TALLINN UNIVERSITY OF TECHNOLOGY

School of Information Technologies Thomas Johann Seebeck Department of Electronics

Gaspar Karm, IVEM153410

Pyha Sequential and Object-Oriented Hardware Description Language based on Python

Master's Thesis

Supervisors:

Muhammad Mahtab Alam PhD

Yannick Le Moullec PhD

Tallinn 2017

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond Thomas Johann Seebecki elektroonikainstituut

Gaspar Karm, IVEM153410 $\,$

Pyha

 $Magistrit\ddot{o}\ddot{o}$

Juhendajad:

Muhammad Mahtab Alam PhD

Yannick Le Moullec PhD

Author's declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Gaspar Karm

[pp.kk.aaaa]

Abstract

The thesis is in English and contains [pages] pages of text, 5 chapters, [figures] figures, [tables] tables.

Annotatsioon

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti [lehekülgede arv töö põhiosas] leheküljel, 5 peatükki, [jooniste arv] joonist, [tabelite arv] tabelit.

Contents:

1	Intro	oduction
	1.1	Problem statement
	1.2	Structure
2	Hard	dware design with Pyha
	2.1	Introduction
	2.2	Sequential logic
		2.2.1 Accumulator
		2.2.2 Block processing and sliding adder
	2.3	Fixed-point designs
		2.3.1 Fixed-point support in Pyha
		2.3.2 Converting sliding adder to fixed-point
	2.4	Summary
3	Synt	hesis 1'
	3.1	Structured, Object-oriented style for VHDL
		3.1.1 Defining registers with variables
		3.1.2 The OOP model
		3.1.3 Use cases
	3.2	Converting Python to VHDL
		3.2.1 Type inference
		3.2.2 Syntax conversion
	3.3	Summary
4	Case	e studies 29
	4.1	Moving average filter
	4.2	Linear-phase DC removal Filter
	4.3	Comparison and limitations
	1.0	4.3.1 vs VHDL/SystemVerilog
		4.3.2 vs other Python based tools
		4.3.3 vs MATLAB/SIMULINK
		4.3.4 vs HLS
		4.3.5 Limitations
		4.3.6 Future perspectives

5 Summary	37
Bibliography	39

Chapter 1

Introduction

Main tools today for designing digital hardware are VHDL and SYSTEMVERILOG (SV). SYSTEMVERILOG is aggressively promoted by the big ELECTRONIC DESIGN AUTOMATION (EDA) companies (Cadence, Mentor, Synopsys) along with UNIVERSAL VERIFICATION METHODOLOGY (UVM). In 2003, Aart de Geus, Synopsys CEO, stated that SYSTEMVERILOG will replace VHDL in 10 years [1]. It is true that tool vendors have stopped improving VHDL support, but advancements are being made in the open source community, where VHDL-2017 standard [2] is being developed. In addition, active development is done on open source simulator GHDL [3], Open Source VHDL Verification Methodology (OSVVM) [4] and unit-testing library VUnit [5]. All of these advancements aim to ease the verification aspects while the synthesis part has mostly stayed the same for the past 10 years.

Numerous projects exist that propose to use higher level HARDWARE DESCRIPTION LANGUAGE (HDL), in order to raise the abstraction level. For example, MyHDL turns Python into a hardware description and verification language [6]. Or $C\lambda$ aSH [7], purely Haskell based functional language developed at University of Twente. Recently Chisel [8] has been gaining some popularity, it is an hardware construction language developed at UC Berkeley that uses Scala programming language, providing functional and object-oriented features. .. Still none of these tools have seen widespread adaption.

On the other front, HIGH-LEVEL SYNTHESIS (HLS) tools aim to automate the refinement from the algorithmic level to REGISTER-TRANSFER LEVEL (RTL) [9]. Lately the Vivado HLS, developed by Xilinx, has been gaining popularity. As of 2015, it is included in the free design suite of Vivado (device limited). Problem with HLS tools is that they are often promoted as direct C to RTL tools but in reality often manual code transformations and guidlines are needed, in order to archive reasonable performance [10]. The designer must already know how the RTL works in order to give these instructions.

The DIGITAL SIGNAL PROCESSING (DSP) systems can be described in previously mentioned HLS or HDL languages, but the most productive way is to use MAT-LAB/Simulink/HDLConverter flow, which allows users to describe their designs in Simulink or MATLAB and use HDL convertible blocks provided by MATLAB or FPGA tool vendor [11].

1.1 Problem statement

There is no doubt that MATLAB based workflow offers an highly productive path from DSP models to hardware. However, these tools can easily cost over tens of thousands of euros and often FPGA vendor tools are required that add additional annual cost [11]. Using these tools is not suitable for reproducible research and is completely unusable for open source designs. Thus, the designers must turn to alternative design flows, for example [12] provides an hardware implementation of an ADS-B (automatic dependent surveillance – broadcast) receiver. First, they did the prototyping in the MATLAB environment, the working model was then translated to C for real-time testing and fixed-point modeling. Lastly, the C model was manually converted to VHDL.

This thesis introduces Pyha, a new Python based hardware description language aimed at simplifying DSP hardware development in an open-source manner and providing an alternative for the MATLAB based design flow. Pyha raises the RTL design abstraction level by enabling sequential and object-oriented style, one of the contributions of this thesis is extension of this style to VHDL language, making the conversion process of Pyha unique. DSP systems can be built by using the fixed-point type and semi-automatic conversion from floating point. In addition, this work makes an effort to simplify the testing process of hardware systems by providing better simulation interface for unit-testing.

Basis of Pyha is Python, a general purpose programming language that is especially well suited for rapid prototyping and modeling. Python has also found its place in scientific projects and academia by offering most of what is familiar from MATLAB, free of charge. Scientist are already shifting from MATLAB to Python in order to conduct research that is reproducible and accessible by everyone [13]. Fig. 1.1 shows the popularity comparison (based on Google searches) of Python, MATLAB and C. Python is far ahead and the only one with positive trend, that should soon push Python past Java to position 1.

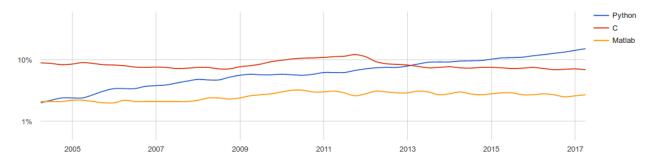


Fig. 1.1: PYPL (PopularitY of Programming Language) [14]. Python 15.1%, C 6.9%, MAT-LAB 2.7%

1.2 Structure

Todo

not done

This thesis is divided into 3 chapters. In chapter 1, main concepts of Pyha are introduced. Following chapter shows First chapter of this thesis gives an overview of the developed tool Pyha and how it can be used for hardware design. Follows the examples that show how Pyha can be used to relatively easly construct moving-average filter and by reusing it the DC-removal filter. Final chapter describes the one of the contributions of this thesis, the sequential VHDL OOP model and how Python is converted to it.

