

Metaohjelmointi Python-kielellä

Mikko Koho

Seminaarityö

HELSINGIN YLIOPISTO

Tietojenkäsittelytieteen laitos

Helsinki, 20. marraskuuta 2014

Tiedekunta — Fakultet — Faculty		Laitos — Institution — Department	
Matemaattis-luonnontieteellinen		Tietojenkäsittelytieteen laitos	
Tekijä — Författare — Author			
Mikko Koho			
Työn nimi — Arbetets titel — Title			
Metaohjelmointi Python-kielellä			
Oppiaine — Läroämne — Subject			
Tietojenkäsittelytiede			
Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages	
Seminaarityö	20. marraskuuta 2014	9	
Tiivistelmä — Referat — Abstract			
<p>Tiivistelmä.</p> <p>ACM Computing Classification System (CCS):</p>			
Avainsanat — Nyckelord — Keywords			
Python, metaohjelmointi			
Säilytyspaikka — Förvaringsställe — Where deposited			
Muita tietoja — Övriga uppgifter — Additional information			

Sisältö

1	Johdanto	1
2	Python-ohjelmointikieli	1
2.1	Syntaksi	2
2.2	Muuttujat	2
2.3	Tietotyytit	3
2.3.1	Sekvenssit	3
2.4	Luokat ja oliot	4
3	Metaohjelmointi Pythonilla	5
3.1	Reflektio	5
3.2	Dynaamiset luokat	6
3.3	Koodin kääntäminen ajon aikana	6
3.4	Tavukoodin tarkastelu ja muokkaus	7
3.5	Käännösaikainen metaohjelmointi	8
	Lähteet	9

1 Johdanto

Metaohjelmoinnilla tarkoitetaan klassisen määritelmän mukaan sellaisen tietokoneohjelman tekemistä, joka kirjoittaa uusia tietokoneohjelmia [HB12]. Tämä on kuitenkin melko yksinkertaistettu määritelmä, eikä metaohjelmointia ole helppo määritellä tarkasti. Toinen yleinen määritelmä esittää metaohjelmoinnin olevan ”tietokoneohjelma, joka manipuloi toisia ohjelmia ajon aikana” [HB12].

Tärkeä osa metaohjelmointia on ohjelman ajonaikainen tilansa tarkastelu ja muuntelu eli reflektio. Reflektioon kuuluu käsite introspektio, jolla tarkoitetaan ajonaikaista muistissa olevien olioiden tarkastelua [Pil04].

Tässä seminaarityössä tarkastellaan Python-ohjelmointikielen tarjoamia työkaluja metaohjelmointiin. Alussa käydään läpi Python-kielen perusteita ja tämän jälkeen tutustutaan metaohjelmointiin Python-kielillä. Metaohjelmoinnista tarkastellaan lähinnä suoritusajasta metaohjelmointia.

2 Python-ohjelmointikieli

Ensimmäinen Python-kielen versio on julkaistu 1991. Python-kielestä on nykyään käytössä eri versioita ja Python 2.7 on edelleen melko suosittu vaikka versio 3 on julkaistu jo 2008. Versio 3 ei ole yhteensopiva aiempien versioiden kanssa. Version 2:n suosion taustalla on se, että monet suositut kirjastot ja sovelluskehikset eivät ole siirtyneet versioon 3. Tämän seminaarityön esimerkit toimivat sekä Python 2.7:llä että Python 3:lla, ellei toisin mainita. Esimerkeissä olevat rivin alusta alkavat kommentit sisältävät ohjelman tulosteita, koodia koskevat kommentit on laitettu kommentoitavan rivin perään.

Pythonin suosio on kasvanut tasaisesti ja se on nykyään käytetyin kieli ohjelmoinnin perusteiden opetukseen Yhdysvaltojen yliopistoissa [Guo14].

Python-ohjelmakoodia voidaan kääntää useilla eri kääntäjillä [Mar06]. Käytetyin kääntäjä on CPython (Classic Python), joka kääntää alkupe-
räisen koodin Python-tavukoodiksi. Muita suosittuja kääntäjiä ovat Java-
tavukoodiksi kääntävä Jython sekä IronPython, joka kääntää Python-koodia
.NET-ympäristön käyttämäksi CIL-tavukoodiksi. PyPy on Python-kielillä

toteutettu useissa eri ympäristöissä toimiva suoraan konekielelle koodia kääntävä ajonaikainen (just-in-time) kääntäjä.

CPythonilla käännettyä tavukoodia voidaan ajaa C-kielellä toteutetulla virtuaalikoneella [Mar06]. Virtuaalikone tulkitsee ajettaessa tavukoodia lause kerrallaan. Pythonin standardikirjasto on toteutettu osittain C:llä ja osittain Pythonilla.

Tässä luvussa käydään läpi Python-kielen perusteita ja metaohjelmoinnin kannalta olennaisia asioita.

