

Szegedi Tudományegyetem
Informatikai Intézet

Ipari JavaScript elemző kiegészítése TypeScript támogatással

Szakdolgozat

Készítette:

Pozsgai Alex

programtervező informatikus BSc

szakos hallgató

Témavezető:

Dr. Antal Gábor

Tudományos munkatárs

Szeged

2023

Feladatkiírás

A jelenlegi projektet SourceMeter JavaScript-nek hívják. A cél az, hogy ez a projekt többet tudjon, mint a piacon a többi kódelemző. Ebből adódóan bővíteni kell TypeScript támogatással.

A hallgató feladata ennek a programnak a fejlesztése úgy, hogy tudjon TypeScript forráskódot, fájlokat és projekteket is kellő pontossággal elemezni. Ezután az eszközcsalád optimalizálása.

Tartalmi összefoglaló

A téma megnevezése:

SourceMeter JavaScript kiegészítése TypeScript támogatással, az eszközcsalád egyes eszközeinek optimalizálása.

A megadott feladat megfogalmazása:

A feladat során el kell érni azt, hogy a SourceMeter for JavaScript elemző eszköz-készlet képes legyen elemezni TypeScript fájlokat és projekteket. Továbbá, a JavaScript elemző optimalizálása.

A megoldási mód:

A megoldás során változtatni kell a nyelvi sémán, amire az eszközök épülnek. A létrejövő változtatásokat minden eszközön elvégezni, tesztekkel bővíteni.

Alkalmazott eszközök, módszerek:

A megoldáshoz Visual Studio Code-t, a nyelvi séma szerkesztéséhez Visual Paradigm-t használtam. A projekt lebuildeléséhez és ellenőrzéséhez Windowson Visual Studio 2017, Linuxon make programot használtam.

Elért eredmények:

A SourceMeter JavaScript képes TypeScript forráskódokat, fájlokat nagy pontossággal elemezni, nagyobb projektekre lényegesen gyorsabban fut le, kevesebb erőforrást igényel. A JSAN2Lim sikeresen alakítja át a JavaScript elemző (JavaScript Analyzer, továbbiakban: JSAN) eredményét nyelvfüggetlen modellre, TypeScript fájlok elemzése során is, amin dolgozik több eszközcsalád.

Kulcsszavak:

JavaScript, TypeScript, Optimalizálás, Visual Paradigm, C++, Statikus kódelemzés, AST

Tartalomjegyzék

Feladatkiírás	1
Tartalmi összefoglaló	2
1. Bevezetés	5
2. SourceMeter	7
2.1. SourceMeter	7
2.2. SourceMeter for JavaScript	8
2.2.1. Nyelvi séma bevezetés	8
2.2.2. JavaScriptAddon	13
2.2.3. JSAN2Lim	17
2.2.4. Regressziós tesztelés	19
3. Oszthatlan közösen készített programok	21
3.1. A nyelvi séma átírása	21
3.2. Nehézségek, problémák	22
4. Általam végzett feladatok az elemzőeszközökön	23
4.1. JavaScriptAddon változások	23
4.2. JSAN2Lim Bővítések	23
4.3. Ast binder optimalizálás	24
4.3.1. Lassúság okai	25
4.3.2. Optimalizálás	26
4.3.3. Eredmények	28

5. Regressziós tesztek frissítése	30
5.1. Tesztek kibővítése	30
6. Összefoglaló	33
6.1. Program javulása	33
6.2. Jövőbeli tervek	33
Nyilatkozat	34
Irodalomjegyzék	35

1. fejezet

Bevezetés

Szakdolgozatomban a SourceMeter for JavaScript továbbfejlesztése volt a cél. A JavaScript elemzése mellett TypeScript nyelvű fájlokat és projekteket is elemezni kellett. Ebből adódik az, hogy gyökerestül át kellett írni az eszközt, annak érdekében, hogy JavaScriptet és TypeScriptet is tudjon egyaránt elemezni.

A SourceMeter for JavaScript projekten többen is fejlesztettek egyszerre, ezáltal voltak átfedések, oszthatatlan részek. A projekt néhány része is egymásra épül, ezért kiemelten fontos volt a csapatmunka egyes részeknél, hogy a projekt eredményesen és hatékonyan haladjon. Az eredményes csapatmunka magában foglalja az eredmények és az előrehaladás folyamatos kommunikálását, ami azt jelenti, hogy mindketten megértjük, hogy milyen célkitűzések vannak a projektben, és rendszeresen jelentést adunk egymásnak az elvégzett munkáról és az elért eredményekről.

Megtalálható számtalan JavaScript és TypeScript fájl elemző eszköz amiknek nyílt a forráskódja. Ezek között megtalálható a Codehawk CLI, Codelyzer vagy a CodeClimmate-Duplication. Mind a három eszköz jól tud elemezni JavaScript és TypeScript fájlokat, viszont csak egy-egy specifikus esetre jók.

Fejlesztés során a csoportos munka elkerülhetetlen volt, hiszen volt olyan rész, amire számtalan eszköz épült. Ez a JavaScript nyelvi séma (továbbiakban:séma) szerkesztése volt. Ezt a 2.2.1 alfejezetben taglalom bővebben.

A séma szerkesztése után a JavaScript Analyzer To Lim (továbbiakban JSAN2Lim) továbbfejlesztése volt a feladatom. A továbbfejlesztés C++, C nyelv, TypeScript nyelvi elemek [8, 6, 11, 5] és a TypeScript séma [16] ismeretét igényelte meg. Emiatt a fejlesztés

előtt tanulmányozni kellett ezeket.

A JSAN2Lim fejlesztése után a JavaScript elemzőnek (JavaScript Analyzer, továbbiakban: JSAN) a binder függvényét kellett optimalizálnom. Először megérteni kellett a kódot, át kellett nézni az Absztrakt Szintaxis Fát (Abstract Syntax Tree, továbbiakban: AST), illetve több adatszerkezetet is, gyorsasági szempontból.

Legvégül bővítettem a regressziós tesztek új projektek behozatalával. Ezt az 5 fejezetben taglalom.

2. fejezet

SourceMeter

2.1. SourceMeter

A SourceMeter egy forráskód-elemző eszköz, amely képes statikus programelemzést végezni a C, C++, Java, Python, C#, JavaScript, TypeScript és RPG (AS/400) [15] nyelvű összetett programok forráskódján. A FrontEndART a Szegedi Tudományegyetem Szoftverfejlesztés Tanszékén kutatott és fejlesztett Columbus technológián [4] alapuló SourceMeter eszközt fejlesztette ki. A statikus kódelemzés egy olyan módszer, amely során a program forráskódját elemezzük, anélkül hogy azt ténylegesen futtatnánk. Az elemzés során különböző eszközök segítségével ellenőrizhetjük a kód helyességét, hatékonyságát, biztonságosságát és karbantarthatóságát. Az ilyen típusú elemzés során gyakran felhasználnak különböző szabályokat és előírásokat, amelyek segítenek az azonosításban és a hibák javításában.

A statikus elemzés során Absztrakt Szemantikus Gráf (ASG) készül a forráskód nyelvi elemeiből. Ezután az ASG-t különböző eszközökkel dolgozzák fel a csomagban annak érdekében, hogy kiszámítsák a metrikákat (LLOC [14, 2], NLE vagy NOA), azonosítsák az ismételt kódrészleteket (másolás-beszúrás; klónok), a kódolási szabályszegéseket, stb. A SourceMeter képes elemzést végezni olyan forráskódon, amely megfelel a Java 8 és korábbi verzióinak, a C/C++, az RPG III és az RPG IV verzióinak (beleértve a szabadon formázottakat), a C# 6.0 és korábbi verzióinak, valamint a Python 2.7.8 és korábbi verzióinak. A C/C++ esetében a SourceMeter támogatja az ISO/IEC 14882:2011 [13] nemzetközi szabványt, amelyet kiegészítettek az ISO/IEC 14882:2014 új funkcióival, és

a C nyelvet az ANSI/ISO 9899:1990, az ISO/IEC 9899:1999 és az ISO/IEC 9899:2011 szabványok határozzák meg. Az alapértelmezett funkciókon túl, a GCC és a Microsoft által meghatározott kiterjesztések is támogatottak.

A SourceMeter a QualityGate [3] eszközben van használva.