1.2. Structure 3

Chapter 2

Hardware design with Pyha

This chapter introduces the main contribution of this thesis, Pyha - a tool to design digital hardware in Python. The fist half of the chapter demonstrates how basic hardware constructs can be defined, using Pyha. Follows the introduction to fixed-point type. All the examples presented in this chapter can be found online HERE, including all the Python sources, unit-tests, VHDL conversion files and Quartus project for synthesis.

Todo

organise examples to web and put link, INT usage!

2.1 Introduction

Conventional HDL languages promote concurrent and entity oriented models which can be confusing. In this thesis, Pyha has been designed as a sequential object-oriented language, that works directly on Python code. Using a sequential design flow is much easier to understand and is equally well synthesizable as shown in this thesis. Object-oriented design helps to better abstract the RTL details and ease design reuse.

For illustration purposes, Listing 2.1 shows an example Pyha design. The main function has been chosen as a top level entry point, other functions can be used as pleased.

Note: First few examples of this chapter use integer types in order to reduce complexity.

Listing 2.1: Simple combinatory design, implemented in Pyha

```
class Basic(HW):
   def main(self, x):
```

```
a = x + 1 + 3
b = a * 314

if a == 9:
    b = 0

return a, b
```

One of the contributions of this thesis is sequential OOP VHDL model which is used to simplify conversion from Pyha to VHDL. Example of the VHDL conversion is shown on Listing 2.2, more details are given in Section 3.

Listing 2.2: Listing 2.1 converted to VHDL, by Pyha conversion routines

Fig. 2.1 shows the synthesis result. The a output is formed by adding '1' and '3' to the x input. Next the a signal is compared to 9, if equal b is outputted as 0, otherwise b = a * 314. That exactly complies with the Python and VHDL descriptions.

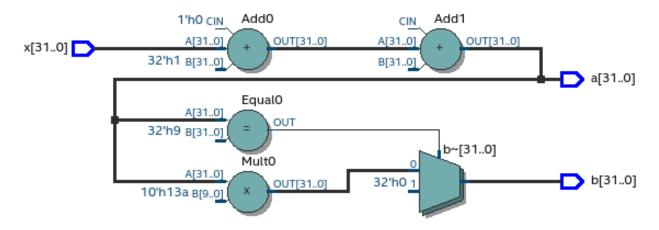


Fig. 2.1: Synthesised RTL of Listing 2.2 (Intel Quartus RTL viewer)

One aspect of hardware design that Pyha aims to improve is testing. Conventional testing flow require the construction of special testbenches that can be executed using simulators.

Pyha has been designed so that the synthesis output is behaviourally equivalent to the Python run output, this means that Pyha designs can use all the Python debugging tools. Fig. 2.2 shows a debugging session on the Listing 2.1 code, this can drastically help the development process.

Fig. 2.2: Debugging using PyCharm (Python editor)

Furthermore, unit testing is accelerated by providing simulate(dut, x) function, that runs the following simulations without any boilerplate code:

- Model: this can be any Python code that fits as an high level model;
- Pyha: like Listing 2.1, Python domain simulation;
- RTL: converts the Pyha model to VHDL and uses the combinition of GHDL and Cocotb for simulation;
- GATE: synthesises the VHDL code, using Intel Quartus, and simulates the resulting gate-level netlist.

This kind of testing function enables test-driven development, where tests can be first defined for the model and fully reused for later RTL implementation. Listing 2.3 shows an example unit test for the Basic() module. Python assert statements can be used for unit

2.1. Introduction 7

test development. Pyha also provides assert_simulate(dut, expected, x) function that automatically compares the output list to the expected list.

Listing 2.3: Unit test for the Basic module

```
x = [1, 2, 3, 4, 5, 6, 7, 8]
dut = Basic()
y = simulation(dut, x) # y contains result of all simulations
# assert something
```

2.2 Sequential logic

The way how registers are inferred is a fundamental difference between the HDL and HLS languages. HDL languages leave the task to the designer, while HLS languages automate the process. In this work, Pyha has been designed to follow the HDL language approach, because this simplifies the conversion to VHDL. Extensions can be considered in future editions.

In conventional programming, state is usually captured by using class variables which can retain values between function calls. Inspired from this, all the class variables in Pyha are handled as registers.

2.2.1 Accumulator

Consider the design of an accumulator (Listing 2.4); it operates by sequentially adding up all the input values of every successive function call.

Listing 2.4: Accumulator implemented in Pyha

```
class Acc(HW):
    def __init__(self):
        self.acc = 0

def main(self, x):
        self.next.acc = self.acc + x
        return self.acc
```

The class structure in Pyha has been designed so that the __init__ function shall define all the memory elements in the design, the function itself is not converted to VHDL, only the variables are extracted. For example __init__ function could be used to call scipy.signal. firwin() to design FIR filter coefficients, initial assignments to class variables are used for register initial/reset values.

Note the self.next.acc = ..., simulates the hardware behaviour of registers, that is delayed assignment. In general this is equivalent to the VHDL <= operator. Values are transferred from next to current just before the main call. In general Pyha abstracts the

clock signal away by denoting that each call to main is a clock edge. Think that the main function is started with the **current** register values known and the objective of the main function is to find the **next** values for the registers.

The synthesis results displayed in the Fig. 2.3 shows the adder and register.

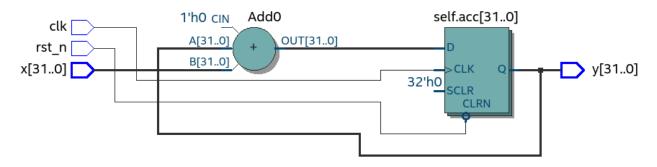


Fig. 2.3: Synthesis result of Listing 2.4 (Intel Quartus RTL viewer)

One inconvenience is that every register on signal path delays the output signal by 1 sample, this is also called pipeline delay or latency. This situation is shown on Fig. 2.4 that shows the simulation results for the Acc module. Note that the model is implemented without register semantics, thus has no pipeline delays. This can be seen from the Fig. 2.4, hardware related simulations are delayed by 1 sample, compared to the software model.

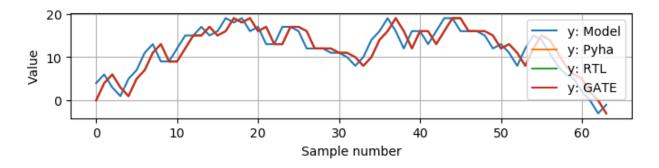


Fig. 2.4: Simulation of the Acc module, input is a random integer [-5;5]

Pyha reserves a self._delay variable, that hardware classes can use to specify their delay. Simulation functions read this variable to compensate the simulation outputs. Setting the self._delay = 1` in the ``__init__ function would shift the hardware simulations left by 1 sample, so that all the simulation would be exactly equal.

2.2.2 Block processing and sliding adder

Common technique required to implement DSP systems is block processing i.e. calculating results on a block of input samples. Until now, the main function has worked with a single

input sample, registers can be used to keep history of samples, so that block processing can be applied.

For an example, consider an algorithm that outputs the sum of last 4 input values. Listing 2.5 shows the Pyha implementation, it works by keeping history of 4 last input samples and summing them for output.

