# LAPORAN TUGAS BESAR IF2124 TEORI BAHASA FORMAL DAN AUTOMATA

# PARSER JAVASCRIPT (NODE.JS)



#### **DISUSUN OLEH:**

KELOMPOK 三位棉兰男神

13521046 Jeffrey Chow

13521054 Wilson Tansil

13521102 Jimly Firdaus

# PROGRAM STUDI TEKNIK INFORMATIKA SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA INSTITUT TEKNOLOGI BANDUNG

# Daftar Isi

BAB I	3
BAB II	4
2.1 Finite State Automata (FSA)	4
2.1 Context Free Grammar (CFG)	5
BAB III	7
3.1 Finite State Automata	7
3.2 Context Free Grammar	8
BAB IV	10
4.1 Module file lexer.py	10
4.2 Module file FA.py	10
4.3 Module file CYK.py	10
4.4 Module file CFG.py	10
4.5 File main.py	11
BAB V	12
BAB VI	16
BAB VII	16
BAB VIII	16

# BAB I Latar Belakang

Dalam proses pembuatan program dari sebuah bahasa menjadi instruksi yang dapat dieksekusi oleh mesin, terdapat pemeriksaan sintaks bahasa atau parsing yang dibuat oleh programmer untuk memastikan program dapat dieksekusi tanpa menghasilkan error. Parsing ini bertujuan untuk memastikan instruksi yang dibuat oleh programmer mengikuti aturan yang sudah ditentukan oleh bahasa tersebut. Baik bahasa berjenis interpreter maupun compiler, keduanya pasti melakukan pemeriksaan sintaks. Perbedaannya terletak pada apa yang dilakukan setelah proses pemeriksaan (kompilasi/compile) tersebut selesai dilakukan.

Dibutuhkan grammar bahasa dan algoritma parser untuk melakukan parsing. Sudah sangat banyak grammar dan algoritma yang dikembangkan untuk menghasilkan compiler dengan performa yang tinggi. Terdapat CFG, CNF-e, CNF+e, 2NF, 2LF, dll untuk grammar yang dapat digunakan, dan terdapat LL(0), LL(1), CYK, Earley's Algorithm, LALR, GLR, Shift-reduce, SLR, LR(1), dll untuk algoritma yang dapat digunakan untuk melakukan parsing.

Pada program parsing ini, kami akan menggunakan grammar CFG yang akan di convert menjadi CNF serta algoritma CYK untuk melakukan parsing.

### BAB II Teori Dasar

#### 2.1 Finite State Automata (FSA)

Finite State Automata (FSA) berupa sistem model matematika dengan masukan dan keluaran diskrit yang dapat mengenali bahasa paling sederhana (bahasa reguler) dan dapat diimplementasikan secara nyata.

FSA dapat menerima input dan mengeluarkan output yang memiliki state yang berhingga banyaknya. Selain itu, FSA memiliki sekumpulan aturan untuk berpindah dari satu state ke state lain berdasarkan input dan fungsi transisi yang diterapkan.

Finite State Automata pada dasarnya digunakan untuk mengenali pola. Dibutuhkan string simbol sebagai input dan statusnya berubah sesuai dengan input tersebut. Ketika ditemukan simbol input yang diinginkan, maka terjadi transisi.

Pada saat transisi, automata dapat berpindah ke keadaan berikutnya atau tetap dalam keadaan yang sama. Finite State Automata memiliki dua status, status Terima atau status Tolak. Ketika string input berhasil diproses, dan automata mencapai state akhir, maka statusnya adalah Terima, demikian juga sebaliknya.

Finite State Automata dapat didefinisikan dengan persamaan berikut:

$$M = (Q, \Sigma, \delta, S, F)$$

Q = himpunan state

 $\Sigma$  = himpunan simbol input

 $\delta$  = fungsi transisi  $\delta$  :  $Q \times \Sigma$ 

 $S = \text{state awal / initial state}, S \in Q$ 

 $F = \text{state akhir}, F \subseteq Q$ 

Finite State Automata memiliki beberapa karakteristik berikut:

• Setiap Finite Automata memiliki keadaan dan transisi yang terbatas.

- Transisi dari satu keadaan ke keadaan lainnya dapat bersifat deterministik atau nondeterministik.
  - Setiap Finite Automata selalu memiliki keadaan awal.
  - Finite Automata dapat memiliki lebih dari satu keadaan akhir.
- Jika setelah pemrosesan seluruh string, keadaan akhir dicapai, artinya otomata menerima string tersebut.

