

Yosys Reverse Engineering

Raphael Biermann

23. August 2021

Inhaltsverzeichnis

1	Lexer	3
1.1	Definiton	3
1.2	Yosys und <i>Flex</i>	3
1.3	Analyse des Lexer-Quellcodes	4
1.3.1	Lexerregeln	4
1.3.2	Zusammenfassung der Lexerregeln	16
1.3.3	Interpretationsprobleme und Ansätze	17
1.3.4	Verbindung zum Parser	17
2	Parser	20
2.1	Analyse des Parser-Quellcodes	20
2.1.1	Analyse der Parser-Grammatik	20
2.1.2	Interpretations-Probleme	28
3	AST-Frontend	29
3.1	Analyse der simplify Funktion	30
3.2	Analyse der genRTLIL Funktion	31
3.3	Prozessgenerator	32
3.3.1	Datenstrukturen und Variablen	34
3.3.2	Non-Blocking Assignments	38
3.3.3	Blocking-Assignments	41
3.3.4	Cases und if-Anweisungen	41
3.3.5	Proc Pass	44

Disclaimer

Dieses Dokument hat keinen Anspruch auf Vollständigkeit und Richtigkeit. Die Grundlagen für die Inhalte entspringen lediglich den Erkenntnissen der Analyse des Yosys Open Source Syntheseprogramms und wurden nicht verifiziert.

Kapitel 1

Lexer

Das Verilog Frontend besteht aus *Preprozessor*, *Lexer* und *Parser*, die den Verilogcode lesen und in seine Bestandteile zerlegen.

1.1 Definiton

Der Lexer (oft auch Scanner oder Tokenizer genannt) ist das erste Glied der Kette. Er liest den Verilogcode und generiert aus den Wörtern und Zeichen kontextunabhängig sogenannte Tokens. Tokens werden als kleinstmögliche Einheit von Text mit semantischer Bedeutung definiert. Damit der Lexer Tokens generieren kann, müssen Regeln festgelegt werden, die Wörter und Zeichen in diese Tokens, die auch wie Kategorien verstanden werden können, einteilen. Diese Regeln liegen in Form von regulären Ausdrücken vor. Der Lexer wird oft mit einem deterministischen Automaten realisiert, da mit diesem leicht reguläre Ausdrücke implementiert werden können.

Um auf eine komplexe manuelle Implementierung zu verzichten, kann von Lexergeneratoren gebrauch gemacht werden, die eine lexikalische Beschreibung in einem kompillierbaren Code implementieren.

Populäre Tools sind *Lex* und *Flex*.

1.2 Yosys und *Flex*

Für das Yosys-Frontend wurde der Lexergenerator *Flex* benutzt. *Flex* benötigt genau wie *Lex* als Eingabe eine *.l* oder *.lex* Datei, die Syntax und Semantik der Eingabe beschreibt.

Eine *.l* oder *.lex* Datei ist folgendermaßen aufgebaut:

```
1 %{  
2 Declarations  
3 %}  
4 Definitions  
5 %%  
6 Rules  
7 %%  
8 Subroutines
```

Die *Declarations* Sektion besteht aus C-Code, der ohne weiteres von *Flex* in die Ausgabedatei kopiert wird. Dort können beispielsweise Variablen definiert werden, die für die Behandlung der Eingabe verwendet werden.

In der *Definitions* Sektion werden Optionen für *Flex* vorgegeben. Oft werden einige Oberbegriffe wie *digit* für den regulären Ausdruck `[0-9]` festgelegt, damit die weitere Spezifizierung einfacher wird.

Die *Rules* Sektion besteht aus Regeln, die in Form von regulären Ausdrücken vorgegeben werden. *Flex* versteht folgende reguläre Ausdrücke:

```

1 x      the character x
2 "x"    an x, even if x is an operator.
3 \x     an x, even if x is an operator.
4 [xy]   the character x or y.
5 [x-z]  the characters x, y or z.
6 [^x]   any character but x.
7 .      any character but newline
8 ^x     an x at the beginning of a line.
9 <y>x    an x when Flex is in start condition y.
10 x$    an x at the end of a line.
11 x?    an optional x.
12 x*    0,1,2, ... instances of x.
13 x+    1,2,3, ... instances of x.
14 x|y   an x or a y.
15 (x)   an x.
16 x/y   an x but only if followed by y.
17 {xx}  the translation of xx from the definitions section.
18 x{m,n} m through n occurrences of x

```

Nach dem regulären Ausdruck folgt C-Code, der als Reaktion auf ein nach den Regeln erkanntes Wort ausgeführt werden soll.

In der *Subroutines* Sektion folgt Code vom Benutzer wie beispielsweise der Aufruf des Lexers mit *yylex()*.

1.3 Analyse des Lexer-Quellcodes

Laut dem Yosys-Manual identifiziert der Lexer einerseits die Wörter und Zeichen, die das Verilog-Frontend identifiziert, erkennt und andererseits die aktuelle Position im Verilog-Code mittels globaler Variablen. Diese werden dem Konstruktor der AST-Knoten, die im Zuge der lexikalischen Analyse erstellt werden, übergeben. Weiterhin erkennt der Lexer spezielle Kommentare beispielsweise für Synopsys und verarbeitet diese entsprechend.

Um den Lexer-Quellcode, der sich in der *verilog_lexer.l* Datei unter *Frontends/Verilog* befindet, zu analysieren, werden die zuvor angedeuteten Scannerfunktionen im Quellcode anhand der Regeln aus regulären Ausdrücken identifiziert.

Das Online-Tool *RegErr* hilft bei der Analyse der regulären Ausdrücke. Es ist unter dem Link <https://regerr.com> zu finden.

Folglich werden die regulären Ausdrücke aufgeführt und deren Bedeutung und Funktion beschrieben:

1.3.1 Lexerregeln

- `<INITIAL,SYNOPSIS_TRANSLATE_OFF>"'file_push "[^\\n]*`

Wenn *Flex* in der Startbedingung `INITIAL,SYNOPSIS_TRANSLATE_OFF` ist, 'file_push folgt und danach jedes Zeichen außer eine neue Zeile folgt, dann..

```

1 fn_stack.push_back(current_filename);

```

```

2 ln_stack.push_back(frontend_verilog_yyget_lineno());
3 current_filename = yytext+11;
4 if (!current_filename.empty() && current_filename.front() == '"')
5     current_filename = current_filename.substr(1);
6 if (!current_filename.empty() && current_filename.back() == '"')
7     current_filename=
8     current_filename.substr(0,current_filename.size()-1);
9 frontend_verilog_yyset_lineno(0);
10 yylloc->first_line = yylloc->last_line = 0;
11 real_location.first_line = real_location.last_line = 0;

```

- weitere Synopsys Terme mit Stacks

```

1 "timescale" [ \t]+ [^ \t\r\n/]+ [ \t]* "/" [ \t]* [^ \t\r\n]*
2 /* ignore timescale directive */

```

Timescale beschreibt die Zeiteinheit und Zeitpräzision des folgenden Verilog-Moduls. Yosys hat keinen Timing-Support. Daher wird die gesamte Anweisung vom Lexer gegen einen leeren Ausdruck getauscht, indem er den ganzen Ausdruck liest, aber keine Reaktion als C-Code folgt. „timescale“ beschreibt das Wort, [\t]+ steht für ein Leerzeichen oder einen Tabulator der ein- oder mehrmals folgt. [^ \t\r\n/]+ steht für alle weiteren Eingaben, die kein Leerzeichen, Tabulator, Return, Zeilenumbruch oder / und deren Wiederholungen sind.

```

1 "celldefine" [^\n]* /* ignore 'celldefine */
2 "endcelldefine" [^\n]* /* ignore 'endcelldefine */

```

Celldefine und alles Nachfolgende, was kein Zeilenumbruch ist, wird durch das Ersetzen mit einem leeren Ausdruck ignoriert. Celldefine ist eine Compileranweisung, die das Modul als Zelle markiert, die Yosys nicht unterstützt.

```

1 "protect" [^\n]* /* ignore 'protect */
2 "endprotect" [^\n]* /* ignore 'endprotect */

```

Keine Unterstützung für protected Verilog.

```

1 "[" [a-zA-Z_$][a-zA-Z0-9_$]* {
2     frontend_verilog_yyerror("Unimplemented compiler
3     directive or undefined macro %s.", yytext);
4 }

```

Mit diesem Ausdruck, der mit dem Zeichen ‘ beginnt, das Compileranweisungen andeutet, werden alle weiteren Anweisungen in jeglicher Form erkannt. Da Yosys diese nicht unterstützt, wird das Bison Error-Handling für Parsererrors mit der Funktion *frontend_verilog_yyerror(char const *fmt, ...)* genutzt, die in *verilog_frontend.cc* deklariert ist. Die Variable *yytext* gibt jeweils den aktuellen Verilogcode für die Fehlerrückverfolgung aus.

```

1 "module"      { return TOK_MODULE; }
2 "endmodule"   { return TOK_ENDMODULE; }
3 "function"    { return TOK_FUNCTION; }
4 "endfunction" { return TOK_ENDFUNCTION; }
5 "task"        { return TOK_TASK; }
6 "endtask"     { return TOK_ENDTASK; }
7 "specify"     { return specify_mode ?
8 TOK_SPECIFY : TOK_IGNORED_SPECIFY; }
9 "endspecify"  { return TOK_ENDSPECIFY; }

```

```

10 "specparam"    { return TOK_SPECPARAM; }
11 "package"     { SV_KEYWORD(TOK_PACKAGE); }
12 "endpackage"  { SV_KEYWORD(TOK_ENDPACKAGE); }
13 "interface"   { SV_KEYWORD(TOK_INTERFACE); }
14 "endinterface" { SV_KEYWORD(TOK_ENDINTERFACE); }
15 "modport"     { SV_KEYWORD(TOK_MODPORT); }
16 "parameter"   { return TOK_PARAMETER; }
17 "localparam"  { return TOK_LOCALPARAM; }
18 "defparam"    { return TOK_DEFPARAM; }
19 "assign"      { return TOK_ASSIGN; }
20 "always"      { return TOK_ALWAYS; }
21 "initial"     { return TOK_INITIAL; }
22 "begin"       { return TOK_BEGIN; }
23 "end"         { return TOK_END; }
24 "if"          { return TOK_IF; }
25 "else"        { return TOK_ELSE; }
26 "for"         { return TOK_FOR; }
27 "posedge"     { return TOK_POSEDGE; }
28 "negedge"     { return TOK_NEGEDGE; }
29 "or"          { return TOK_OR; }
30 "case"        { return TOK_CASE; }
31 "casex"       { return TOK_CASEX; }
32 "casez"       { return TOK_CASEZ; }
33 "endcase"     { return TOK_ENDCASE; }
34 "default"     { return TOK_DEFAULT; }
35 "generate"    { return TOK_GENERATE; }
36 "endgenerate" { return TOK_ENDGENERATE; }
37 "while"       { return TOK_WHILE; }
38 "repeat"      { return TOK_REPEAT; }
39 "automatic"   { return TOK_AUTOMATIC; }
40
41 "unique"       { SV_KEYWORD(TOK_UNIQUE); }
42 "unique0"      { SV_KEYWORD(TOK_UNIQUE0); }
43 "priority"     { SV_KEYWORD(TOK_PRIORITY); }
44
45 "always_comb"  { SV_KEYWORD(TOK_ALWAYS_COMB); }
46 "always_ff"    { SV_KEYWORD(TOK_ALWAYS_FF); }
47 "always_latch" { SV_KEYWORD(TOK_ALWAYS_LATCH); }

```

In dem C-Code für diese Verilog Direktive werden entweder Tokens direkt zurück gegeben oder der Funktion *SV_KEYWORD()* übergeben. Diese Funktion ist ein C-Makro, der sich in der zuvor angesprochenen *Declarations* Sektion befindet:

```

1 #define SV_KEYWORD(_tok) \
2 if (sv_mode) return _tok; \
3 log("Lexer warning: The SystemVerilog keyword '%s' \
4 (at %s:%d) is not " \
5 "recognized unless read_verilog is called with -sv!\n", yytext, \
6 AST::current_filename.c_str(), frontend_verilog_yyget_lineno()); \
7 yylval->string = new std::string(std::string("\\") + yytext); \
8 return TOK_ID;

```

Es handelt sich bei den Tokens, die hier übergeben werden, also um SystemVerilog Schlüsselwörter, die nur verarbeitet werden können, wenn der SystemVerilog Modus beim Einlesen der Verilogdatei im Yosys Terminal mit *read -sv verilogdatei.v* eingeschaltet wurde. Wenn der Modus aktiv ist, wird ein Token zurückgegeben. Ansonsten erfolgt Error-logging mit aktuellem Code, Dateinamen und Codeposition. Der Integer *yylval* speichert den Wert eines eingelesenen Ausdrucks. Dieser kann aus dem String *yytext* mit einer Typumwandlung gewonnen werden. Diese beiden Variablen sind in *Flex* implementiert und

werden später für den Parser benötigt.

Es gibt also zwei unterschiedliche Return-Typen; *TOK_%Direktiv%* und *TOK_ID*, deren Bedeutung analysiert werden muss.

Laut Yosys-Manual werden die Tokens dem AST Konstruktor übergeben. An anderer Stelle muss die Lexerfunktion aufgerufen werden und die Rückgabe verarbeitet werden.

Die originale *Flex* Funktion für den Aufruf des Lexers auf eine Eingabe ist *yylex()*. Diese hat als return-Wert einen Integer für den Token. Bei der Ausführung des Lexers werden also Ganzzahlen zurückgegeben. Da die Integervariablen für die Tokens (beispielsweise *TOK_MODULE*) nicht in der .lex Datei selbst definiert sind, muss auf die jeweiligen Header zurückgegriffen werden. In der *verilog_parser.tab.hh* Header Datei befindet sich eine Liste mit allen Tokenvariablen und deren jeweiligen Integerwert. Hier ein Ausschnitt:

```
1 TOK_STRING = 258,          /* TOK_STRING */
2 TOK_ID = 259,             /* TOK_ID */
3 TOK_CONSTVAL = 260,       /* TOK_CONSTVAL */
4 TOK_REALVAL = 261,        /* TOK_REALVAL */
5 TOK_PRIMITIVE = 262,      /* TOK_PRIMITIVE */
6 TOK_SVA_LABEL = 263,      /* TOK_SVA_LABEL */
7 TOK_SPECIFY_OPER = 264,   /* TOK_SPECIFY_OPER */
```

Es zeigt sich in dieser Liste, dass *TOK_ID* nur ein weiteres Token ist. Insgesamt gibt es also nur eine Rückgabeart.

