

Szegedi Tudományegyetem
Informatikai Intézet

**Ipari JavaScript elemző kiegészítése TypeScript
támogatással**

**Industrial JavaScript analyzer enhancement with
TypeScript support**

Szakedolgozat

Készítette:

Pozsgai Alex

programtervező informatikus BSc
szakos hallgató

Témavezető:

Dr. Antal Gábor

Tudományos munkatárs

Szeged

2023

Feladatkiírás

Manapság kezd a piacon egyre több kódelemző megjelenni, többek között JavaScript programozási nyelvre is. Jelenlegi projekt is egy JavaScript Analyzer. A cél az, hogy a projekt több legyen, mint a piacon a többi kódelemző, ezért bővíteni kell TypeScript támogatással.

A hallgató feladata ennek a programnak (JSAN) a fejlesztése úgy, hogy tudjon TypeScript fileokat és projekteket is kellő pontossággal elemezni. Ezután pedig a meglévő programot optimalizálni, hogy kevesebb erőforrást vegyen igénybe a futtatása, és hogy gyorsabban fusson le.

A megoldási módszerek a hallgató kreativitására vannak bízva.

Tartalmi összefoglaló

A téma megnevezése:

JavaScript Analyzer kiegészítése TypeScript supporttal, JSAN optimalizálása.

A megadott feladat megfogalmazása:

A feladat során el kell érni azt, hogy a JavaScript Analyzer Tool tudjon TypeScript fileokra lefutni, és azokat nagy pontossággal elemezni. Emellett a JSAN működését optimalizálni.

A megoldási mód:

A megoldás során kell változtatni a JavaScriptSchémán, JSAN programban azon belül is az astTransformer.js fileban és a JSAN2Limben.

Alkalmazott eszközök, módszerek:

A megoldáshoz Visual Studio Code-t, a JavaScriptSchema szerkesztéséhez Visual paradigm-t használtam. A projekt lebuildeléséhez és ellenőrzéséhez Visual Studio 2017 programot használtam.

Elért eredmények:

JSAN képes TypeScript fileokat nagy pontossággal elemezni, nagyobb projektre lényegesen gyorsabban fut le, kevesebb erőforrást igényel, mint előtte. JSAn2Lim sikeresen alakítja át a JSAN outputot lim formátumba, typescriptes elemzések során is.

Kulcsszavak:

JavaScript, TypeScript, Optimalizálás, Visual Paradigm, C++

Tartalomjegyzék

Feladatkiírás	1
Tartalmi összefoglaló	2
Bevezetés	5
1. JavaScriptSchema átírása	7
1.1. A JavaScriptSchemáról	7
1.2. JavaScriptSchema átírása	12
1.3. Nehézségek, problémák	13
2. JavaScriptAddon	16
2.1. JavaScriptAddonról	16
2.2. JavaScriptAddon változások	20
3. JSAN2Lim átírása	21
3.1. JSAN2Limről	21
3.2. Bővítések	24
4. AST Binder Optimalizálás	26
4.1. A binderről	26
4.2. Lassúság okai	27
4.3. Optimalizálás	28
5. Regteszt frissítés	31
5.1. A regressziós tesztelésről	31
5.2. Tesztekről	32

5.3. Tesztek kibővítése	32
6. Összefoglaló	36
6.1. Program javulása	36
6.2. Jövőbeli tervek	37
Nyilatkozat	38
Irodalomjegyzék	39

Bevezetés

Szakdolgozatomban a JavaScript Analyzer továbbfejlesztése a cél. A továbbfejlesztés arról szól, hogy a javascript mellett előtérbe jön a typescript is. Ezáltal a typescript fileokat és projekteket is elemezni kell. Ez azt eredményezi, hogy a Javascript Analyzert gyökerestül át kell írni, annak érdekében, hogy JavaScriptet és TypeScriptet is tudjon egyaránt elemezni.

A JavaScript Analyzer az egy nagyobb projektnek az egyik alprojektje. Ezt a nagyobb projektet Analyzer-JavaScriptnek hívják. Az Analyzer-Javascriptnek több alprojektje is van, amiben sok minden egymásra épül. A szakdolgozatomban a következő alprojekteken fogok változtatni és bővíteni: JavaScriptSchema, JSAN (ezen belül is csak az astTransformerben), és az erre épülő JSAN2Lim. Emellett még megtalálhatóak a következő alprojekteket is: ESLintRunner, ESLint2Graph, LIM2Metrics, LIM2Patterns, DuplicatedCodeFinder és a ChangeTracker. Azért kell a JSAN2Limben is változtatni, mivel ha gyökerestül megváltoztatom a JSAN-t, akkor a JSAN2Lim rossz eredményeket fog visszaadni a typescript fileok elemzése közben.

A projekten amin fejlesztettem, többen is fejlesztettek egyszerre, ezáltal voltak átfedések, oszthatatlan részek. A projekt néhány része 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.

Fejlesztés során JavaScriptSchema file szerkesztése csoportos munka volt, hiszen minden erre épült, ez volt az alapja mindennek. Sok időt vett igénybe a szerkesztés. Nulláról kellett újraírni az egész vpp filet. Sajnos dokumentáció nem készült az előző

filehoz, ezért nekünk kellett kitalálni, hogy mi mit csinált, hiszen akik ezt írták, ők már nem foglalkoztak ezzel és elérhetetlenek voltak. A szerkesztéshez szükséges volt a Visual Paradigm alkalmazást használni. Ezzel egyikőnk sem találkozott még, szóval először ezt kellett tanulmányozni, megérteni. Ezután értelmezni kellett a meglévő schemát, hiszen eddig az jól működött, csak nem lehetett könnyen bővíteni. Mérlegeltük a két opciót, ahol vagy megpróbáljuk bővíteni a jelenlegi schemát, vagy nulláról elkezdjük újraírni. Végül az újraírás mellett döntöttünk, hiszen ezt láttuk gyorsabb és könnyebb megoldásnak. Egy könnyen bővíthető, dokumentál schema volt az elképzelés. Ketten fejlesztettük le végül ezt a schemát.

Ezután változtattam a JSAN2Lim-et, hogy képes legyen értelmezni a JSAN outputjait a typescriptes elemzések során. Itt át kellett néznem a c++ programozási nyelvet, hiszen az egész program ebben íródott, ugyanúgy dokumentáció nélkül. Több nehézségbe is ütköztem, de a legvégén sikerült olyan állapotra hozni a programot, hogy az output nagy százalékát sikeresen átkonvertálja lim fileba.

Ezután a JSANnak az astTransformer filejában a bindert optimalizáltam. Értelmezni kellett a kódot, át kellett néznem az ast-t, meg több adatszerkezetet is, gyorsasági szempontból.

Legvégül pedig miután minden készen volt, bővítettem a regteszteket új projektek behozatalával, mind javascripteseket és typescripteseket. Ezeknek az outputjait megnéztem, hogy jó eredményt ad-e, itt még volt egy kis bugfixing a JSAN2Limben és a Schémában.