#### 2.1 Context Free Grammar (CFG)

Context Free Grammar (CFG) / Bahasa Bebas Konteks adalah sebuah tata bahasa dimana tidak terdapat pembatasan pada hasil produksinya, Contoh Pada aturan produksi :

$$\alpha \rightarrow \beta$$

batasannya hanyalah ruas kiri  $(\alpha)$  adalah sebuah simbol variabel. Sedangkan contoh aturan produksi yang termasuk CFG adalah seperti di bawah :

$$B \rightarrow CDeFg$$

$$D \rightarrow BcDe$$

Context Free Grammar (CFG) adalah tata bahasa yang mempunyai tujuan sama seperti halnya tata bahasa regular yaitu merupakan suatu cara untuk menunjukkan bagaimana menghasilkan suatu untai-untai dalam sebuah bahasa.

Context Free Grammar (CFG) menjadi dasar dalam pembentukan suatu parser/proses analisis sintaksis. Bagian sintaks dalam suatu kompilator kebanyakan di definisikan dalam tata bahasa bebas konteks. Pohon penurunan (derivation tree/parse tree) berguna untuk menggambarkan simbol-simbol variabel menjadi simbol-simbol terminal setiap simbol variabel akan di turunkan menjadi terminal sampai tidak ada yang belum tergantikan.

Proses penurunan / parsing bisa dilakukan dengan cara sebagai berikut :

- Penurunan terkiri (leftmost derivation): simbol variabel terkiri yang di perluas terlebih dahulu.
- Penurunan terkanan (rightmost derivation) : simbol variabel terkanan yang diperluas terlebih dahulu.

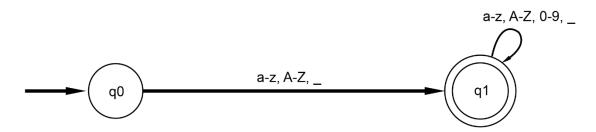
#### 2.3 Syntax JavaScript dalam pembuatan CFG dan FA

while	break	default	for	return	var
const	delete	function	switch	while	case
else	if	throw	catch	false	let
try	continue	finally	null	true	import
export	else if	export default	in	from	instanceof
as					

Pada syntax Javascript ini, titik koma merupakan hal yang wajib dalam membatasi sebuah command program. Hal ini diimplementasikan agar program Javascript dapat memiliki pembatasan yang signifikan guna membantu dalam hal debugging dan penulisan program. Keterbatasan grammar juga terletak pada penggunaan atau pemanggilan struct, fungsi, atau library pada sebuah operasi aritmatika, misalnya (var1/var2).tofixed(x). Setiap elemen yang berada di dalam {} dan [] dianggap sebagai list dan array sehingga syntax tersebut haruslah {x1, x2, x3, ..., xn} atau [x1, x2, x3, ..., xn]. Dalam hal ini, x adalah kumpulan kemungkinan yang dapat diinput ke dalam baik merupakan sebuah expression maupun variabel. Pada tahap pengisian kurung {} dan [], elemen paling akhir tidak dapat diakhiri dengan koma. Contohnya program akan menyatakan {x1, x2, x3,} salah dikarenakan terdapat koma pada elemen akhir, sebaliknya program akan menyatakan {x1, x2, x3} benar dikarenakan tidak terdapat koma pada elemen terakhir. Syntax import harus selalu terletak pada bagian atas, sebagai header.

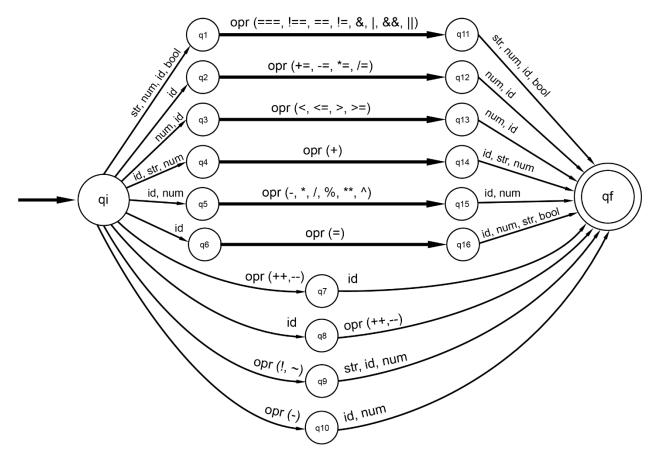
# BAB III Finite Automata dan Context Free Grammar

#### 3.1 Finite State Automata



Gambar 3.1.1 DFA untuk pengecekan variabel

Program dapat melakukan pengecekan variabel dengan prinsip Finite State Automata. Contoh variabel yang tidak diterima adalah 123cyk, 456FA\_, atau 23\_abc. Contoh variabel yang diterima adalah TBFO\_Mantap, tubesTbfo, atau pembuatProgramIniGanteng. DFA ini diimplementasi pada file FA.py modul checkVar.



