

Programmierpraktikum

# **Einfacher? Crosscompiler von Python nach C++**

Felix Hensch  
Julian Buchhorn

8. Februar 2015

**Betreuer**  
Dr. Holger Arndt

## Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>2</b>
<b>Listings</b>	<b>2</b>
<b>1 Einleitung</b>	<b>3</b>
<b>2 Verwendung</b>	<b>4</b>
<b>3 Installation</b>	<b>5</b>
3.1 Benötigte Programme . . . . .	5
3.2 Optionale Programme . . . . .	5
<b>4 Features</b>	<b>6</b>
4.1 Unterstützte Syntaxelemente . . . . .	6
4.2 Fehlererkennung . . . . .	6
<b>5 Codestruktur</b>	<b>6</b>
<b>A Installation und Verwendung von Antlr</b>	<b>7</b>
A.1 Installation . . . . .	7
A.2 Rekompilieren und grafische Synatxtree-Ansicht . . . . .	8
<b>Literatur</b>	<b>9</b>

## Abbildungsverzeichnis

1 Beispiel Parsetree . . . . .	8
--------------------------------	---

## Listings

1 Python Testprogramm für den Compiler . . . . .	3
2 Vom Compiler erzeugtes C++ Programm . . . . .	3

## 1 Einleitung

Im Rahmen eines Programmierpraktikums an der Bergischen Universität Wuppertal haben wir einen Compiler von Python zu C++ geschrieben. Dabei ging es in erster Linie darum, zu verstehen wie ein Compiler funktioniert und nicht darum einen vollständigen Compiler zu schreiben.

Als vorläufiges Ziel hatten wir uns gesetzt ein einfaches Testprogramm mit grundlegenden Syntaxelementen zu übersetzen (vgl. Lst. 1).

**Listing 1:** *Python Testprogramm für den Compiler*

```
1 from __future__ import division
2
3 """
4 Test program for codegeneration
5 """
6
7 def f(x):
8     """calculate"""
9     return x**2 + x*4 - (x-3) /x -1.3e2+2\
10         +5 #comment
11
12
13
14 if __name__ == '__main__':
15
16     L= [0]*10
17
18     for i in range(len(L)):
19         if i>3:
20             L[i]= f(i+1)
21         else:
22             L[i]+=1
23
24     for x in L:
25         print x
26
27     print 'program', 'end'
```

Das Ziel haben wir erreicht. Unser Compiler kann in seinem momentanen Zustand das Python Programm übersetzen und liefert das C++ Programm in Listing 2. Dabei mussten wir, wie erwartet, einige Annahmen über das Programm machen, um die Übersetzung zu vereinfachen.

**Listing 2:** *Vom Compiler erzeugtes C++ Programm*

```
1 #include <array>
2 #include <cmath>
3 #include <iostream>
4
5 using namespace std;
6
```

```
7  /*
8  Test program for codegeneration
9  */
10
11 double f(double x) {
12     /*calculate*/
13     return pow(x, 2)+ x*4 - (x-3) /(double) x-1.3e2+2
14         +5; //comment
15 }
16
17
18
19 int main() {
20
21     array<double,10> L={};
22
23     for (int i= 0; i<L.size(); i++) {
24         if (i>3) {
25             L[i]= f(i+1);
26         }
27         else {
28             L[i]++; // com
29         }
30     }
31
32     for (auto x : L) {
33         cout<< x << endl;
34     }
35
36     cout<< "program" << ' ' << "end" << endl;
37
38     return 0;
39
40 }
```

In den folgenden Kapiteln werden wir erklären wie man den Compiler verwenden kann, welche Programme benötigt werden und welche Features der Compiler im Moment unterstützt. Zum Schluss werden wir noch kurz auf die Struktur des Codes eingehen, den wir geschrieben haben.

Für eine detailliertere Erklärung verweisen wir auf die Bachelorarbeit von Felix Hensch verwiesen..., in welcher die Erstellung des Compilers fortgeführt wird.

## 2 Verwendung

Um ein Python Programm nach C++ zu compilieren, muss man nur das Main-Programm `py2cpp.py` aufrufen und dabei den Namen des zu übersetzenden Programms als Kommandozeilenparameter angeben. Das heißt, dass zum Beispiel unser Testprogramm `program.py` mit dem folgenden Konsolenbefehl übersetzt werden kann:

```
python py2cpp.py program.py
```

Dafür muss man allerdings die benötigten Programme installiert haben (s. Kapitel 3) und beachten, dass das zu übersetzende Programm nur die momentan unterstützten

Features enthält (s. Kapitel 4).