Damit der Parser auf diese Tokens reagieren kann, muss eine Interpretation der Integer durchgeführt werden. Darüber wird im nächsten Kapitel diskutiert.

```
1 /* use special token for labels on assert,
2 assume, cover, and restrict because it's insanley complex
3 to fix parsing of cells otherwise.
4 (the current cell parser forces a reduce very early to update some
5 global state.. its a mess) */
6 [a-zA-Z_$][a-zA-Z0-9_$]*/[ \t\r\n]*:[ \t\r\n]*
7 (assert|assume|cover|restrict)[^a-zA-Z0-9_$\.] {
8 if (!strcmp(yytext, "default"))
9     return TOK_DEFAULT;
10 yylval->string = new std::string(std::string("\\") + yytext);
11 return TOK_SVA_LABEL;
12 }
```

Besonders komplizierte Ausdrücke bekommen hier entweder einen Standardtoken oder ein Label, falls die aktuelle Eingabe nicht gleich *default* ist (*strcmp* = 0 für gleiche Strings). Dies gilt für *assert*, *assume*, *cover*, *restrict*. Diese Ausdrücke sind außerdem SystemVerilog Direktive, die von Yosys nur sehr limitiert unterstützt werden. Daher ist das folgende weitere Handling abzusehen:

```
1 "assert"      { if (formal_mode)
2 return TOK_ASSERT; SV_KEYWORD(TOK_ASSERT); }
3 "assume"      { if (formal_mode)
4 return TOK_ASSUME; SV_KEYWORD(TOK_ASSUME); }
5 "cover"       { if (formal_mode)
6 return TOK_COVER; SV_KEYWORD(TOK_COVER); }
7 "restrict"    { if (formal_mode)
8 return TOK_RESTRICT; SV_KEYWORD(TOK_RESTRICT); }
9 "property"    { if (formal_mode)
10 return TOK_PROPERTY; SV_KEYWORD(TOK_PROPERTY); }
11 "rand"       { if (formal_mode)
12 return TOK_RAND; SV_KEYWORD(TOK_RAND); }
```



```

13 "const"      { if (formal_mode)
14 return TOK_CONST; SV_KEYWORD(TOK_CONST); }
15 "checker"    { if (formal_mode)
16 return TOK_CHECKER; SV_KEYWORD(TOK_CHECKER); }
17 "endchecker" { if (formal_mode)
18 return TOK_ENDCHECKER; SV_KEYWORD(TOK_ENDCHECKER); }
19 "final"      { SV_KEYWORD(TOK_FINAL); }
20 "logic"      { SV_KEYWORD(TOK_LOGIC); }
21 "var"        { SV_KEYWORD(TOK_VAR); }
22 "bit"        { SV_KEYWORD(TOK_LOGIC); }
23 "int"        { SV_KEYWORD(TOK_INT); }
24 "byte"       { SV_KEYWORD(TOK_BYTE); }
25 "shortint"   { SV_KEYWORD(TOK_SHORTINT); }
26 "longint"    { SV_KEYWORD(TOK_LONGINT); }
27
28 "eventually" { if (formal_mode)
29 return TOK_EVENTUALLY; SV_KEYWORD(TOK_EVENTUALLY); }
30 "s_eventually" { if (formal_mode)
31 return TOK_EVENTUALLY; SV_KEYWORD(TOK_EVENTUALLY); }

```

Diese Ausdrücke werden alle als SystemVerilog interpretiert. Die Variable *formal_mode* wird in *verilog_frontend.cc* definiert und ergibt sich aus dem Lesemodus im Yosys Terminal mit dem Befehl *read -formal verilogdatei.v*.

```

1 "input"      { return TOK_INPUT; }
2 "output"     { return TOK_OUTPUT; }
3 "inout"      { return TOK_INOUT; }
4 "wire"       { return TOK_WIRE; }
5 "wor"        { return TOK_WOR; }
6 "wand"       { return TOK_WAND; }
7 "reg"        { return TOK_REG; }
8 "integer"    { return TOK_INTEGER; }
9 "signed"     { return TOK_SIGNED; }
10 "unsigned"   { SV_KEYWORD(TOK_UNSIGNED); }
11 "genvar"     { return TOK_GENVAR; }
12 "real"       { return TOK_REAL; }
13 "enum"       { SV_KEYWORD(TOK_ENUM); }
14 "typedef"    { SV_KEYWORD(TOK_TYPEDEF); }
15 "struct"     { SV_KEYWORD(TOK_STRUCT); }
16 "union"      { SV_KEYWORD(TOK_UNION); }
17 "packed"     { SV_KEYWORD(TOK_PACKED); }

```

Weitere Tokens und SystemVerilog Ausdrücke.

```

1 [0-9][0-9_]* {
2   yylval->string = new std::string(yytext);
3   return TOK_CONSTVAL;
4 }

```

Alle konstanten Ganzzahlen werden in *yylval* überführt und ein Token *TOK_CONSTVAL* zurückgegeben.

```

1 \'[01zxZX] {
2   yylval->string = new std::string(yytext);
3   return TOK_UNBASED_UNSIZEED_CONSTVAL;
4 }

```

Ein SystemVerilog Ausdruck. Erkennt einige Signalzuweisungen, die einen Bitvektor füllen. Ein Beispiel dafür ist der Verilogausdruck *a <= '1;*. Es werden die Bitzustände 0,1,X,Z (Groß- und Kleinschreibung) erkannt.

```

1  \'[sS]?[bodhBODH] {
2      BEGIN(BASED_CONST);
3      yylval->string = new std::string(yytext);
4      return TOK_BASE;
5  }

```

Erkennt einige Signалуweisungen, die einen Bitvektor füllen. In diesem Fall werden Zahlenprefixe für signed (oder unsigned bei ausbleibendem *s*) und Binär-, Oktal-, Dezimal- und Hexadezimalzahlen (Groß- und Kleinschreibung) erkannt. Dieser Ausdruck läutet eine Zahl mit der festgelegten Basis ein. Daher wird hier die Startbedingung `BASED_CONST` für die nachfolgende Eingabe festgelegt. Nur Reguläre Ausdrücke, die mit `<BASED_CONST>` in den Regeln eingeleitet werden

```

1  <BASED_CONST>[0-9a-fA-FzxZX?][0-9a-fA-FzxZX?_]* {
2      BEGIN(0);
3      yylval->string = new std::string(yytext);
4      return TOK_BASED_CONSTVAL;
5  }

```

Erkennt alle Zahlen bei Startbedingung `BASED_CONST` für eine vorausgehende Basisfestlegung.

```

1  [0-9][0-9_]*\.[0-9][0-9_]*([eE][+-]?[0-9_]+)? {
2      yylval->string = new std::string(yytext);
3      return TOK_REALVAL;
4  }

```

Eingabemöglichkeiten von reellen Zahlen mit Komma (bspw. `1.5e+6`) werden erkannt.

```

1  [0-9][0-9_]*[eE][+-]?[0-9_]+ {
2      yylval->string = new std::string(yytext);
3      return TOK_REALVAL;
4  }

```

Eingabemöglichkeiten von reellen Zahlen ohne Komma werden erkannt (bspw. `1e+6` oder `1e-6`).

```

1  \" { BEGIN(String); }
2  <String>\\. { yymore(); real_location = old_location; }
3  <String>\" {
4      BEGIN(0);
5      char *yystr = strdup(yytext);
6      yystr[strlen(yytext) - 1] = 0;
7      int i = 0, j = 0;
8      while (yystr[i]) {
9          if (yystr[i] == '\\') {
10             i++;
11             if (yystr[i] == 'a')
12                 yystr[i] = '\\a';
13             else if (yystr[i] == 'f')
14                 yystr[i] = '\\f';
15             else if (yystr[i] == 'n')
16                 yystr[i] = '\\n';
17             else if (yystr[i] == 'r')
18                 yystr[i] = '\\r';
19             else if (yystr[i] == 't')
20                 yystr[i] = '\\t';
21             else if (yystr[i] == 'v')
22                 yystr[i] = '\\v';

```

```

23     else if ('0' <= yystr[i] && yystr[i] <= '7') {
24         yystr[i] = yystr[i] - '0';
25         if ('0' <= yystr[i + 1] &&
26             yystr[i + 1] <= '7') {
27             yystr[i + 1] =
28                 yystr[i] * 8 + yystr[i + 1] - '0';
29             i++;
30         }
31         if ('0' <= yystr[i + 1] &&
32             yystr[i + 1] <= '7') {
33             yystr[i + 1] =
34                 yystr[i] * 8 + yystr[i + 1] - '0';
35             i++;
36         }
37     }
38 }
39 yystr[j++] = yystr[i++];
40 }
41 yystr[j] = 0;
42 yylval->string = new std::string(yystr, j);
43 free(yystr);
44 return TOK_STRING;
45 }

```

Mit einem Anführungszeichen " wird der Beginn eines Strings eingeleitet. Der Lexer geht in die Startbedingung STRING. Alle weiteren geltenden Regeln beginnen mit <STRING>.

Die nächste Regel erkennt alle Strings, die mit \ beginnen und ein weiteres Zeichen haben (beispielsweise \h). In dem folgenden Quellcode wird mit den Objekten *real_location* und *old_location* vom Typ *YYLTYPE* durch eine Zuweisung die Position intern gespeichert. Außerdem wird die Flex Funktion *yymore()* aufgerufen. Diese wird benutzt, wenn die folgende Eingabe an die aktuelle Eingabe angehängt werden soll, weil ein regulärer Ausdruck, der auf die gesamte Eingabe zutreffen soll, zu kompliziert ist.

Die folgende Regel hat auch die Bedingung STRING und erkennt ein Anführungszeichen ".

In dem Quellcode werden die Startbedingungen zurückgesetzt. Mit *strdup(yytext)* wird der aktuelle String dupliziert und ein Pointer auf das Duplikat auf *yystr* zurückgegeben. Anschließend wird das letzte Zeichen von *yystr* auf 0 gesetzt. Damit kann mit einer while-Schleife mit der Bedingung *yystr[i]* durch das Wort iteriert werden, da die 0 das Unterbrechungskriterium ist. Durch die Verwendung eines Pointers auf ein Duplikat wird vermieden, dass *yytext* verändert wird.

```

● <STRING>. { yymore(); real_location = old_location; }

● and|nand|or|nor|xor|xnor|not|buf|bufif0|bufif1|notif0|notif1 {
2   yylval->string = new std::string(yytext);
3   return TOK_PRIMITIVE;
4 }

```

Diese Regel beschreibt alle sogenannten *Built-In Primitives* von Verilog für Hardware-Beschreibung auf Gatterebene. Dazu gehören die Standardgatter, Inverter und Tri-State Buffer ohne und mit Steuereingang High und Low-Aktiv mit einem weiteren Zustand Z. Alle Ausdrücke sind Oder-Verknüpft. Es wird jeweils der Token TOK_PRIMITIVE, sowie die Art des Tokens über *yytext* in *yylval* zurückgegeben.

```

1 supply0 { return TOK_SUPPLY0; }
2 supply1 { return TOK_SUPPLY1; }

```

supply1 und supply0 werden für die Modellierung von Spannungsversorgung und Masse verwendet.

```

1 "$"(display|write|strobe|monitor|time
2 |stop|finish|dumpfile|dumpvars|dumpon|dumpoff|dumpall) {
3   yylval->string = new std::string(yytext);
4   return TOK_ID;
5 }

```

Diese Regel erkennt einige Ausdrücke, die mit \$ beginnen. Dies sind hauptsächlich Testbench-Funktionen von Verilog. Es wird der Token TOK_ID zurückgegeben, der zuvor zurückgegeben wurde, wenn ein SystemVerilog Ausdruck nicht unterstützt wurde. Da Yosys kein Simulator ist, werden auch Testbench Funktionen nicht unterstützt.

```

1 "$"(setup|hold|setuphold|removal|recovery|recrem
2 |skew|timeskew|fullskew|nochange) {
3   if (!specify_mode) REJECT;
4   yylval->string = new std::string(yytext);
5   return TOK_ID;
6 }

```

Bei diesen Funktionen handelt es sich weiterhin um Testbench Funktionen, die nicht unterstützt werden.

```

1 "$"(info|warning|error|fatal) {
2   yylval->string = new std::string(yytext);
3   return TOK_MSG_TASKS;
4 }

```

Weitere SystemVerilog Testbench Funktionen. Hier wird jedoch anstelle des TOK_ID Tokens TOK_MSG_TASKS zurückgegeben.

```

1 "$signed" { return TOK_TO_SIGNED; }
2 "$unsigned" { return TOK_TO_UNSIGNED; }

```

Operatoren für Konvertierungen zwischen unsigned und signed.

```

1 [a-zA-Z_][a-zA-Z0-9_]*::[a-zA-Z_$][a-zA-Z0-9_$]* {
2   // package qualifier
3   auto s = std::string("\\") + yytext;
4   if (pkg_user_types.count(s) > 0) {
5     // package qualified typedefed name
6     yylval->string = new std::string(s);
7     return TOK_PKG_USER_TYPE;
8   }
9   else {
10    // backup before :: just return first part
11    size_t len = strchr(yytext, ':') - yytext;
12    yyless(len);
13    yylval->string =
14      new std::string(std::string("\\") + yytext);
15    return TOK_ID;
16  }
17 }

```

Für das Importieren von SystemVerilog Packages. `pkg_user_types` ist ein Objekt der Klasse `dict`, die in `hashlib.h` definiert ist. Die Funktion `count(const K &key)` dieser Klasse gibt eine 1 zurück, wenn ein Hash-Lookup mit `\\yytext` einen Wert `>0` (1 bei erfolgreichem Lookup) ergibt. Während eines Lookups wird der übergebene Key in einen Hash umgewandelt, um im Speicher einen passenden Eintrag zu finden (`lookup`). Wenn der Eintrag gefunden wurde, wird der Token `TOK_PKG_USER_TYPE` zurückgegeben und der Hash-key `s` in `yylval` gespeichert.

Ansonsten wird nur der Teil der Package-Beschreibung vor dem `::` Zeichen in `yylval` gespeichert und der Token `TOK_ID` wird wie bei anderen SystemVerilog Behandlungen zurückgegeben.

```

1 [a-zA-Z_$][a-zA-Z0-9_$]* {
2     auto s = std::string("\\") + yytext;
3     if (isUserType(s)) {
4         // previously typedefed name
5         yylval->string = new std::string(s);
6         return TOK_USER_TYPE;
7     }
8     else {
9         yylval->string =
10         new std::string(std::string("\\") + yytext);
11         return TOK_ID;
12     }
13 }

```

Erkennt alle Buchstaben, `_` und `$`, sowie diese nachfolgend ergänzt mit Zahlen. `isUserType(s)` benutzt die Funktion `count(const key_type& __x)` aus der C++ Standardbibliothek `std::map.h`, die 1 zurückgibt, wenn `key` in dem `map`-Container vorhanden ist.

```

1 static bool isUserType(std::string &s)
2 {
3     // check current scope then outer scopes for a name
4     for (auto it = user_type_stack.rbegin();
5     it != user_type_stack.rend(); ++it) {
6         if ((*it)->count(s) > 0) {
7             return true;
8         }
9     }
10     return false;
11 }

```

Der `map`-Container ist in diesem Fall ein Container aus `std::string` und `AST::AstNode`. Diese Container sind in einem Vektor `user_type_stack`, der iteriert wird. Wenn also `\\yytext` in einem `map`-Container gefunden wird, bedeutet dies, dass vorher ein benutzerdefinierter Ausdruck beziehungsweise AST-Knoten über die Funktion `static void addTypeDefNode(std::string *name, AstNode *node)` hinzugefügt wurde. In diesem Fall wird der Token `TOK_USER_TYPE` zurückgegeben und `yylvalue` auf `\\yytext` gesetzt. Andernfalls wird der Token `TOK_ID` zurückgegeben.

```

1 [a-zA-Z_$][a-zA-Z0-9_$\.]* {
2     yylval->string = new std::string(std::string("\\") + yytext);
3     return TOK_ID;
4 }

```

Ähnliche Regel wie zuvor, nur mit optionalem Punkt `.` nach dem ersten Zeichen. Damit werden benutzerdefinierte Ausdrücke exkludiert und es wird wie zuvor `TOK_ID` zurückgegeben.

```

1  "/*[ \t]*(synopsys|synthesis)[ \t]*translate_off[ \t]*"*/" {
2      static bool printed_warning = false;
3      if (!printed_warning) {
4          log_warning(
5              "Encountered 'translate_off' comment!
6              Such legacy hot "
7              "comments are supported by Yosys,
8              but are not part of "
9              "any formal language specification.
10             Using a portable "
11             "and standards-compliant
12             construct such as 'ifdef is "
13             "recommended!\n"
14         );
15         printed_warning = true;
16     }
17     BEGIN(SYNOPSYS_TRANSLATE_OFF);
18 }

```

Wenn ein Kommentar wie */*synopsys translate_off*/* im Verilogcode steht, wird eine Warnung geloggt, die darauf hinweist, dass Yosys dieses Konstrukt zwar unterstützt, jedoch `ifdef` empfohlen wird, um die Sprachspezifizierungen nicht zu überschreiten.