A SourceMeternek található egy plug-in a SonarQubehoz. A SourceMeter plug-in a SonarQube platformhoz egy kiterjesztése az nyílt forráskódú SonarQube platformnak, amelyet a kód minőségének kezelésére használnak. A plug-in a SourceMeter-t futtatja a SonarQube platformról, és feltölti a forráskód elemzésének eredményeit a SourceMeter-től a SonarQube adatbázisába. A plug-in nyílt forráskódú, és az összes szokásos SonarQube kódelemzési eredményt biztosítja, kiegészítve sok további metrikával és problémakeresővel, amelyeket a SourceMeter eszköz biztosít. A plug-in támogatja a C/C++, a Java, a C#, a Python és a RPG nyelveket. [10]

2.2. SourceMeter for JavaScript

A SourceMeter for JavaScript a SourceMeternek egy nagyobb alprojektje. A SourceMeter for JavaScript egy olyan eszköz, amely lehetővé teszi a statikus forráskód elemzést a bonyolult JavaScript és TypeScript rendszerekben. Képes felismerni a kód hibáit, mint például a nem definiált változók vagy függvények használata, a nem biztonságos kódrészletek, a nem hatékony kódrészletek, valamint a redundáns és ismétlődő kódok. Ezenkívül az eszköz képes összehasonlítani a kódot az általános gyakorlatokkal és a meghatározott szabályokkal, és jelezni az eltéréseket. Az ilyen eszközök használata segíthet az észlelt hibák javításában és a kód minőségének javításában, ami végső soron javíthatja a rendszer biztonságát és hatékonyságát.

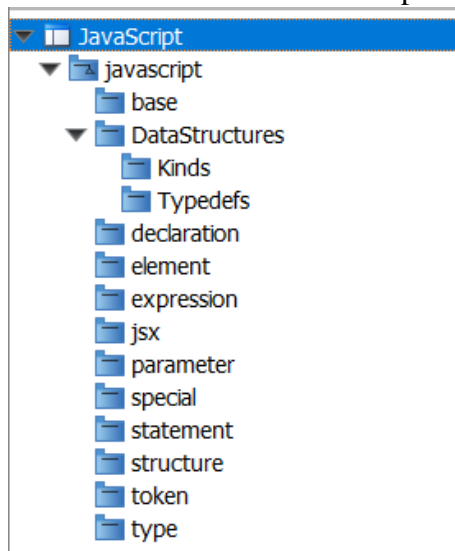
A SourceMeter for JavaScript projekt több alprojektet is magába foglal, amelyek különböző részfeladatokra specializálódnak.

2.2.1. Nyelvi séma bevezetés

A JavaScript nyelvi séma(továbbiakban: séma) [9] egy UML Diagramhoz hasonló séma. A Visual Paradigm(továbbiakban: vpp) alkalmazással szerkeszthető. Azért szerkesztjük a sémát vpp-ben, mivel a többi nyelvi változat esetében is így volt, és van konver-

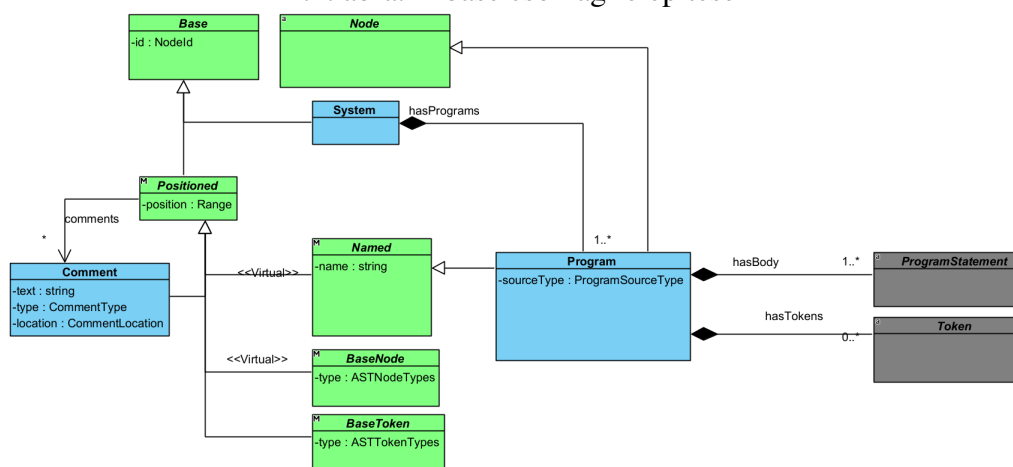
ter, ami ebből a vpp fájlból létrehozza a C++ alap fájlokat, amiket az elemző programok tudnak használni. A felépítése a következőképpen néz ki:

2.1. ábra. Séma strukturális felépítése



A 2.1 ábrán a ProjektNév/Model/Csomagok struktúra figyelhető meg. A projekt neve JavaScript, ezen belül található egy model, amit javascript-nek hívnak. A modellen belül találhatóak meg a csomagok. A csomagok a TypeScript séma [16] alapján lettek elnevezve. Átláthatóság és könnyebb bővíthetőség szempontjából lett létrehozva több csomag. Követtük a TypeScript-eslint [16] projekt strukturális felépítését. A csomagok a következőképpen épülnek fel:

2.2. ábra. A base csomag felépítése



A 2.2 ábrán a Base, Node, Positioned, Comment, System, Named, BaseNode, BaseToken, Program, ProgramStatement és a Token osztályok találhatóak. A Base osztály-

ból öröklődik minden osztály. A Base osztály rendelkezik id attribútummal, ami NodeId típusú.

Az absztrakciót zöld háttérszínnel jelöltük. A séma könnyebb bővíthetősége és olvashatósága miatt jeleztük így, illetve más nyelvi sémákban is ezt a logikát követjük. Más csomagokban lévő definiálást szürke háttérszínnel, ugyanabban a csomagban lévő definiálást világosszürke háttérszínnel jelöltük. Az alapértelmezett osztályt kék háttérszínnel jelöltük. A Positioned osztályból öröklődnek a Named, BaseNode és a BaseToken osztályok. A BaseNode osztályból több minden öröklődik, mint például a DeclarationStatement, ImportDeclaration, Expression és több minden is. Ez látható a 2.3 ábrán. A Positioned osztály rendelkezik position attribútummal, ami Range típusú. A Positioned osztályhoz tartozhatnak kommentek is. A Comment osztály öröklődik a Positioned osztályból, ezzel biztosítva azt, hogy minden kommentnek lesz pozíciója. A Comment osztály rendelkezik text, type és location attribútummal. A CommentType és a CommentLocation a DataStructures csomagban lett definiálva. Végül, Program osztály öröklődik a Named osztályból, ezáltal rendelkezik name attribútummal, ami string típusú. Minden Program osztályhoz tartozik legalább 1 System, ezt az 1..*-al jelöltük. A hasPrograms-nak a JavaScriptAddon-ban van jelentősége. A ProgramSourceType is a DataStructures csomagban lett definiálva. A ProgramSourceType értéke lehet source, vagy module. A Base csomag eltér a TypeScript-eslint [16] projektben lévő Base csomagtól. Régebbi sémának a Base csomagét használjuk, BaseNode és BaseToken osztályok kibővítésével.

Mivel a Base csomag egyedi a mi esetünkben, ezért kitérnék a Declaration csomagra is. A Declaration csomagon belül az ExportAllDeclaration osztályt mutatom be.