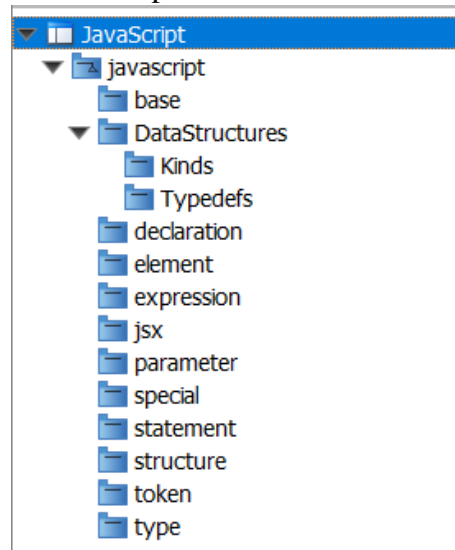
1. fejezet

JavaScriptSchema átírása

1.1. A JavaScriptSchemáról

A JavaScriptSchéma egy UML Diagramhoz hasonlító schéma. Jelen esetben a Visual Paradigm alkalmazással szerkeszthető. A felépítése a következőképpen alakul:

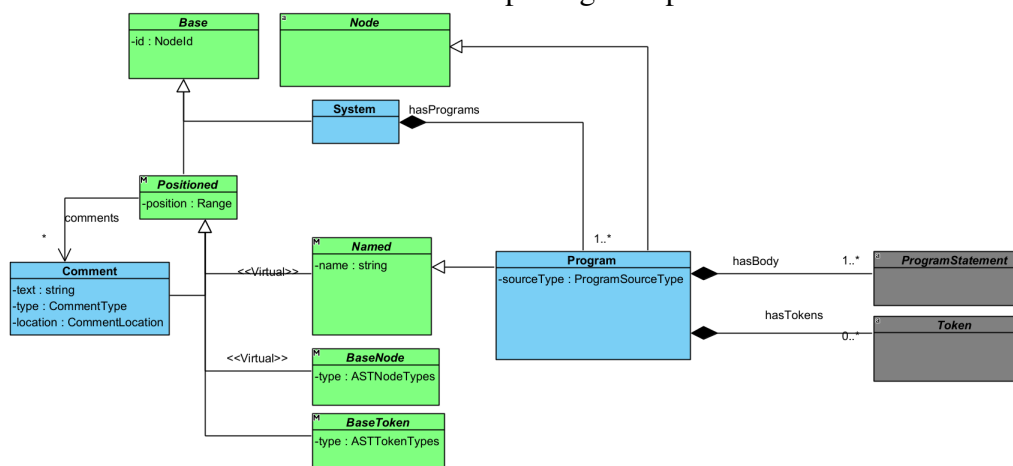
1.1. ábra. JavaScriptSchema struktúrális felépítése



Az 1.1 ábrán a következő struktúra figyelhető meg: ProjektNév/Model/Packagek. A projekt neve a JavaScript, ezen belül található egy model, amit javascript-nek hívnak. A modellen belül találhatóak meg a packagek. A Packagek nem véletlenül így lettek elnevezve. Az alábbi hivatkozáson található meg, hogy milyen logika alapján neveztük el a packageket: [1] Nem feltétlenül muszáj több package-t létrehozni, ez csupán az átlátható-

ság és a könnyen bővíthetőség céljából lett így megvalósítva. Követtük a typescript-eslint githubon lévő projekt struktúráis logikáját, de több helyen is eltértünk tőle, mivel a mi projektünk másabb, illetve speciális megoldásokat igényelt egyes helyeken. A következő oldalakon bemutatom a packagek felépítését néhány egyszerűbb példán szemlélítve. A packagek a következőképpen épülnek fel:

1.2. ábra. A base package felépítése



Az 1.2 ábrán különböző osztályok láthatóak. A Base osztály mindenek az alapja, ebből öröklődik minden más. Látható, hogy egy id attribútummal rendelkezik, aminek a típusa NodeId. A zöld háttérszínű osztályok az absztrakciót jelentik. Az asg fileban jelentősége, illetve a schéma szerkeszthetősége és olvashatósága miatt van így jelezve. A szürke háttérszínű osztályok pedig csupán annyit jelentenek, hogy más packageben lettek definiálva. A kék a default, normális osztályt jelölik. A Positioned és a System származik le a Baseből, értelemszerűen a system az maga a program lesz. A Positioned az azért absztrakt, mivel majd ebből fognak leszármazni a kisebb osztályok, mint például az expression, statement és a többi. A Positioned osztálynak van egy attribútuma, a position, ami egy Range típus. Emellett még tartozhatnak hozzá kommentek is. A komment egyaránt leszármazik a positioned-ből, ezzel garantálva azt, hogy a kommentnek van pozíciója. Látható, hogy a kommentnek van egy text, type és location attribútuma. A CommentType és a CommentLocation a DataStructures-ben van definiálva. Emellett a Named, BaseNode és a BaseToken származik le a Positionedből. A Named osztály azt jelenti, hogy egy nodenak van-e name attribútuma vagy sem. A BaseNodeból és a BaseTokenből fog nagyon sok minden leszármazni aminek van typeja. Végül, megtalálható a Program osztály,

aminek van name attribútuma, mivel Named-ből származik le, és van Systeme. Az 1.* jelenti azt, hogy legalább 1 Systeme van, de lehet több is. A hasPrograms-nak majd máshol lesz jelentősége, a mi esetünkben majd a javascriptben. Tetszőlegesen el lehet nevezni, de konzisztencia miatt, minden attribútum ami osztály (és nem a DataStructuresben azon belül is a Kindsban van definiálva, hanem van neki egy osztály, mint pl Comment) azt a hasOsztály névvel fogjuk ellátni. A ProgramSourceType is a Kinds packageben van definiálva, ami csupán azt mondja meg, hogy az adott program vagy script vagy module típusú. A base package nem követi a typescript-eslint githubon lévő projekt base map-páját, ez teljesen egyedi, átememeltük az egészet apróbb módosítással az előző projekt verzióból.

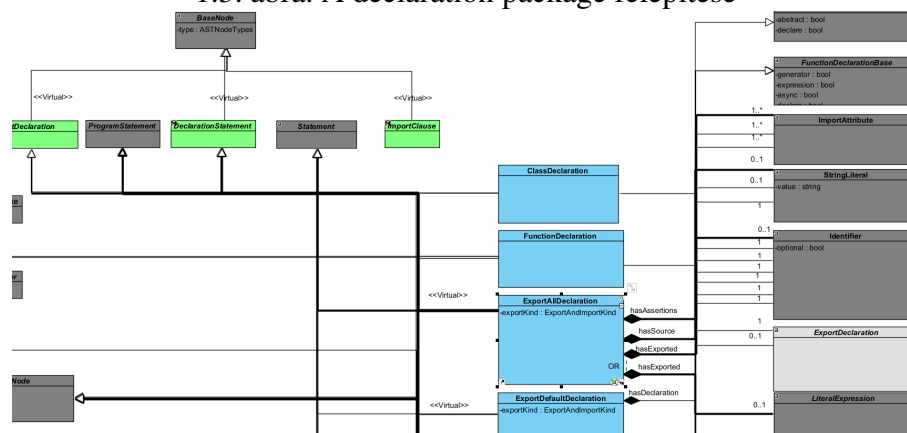
A base packagen kívül bemutatom a declaration package-t, hogy lássuk milyen logikát követtünk a megírás során. Ezen belül is az ExportAllDeclaration bemutatása:

Az 1.1 kódrészlet így lett megvalósítva a JavaScriptSchémában:

```
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 1.1. ExportAllDeclaration typescriptes megvalósítása

1.3. ábra. A declaration package felépítése



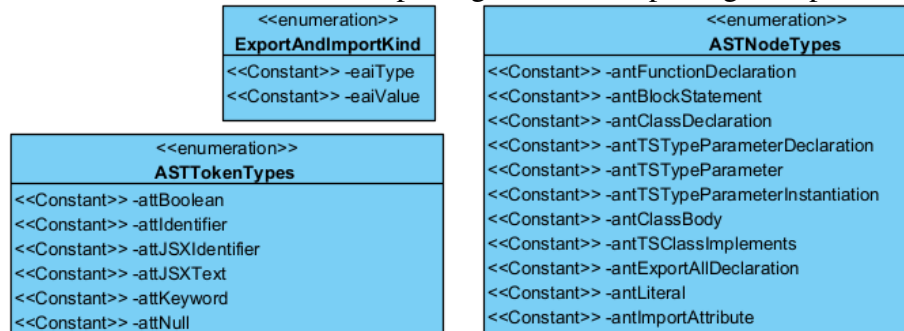
Az 1.3 ábrán látható az, hogy minden a BaseNodeből származik. Az 1.1 kódrészlet-

ben ez az extends BaseNode. Itt maga a declaration osztály az a DeclarationStatement névre hallgat, csupán azért, mert követtük a typescriptes kódot. A főbb osztályok a BaseNode alatt találhatók. Jobb oldalt a sötétszürke és a világosszürke osztályok azok az attribútumok. Ebben az esetben az ExportAllDeclaration osztály leimplementálását mutatom be a JavaScriptSchemában. Az ExportAllDeclaration öröklődik a Statement, DeclarationStatement, Node és a ProgramStatementből. Ezeket az öröklődéseket ugyanúgy a typescript-eslint githubról néztük, unions mappában találhatóak el. Ezáltal megkapja az összes szülőnek a tulajdonságait. A DeclarationStatement öröklődik a BaseNodeból, ami azt jelenti, hogy a BaseNode tulajdonságait is megkapja az ExportAllDeclaration. A BaseNode ugye meg származik a Positionedből, ez látható az 1.2 ábrán. Emiatt az ExportAllDeclaration-nek lesz pozíciója, kommentje és NodeId-ja is. Assertions attribútum típusa az ImportAttribute, látható, hogy egy tömböt vár, ezért 1..* a multiplicityje. Exported attribútumnál egy Identifier típusú attribútumot vár, itt mi kiegészítettük még egy LiteralExpressiönnel is, ami maga a Literal. A vagyolást egy OR-al jeleztük a JavaScriptSchemában. Mivel lehet null-is, ezért 0..1 a multiplicity, szóval vagy 0 vagy 1. Az ExportKind az a Kindsban található meg, így szimplán csak arra hivatkozunk, mint ExportAndImportKind. Végül a source attribútuma egy StringLiteral, nálunk is így szerepel. Természetesen ami sötétszürkével van jelölve, az máshol létre van hozva és vannak neki attribútumai. A világosszürke annyit jelöl, hogy ebben a packageben lett létrehozva az osztály, de láthatóság szempontból többször szerepel, mivel lehet attribútum is. Így lett minden egyes osztály felépítve, természetesen Az 1.3 ábra csak egy részlete a declaration package-nek, ennél jóval nagyobb. Vannak úgynevezett gyűjtő osztályok is, mint pl a Statement, ezekre azért van szükség, mert majd később ha le lett generálva minden, akkor tudunk majd vizsgálni különböző nodeokra, mint például az isStatement.

Sokat említettem a DataStructurest. Hadd mutassam be egy példán keresztül, hogy ez hogy néz ki.

A DataStructures packageben található 2 package, ez az 1.1 ábrán látható. Ebből a Kinds paketet mutatom be. Az 1.4 ábrán látható egy kis szelet a Kinds packageből. Enumok találhatóak ebben a packageben. Például ha az ExportAndImportKind-ot adjuk meg típusnak az egyik attribútumnak, akkor az lehet vagy eaiType vagy eaiValue. Minden constant előtt van 3 karakter, ez azért szükséges, mert később ezekkel még foglalkozni

1.4. ábra. A DataStructures packageben a Kind package felépítése



fogunk.

A JavaScriptSchema magában még semmit sem csinál. Először ki kell exportálni az egész projectet xml formátumba. Ezután az xml filet átkonvertáljuk egy asg filera. Ez úgy történik, hogy a JavaScript Analyzer projektnek van egy kisebb alprojektje, amit UmlToAsg-nek hívnak. Ez a java kód átkonvertálja az xml fileban lévő adatot egy asg fileba. Jobban nem térnék ki az UmlToAsg projektre, mivel nem szerkesztettük. Az UmlToAsg projektet a SchemaGenerator generálta le. A SchemaGenerator c++ nyelven megírt program. Több mindent is generál c, c++, vagy java nyelven. Ez a projekt is az Analyzer JavaScript alprojektei közé tartozik. A file elején a következő található meg:

```

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

Kódrészlet 1.2. Asg file első sorai

A verziókat kézzel tudjuk átírni abban a fileban ami generálja ezt, ez egy c++ file, a SchemaGeneratorban található meg. Ezután a Kinds mappa tartalmát írja bele a következőképpen:

```

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

Kódrészlet 1.3. Asg file kind

Az a 3 karakter amit minden constant elé tettünk azt kitette paraméterbe és csak az utáni stringet írta át. Ugyanígy van a többi kindnál is. Ha az összes kindot beleírta, akkor kezd

írni sorban a többi package-t. Az 1.1 alapján megy sorba. Példának a declaration package-t mutatom be, azon belül is az ExportAllDeclaration-t.

```
1 SCOPE declaration {
2
3     NODE DeclarationStatement : virtual base::BaseNode [ABSTRACT] {
4     }
5
6     NODE ExportAllDeclaration : DeclarationStatement, statement::
        Statement, virtual statement::ProgramStatement, special::Node
        {
7         ATTR ExportAndImportKind exportKind;
8         EDGE TREE 1 hasExported (expression::Identifier |
            expression::LiteralExpression);
9         EDGE TREE 1 hasSource (structure::StringLiteral);
10        EDGE TREE * hasAssertions (special::ImportAttribute);
11    }
12 }
```

Kódrészlet 1.4. Asg file ExportAllDeclaration

Látható, hogy a Package SCOPE-nak értelmezi, és ezen belül NODE-ok találhatóak. A Nodeoknak ATTR és EDGE TREE van. A vagyolás is látható. Az öröklődik egy kettőspont után, felsorolás szerűen írta át, ha valami másból származik, akkor package-Nev::osztalyNev szerint. Az kapott ATTR jelölést ami a Kinds-ban megtalálható vagy egy szimpla típus (mint pl string, int). Minden mást EDGE TREE-nek nevezett el. Az Edge tree utáni szám vagy csillag az a multiplicitást jelenti. Ha 1es, akkor a multiplicitás 0..1, ha *, akkor vagy 0..* vagy 1..* a multiplicitás. Az 1.3as ábrán látható, hogy mit hogyan írt át.

1.2. JavaScriptSchema átírása

A JSAN a JavaScriptSchemára épül. Ha valamit meg akarunk változtatni gyökerestül a JSANba, akkor a schémát is változtatni kell. Azt elérni, hogy javascript mellett még typescriptes kódokat is elemezzen a JSAN, ahhoz gyökerestül meg kellett változtatni a JavaScript Analyzert. Az előző JavaScriptSchema (ami csak javascriptet elemzett) az a

javascript hivatalos oldala alapján készült. Ami ábrák fentebb megtalálhatóak, azok már az átírt schémából vannak. Több opció is volt, hogy most vagy legyen átírva a schéma, vagy legyen egy külön typescriptre is írva. Végül rájöttünk, hogy ha van egy typescriptes schémánk, az képes javascriptet ugyanúgy elemezni ha megfelelően van lefejlesztve. Így egy schémát fejlesztettünk le. Az előző schéma átláthatatlan volt, ezért úgy döntöttünk, hogy majdnem a nulláról újraírjuk. Egyedül a base package emeltük át a régiből, BaseNode-al és BaseTokennel kibővívte (Az 1.2 ábrán látható). Emellett el kellett dönteni, hogy milyen struktúrát kövessünk, ami jól átlatható és később könnyebben bővíthető. Végül a typescript-eslint official github alapján haladtunk. Annyi változtatással, hogy ami nekik a base mappában volt, mi arra létrehoztunk egy külön structure package-t. Mindenhez készítettünk dokumentációt, jól érthetően leírtuk, hogy mit miért kellett csinálni. Mivel minden is egymásra épül a typescriptes schémában, ezért nem tudtuk tesztelni minden egyes package után, hogy működik-e vagy sem. Mint például, látható az 1.4 kódrészleten, hogy az ExportAllDeclaration mennyi mindenből származik le vagy mennyire sok attribútuma van. Az ExportAllDeclaration egyik attribútuma, az assertions típusa az ImportAttribute. Ahhoz, hogy jól észrevegye az ExportAllDeclarationt az Analyzer, ahhoz jól le kellett fejleszteni az ImportAttributet.