Anschließend wird die Anfangsbedingung `SYNOPSIS_TRANSLATE_OFF` für die nächsten Ausdrücke festgelegt.

```

1 <SYNOPSISYS_TRANSLATE_OFF>. /* ignore synopsys translate_off body */
2 <SYNOPSISYS_TRANSLATE_OFF>\n /* ignore synopsys translate_off body */
3 <SYNOPSISYS_TRANSLATE_OFF>>"/*" [ \t]*(synopsys|synthesis)[ \t]*
4 "translate_on"[ \t]*"/" { BEGIN(0); }

```

Diese Regeln erfüllen die Anfangsbedingung `SYNOPSIS_TRANSLATE_OFF`. Die erste Regel erkennt ein nachfolgendes Zeichen außer einen Zeilenumbruch, der `synopsis translate_off` body wird mangels folgendem C++-Code ignoriert. Das gleiche passiert in der zweiten Regel für einen Zeilenumbruch.

Die letzte Regel erkennt beispielsweise den Ausdruck */*synopsys translate_on*/*, mit dem der `translate_off` body zu ende ist. Damit wird die Anfangsbedingung `SYNOPSIS_TRANSLATE_OFF` mit `BEGIN(0)` überschrieben und beendet.

```
1  "/*" [ \t]*(synopsys|synthesis)[ \t]+ {
2      BEGIN(SYNOPSYS_FLAGS);
3  }
```

Der Ausdruck `/*synopsis` steht vor Synopsys-Flags, die anschließend erkannt werden. Daher wird die Anfangsbedingung `SYNOPSIS_FLAGS` für die folgenden Ausdrücke festgelegt.

```

1  <SYNOPSIS_FLAGS>full_case {
2      static bool printed_warning = false;
3      if (!printed_warning) {
4          log_warning(
5              "Encountered 'full_case' comment! Such legacy hot "
6              "comments are supported by Yosys, but are not part of "
7              "any formal language specification. Using the Verilog "
8              "'full_case' attribute or the SystemVerilog 'unique' "
9              "or 'unique0' keywords is recommended!\n"
10             );
11         printed_warning = true;
12     }

```

```

13     return TOK_SYNOPSIS_FULL_CASE;
14 }

```

Für Ausdrücke, die mit */*synopsys full_case* beginnen, wird eine Warnung geloggt, in der erneut auf das Synopsysproblem hingewiesen wird und das Verilog Schlüsselwort *full_case* oder SystemVerilog Schlüsselwörter *unique* oder *unique0* empfohlen werden.

Der Token TOK_SYNOPSIS_FULL_CASE wird zurückgegeben. Hier sollte später beim Parser analysiert werden, ob TOK_SYNOPSIS_FULL_CASE und TOK_UNIQUE gleichbehandelt werden.

```

❶ <SYNOPSIS_FLAGS>parallel_case {
2     static bool printed_warning = false;
3     if (!printed_warning) {
4         log_warning(
5             "Encountered 'parallel_case' comment! Such legacy hot "
6             "comments are supported by Yosys, but are not part of "
7             "any formal language specification. Using the Verilog "
8             "'parallel_case' attribute or the SystemVerilog "
9             "'unique' or 'priority' keywords is recommended!\n"
10        );
11        printed_warning = true;
12    }
13    return TOK_SYNOPSIS_PARALLEL_CASE;
14 }

```

Für das Synopsys-Flag *parallel_case* wird auch eine Warnung mit Empfehlung für *parallel_case*, *unique* oder *priority* geloggt.

```

❶ <SYNOPSIS_FLAGS>. /* ignore everything else */
2 <SYNOPSIS_FLAGS>"/" { BEGIN(0); }

```

Alle anderen Synopsys-Flags werden ignoriert. Mit **/* ist die Eingabe von Synopsys-Flags beendet und die Anfangsbedingung wird mit BEGIN(0) zurückgesetzt.

```

❶ import [ \t\r\n]+\"(DPI|DPI-C)\"[ \t\r\n]+function [ \t\r\n]+ {
2     BEGIN(IMPORT_DPI);
3     return TOK_DPI_FUNCTION;
4 }

```

Diese Regel berücksichtigt die Import-Methoden von SystemVerilog, mit denen Funktionen aus fremden Sprachen importiert und aufgerufen werden können. Damit wird C, C++, SystemC und weiteres unterstützt.

Wird dieser Ausdruck erkannt, so wird die Startbedingung IMPORT_DPI festgelegt und der Token TOK_DPI_FUNCTION wird festgelegt.

```

❶ <IMPORT_DPI>[a-zA-Z_$][a-zA-Z0-9_$]* {
2     yyval->string = new std::string(std::string("\\") + yytext);
3     return TOK_ID;
4 }

```

Unter der Startbedingung IMPORT_DPI werden nun Funktionen anderer Sprachen erkannt. Die Funktion, die importiert wird, wird in yyval gespeichert und TOK_ID wird zurückgegeben.

```

❶ <IMPORT_DPI>[ \t\r\n] /* ignore whitespaces */

```

Tabulatoren, Returns und Zeilenumbrüche werden ignoriert.

```

1 <IMPORT_DPI>" ;" {
2     BEGIN(0);
3     return *yytext;
4 }

```

Da nur eine Funktion erkannt wird, wird der Import ab dem Ende der Funktion bei einem Semikolon beendet. Die Anfangsbedingung wird zurückgesetzt und ein Zeiger auf den aktuellen Ausdruck wird zurückgegeben. Dies ist unüblich, da ansonsten immer Tokens zurückgegeben werden. Der Return-Wert wäre hier das Semikolon.

```

1 <IMPORT_DPI>. {
2     return *yytext;
3 }

```

Alle weiteren Zeichen unter der IMPORT_DPI Bedingung.

```

1 "\\\" [^ \\t\\r\\n]+ {
2     yylval->string = new std::string(yytext);
3     return TOK_ID;
4 }

```

Alles außer Whitespaces nach \\. Dies sind Kommentare.

```

1 "(" { return ATTR_BEGIN; }
2 ")" { return ATTR_END; }
3
4 "{" { return DEFATTR_BEGIN; }
5 "}" { return DEFATTR_END; }
6
7 "*" { return OP_POW; }
8 "||" { return OP_LOR; }
9 "&&" { return OP_LAND; }
10 "==" { return OP_EQ; }
11 "!=" { return OP_NE; }
12 "<=" { return OP_LE; }
13 ">=" { return OP_GE; }
14
15 "===" { return OP_EQX; }
16 "!==" { return OP_NEX; }
17
18 "~&" { return OP_NAND; }
19 "~|" { return OP_NOR; }
20 "^~" { return OP_XNOR; }
21 "^~" { return OP_XNOR; }
22
23 "<<" { return OP_SHL; }
24 ">>" { return OP_SHR; }
25 "<<<" { return OP_SSHL; }
26 ">>>" { return OP_SSHR; }
27
28 "'" { return OP_CAST; }
29
30 "::" { return TOK_PACKAGESEP; }
31 "++" { return TOK_INCREMENT; }
32 "--" { return TOK_DECREMENT; }
33
34 "+:" { return TOK_POS_INDEXED; }
35 "-:" { return TOK_NEG_INDEXED; }
36
37 ".*" { return TOK_WILDCARD_CONNECT; }

```



```

38
39 "|=" { SV_KEYWORD(TOK_OR_ASSIGN); }
40 "&=" { SV_KEYWORD(TOK_AND_ASSIGN); }
41 "+=" { SV_KEYWORD(TOK_PLUS_ASSIGN); }
42 "-=" { SV_KEYWORD(TOK_SUB_ASSIGN); }
43 "^=" { SV_KEYWORD(TOK_XOR_ASSIGN); }
44
45 [-+]?[=*]> {
46     if (!specify_mode) REJECT;
47     yylval->string = new std::string(yytext);
48     return TOK_SPECIFY_OPER;
49 }
50
51 "&&&" {
52     if (!specify_mode) return TOK_IGNORED_SPECIFY_AND;
53     return TOK_SPECIFY_AND;
54 }

```

Rückgabe von Tokens für alle Verilog Operatoren. SystemVerilog Operatoren werden mit der Funktion SV_KEYWORD behandelt. Wenn Yosys nicht im -specify Modus läuft, wird REJECT aufgerufen. Die REJECT Funktion von *Flex* bewirkt, dass die nächstbeste Regel auf den Ausdruck angewendet wird. Ansonsten wird der aktuelle Ausdruck in yylval gespeichert und der Token TOK_SPECIFY_OPER zurückgegeben.

Es werden scheinbar neue (System)Verilog Funktionen eingeführt, die nur erkannt werden, wenn der *specify_mode* eingeschaltet ist.

```

1 • <INITIAL,BASED_CONST>"/*" { comment_caller=YY_START; BEGIN(COMMENT); }
2 <COMMENT>.* /* ignore comment body */
3 <COMMENT>\n /* ignore comment body */
4 <COMMENT>"/*" { BEGIN(comment_caller); }
5
6 <INITIAL,BASED_CONST>[ \t\r\n] /* ignore whitespaces */
7 <INITIAL,BASED_CONST>\\[\r\n] /* ignore continuation sequence */
8 <INITIAL,BASED_CONST>"/" [^\r\n]* /* ignore one-line comments */
9
10 <INITIAL>.* { return *yytext; }
11 <*>. { BEGIN(0); return *yytext; }

```

Weitere Regeln für die Startbedingungen INITIAL und BASED_CONST. INITIAL ist die Startbedingung von *Flex*. Wenn /* erkannt wird, wird die Startbedingung COMMENT aktiviert. Danach werden alle Zeichen und Zeilenumbrüche erkannt. Diese Regel berücksichtigt also Kommentare über mehrere Zeilen. Mit */ wird der Kommentar beendet. *Flex* geht in die Startbedingung comment_caller, die vorher auf YY_START gesetzt wurde.

Weitere Regeln für die Startbedingungen INITIAL und BASED_CONST sind für Leerzeichen, Tabulatoren, Returns und Zeilenumbrüche, die ignoriert werden, Continuation Sequences und einzeilige Kommentare.

1.3.2 Zusammenfassung der Lexerregeln

Für die meisten Verilogausdrücke werden Tokens über return zurückgegeben, die ähnlich benannt sind, wie der eingelesene Ausdruck. Gleichzeitig wird in der *Flex* Klasse der aktuelle Wert des Ausdrucks gespeichert, wenn es sich nicht um eine Regel handelt, die nur für einen einzigen Ausdruck gilt.

SystemVerilog wird nur teilweise unterstützt. Die Tokens werden der Funktion SV_KEYWORD übergeben, die nur den passenden Token zurückgibt, wenn der SystemVerilog Modus von Yosys

sv_mode aktiv ist.

Ansonsten wird TOK_ID zurückgegeben. Dies scheint eine Art Fehlertoken zu sein, der in allen Fällen benutzt wird, in denen es Kompatibilitätsprobleme gibt.

Weiterhin gibt es den Lesemodus *formal_mode*, der für einige Ausdrücke entscheidet, ob sie als Verilog oder SystemVerilog Ausdrücke behandelt werden sollen.

Der Modus *specify_mode* scheint für die Behandlung von spezifischen benutzerdefinierten Verilogausdrücken zu sein.

1.3.3 Interpretationsprobleme und Ansätze

- Was macht die *Flex* Funktion yymore() genau?
- Die String-Regel ist unverständlich. Es scheint als würden die Buchstaben der Backslash Sequenzen wie \n, \t .. gelesen werden und durch die eigenen Backslash Sequenzen ersetzt zu werden. Weiterhin ist unklar, warum mit den Zeichen 1 bis 6 am Ende der Regel gerechnet wird (ASCII?).
- TOK_ID scheint ein Fehlertoken zu sein. Ist das richtig?
- Werden die Synopsys spezifischen Tokens und die jeweils passenden empfohlenen SystemVerilog Tokens vom Parser gleichbehandelt?
- Was sind Continuation-Sequences (Mit \\\)?

1.3.4 Verbindung zum Parser

Die Regeln, die vorher analysiert wurden, befinden sich in der Datei verilog_lexer.l, die von *Flex* in einen C++ Code verilog_lexer.cc überführt werden.

In dieser Datei befinden sich einige Makrofunktionen. Ein Kommentar, sowie eine lange Case-Anweisung mit Integercases und Token>Returns weisen darauf hin, dass die Makrofunktion YY_DECL die Haupt-Scannerfunktion ist.

In die Scanner Klasse verilog_lexer.cc wird die Parserbibliothek verilog_parser.tab.hh mit den Tokendefinitionen importiert. *Flex* hat eine Funktion, yylex(), die ein Wertepaar Token und Wert zurückgibt. Wir haben bereits in der Tokenanalyse gesehen, dass die Tokens Integer Werte haben. Diese Werte sind in verilog_parser.tab.hh festgelegt.