Gambar 3.1.2 NFA untuk pengecekan ekspresi

Program dapat melakukan pengecekan ekspresi dengan prinsip Finite State Automata. Contoh ekspresi yang tidak diterima adalah a = b = 2 + atau "tbfo"\*\*2 . Contoh ekspresi yang diterima adalah nama = "teori" + "bahasa", <math>a < b, atau i++. NFA ini diimplementasi pada file FA.py modul checkExpr.

#### **3.2 Context Free Grammar**

$$G = (V, T, P, S)$$

# Non Terminal Symbol (V)

S	BODY	IMPORT_STM T	EXPORT_STM T	EXPORT_DEF AULT_STMT	PART
VALUE_BOOL EAN	EXPR_BOOLE AN	VARIOUS_EX PR_BOOLEAN	LOGOP	ISSTRICTEQ	NOTSTRICTE Q
ISEQ	ISNEQ	LEQ	L	GEQ	G
LOGOPB	AND	OR	NOTB	IF_STMT	ELSE_STMT
ELSEIF_STMT	IF	ELSE	WHILE_STMT	FUNCTION_S TMT	RETURN_STM T
FOR_STMT	DELETE_STM T	THROW_STM T	THROW	SWITCH_STM T	SWITCH_BOD Y
SWITCH_PAR AM	CASE_STMT	CASE_BODY	DEFAULT_ST MT	SWITCH	CASE
DELETE	FOR	FOR_PARAM	TYPE_FOR	IN	OF
INSTANCEOF	RETURN	CONTINUE_S TMT	CONTINUE	BREAK_STMT	BREAK
WHILE	TRY_STMT	CATCH_PARA M	TRY	CATCH	FINALLY
TITIK_STMT	TITIK_VALUE	ARRAY_CALL	FUNCTION	FUNCTION_C ALL	FUNCTION_P ARAM
VARIOUS_VA RIABLE_FUN CTION	TYPE_STMT	BODY_OBJEC T	BODY_ARRA Y	VARIOUS_VA RIABLE_BOD Y	VALUE
ASSIGN_VARI ABLE	ASSIGNMENT	NULL	TD	PMEQ_EXPR	PMEQ_VALUE
PMEQ	PEQ	MEQ	MULEQ	DIVEQ	INCDEC
INC	DEC	EXPR_ARIT	EXPR_ARIT_V ALUE	ARIT_OPERA TOR	POW
XOR	ANDB	ORB	SUM	SUB	DIV
MOD	HEAD_OBJEC T	VARIOUS_VA RIABLE	HEAD_VALUE	DEFAULT	EXPORT
IMPORT	KBKI	KBKA	KKKA	KKKI	MUL
FROM	STRING	AS	ID	TK	KM
LET	VAR	CONST	BOOL	TRUE	FALSE
NUMBER	IS	KSKI	KSKA	TITIK	TENARY_STM T

# Terminal Symbol (T)

isstricteq	notstricteq	iseq	isneq	leq	1
geq	g	and	or	notb	if
else	return_stmt	throw	default_stmt	switch	case
delete	for	in	of	instanceof	return

continue_stmt	continue	break_stmt	break	while	catch_param
try	catch	finally	function	function_param	null
pmeq_expr	peq	meq	muleq	diveq	inc
dec	pow	xor	andb	orb	sum
sub	div	mod	head_object	default	export
import	kbki	kbka	kkka	kkki	mul
from	string	as	id	tk	km
let	var	const	true	false	number
is	kski	kska	titik		

#### **Production (P)**

Terdapat 337 baris production yang kami implementasikan pada CFG.

#### Start Symbol (S)

Pada CFG, start symbol yang kami gunakan adalah S. Syntax Javascript akan dianggap benar apabila grammar bisa dicapai dari S.