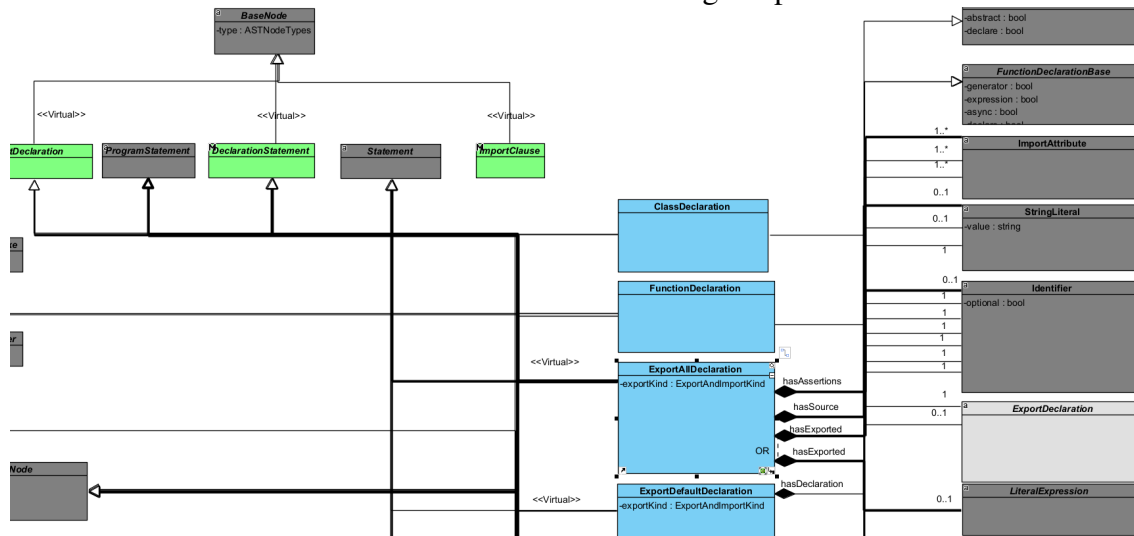
```
1 export interface ExportAllDeclaration extends BaseNode {  
2     type: AST_NODE_TYPES.ExportAllDeclaration;  
3     assertions: ImportAttribute[];  
4     exported: Identifier | null;  
5     exportKind: ExportKind;  
6     source: StringLiteral;  
7 }
```

Kódrészlet 2.1. ExportAllDeclaration TypeScript megvalósítása

A 2.1 kódrészlet megvalósítása a sémában:

Az ExportAllDeclaration osztály öröklődik a DeclarationStatement osztályból, ami

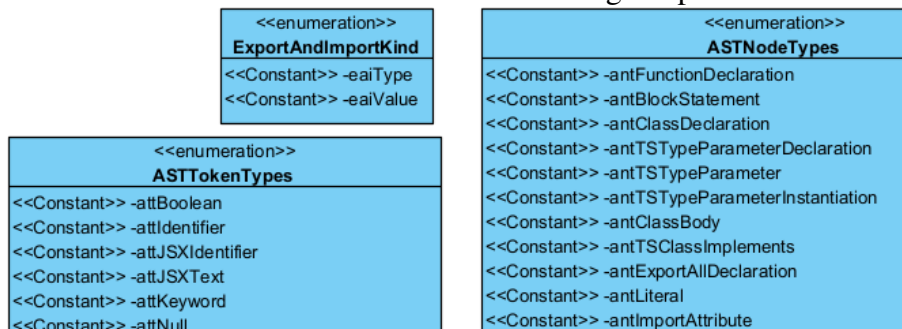
2.3. ábra. A Declaration csomag felépítése



öröklődik a `BaseNode` osztályból. Ezért az `ExportAllDeclaration` egyaránt öröklődik a `BaseNode` osztályból. A 2.1 kódrészletben ez az `extends BaseNode`-al van jelezve. A főbb osztályok a `BaseNode` alatt találhatók az ábrán. Az ábra jobb oldalán lévő osztályok, amik sötét- és világosszürkével vannak jelölve, az attribútumok. Az `ExportAllDeclaration` osztály öröklődik a `Statement`, `DeclarationStatement`, `Node` és a `ProgramStatement` osztályokból. Az összes öröklődés a TypeScript-eslint [16] projekt unions mappájában találhatóak meg. Az `ExportAllDeclaration` osztály rendelkezik `pozíció`, `Comment`, `type` és `NodeId` attribútummal, mivel öröklődik a `BaseNode` osztályból. Emellett még rendelkezik `Assertions`, `Source` és `Exported` attribútummal is. Az `Assertions` attribútum típusa `ImportAttribute`, amihez legalább egy `ImportAttribute` tartozik. Az `Exported` attribútumnál típusa lehet vagy `Identifier` vagy `LiteralExpression`. A vagyolást egy OR-ral jeleztük a sémában. Az `Exported` attribútum opcionális, ezért az értéke lehet null is. Ezt a sémában a 0..1-el jelöltük. A `Source` attribútum is opcionális, típusa `StringLiteral`. Végül, `ExportKind` attribútummal is rendelkezik, típusa `ExportAndImportKind`. Az `ExportAndImportKind` az a `DataStructures` csomagban lett definiálva. Így lett minden csomag és osztály felépítve. A `DeclarationStatement`, `Statement`, `ProgramStatement` és több osztály, ami a unions mappában található, gyűjtő osztály, aminek a jelentősége a JavaScriptAddon-nál fog mutatkozni.

A `DataStructures` csomag többször volt említve, hadd mutassam be a következő ábrán a felépítését:

2.4. ábra. A DataStructures csomag felépítése



A DataStructures csomagban található 2 csomag, Kinds és Typedefs. A 2.1 ábrán látható. Ebből a Kinds csomagot mutatom be. A 2.4 ábrán a Kinds csomagnak egy kisebb része látható. Enumok vannak deklarálva ebben a csomagban. Az enumokban konstansok találhatóak. Minden konstans előtt található 3 karakter, ezek a karakterek ASG konvertáláskor el fognak tűnni. Ha ExportAndImportKind-ot adunk meg egy attribútum típusának, akkor az attribútum típusa lehet Type vagy Value.

A séma használata előtt először ki kell exportálni a projektet. A vpp segítségével tehető meg, egy xml formátumú fájlba exportálódik egész projekt. Ezután ezt a fájlt át kell konvertálni ASG kiterjesztésű fájlra. UmlToAsg eszközzel tehető ez meg, ami a SourceMeter alprojektjei közé tartozik. Ez a projekt egy xml fájlból egy ASG fájlt generál. Az UmlToAsg projekt a SchemaGenerator által generálódik. A SchemaGenerator C++ nyelven megírt eszköz, a SourceMeter alprojektjei közé tartozik. Több mindent is generál C, C++, vagy Java nyelven. A generált ASG fájl elején a következő sorok láthatóak:

```
1 NAME = javascript;
2 APIVERSION = 0.3.1;
3 BINARYVERSION = 0.3.1;
4 CSIHEADERTEXT = JavaScriptLanguage;
```

Kódrészlet 2.2. ASG fájl első sorai

A verziókat kézzel tudjuk átírni abban a fájlban ami generálja ezt, ez egy C++ fájl, a SchemaGeneratorban található meg. Ezután a Kinds csomag tartalmát írja bele a következőképpen:

```
1 KIND ASTNodeTypes (ant) {
2     FunctionDeclaration;
3     BlockStatement;
```

```
4      ClassDeclaration; }
```

Kódrészlet 2.3. A KIND típus átalakítása

A 3 karakter, amit minden konstans előtt szerepel, azok kikerülnek paraméterbe és csak az utáni stringet írta át az ASG fájlba. Ugyanígy jár el a többi enumnál is. Ha az összes enumot beleírta, akkor a többi csomagot kezdi el bele írni a fájlba. A 2.1 alapján halad sorban. Példának a Declaration csomagot mutatom be, azon belül is az ExportAllDeclaration osztályt.

```
1 SCOPE declaration {
2     NODE DeclarationStatement : virtual base::BaseNode [ABSTRACT] {
3     }
4     NODE ExportAllDeclaration : DeclarationStatement, statement::
        Statement, virtual statement::ProgramStatement, special::Node
        {
5         ATTR ExportAndImportKind exportKind;
6         EDGE TREE 1 hasExported (expression::Identifier |
            expression::LiteralExpression);
7         EDGE TREE 1 hasSource (structure::StringLiteral);
8         EDGE TREE * hasAssertions (special::ImportAttribute);
9     }
10 }
```

Kódrészlet 2.4. ASG fájl ExportAllDeclaration

Az UmlToAsg minden egyes csomagot Scope-nak értelmez. A Scope-on belül az osztályokat gráf csomópontnak (továbbiakban: node) értelmezi. A node-nak kettő különféle attribútuma lehet, ATTR és EDGE TREE. Akkor konvertálja ATTR-ra az attribútumot, ha az a DataStructures csomagban deklarálva lett. Egyéb esetben EDGE TREE-vé konvertál. Az EDGE TREE után álló szám vagy csillag, az adott attribútum multiplicitását jelenti. A vagyolást egy | jellel jelzi. Az öröklődést kettősponttal jelzi. Felsorolásszerűen írta át, ha valami másból származik, akkor csomagNev::osztalyNev alapján. A 2.3as ábrán látható, hogy mit hogyan írt át.