Ursprünglich wird dem Bison Parser eine Liste aller Tokens in seiner Grammatikdatei .y übergeben.

```
1 %token TOK_ASSERT TOK_ASSUME TOK_RESTRICT TOK_COVER TOK_FINAL
2 %token ATTR_BEGIN ATTR_END DEFATTR_BEGIN DEFATTR_END
3 %token TOK_MODULE TOK_ENDMODULE TOK_PARAMETER TOK_LOCALPARAM TOK_DEFPARAM
4 %token TOK_PACKAGE TOK_ENDPACKAGE TOK_PACKAGESEP
```

Code: Ausschnitt aus der Grammatikdatei: Tokenliste

Außerdem wird in der Bison Parser Grammatikdatei, die ähnlich aufgebaut ist, wie die am Anfang beschriebene Lexerdatei (Code + Regeln), unter anderem ein Objekt vom Typ *std::istream *lexin* definiert. Dies scheint der Lexer Inputstream zu sein, der vom Preprozessor kommt. Folgende Abbildung verdeutlicht den Ablauf:

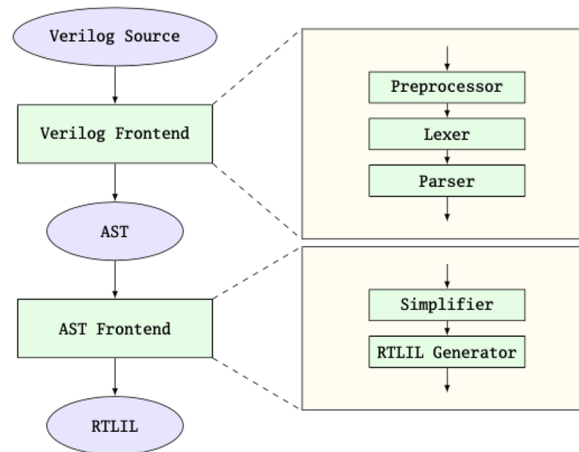


Figure 7.1: Simplified Verilog to RTLIL data flow

Das Inputstreamobjekt *lexin* wird in `verilog_frontend.cc` genutzt.

```

1 if (!flag_nopp) {
2   code_after_preproc = frontend_verilog_preproc(*f, filename, defines_map,
3   *design->verilog_defines, include_dirs);
4   if (flag_ppdump)
5     log("-- Verilog code after preprocessor
6     --\n%s-- END OF DUMP --\n", code_after_preproc.c_str());
7   lexin = new std::istringstream(code_after_preproc);
8   }

```

Das Streamobjekt *f*, das dem Preprozessor übergeben wird, wird im Konstruktor der hierarchisch höheren Funktion übergeben:

```

1 void execute(std::istream *f, std::string filename,
2 std::vector<std::string> args, RTLIL::Design *design) override

```

code_after_preproc ist also die Ausgabe des Preprozessors, deren Wert *lexin* bekommt. *lexin* findet sich im Lexercode wieder:

```

1 #define YY_INPUT(buf,result,max_size) \
2   result = readsome(*VERILOG_FRONTEND::lexin, buf, max_size)

```

Die Funktion *readsome* ist folgende:

```

1 int readsome(std::istream &f, char *s, int n)
2 {
3   int rc = int(f.readsome(s, n));
4
5   // f.readsome() sometimes returns 0 on a non-empty stream..
6   if (rc == 0) {
7     int c = f.get();
8     if (c != EOF) {
9       *s = c;
10      rc = 1;
11    }
12  }
13
14  return rc;
15 }

```

Es wird ein `istream` Objekt *f*, ein Char-Array **s* und ein Integer *n* übergeben. Aus *f* wird mit `readsome` eine Anzahl *n* Zeichen in ein Char-Array *s* geladen. Die Rückgabe ist die Anzahl an Zeichen, die geladen wurden. Danach kommt eine Behandlung für den Fall, dass `rc == 0`,

wobei es sein kann, dass das istream Objekt nicht wirklich leer ist.

In der Funktion YY_INPUT wird also der Inputstream *lexin* in den Buffer *buf* geladen, bis *max_size* erreicht ist. In result steht später eine 1, wenn ein Lesevorgang erfolgreich war.

Im *Flex* Manual steht zu YY_INPUT, dass, wenn diese Funktion auf einen Eingang definiert ist (hier *lexin*), der Funktionsaufruf des Lexers mit einem Nullpointer als Argument durchgeführt werden kann.

Später erfolgen folgende Funktionsaufrufe:

```
1 frontend_verilog_yyset_lineno(1);  
2 frontend_verilog_yyrestart(NULL);  
3 frontend_verilog_yyparse();  
4 frontend_verilog_yylex_destroy();
```

Die Funktion *frontend_verilog_yyrestart* wird mit einem Nullpointer wie vorher diskutiert aufgerufen. Damit wird *Flex* also mit der Eingabe *lexin* aufgerufen.

Normalerweise ist die *Flex* Hauptfunktion *yylex()*. Hier wird *yyrestart()* benutzt. Dies könnte damit zusammen hängen, dass ein Inputstream eingestellt wird. Außerdem bleiben bei *yyrestart()* die Anfangsbedingungen der Regeln erhalten.

Mit *yyparse()* wird dann der Parser aufgerufen. In dem Parser Quellcode wird *yylex()* aufgerufen. Die Integer Rückgabe wird in *ychar* gespeichert. Danach wird die Funktion *yytranslate()* auf *ychar* aufgerufen, die eine zu dem Integer korrespondierende Symbolnummer zurückgibt. Diese wird in *yytoken* gespeichert. Diese Variable ist vom Typ *yysymbol_kind_t*.

In dem Parser Quellcode existiert eine lange Case-Anweisung für die Integer Variable *yyn*. Hier wird die AST-Struktur gebaut. *yyn* wird an einer Stelle mit *yytoken* inkrementiert.

Kapitel 2

Parser

Der Parser wird wie der Lexer automatisiert anhand einer Grammatikdatei .y gebaut. Hier wird anstatt *Flex* der Open-Source Parsergenerator *Bison* genutzt.

2.1 Analyse des Parser-Quellcodes

Die Parsergrammatik ist ähnlich wie die *Flex* Beschreibung aufgebaut:

```
1 %{  
2 C declarations  
3 %}  
4 Bison declarations  
5  
6 %%  
7 Grammar rules  
8 %%  
9  
10 Additional C code
```

Im folgenden werden die Grammatikregeln analysiert:

2.1.1 Analyse der Parser-Grammatik

Um die Grammatikregeln zu verstehen, ist es notwendig, deren Zusammensetzung nachzuvollziehen.

Eine Grammatikregel in der Backus-Naur Form sieht so aus:

```
1 result: components ...  
2 ;
```

Result ist ein Nicht-Terminalsymbol. Component sind Terminal- und Nicht-Terminalsymbole. Terminalsymbole sind Symbole, die einzeln nicht weiter durch Produktionsregeln ersetzt werden können. Eine Beispielregel wäre folgende:

```
1 exp: exp '+' exp  
2 ;  
3  
4
```

exp '+' exp kann also in einer Produktion mit exp ersetzt werden.

Es folgt ein Beispiel für einen Taschenrechner, der mit Klammern Addieren und Multiplizieren kann. Zuerst wird ein Lexer mit *Flex* gebaut, der die Eingabe in Tokens zerlegt.

```

1 %{
2 #include <stdio.h>
3 #include <string.h>
4 #include "add.tab.h"
5
6
7
8 %}
9
10 %option noyywrap
11
12 %%
13
14 [0-9]+ { yylval.zahl = atoi(yytext); return(ZAHL);}
15 "+" {return(PLUS);}
16 "=" {return(GLEICH);}
17 ";" {return(SEMIKOLON);}
18 "*" {return(MAL);}
19 "(" {return(KLOFFEN);}
20 ")" {return(KLZU);}
21 "\n" {return(RETURN);}
22 . {return(OTHER);}
23
24
25
26 %%

```

Hier ist es wichtig, dass der Lexer in seiner *Declaration* Sektion den Header des zukünftigen Bison Parsers übergeben bekommt. Die Header-Datei wird Standardmäßig immer so bezeichnet: "named.ydatei".tab.h. In dieser Headerdatei stehen die Integerwerte für die Tokenrückgabe für die Interpretation des Parsers.

Es folgt die Parser Grammatik:

```

1
2 %{
3
4 #include <stdio.h>
5
6 extern int yylex();
7 extern int yyparse();
8 int yyerror(char *s);
9
10
11 %}
12
13 //alle Tokens
14 %token ZAHL PLUS GLEICH SEMIKOLON OTHER MAL RETURN KLOFFEN KLZU
15
16 //ZAHL bekommt einen Wert vom Typ zahl
17 %type <zahl> ZAHL
18 %type <zahl> faktor
19 %type <zahl> term
20 %type <zahl> ausdruck
21
22
23 //Definition des Typs zahl
24 %union{
25     int zahl;
26 }

```

```

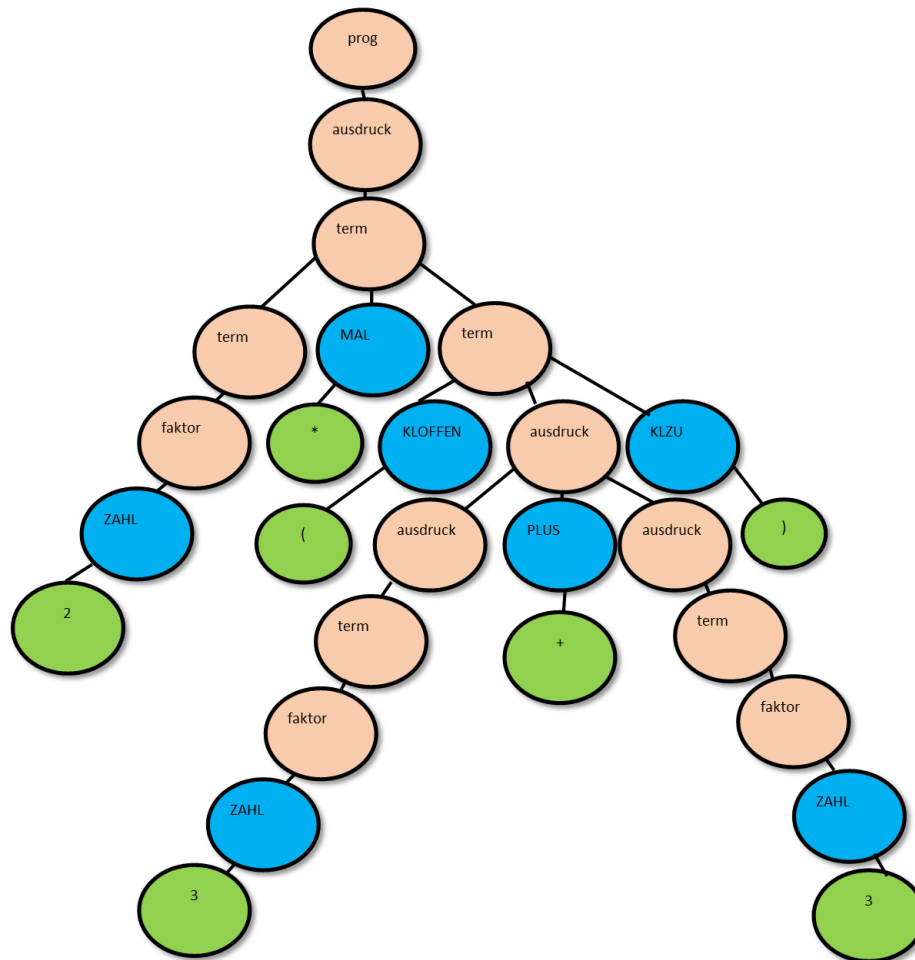
27
28 %%
29
30 prog:
31     ausdruck {
32         printf("Ergebnis: %d", $1);
33     }
34 ;
35 ausdruck:
36     ausdruck PLUS ausdruck{
37         $$=$1+$3;
38     }
39     | term{
40         $$=$1;
41     }
42 ;
43 term:
44     term MAL term{
45         $$=$1*$3;
46     }
47     | faktor{
48         $$=$1;
49     }
50 ;
51 faktor:
52     KLOFFEN ausdruck KLZU{
53         $$=$2;
54     }
55     | ZAHL{
56         $$=$1;
57     }
58 ;
59
60 %%
61
62 int yyerror(char *s)
63 {
64     printf("Syntax Fehler in Zeile %s\n", s);
65     return 0;
66 }
67
68 int main()
69 {
70     yyparse();
71     return 0;
72 }

```

In der Declarationssektion wird dem Parser übergeben, dass es eine `yylex()`, eine `yyparse()` und eine `yyerror()` Funktion gibt. Danach werden in der *Definitions* alle Tokens aufgelistet, die von dem Lexer übergeben werden können.

In dieser Sektion werden außerdem die Datentypen aller wertbehafteten Tokens sowie der im Parser definierten Nichtterminalsymbole mit `%type` definiert. In `%union` werden die Parser-Datentypen mit C-Datentypen verknüpft.

Anschließend folgen die Grammatik-Regeln.



Parserbaum für die Eingabe $2*(3+3)$. Zeichen sind grün, Tokens blau und die definierten Nichtterminalsymbole orange dargestellt.

Dieses Verständnis wird nun auf den Verilogparser übertragen. Aufgrund des großen Umfangs der Grammatik, dessen Ableitungen nicht so einfach analysiert werden können, wie die Tokengeneration im Lexer, wird die Parsergrammatik anhand eines Beispiels analysiert.

Die Analyse findet anhand des Verilog Ausdrucks *assign* statt. Ein einfaches Beispielm Modul wäre folgendes:

```

1  module behave;
2      wire i1,i2;
3      wire out;
4      assign out = i1 & i2;
5  endmodule

```

1. Vom Lexer wird für *assign* der Token TOK_ASSIGN zurückgegeben.
2. Im Parser steht folgende Grammatikregel:

```

1  assign_stmt:
2      TOK_ASSIGN delay assign_expr_list ';' ;
3

```

Für delay gibt es folgende Regel:


```

1      delay:
2      non_opt_delay | %empty;
3

```

delay kann also auch ein leeres Zeichen sein, was zu dem Beispiel passt. Zusätzlich dazu kommt non_opt_delay:

```

1      non_opt_delay:
2      '#' TOK_ID { delete $2; } |
3      '#' TOK_CONSTVAL { delete $2; } |
4      '#' TOK_REALVAL { delete $2; } |
5      '#' '(' mintypmax_expr ')' |
6      '#' '(' mintypmax_expr ',' mintypmax_expr ')' |
7      '#' '(' mintypmax_expr ',' mintypmax_expr ',' mintypmax_expr ')';
8

```

Da nicht weiter auf den Wert der Eingabe reagiert wird, kann interpretiert werden, dass dies eine Inkompatibilität von Yosys ist, die ignoriert wird.

Interessant wird hier assign_expr_list:

```

1      assign_expr_list:
2      assign_expr | assign_expr_list ',' assign_expr;
3

```

Mit folgender Erweiterung:

```

1      assign_expr:
2      lvalue '=' expr {
3          AstNode *node = new AstNode(AST_ASSIGN, $1, $3);
4          SET_AST_NODE_LOC(node, @$, @$);
5          ast_stack.back()->children.push_back(node);
6      };
7

```

Hier lässt sich die konkrete Zuweisung erkennen. Als Reaktion wird eine AstNode mit dem Typ AST_ASSIGN erstellt, dessen Children die Werte von lvalue und expr bekommen. Für die weitere Auswertung ist eine weitere Erweiterung möglich:

```

1      expr:
2      basic_expr {
3          $$ = $1;
4      } |
5      basic_expr '?' attr expr ':' expr {
6          $$ = new AstNode(AST_TERNARY);
7          $$->children.push_back($1);
8          $$->children.push_back($4);
9          $$->children.push_back($6);
10         SET_AST_NODE_LOC($$, @1, @$);
11         append_attr($$, $3);
12     };
13

```

Mit basic_expr: (Ausschnitt wegen langer Anweisung)

```

1      } |
2      basic_expr OP_LAND attr basic_expr {
3          $$ = new AstNode(AST_LOGIC_AND, $1, $4);
4          SET_AST_NODE_LOC($$, @1, @4);
5          append_attr($$, $3);
6      } |
7      basic_expr OP_LOR attr basic_expr {

```

```

8    $$ = new AstNode(AST_LOGIC_OR, $1, $4);
9    SET_AST_NODE_LOC($$, @1, @4);
10   append_attr($$, $3);
11   } |
12   '!' attr basic_expr %prec UNARY_OPS {
13     $$ = new AstNode(AST_LOGIC_NOT, $3);
14     SET_AST_NODE_LOC($$, @1, @3);
15     append_attr($$, $2);
16   } |

```

expr besteht also aus allen möglichen logischen Verknüpfungen, mathematischen Funktionen und Vergleichen.