#### **Chomsky Normal Form (CNF)**

Pada program ini, Context Free Grammar (CFG) yang telah dibuat akan diubah ke dalam bentuk Chomsky Normal Form (CNF) yang menjadi prasyarat penggunaan algoritma Cocke-Younger-Kasami dalam pengecekan syntax. Program pengubahan CFG ke CNF diimplementasikan dalam file CFG.py modul CFGtoCNF.

# BAB IV Implementasi

#### 4.1 Module file lexer.py

No	Fungsi/Prosedur	Deskripsi
1	lexer	Mem- <i>parsing</i> semua <i>code</i> JavaScript ke dalam token-token yang sudah di- <i>define</i> agar bisa dibaca dan diproses dengan CFG yang dibuat.
2	parseToToken	Mem- <i>parse</i> setiap <i>code by code</i> JavaScript ke dalam token- token (terminal) untuk di- <i>compare</i> dengan <i>grammar</i> CFG yang dibuat.

#### 4.2 Module file FA.py

Modul ini mengimplementasikan Finite Automata untuk melakukan pengecekan variabel dan ekspresi pada *syntax javascript*.

No	Fungsi/Prosedur	Deskripsi
1	checkVar	Melakukan pengecekan nama variabel agar sesuai dengan kaidah penulisan variabel. Mengirimkan <i>True</i> jika penamaan variabel benar, vice versa.
2	checkExpr	Melakukan pengecekan ekspresi pada <i>syntax javascript</i> . Mengirimkan <i>True</i> jika ekspresi benar dan mengirimkan indeks token ekspresi yang salah jika ekspresi salah.

# 4.3 Module file CYK.py

No	o Fungsi/Prosedur	Deskripsi
1	cykParse	Melakukan parsing token syntax javascript terhadap CNF yang telah dibuat. Bernilai True jika token bisa dicapai dari start symbol.

#### 4.4 Module file CFG.py

Modul ini diimplementasikan dalam bentuk *Object Oriented Programming* (OOP) untuk memudahkan pengubahan CFG. Penulisan dalam OOP juga membuat program lebih rapi dan terstruktur. *Class* CFG ini diawali dengan memanggil fungsi parseCFG lalu akan memanggil fungsi CFGtoCNF dan diakhiri dengan mengembalikan CFG yang sudah berbentuk CNF.

No	Fungsi/Prosedur	Deskripsi	
1	updateKeyVal	Mengupdate Key dan Value dari dictionary CFG	
2	createProduction	Membuat production baru untuk ditambahkan ke CFG	
3	newVar	Membuat nama variabel baru yang digunakan dalam memproduksi rules CNF	
4	CFGtoCNF	Mengubah CFG menjadi CNF	
5	parseCFG	Melakukan parsing CFG dari textfile	

#### 4.5 File main.py

Program utama dari *parser javascript* ini diimplementasikan pada file main.py. File ini menggunakan modul-modul yang telah dibuat di berbagai file yang sudah dijelaskan di poin 4.1-4.4. Program dimulai dengan melakukan parsing terhadap argumen yang merupakan nama file javascript yang akan di cek. Selanjutnya program akan menampilkan splash screen program. Program akan melakukan parsing terhadap file grammar.txt terlebih dahulu sampai terbentuknya CNF. Lalu program akan melakukan parsing terhadap file javascript yang akan dicek sehingga berbentuk seperti token. Kemudian token dan CNF akan dimasukkan ke algoritma CYK dan akan mengembalikan nilai berupa boolean. Jika True maka akan ditampilkan pesan "Compile Success!!!", dan jika False akan ditampilkan pesan "Compile Error...". Jika error sudah ditemukan disaat parsing kode javascript menggunakan lexer dan FA, maka akan menampilkan pesan "Invalid Syntax". Diakhir program akan ditampilkan waktu eksekusi pengecekan syntax.