2.1 Syntaksi

Python ohjelma koostuu loogisista riveistä, jotka ovat yhden tai useamman ”fyysisen” rivin mittaisia. [Mar06]. Loogisten rivien päättämiseen ei käytetä mitään merkkiä. Rivien sisennyksen perusteella erotetaan ohjelmakoodin lohkot toisistaan. Suositeltu tapa sisentää on käyttää ensimmäisen tason sisentämiseen 4 välilyöntiä ja seuraavaan 8 ja niin edelleen [VWC13].

Pythonissa on 30 avainsanaa (keyword), jotka ovat kielen varattuja sanoja. Näitä ovat esimerkiksi funktio `print` ja kielen rakenteissa käytetyt sanat kuten `if`, `and` ja `class`. Pythonin standardikirjasto koostuu sisäänrakennettujen funktioiden lisäksi kokoelmasta eri tarkoituksiin soveltuvia moduuleita (module), jotka pitää tarvittaessa tuoda erikseen osaksi suoritettavaa ohjelmaa komennolla `import`.

2.2 Muuttujat

Python on dynaamisesti tyyppitetty kieli eli muuttujien arvon tyyppiä ei tarvitse eksplisiittisesti määrittää vaan tyyppi määräytyy sen perusteella minkälainen arvo muuttujaan sijoitetaan. Muuttujan tyyppiä voi myös vaihtaa sijoittamalla siihen uuden eri tyyppisen arvon. Listaus 1 sisältää yksinkertaisen esimerkin Python-kielen syntaksista. Rivillä 1 asetetaan muuttujan `a` arvoksi merkkijono `Hello world!`, joka tulostetaan rivillä 2. Rivillä 5 tulostetaan merkkijonon `a` pituus.

Modernit IDE:t kuten PyCharm¹ pystyvät koodin perusteella usein päättelemään muuttujan tyyppin.

¹<https://www.jetbrains.com/pycharm/>

```

1  a = 'Hello world!'
2  print(a)
3  # Hello world!
4
5  print(len(a))
6  # 12

```

Listaus 1: Yksinkertainen esimerkki Python-kielen syntaksista.

2.3 Tietotyypit

Pythonissa kaikki arvot, muuttujat ja funktiot ovat olioita. Olion tyyppi määrittää mitä metodeja ja ominaisuuksia olio tarjoaa. Osa olioista on muuttumattomia (immutable) ja osa muutettavia (mutable). Python-kielessä ei ole erikseen vakioita, mutta käytäntönä on käyttää muuttujan nimessä pelkästään isoja kirjaimia, jos muuttujan arvoa ei ole tarkoitus muuttaa.

Pythonin sisäänrakennettuja tietotyyppiejä on mm. numeeriset `int` ja `float`, sekvenssityypit `list`, `str` ja `tuple`, joukko `set`, "sanakirja" `dict` sekä tiedosto `file`. Näiden lisäksi standardikirjaston moduuleista löytyy lisää tietotyyppiejä kuten `datetime` ja `array`.

2.3.1 Sekvenssit

Sisäänrakennetuista tietotyypeistä mm. lista, merkkijono ja monikko (`tuple`) ovat sekvenssejä. Sekvenssit ovat iteroitavia (iterable) olioita eli ne kykenevät palauttamaan jäseniään yksi kerrallaan. Iteroitavia olioita ovat myös muut oliot, joissa on toteutettu jäseniä palauttava `__iter__`-metodi. Iteroitavia olioita voidaan käyttää suoraan osana esimerkiksi `for`-toistolauseissa. kuten esimerkiksi lauseessa `for x in [1,2,3]: print(x)`.

Listakehitelmä (list comprehension), joukkokehiteelmä (set comprehension) ja sanakirjakehitelmä (dictionary comprehension) ovat helppoja tapoja luoda lista-, joukko- tai sanakirjaolioita jonkin syötteen perusteella. Esimerkkejä listakehitelmän käytöstä sekä joidenkin Python-kielen funktioiden käytöstä on listauksessa 2. Funktio `range` palauttaa Python 2:ssa listan ja Python 3:ssa generaattoriolion, jota voidaan käyttää listan tapaan. Funktio `all` tarkastaa kaikkien listan (tai muun iteroitavan olion) totuusarvon. Pythonissa kaikki

oliot voidaan evaluoida totuusarvoina, jolloin lukujen tapauksessa aina luku "0" evaluoituu epätodeksi ja muut luvut todeksi.

```
1 print(range(10))
2 # [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
3
4 print( [x**2 for x in range(11)] )
5 # [0, 1, 4, 9, 16, 25, 36, 49, 64, 81, 100]
6
7 some_list = [1, 4, 7, 'foo', 12, 'bar']
8 print( [x for x in some_list if str(x).isalpha()] )
9 # ['foo', 'bar']
10
11 print( # Alkulukuja
12     [x for x in range(2, 50) if all([x % y for y in range(2, x-1)])]
13 )
14 # [2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47]
```

Lista 2: Esimerkki listakehittelystä ja funktion range käytöstä.