2.2.2. JavaScriptAddon

A JavaScriptAddon a SchemaGenerator által generált fájl. A 2.4 kódrészleten lát-

ható, hogy node-okat generált az UmlToAsg az osztályokból. Ezeknek a node-oknak a SchemaGenerator generál külön egy WrapperHeader és egy WrapperCC fájlt. A SchemaGenerator továbbá generál Factory.cc és a hozzá tartozó header fájlt. A Factory.cc-t bővebben taglalom a 2.11 kódrészletnél. A WrapperCC, Factory.cc és a hozzájuk tartozó header fájlok összeolvasztása eredményezi a JavaScriptAddon fájlt. A JavaScriptAddon fájl generálás menetét a NodeAddonGenerator fájlban található. A SchemaGenerator main.c fájlba beimportálja a NodeAddonGenerator összes metódusát, aztán ezeket a metódusokat használva legenerálja a JavaScriptAddon fájlt. A `-genNodeAddon` kapcsolóval fogja legenerálni a JavaScriptAddon fájlt a SchemaGenerator. A SchemaGenerator main.c fájl elején először vizsgál a megadott kapcsolókra az alábbi módon:

```
1 if(!strcmp(argv[i], "-genNodeAddon")) {
2     options.generateNodeAddon = true;
3 }
```

Kódrészlet 2.5. SchemaGenerator kapcsoló vizsgálás

Ha az argumentumban megtalálható a `-genNodeAddon` kapcsoló, akkor az `options.generateNodeAddon`-t igazra állítja. Az alapértelmezett értéke az `options.generateNodeAddon`-nak hamis. Több fájl generálás után, ami kell több program működéséhez, megvizsgálja az `options.generateNodeAddon`-t. Ha az `options.generateNodeAddon` hamis, akkor nem fog semmi se történni.

```
1 if (createAndEnterDirectory(SOURCE_NODE_ADDON_DIR_NAME)) {
2     generatePackageJson();
3     generateBindingGyp();
4     generateAddonCC();
5     generateFactoryWrapper();
6     if (createAndEnterDirectory("inc")) {
7         generateWrapperHeaders();
8         leaveDirectory();
9     }
10    if (createAndEnterDirectory("src")) {
11        generateWrapperSources();
12        leaveDirectory();
13    }
14    leaveDirectory();
15 }
```

Kódrészlet 2.6. SchemaGenerator JavaScriptAddon generálás

A `createAndEnterDirectory` metódus először létrehoz egy `addon` nevű mappát és `chdir` parancsot használva belelép. Ha ez a mappa már létezik, akkor a létrehozást kihagyja és a `chdir` paranccsal belelép. Létrehozás és `chdir` parancs használat után az `addon` mappába generál `package.json` fájlt. Ebben a fájlban beállítja a projekt nevét, verziószámát, függőségeket, szkripteket és a gypfájl kapcsolónak igaz értéket beállítja.

```
1 fprintf(f, "    \"rebuild\": \"node-gyp configure && node-gyp rebuild -\n    j 8\",\\n");\n2 fprintf(f, "    \"install\": \"node-gyp configure && node-gyp build -j\n    8\",\\n");
```

Kódrészlet 2.7. NodeAddonGenerator package.json szkriptek

A 2.7 kódrészleten látható, hogy a `node-gyp` build segítségével fog történni a buildelés. A `node-gyp` egy eszköz, amely összeállítja a Node.js Addonokat. A Node.js Addonok natív Node.js modulok, amelyeket C vagy C++ nyelven írnak. Miután az ilyen modulokat eszközökkel, például a `node-gyp` segítségével lefordították, funkcióik elérhetők lesznek a `require()` segítségével, ahogy bármely más Node.js modul esetében.

A `package.json` fájl után generálja le a `binding.gyp` fájlt. A `binding.gyp` fájlban találhatóak meg az összes node fájl relatív útvonala a `package.json` fájlhoz képest. A `binding.gyp` fájl után generálja le az `addonCC` fájlt.

```
,\n1 if (!traversalDescendantBFT(rootNode, generateWrapIncludes, false)) {\n2     debugMessage(0, " failed\\n");\n3     fclose(f);\n4     return false;\n5 }
```

Kódrészlet 2.8. Addon.cc Wrapperek importolása

A `traversalDescendantBFT` metódus egy bejárás, ami az összes node-ra lefut. Ezekre a node-okra meghívja a `generateWrapIncludes` metódust. A `generateWrapIncludes` metódus a `NodeAddonGenerator` fájlban van leimplementálva a következőképp:

```
,\n1 if (node->type.abstract) {\n2     return true;
```



```
3 }
4 fprintf(f, "#include \"");
5 fprintf(f, "inc/%sWrapper.h\\\"\\n", node->name );
6 return true;
```

Kódrészlet 2.9. generateWrapIncludes leimplementálása

Ha a jelenlegi node absztrakt, akkor nem ír semmit a fájlba. Ha nem absztrakt, akkor a 2.9 kódrészlet alapján importálja a node-nak a header fájl relatív útvonalát. A `node->name` több értéket is vehet fel, például `FunctionDeclaration`, `ClassDeclaration`, ezek megtalálhatóak az ASG fájlban. Ezután ez a bejárás még egy alkalommal végrehajtódik. Ebben az esetben a `wrapperInit`-eket ír az `addonCC` fájlba.

```
1 fprintf(f, "    columbus::%s::asg::addon::%sWrapper::Init(env, exports);\\n", schemaName, node->name);
```

Kódrészlet 2.10. generateWrapInit leimplementálása

Az `addonCC` fájl generálása után a `generateFactoryWrapper` függvényhívás történik. Ebben a függvényben 2 metódus függvényhívás található, a `generateFactoryWrapperHeader` és `generateFactoryWrapperSource`.

A `generateFactoryWrapperHeader` a `Factory.h` fájlt generálja le. Először importálja az összes `nodeWrapper` header fájlt, és utána létrehoz egy `Factory` osztályt, a publikus metódusait, ami az `Init` és `Destructor`, illetve a `private` metódusait, ami a `destructor`, `New`, `SaveAST`, `LoadAST`, `Clear`, `getRoot` és az összes `nodeWrapper` `create` metódusai.

A `generateFactoryWrapperSource` a `Factory.cc` fájlt generálja le. `DECLARE_NAPI_METHOD`-okat hoz létre, az összes `nodeWrapper` fájlnek.

```
1 napi_property_descriptor props [] = {
2     DECLARE_NAPI_METHOD("getRoot", getRoot),
3     DECLARE_NAPI_METHOD("createCommentWrapper", createCommentWrapper)
4     ...}
```

Kódrészlet 2.11. Factory.cc fájl

Legvégül, a `SchemaGenerator` legenerálja az `inc` és az `src` mappákat. Az `inc` mappában találhatóak a header fájlok minden egyes `nodeWrapper`-nek, az `src` mappában maguk a `nodeWrapper`-ek vannak megvalósítva.

A SchemaGenerator legenerálja az ExportAllDeclarationWrapper.h és az ExportAllDeclaration.cc fájlokat. A header fájlban létrehozza az ExportAllDeclarationWrapper osztályt, ami öröklődik a BaseWrapperből. A BaseWrapper alap Wrapper osztály. Három publikos metódust generál le, az Init-et, a Destructor-t és a NewInstance-t. Ezen felül több private metódust is.

```
1 napi_property_descriptor props [] = {  
2     static napi_value setPosition(...);  
3     static napi_value addAssertions(...);  
4     static napi_value setType(...);  
5 ...}
```

Kódrészlet 2.12. ExportAllDeclarationWrapper.h fájl

A 2.12 kódrészleten megfigyelhető, hogy létrehozott az ExportAllDeclarationWrapper-nek egy setPosition, addAssertions, setType metódusokat. Ezek az ExportAllDeclaration attribútumai, ami a sémában be lett állítva. A 2.3 ábrán látható, hogy az ExportAllDeclaration hasAssertions attribútumának a multiplicitása 1.*, ezt átkonvertálta addAssertions-re. Ahol 0.* a multiplicitás, azt átkonvertálta setAttribute-ra, például setSource vagy setExported-re. Minden nodeWrapper rendelkezik setPath, setPosition, setType és addComments metódusokkal.

Az ExportAllDeclaration.cc fájlt ugyanúgy generálja le, mint a Factory.cc fájlt. DECLARE_NAPI_METHOD-okat generál, a setExported, setSource, setType és a többi metódushoz. Ezután az összes metódust megvalósításra kerül.

