i Informatyki

stacjonarne studia I stopnia

kierunek: informatyka

nr albumu: 272220

Oświadczenie o autorskim wykonaniu pracy dyplomowej

Niniejszym oświadczam, że złożoną do oceny pracę zatytułowaną *Implementacja protokołu LSP dla wybranego środowiska zintegrowanego* wykonałem/am samodzielnie pod kierunkiem promotora, dr Wiktora Zychli. Oświadczam, że powyższe dane są zgodne ze stanem faktycznym i znane mi są przepisy ustawy z dn. 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (tekst jednolity: Dz. U. z 2006 r. nr 90, poz. 637, z późniejszymi zmianami) oraz że treść pracy dyplomowej przedstawionej do obrony, zawarta na przekazanym nośniku elektronicznym, jest identyczna z jej wersją drukowaną.

Wrocław, 30 stycznia 2018 r.

(czytelny podpis)

Streszczenie

Visual Studio Code to na pierwszy rzut oka prosty edytor kodu, jednakże system rozszerzeń pozwala rozbudować go do pełnoprawnego środowiska programistycznego. Ponieważ istnieje wiele edytorów, a każdy z nich posiadał swój interfejs programistyczny, powstał protokół Language Server Protocol (LSP), który umożliwia napisanie logiki wspomagającej pisanie programu raz i użycie jej w wielu edytorach. Tematem tej pracy jest implementacja rozszerzenia do edytora kodu wykorzystującego ww. protokół.

Visual Studio Code at first seems like a simple code editor, though its extension system allows to expand it into full-featured IDE. As there are many editors, and each of them has its own application interface, Language Server Protocol (LSP) was created, to allow writing helper logic once and using it in many editors. The topic of this thesis is the implementation of code editor extension which uses said protocol.

Spis treści

1. Preliminaria	7
1.1. Visual Studio (Code)	7
1.2. Struktura rozszerzenia Visual Studio Code	7
1.3. Protokół Language Server Protocol	8
1.4. Biblioteka Luaparse i drzewa rozbioru	10
2. Część kliencka rozszerzenia	13
2.1. Komunikacja z serwerem	13
2.2. Dodatkowe funkcjonalności	13
3. Implementacja serwera LSP	15
3.1. Zapytanie Initialize	15
3.2. Komunikaty generujące drzewa	15
3.2.1. Komunikat DidChangeTextDocument	15
3.2.2. Komunikat DidChangeConfiguration	16
3.3. Zapytania przechodzące po drzewie	16
3.3.1. Funkcja TraverseTreeDown	16
3.3.2. Zapytanie Hover	16
3.3.3. Zapytanie GotoDefinition	16
3.3.4. Zapytanie Completion	16
4. Podsumowanie	19
Bibliografia	21

Dodatki	23
A Instrukcja uruchomienia rozszerzenia	23

Rozdział 1.

Preliminaria

Aby zrozumieć jak działa dostarczony serwer, należy wpierw zrozumieć architekturę rozszerzenia w systemie edytora Visual Studio Code, a także na czym polega omawiany protokół. W tym rozdziale poruszone zostaną:

- Visual Studio a Visual Studio Code.
- Struktura wtyczki rozszerzającej działanie edytora.
- Opis protokołu LSP.
- Parser Luaparse.

1.1. Visual Studio (Code)

Wiele osób nie rozróżnia od siebie dwóch produktów Microsoftu. Visual Studio to zintegrowane środowisko programistyczne, nastawione głównie na pisanie programów w języku C#. Visual Studio Code jest natomiast edytorem kodu o otwartym kodzie opartym na silniku renderującym Electron od portalu Github, co sprawia, że jest on dosyć podobny do edytora Atom (również pod względem metodologii wtyczek). Dużą zaletą VS Code jest wbudowane wsparcie dla języków JavaScript i TypeScript, oraz środowiska Node.js, co przełożyło się na znaczny wzrost liczby użytkowników w ciągu 2 lat od powstania edytora (ponad 2,6 miliona aktywnych użytkowników [1])