```
1 export interface ImportAttribute extends BaseNode {  
2     type: AST_NODE_TYPES.ImportAttribute;  
3     key: Identifier | Literal;  
4     value: Literal;  
5 }
```

Kódrészlet 1.5. ImportAttribute

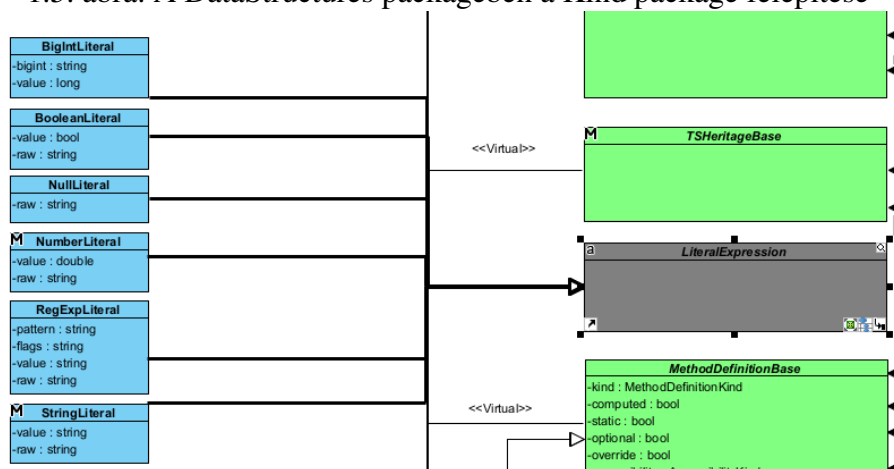
Ahhoz, hogy az ImportAttributet letudjuk fejleszteni a schémában, ahhoz le kell fejleszteni az Identifiert, és a Literalt. Kicsit összetett, hogy mi minden épül egymásra. Ebből adódik a következő alfejezet mondandója is.

1.3. Nehézségek, problémák

A schéma átírása során több nehézségbe is ütköztünk. A legelső nehézség az a Visual Paradigm program korrekt használata volt. Egy kisebb időbe telt rájönni arra, hogy egy

adott classnak hogyan kell megváltoztatni a háttérszínét, packagek közötti mozgást, hogyan kell egy adott classra hivatkozni ami másik packageben volt. Ezután ami szerintem a legnagyobb nehézség volt, az az, hogy nem volt dokumentáció az előző schémához. Mindent nekünk kellett kitalálni, hogy mit miért csináltak az előző fejlesztők. Akik ezt a schémát fejlesztették le, ők már nem foglalkoztak ezzel, és más nem nagyon mélyedt bele ebbe az egészbe azóta. A következő nagyobb nehézség, az az volt, hogy összehozunk egy olyan schémát ami működőképes, és az analyzer ezt tudja is használni. Előző alfejezetben említettem, hogy egy adott classt (mint például ExportAllDeclaration) lefejllesszünk, ahhoz sok minden mást is le kell fejleszteni, amihez meg még több minden kellett. Ebből adódóan tesztelni nagyon nem tudtunk, mivel az egésznek működnie kellett ahhoz, hogy az analyzer egyáltalán lefusson. Mivel mi majdnem nulláról írtuk újra, ez volt a hátrány. Amikor elérkeztünk ahhoz az állapothoz, hogy mindenre rámondjuk azt, hogy működik, akkor jött egy nagyobb probléma. Valahol Segmentation Fault-ot kapott a program, és nem nagyon tudtuk ezt debugolni. Több sejtésünk is volt, hogy mi lehet a baj. Emiatt az egész schémát át kellett nézni alaposan, hogy hol vétettünk hibákat. Sok helyen voltak pontatlanságok, rossz osztályból származtattunk le, rossz típus volt megadva attribútumnak. Ezeket mind kijavítottuk, de most már más hibát kaptunk. A JavaScript Analyzerben kiderítettük, hogy pontosan hol szállt el a program. A Literal volt a hiba, mivel ezt nem a github alapján írtuk meg, hanem egyedi ötlettel. Előző schémában is máshogy volt megoldva.

1.5. ábra. A DataStructures packageben a Kind package felépítése



Az 1.5 ábrán látható, hogy a LiteralExpressionből (Ami a Literal) több literal is örök-

lődik.

```
1 export interface LiteralBase extends BaseNode {  
2   type: AST_NODE_TYPES.Literal;  
3   raw: string;  
4   value: RegExp | bigint | boolean | number | string | null;  
5 }
```

Kódrészlet 1.6. Literal

Próbáltuk úgy megoldani, hogy 4 vagyolással 4 különböző literál tartalmazza, de az analyzerben máshogy kezeltük le a literalt, mint nodeot. Ezért a tartalmazás helyett inkább örököltettünk a literalból, így megoldva ezt a problémát. Még az is probléma volt, hogy volt egy LiteralBaseünk, amiből származott ez az 5 literal. A LiteralBase származott a LiteralExpressionből, de valamiért Segmentation fault lett a vége ha így próbáltuk megoldani. Ezért a LiteralBase-t kivettük, és LiteralExpressionből származik minden literal. Végül még az volt a nehézség, hogy kódrészletet keressünk az analyzernek, amin le tudjuk tesztelni, hogy az analyzer jó outputot ad-e.

Mivel nem nagyon volt typescriptes háttértudásunk, először a typescriptet kellett átnézni, hogy mit hogyan tudunk megvalósítani. Emellett még a javascriptes outputokat is át kellett nézni, mivel a mi célunk a fejlesztés volt. A jelenlegi javascriptes projektekre ugyanolyan eredményt adott, sőt néhány helyen jobbat is, mert az előző schémában is voltak pontatlanságok. A typescriptes projektek nagy részét tudja elemezni az analyzer, de közel sem tökéletes, mivel egyfolytában kellene fejleszteni a schémát, mivel a typescript nem annyira régi.

2. fejezet

JavaScriptAddon

2.1. JavaScriptAddonról

A JavaScriptAddon az egy node kiterjesztésű file, amit a SchemaGenerator generál. Minden egyes nodenak külön generál nodeWrapperheader és egy nodeWrapperCC file. A SchemaGenerator által generált Factory.cc, Factory.h és az összes nodeWrapper.cc és a hozzá tartozó nodeWrapper.h fileok összemergelése eredményezi a javascriptAddon.node file. A SchemaGenerator a JavaScriptSchemából és az ebből generált asgből generálja le a javascriptAddon.node file. A SchemaGeneratorknak több generálása is definiálva van, hiszen nem csak erre használják. Nekünk a NodeAddonGenerator.c fileban és a hozzá tartozó header fileban van minden. A main.c-be ezt beincludeolja, és ennek hívja meg egyes metódusait, ami által elkészül a javascriptAddon. A következő pár oldalon bemutatnám, hogy a generálás hogyan zajlik. A SchemaGeneratorknak rengeteg kapcsolója van, köztük a generateNodeAddon is. Ezekre a kapcsolókra vizsgál egyesével, egy nagy if-elseben.

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

Kódrészlet 2.1. SchemaGenerator kapcsoló vizsgálás

A kód elején történik meg ez, ha meg van adva paraméternek a genNodeAddon string, akkor az options.generateNodeAddon-t igazra állítja. A default értéke false. Ezután jóval lentebb miután rengeteg mindent legenerált ami kell alapból is a normális működéshez, vizsgál egyet az options.generateNodeAddon-ra. Ez látható a 2.2 kódrészleten is. Else

ága nincs, szóval semmi nem történik ha nincs megadva a genNodeAddon string a paramétereknél.