Listing 2.5: Sliding adder algorithm, implemented in Pyha

```
class SlidingAdder(HW):
    def __init__(self):
        self.shr = [0, 0, 0, 0] # define list of registers
        self.y = 0 # output register

def main(self, x):
    # add new 'x' to list, throw away last element
    self.next.shr = [x] + self.shr[:-1]

# add all elements in 'shr'
    sum = 0
    for x in self.shr:
        sum = sum + x

# register the output
    self.next.y = sum
    return self.y
```

The self.next.shr = [x] + self.shr[:-1] implements an 'shift register', because on every call it shifts the list contents to the right and adds new x as the first element. Sometimes the same structure is used as a delay-chain, because the sample x takes 4 updates to travel from shr[0] to shr[3]. This is a very common element in hardware designs. Fig. 2.5 shows the synthesis results.

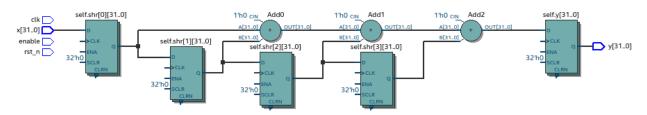


Fig. 2.5: Synthesis result of Listing 2.5 (Intel Quartus RTL viewer)

This design can be made generic by changing the __init__ function to take the window length as a parameter (Listing 2.6), so that SlidingAdder(window_len=4) would add 4 last elements, while SlidingAdder(window_len=6) would add 6.

Listing 2.6: Generic sliding adder, window_len controls the shr list length

```
class SlidingAdder(HW):
    def __init__(self, window_len):
        self.shr = [0] * window_len
    ...
```

This design has a few issues when the window_len gets higher (Fig. 2.6). First, every stage requires a separate adder increasing the resource cost, this also forms an long critical path which in turn decreases the maximum clock rate of the design.

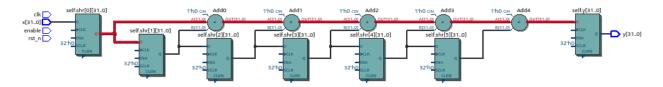


Fig. 2.6: Synthesis result of ``SlidingAdder(window_len=6)``, the red line shows the critical path (Intel Quartus RTL viewer)

Conveniently, the algorithm can be optimized to use only 2 adders, no matter the window length. Listing 2.7 shows that instead of summing all the elements, the overlapping part of the previous calculation can be used to significantly optimize the algorithm.

Listing 2.7: Optimizing the sliding adder algorithm by using recursive implementation

```
y[4] = x[4] + x[5] + x[6] + x[7] + x[8] + x[9]

y[5] = x[5] + x[6] + x[7] + x[8] + x[9] + x[10]

y[6] = x[6] + x[7] + x[8] + x[9] + x[10] + x[11]

# optimized way to calculate by reusing previous results (recursive)

y[5] = y[4] + x[10] - x[4]

y[6] = y[5] + x[11] - x[5]
```

Listing 2.8 gives the implementation of the optimal sliding adder; it features a new register sum, that keeps track of the previous output. The shr now serves the purpose of delay-chain.

Listing 2.8: Optimal sliding adder, implemented in Pyha

```
class OptimalSlideAdd(HW):
    def __init__(self, window_len):
        self.shr = [0] * window_len
        self.sum = 0 # register to remember the 'last' sum

    self._delay = 1

def main(self, x):
```

```
self.next.shr = [x] + self.shr[:-1]

# add new 'x' to sample and subtract the delayed 'x'
self.next.sum = self.sum + x - self.shr[-1]
return self.sum
...
```

Fig. 2.7 shows the synthesis result. Now the critical path is 2 adders, no matter the window_len. In addition, noteice how the shr is just a stack of registers to delay the input signal.

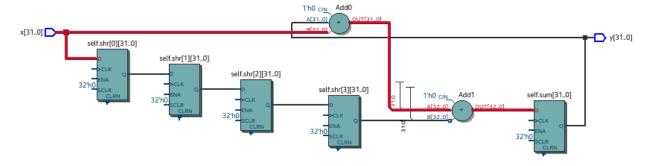


Fig. 2.7: Synthesis result of OptimalSlideAdd(window_len=4) (Intel Quartus RTL viewer)

2.3 Fixed-point designs

DSP systems are commonly described in floating-point arithmetic, which are supported by all conventional programming languages. Floating-point arithmetic can also be used in RTL languages, but the problem is high resource usage [15]. Alternative is to use fixed-point numbers, that work with integer arithmetic. Another benefit of fixed-point numbers is that they can map to FPGA DSP blocks, thus providing higher clocks speed and reduced resource use¹.

Common workflow is to experiment and write model using the floating-point arithmetic, then convert to fixed-point for hardware implementation. In this work Pyha has been designed to simplify the conversion and equivalence testing operations.

2.3.1 Fixed-point support in Pyha

In this work, Pyha has been designed to support signed fixed-point type by providing the Sfix class. The implementation maps directly to the VHDL fixed-point library $[17]^2$, that is already known in the VHDL community and proven to be well synthesizable.

¹ Some high-end FPGAs also include floating-point DSP blocks [16]

² https://github.com/FPHDL/fphdl.

Sfix class works by allocating bits to the left and right side of the decimal point. Bits to the left determine the integer bounds (sign bit is implicit), while the right bits determine the minimum resolution of the number. For example, Sfix(left=0, right=-17) represents a number between [-1;1] with resolution of 0.000007629 (2**-17). Listing 2.9 shows a few examples on how reducing the right reduces the number precision.

Listing 2.9: Example of Sfix type, more bits give better accuracy

```
>>> Sfix(0.3424, left=0, right=-17)
0.34239959716796875 [0:-17]
>>> Sfix(0.3424, left=0, right=-7)
0.34375 [0:-7]
>>> Sfix(0.3424, left=0, right=-4)
0.3125 [0:-4]
```

The default and recommended fixed-point type in Pyha has been chosen to be Sfix(left=0, right=-17), because it can represent normalized numbers and fits into FPGA DSP blocks [18] [15]. Keeping block inputs and outputs in the normalized range can simplify the overall design process.

2.3.2 Converting sliding adder to fixed-point

Consider converting the sliding window adder (developed in Section 2.2.2) to fixed-point implementation. This requires changes only in the __init__ function (Listing 2.10).

Listing 2.10: Fixed-point sliding adder, the rest of the code is identical to one in Section 2.2.2

```
def __init__(self, window_size):
    self.shr = [Sfix()] * window_size # lazy type
    self.sum = Sfix(left=0) # always resize left to 0
```

The first line sets self.shr to store Sfix() elements, this is a lazy statement as it does not specify the fixed-point bounds i.e. it will take bounds from the first assignment to the self.shr variable. The Sfix(left=0) forces left to 0 bits, while the fractional part is determined by the first assign. One problem with the VHDL fixed-point library is that the designer is constantly forced to resize the value to desired format, in this work Pyha has been designed to automate this step i.e. every assign to fixed-point variable is resized to the initial format, the bounds may be taken from the assigned value if initial value is lazy.

Synthesis results Fig. 2.8 show that inputs and outputs are now 18-bits wide, this is due the use of default fixed-point type. Another big addition is the saturation logic, which prevents the wraparound behaviour by saturating the value instead. Wraparound related bugs can be very hard to find, thus it is suggested to keep saturation logic enabled when the overflows are possible.