1.2. Struktura rozszerzenia Visual Studio Code

Sercem każdego rozszerzenia jest plik `package.json`, który przechowuje informacje na temat autora pakietu, warunki jego uruchomienia, a także wszystkie jego zależności:

```

{
  "name": "lua-lang",
  "description": "Lua language support",
  "author": "Wiktor Adamski",
  "license": "MIT",
  "version": "0.0.1",
  "engines": {
    "vscode": "^1.16.0"
  },
  "categories": [
    "Languages"
  ],
  "activationEvents": [
    "onLanguage:lua"
  ],
  "main": "./out/src/extension",
  "contributes": {
    "languages": [
      {
        "id": "lua",
        "aliases": [
          "Lua",
          "lua"
        ],
        "extensions": [
          ".lua",
          ".p8",
          ".rockspec"
        ],
        "configuration": "./language-configuration.json"
      }
    ],
  },
  "dependencies": {
    "vscode": "^1.1.5",
    "vscode-languageclient": "^3.4.2"
  }
}

```

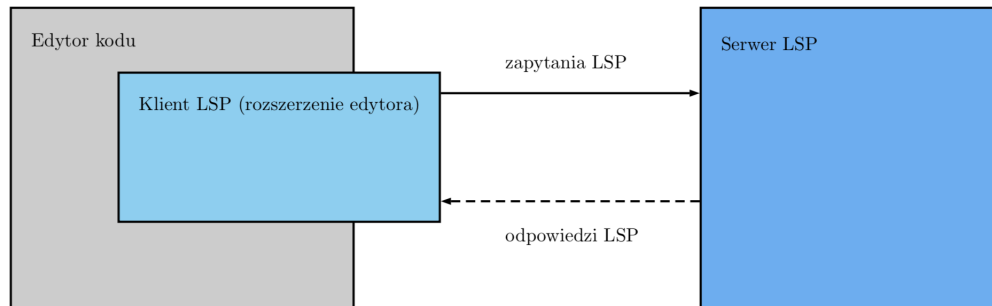
O ile serwer LSP może być napisany w dowolnym języku, część bezpośrednio łącząca się z edytorem (aktualna wtyczka) musi być w języku JavaScript dla środowiska uruchomieniowego Node.js.

Duża część kodu który jest wspólny dla wszystkich rozszerzeń jest możliwa do automatycznego stworzenia przez generator kodu Yeoman. Polecenie `yo code` przeprowadzi nas przez kreator wtyczek i utworzy dodatkowe pliki konfiguracyjne, które pozwolą korzystać z edytora VS Code jako środowiska developerskiego.

1.3. Protokół Language Server Protocol

Protokół LSP [2] jest specjalizacją protokołu JSON-RPC, który przesyła dane między stronami komunikacji za pomocą obiektów JSON. Klient (rozszerzenie edytora kodu) wysyła zapytania do serwera (program wspomagający) odpowiadające

różnym akcjom podejmowanym przez programistę, np. zapytanie się o miejsce deklaracji danej zmiennej.



Pierwszą wiadomością w trakcie połączenia jest wymiana możliwości zarówno klienta (np. czy edytor wspiera przemianowanie zmiennej), jak i serwera (np. wskazanie definicji danego symbolu lub automatyczne uzupełnianie pisanego tekstu). Twórcy protokołu udostępnili bibliotekę ułatwiającą korzystanie z niego w języku TypeScript. Poniżej przedstawiam przykładowy komunikat klienta informujący o otwarciu pliku w edytorze:

```
{
  "jsonrpc": "2.0",
  "id": 1,
  "method": "textDocument/didOpen",
  "params": {
    "uri": "/Users/wiktor/Desktop/test.lua",
    "languageId": "lua",
    "version": 1,
    "text": "w = 'hello world'\nprint(n)"
  }
}
```