```
1 if(options.generateNodeAddon ) {
2     debugMessage(0, "Generating Node.JS Addon sources\n");
3     if (createAndEnterDirectory(SOURCE_NODE_ADDON_DIR_NAME)) {
4         generatePackageJson();
5         generateBindingGyp();
6         generateAddonCC();
7         generateFactoryWrapper();
8
9         if (createAndEnterDirectory("inc")) {
10            generateWrapperHeaders();
11            leaveDirectory();}
12        if(createAndEnterDirectory("src")){
13            generateWrapperSources();
14            leaveDirectory();}
15
16        leaveDirectory();}
17 }
```

Kódrészlet 2.2. SchemaGenerator javascriptAddon generálás

A createAndEnterDirectory metódus annyit csinál, hogy létrehoz egy mappát és ch-dirrel belelép. Ha az adott mappa már létezik, akkor csak szimplán belelép. A SOURCE_NODE_ADDON_DIR_NAME változó jelen esetben addon értéket fog kapni, hiszen a javascriptAddon dolgai ebbe fognak generálódni. Miután létrehozta és/vagy belelépett az addon mappába először legenerálja a package.json file-t a projekthez. A package.json fileban beállítja a projektnevét, verziószámát, dependencyket, scripteket és a végén a gyp-file kapcsolónak egy true-t beállít.

```
1 fprintf(f, "    \"rebuild\": \"node-gyp configure && node-gyp rebuild -\n\nej 8\\",\n");
2 fprintf(f, "    \"install\": \"node-gyp configure && node-gyp build -j\n8\\",\n");
```

Kódrészlet 2.3. NodeAddonGenerator package.json scripts

A 2.3 kódrészleten látható, hogy majd gyp segítségével fog történni a generálás.

Ezután legenerálja a binding.gyp fílet. Ez a fíle fog felelni azért, hogy a javascriptAddonhoz sikeresen generálódjanak le a wrapperek és azok headerjei. Ha ez megtörtént, utána fogja legenerálni az addonCC-t. Az addon.cc fíleban beimportálja a Factory.h header, amiben több metóduis is megtalálható. ,

```
1 if (!traversalDescendantBFT(rootNode, generateWrapIncludes, false)) {
2     debugMessage(0, " failed\n");
3     fclose(f);
4     return false;
5 }
```

Kódrészlet 2.4. Addon.cc Wrapperek includolása

A 2.4 kódrészletben a traversalDescendantBFT metóduis egy másik projektből includoltuk. Ez egy bejárás, ami az összes nodera lefut, és az összes nodera meghív egy metóduis, jelen esetben a generateWrapIncludes metóduis. Ez a generateWrapIncludes a nodeAddonGenerator.c fíleban van leimplementálva a következőképp: ,

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

Kódrészlet 2.5. generateWrapIncludes leimplementálása

Először megvizsgálja, hogy az adott node abstract-e, ha nem akkor tovább megy, és az addon.cc-be includolja az adott nodeWrappernek a headerjét. A node->name lehet például FunctionDeclaration, ClassDeclaration, amik megtalálhatóak az asg fíleban.

Ezután ezt a bejárást még egyszer végrehajtja, de most a wrapperIniteket generálja az addon.cc fíleba. Annyi változtatással, hogy picit mást ír az addon.cc fíleba.

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

Kódrészlet 2.6. generateWrapInit leimplementálása

Ha ezek megtörténtek, akkor az addon.cc fílet legeneráltuk sikeresen, és így benne találhatóak a wrapperek headerjének includolása és a wrapperek Initjei.

Ezután következik egy `generateFactoryWrapper` hívás. A `generateFactoryWrapper`-ben 2 metódus hívás található, a `generateFactoryWrapperHeader` és `generateFactoryWrapperSource`.

A `generateFactoryWrapperHeader` a `Factory.h` filet generálta le a következőképp: Először beincludolja az összes `nodeWrapper`-nek a headerjét, és utána létrehoz egy `Factory` osztályt, publikus metódusait, ami `Init` és `Destructor`, illetve a `private` metódusait, ami a `destructor`, `New`, `SaveAST`, `LoadAST`, `Clear`, `getRoot` és az összes `nodeWrapper` `create` metódusa. Erre azért van szükség, mivel majd a JSAN-ban ezeket a wrappereket fogjuk létrehozni és szerkeszteni.

A `generateFactoryWrapperSource` pedig a `Factory.cc` filet generálja le. Ebben a file-ban pedig meg van az összes metódus valósítva, ami a headerjében található. Emellett a `Factory` `Init`-jének a `props` változója a következőképpen néz ki:

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

Kódrészlet 2.7. `Factory.cc` file

`DeclareNapiMethodokat` hoz létre, az összes `nodeWrapper`-nek, ezek majd a JSAN-ban lesznek használatosak.

Utolsó lépésként, először az `inc` mappát majd az `src` mappát generálja le a `SchemaGenerator`. Az `inc` mappában találhatóak a header fileok egyes wrappereknél, az `src` mappában maguk a wrappereket vannak megvalósítva. Minden nodelnek hoz létre wrapper file-t, és ezekben valósít meg néhány metódust. Vegyük most egy példának az `ExportAllDeclaration`-t. Létrehoz egy `ExportAllDeclarationWrapper.h` és egy `ExportAllDeclaration.cc` file-t. A header fileban létrehoz egy `ExportAllDeclarationWrapper` osztályt, ami származik a `BaseWrapper`-ből. A `BaseWrapper` az egy alap `Wrapper` osztály, amit bővíteni fogunk. Három publikus metódust generál, `Init`, `Destructor` 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 2.8. ExportAllDeclaration.h file

Látható a 2.8 kódrészleten, hogy létrehozott az ExportAllDeclarationWrappernek egy setPosition, addAssertions, setType és még több metódust. Ezek az ExportAllDeclaration attribútumai, amit a JavaScriptSchemában beállítottunk. Az 1.3 ábrán látható, hogy az ExportAllDeclaration hasAssertions attribútumának a multiplicitása 1..*, ezt átkonvertálta addAssertionsre. Ahol 0..* a multiplicitás, azt átkonvertálta setAttribute-ra, pl setSource vagy setExportedre. Minden nodeWrappernek van olyan metódusa, hogy setPath, setPosition, setType, addComments, hiszen minden a BaseNode vagy a BaseTokenből öröklődik le.

Az ExportAllDeclaration.cc fileban ugyanaz történik, mint ami volt a Factory.cc fileban. Declare napi methodokat generál, a setExported, setSource, setType és a többi metódushoz. Ezután az összes metódust megvalósítja.

2.2. JavaScriptAddon változások

A JavaScriptSchema változtatások után, a NodeAddonGeneratort nem módosítottuk. Ennek ellenére a javascriptAddon.node mérete megduplázódott, csupán azért, mert a schéma ekkora bővítésen esett át. Sokkal több nodeWrapper lett, mivel bejöttek a typescript általi dolgok is a javascript mellett. Természetesen a javascriptes dolgok megmaradtak, így javascriptet ugyanolyan jól tud elemezni, sőt néhány esetben jobban is, mert voltak figyelmetlenségek előző verzióban.

Mivel a javascriptAddon mérete nőtt, ezáltal a sebesség csökkent. Ezt majd egy másik fejezetben jobban kifejtem, célunk az volt, hogy először tudjon typescriptes fileokat és projekteket elemezni a jsan.

3. fejezet

JSAN2Lim átírása

3.1. JSAN2Limről

A JSAN2Lim egy c++-ban megírt program. A JSAN általi outputot átalakítja lim formátumú fileokra. Azért lim formátumú fileokra, mivel a lim az egy környezetfüggetlen nyelv, és sokkal könnyebb ezen metrikákat mérni. A JSAN2Lim is egy alprojektje az Analyzer JavaScriptnek. Fontos, hogyha a JSAN programot fejlesztik, akkor a JSAN2Lim-et is fejleszteni kell, mivel ha nem, akkor a metrikák nem fognak jól számolódni. A JSAN2Lim használja a LimSchemat, ez is ugyanúgy egy *.vpp* kiterjesztésű file. Más logikával van felépítve, mint a JavaScriptSchema, ezért nem is nagyon mennék bele az elemzésébe, csak csupán néhány dolgot fogok használni belőle. A JSAN2Lim visiteli az összes nodeot amit a JSAN outputba kirak, ezért JSAN2LimVisitor a fő file neve. Minden egyes nodera külön meg kell írni a visit függvényt, hogy tudja a c++ program, hogy mit csináljon ha olyan nodeot kap.

A JSAN2Lim működését egy kódrészleten keresztül mutatom be.