3. assign_stmt kann weiter reduziert werden:

```

1   module_body_stmt:
2   task_func_decl | specify_block | param_decl | localparam_decl
3   | typedef_decl | defparam_decl | specparam_declaration | wire_decl
4   | assign_stmt | cell_stmt
5   | enum_decl | struct_decl
6   | always_stmt | TOK_GENERATE module_gen_body TOK_ENDGENERATE
7   | defattr | assert_property | checker_decl
8   | ignored_specify_block;

```

ODER:

```

1   interface_body_stmt:
2   param_decl | localparam_decl | typedef_decl
3   | defparam_decl | wire_decl | always_stmt | assign_stmt |
4   modport_stmt;

```

Hier sind zwei also zwei Reduktionen möglich! Dies führt zu einem Konflikt. Bison meldet in einem mehrdeutigen Fall wie diesem eine **reduce/reduce conflict** Warnung.

4. Für module_body_stmt ist folgende Reduktion möglich:

```

1   module_body:
2   module_body module_body_stmt |
3   /* the following line makes the generate..
4   endgenerate keywords optional */
5   module_body gen_stmt |
6   module_body gen_block |
7   module_body ';' |
8   %empty;

```

Für interface_body_stmt:

```

1   interface_body:
2   interface_body interface_body_stmt | %empty;

```

Auffällig ist, dass diese Reduzierungen jeweils nur eine weitere Stufe ohne weitere Reaktionen / Verknüpfungen sind.

5. module_body wird zu module reduziert:

```

1   module:
2   attr TOK_MODULE {
3     enterTypeScope();
4   } TOK_ID {
5     do_not_require_port_stubs = false;
6     AstNode *mod = new AstNode(AST_MODULE);
7     ast_stack.back()->children.push_back(mod);

```

```

8     ast_stack.push_back(mod);
9     current_ast_mod = mod;
10    port_stubs.clear();
11    port_counter = 0;
12    mod->str = *$4;
13    append_attr(mod, $1);
14    delete $4;
15 } module_para_opt module_args_opt ';' module_body
16 TOK_ENDMODULE opt_label {
17     if (port_stubs.size() != 0)
18         frontend_verilog_yyerror(
19             "Missing details for module port '%s'.",
20             port_stubs.begin()->first.c_str());
21     SET_AST_NODE_LOC(ast_stack.back(), @2, @$);
22     ast_stack.pop_back();
23     log_assert(ast_stack.size() == 1);
24     current_ast_mod = NULL;
25     exitTypeScope();
26 };

```

6. module wird auf design reduziert. Hier ist sichtbar, dass ein design durch den Wiederaufruf aus mehreren Modulen oder anderen Bestandteilen entstehen kann.

```

1     design:
2     module design |
3     defattr design |
4     task_func_decl design |
5     param_decl design |
6     localparam_decl design |
7     typedef_decl design |
8     package design |
9     interface design |
10    %empty;

```

7. Zuletzt wird design auf inout reduziert.

```

1     input: {
2     ast_stack.clear();
3     ast_stack.push_back(current_ast);
4 } design {
5     ast_stack.pop_back();
6     log_assert(GetSize(ast_stack) == 0);
7     for (auto &it : default_attr_list)
8         delete it.second;
9     default_attr_list.clear();
10 };

```

In einigen der Reduktionen wird mit AST-Strukturen gearbeitet. AST-Stack ist eine Datenstruktur vom Typ Vector:

```

1     std::vector<AstNode*> ast_stack;

```

Vektoren sind sequentielle Datenstrukturen wie Arrays, jedoch mit dynamischer Länge. Die Vektorstruktur wird hier genutzt, um einen Stack zu repräsentieren.

Ein Stack ist eine LIFO-Struktur, die als C-Vektor nach unten wächst. Das letzte Element was auf den Stack gelegt wird (push_back(objekt)) wird als erstes ausgegeben (back()). Es ist üblich, das erste Stackobjekt nach dem lesen zu entfernen (pop_back()).

Die Objekte, die auf diesen Stack gelegt werden, sind vom Typ `AstNode` und repräsentieren die Knoten des Abstrakten-Syntax-Baums. Hier ist ein Ausschnitt der Struktur.

```

1  struct AstNode
2  {
3      // for dict<> and pool<>
4      unsigned int hashidx_;
5      unsigned int hash() const { return hashidx_; }
6
7      // this nodes type
8      AstNodeType type;
9
10     // the list of child nodes for this node
11     std::vector<AstNode*> children;
12
13     // the list of attributes assigned to this node
14     std::map<RTLIL::IdString, AstNode*> attributes;
15     bool get_bool_attribute(RTLIL::IdString id);
16
17     // node content - most of it is unused in most node types
18     std::string str;
19     std::vector<RTLIL::State> bits;
20     bool is_input, is_output, is_reg, is_logic, is_signed,
21     is_string, is_wand, is_wor, range_valid, range_swapped,
22     was_checked, is_unsized, is_custom_type;
23     int port_id, range_left, range_right;
24     uint32_t integer;
25     double realvalue;
26     // set for IDs typed to an enumeration, not used
27     bool is_enum;

```

Gespeichert sind also Eigenschaften wie beispielsweise Typ, children (hierarchisch niedrigere Knoten im Baum), Attribute und Inhalt.

```

1  AstNode *mod = new AstNode(AST_MODULE);
2      ast_stack.back()->children.push_back(mod);
3      ast_stack.push_back(mod);

```

In der module-Regel wird ein neuer Knoten `AST_MODULE` erstellt. Das letzte Stackobjekt bekommt diesen Knoten als Children angehängen. Danach wird der Knoten selbst auf den Stack gelegt. `ast_stack` und `children` sind also separate Stacks. Jede Node hat einen stack children, wie bereits zuvor in der AST-Struktur zu sehen war.

```

1  current_ast_mod = mod;
2      port_stubs.clear();
3      port_counter = 0;
4      mod->str = *$4;
5      append_attr(mod, $1);
6      delete $4;

```

Im Yosys Manual wird folgender Parserbaum für ein assign Statement dargestellt:

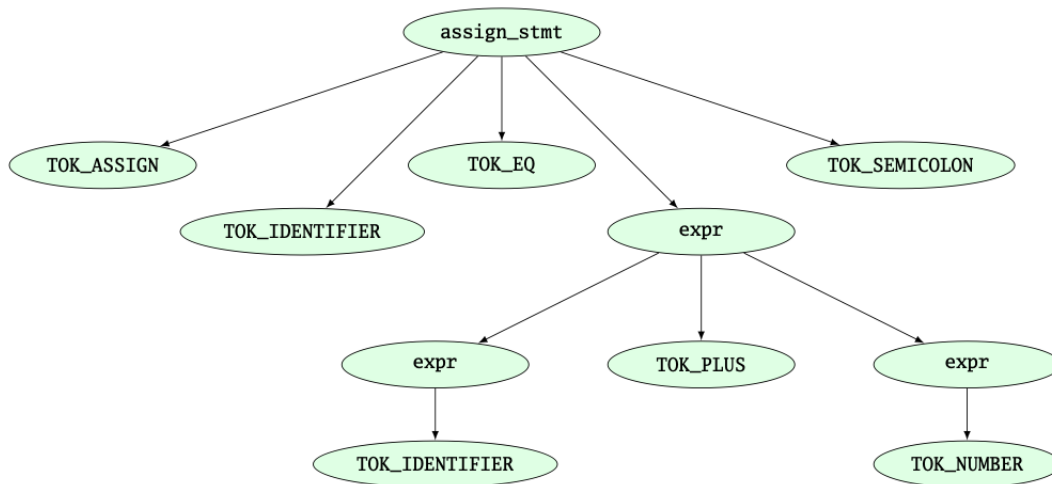


Figure 2.3: Example parse tree for the Verilog expression “**assign** foo = bar + 42;”.

Bei Betrachtung der Reduzierungen unter Beachtung der Parserregeln fällt jedoch auf, dass diese Grafik nicht mehr aktuell sein kann. Es sind einige Reduktionen notwendig, die hier nicht berücksichtigt werden. Ein Problem dabei sind Konflikte, da teilweise verschiedene Reduktionen möglich sind.

2.1.2 Interpretations-Probleme

1. Reduzierungskonflikte

Wie priorisiert der Parser Regeln. Woher kommt der Determinismus? Welche Rolle spielen der Parsertyp LALR (Look-Ahead-Left-Right) und die Bison precedence Regeln?

2. Stackaufbau

Wie wird der Masterstack `ast_stack` aufgebaut? Wann werden Objekte auf den Stapel gelegt?

Unter Vernachlässigung dieser Probleme wird die Parseranalyse an dieser Stelle beendet und es wird davon ausgegangen, dass AST-Stack Strukturen gebildet wurden.

Diese Strukturen werden von dem AST-Frontend weiterverarbeitet.

Kapitel 3

AST-Frontend

Das AST-Frontend überführt die AST-Strukturen in die interne Registertransfer-Level Sprache RTLIL. Laut Yosys-Manual findet die Überführung in RTLIL in zwei Schritten statt:

1. Vereinfachung (Simplification)
2. RTLIL Generation

In der Struktur `AstNode` in der Headerdatei `ast.h` ist eine Funktion für die Vereinfachung deklariert:

```
1  bool simplify(bool const_fold, bool at_zero,
2  bool in_lvalue, int stage, int width_hint,
3  bool sign_hint, bool in_param);
```

Laut Manual werden folgende Vereinfachungen durchgeführt:

1. Inline all task and function calls.
2. Evaluate all generate-statements and unroll all for-loops.

Generate-Blöcke erzeugen mehrere Instanzen eines Moduls oder ermöglichen die bedingte Instanziierung eines Moduls. Hier werden also alle Instanzen evaluiert.
For-Schleifen werden abgewickelt, indem die Iterationsanzahl verringert wird

3. Perform const folding where it is necessary (e.g. in the value part of `AST_PARAMETER`, `AST_LOCALPARAM`, `AST_PARASET` and `AST_RANGE` nodes).

Folding (oder auch Falten) ist ein Prozess, bei dem die Anzahl von funktionalen Blöcken reduziert wird, indem Register und Multiplexer eingesetzt werden.

4. Replace `AST_PRIMITIVE` nodes with appropriate `AST_ASSIGN` nodes.

Verilog Primitives sind von folgender Form:

```
1  and (out, in1, in2, in3);
```

Diese Form entspricht nicht den gewöhnlichen assign-Statements, kann jedoch einfach umgewandelt werden. Dies vereinfacht die Konvertierung nach RTLIL, da die Primitives mit assign-Statements abgedeckt werden.

5. Replace dynamic bit ranges in the left-hand-side of assignments with AST_CASE nodes with AST_COND children for each possible case.

Dynamische Wortgrößen müssen deterministisch festgelegt werden. Dafür werden AST_CASE nodes für die Fallunterscheidung eingesetzt. Die children dieser Nodes sind die Fälle.

6. Detect array access patterns that are too complicated for the RTLIL::Memory abstraction and replace them with a set of signals and cases for all reads and/or writes.

Komplizierte Speicherzugriffsmuster werden mit Signalen und Fällen für alle Lese- und Schreibzugriffe ersetzt.

7. Otherwise replace array accesses with AST_MEMRD and AST_MEMWR nodes.

Standardbehandlung für Lese- und Schreibzugriffe auf Speicher ist mit AST_MEMRD und AST_MEMWR Nodes.

3.1 Analyse der simplify Funktion

Die simplify-Funktion ist in der Datei simplify.cc enthalten.

```

1  static int recursion_counter = 0;
2  static bool deep_recursion_warning = false;
3
4  if (recursion_counter++ == 1000 && deep_recursion_warning) {
5      log_warning("Deep recursion in AST simplifier.\n
6      Does this design contain insanely long expressions?\n");
7      deep_recursion_warning = false;
8  }
9
10 AstNode *newNode = NULL;
11 bool did_something = false;
```

Der Anfang der Funktion beinhaltet einen recursion counter, der bei Aufruf der Bedingung inkrementiert wird. Die Vereinfachungen scheinen also rekursiv aufgerufen zu werden. Ab 1000 Aufrufen zählt dies hier als *deep recursion* (tiefe Rekursion).

Es werden die angesprochenen Vereinfachungen im Quellcode gesucht.

```

1 // unroll for loops and generate-for blocks
2 if ((type == AST_GENFOR || type == AST_FOR) && children.size() != 0)
3 {
4     AstNode *init_ast = children[0];
5     AstNode *while_ast = children[1];
6     AstNode *next_ast = children[2];
7     AstNode *body_ast = children[3];
8
9     while (body_ast->type == AST_GENBLOCK &&
10    body_ast->str.empty() &&
11    body_ast->children.size() == 1 &&
12    body_ast->children.at(0)->type == AST_GENBLOCK)
13    body_ast = body_ast->children.at(0);
```

Ein generate-for Block generiert Instanzen eines Moduls bei jedem Aufruf der for-Schleife.

Die Reihenfolge der children jeder AST-Node scheint eine große Bedeutung zu haben, da diese Nodes in Kategorien eingeteilt werden.

Das Ausrollen geschieht durch das Auswerten von drei Ausdrücken.

```

1  // eval 1st expression
2  AstNode *varbuf = init_ast->children[1]->clone();
3  {
4      int expr_width_hint = -1;
5      bool expr_sign_hint = true;
6      varbuf->detectSignWidth(expr_width_hint, expr_sign_hint);
7      while (varbuf->
8          simplify(true, false, false, stage, 32, true, false)) { }
9      }
10
11 if (varbuf->type != AST_CONSTANT)
12     log_file_error(filename, linenum, "Right hand side of
13     1st expression of generate for-loop is not constant!\n");
14
15     varbuf = new AstNode(AST_LOCALPARAM, varbuf);
16     varbuf->str = init_ast->children[0]->str;
17
18     AstNode *backup_scope_varbuf = current_scope[varbuf->str];
19     current_scope[varbuf->str] = varbuf;
20
21     size_t current_block_idx = 0;
22     if (type == AST_FOR) {
23         while (current_block_idx < current_block->children.size() &&
24             current_block->
25             children[current_block_idx] != current_block_child)
26             current_block_idx++;
27     }

```

Die Breite oder hier *Width* eines Signals ist Teil der Range.

3.2 Analyse der genRTLIL Funktion

genRTLIL ist Teil der Strukturdefinition für Syntaxbaum-Knoten AstNode.

Die genRTLIL Funktion besteht zu einem Teil aus einer großen Case-Bedingung.

```

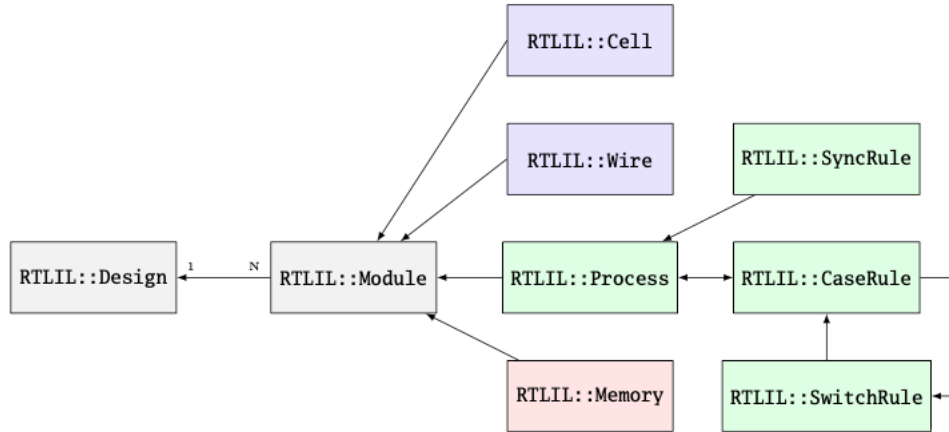
1  switch (type)
2  {
3      // simply ignore this nodes.
4      // they are either leftovers from simplify() or
5      // are referenced by other nodes
6      // and are only accessed here thru this references
7      case AST_NONE:
8      case AST_TASK:
9      case AST_FUNCTION:
10     case AST_DPI_FUNCTION:
11     case AST_AUTOWIRE:
12     case AST_DEFPARAM:
13     case AST_GENVAR:
14     case AST_GENFOR:
15     case AST_GENBLOCK:
16     case AST_GENIF:
17     case AST_GENCASE:
18     case AST_PACKAGE:
19     case AST_MODPORT:
20     case AST_MODPORTMEMBER:
21         break;

```


Diese Blöcke werden ignoriert, da sie Überbleibsel der simplify-Funktion sind oder mit anderen Knoten behandelt werden. Hier ist auffällig, dass unter anderem die Generate-Blöcke enthalten sind, die in der simplify-Funktion ausgerollt werden.