## 3 Installation

Zuerst muss man das Archiv mit den Compilerdateien herunterladen [1] und an einer beliebigen Stelle auspacken. Außerdem müssen Python und C++ installiert sein (siehe nächster Abschnitt). Innerhalb dieses Verzeichnisses kann der Compiler dann wie beschrieben aufgerufen werden, wenn sich die zu übersetzende Datei ebenfalls im selben Verzeichnis befindet. Sonst sollte man entsprechend den absoluten Pfad angeben.

### 3.1 Benötigte Programme

Folgende Programme werden zur Verwendung des Compilers in jedem Fall benötigt, jeweils in der angegebenen Version oder einer neueren:

- Python 2.7
  - pandas 0.15.2
- C++ 11
- Python-Runtime für Antlr 4.4.1

Die `Python-Runtime` ist in dem Compiler Archiv schon enthalten, braucht also nicht extra installiert werden. Falls man sie doch runterladen möchte findet man sie in der entsprechenden Sektion der `Antlr` Website [2].

`pandas` ist ein Package für Python welches der Compiler ebenfalls benötigt. Da `pandas` wiederum von den Packages `numpy`, `python-dateutil` und `pytz` abhängt, müssen diese ebenfalls installiert werden. Falls man diese Installationen nicht alle manuell durchführen will, bietet es sich hier an einen Package Manager zu benutzen (für Python unter Windows z.B. `Anaconda`). Für eine detaillierte Installationanleitung verweisen wir auf die Website von Pandas [3].

### 3.2 Optionale Programme

Die nachfolgenden Programme werden prinzipiell nur benötigt, wenn man die Grammatik `py.g4` neukompillieren möchte:

- Antlr 4.4
  - Java 1.6

`Antlr` bietet allerdings auch noch zusätzliche Funktionen für die sich eine Installation lohnen kann. Zum Beispiel kann `Antlr` auch grafische Ansichten des Parsetrees erzeugen (s. Abbildung 1), welche sehr nützlich sind zur Fehlersuche, oder einfach nur um die Struktur des Parsetrees besser zu verstehen. Außerdem kann die Grammatik dann nicht nur für Python kompiliert werden, sondern auch für die anderen Sprachen welche `Antlr` unterstützt.

Da **Antlr** in **Java** geschrieben ist, benötigt es auch eine **Java** Installation. Insgesamt ist die Installation von **Antlr** etwas aufwändiger (insbesonder unter Windows), weswegen wir dazu noch eine ausführlichere Anleitung verfasst haben, welche sich in Abschnitt [A](#) des Anhangs befindet. Insbesondere für die Installation unter Linux wollen wir aber auch auf die Anleitung auf der **Antlr** Website verweisen [\[4\]](#).

## 4 Features

### 4.1 Unterstützte Syntaxelemente

### 4.2 Fehlererkennung

## 5 Codestruktur

## A Installation und Verwendung von Antlr

**Antlr** ist prinzipiell ein Parser-Generator mit ähnlicher Syntax wie **Yacc** oder **Bison**. Der hauptsächliche Unterschied liegt darin, dass im Gegensatz zu **Yacc** einen recursiv-descent Parser erzeugt. Daraus ergibt sich der Vorteil, dass **Antlr** bessere Fehlermeldungen erzeugen kann, aber ein Nachteil ist, dass keine indirekte Linksrekursion möglich ist.

Für eine Erklärung der Syntax verweisen wir auf die **Antlr** Homepage [5], in diesem Kapitel werden nur die Installation beschrieben und die benötigten Kommandozeilenaufrufe für **Antlr** bzw. den erzeugten Parser.