W odpowiedzi serwer wyśle następujący komunikat z informacją o braku błędów w pliku:

```
{
  "jsonrpc": "2.0",
  "id": 2,
  "method": "textDocument/publishDiagnostics",
  "params": {
    "uri": "/Users/wiktor/Desktop/test.lua",
    "diagnostics": []
  }
}
```

Kontynuując, klient może wysłać zapytanie na temat definicji danego symbolu:

```
{
  "jsonrpc": "2.0",
  "id": 3,
  "method": "textDocument/definition",
  "params": {
    "uri": "/Users/wiktor/Desktop/test.lua",
    "position": {
      "line": 1,
      "character": 7
    }
  }
}
```

Przy odpowiadaniu serwer musi stwierdzić co znajduje się na podanej w zapytaniu pozycji (w tym przykładzie zmienna `n`) i zwrócić miejsce w pliku gdzie została zadeklarowana:

```
{
  "jsonrpc": "2.0",
  "id": 3,
  "result": {
    "uri": "/Users/wiktor/Desktop/test.lua",
    "range": {
      "start": {
        "line": 0,
        "character": 0
      },
      "end": {
        "line": 0,
        "character": 0
      }
    }
  }
}
```

W powyższej odpowiedzi został zwrócony fragment długości 0, co daje znać edytorowi by zaznaczył całe słowo zawierające daną pozycję.

1.4. Biblioteka Luaparse i drzewa rozbioru

Aby dostarczać jakiegokolwiek sensowne informacje na temat kodu, potrzebne jest jego sparsowanie. Zajmuje się tym biblioteka Luaparse [3], która produkuje abstrakcyjne drzewa rozbioru programów napisanych w języku Lua. Drzewa reprezentowane

za pomocą obiektów JavaScript są inspirowane na specyfikacji Mozilla Parser API. Przykładowo wyrażenie:

```
foo = "bar"
```

zostanie przełożone na następujące drzewo:

tu diagram drzewa rozbioru

Powstałe drzewo jest później wykorzystywane do wyliczania odpowiedzi na zapytania klienta LSP, np. posłuży do odnalezienia miejsca definicji zmiennej o którą pyta się użytkownik rozszerzenia.

Rozdział 2.

Część kliencka rozszerzenia

W przypadku edytora Visual Studio Code, komunikacja między serwerem LSP a edytorem następuje poprzez dodatkowy adapter w postaci osobnego małego rozszerzenia. Dodatkowo, nie wszystkie możliwości edytora są wspierane przez protokół. Część kliencka rozszerzenia zatem ma 2 funkcjonalności:

1. Uruchomienie serwera i komunikacja z nim.
2. Udostępnienie funkcjonalności które nie są wspierane przez protokół LSP.

2.1. Komunikacja z serwerem

Ponieważ zarówno klient jak i serwer LSP są napisane w środowisku Node.js, możliwe było wykorzystanie bibliotek udostępnionych przez twórców edytora, przez co nawiązanie połączenia i jego obsługa ogranicza się do wskazania pliku serwera. Dodatkowo moduł serwera zostaje przy utworzeniu dodany do listy obiektów usuwanych przy zamknięciu rozszerzenia (np. zamknięte zostały wszystkie pliki Lua lub został wyłączony edytor).