```
1 void JSAN2LimVisitor::visit(const javascript::asg::structure::  
    ClassDeclarationBase& clNode, bool callVirtualBase)  
2 {  
3     VISIT_BEGIN(clNode, callVirtualBase);  
4     lim::asg::logical::Class& classLimNode = dynamic_cast<lim::asg::  
        logical::Class&>(createLimNode(clNode, callVirtualBase));  
5     fillData(classLimNode, clNode); //sets the name  
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.
getSuperClass())) {
12             javascript::asg::expression::Identifier& superClass =
dynamic_cast<javascript::asg::expression::Identifier>(*clNode.
getSuperClass());
13             classStack.top().isSubclass = superClass.getId();
14         }
15     }
16     classStack.top().TLOC = clNode.getPosition().getEndLine() -
clNode.getPosition().getLine() + 1;
17     classStack.top().LOC = clNode.getPosition().getEndLine() - clNode.
getPosition().getLine() + 1;
18     if (!packageStack.empty()) {
19         lim::asg::logical::Package& packageLimNode = dynamic_cast<lim::
asg::logical::Package>(limFactory.getRef(packageStack.top().
packageNodeId));
20         SAFE_EDGE(packageLimNode, Member, lim::asg::logical::Member,
classLimNode.getId());
21     }
22     else {
23         common::WriteMsg::write(common::WriteMsg::mlDebug, "Empty
packageStack while visiting node %d.", clNode.getId());
24     }
25     addIsContainedInEdge(classLimNode, clNode);
26 }

```

Kódrészlet 3.1. ClassDeclaration Visitor

A VISIT_BEGIN definiálva van a file elején, ami a VisitorAbstractNodes-nak a visit metódusa. Ezután létrehozok egy classLimNode nevű változót, ami *lim::asg::logical::Class* típusú, ez a limFactoryba található meg. Ennek a változónak adunk egy értéket, ami a createLimNode által visszaadott node lesz, és ezt castoljuk erre a típusra.

A `createLimNode` meghívja a `limFactory`-ból a `createNode`-ot, ami vár paraméterbe egy `nodeKind`-ot. Ezt a `nodeKind`-ot mi adjuk meg neki, az alapján, hogy az adott `node` milyen típusú. A `nodeKind`-ot a `getLimKind` függvény határozza meg a következők szerint. Megvizsgáljuk, hogy az adott `node` `getIsClassDeclarationBase()` vagy `getIsMethodDefinition` vagy `getIsFunctionDeclarationBase` vagy sorolhatnám még. Ebből van 10-15db. Típustól szerint adjuk meg a `nodeKind`-nak az értékeket, amik lehetnek a következők: `ndkClass`, `ndkMethod`, `ndkParameter`, `ndkMethodCall`, `ndkAttribute`, `ndkComment`, `ndkPackage`, `ndkAttributeAccess`, `ndkComponent`. Ezek mind a `lim`-nek az asgben való típusok és mindet tudja kezelni. Miután beállítottam a `nodeKind` értékét, szimplán `return`olom. Ha megkaptuk a `nodeKind` értékét, akkor ezt létrehozza a `createLimNode` függvény. Állít be ennek pozíciót és `return`oli a `*limNode`-ot.

Ezáltal most megkaptuk a `classLimNode`-ot, hiszen a `nodeKind` az `ndkClass` lett és ennek megfelelően hozta létre a `lim`. Ezután meghívjuk a `filldata` függvényt, első paraméternek magát a létrehozott `classLimNode`-ot, másodiknak magát a `node`-ot adjuk meg. A JSAN által kiadott `output`-ot `.jssi` kiterjesztésű fileon végig megy. Visitorok segítségével oldja meg ezt. Ebben a `filldata`-ban beállítom a `limnode` láthatóságát, nyelvét (ami konstans javascript), azt, hogy `abstract`-e. Lekérem a `nodeId`-ját és a `lastLimMemberNodeId` illetve a `lastLimScopeNodeId`-ba pusholom. Ezután beállítom a `classKind`-ját, ami `clkClass` lesz (ezt is a `factory`-ból). Aztán a nevét szeretném beállítani, a `node`-nak lekérem azt, hogy `identifier`-e, vagy `sem`. Ha nem `null`, akkor lekérem az `identifier`-t és `identifier.getName()`-el lekérem a nevét és be is állítom a `limNode`-nak a `setName()` függvénnyel. Ha az `identifier` `null`, akkor a `limNode.setIsAnonymous()`-t beállítom `true` értékre és adok neki egy `"anonymousClass"` + `ss.str()` ahol az `ss` az egy `stringstream`, szimplán egy `counter`, hogy eddig hány `anonymous class` volt. Ezután megvizsgálom, hogy a `methodstack` `empty`-e, ha nem, akkor van szülője, ezért a `SAFE_EDGE` function hozzáadok egy `edge`-t a `parent` és a jelenlegi `limnode`-hoz. Ezt ugyanúgy megvizsgálom `packageStack`-re, ha nem `empty`, akkor ugyanezt elvégzem csak a `parent` most más lesz. Végezetül az `addIsContainedInEdge` függvényt meghívom a `limnode` és a `jsNode` paraméterekkel.

Ezután a `fillMemberData()` függvényt hívjuk meg, ami igazából csak pozíciót állít be, illetve a neven picit változtat. Utána a `classStack`-hez hozzáadja az adott osztály

tulajdonságait, beállítja a parent classt ha van, pozíciót állít be, és a végén ugyanúgy meghívjuk a *addIsContainedInEdge* a *classLimNode* és a *clNodeal*. Ez a procedúra történik minden egyes nodenál a JSAN2Limben.

3.2. Bővítések

Mivel a JSAN mostmár tud typescriptet is elemezni, emiatt a JSAN2Lim-et is fel kell erre készíteni, hogyha typescriptes nodeokat kap, akkor tudja, hogy mit kell azzal csinálni. A javascript asg-t is használja a JSAN2Lim, ezért sok helyen először változtatni kellett néhány dolgon.

- Néhány visitornál a várt paraméter típusát átírni, például *Class*-ról *ClassDeclarationBase*-re, hiszen a schémában változtatás volt.
- Néhány visitornál a várt paraméternél a helyét az adott nodenak. Például a *RestElement* már nem a *statement* package-n belül van, hanem a *parameter*-en belül.
- A *Pattern* mint node megszűnt, *Parameter*-re lett átnevezve, és a helye is megváltozott, *statement*-ből átkerült *parameter* package-be
- Típus lekérdezésnél a *getIsClass()* helyett például *getIsClassDeclarationBase()* lett, változott ez is a *JavaScriptSchema* miatt.

A bővítések, hogy miket adtam hozzá a JSAN2Limhez:

- Az új Kindokat amiket felvettünk a *JavaScriptSchema*-ban. A következőképp: *std :: stringgetText(javascript :: asg :: ASTNodeTypes);*
- Új visitorok írása. *TSEnumDeclaration*, *TSImportEqualsDeclaration*, *TSInterfaceDeclaration*, és még sok másik, ezekhez a filldatákat is megírni
- *kindStrings* tömb kiegészítése a typescriptes nodeokkal
- Bugok kijavítása, a *VariableDeclaration*-nél nem vette észre az összes *declaration*-t, főleg metóduson belül.

- A *getLimKind()* function bővítése *getIsTSTypeAliasDeclaration()*, *getIsTSInterfaceDeclaration()*, *getIsTSAbstractMethodDefinition()* és *getIsTSAbstractPropertyDefinition()* else if ágakkal.
- TSEmptyBodyFunctionExpression debugolása, sok helyen lehalt a program emiatt.

4. fejezet

AST Binder Optimalizálás

4.1. A binderről

Az AST binder referenciákat bindol. Ezek a referenciák a JSAN2Lim programhoz kellenek. Binder néven van definiálva az astTransformer.js fileban. Kétszer van használva, ezért is kulcsfontosságú, hogy gyors legyen. Egyszer VariableUsages referenciákat bindol és egyszer ACG referenciákat bindol.

A binder 4 argumentumot vár:

- Egy stringet, ami lehet vagy VU (ami VariableUsages-t rövidíti) vagy ACG (egy javascript callgraph).
- Egy Abstract Syntax Tree-t (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 javascriptAddonban található meg és onnan hívódik meg. Akkor addCalls ha ACG referenciákat bindolunk és akkor setRefersTo ha VU referenciákat.

A linkeket tartalmó tömb egyik JSON objektuma a következőképp 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 argumentuma

A 4.1 kódrészletben egy JSON elemét láthatjuk a tömbnek. Ilyen JSON elemekből épül fel a tömb. A source és a target felépítése ugyanaz, csupán másak az értékek bennük. A kódrészletben csak a source-t fejtettem ki kicsit bővebben. A label az egy Identifier vagy PrivateIdentifier nevét fogja jelölni. A file egy abszolút útvonalat, ami az Identifiert tartalmazó filet jelöli. A start az egy object, 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 object, 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 object ami maga az Identifier vagy PrivateIdentifier.

4.2. Lassúság okai

A binder nagyobb projektekre lassan fut le. Ennek több oka is van:

- Minden hívásnál meghívja a javascriptAddon.nodeot, ami már magában nagy file. TypeScript támogatás után kétszeres lett a mérete, ezáltal sokkal lassabb lett.
- Mivel nagyobb a projekt, ezért sokkal több a függvényhívás is az elemzett kódban, emellett sok a változószám is.
- Az AST nagyon nagy, és ezt bejárjuk többször is bindolás alatt. Emiatt nagyon lassú a program nagy projektek esetén.
- A JSONokat tartalmazó tömb is nagyon nagy lesz, de annyira ez nem lassítja le a programot.

4.3. Optimalizálás

Először ki kellett derítenem, pontosan melyik funkció mennyi memóriát vesz igénybe és milyen gyorsan fut le. Erre találtam egy javascript profilert, ami kiírta minden függvényhívásnál a lefoglalt memóriát és a memória használatot. Ez nekem nem volt a legjobb, mivel nagyon egybe van ágyazva minden, és eléggé mélyre mentek a functionok. Ezután próbáltam picit egyszerűbbet, a binder kódját 3 részre osztani, `console.time()` és `console.timeEnd()` használatával. Ez a `console.time()` parancs kiírja nekem ms-ben, hogy mennyi idő telt el amíg eljutott a `console.timeEnd()`ig. Így megkaptam, hogy melyik rész nagyjából mennyi idő alatt fut le, könnyebb volt szűkíteni, hogy melyik function lesz a bajos. Végül kiderült, hogy a következő sor lassítja be nagyon a programot:

```
1 globals.getWrapperOfNode(resolveNode(astSet, sourceFile, element.source
    .range.start, element.source.range.end, true));
```

Kódrészlet 4.2. Problémás function

Ezen belül is a `resolveNode` function volt a probléma. A `resolveNode` kapott egy `ast`-t (amit a binder kapott meg paraméterbe), egy filenevet, kezdő- és végPozíciót, illetve egy `true` vagy `false` értéket, ami attól függ, hogy `sourcenodeot` vagy `targetnodeot` keresünk. Az `ast`-n `forEach`-el végigmentünk, amin belül az `astNodeon` walkoltunk addig amíg nem találtuk meg a számunkra megfelelő `nodeot`. Rosszabb esetben a legvégén volt a `node`, mivel már a végén voltunk a bindolásnak, és ebből is látható volt, hogy ez kellően sok időbe is kerülhet, ha rengeteg `nodeunk` van. Kisebb projekt esetében ez nem baj, mivel az `ast` abban az esetben nem annyira nagy. Ahol már van 100 vagy 200 ezer változó és vagy függvényhívás, ott már nagyobb a baj, mivel ezt rengetegszer meg kell ismételni. Ezután ötleteltem, hogy hogyan tudnám jobbra átírni a `resolveNodeot` functiont, vagy egy másik logika alapján megközelíteni a problémát. Végül eszembe jutott egy sokkal könnyebb megoldás erre, ami nem igényli a sokszori bejárását az `ast`nek.

Először is, létrehoztam egy `indexAST` nevű functiont, ami a következőket hajtja végre:

```
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) } } } }
    return globals.indexedAST}
```

Kódrészlet 4.3. indexAST function

A 4.3 kódrészletben látható, hogy egy ast-t várunk paraméterben. Ezt a bindertől fogja kapni, ez az egyénileg létrehozott ast. A functionben forEach-el végig megyünk az ast elemein, amik az astNodeok. Itt beállítjuk a *globals.setActualFile()* függvénnyel az aktuális filenevet. Ezután az adott astNode-on elkezdünk walkolni. Csak az enter metódusát írtam meg. Meghívunk egy függvényt ami a globalsban lett definiálva, setIndexed a neve. Ennek a függvénynek megadjuk a filevenet, a node rangenek a kezdő- és a végparaméterét (tehát ahol kezdődik az adott node és hol végződik, karakterpontosan), és magát a nodeot.

```
1 const setIndexed = function(filename, range_start, range_end, node){
2     let actualfilename = getFilePathVariable(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 function

A setIndexed először végez egy vizsgálatot arra, hogy az adott node be van már indexelve. Ezt úgy teszi meg, hogy az indexedAST tömbben keres egy indexre (Ez az index a következőképp néz ki: fileNev-StartPosition-Endposition), és még vizsgál is arra, hogy ha van ilyen index, akkor ezen van emár olyan node. Ha van akkor nem állít be semmit, csak returnol, mivel már be van indexelve az adott node. Ha nincs, akkor beállítja az indexedAST tömbnek az adott indexre az adott nodeot.

Ezáltal az adott astn csupán csak egyszer megyünk végig foreach-el és egyszer walkolunk, ekkor egy tömbbe beindexelünk minden egyes létező nodeot ami kellhet nekünk. Ha ezzel megvoltunk akkor utána kezdődhet a binder része. A 4.2 látható, hogy először megkerestük a nodeot index alapján és az-

tán hívtuk meg rá a *getWrapperOfNode()* függvényt. Ezt is megváltoztattam, írtam rá egy *getIndexed* függvényt. A *getIndex()* függvény megvizsgálja, hogy *source* vagy *targetNode*ot keresünk. Ha *targetNode*ot, akkor returnoljuk a *getWrapperOfNode(indexedAST[filename – StartPosition – Endposition])*-t. Ha *targetNode*ot keresünk, akkor walkolunk ismét, de nem az *ast*-ben, hanem már a beindexelt tömbben. Ez lényegesen gyorsabb, mint a *resolveNode* megoldás, mert ott minden egyes esetben az *ast*-n walkoltunk, itt meg csak egy beindexelt elemén a tömbnek. Ha megkaptuk a *sourceNode*ot akkor returnoljuk a *getWrapperOfNode(result)*-t.