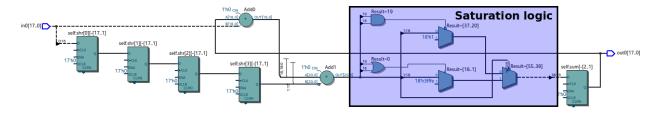


Fig. 2.8: RTL of fixed-point sliding adder, default fixed-point type (Intel Quartus RTL viewer)

The simulate function in Pyha has been designed to automatically convert floating-point inputs to fixed-point, same goes for outputs. This way the unit-test can be kept simple, Listing 2.11 gives an example.

Listing 2.11: Pyha enables testing of fixed-point design with floating-point numbers

```
dut = OptimalSlidingAddFix(window_len=4)
x = np.random.uniform(-0.5, 0.5, 64) # random signal in [-0.5, 0.5] range
y = simulate(dut, x) # all outputs are floats
# assert or plot results
```

The simulation results shown on Fig. 2.9, show that the hardware related simulations differ from the model. This is because model is implemented in floating-point arithmetic while hardware typing is limited to [-1;1] range. Notice that the mismatch starts when the value rises over 1.0.

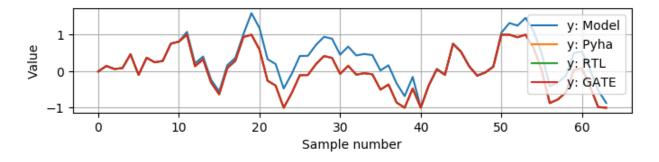


Fig. 2.9: Simulation results of FP sliding sum, input is random signal in [-0.5; 0.5] range

2.4 Summary

This chapter has demonstrated the major features of Pyha and the motivation behind them. It was shown that Pyha is an sequential object-oriented programming language based on Python. It falls to the category of behavioral languages, meaning that the output of Python program is equal to the output of generated hardware. Pyha provides simulate functions to automatically and without any boilerplate code run model and hardware related simulations,

this helps the design of unit-tests. In addition, Pyha designs are fully debuggable in Python ecosystem. Class variables are used to define registers, this has been inspired by traditional programming languages. DSP systems can be implemented by using the fixed-point type. Pyha has 'semi-automatic conversion' from floating point to fixed point numbers. Verifying against floating point model helps the design process.

2.4. Summary 15

Chapter 3

Synthesis

Majority of the hardware related tools end up converting to either VHDL or SystemVerilog, because these are supported by the synthesis tools. Most often the higher level language is converted to very low level VHDL/ SystemVerilog code, resulting in a confusing conversion process and unreadable code.

This thesis tests an alternative path by contributing the sequential synthesizable object-oriented (OOP) programming model for VHDL. The main motivation is to use it as an conversion target for higher level languages. Major advantages are that the conversion process is simple, output VHDL readable and structured.

VHDL has been chosen over SystemVerilog because it is a strict language and forbids many mistakes during compile time. SystemVerilog on the other hand is much more permissive, for example allowing out-of-bounds array indexing [19]. In future both could be supported.

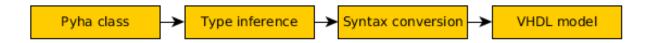


Fig. 3.1: Process of converting Pyha to VHDL.

Fig. 3.1 shows the process of converting Pyha to VHDL. This relies heavily on the final VHDL model (Section 3.1). The process of type inference and syntax conversion is described in Section 3.2.

3.1 Structured, Object-oriented style for VHDL

Todo

Give links to examples, top levels etc, redo images (in0, out0), why int types?

Structured programming in VHDL has been studied by Jiri Gaisler in [20]. He showed that combinatory logic is easily described by functions with sequential statements and proposed the 'two-process' design method, where the first process is used to describe combinatory logic and the second process describes registers. This thesis contributes to the 'two process' model by adding an object-oriented approach, which allows fully sequential designs, easier reuse and removes the one clock domain limitation.

The basic idea behind OOP is to strictly define functions that can perform actions on some group of data. This idea fits well with hardware design, as 'data' can be thought as registers and combinatory logic as functions that perform operations on the data. VHDL has no direct support for OOP but it can still be used by grouping data in record (same as C struct) and passing it as parameter to functions. This is essentially the same way how C programmers do it.

Listing 3.1 demonstrates pipelined multiply-accumulate (MAC), written in OOP VHDL. Recall that all the items in the self_t record are expected to synthesise as registers.

Listing 3.1: OOP style multiply-accumulate in VHDL

```
type self_t is record
    mul: integer;
    acc: integer;
    coef: integer;
end record;

procedure main(self: inout self_t; a: in integer; ret_0: out integer) is
begin
    self.mul := a * self.coef;
    self.acc := self.acc + self.mul;
    ret_0 := self.acc; -- return via 'out' argument
end procedure;
```

The synthesis results (Fig. 3.2) show that a functionally correct MAC has been implemented. However, in terms of hardware, it is not quite what was wanted. The data model specified 3 registers, but only the one for acc is present, and even that is not placed on the critical path.

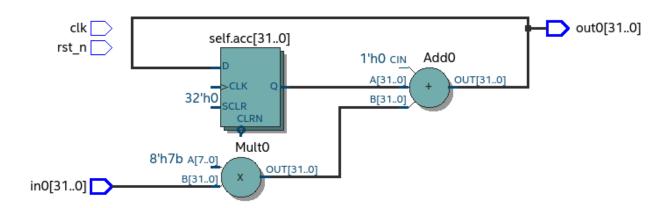


Fig. 3.2: Unexpected synthesis result of Listing 3.1, self.coef=123 (Intel Quartus RTL viewer)

3.1.1 Defining registers with variables

Previous section made a mistake of expecting the registers to work in the same way as 'class variables' in traditional programming languages. Actually registers are delayed elements, so to say they take the next value. VHDL defines a special 'signal assignment' operator for this kind of delayed assignment, that can be used on VHDL signal objects like like a <= b. These objects are hard to map to higher level languages and have limited usage in VHDL structured code constructs. Cannot map to record.

Conveniently, the signal assignment can be mimicked with two variables, to represent the **next** and **current** values. Signal assignment operator sets the value of **next** variable. On the next simulation delta, all the signals are updated i.e. **next** written to **current**. This way of writing sequential logic has been suggested by Pong P. Chu in his VHDL book [21] and is also used in MyHDL signal objects [22].

Adapting this style for the MAC example is shown in Listing 3.2, the data model now includes the nexts member, that should be used to write register values.

Listing 3.2: Data model with **next**, in OOP-style VHDL

```
type next_t is record -- new record to hold 'next' values
    mul: integer;
    acc: integer;
    coef: integer;
end record;

type self_t is record
    mul: integer;
    acc: integer;
    coef: integer;
    roef: integer;
```

VHDL signal assignment automatically updates the signal values, now with the variables method, this has to be done manually. Listing 3.3 defines new function 'update_registers', taking care of this task.

Listing 3.3: Function to update registers, in OOP-style VHDL

```
procedure update_register(self: inout self_t) is
begin
    self.mul := self.nexts.mul;
    self.acc := self.nexts.acc;
    self.coef:= self.nexts.coef;
end procedure;
```

Note: Function 'update_registers' is called on clock raising edge. While the 'main' is called as combinatory function.