2.2. Dodatkowe funkcjonalności

Klient poza przekazywaniem pracy do serwera, może implementować wiele osobnych funkcjonalności. Są to między innymi rejestracja danego języka programowania w słowniku edytora, co pozwala wielu wtyczkom dotyczącym jednego języka na korzystanie ze swoich funkcji nawzajem. Inną ważną funkcjonalnością jest dodanie opisu kolorowania składni. W przypadku VS Code robi się to przez załączenie pliku gramatyki dla edytora TextMate (podejście identyczne jak w przypadku rozszerzeń do edytora Atom). Przykładowo reguła składni:

```

<dict>
  <key>match</key>
  <string>(?!<![\w\d.])\d+(?![pPeE.0-9])</string>
  <key>name</key>
  <string>constant.numeric.integer.lua</string>
</dict>

```

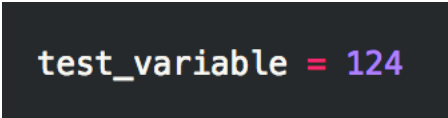
Spowoduje, że fragmentom tekstu pasującym do podanego wyrażenia regularnego (w tym przypadku liczbom całkowitym w systemie dziesiętnym) zostanie przypisana etykieta `constant.numeric.integer`, która zostanie odpowiednio pokolorowana na podstawie wybranego przez użytkownika motywu kolorystycznego. W przypadku prezentowanego motywu:

```

<dict>
  <key>name</key>
  <string>Number</string>
  <key>scope</key>
  <string>constant.numeric</string>
  <key>settings</key>
  <dict>
    <key>foreground</key>
    <string>#AE81FF</string>
  </dict>
</dict>

```

wszystkie fragmenty tekstu, których etykieta zawiera prefiks `constant.numeric` zostaną pokolorowane na kolor fioletowy.



```
test_variable = 124
```

Rozdział 3.

Implementacja serwera LSP

Dysponując parserem kodu Lua, wystarczy odpowiednio przechodzić generowane przez niego drzewa, w celu odpowiedzi na poszczególne zapytania klienta. Punktem wejścia dla projektu będącego częścią niniejszej pracy jest artykuł [4] opisujący utworzenie prostego serwera LSP.

3.1. Zapytanie Initialize

W odpowiedzi na to zapytanie serwer zwraca informacje na temat jego możliwości. W przypadku tej pracy są to:

- Znalezienie definicji danej zmiennej lub funkcji.
- Automatyczne sugestie pisanego kodu.
- Wyświetlenie informacji na temat danego symbolu.

3.2. Komunikaty generujące drzewa

3.2.1. Komunikat `DidChangeTextDocument`

Komunikat `DidChangeTextDocument` informuje serwer, że użytkownik dokonał zmian w danym pliku. Uruchomiony zostaje wtedy parser, który czyta treść rzeźzonego pliku (możliwością jest pracowanie na samej treści zmiany, jednakże wiązałoby się to z koniecznością napisania własnego parsera wspierającego inkrementalne zmiany w parsowanym tekście). Jeżeli parser napotkał jakiś błąd, jest on zwracany z powrotem do klienta i wyświetlany użytkownikowi pod postacią czerwonego podkreślenia problematycznego fragmentu kodu. Przy udanym parsowaniu otrzymane drzewo jest zapisywane w pamięci, aby można było je odwiedzić przy odpowiadaniu na inne zapytania.

3.2.2. Komunikat `DidChangeConfiguration`

Komunikat `DidChangeConfiguration` informuje nas, że użytkownik zmienił ustawienia edytora, co mogło wpłynąć na pliki w niekontrolowany przez nas sposób. W takim przypadku następuje ponowne parsowanie wszystkich otwartych dokumentów analogicznie do komunikatu `DidChangeTextDocument`.

3.3. Zapytania przechodzące po drzewie

3.3.1. Funkcja `TraverseTreeDown`

Każde z zapytań które mają w efekcie przejść się po drzewie rozbiórki dostarcza nam informacje na temat pozycji w pliku na której się znajduje kursor. W takim razie wydzielona została funkcjonalność tłumaczenia pozycji na węzeł drzewa. Ponieważ każdy z węzłów drzewa rozbiórki zawiera informację na temat zakresu tegoż węzła, wystarczy przejść się wgłąb drzewa tak długo jak szukana pozycja znajduje się wewnątrz zakresu przeszukiwanego wierzchołka. Funkcja zwraca odnaleziony wierzchołek.