BAB V Eksperimen

Kata Kunci Javascript	Syntax	Result
While Loop	<pre>while (condition) {   console.log(i);   i++;   if (i == 1) {     break;   } else {     continue;   } }</pre>	JAVASCRIPT SYNTAX LINTER 三位棉兰男神
For Loop	<pre>for (let i = 0; i &lt; 5; i++) {   for (let j = 0; j &lt; i; j++) {     console.log(i + j); } } </pre>	JAVASCRIPT SYNTAX LINTER 三位棉兰男神
Conditional	<pre>const sum = 10; if (sum &gt; 30) {     if (sum &gt; 25 &amp;&amp; sum &lt; 28) {         console.log(1);     } else {         console.log(2);     } } else if (sum &gt; 15) {     if (sum &gt; 18 &amp;&amp; sum &lt; 20) {         console.log(3);     } else if (sum &gt; 20 &amp;&amp; sum &lt; 25) {         console.log(4);     } } else if (sum &gt; 5) {         console.log(sum);     } }  // console.log(sum); // console.log("failed"); //</pre>	The secution time: 27.05 seconds

```
JAVASCRIPT SYNTAX LINTER
                                                                       三位棉兰男神
                           const temp = "jimly";
                           switch (temp) {
                             case "jimly":
                                                               Checking test/tc4.js...
Switch, Case,
                              console.log("hello");
                                                               Please wait for a moment...
   Default
                             case "firdaus":
                              console.log("world");
                                                                Compile Success!!!
                               break;
                                                               Execution time : 0.84 seconds
                                                                  JAVASCRIPT SYNTAX LINTER
                                                                      三位棉兰男神
 Try, Catch,
   Finally,
                                                                Checking test/tc5.js...
                          finally {
console.log("you have passed the last stage...");
                                                                Please wait for a moment...
   Export,
  Function,
                          let a = b + c;
console.log(a);
   Default
                                                                 Compile Success!!!
                                                                Execution time : 6.42 seconds
                                                                 JAVASCRIPT SYNTAX LINTER
三位棉兰男神
                           arr = ["a", "b", "c"];
                                                               Checking test/tc6.js...
                           delete arr[0];
Array, Delete
                                                               Please wait for a moment...
                           arr = ["e", "f", "g"];
                           delete arr[2];
                                                                 Compile Success!!!
                                                               Execution time : 0.64 seconds
                                                                  JAVASCRIPT SYNTAX LINTER
三位棉兰男神
                          let a = true;
  Let, var,
const, throw,
ternary, true,
  false, null
                                                                  Compile Success!!!
                          throw "this is an error...";
```

```
JAVASCRIPT SYNTAX LINTER
三位棉兰男神
    Simple
 program 1
                                                                                    Please wait for a moment...
   (import)
                                                                                        JAVASCRIPT SYNTAX LINTER
三位棉兰男神
    Simple
  program 2
                                  • • •
                                                                                        JAVASCRIPT SYNTAX LINTER
三位棉兰男神
                                                                                     Please wait for a moment...
inputAcc.js
                                                                                        Compile Success!!!
                                          return 4;
} else {
return "Momen";
                              • • •
                                                                                        JAVASCRIPT SYNTAX LINTER
                                     // This is a sample multiline comment if (x == 0) {
                                     } else if x + 4 == 1 {
  if (true) {
    return 3;
InputReject.j
```

```
三位棉兰男神
syntaxFail.js
                      let `a = 3;
                                                              三位棉兰男神
                      if (x < ){
 exprFail.js
                        x = true;
                                                     JAVASCRIPT SYNTAX LINTER
三位棉兰男神
  Simple
 Program 3
                   for (let i=0, i < a, i++){
  (failed)
                        console.log(i)
                                                     JAVASCRIPT SYNTAX LINTER
三位棉兰男神
                      while (i){
                           console.log(i);
  Simple
                                                    Checking test/will_fail3.js...
                                                    Please wait for a moment...
 Program 4
                              break;
  (failed)
                                                     Compile Error...
                          i++;
```

# BAB VI Github Repository

https://github.com/JeffreyChow19/TubesTBFO.git

# BAB VII Pembagian Tugas

PIC (Person in Charge)	Job Description
13521046 Jeffrey Chow	Grammar parser, CFG to CNF, CYK, main,
	FA (CFG.py, CYK.py, FA.py, main.py)
13521054 Wilson Tansil	Grammar and Lexer (grammar.txt, lexer.py)
13521102 Jimly Firdaus	Grammar, Lexer, FA (grammar.txt, lexer.py,
	FA.py)

# BAB VIII Referensi

https://www.geeksforgeeks.org/cocke-younger-kasami-cyk-algorithm/

https://www.geeksforgeeks.org/converting-context-free-grammar-chomsky-normal-form/

https://gomakethings.com/javascript-

 $\underline{linters/\#:\sim:text=JavaScript\%20linters\%20are\%20tools\%20that, can\%20use\%20to\%20fix\%20thin}\underline{gs}.$ 

https://fairuzelsaid.wordpress.com/2011/06/16/tbo-context-free-grammar-cfg/

https://www.trivusi.web.id/2022/08/finite-state-automata.html