3.3 Prozessgenerator

Die gesamte Prozessstruktur sieht so aus:



Ein Objekt *proc* vom Typ `RTLIL::Process` wird bei Ausführung der Funktion *ProcessGenerator()* erstellt. Die Funktion wird ausgeführt, wenn die *genRTLIL* Funktion auf eine AST-Node des Typs *AST_ALWAYS* oder *AST_INITIAL* trifft:

```

1  // use ProcessGenerator for always blocks
2  case AST_ALWAYS: {
3      AstNode *always = this->clone();
4      ProcessGenerator generator(always);
5      ignoreThisSignalsInInitial.append(generator.outputSignals);
6      delete always;
7  } break;
8
9  case AST_INITIAL: {
10     AstNode *always = this->clone();
11     ProcessGenerator generator(always, ignoreThisSignalsInInitial);
12     delete always;
13 } break;

```

Der Datentyp `RTLIL::Process` besitzt die Elemente aus der Grafik:

```

1 struct RTLIL::Process : public RTLIL::AttrObject
2 {
3     RTLIL::IdString name;
4     RTLIL::CaseRule root_case;
5     std::vector<RTLIL::SyncRule*> syncs;
6
7     ~Process();
8
9     template<typename T> void rewrite_sigspecs(T &functor);
10    template<typename T> void rewrite_sigspecs2(T &functor);
11    RTLIL::Process *clone() const;
12 };

```

Ein `RTLIL::CaseRule` Objekt für den hierarchisch höchsten *root_case*, sowie einen Vektorstapel *syncs* für Synchronisationsregeln mit dem Takt. Jedes `RTLIL::CaseRule` Objekt besitzt ein oder

mehrere RTLIL::SwitchRule Objekte in einem Vektorstapel, einen Vektor *actions* für Zuweisungen (später detaillierter) sowie einen Vektor mit Signalen *compare*.

Die RTLIL::SyncRule Objekte enthalten die Zuweisungen, die getaktet erfolgen, in einem Vektorstapel *actions*:

```

1 struct RTLIL::SyncRule
2 {
3     RTLIL::SyncType type;
4     RTLIL::SigSpec signal;
5     std::vector<RTLIL::SigSig> actions;
6
7     template<typename T> void rewrite_sigspecs(T &functor);
8     template<typename T> void rewrite_sigspecs2(T &functor);
9     RTLIL::SyncRule *clone() const;
10 };

```

Der Vektorstapel *actions* existiert auch in RTLIL::CaseRule.

Laut Yosys-Manual passiert zuerst folgendes: *On startup the ProcessGenerator generates a new RTLIL::Process object with an empty root case and initializes its state variables as described above. Then the RTLIL::SyncRule objects are created using the synchronization events from the AST_ALWAYS node and the initial values of subst_lvalue_from and subst_lvalue_to. Then the AST for this process is evaluated recursively.*

Dies deckt sich mit dem Quellcode:

```

1 RTLIL::CaseRule *current_case;
2 stackmap<RTLIL::SigBit, RTLIL::SigBit> subst_rvalue_map;
3 stackmap<RTLIL::SigBit, RTLIL::SigBit> subst_lvalue_map;

```

Später:

```

1 // create syncs for the process
2 bool found_clocked_sync = false;
3 for (auto child : always->children)
4     if (child->type == AST_POSEDGE || child->type == AST_NEGEDGE) {
5         if (GetSize(child->children) == 1 &&
6             child->children.at(0)->type ==
7             AST_IDENTIFIER && child->children.at(0)->id2ast &&
8             child->children.at(0)->id2ast->type == AST_WIRE &&
9             child->children.at(0)->id2ast->get_bool_attribute("\\gclk"))
10            continue;
11         found_clocked_sync = true;
12         if (found_global_syncs || found_anyedge_syncs)
13             log_file_error(always->filename, always->linenum,
14                 "Found non-synthesizable event list!\n");
15         RTLIL::SyncRule *syncrule = new RTLIL::SyncRule;
16         syncrule->type = child->type == AST_POSEDGE ? RTLIL::STp : RTLIL::STn;
17         syncrule->signal = child->children[0]->genRTLIL();
18         if (GetSize(syncrule->signal) != 1)
19             log_file_error(always->filename, always->linenum,
20                 "Found posedge negedge event on a signal that is not 1 bit wide!\n");
21         addChunkActions(syncrule->actions, subst_lvalue_from,
22             subst_lvalue_to, true);
23         proc->syncs.push_back(syncrule);
24     }

```

Hier ist sichtbar, dass die Funktion *addChunkActions* hier zu dem Vektorstapel *actions* die Zuweisungen aus *subst_lvalue_from* und *subst_lvalue_to* hinzufügt:

```

1 void addChunkActions(std::vector<RTLIL::SigSig> &actions, RTLIL::SigSpec lvalue,
    RTLIL::SigSpec rvalue, bool inSyncRule = false)

```

```

2  {
3      if (inSyncRule && initSyncSignals.size() > 0) {
4          init_lvalue.append(lvalue.extract(initSyncSignals));
5          init_rvalue.append(lvalue.extract(initSyncSignals, &rvalue));
6          lvalue.remove2(initSyncSignals, &rvalue);
7      }
8      log_assert(lvalue.size() == rvalue.size());
9
10     int offset = 0;
11     for (auto &lvalue_c : lvalue.chunks()) {
12         RTLIL::SigSpec lhs = lvalue_c;
13         RTLIL::SigSpec rhs = rvalue.extract(offset, lvalue_c.width);
14         if (inSyncRule && lvalue_c.wire && lvalue_c.wire->get_bool_attribute("\\nosync"))
15             rhs = RTLIL::SigSpec(RTLIL::State::Sx, rhs.size());
16         remove_unwanted_lvalue_bits(lhs, rhs);
17         actions.push_back(RTLIL::SigSig(lhs, rhs));
18         offset += lhs.size();
19     }
20 }

```

Die Zuweisung besteht aus zwei Signalen *SigSpec*, die zu einer Signalverknüpfung *SigSig* werden, die auf den Stapel *actions*, der als Parameter übergeben wird, gelegt wird.

Ob der *actions*-Stapel aus der SyncRule oder aus der CaseRule stammt, hängt davon ab, ob die Zuweisung innerhalb eines Prozesses (getaktet oder Synchronisation mit Signal) oder außerhalb stattfindet. Beispielhaft ist eine Zuweisung mit Inertialwerten, die ohne Synchronisation stattfindet:

```

1 // create initial assignments for the temporary signals
2 if ((flag_nolatches || always->get_bool_attribute("\\nolatches") ||
3     current_module->get_bool_attribute("\\nolatches")) && !found_clocked_sync) {
4     subst_rvalue_map = subst_lvalue_from.to_sigbit_dict(RTLIL::SigSpec(RTLIL::
5     State::Sx, GetSize(subst_lvalue_from)));
6 } else {
7     addChunkActions(current_case->actions, subst_lvalue_to, subst_lvalue_from)
8 ;
9 }

```

Wo wird der *actions*-Stapel weiterverarbeitet?

3.3.1 Datenstrukturen und Variablen

1. *lvalue*

Bekommt Rückgabe der `genRTLIL()` Funktion zugewiesen. Rückgabotyp ist `RTLIL::SigSpec`

2. *SigSpec* Aus dem Manual: *The RTLIL::SigSpec data type is used to represent signals. The RTLIL::Cell object contains one RTLIL::SigSpec for each cell port. In addition, connections between wires are represented using a pair of RTLIL::SigSpec objects. Such pairs are needed in different locations. Therefore the type name RTLIL::SigSig was defined for such a pair.*

```

1 struct RTLIL::SigSpec
2 {
3     private:
4         int width_;
5         unsigned long hash_;
6         std::vector<RTLIL::SigChunk> chunks_; // LSB at index 0
7         std::vector<RTLIL::SigBit> bits_; // LSB at index 0
8
9         void pack() const;
10        void unpack() const;

```

```

11 void updhash() const;
12
13 inline bool packed() const {
14     return bits_.empty();
15 }
16
17 inline void inline_unpack() const {
18     if (!chunks_.empty())
19         unpack();
20 }

```

...

SigSpec besitzt remove Funktionen, um ein Muster zu löschen. Das Muster kann aus einzelnen Signal-Bits SigBit, Signalen SigSpecs, die generell Vektoren aus Sigbits sind und Mischungen bestehen.

```

1 void replace(const RTLIL::SigSpec &pattern, const RTLIL::SigSpec &with);
2 void replace(const RTLIL::SigSpec &pattern, const RTLIL::SigSpec &with,
   RTLIL::SigSpec *other) const;
3
4 void replace(const dict<RTLIL::SigBit, RTLIL::SigBit> &rules);
5 void replace(const dict<RTLIL::SigBit, RTLIL::SigBit> &rules, RTLIL::
   SigSpec *other) const;
6
7 void replace(const std::map<RTLIL::SigBit, RTLIL::SigBit> &rules);
8 void replace(const std::map<RTLIL::SigBit, RTLIL::SigBit> &rules, RTLIL::
   SigSpec *other) const;
9
10 void replace(int offset, const RTLIL::SigSpec &with);
11
12 void remove(const RTLIL::SigSpec &pattern);
13 void remove(const RTLIL::SigSpec &pattern, RTLIL::SigSpec *other) const;
14 void remove2(const RTLIL::SigSpec &pattern, RTLIL::SigSpec *other);
15
16 void remove(const pool<RTLIL::SigBit> &pattern);
17 void remove(const pool<RTLIL::SigBit> &pattern, RTLIL::SigSpec *other)
   const;
18 void remove2(const pool<RTLIL::SigBit> &pattern, RTLIL::SigSpec *other);
19 void remove2(const std::set<RTLIL::SigBit> &pattern, RTLIL::SigSpec *
   other);

```

3. *SigBit* Ein einzelnes Bit eines Signals. Ist als Vektor in *SigSpec* enthalten.

```

1 struct RTLIL::SigBit
2 {
3     RTLIL::Wire *wire;
4     union {
5         RTLIL::State data; // used if wire == NULL
6         int offset;        // used if wire != NULL
7     };
8
9     SigBit();
10    SigBit(RTLIL::State bit);
11    SigBit(bool bit);
12    SigBit(RTLIL::Wire *wire);
13    SigBit(RTLIL::Wire *wire, int offset);
14    SigBit(const RTLIL::SigChunk &chunk);
15    SigBit(const RTLIL::SigChunk &chunk, int index);
16    SigBit(const RTLIL::SigSpec &sig);
17    SigBit(const RTLIL::SigBit &sigbit);

```

```

18 RTLIL::SigBit &operator =(const RTLIL::SigBit &other) = default;
19
20 bool operator <(const RTLIL::SigBit &other) const;
21 bool operator ==(const RTLIL::SigBit &other) const;
22 bool operator !=(const RTLIL::SigBit &other) const;
23 unsigned int hash() const;
24 };

```

Ein SigBit besteht aus einem *RTLIL::State* Objekt für die logischen Zustände 0,1,X,Z, *don't care* und dem internen Zustand *marker*.

```

1 namespace RTLIL
2 {
3     enum State : unsigned char {
4         S0 = 0,
5         S1 = 1,
6         Sx = 2, // undefined value or conflict
7         Sz = 3, // high-impedance / not-connected
8         Sa = 4, // don't care (used only in cases)
9         Sm = 5 // marker (used internally by some passes)
10 };

```

4. *RTLIL::SigChunk*

```

1 struct RTLIL::SigChunk
2 {
3     RTLIL::Wire *wire;
4     std::vector<RTLIL::State> data; // only used if wire == NULL, LSB at
        index 0
5     int width, offset;
6
7     SigChunk();
8     SigChunk(const RTLIL::Const &value);
9     SigChunk(RTLIL::Wire *wire);
10    SigChunk(RTLIL::Wire *wire, int offset, int width = 1);
11    SigChunk(const std::string &str);
12    SigChunk(int val, int width = 32);
13    SigChunk(RTLIL::State bit, int width = 1);
14    SigChunk(RTLIL::SigBit bit);
15    SigChunk(const RTLIL::SigChunk &sigchunk);
16    RTLIL::SigChunk &operator =(const RTLIL::SigChunk &other) = default;
17
18    RTLIL::SigChunk extract(int offset, int length) const;
19    inline int size() const { return width; }
20
21    bool operator <(const RTLIL::SigChunk &other) const;
22    bool operator ==(const RTLIL::SigChunk &other) const;
23    bool operator !=(const RTLIL::SigChunk &other) const;
24 };

```

Ein Ausschnitt eines Signals mit bestimmter Länge *width* und *offset*.

5. *compare* Vektor in *RTLIL::CaseRule* In der *CaseRule* ist dieser Vektor enthalten. Vermutlich werden dort Signale für den Vergleich mit der *RTLIL::SwitchRule* gespeichert.

```

1 struct RTLIL::CaseRule : public RTLIL::AttrObject
2 {
3     std::vector<RTLIL::SigSpec> compare;

```

In der *RTLIL::SwitchRule* ist ein Signal als *RTLIL::SigSpec* Objekt gespeichert:

```

1 struct RTLIL::SwitchRule : public RTLIL::AttrObject
2 {
3     RTLIL::SigSpec signal;

```

6. Ersetzungsmuster in *subst_rvalue_map* und *subst_lvalue_map*.

```

1 // This map contains the replacement pattern to be used in the right hand
2 // side
3 // of an assignment. E.g. in the code "foo = bar; foo = func(foo);" the
4 // foo in the right
5 // hand side of the 2nd assignment needs to be replace with the temporary
6 // signal holding
7 // the value assigned in the first assignment. So when the first
8 // assignment is processed
9 // the according information is appended to subst_rvalue_from and
10 // subst_rvalue_to.
11 stackmap<RTLIL::SigBit, RTLIL::SigBit> subst_rvalue_map;
12
13 // This map contains the replacement pattern to be used in the left hand
14 // side
15 // of an assignment. E.g. in the code "always @(posedge clk) foo <= bar"
16 // the signal bar
17 // should not be connected to the signal foo. Instead it must be
18 // connected to the temporary
19 // signal that is used as input for the register that drives the signal
20 // foo.
21 stackmap<RTLIL::SigBit, RTLIL::SigBit> subst_lvalue_map;

```

Stackmaps (Yosys-Customs) sind ähnlich wie Maps in C++ mit der Erweiterung, dass Zustände gespeichert und wiederhergestellt werden können. Elemente können mit einem Schlüssel gespeichert werden und mit diesem Schlüssel wieder ausgelesen werden (vgl. Hashmap).