5. fejezet

Regtest frissítés

5.1. A regressziós tesztelésről

A Sourcemeter Javascript projektben regressziós tesztelés folyik, mint tesztelési folyamat. A regressziós teszt segítségével hamar tudunk hibákat kiszűrni fejlesztés során. Ez a következőképpen zajlik a projekten:

- Cmake segítségével először legeneráljuk a megfelelő fileokat. Vcxproj illetve make fileokat, használt operációs rendszertől függően.
- Visual Studio 2017 vagy make segítségével lebuildeljük a Regtest_javascript targetet.
- A Regtest_javascript buildelése során ellenőrizzük, hogy az adott projekt le van-e már buildelve. Ha nem, akkor lebuildeljük tesztelés előtt.
- Tesztelés közben, programonként írja ki a konzolra, hogy sikeres vagy sem az adott teszt.
- Tesztelés végén egy regtest.xml fileba írja az összesített eredményeket fileokra lebontva.
- Ha valami differencia van a referenciához képest, akkor azt egy külön diff kiterjesztésű fileban jelzi a rendszer nekünk, az elvárt és a kapott eredményt beleírva.

5.2. Tesztekőről

A Regtest_javascript target több programot foglal magában, nem csak a JSAN-t. Egészen pontosan a következőket: JSAN, JSAN2Lim, LIM2Metrics, LIM2Patterns, ESLintrunner, ESLint2Graph, ChangeTracker, DuplicatedCodefinder és SourceMeter projekteknek tartalmazza a tesztjeit javascriptes (és mostmár typescriptes) fileokra. Én a JSAN, JSAN2Lim, ESLintrunner és a SourceMeter projekteknek változtattam a tesztelésén. Eddig a tesztelés úgy zajlott, hogy beadtunk inputnak pár nagyobb filet, és lefuttattuk rá a programot amit tesztelni szerettünk volna, és megnéztük az outputot. Azt kaptam feladatnak, hogy a tesztelési logikát írjam át arra, hogy kapcsolókra is teszteljünk. Ez azt jelentette, hogy minden kisebb projektnek vannak külön kapcsolói. Eddig mindig csak egy adott sorral futtattuk le a programot, és nem volt az tesztelve, hogy pl useRelativePath működik-e az elvártak szerint vagy sem.

5.3. Tesztek kibővítése

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

- -i: Jelentése input, lehet relatív vagy abszolút útvonal a filehoz vagy projekthez.
- -o: Jelentése output neve, lehet relatív vagy abszolút útvonal.
- -d: Jelentése dumpjsml, a JSAN outputját átgenerálja XML stílusú fileba és ezt egy jsml fileba kiírja.
- -e: Jelentése ExternalHardFilter, relatív vagy abszolút útvonal egy olyan filehoz, ami szövegalapú és olyan syntax található benne, ami kell az externalHardFilternek
- -help: Jelentése help, kiírja minden kapcsolóhoz tartozó descriptiont.
- -r: Jelentése useRelativePath, outputban az útvonalakat átírja relatív útvonalra.
- -h: Jelentése html, a JSAN html fileokra is lefut, bennük keresve javascriptes scripteket és azokat tesztelni.

- -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 javascriptes tesztek mellé kell majd typescriptes tesztek is keresnem. A projekteket python scriptek 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_pathban van a mappa is, és ugye 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 switchet annak függvényében, hogy igazat kaptunk e vagy sem. Utána kellett néznem, hogy egy ExternalHardFilter file hogy is néz ki, hogyan kell használni. A használata a következő: létrehozuk az externalHardFilter file-t a tesztelendő file mellé vagy a tesztelendő projekt gyökerébe, és attól függően, hogy ki akarjuk hagyni az adott file-t vagy hozzáadni, írunk egy + vagy egy – jelet a sor elejére és

utána relatív útvonal és a file neve. Alapértelmezetten minden filet leelemez az adott program, JSAN esetében ezek azok a fileok amiknek a kiterjesztése js,jsx. (külön kapcsolóval megadhatjuk neki, hogy a html kiterjesztésű fileokat elemezze-e vagy se.) Typescriptes bővítés után már a ts és a tsx kiterjesztésű fileokat is elemzi. Ezért írtam több tesztet is. Egy példa, hogy hogyan néz ki ez a file:

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

Kódrészlet 5.3. ExternalHardFilter file

Az 5.3 kódrészletben látható, hogy filtered01 02 és 03 at kivettük, hogy azokat ne elemezze a jsan. Ezután visszavettük a filtered01et. A filtered fileok azok javascriptes fileok, javascriptes kóddal. Ezután referenciába csak a filtered01 file outputját raktam be, hiszen a 02 és 03at nem elemzi, ha elemezné, akkor szólna a program, hogy missing reference file. Természetesen lehet regexpet is használni filterezésnél.

Ezután teszteltem a *-i* kapcsolót, itt ugye a programnak vissza kellene adnia, hogy üres az input ha nincs megadva. Ezután a *-o* kapcsolóra teszteltem, itt ebben az esetben default értékben *out.jssi* filet kellene visszaadnia. Végül a *-h* kapcsolót néztem meg, itt ha megvan adva ez a kapcsoló, akkor az adott projektben a html fileokban a javascriptet 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 typescriptes supportal és a JSAN2Limet átírtam, hogy a JSAN általi typescript elemzéseket jól olvassa be, ideje volt teszteket keresni, mind JSAN-nak és mind JSAN2Limnek. Kerestem egyszerűbb typescript és picit összetettebb typescriptes projekteket is, hogy lássam a hiányosságokat. Természetesen ami outputokat adtak a programok, azokat át kellett néznom egyesével, hiszen csak így tudom meg, hogy jól tesztelte-e a megadott filet vagy sem. Több hiányosságot is észrevettem, mind JSAN

oldalról, mind JSAN2Lim oldalról, ezek kisebb hiányosságok voltak. Például, hogy egy nodenak 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 az Ast binder közel tízszer gyorsabban fusson le nagyobb projektekre, mint ezelőtt. Leteszteltem ugyanarra a nagy projektekre az eredeti JSAN lefutását és az én általam átírt lefutását. Ezt A 6.1 kódrészleten láthatjuk. Kisebb projektekre körülbelül ötszörös gyorsulás van. Lefuttattam egy nagyobb projektekre a jsan-t, az optimalizált és az optimalizálatlan verzióval, hogy jobban lássuk a különbséget. A következő eredmények születtek:

```
1 Optimalizalt_verzio:  
2 Binding VU: 71599/71599  
3 VU binding: 1:39.140 (m:ss.mmm)  
4  
5 Optimalizalatlan_verzio:  
6 Binding VU: 71599/71599  
7 VU binding: 42:31.211 (m:ss.mmm)
```

Kódrészlet 6.1. JSAN lefutási idő előtte és utána

Ugyanazt az outputot adja mind a kettő program, szóval csak gyorsaságban és memóriahasználatban változott sokat.

Emellett a JavaScriptSchema újraírása is sikeres volt, emiatt a JSAN tud typescriptes kódokat elemezni. JSAN2Limet sikeresen módosítottam, hogy a JSAN által kiadott outputot sikeresen limmé alakítsa.

6.2. Jövőbeli tervek

JavaScriptSchema bővítése sok időbe telt, mivel se dokumentáció, se tapasztalat nem volt. Ezután még a JSAN2Lim átírás is sok időbe került. Eközben a typescript-eslint github repo amit használtunk a bővítésre, frissült, sok új funkciót hoztak be. Rengeteg változtatás volt a typescript részénél, de a schémát nem tudtuk még naprakészre hozni, mivel voltak ennél fontosabb feladatok. Egyik jövőbeli terv az, hogy a schémát up to date-re hozzam, mivel már van hozzá dokumentáció is, meg nagyjából én is írtam, ezért nem lesz ez annyi idő, mint volt az elején az újraírása.

Végül a JSAN programra még bőven ráfér az optimalizálás, mivel csak az ast bindert optimalizáltuk, rengeteg helyen feleslegesen van bejárva az ast. Ez a másik jövőbeli terv.

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 8.

.....
aláírás

Irodalomjegyzék

[1] Typescript-eslint github repository. <https://github.com/typescript-eslint/typescript-eslint/tree/main/packages/ast-spec/src>.