2.2.3. JSAN2Lim

A JSAN2Lim C++ programozási nyelvben megírt program. A JSAN általi eredményt alakítja át nyelvfüggetlen (továbbiakban: LIM) formátumú fájlokra. Azért LIM formátumú fájlokra alakít át, mivel egy nyelvfüggetlen modellen sokkal könnyebb metrikákat mérni. A JSAN2Lim függ a JSAN programtól, hiszen a JSAN eredményeit alakítja át. Ha a JSAN programot változás éri, akkor a JSAN2Lim programot is változtatni kell, különben nem lesznek pontosak a metrika mérések. A JSAN2Lim használja a LimSchema-t. A LimSchema hasonlóan van felépítve, mint a JavaScript nyelvi sémája. A JSAN2Lim bejárja az összes node-ot ami megtalálható a JSAN általi eredmény fájlban. Minden node-ra

külön meg kell írni a bejáró (továbbiakban: visit) függvényt.

A JSAN2Lim működését egy kódrészleten keresztül mutatom be.

```
1 VISIT_BEGIN(clNode, callVirtualBase);
2 lim::asg::logical::Class& classLimNode =
3 dynamic_cast<lim::asg::logical::Class&>(createLimNode(clNode,
4     callVirtualBase));
5 fillData(classLimNode, clNode);
6 fillMemberData(classLimNode);
7 demangledNameParts.push_back(classLimNode.getName());
8 classStack.push(ClassInfo());
9 classStack.top().classNodeId = classLimNode.getId();
10 if (clNode.getSuperClass() != NULL) {
11     if (javascript::asg::Common::getIsIdentifier(*clNode.
12         getSuperClass())) {
13         javascript::asg::expression::Identifier& superClass =
14         dynamic_cast<javascript::asg::expression::Identifier&>(*
15             clNode.getSuperClass());
16         classStack.top().isSubclass = superClass.getId();
17     }
18 }
19 if (!packageStack.empty()) {
20     lim::asg::logical::Package& packageLimNode =
21     dynamic_cast<lim::asg::logical::Package&>(limFactory.getRef(
22         packageStack.top().packageNodeId));
23     SAFE_EDGE(packageLimNode, Member, lim::asg::logical::Member,
24         classLimNode.getId());
25 }
26 addIsContainedInEdge(classLimNode, clNode);
```

Kódrészlet 2.13. ClassDeclaration Visitor

A VISIT_BEGIN a VisitorAbstractNodes programnak a visit metódusa. Ezután létrejön a classLimNode változó, ami lim::asg::logical::Class& típusú. Ez a típus a LimSchema-ban található meg, a logical csomagban. A változónak a createLimNode metódus által visszaadott node lesz az értéke. Ezt az értéket konvertáljuk át a erre a lim::asg::logical::Class& típusra.

A createLimNode függvény meghívja a createNode függvényt. Ez a függvény vár

paraméterbe egy `nodeKind`-ot. A `nodeKind` a `node` típusát jelöli. A `nodeKind`-ot a `getLimKind` függvény határozza meg. A `getLimKind` függvény megvizsgálja, hogy a `node` `getIsClassDeclarationBase` vagy `getIsMethodDefinition` vagy `getIsFunctionDeclarationBase`. A `nodeKind` értéke típustól függően adódik meg. Ezek az értékek a következők lehetnek: `ndkClass`, `ndkMethod`, `ndkParameter`, `ndkMethodCall`, `ndkAttribute`, `ndkComment`, `ndkPackage`, `ndkAttributeAccess`, `ndkComponent`. Ezek az értékek mind a `LimSchema`-ban vannak deklarálva. A `nodeKind` értékadása után a függvény visszaadja ezt a változót és tovább használja a `visitor`-ban. Ezután, a `createLimNode` függvény létrehozza a megfelelő `node`-ot a `nodeKind`-tól függően. Beállítja a pozícióját a `node`-nak és visszaadja a `*limNode`-ot.

Ebben az esetben a `nodeKind` az `ndkClass` értéket kapja meg. A `classLimNode` értékadás után meghívódik a `fillData` függvény. Kettő paramétert vár, `classLimNode`-ot, és a `node`-ot. A `fillData` a JSAN általi adott eredményt bejárja, `visitor`ok segítségével. A `fillData`-ban beállítódik a `limnode` láthatósága, nyelve, az hogy absztrakt-e és a `classKind`-ja, ami jelen esetben `clkClass` értéket fog felvenni.

Ezután beállítódik a `limnode` neve. Ha a `node`-nak volt `identifier`-e, akkor a `getName` függvénnyel beállítódik a `limnode` neve különben `anonymousClass` érték állítódik be. Ezután megvizsgálja, hogy a `methodstack` (ami egy `stack` és metódusokat tartalma) üres-e. Ha nem üres, akkor létezik szülője, ezért a `SAFE_EDGE` metódussal hozzáadódik egy csúcs a szülő és a jelenlegi `limnode` között. Ezt ugyanúgy megvizsgálja a `packageStack`-re (ami egy `stack` és csomagokat tartalmaz), ha nem üres, akkor ugyanezt elvégzi csak a szülő más lesz. Legvégül az `addIsContainedInEdge` függvény meghívódik a `limnode` és a `jsNode` paraméterekkel.

Ezután a `fillMemberData` függvény hívódik meg. Ez a függvény a `limnode` pozícióját állítja be, és a nevéen változtat. Utána a `classStack`-hez hozzáadódik az adott osztály tulajdonságai, beállítja a szülő osztályt ha van, a pozícióját is beállítja, és a végén meghívódik az `addIsContainedInEdge` a `classLimNode` és a `clNode` paraméterekkel. Minden egyes `node`-nál ez az eljárás fut le.

2.2.4. Regressziós tesztelés

A regressziós tesztelés általában egy olyan szoftvertesztelési típus, amely akkor hasz-

nálatos, amikor megerősítjük, hogy a legfrissebb program- vagy kódmódosítások nem érintették negatív módon a meglévő funkciókat. A regressziós tesztelés során megállapítják, hogy a szoftver vagy alkalmazás megfelelően működik-e az új változások és hibajavítások tekintetében. A regressziós tesztelés főként arról szól, hogy az már végrehajtott teszteseteket újra lefuttatjuk annak érdekében, hogy megerősítsük, az alkalmazás megfelelően működik-e. A regressziós tesztelésnek számos különböző típusa van, amelyből csak a mindent újratestelni (Re-test All) módszert használ a SourceMeter for JavaScript. Minden nagyobb módosításnál újra megtörténik a mindent újratestelés.

A Re-test All (Mindent újratestelni) módszer egy módja a regressziós tesztelésnek az, hogy az összes tesztesetet újra futtatjuk. Ez valójában egy költséges folyamat, mivel több időt, erőforrást és erőfeszítést igényel az egész teszteset újbóli végrehajtása. Az ebben a folyamatban észlelt hibákat előre jelentik a javítás érdekében.

3. fejezet

Oszthatlan közösen készített programok

3.1. A nyelvi séma átírása

Az eddigi JavaScript nyelvi séma (továbbiakban: séma), ami csak JavaScript nyelvű projekteket és fájlokat elemzett, a JavaScript hivatalos oldala [1] alapján készült. A 2.2.1 fejezetben ami ábrák láthatóak, azok már az átírt sémából lettek kifelé fordítva.

A séma átírásánál arra a döntésre jutottunk, hogy előlről kezdjük az egészet. Átláthatatlan volt az előző séma, és nem volt hozzá megfelelő dokumentáció, ezért a bővítés nehezebbnek bizonyult, mint az újírás a kezdetekről. Továbbá, egy séma ami képes TypeScript fájlokat és projekteket elemezni, az képes JavaScript fájlokat és projekteket is. A Base csomagot emeltük át az előző sémából, BaseNode-al és BaseToken-el kibővítve. Ez látható a 2.2 ábrán is. A TypeScript-eslint github [16] alapján hoztuk létre a struktúrális felépítést, annyi változtatással, hogy mi hoztuk létre még egy structure nevű csomagot, amiben azok az osztályok találhatóak meg, amelyek a TypeScript-eslint github base mappájában voltak. Fejlesztés során dokumentáltuk a lépések nagy részét.