7. *current_state*

Ein konstantes Hashlib *dict* Objekt. Ein Dictionary ist ähnlich wie eine Hashmap. Es können jedoch nur Paare von gleichem Datentyp gespeichert werden. Ein Dictionary ist jedoch schneller und behält seine Ordnung.

8. *current_case*

```

1 struct RTLIL::CaseRule : public RTLIL::AttrObject
2 {
3     std::vector<RTLIL::SigSpec> compare;
4     std::vector<RTLIL::SigSig> actions;
5     std::vector<RTLIL::SwitchRule*> switches;
6
7     ~CaseRule();
8     void optimize();
9
10    bool empty() const;
11
12    template<typename T> void rewrite_sigspecs(T &functor);
13    template<typename T> void rewrite_sigspecs2(T &functor);
14    RTLIL::CaseRule *clone() const;
15 };

```

Laut Manual der aktuelle Case der gerade gefüllt wird. Actions sind Zuweisungen. Ein Case kann außerdem weitere Switches besitzen. Wenn die Methode *ProcessGenerator()* ausgeführt wird, wird der root-case der Anweisung zum *current_case*.

9. *SigSig*

Verbindungen von zwei Signalen werden als Paar aus zwei SigSpec Signalen SigSig implementiert:

```
1 typedef std::pair<SigSpec, SigSpec> SigSig;
```

10. *pool*

Laut Manual:

Pool ist eine Container Datenstruktur und ersetzt std::unordered_set.

3.3.2 Non-Blocking Assignments

Die Behandlung von Non-Blocking Assignments wird im Yosys Manual beschrieben. Hier werden die entsprechenden Einträge mit dem Quellcode aus der genRTLIL.cc Datei verglichen:

When an AST_ASSIGN_LE node is discovered, the following actions are performed by the ProcessGenerator:

1. The left-hand-side is evaluated using AST::AstNode::genRTLIL() and mapped to a temporary signal name using subst_lvalue.from and subst_lvalue.to.

```
1 case AST_ASSIGN_LE:
2 {
3 RTLIL::SigSpec unmapped_lvalue =
4 ast->children[0]->genRTLIL(), lvalue = unmapped_lvalue;
5
6
```

Das nullte Objekt auf dem children Stack einer ASSIGN Node ist die linke Seite einer Zuweisung. Nochmal zur Erinnerung: Der AstNode Struktur werden folgende Parameter übergeben:

```
1 AstNode::AstNode(AstNodeType type, AstNode *child1,
2 AstNode *child2, AstNode *child3)
```

Daraus wird der children-Stack befüllt:

```
1 if (child1)
2     children.push_back(child1);
3 if (child2)
4     children.push_back(child2);
5 if (child3)
6     children.push_back(child3);
```

Im Parser werden die Werte der linken und rechten Seite der Zuweisung wie folgt übergeben:

```
1 assign_expr:
2 lvalue '=' expr {
3     ast_stack.back()->
4     children.push_back(new AstNode(AST_ASSIGN, $1, $3));
5 }
```

Child1 bekommt den Wert von lvalue, Child2 den Wert von expr.

2. The right-hand-side is evaluated using AST::AstNode::genRTLIL(). For this call, the values of subst_rvalue_from and subst_rvalue_to are used to map blocking-assigned signals correctly.

```
1 RTLIL::SigSpec rvalue =
2 ast->children[1]->genWidthRTLIL(lvalue.size(),
3 &subst_rvalue_map.stdmap());
```

Da eine Zuweisung nur möglich ist, wenn der Ausdruck auf der rechten Seite die gleiche Wortbreite wie der Ausdruck auf der linken Seite hat, wird hier ein Wrapper der RTLIL Funktion verwendet, der mit einem Parameter für die Breite (Width) Rücksicht darauf nimmt. Als Parameter wird `lvalue.size()`, also die Breite der linken Seite übergeben.

`subst_rvalue_map` ist eine `stackmap`:

```
1 stackmap<RTLIL::SigBit, RTLIL::SigBit> subst_rvalue_map;
```

`Stackmap` ist ein eigener Datentyp, der wie eine `map` funktioniert wobei zusätzlich der aktuelle Zustand gespeichert und wiederhergestellt werden kann.

3. *Remove all assignments to the same left-hand-side as this assignment from the current_case and all cases within it.*

Alle vorherigen Zuweisungen werden bei einem Non-Blocking Assignment überschrieben.

```
1 removeSignalFromCaseTree(lvalue, current_case);
```

Die Funktion `removeSignalFromCaseTree()` löscht alle Zuweisungen in einem `root-case` und seinen Ableitungen. Beispiel: `a = 23; if (b) a = 42; a = 0;` Die ersten beiden Zuweisungen werden überschrieben. Die dritte Zuweisung wird ausgeführt. Quellcode der Funktion:

```
1 void removeSignalFromCaseTree(const RTLIL::SigSpec &pattern, RTLIL::
   CaseRule *cs)
2 {
3     for (auto it = cs->actions.begin(); it != cs->actions.end(); it++)
4         it->first.remove2(pattern, &it->second);
5
6     for (auto it = cs->switches.begin(); it != cs->switches.end(); it++)
7         for (auto it2 = (*it)->cases.begin(); it2 != (*it)->cases.end(); it2
          ++)
8             removeSignalFromCaseTree(pattern, *it2);
9 }
```

Der Vektorstapel `actions` wird durchiteriert. Die `RTLIL::SigSpec` Objekte sind C++ `std::pair` Objekte mit den beiden verknüpften Signalen als Paar. Hier wird also mit `first` auf das erste Paarobjekt zugegriffen und darauf die `RTLIL::SigSpec remove2()` Funktion angewendet.

Die Funktion `remove2()`:

```
1 void RTLIL::SigSpec::remove2(const RTLIL::SigSpec &pattern, RTLIL::SigSpec
   *other)
2 {
3     if (other)
4         cover("kernel.rtlil.sigspec.remove_other");
5     else
6         cover("kernel.rtlil.sigspec.remove");
7
8     unpack();
9     if (other != NULL) {
10         log_assert(width_ == other->width_);
11         other->unpack();
12     }
13
14     for (int i = GetSize(bits_) - 1; i >= 0; i--)
15     {
16         if (bits_[i].wire == NULL) continue;
17
18         for (auto &pattern_chunk : pattern.chunks())
```



```

19     if (bits_[i].wire == pattern_chunk.wire &&
20         bits_[i].offset >= pattern_chunk.offset &&
21         bits_[i].offset < pattern_chunk.offset + pattern_chunk.width) {
22         bits_.erase(bits_.begin() + i);
23         width_--;
24         if (other != NULL) {
25             other->bits_.erase(other->bits_.begin() + i);
26             other->width_--;
27         }
28         break;
29     }
30 }
31
32 check();
33 }

```

Die Funktion *unpack()* erstellt einen Bitvektor aus dem aktuellen SigSpec Objekt:

```

1 void RTLIL::SigSpec::unpack() const
2 {
3     RTLIL::SigSpec *that = (RTLIL::SigSpec*)this;
4
5     if (that->chunks_.empty())
6         return;
7
8     cover("kernel.rtlil.sigspec.convert.unpack");
9     log_assert(that->bits_.empty());
10
11     that->bits_.reserve(that->width_);
12     for (auto &c : that->chunks_)
13         for (int i = 0; i < c.width; i++)
14             that->bits_.push_back(RTLIL::SigBit(c, i));
15
16     that->chunks_.clear();
17     that->hash_ = 0;
18 }

```

Iteriert durch *chunks_* Stapel. Wenn der Stapel leer ist, wird nichts zurück gegeben. Ansonsten werden die *RTLIL::SigChunk* Objekte, die auf dem Stapel liegen, auf die Länge ihrer Wortbreite iteriert, und auf einen Stapel *bits_* des aktuellen *RTLIL::SigSpec* Objekts gelegt.

Danach wird der Bitvektor durchiteriert. Die Objekte, die auf dem Vektorstapel liegen, sind vom Typ *RTLIL::SigBit*. Einzelne Bits eines Signals können einem Verilog *Wire* zugewiesen werden. Daher besitzt ein *RTLIL::SigBit* Objekt die Möglichkeit, es mit einem *Wire* zu initialisieren.

Beim Löschen der vorherigen Zuweisungen wird verglichen, ob Wert, Offset und Offset + Wortbreite identisch mit einem Pattern sind, das im Methodenkopf der *remove* Funktion übergeben wird. Mit *_chunk* wird auf Das Pattern, das übergeben wird, ist *lvalue*, also das Signal der linken Zuweisungsseite, dessen ältere Zuweisungen gelöscht werden sollen. Sofern dies zutrifft, werden mit der *erase()* Funktion des *bits_*-Stapels

Zusammenfassung:

- Zuweisungsstapel des aktuellen Falls *actions* durchiterieren
- Auf Objekte des Stapels zugreifen (*RTLIL::SigSig* Objekte)
- Zweiter Stapel *bits_* enthält die Bits des Signals von *first*, also der linken Seite.
- Mit Pattern, hier *lvalue*, vergleichen

- (e) Wenn Pattern-Inhalt gleich *bits_* von linker Zuweisungsseite der Stapelobjekte, dann Löschen der relevanten Bits und Wortlänge verkürzen.

Also: Zuweisungen mit gleichem *first* Objekt wie lvalue löschen durch vergleichen der actions Stapel Objekte mit lvalue!

4. Add the new assignment to the *current_case*.

```
1 current_case->actions.push_back(RTLIL::SigSig(lvalue, rvalue));
```

Die *RTLIL::SigSpec* Objekte der Zuweisung werden zu *RTLIL::SigSig* verknüpft und auf den Stapel *actions* von *current_case* gelegt.

3.3.3 Blocking-Assignments

When an *AST_ASSIGN_EQ* node is discovered, the following actions are performed by the *ProcessGenerator*:

1. Perform all the steps that would be performed for a nonblocking assignment (see above).

```
1 case AST_ASSIGN_EQ :
2 case AST_ASSIGN_LE :
```

Da kein break; Abbruch der case-Bedingung erfolgt, geht das Programm in den nachfolgenden Case-Fall, der das Non-Blocking-Assignment behandelt.

2. Remove the found left-hand-side (before lvalue mapping) from *subst_rvalue_from* and also remove the respective bits from *subst_rvalue_to*.
3. Append the found left-hand-side (before lvalue mapping) to *subst_rvalue_from* and append the found right-hand-side to *subst_rvalue_to*.

In dem *AST_ASSIGN_LE* Case ist eine if-Bedingung für die Behandlung des Blocking Assignments.

```
1 if (ast->type == AST_ASSIGN_EQ) {
2     for (int i = 0; i < GetSize(unmapped_lvalue); i++)
3         subst_rvalue_map.set(unmapped_lvalue[i], rvalue[i]);
4 }
```

Bevor temporäre Signale gelöscht werden, wie es bei einem Non-Blocking Assignment der Fall ist, werden diese zwischengespeichert.

3.3.4 Cases und if-Anweisungen

When an *AST_CASE* node is discovered, the following actions are performed by the *ProcessGenerator*:

1. The values of *subst_rvalue_from*, *subst_rvalue_to*, *subst_lvalue_from* and *subst_lvalue_to* are pushed to the stack.
2. A new *RTLIL::SwitchRule* object is generated, the selection expression is evaluated using *AST::AstNode::genRTLIL()* (with the use of *subst_rvalue_from* and *subst_rvalue_to*) and added to the *RTLIL::SwitchRule* object and the object is added to the *current_case*.

```

1  case AST_CASE:
2      {
3          RTLIL::SwitchRule *sw = new RTLIL::SwitchRule;
4          sw->attributes["\\src"] = stringf("%s:%d", ast->filename.c_str(),
5          ast->linenum);
6          sw->signal = ast->children[0]->genWidthRTLIL(-1, &subst_rvalue_map.
7          stdmap());
8          current_case->switches.push_back(sw);
9      }

```

Ein SwitchRule Objekt sw wird erstellt. Die SwitchRule Struktur sieht so aus:

```

1  struct RTLIL::SwitchRule : public RTLIL::AttrObject
2  {
3      RTLIL::SigSpec signal;
4      std::vector<RTLIL::CaseRule*> cases;
5
6      ~SwitchRule();
7
8      bool empty() const;
9
10     template<typename T> void rewrite_sigspecs(T &functor);
11     template<typename T> void rewrite_sigspecs2(T &functor);
12     RTLIL::SwitchRule *clone() const;
13 };
14

```

Die SwitchRule Struktur besitzt also ein Signal und einen Vektor (Stapel) cases. Das Signal bekommt den RTLIL Code von subst_rvalue_map.stdmap(), die Ersetzung für Blocking-Assignments.

3. All lvalues assigned to within the AST_CASE node using blocking assignments are collected and saved in the local variable *this_case_eq_lvalue*.

```

1 RTLIL::SigSpec this_case_eq_lvalue;
2 collect_lvalues(this_case_eq_lvalue, ast, true, false);

```

Erinnerung: AST_ASSIGN_EQ sind AstNodes für Blocking-Assignments. Der Funktion collect_lvalues hat Parameter für Blocking- und Non-Blocking Assignments:

```

1 void collect_lvalues(RTLIL::SigSpec &reg, AstNode *ast, bool type_eq, bool
2 type_le, bool run_sort_and_unify = true)

```

Über den Boolean *type_eq* kann also übergeben werden, dass nur lvalues von Blocking-Assignments gesammelt werden sollen.

4. New temporary signals are generated for all signals in *this_case_eq_lvalue* and stored in *this_case_eq_ltemp*.

```

1 RTLIL::SigSpec this_case_eq_ltemp = new_temp_signal(this_case_eq_lvalue);

```

5. The signals in *this_case_eq_lvalue* are mapped using *subst_rvalue_from* and *subst_rvalue_to* and the resulting set of signals is stored in *this_case_eq_rvalue*.

```

1 RTLIL::SigSpec this_case_eq_rvalue = this_case_eq_lvalue;
2 this_case_eq_rvalue.replace(subst_rvalue_map.stdmap());

```

Then the following steps are performed for each AST_COND node within the AST_CASE node:

1. Set *subst_rvalue_from*, *subst_rvalue_to*, *subst_lvalue_from* and *subst_lvalue_to* to the values that have been pushed to the stack.
2. Remove *this_case_eq_lvalue* from *subst_lvalue_from* *subst_lvalue_to*.
3. Append *this_case_eq_lvalue* to *subst_lvalue_from* and append *this_case_eq_ltemp* to *subst_lvalue_to*.

```
1 for (int i = 0; i < GetSize(this_case_eq_lvalue); i++)
2   subst_lvalue_map.set(this_case_eq_lvalue[i], this_case_eq_ltemp[i]);
```

4. Push the value of *current_case*.
5. Create a new *RTLIL::CaseRule*. Set *current_case* to the new object and add the new object to the *RTLIL::SwitchRule* created above.

```
1 current_case = new RTLIL::CaseRule;

1 if (default_case != current_case)
2   sw->cases.push_back(current_case);
```

6. Add an assignment from *this_case_eq_rvalue* to *this_case_eq_ltemp* to the new *current_case*.

```
1 addChunkActions(current_case->actions, this_case_eq_ltemp,
   this_case_eq_rvalue);
```

7. Evaluate the compare value for this case using *AST::AstNode::genRTLIL()* (with the use of *subst_rvalue_from* and *subst_rvalue_to*) modify the new *current_case* accordingly.

```
1 current_case->compare.push_back(node->genWidthRTLIL(sw->signal.size(), &
   subst_rvalue_map.stdmap()));
```

8. Recursion into the children of the *AST_COND* node.

```
1 for (auto child : ast->children)
2   {
```

Oben genannte Aufgaben stehen in einem range-based for-loop (Iteration über *children* Vektor).