Synthesising the revised code shows that the pipelined MAC has been implemented (Fig. 3.3)..



Fig. 3.3: Synthesis result of the revised code (Intel Quartus RTL viewer)

3.1.2 The OOP model

The OOP model, developed in this thesis, consists of following elements:

- Record for 'next',
- Record for 'self',
- User defined functions (like 'main'),

- 'Update registers' function,
- 'Reset' function.

VHDL supports 'packages', that can be used to group all these elements into common namespace. Listing 3.4 shows the template package for VHDL 'class'. All the class functionality is now in common namespace.

Listing 3.4: Class template for OOP style VHDL

```
package MAC is
    type next_t is record
    ...
    end record;

type self_t is record
    ...
    nexts: next_t;
    end record;

-- function prototypes
end package;

package body MAC is
    procedure reset(self: inout self_t) is
    ...
    procedure update_registers(self: inout self_t) is
    ...
    procedure main(self:inout self_t) is
    ...
-- other user defined functions
end package body;
```

The 'reset' function sets the initial values for registers. Listing 3.5 shows a reset function for the MAC circuit. Note that the hardcoded self.nexts.coef := 123; could be replaced with VHDL package generics.

Listing 3.5: Reset function for MAC, in OOP-style VHDL

```
procedure reset(self: inout self_t) is
begin
    self.nexts.coef := 123;
    self.nexts.mul := 0;
    self.nexts.sum := 0;
    update_registers(self);
end procedure;
```

The hardcoded self.nexts.coef := 123; could be replaced with package generic, for example coef. Then each new package could define a new reset value for it (Listing 3.6).

Listing 3.6: Initialize new package MAC₋0, with 'coef' 123

```
package MAC_0 is new MAC
generic map (COEF => 123);
```

3.1.3 Use cases

This section demonstrates how instances of VHDL 'classes' can be used for design reused. Consider an example that consists of two MAC instances and aims to connect them in series (Listing 3.7). In main,

Listing 3.7: Series MACs in OOP-style VHDL

```
type self_t is record
    mac0: MAC_0.self_t; -- define 2 MACs as part of data model
    mac1: MAC_1.self_t;

    nexts: next_t;
end record;

procedure main(self:inout self_t; a: integer; ret_0:out integer) is
    variable out_tmp: integer;
begin
    MAC_0.main(self.mac0, a, ret_0=>out_tmp); -- connect MAC_0 output to MAC_1.
    input
    MAC_1.main(self.mac1, out_tmp, ret_0=>ret_0); -- connect MAC_1 to output
end procedure;
```

Synthesis result shows that two MACs are connected in series Fig. 3.4.

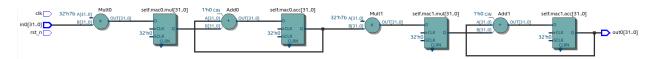


Fig. 3.4: Synthesis result of the new class (Intel Quartus RTL viewer)

Connecting two MAC's instead in parallel can be done with simple modification to main function to return both outputs (Listing 3.8).

Listing 3.8: Main function for parallel instances, in OOP-style VHDL

```
MAC_1.main(self.mac1, a, ret_0=>ret_1); -- return MAC_1 output
end procedure;
```

Two MAC's are synthesized in parallel, as shown in Fig. 3.5.

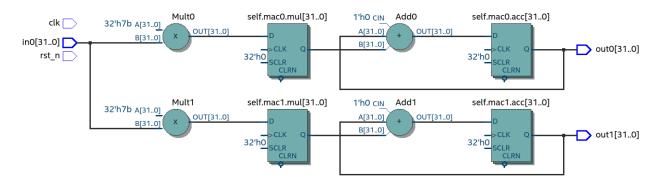


Fig. 3.5: Synthesis result of Listing 3.8 (Intel Quartus RTL viewer)

Multiple clock domains can be easily supported by updating registers at different clock edges. By reusing the parallel MAC's example, consider that MAC_0 and MAC_1 are specified to work in different clock domain. For this only the top level process must be modified (Listing 3.9), rest of the code stays the same.

Listing 3.9: Top-level for multiple clocks, in OOP-style VHDL

```
if (not rst_n) then
    ReuseParallel_0.reset(self); -- reset everything
else
    if rising_edge(clk0) then
        MAC_0.update_registers(self.mac0); -- update 'mac0' on 'clk0' rising edge
    end if;

    if rising_edge(clk1) then
        MAC_1.update_registers(self.mac1); -- update 'mac1' on 'clk1' rising edge
    end if;
end if;
```

Synthesis result (Fig. 3.6) show that registers are clocked by different clocks. The reset signal is common for the whole design.

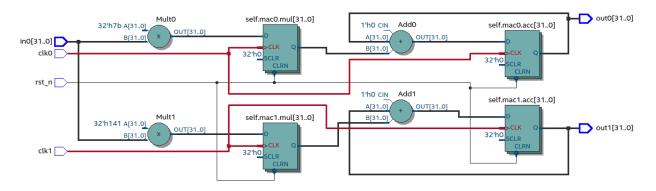


Fig. 3.6: Synthesis result with modified top-level process (Intel Quartus RTL viewer)

3.2 Converting Python to VHDL

The conversion process requires no major transformations or 'understanding' of the source code, this is made possible by the OOP VHDL model, that allows easy mapping of Python constructs to VHDL. Even so, the conversion process poses some challenges like type inference and syntax conversion.

3.2.1 Type inference

One of the biggest difference between Python and VHDL is the typing system. Python uses dynamic typing i.e. types are determined during code execution, while VHDL is statically typed. This poses a major problem for conversion, as the missing type info in Python sources must be somehow inferred in order to produce VHDL code. Naive way to tackle this problem

is to try inferring the types directly from code, for example clearly the type of 'a = 5' is integer. However typically the task is more complex, consider Listing 3.10 as an example, no types can be inferred from this code.

Listing 3.10: What are the types of self.coef, a and local_var?

```
class SimpleClass(HW):
    def __init__(self, coef):
        self.coef = coef

    def main(self, a):
        local_var = a
```

Alternative is to follow the definition of dynamic typing and execute the code, after what the value can be inspected and type inferred. Listing 3.11 shows this method applied on the class variable, the Python function "type()" can be used to query the variable type.

Listing 3.11: Solving the problem for class variables

```
>>> dut = SimpleClass(coef=5)
>>> dut.coef
5
>>> type(dut.coef)
<class 'int'>
```

This solves the problem for class values. The same method cannot be applied for the local variables of functions, because these only exist in the stack. This problem has been encountered before in [23], which proposes to modify the Python profiling interface in order to keep track of function local variables. Pyha has applied this method, usage example is shown on Listing 3.12.