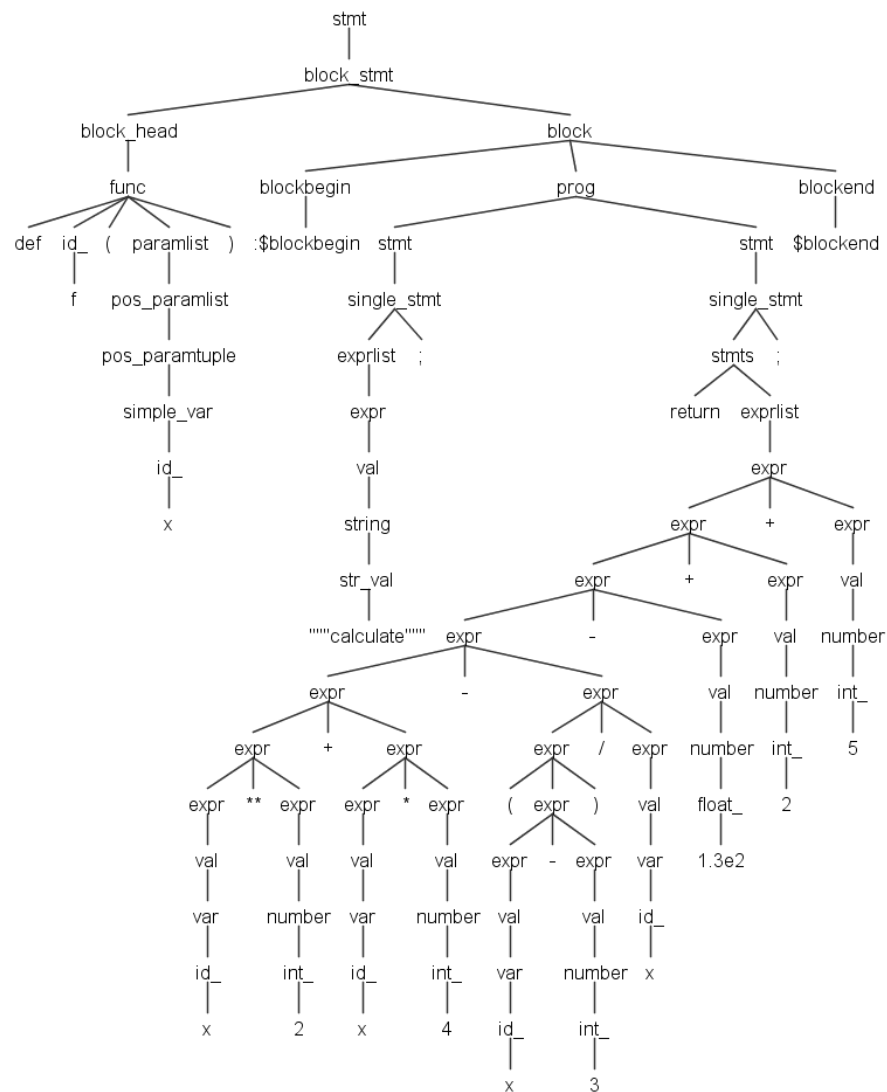
### A.1 Installation

Um den Compiler ausführen zu können, muss Python in der Version 2.7 installiert sein. Außerdem werden die Python-Bibliotheken Numpy und Pandas, sowie deren Abhängigkeiten, benötigt. Am einfachsten ist es, hierfür einen Package-Manager wie zum Beispiel Anaconda zu installieren.

Wenn die enthaltene Python Grammatik geändert werden soll und diese somit neu compiliert werden muss, ist die Installation etwas aufwändiger. Im Folgenden werden die Installationsschritte für Windows und Linux gezeigt.

1. Installation des Java Development Kit (SDK/JDK) in der Version 1.6 oder höher
2. Download des „Complete ANTLR 4.4 Java binaries jar“-Paketes (oder neuer) von der [Download-Seite des ANTLR-Projektes](#)
3. Kopieren der heruntergeladenen Java-Bibliothek in den „jre/lib/ext“-Ordner des JDK (zum Beispiel: „*C : \ProgramFiles(x86)\Java\jdk1.8.0\_25\jre\lib\ext*“ bzw. unter Linux: */usr/lib64/jdk1.8.0\_25/jre/lib/ext/*)
4. Hinzufügen der jar-Datei zum Klassenpfad

## A.2 Rekompilieren und grafische Synataxtree-Ansicht



### Abbildung 1: Beispiel Parsetree



## Literatur

- [1] *Compiler*. URL: ??.
- [2] *antlr, python-runtime*. URL: <https://theantlguy.atlassian.net/wiki/display/ANTLR4/Python+Target>.
- [3] *pandas, Installation*. URL: <http://pandas.pydata.org/pandas-docs/stable/install.html>.
- [4] *antlr, Installation*. URL: <https://theantlguy.atlassian.net/wiki/display/ANTLR4/Getting+Started+with+ANTLR+v4>.
- [5] *antlr, Homepage*. URL: <http://www.antlr.org/>.