9. Restore *current_case* by popping the old value from the stack. Finally the following steps are performed:
10. The values of *subst_rvalue_from*, *subst_rvalue_to*, *subst_lvalue_from* and *subst_lvalue_to* are popped from the stack.
11. The signals from *this_case_eq_lvalue* are removed from the *subst_rvalue_from/subst_rvalue_to*-pair.
12. The value of *this_case_eq_lvalue* is appended to *subst_rvalue_from* and the value of *this_case_eq_ltemp* is appended to *subst_rvalue_to*.

```
1 for (int i = 0; i < GetSize(this_case_eq_lvalue); i++)
2   subst_rvalue_map.set(this_case_eq_lvalue[i], this_case_eq_ltemp[i]);
```

13. Map the signals in *this_case_eq_lvalue* using *subst_lvalue_from/subst_lvalue_to*.

14. Remove all assignments to signals in *this_case_eq_lvalue* in *current_case* and all cases within it.
15. Add an assignment from *this_case_eq_ltemp* to *this_case_eq_lvalue* to *current_case*.

```
1 addChunkActions(current_case->actions, this_case_eq_lvalue,
   this_case_eq_ltemp);
```

3.3.5 Proc Pass

Die genannten Prozesse sind in dieser Form noch nicht synthetisierbar. Prozesse werden in erster Linie von einem Behavioural Model in AST Ebene in ein Behavioural Model in RTLIL Ebene überführt. Das Prozessstiming erfolgt in der Realität beispielsweise mit Flipflops, die entsprechend eingesetzt werden müssen. Hierfür sind weitere Passes, also globale Aktionen, die auf das Design angewendet werden, erforderlich.

Diese werden in *proc.cc* aufgerufen:

```
1 Pass::call(design, "proc_clean");
2 if (!ifxmode)
3     Pass::call(design, "proc_rmdead");
4 Pass::call(design, "proc_init");
5 if (global_arst.empty())
6     Pass::call(design, "proc_arst");
7 else
8     Pass::call(design, "proc_arst -global_arst " + global_arst);
9 Pass::call(design, ifxmode ? "proc_mux -ifx" : "proc_mux");
10 Pass::call(design, "proc_dlatch");
11 Pass::call(design, "proc_dff");
12 Pass::call(design, "proc_clean");
```

Aus dem Manual zu den einzelnen Passes:

1. *proc_clean* and *proc_rmdead*

These two passes just clean up the RTLIL::Process structure. The proc_clean pass removes empty parts (eg. empty assignments) from the process and proc_rmdead detects and removes unreachable branches from the process's decision trees.

```
1 void proc_clean(RTLIL::Module *mod, RTLIL::Process *proc, int &total_count)
2 {
3     int count = 0;
4     bool did_something = true;
5     for (size_t i = 0; i < proc->syncs.size(); i++) {
6         for (size_t j = 0; j < proc->syncs[i]->actions.size(); j++)
7             if (proc->syncs[i]->actions[j].first.size() == 0)
8                 proc->syncs[i]->actions.erase(proc->syncs[i]->actions.begin() + (j
9                 --));
10         if (proc->syncs[i]->actions.size() == 0) {
11             delete proc->syncs[i];
12             proc->syncs.erase(proc->syncs.begin() + (i--));
13         }
14     }
15     while (did_something) {
16         did_something = false;
17         proc_clean_case(&proc->root_case, did_something, count, -1);
18     }
19     if (count > 0)
```

```

19     log("Found and cleaned up %d empty switch%s in '%s.%s'.\n", count,
        count == 1 ? "" : "es", mod->name.c_str(), proc->name.c_str());
20     total_count += count;
21 }

```

Der Stapel *syncs* wird von dem Prozessgenerator bei einer erkannten Synchronisation gefüllt:

```

1 addChunkActions(syncrule->actions, subst_lvalue_from, subst_lvalue_to, true
   );
2 proc->syncs.push_back(syncrule);

```

Auf den Stapel wird ein Objekt *syncrule* gelegt. Eine *syncrule* hat einen weiteren Vektorstapel *actions*. Die Funktion *addChunkActions* ermöglicht das Aufsplitten von Zuweisungen in Teile *chunks*, um natürlich große Multiplexer zu generieren.

Die *proc_clean* Funktion iteriert beide Vektoren *syncs* und *actions*. Jedes Objekt auf dem Vektor *syncs* hat einen Vektor *actions* mit Zuweisungen für jede Synchronisation. Daher werden mit einer verschachtelten for-schleife die *action*-Vektoren der *RTLIL::SyncRule* Objekte iteriert und, falls leere Zuweisung, das aktuelle Objekt (an Stelle *actions.begin()+j*) gelöscht. Zusätzlich wird für jedes *RTLIL::SyncRule* Objekt geprüft, ob der zugehörige *actions*-Vektor komplett leer ist (*size()==0*). Dann ist die Synchronisierung überflüssig und wird entfernt.

Der *proc_rmdead* Pass: Dieser Pass löscht unerreichbare Äste aus Entscheidungsbäumen. Hier müssen also *if*- und *case*-Anweisungen untersucht werden.

```

1 void proc_rmdead(RTLIL::SwitchRule *sw, int &counter, int &
   full_case_counter)
2 {
3     BitPatternPool pool(sw->signal);
4
5     for (size_t i = 0; i < sw->cases.size(); i++)
6     {
7         bool is_default = GetSize(sw->cases[i]->compare) == 0 && (!pool.empty()
           || GetSize(sw->signal) == 0);
8
9         for (size_t j = 0; j < sw->cases[i]->compare.size(); j++) {
10             RTLIL::SigSpec sig = sw->cases[i]->compare[j];
11             if (!sig.is_fully_const())
12                 continue;
13             if (!pool.take(sig))
14                 sw->cases[i]->compare.erase(sw->cases[i]->compare.begin() + (j--));
15         }
16
17         if (!is_default) {
18             if (sw->cases[i]->compare.size() == 0) {
19                 delete sw->cases[i];
20                 sw->cases.erase(sw->cases.begin() + (i--));
21                 counter++;
22                 continue;
23             }
24             // if (pool.empty())
25             // sw->cases[i]->compare.clear();
26         }
27
28         for (auto switch_it : sw->cases[i]->switches)
29             proc_rmdead(switch_it, counter, full_case_counter);

```

```

30
31     if (is_default)
32         pool.take_all();
33     }
34
35     if (pool.empty() && !sw->get_bool_attribute("\\full_case")) {
36         sw->set_bool_attribute("\\full_case");
37         full_case_counter++;
38     }
39 }

```

Der Vektor *cases* jedes *RTLIL::SwitchRule* Objekts wird iteriert. *bool is_default* ist *true*, wenn der Vektor *compare* leer ist, der *BitPatternPool* mit den Signalen des *SwitchRule* Objekts nicht leer ist, oder der Vektor mit Signalen leer ist *size_t* ist ein *unsigned* Datentyp, der nie kleiner als 0 werden kann. Daher wird er für die Indexierung von Arrays genutzt. Hier wird über den Vektor *compare* iteriert.

In *compare* werden Signale gespeichert, die mit dem Signal in der *RTLIL::SwitchRule* verglichen werden sollen.

Auffällig ist, dass das Löschen hier genau so funktioniert wie vorher. Wenn in dem Vektor *compare* an der auf der Stelle liegendem Signal ein Wert liegt (Funktion *is_fully_const()* gibt *false* zurück, *negiert* zu *true*, wird das Objekt übersprungen.

Ausschnitt der Funktion:

```

1 bool RTLIL::SigSpec::is_fully_const() const
2 {
3     cover("kernel.rtlil.sigspec.is_fully_const");
4
5     pack();
6     for (auto it = chunks_.begin(); it != chunks_.end(); it++)
7         if (it->width > 0 && it->wire != NULL)
8             return false;
9     return true;
10 }

```

Ansonsten wird, wenn *take()* *true* zurückgibt, das entsprechende Objekt aus dem *compare* Vektor gelöscht.

Bei dem Fall, dass der Vektor *compare* komplett leer ist (*compare.size() == 0*), kann das ganze *RTLIL::CaseRule* Objekt gelöscht werden. Hier muss jedoch vorher überprüft werden, ob der Fall nicht der *default*-Case ist;

Ein Verilog *case*-Statement hat folgende Struktur:

```

1 case (<expression>)
2     case_item1 : <single statement>
3     case_item2,
4     case_item3 : <single statement>
5     case_item4 : begin
6                     <multiple statements>
7                     end
8     default    : <statement>
9 endcase

```

Hier ist sichtbar, dass der *default*-Fall kein *case_item* (in Yosys vgl. *RTLIL::CaseRule*) für den Vergleich hat, sondern ein Ausweichen bei nicht-Zutreffen der anderen Fälle ermöglicht. Ein leerer *compare*-Vektor bedeutet bei dem *default*-Case also nicht, dass der Ast unge-nutzt ist, sondern ist eine Folgerung aus dem Verilog-Syntax.

Zusammenfassung: Wenn ein Signal zum Vergleich für eine Case-Bedingung mit seinem Switch leer ist, kann es aus der entsprechenden Vektor für die Vergleiche *compare* gelöscht werden.

Ist der gesamte Vektor mit Vergleichssignalen leer, so ist der Fall nicht mehr erreichbar, da er nicht mehr eintreten kann. Dann wird er gelöscht, wenn es sich nicht um den Verilog *default*-Case handelt.

2. *proc_arst*

This pass detects processes that describe d-type flip-flops with asynchronous resets and rewrites the process to better reflect what they are modelling: Before this pass, an asynchronous reset has two edge-sensitive sync rules and one top-level RTLIL::SwitchRule for the reset path. After this pass the sync rule for the reset is level-sensitive and the top-level RTLIL::SwitchRule has been removed.

Der Prozess-Generator in *genRTLIL.cc* generiert Zuweisungsereignisse für fallende und steigende Flanken. Bei einem asynchronen Reset wird bei aktivem Reset das Ausgangssignal *Q* auf 0 zugewiesen.

```

1 void proc_arst(RTLIL::Module *mod, RTLIL::Process *proc, SigMap &assign_map
2 )
3 {
4 restart_proc_arst:
5     if (proc->root_case.switches.size() != 1)
6         return;
7
8     RTLIL::SigSpec root_sig = proc->root_case.switches[0]->signal;
9
10    for (auto &sync : proc->syncs) {
11        if (sync->type == RTLIL::SyncType::STp || sync->type == RTLIL::SyncType
12            ::STn) {
13            bool polarity = sync->type == RTLIL::SyncType::STp;
14            if (check_signal(mod, root_sig, sync->signal, polarity)) {
15                if (proc->syncs.size() == 1) {
16                    log("Found VHDL-style edge-trigger %s in '%s.%s'.\n", log_signal(
17                        sync->signal), mod->name.c_str(), proc->name.c_str());
18                } else {
19                    log("Found async reset %s in '%s.%s'.\n", log_signal(sync->signal
20                        ), mod->name.c_str(), proc->name.c_str());
21                    sync->type = sync->type == RTLIL::SyncType::STp ? RTLIL::SyncType
22                        ::ST1 : RTLIL::SyncType::ST0;
23                }
24                for (auto &action : sync->actions) {
25                    RTLIL::SigSpec rspec = action.second;
26                    RTLIL::SigSpec rval = RTLIL::SigSpec(RTLIL::State::Sm, rspec.size
27                        ());
28                    for (int i = 0; i < GetSize(rspec); i++)
29                        if (rspec[i].wire == NULL)
30                            rval[i] = rspec[i];
31                    RTLIL::SigSpec last_rval;
32                    for (int count = 0; rval != last_rval; count++) {
33                        last_rval = rval;
34                        apply_const(mod, rspec, rval, &proc->root_case, root_sig,
35                            polarity, false);
36                        assign_map.apply(rval);
37                        if (rval.is_fully_const())
38                            break;
39                        if (count > 100)
40                            log_error("Async reset %s yields endless loop at value %s for

```



```

    signal %s.\n",
34     log_signal(sync->signal), log_signal(rval), log_signal(
    action.first));
35     rspec = rval;
36 }
37 if (rval.has_marked_bits())
38     log_error("Async reset %s yields non-constant value %s for
    signal %s.\n",
39     log_signal(sync->signal), log_signal(rval), log_signal(
    action.first));
40     action.second = rval;
41 }
42 eliminate_const(mod, &proc->root_case, root_sig, polarity);
43 goto restart_proc_arst;
44 }
45 }
46 }
47 }

```

Die Funktion wird so aufgerufen:

```

1 for (auto mod : design->modules())
2     if (design->selected(mod)) {
3         SigMap assign_map(mod);
4         for (auto &proc_it : mod->processes) {
5             if (!design->selected(mod, proc_it.second))
6                 continue;
7             proc_arst(mod, proc_it.second, assign_map);

```

Es wird durch alle Module des Designs iteriert. Für jedes Modul wird eine *SigMap assign_map()* erstellt. Zu SigMap im Manual:

```

1 /*
2  When a design has many wires in it that are connected to each other, then a
3  single signal bit can have multiple valid names. The SigMap object can be
4  used
5  to map SigSpecs or SigBits to unique SigSpecs and SigBits that consistently
6  only use one wire from such a group of connected wires. For example:
7  */
8 SigBit a = module->addWire(NEW_ID);
9 SigBit b = module->addWire(NEW_ID);
10 module->connect(a, b);
11 log("%d\n", a == b); // will print 0 155
12 SigMap sigmap(module);
13 log("%d\n", sigmap(a) == sigmap(b)); // will print 1

```

3. *proc_mux*

This pass converts the RTLIL::CaseRule/RTLIL::SwitchRule-tree to a tree of multiplexers per written signal. After this, the RTLIL::Process structure only contains the RTLIL::SyncRules that describe the output registers. Anschließend wird durch die Prozesse der Module iteriert.

```

1 void proc_mux(RTLIL::Module *mod, RTLIL::Process *proc, bool ifxmode)
2 {
3     log("Creating decoders for process '%s.%s'.\n", mod->name.c_str(), proc->
        name.c_str());
4
5     SigSnippets sigsnip;
6     sigsnip.insert(&proc->root_case);
7

```

```

8 SnippetSwCache swcache;
9 swcache.snippets = &sigsnip;
10 swcache.insert(&proc->root_case);
11
12 dict<RTLIL::SwitchRule*, bool, hash_ptr_ops> swpara;
13
14 int cnt = 0;
15 for (int idx : sigsnip.snippets)
16 {
17     swcache.current_snippet = idx;
18     RTLIL::SigSpec sig = sigsnip.sigidx[idx];
19
20     log("%6d/%d: %s\n", ++cnt, GetSize(sigsnip.snippets), log_signal(sig));
21
22     RTLIL::SigSpec value = signal_to_mux_tree(mod, swcache, swpara, &proc->
root_case, sig, RTLIL::SigSpec(RTLIL::State::Sx, sig.size()), ifxmode);
23     mod->connect(RTLIL::SigSig(sig, value));
24 }
25 }

```

Die Hauptfunktion `proc_mux`

4. *proc_dff*

This pass replaces the `RTLIL::SyncRules` to d-type flip-flops (with asynchronous resets if necessary).