Listing 3.12: Solving the problem for local variables

```
>>> dut.main.locals # locals are unknown before call
{}
>>> dut.main(1) # call function
>>> dut.main.locals # locals can be extracted
{'a': 1, 'local_var': 1}
>>> type(dut.main.locals['local_var'])
<class 'int'>
```

In sum, this method requires the execution of the Python code before types can be inferred. Main advantage of this is very low complexity. In addition this allows the usage of 'lazy' fixed point types as shown in Section 2.3. This method can also be used to keep track of all the values a variable takes, this can enable automatic conversion from floating-point to fixed-point. The code execution needed for conversion is automated in the simulate functions by running the Python domain simulation.

3.2.2 Syntax conversion

Python provides tools that simplify the traversing of source files, like abstract syntax tree (AST) module, that works by parsing the Python file into a tree structure, which can then be modified. Using AST for syntax conversion is known to work but it has very low abstraction level, thus most of the time resulting in complex conversion process. RedBaron [24] is a recent high-level AST tool, that aims to simplify operations with Python source code, unlike AST it also keeps the code formatting and comments.

RedBaron parses the source code into rich objects, for example the 'a = 5' would result in an AssignmentNode. Nodes can be overwritten to change some part of the behaviour. For example, the AssignmentNode can be modified to change = to := and add; to the end of statement. Resulting in a VHDL compatible statement 'a := 5;'. This simple modification turns all the assignments in the code to VHDL style assignments.

Listing 3.13 shows a more complex Python code that is converted to VHDL (Listing 3.14), by Pyha. Most of the transforms are obtained by the same method described above. Some of the transforms are a bit more complex, like return statement to output argument conversion.

Listing 3.13: Python function to be converted to VHDL

```
def main(self, x):
    y = x
    for i in range(4):
        y = y + i
    return y
```

Listing 3.14: Conversion of Listing 3.13, assuming integer types

```
procedure main(self:inout self_t; x: integer; ret_0:out integer) is
    variable y: integer;
begin
    y := x;
    for i in 0 to (4) - 1 loop
        y := y + i;
    end loop;

ret_0 := y;
end procedure;
```

3.3 Summary

This chapter has shown that Pyha achieves synthesizability by converting the Python code to VHDL. The sequential object-oriented VHDL model is one of the contributions of this thesis, it has been developed to provide simpler conversion from Python to VHDL, enabling almost direct conversion from Python the VHDL by using RedBaron based syntax transformations. Type information is acquired trough the simulation. Pyha provides simulate functions that automate the simulation and conversion parts.

The conversion process is one of the advantages of Pyha, compared to other similar tools. The process has low complexity and produces well formatted and readable VHDL. In addition, syntax conversion could be easily extended to support other conversion targets like SytemVerilog or C.

3.3. Summary 27

Chapter 4

Case studies

This chapter demonstrates that Pyha is already usable for real designs. First designs an moving average filter and later reuses this for linear-phase DC removal filter. Last chapter also compares the developed tool to other available toolsets.

Todo

fix

Pyha is based on the object-oriented design practices, this greatly simplifies the design reuse as the classes can be used to initiate objects. Another benefit is that classes can abstract away the implementation details, in that sense Pyha can become a high-level synthesis (HLS) language.

Lastly we showed that Pyha is already usable to convert some mdeium complexity designs, FSK demodulator, that was used on Phantom 2 stuff..

4.1 Moving average filter

The moving average (MA) is the easiest digital filter to understand and use. It is optimal filter for for reducing random noise while retaining a sharp step response [25]. In communication systems, MA is widely used as an matched filter for rectangular pulses. Fig. 4.1 shows an example of applying MA filter to reduce noise on harmonic signal. Higher window length (averaged over more elements) reduces more noise but also increases the complexity and delay of the filter (MA is a special case of FIR filter [25]).

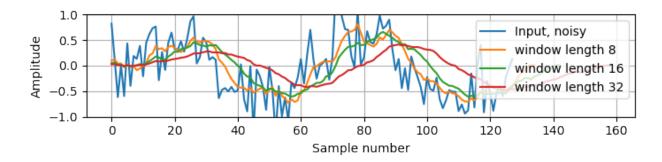


Fig. 4.1: Moving average filter applied on noisy signal

Good noise reduction performance can be explained by the frequency response of the MA filter (Fig. 4.2), showing that it is a low-pass filter. Passband width and stopband attenuation are controlled by the window length.

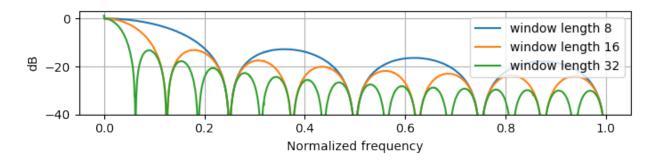


Fig. 4.2: Frequency response of moving average filter

MA filter is implemented by sliding sum, that is divided by the sliding window length. The division can be carried out by a shift operation if divisor is power of two. In addition, division can be performed on each sample instead of on the sum, that is (a + b) / c = a/c + b/c. This guarantees that the sum is always in the [-1;1] range and no saturation logic is needed.

Listing 4.1 shows the MA filter implementation in Pyha. It is based on the sliding sum, that was implemented in Section 2.3.2. Minor modifications are commented in the code.

Listing 4.1: MA implementation in Pyha

```
class MovingAverage(HW):
    def __init__(self, window_len):
        # calculate power of 2 value of 'window_len', used for division
        self.window_pow = Const(int(np.log2(window_len)))

# 'overflow_style' turns the saturation off
        self.sum = Sfix(0, 0, -17, overflow_style=fixed_wrap)
        self.shr = [Sfix()] * window_len
        self._delay = 1
```

```
def main(self, x):
    # divide by shifting
    div = x >> self.window_pow

self.next.shr = [div] + self.shr[:-1]
self.next.sum = self.sum + div - self.shr[-1]
return self.sum
```

Fig. 4.3 shows the synthesized result of this work; as expected it looks very similar to the sliding sum RTL schematics. In general, shift operators are hard to notice on the RTL schematics because they are implemented by routing semantics.

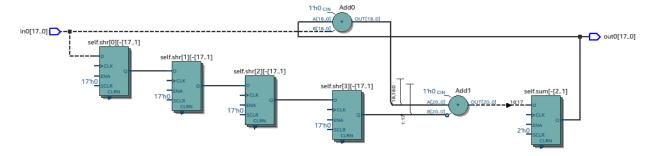


Fig. 4.3: RTL view of moving average (Intel Quartus RTL viewer)

Fig. 4.4 shows simulation results of MA filter used for matched filtering. The plot on (a) shows digital input signal that is corrupted by noise. Plot (b) shows that the MA with window length equal to samples per symbol can recover (optimal result) the signal from the noise. Next the signal could be sampled to recover bit values (0.5=1, -0.5=0).

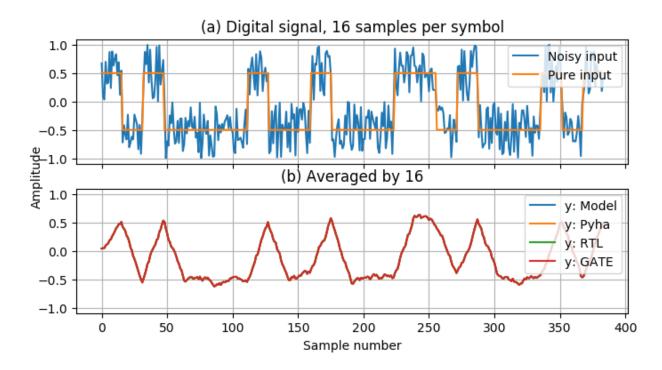


Fig. 4.4: Moving average as matched filter. (b) noisy input signal, (a) averaged by 16, Pyha simulations

4.2 Linear-phase DC removal Filter

This design demonstrates how the object-oriented nature of Pyha can be used for simple design reuse by chaining multiple MA filters to implement linear-phase DC removal filter.