Minden egymásra épült a sémában, emiatt a működést nem tudtuk az elején letesztelni, csak miután több csomag is készen lett. Megfigyelhető a 2.4 kódrészleten, hogy az ExportAllDeclaration több mindenből öröklődik és sok attribútuma van. Az ExportAllDeclaration egyik attribútumának a típusa az ImportAttribute. Ehhez az ImportAttribute osztályt is le kellett fejleszteni a sémában.

```
1 export interface ImportAttribute extends BaseNode {  
2     type: AST_NODE_TYPES.ImportAttribute;
```

```
3     key: Identifier | Literal;  
4     value: Literal;  
5 }
```

Kódrészlet 3.1. ImportAttribute

Az ImportAttributet lefejlesztéséhez le kell fejleszteni az Identifiert, és a Literalt is. Ezek lefejlesztéséhez is kellett más-más osztályokat lefejleszteni.

3.2. Nehézségek, problémák

A séma újból írása során több nehézségbe ütköztünk. Dokumentáció hiánya miatt nem tudtuk az előző sémának minden részét jól értelmezni, néhány félreértés volt fejlesztés közben. Fejlesztés végén, miután az összes csomag és osztály le volt fejlesztve, próbáltuk tesztelni. Tesztelés során Segmentation Fault hibát kapott a program, és nem tudtunk rájönni, hogy miért történik ez. Emiatt az egész sémát átnéztük, biztosra mentünk, hogy az osztályok öröklődési jól voltak a sémában lefejlesztve, illetve az összes attribútumot is ellenőriztük. Több kisebb figyelmetlenséget is észrevettünk, és javítottuk. Több javítás után, újra próbáltunk tesztelni. Ebben az esetben egy másik, jobban beazonosítható hibát kaptunk futtatás során. Már nem kaptunk Segmentation Fault hibát, hanem futás közben egy node-nál minden hibaüzenet nélkül leállt a program. Kiderítettük, hogy melyik node-nál áll le a program mindig, ez volt maga a Literal node. Ismét átnéztük a LiteralExpression osztályt, mivel ez maga a Literal, és egy rossz osztályból öröklődött le, ez okozta a program leállítását. Javítottuk a problémát, és ezek után hibamentesen és jól letudtuk futtatni.

Továbbá sok nyelvi elemet a TypeScript-ben nem ismertünk. Tanulmányozni kellett jobban a TypeScript programozási nyelvet [8, 6, 11, 5], hogy tesztek átnézése során tudjuk mit elemzett le jól és mit nem.

4. fejezet

Általam végzett feladatok az elemzőeszközökön

4.1. JavaScriptAddon változások

A séma változtatás után, magát a NodeAddonGenerator fájlt nem módosítottam. Ennek ellenére a JavaScriptAddon mérete lényegesen megnőtt, mivel a séma nagy bővítésen esett át. A JavaScript-es nyelvi elemek mellett a TypeScript-es nyelvi elemek is megtalálhatóak benne. Emiatt sokkal több node, és hozzájuk a wrapper-ek generálódtak le. A JavaScript-es nyelvi elemek megmaradtak, néhány helyen még javítottunk hibákat is.

4.2. JSAN2Lim Bővítések

A JSAN már TypeScript fájlokat és projekteket is tud elemezni, emiatt a JSAN2Lim-et is fejleszteni kellett. A JSAN2Lim használja a JavaScript és a Lim ASG fájlokat egyaránt.

Több helyen is át kellett írni a JSAN2Lim-et:

- Több visitor-nál a várt paraméter típusát átírni, például `Class`-ról `ClassDeclarationBase`-re, mivel a sémában már nem `Class`-ként, hanem `ClassDeclarationBase`-ként szerepel.
- Több visitor-nál a várt paraméternél helyét meg kellett változtatni egyes node-nál. Például a `RestElement` eddig a `Statement` csomagon belül helyezkedett el,

de a bővítés során átkerült a `Parameter` csomagba.

- A `Pattern`, mint `node`, `Parameter`-re lett átnevezve. A helye is megváltozott, `Statement` csomagból átkerült a `Parameter` csomagba.
- Típus lekérdezésnél a `getIsClass` helyett `getIsClassDeclarationBase` kellett használni.

A bővítések, hogy mikkel bővítettem a `JSAN2Lim`-et:

- Több `kind`-ot is létrehoztunk a sémában, ezekkel bővíteni kellett a `JSAN2Lim`-et.
- Új `node visitor`-ok írása. Mint például a `TSEnumDeclaration`, `TSImportEqualsDeclaration`, `TSInterfaceDeclaration`, és még `TypeScript node` amihez kell `visitor`. A `visitor`-hoz a `fillData` függvényeket is megírni.
- `JSAN2Lim`-ben a `kindStrings` tömb kiegészítése `TypeScript`-es `node`-okkal.
- Hibák kijavítása, a `VariableDeclaration`-nél nem detektálta az összes változó deklarálást, főleg metódusokon belül.
- A `getLimKind` metódus bővítése `getIsTSTypeAliasDeclaration`, `getIsTSInterfaceDeclaration`, `getIsTSAbstractMethodDefinition` és `getIsTSAbstractPropertyDefinition` esetekkel.
- `TSEmptyBodyFunctionExpression` osztály hibajavítása, hiszen több helyen is leállt a program emiatt.

4.3. Ast binder optimalizálás

Az AST binder referencia kötésekre (továbbiakban: bindolásra) alkalmas. Ezek a referenciák a `JSAN2Lim` programhoz szükségesek. Binder néven van definiálva a metódus a `JSAN`-ban. Kétszer van használva, ezért is kulcsfontosságú, hogy gyors legyen. Egyszer `VariableUsages` referenciákat bindol és egyszer `ACG` [7] referenciákat bindol.

A binder 4 argumentumot vár:

- Egy `stringet`, ami lehet vagy `VariableUsages` (továbbiakban: `VU`) vagy `ACG` [7].

- Absztrakt Szintaxis Fát (Abstract Syntax Tree, továbbiakban: AST), ami egyedien van felépítve.
- Egy tömböt, amiben JSON objektumok találhatóak, amik a linkeket tartalmazzák.
- Végül még egy stringet, ami lehet addCalls, vagy setRefersTo. Az AddCalls és a SetRefersTo a JavaScriptAddon-ban található meg és onnan hívódik meg. Abban az esetben kap addCalls értéket, ha ACG referenciákat bindol a binder, és akkor setRefersTo, ha VU referenciákat.

A linkeket tartalmazó tömb egy JSON objektuma a következőképpen néz ki:

```
1 source: {  
2     label: IdentifierNeve,  
3     file: AbsPath,  
4     start: { row: <int>, column: <int>},  
5     end: { row: <int>, column: <int>},  
6     range: { start: <int>, end: <int>},  
7     node: [Object]  
8 },  
9 target: {...}
```

Kódrészlet 4.1. Binder JSON objektuma

Ilyen JSON elemekből épül fel a tömb. A forrás és a cél objektum felépítése ugyanaz, csak másak az értékek. A kódrészletben a cél (source) van kifejtve bővebben. A `label` az egy Identifier vagy PrivateIdentifier node nevét fogja jelölni. A `file` egy abszolút útvonalat kap, ami az Identifiert tartalmazó fájlt jelöli. A `start` az egy objektum, amiben külön van sor és oszlop meghatározva. Ez az Identifier első karakter pozíciójának a sor és oszlop értékei. Az `end` ugyanaz, mint a `start`, csak itt az utolsó karakter pozíciójának a sor és oszlop értékei. A `range` is egy objektum, a `start` ennél a kezdő karakter hanyadik karakter volt a kódban, és az `end` pedig az Identifier utolsó karakterének a karakterszáma. A `node` is egy objektum ami maga az Identifier vagy PrivateIdentifier.

4.3.1. Lassúság okai

Binder-nek a futásideje lényegesen megnő, ha nagyon nagy projektekre futtatódik le. Ennek több oka is van:

- Minden node hívásnál meghívja a JavaScriptAddon-ból egy metódust. A JavaScriptAddon már magában nagy terjedelmű fájl. TypeScript támogatás után kétszeresére nőtt a mérete, mint ezelőtt, emiatt lassult is.
- Nagyobb projektekben több függvényhívás és változószám található az elemzett kódokban.
- Az AST nagyobb projektekénél nagyon nagy is lehet. Ezt az AST-t bejárjuk többször is bindolás alatt. Emiatt lesz nagyon lassú a binder.
- A JSON objektumokat tartalmazó tömb is nagy méretű lesz, de ez kevésbé lassítja a bindert, mint az AST bejárás.