3.3.2. Zapytanie `Hover`

Zapytanie `Hover` pyta się o możliwe informacje do wyświetlenia gdy użytkownik najedzie kursorem myszy na daną pozycję w pliku. Po znalezieniu najlepiej pasującego wierzchołka za pomocą funkcji `TraverseTreeDown`, zwracany jest komunikat którego dokładna treść jest określana na podstawie typu wierzchołka (np. dla funkcji będzie to jej nazwa i lista argumentów, a dla zmiennej jej typ).

3.3.3. Zapytanie `GotoDefinition`

Zapytanie `GotoDefinition` odpowiada akcji polegającej na szukaniu miejsca zdefiniowania danego symbolu w kodzie. Po znalezieniu szukanego symbolu następuje ponowne przejście po drzewie w celu odnalezienia najpóźniejszej definicji tegoż symbolu (w języku Lua symbole mogą być definiowane na nowo w trakcie działania programu, a także przesłaniane za pomocą słowa kluczowego `local`). Zwracana jest pozycja odnalezionej definicji.

3.3.4. Zapytanie `Completion`

Zapytanie `Completion` jest wysyłane przez klienta w celu odpytania serwera na temat możliwego dokończenia aktualnie pisanego tekstu. Również i w tym przypadku dostarczana jest pozycja kursora, jednakże serwer nie szuka aktualnie edytowanego

wierzchołka, tylko listę symboli które zostały zdefiniowane i są dostępne w danym kontekście. Lista odnalezionych symboli jest później rozszerzana o funkcje i zmienne zdefiniowane w bibliotece standardowej Lua (informacje na ich temat znajdują się w osobnym pliku JSON, który został utworzony na podstawie dokumentacji języka [5]).

Rozdział 4.

Podsumowanie

W ramach niniejszej pracy powstało rozszerzenie programu Visual Studio Code, które wspomaga programistę przy pisaniu kodu. Główna część programu, mianowicie serwer LSP, może zostać użyta przy implementacji analogicznego rozszerzenia dla innych edytorów, bez potrzeby wprowadzania zmian w kodzie. Powstałe rozszerzenie implementuje znaczną część funkcjonalności protokołu LSP, dodanie do niego pozostałych funkcji nie będzie zadaniem trudnym, jedynie czasochłonnym. Zadanie było rozwijające zarówno pod względem pracy z parserem kodu i interpretowaniem jego wyników, ale również pozwoliło prześledzić cały proces rozszerzania funkcjonalności istniejącego programu za pomocą udostępnionego interfejsu.

Bibliografia

- [1] Sean McBreen Visual Studio Code at **Connect(); 2017**, 2017
- [2] Microsoft Language Server Protocol documentation, 2018
- [3] Oskar Schöldström Luaparse, 2013 - 2017
- [4] Microsoft Creating Language Servers for Visual Studio Code, 2018
- [5] Roberto Ierusalimsky, Luiz Henrique de Figueiredo, Waldemar Celes Lua 5.3 Reference Manual, 2015 - 2017

Dodatek A

Instrukcja uruchomienia rozszerzenia

Niniejszy dodatek opisuje instrukcję uruchomienia rozszerzenia w programie Visual Studio Code. Wymagana jest dodatkowo instalacja środowiska Node.js, które musi być dostępne z wiersza poleceń. Kroki do wykonania:

1. Otworzyć katalog **lua** w edytorze VS Code.
2. Nacisnąć kombinację klawiszy **Ctrl+‘**, otwierając tym samym wbudowane okno terminala.
3. Wykonać polecenie **npm install**, które zainstaluje brakujące pakiety od których zależy rozszerzenie.
4. W sekcji **Debug** wybrać konfigurację **Launch Client** i nacisnąć przycisk zielonej strzałki (ewentualnie nacisnąć klawisz **F5**).
5. Zostanie otwarta druga instancja edytora, w której rozszerzenie jest aktywne.