Direct conversion (homodyne or zero-IF) receivers have become very popular recently especially in the realm of software defined radio. There are many benefits to direct conversion receivers, but there are also some serious drawbacks, the largest being DC offset and IQ imbalances [26]. DC offset looks like a peak near the 0 Hz on the frequency response. In time domain it manifests as a constant component on the harmonic signal.

In [27], Rick Lyons investigates the use of moving average algorithm as a DC removal circuit. This works by subtracting the MA output from the input signal. The problem of this approach is the 3 dB passband ripple. However, by connecting multiple stages of MA's in series, the ripple can be avoided (Fig. 4.5) [27].

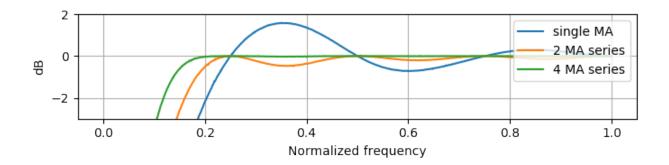


Fig. 4.5: Frequency response of DC removal filter (MA window length is 8)

The algorithm is composed of two parts. First, four MA's are connected in series, outputting the DC component of the signal. Second, the MA's output is subtracted from the input signal, thus giving the signal without DC component. Listing 4.2 shows the Pyha implementation.

Listing 4.2: DC-Removal implementation

This implementation is not exactly following that of [27]. They suggest to delay-match the step 1 and 2 of the algorithm, but since the DC component is more more or less stable, this can be omitted.

Fig. 4.6 shows that the synthesis generated 4 MA filters that are connected in series, output of this is subtracted from the input.

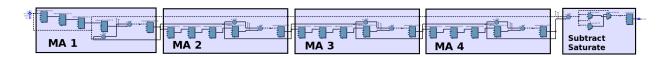


Fig. 4.6: Synthesis result of DCRemoval(window_len=4) (Intel Quartus RTL viewer)

In a real application, one would want to use this component with larger window_len. Here 4 was chosen to keep the synthesis result simple. For example, using window_len=64 gives much better cutoff frequency (Fig. 4.7); FIR filter with the same performance would require hundreds of taps [27].

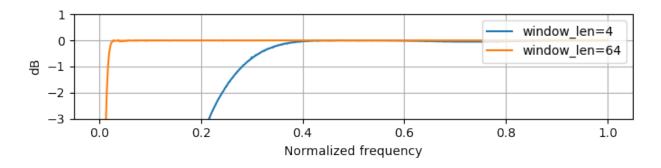


Fig. 4.7: Comparison of frequency response

This implementation is also very light on the FPGA resource usage (Listing 4.3).

Listing 4.3: Cyclone IV FPGA resource usage for DCRemoval(window_len=64)

```
Total logic elements 242 / 39,600 ( < 1 % )
Total memory bits 2,964 / 1,161,216 ( < 1 % )
Embedded Multiplier 9-bit elements 0 / 232 ( 0 % )
```

Fig. 4.8 shows the simulation results for input signal with DC component of +0.5, the output of the filter starts countering the DC component until it is removed.

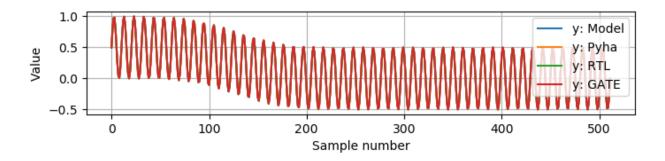


Fig. 4.8: Simulation of DC-removal filter in the time domain

34

4.3 Comparison and limitations

At this point all the majord features of Pyha has been introduced so it is suitable to compare this solution to other languages.

Warning: this section very very in progress

4.3.1 vs VHDL/SystemVerilog

Naturally, as Pyha converts to VHDL, everything in doable in Pyha can be also done in VHDL. However Pyha can be considered as an higher level abstraction for VHDL language, simplifying many tasks. In addition the simulations processes are greatly improved.

4.3.2 vs other Python based tools

MyHDL is following the event-driven approach which is a trait of the classical HDL's. It features an function based design that is very similar to Verilog processes. In general the synthesizable subset of MyHDL is very limited, it has been found that the tool is more useful for high-level modeling purposes [28]. Another package in the Python ecosystem is Migen, that replaces the event-driven paradigm with the notions of combinatorial and synchronous statements [29]. Migen can be considered as meta-programming in Python so it is a bit complicated. Both Migen and MyHDL are more aimed at the control logic, neither implements the fixed-point data type, that is a standard for hardware DSP designs.

Overall i would say that both MyHDL and Migen are awesome tools, in the future merging of Pyha to either MyHDL or Migen can defiantly be considered.

4.3.3 vs MATLAB/SIMULINK

Pyha is already comparing well against MATLAB to HDL based workflow. MATLAB also requires the user to handle the register assignments like Pyha, but provides a much counterintuitive method for this. Reuse in MATLAB is complicated as the designs are function based.

SIMULINk flow is based on connecting togather already exsisting blocks. In current state, Pyha cannot compare directly to the SIMULINK based design flows because there is no graphical user interface. However as shown in this chapter, exsisting Pyha blocks can be connected togather in purely Pythonic way, thus the designer needs no knowledge aout the underlying RTL implementation. In the future GUI could be easily made. GNURADIO example?

Pyha aims to raise the abstraction level by using sequential object-oriented style, major advantage of this is that existing blocks can be connected together in purely Pythonic way, the designer needs to know nothing about the underlying RTL implementation.

4.3.4 vs HLS

In some seneses the Python part could be considers as Python binding to VHDL OOP model. Convert to HLS language instead of VHDL, then the designer could choose to to either design for RTL or HLS, this is more as an futures perspective, this thesis works only with the RTL part.

The design choices done in the process of Pyha design have focused on simplicity. The conversion process of Python code to VHDL is straight-forward as the synthesis tools are already capable of elaborating sequential VHDL code. This work contributes the object-oriented VHDL desing way that allows defining registers in sequential code. Thanks to that, the OOP Python code can be simply mapped to OOP VHDL code. Result is readable (keeps hirarchy) VHDL code that may provide an bridge for people that already know VHDL.

4.3.5 Limitations

Pyha has been developed during the period of last year by me, side project, so it is clear it has some limitations and that it cannot compare to the library support of other tools that have exsisted for years and are developed by bigger community.

One of the limitation emerging from the Python domain simulation is the support of support of one clock domain. In general since Pyha has been applied to SDR applications this has not been a problem as all the processing happens in baseband and generally no sample rate conversions are required. In addition Pyha lacks the support for interfaces like AXI or Avalon. Also most of the tools support converting to Verilog and VHDL, while Pyha only supports VHDL.