4.3.2. Optimalizálás

Optimalizálás során először a JavaScript Profiler-t használtam memória és futásidő vizsgálatra. Nekem ez nem felelt meg, mivel a JavaScriptAddon-ban rengeteg minden egymásra épül, és 50100 mélységű függvényhívásnál nem tudta kiírni a futásidejét és a memóriahasználatot. Ezután a binder metódust 3 részre bontottam `console.time` és `console.timeEnd` beépített függvények segítségével. Futásidőt írta ki a függvény a `console.time` és a `console.timeEnd` sorok között milliszekundumban. Kiderült, hogy a következő sor felettébb nagy futási idővel rendelkezik:

```
1 getWrapperOfNode (
2     resolveNode (
3         astSet,
4         sourceFile,
5         element.source.range.start,
6         element.source.range.end,
7         true
8     )
9 );
```

Kódrészlet 4.2. Lassú metódus

A `resolveNode` metódus első paraméterként vár egy AST-t, utána egy fájlnevet, kezdő- és végPozíciót, illetve egy igaz vagy hamis értéket.

Az AST-t `forEach`-cel bejárja a program. Az AST egyes elemei az AST node-ok. A `forEach`-csen belül az AST node-on `walk` [12] metódus segítségével bejártuk, addig amíg nem találtuk meg a keresett node-ot. Rosszabb esetben az AST legvégén volt a keresett node, mivel már a végén voltunk a bindolásnak. Kisebb projektek esetén ez nem baj, mivel az AST mérete nem akkora, mint egy nagyobb projekt esetében.

Optimalizálás során, létrehoztam `indexAST` metódust, ami indexeli az AST-t.

```
1 let indexAST = function (ast) {
2     ast.forEach(astNode =>{
3         globals.setActualFile(astNode.filename)
4         walk(astNode, {
5             enter: function(node) {
6                 globals.setIndexed(astNode.filename, node.range[0],
7                                     node.range[1], node)}}))
8     return globals.indexedAST}
```

Kódrészlet 4.3. `indexAST` metódus

A 4.3 kódrészletben látható, hogy egy AST-t várunk paraméterben. Ezt a bindertől fogja kapni. A metódusban `forEach`-cel bejárjuk az AST elemeit, amik az AST node-ok. Beállítom a `globals.setActualFile()` függvénnyel az aktuális fájlnévet. Ezután az adott AST node-ot `walk` függvény segítségével bejárjuk. Csak az `enter` metódusát kellett szerkeszteni. Ezután a `setIndexed` metódus meghívódik. A `setIndexed` függvénynek megkapja a fájlnévet, a node `range`-nek a kezdő- és a végparaméterét, ahol kezdődik az adott node és hol végződik, karakterpontosan, és magát a node-ot.

```
1 const setIndexed = function(filename, range_start, range_end, node){
2     let actualfilename = getFilePathAlt(filename)
3     if (indexedAST[actualfilename + "-" + range_start + "-" +
4         range_end] !== undefined && indexedAST[actualfilename + "-" +
5         range_start + "-" + range_end] !== node){
6         return
7     }
8     indexedAST[actualfilename + "-" + range_start + "-" + range_end]
9     = node
10 }
```

Kódrészlet 4.4. `setIndexed` metódus

A `setIndexed` először vizsgál arra, hogy az adott node be van-e már indexelve. Ezt úgy teszi meg, hogy az `indexedAST` tömbben keres egy indexre. Ez az index a következőképp néz ki: `FájlNév-KezdőPozíció-VégPozíció`. Továbbá vizsgál arra is, hogy ha van ilyen indexű elem a tömbben, akkor ezen az indexen található-e már ilyen node. Ha van akkor nem állít be semmit, csak kilép, mivel már be van indexelve az adott node. Ha nincs, akkor beállítja az `indexedAST` tömbnek az adott indexre az adott node-ot. Az AST-t csak egyszer járjuk be a `foreach`-csel. A bejárás után minden node be lesz indexelve a tömbbe.

A 4.2 kódrészleten látható, hogy először a `resolveNode` függvény segítségével megkerestük a keresett node-ot kezdő- és végPozíció alapján. Ezután a megkapott node-ra meghívtuk a `getWrapperOfNode` függvényt, hogy megkapjuk a wrapperjét a `JavaScriptAddon`-ból. A keresést megváltoztattam, írtam rá egy `getIndexed` metódust.

A `getIndexed` függvény megvizsgálja, hogy forrás vagy cél node-ot keresünk. Ha cél node-ot, akkor visszaadjuk a `getWrapperOfNode(indexedAST[FájlNév-KezdőPozíció-VégPozíció])`-t. Ha cél node-ot keresünk, akkor a `walk` függvény segítségével bejárjuk a beindexelt tömböt. Ez lényegesen gyorsabb, mint a `resolveNode`-os megoldás, mert ott minden esetben az egész AST-t bejártuk a `walk` függvény segítségével, jelen esetben csak a beindexelt tömb elemét járjuk be. Ha megkaptuk a cél node-ot akkor visszaadjuk azt a `getWrapperOfNode(result)` hívással.

4.3.3. Eredmények

4.1. táblázat. Eredmények összehasonlítása

Optimalizálás előtt (m:ss.mmm)	Optimalizálás után (m:ss.mmm)
42:31.211	1:39.140
2:51.045	0:29.922
1:17.825	0:3.949
0:32.5	0:0.614
0:8.2	0:0.120
0:4.830	0:1.423
0:4.8	0:0.153
0:0.4	0:0.2

Egy nagyobb projektre futtattam le a JSAN-t, 2 közepesre és 5 kisebbre. Az eredmények láthatóak a 4.1 táblázatban. Eredménynek ugyanazt adta mind a kettő futás, viszont futásidőben nagyon eltértek egymástól.

5. fejezet

Regressziós tesztek frissítése

5.1. Tesztek kibővítése

Először kezdtem a JSAN tesztek átírásával. A JSAN-nak a következő kapcsolói voltak:

- -i: Jelentése input, lehet relatív vagy abszolút útvonal a fájlhoz vagy projekthez.
- -o: Jelentése output neve, lehet relatív vagy abszolút útvonal.
- -d: Jelentése dumpjsml, a JSAN eredményét átgenerálja XML stílusú fájlba és ezt egy jsml fájlba kiírja.
- -e: Jelentése ExternalHardFilter, relatív vagy abszolút útvonal egy olyan fájlhoz, ami szövegalapú és olyan szintaxis található benne, ami kell az externalHardFilter-nek.
- -help: Jelentése help, kiírja minden kapcsolóhoz tartozó leírást.
- -r: Jelentése useRelativePath, eredményben az útvonalakat átírja relatív útvonalra.
- -h: Jelentése html, a JSAN html fájlokra is lefut, bennük keresve JavaScript-es szkripteket és azokat teszteli.
- -stat: Jelentése statistics, kiírja a memóriahasználatot és a futásidőt amit a JSAN vett igénybe.

Az input, output, dumpjsml, ExternalHardFilter, és html kapcsolókra tudtam tesztelést írni. A logika az volt, hogy mappanév alapján tesztelek egyes kapcsolókra. ProgramozásiNyelv-kapcsolónév logikát követtem, mivel JavaScript-es tesztek mellé kell majd TypeScript-es tesztek is keresnem. A projekteket python szkriptek segítségével futtattam le.

```
1 if "externalHardFilter" in input_path:
2     ret_val = self._execute_one_test(input_path, external_hard_filter
    =True) and ret_val
3     return ret_val
```

Kódrészlet 5.1. JSAN kapcsoló vizsgálat pythonban

Az 5.1 kódrészleten látható, hogy hogyan keresek egy adott kapcsolóra. Az input_path-ban van a mappa is, és a js-externalHardFilter-ben megtalálható az externalHardFilter szó. Az execute_one_test függvényemben beállítom a teszteléshez az adott dolgokat.

```
1 if external_hard_filter:
2     input_dir = os.path.dirname(input_path)
3     external_hard_filter_path = os.path.join(input_dir, "
    externalHardFilter.txt")
4     external_hard_filter_switch = "-e"
5 else:
6     external_hard_filter_path = ""
7     external_hard_filter_switch = ""
```