Integration to bus structures is another item in the wish-list. Streaming blocks already exist in very basic form. Ideally AvalonMM like buses should be supported, with automatic HAL generation, that would allow design of reconfigurable FIR filters for example.

4.3.6 Future perspectives

Long term goal of the project is to develop enough blocks that are functionally equal to GNURadio blocks, so that flow-graphs could be converted to FPGA designs, thus providing an open source alternative for Simulink based flows.

Chapter 5

Summary

This work introduced new Python based HDL language called Pyha. That is an sequential object-oriented HDL language. Overview of Pyha features have been given and shown that the tool is suitable to use with model based DSP systems. It was also demonstrated that Pyha supports fixed point types and semi-automatic conversion from floating point types. Pyha also provides good support for unit test and designs are debuggable.

Two case studies were presented, first the moving average filter that was implemented in RTL level. Second example demonstrated that blocks written in Pyha can be reused in pure Python way, developed linear phase DC removal filter by reusing the implementation of moving average filter. Synthesisability has been demonstrated on Cyclone IV device.

Another contribution of this thesis is the sequential object-oriented VHDL model. This was developed to enable simple conversion of Pyha to VHDL. One of the advantages of this work compared to other tools is the simplicity of how it works.

Future perspectives are to implement more DSP blocks, especially by using GNURadio blocks as models. That may enable developing an work-flow where GNURadio designs can easily be converted to FPGA. In addition, the Pyha system could be improved to add automatic fixed point conversion and support for multiple clock domains.

In this work Pyha has been designed to add DSP related features to the Python conversion scope, this includes fixed-point type and semi-automatic conversion from floating point. In addition Pyha integrates the model to designs and test functions simplification. Pyha includes functions that help verification by automatically running all the simulations, asserting that model is equivalent to the synthesis result, tests defined for model can be reused for RTL, model based verification. Pyha designs are also simulatable and debuggable in Python domain. The design of Pyha also supports fully automatic conversion but currently this is left as a future work.

Pyha is a fully sequential language that works on purely Python code. However Pyha resides in the RTL level, allowing to define each and every register. In that sense Pyha is at somewhere between the HLS and HDL language. Pyha aims to raise the abstraction level by using the object-oriented style, so that the RTL details can be easily abstracted away.

One major advantage of Pyha is that existing blocks can be connected together in purely Python way, the designer needs to know nothing about hardware design or underlying RTL implementation.

The design choices done in the process of Pyha design have focused on simplicity. The conversion process of Python code to VHDL is straight-forward as the synthesis tools are already capable of elaborating sequential VHDL code. This work contributes the object-oriented VHDL design way that allows defining registers in sequential code. Thanks to that, the OOP Python code can be simply mapped to OOP VHDL code. Result is readable (keeps hierarchy) VHDL code that may provide an bridge for people that already know VHDL.

Bibliography

- [1] Richard Goering. Designers debate 'vhdl is dead' assertion. URL: http://www.eetimes.com/document.asp?doc_id=1216820.
- [2] Ieee p1076 working group vhdl analysis and standardization group (vasg). URL: http://www.eda-twiki.org/cgi-bin/view.cgi/P1076/WebHome.
- [3] Tristan Gingold. Vhdl 2008/93/87 simulator. 2017. URL: https://github.com/tgingold/ghdl.
- [4] Open source vhdl verification methodology (osvvm). URL: http://osvvm.org/.
- [5] Lars Asplund. Vunit. URL: http://vunit.github.io/.
- [6] Myhdl. URL: http://www.myhdl.org.
- [7] Clash. URL: http://www.clash-lang.org/.
- [8] Jonathan Bachrach. Chisel: constructing hardware in a scala embedded language. 2012.
- [9] Wim Meeus, Kristof Van Beeck, Toon Goedemé, Jan Meel, and Dirk Stroobandt. An overview of today's high-level synthesis tools. *Design Automation for Embedded Systems*, 16(3):31–51, 2012. URL: http://dx.doi.org/10.1007/s10617-012-9096-8, doi:10.1007/s10617-012-9096-8.
- [10] F. Winterstein, S. Bayliss, and G. A. Constantinides. High-level synthesis of dynamic data structures: a case study using vivado hls. In 2013 International Conference on Field-Programmable Technology (FPT), 362–365. Dec 2013. doi:10.1109/FPT.2013.6718388.
- [11] Hayden Kwok-Hay So. BORPH: An Operating System for FPGA-Based Reconfigurable Computers. PhD thesis, UNIVERSITY OF CALIFORNIA, BERKELEY, 2007.
- [12] Robert Ghilduta and Brian Padalino. Bladerf vhdl ads-b decoder. URL: https://www.nuand.com/blog/bladerf-vhdl-ads-b-decoder/.
- [13] Neil Lawrence. Gpy: moving from matlab to python. URL: http://inverseprobability.com/2013/11/25/gpy-moving-from-matlab-to-python.
- [14] Pierre Carbonnelle. Pypl popularity of programming language. URL: http://pypl.github.io/PYPL.html.

- [15] Ambrose Finnerty and Hervé Ratigner. Reduce power and cost by converting from floating point to fixed point. 2017.
- [16] Amulya Vishwanath. Enabling high-performance floating-point designs. URL: https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/wp/wp-01267-fpgas-enable-high-performance-floating-point.pdf.
- [17] David Bishop. Fixed point package user's guide. 2016.
- [18] Altera. Cyclone iv fpga device family overview. 2016.
- [19] Aurelian Ionel Munteanu. Gotcha: access an out of bounds index for a system verilog fixed size array. URL: http://www.amiq.com/consulting/2016/01/26/gotcha-access-an-out-of-bounds-index-for-a-system verilog-fixed-size-array/.
- [20] Jiri Gaisler. A structured vhdl design method. URL: http://www.gaisler.com/doc/vhdl2proc.pdf.
- [21] Pong P. Chu. RTL Hardware Design Using VHDL: Coding for Efficiency, Portability, and Scalability. Wiley, 2006.
- [22] Jan Decaluwe. Why do we need signal assignments? URL: http://www.jandecaluwe.com/hdldesign/signal-assignments.html.
- [23] Pietro Berkes and Andrea Maffezoli. Decorator to expose local variables of a function after execution. URL: http://code.activestate.com/recipes/577283-decorator-to-expose-local-variables-of-a-function-/.
- [24] Laurent Peuch. Redbaron: bottom-up approach to refactoring in python. URL: http://redbaron.pycqa.org/.
- [25] Steven W. Smith. The Scientist & Engineer's Guide to Digital Signal Processing. California Technical Pub, 1997.
- [26] BladeRF community. Dc offset and iq imbalance correction. 2017. URL: https://github.com/Nuand/bladeRF/wiki/DC-offset-and-IQ-Imbalance-Correction.
- [27] Rick Lyons. Linear-phase dc removal filter. 2008. URL: https://www.dsprelated.com/showarticle/58.php.
- [28] Jan Decaluwe. It's a simulation language! URL: http://www.jandecaluwe.com/blog/its-a-simulation-language.html.
- [29] Migen. URL: https://m-labs.hk/gateware.html.

40 Bibliography