Kódrészlet 5.2. JSAN kapcsoló beállítása pythonban

Az 5.2 kódrészletben az execute_one_test függvénynek egy részét láthatjuk, ahol beállítjuk az external_hard_filter_path-t és a kapcsolót annak függvényében, hogy igazat kaptunk e vagy sem. Létrehozzuk az externalHardFilter fájlt a tesztelendő fájl mellé vagy a tesztelendő projekt gyökerébe, és attól függően, hogy ki akarjuk hagyni az adott fájlt vagy hozzáadni, írunk egy + vagy egy – jelet a sor elejére és utána relatív útvonal és a fájl neve. Alapértelmezetten minden fájl elemez az adott program, JSAN esetében ezek azok a fájlok amiknek a kiterjesztése js, jsx. (külön kapcsolóval megadhatjuk neki, hogy a html kiterjesztésű fájlokat elemezze-e vagy se.) TypeScript-es bővítés után már a ts és a tsx kiterjesztésű fájlokat is elemzi. Ezért írtam több tesztet is. Egy példa, hogy

hogyan néz ki ez a fájl:

```
1 -filtered01
2 -filtered02
3 -filtered03
4 +filtered01
```

Kódrészlet 5.3. ExternalHardFilter fájl

Az 5.3 kódrészletben látható, hogy filtered01 02 és 03 at kivettük, hogy azokat ne elemezze a jsan. Ezután visszavettük a filtered01et. A filtered fájlok azok JavaScript-es fájlok, JavaScript-es kóddal. Ezután referenciába csak a filtered01 fájl eredményét raktam be, hiszen a 02 és 03at nem elemzi, ha elemezné, akkor szólna a program, hogy hiányzó referencia fájl. Természetesen lehet reguláris kifejezést is használni filterezésnél.

Ezután teszteltem a `-i` kapcsolót, itt a programnak vissza kellene adnia, hogy üres az input ha nincs megadva. Ezután a `-o` kapcsolóra teszteltem, itt ebben az esetben alapértelmezett értékben `out.jssi` fájlt kellene visszaadnia. Végül a `-h` kapcsolót néztem meg, itt ha megvan adva ez a kapcsoló, akkor az adott projektben a html fájlokban a JavaScript-et kellene tesztelni. A `-d`, `-help`, `-stat` kapcsolókra nem tudtam tesztelni, még a `-useRelativePath` kapcsolóra sem, hiszen ha abszolút utat kérek, akkor a referenciákban másnál rossz lesz az elvárt eredmény.

Ezután a JSAN2Lim és az ESLintrunner programoknak írtam át a tesztelési menetét, mivel nekik volt még olyan kapcsoló, amit lehetett tesztelni. A többi projektnek 1 vagy 2 volt, amik kellettek a működésükhöz, nem voltak opcionális kapcsolók.

Miután a JSAN ki lett egészítve TypeScript támogatással és a JSAN2Limet átírtam, elkezdtem teszteket keresni, mind a JSAN-nak és mind JSAN2Lim-nek. Kerestem egyszerűbb TypeScript és picit összetettebb TypeScript projekteket is, hogy lássam az esetleges hiányosságokat. Természetesen ami eredményeket adtak a programok, azokat át kellett néznem egyesével, hiszen csak így tudom meg, hogy jól tesztelte-e a megadott fájlt vagy sem. Több hiányosságot is észrevettem, mind JSAN oldalról, mind JSAN2Lim oldalról, ezek kisebb hiányosságok voltak. Például, hogy egy node-nak nem volt beállítva pozíció, vagy nem volt jól beállítva a paramétere.

6. fejezet

Összefoglaló

6.1. Program javulása

A szakdolgozatom során sikerült elérnem azt, hogy a JavaScript-es nyelvi séma tartalmazza a TypeScript-es nyelvi elemeket is. Továbbá a JSAN2Lim-et sikerült kibővíteni TypeScript támogatással. Mindeközben a regressziós teszteléseket sikeresen frissítettem és validáltam, amit később mások is jóvá hagytak. Legvégül sikerült az AST binder metódusát optimalizálnom, sokkal gyorsabban fut le, mint ezelőtt.

6.2. Jövőbeli tervek

Rengeteg változtatás volt a TypeScript résznél, de a sémát nem tudtuk még naprakészre hozni, mivel voltak ennél fontosabb feladatok. Mint például az AST binder optimalizálása, vagy a JSAN2Lim továbbfejlesztése. Ebből kifolyólag jelenleg a séma nem naprakész. Mivel a szakdolgozatom elkészítése során megfelelő jártasságra tettem szert, így a későbbi bővítést az eddigiektől gyorsabban el tudnám végezni.

Nyilatkozat

Alulírott Pozsgai Alex programtervező informatikus BSc szakos hallgató, kijelentem, hogy a dolgozatomat a Szegedi Tudományegyetem, Informatikai Intézet Szoftverfejlesztés Tanszékén készítettem, programtervező informatikus BSc diploma megszerzése érdekében.

Kijelentem, hogy a dolgozatot más szakon korábban nem védtem meg, saját munkám eredménye, és csak a hivatkozott forrásokat (szakirodalom, eszközök, stb.) használtam fel.

Tudomásul veszem, hogy szakdolgozatomat a Szegedi Tudományegyetem Informatikai Intézet könyvtárában, a helyben olvasható könyvek között helyezik el.

Szeged, 2023. május 12.

.....
aláírás

Irodalomjegyzék

- [1] The javascript language. <https://javascript.info/js>.
- [2] Tibor Bakota, Péter Hegedűs, Péter Körtvélyesi, Rudolf Ferenc, and Tibor Gyimóthy. A probabilistic software quality model. In *2011 27th IEEE International Conference on Software Maintenance (ICSM)*, pages 243–252. IEEE, 2011.
- [3] Tibor Bakota, Péter Hegedűs, István Siket, Gergely Ladányi, and Rudolf Ferenc. Qualitygate sourceaudit: A tool for assessing the technical quality of software. In *2014 Software Evolution Week-IEEE Conference on Software Maintenance, Reengineering, and Reverse Engineering (CSMR-WCRE)*, pages 440–445. IEEE, 2014.
- [4] Árpád Beszédes, Rudolf Ferenc, and Tibor Gyimóthy. Columbus: A reverse engineering approach. *STEP 2005*, page 93, 2005.
- [5] Gavin Bierman, Martín Abadi, and Mads Torgersen. Understanding typescript. In Richard Jones, editor, *ECOOP 2014 – Object-Oriented Programming*, pages 257–281, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.
- [6] B. Cherny. *Programming TypeScript: Making Your JavaScript Applications Scale*. O’Reilly Media, 2019.
- [7] Asger Feldthaus, Max Schäfer, Manu Sridharan, Julian Dolby, and Frank Tip. Efficient construction of approximate call graphs for javascript ide services. In *2013 35th International Conference on Software Engineering (ICSE)*, pages 752–761. IEEE, 2013.
- [8] Steve Fenton, Fenton, and Spearing. *Pro TypeScript*. Springer, 2014.

- [9] Rudolf Ferenc, Árpád Beszédes, Mikko Tarkiainen, and Tibor Gyimóthy. Columbus-reverse engineering tool and schema for c++. In *International Conference on Software Maintenance, 2002. Proceedings.*, pages 172–181. IEEE, 2002.
- [10] Rudolf Ferenc, László Langó, István Siket, Tibor Gyimóthy, and Tibor Bakota. Source meter sonar qube plug-in. In *2014 IEEE 14th International Working Conference on Source Code Analysis and Manipulation*, pages 77–82. IEEE, 2014.
- [11] C. Nance. *TypeScript Essentials*. Community experience distilled. Packt Publishing, 2014.
- [12] Harry Porter. The abstract syntax tree. 2006.
- [13] Ferenc Rudolf. Sourceter for c/c++ with open-source sonarqube plugin released. 2015.
- [14] István Siket, Árpád Beszédes, and John Taylor. Differences in the definition and calculation of the loc metric in free tools. *Dept. Softw. Eng., Univ. Szeged, Szeged, Hungary, Tech. Rep. TR-2014-001*, 2014.
- [15] Gábor Szőke, Csaba Nagy, Rudolf Ferenc, and Tibor Gyimóthy. A case study of refactoring large-scale industrial systems to efficiently improve source code quality. In *Computational Science and Its Applications–ICCSA 2014: 14th International Conference, Guimarães, Portugal, June 30–July 3, 2014, Proceedings, Part V 14*, pages 524–540. Springer, 2014.
- [16] Typescript-eslint github repository. <https://github.com/typescript-eslint/typescript-eslint/tree/main/packages/ast-spec/src>.