2.4 Luokat ja oliot

Python on olio-ohjelmointikieli, joka ei kuitenkaan pakota käyttämään olio-ohjelmointia ohjelmien toteutuksessa.

Pythonissa myös luokat ovat olioita. Luokilla on joitain piirteitä kuten muodostin (constructor) metodissa `__init__`, jotka mahdollistavat niiden käyttämisen luokkina. Luokka voi periä yhden tai useamman luokan. Luokka määritellään lausekkeella `class ClassName(InheritedClass, AnotherInherited)`. Python 2:ssa luokan tulisi periä ainakin `object`, jolloin se muodostaa ns. uudentyyppisen luokan [Mar06], joka on Python 3:ssa oletuksena. Uudentyyppisten luokkien metaluokka on aina `type` [Mar06]. *Metaluokkaa* käytetään muodostimena luotaessa uusi luokkaolio [pyt14]. Metaluokkia voi luoda itse, mutta tälle on harvoin tarvetta [pyt14].

Kahdella alaviivalla metodin nimen alussa ja lopussa merkitään luokan

”maagisia” metodeita ja attribuutteja [Mar06], joita Python käyttää joissain tietyissä tilanteissa. Python-kielessä käytetään usein sanaa ”type” luokan synonyyminä [pyt14].

3 Metaohjelmointi Pythonilla

Metaohjelmointi on Pythonilla hyvin luontevaa, koska olioita voidaan tutkailla ja yleensä muokata ajon aikana vapaasti ja niistä saadaan paljon metatietoa ulos Pythonin peruskirjaston työkaluilla.

3.1 Reflektio

Funktio `type` palauttaa olion tietotyyppin [Pil04]. Tämä on yleensä luokka, mutta vanhan tyyliä Python 2 -luokilla tämä on merkkijono ”instance”. Varsinainen luokka löytyy aina olion attribuutista `__class__`, mutta Python 3:ssa `type:n` käyttö on suositus.

Funktio `dir` palauttaa listan olion attribuuteista, joihin kuuluu myös olion metodit. Funktio `getattr` ottaa parametrina merkkijonon ja palauttaa parametrin nimisen attribuutin. Tällöin voidaan esimerkiksi kutsua metodeja oliosta, jonka rakennetta ei tunnetta vielä käännösvaiheessa [Pil04]. Funktiolla `isinstance` voidaan tarkistaa onko joku olio tietyn tyyppinen.

Olioiden lyhyet kuvaukset (docstring) saa ajon aikana haettua niiden `__doc__`-attribuutista.

Olioita, luokkia ja funktioita voidaan muokata ajon aikana melko vapaasti. Esimerkki peruskirjaston `dir` funktion ylikirjoittamisesta omalla funktiolla on listauksessa 3. Esimerkin oma funktio palauttaa sille annetut argumentit tekemättä niille mitään. Tämänkaltaisesta kirjastojen ja moduuleiden osien ajonaikaisesta muokkaamisesta käytetään myös tsesta käytetään myös termiä ”monkey patching”.

Pythonin standardikirjaston `inspect`-moduuli tarjoaa työkaluja olioiden tilan tutkimiseen ohjelman ajon aikana. Moduuli `sys` tarjoaa työkaluja käytetyn Python-tulkin ja käyttöjärjestelmän tarkasteluun.


```

1  # -*- coding: utf-8 -*-
2
3  import __builtin__
4
5  def x(*args):
6      return args
7
8  print(x(1, 5, 'kissa'))
9  # (1, 5, 'kissa')
10
11 print(dir(__builtin__))
12 # ['ArithmeticError', 'AssertionError', 'AttributeError', ... ]
13
14 __builtin__.dir = x
15
16 print(dir(__builtin__))
17 # (<module '.__builtin__' (built-in)>,)

```

Listaus 3: Standardikirjaston funktion ylikirjoittaminen omalla funktiolla.

3.2 Dynaamiset luokat

Luokkia voidaan muokata ohjelman ajon aikana, jolloin vaikutus ulottuu kaikkiin jo luotuihin luokan instansseihin. On myös mahdollista luoda uusia luokkia dynaamisesti ohjelman ajon aikana `type`-metaluokan avulla [pyt14]. Kutsumalla `type`:ä kolmella parametrilla saadaan paluuarvona luokka. Parametrit ovat järjestyksessä luokan nimi, perittävät luokat ja sanakirja-tyyppinen muuttuja luokan sisällöstä [pyt14]. Esimerkiksi tyhjä luokka voidaan luoda dynaamisesti lausekkeella `C = type('C', (), {})`.

3.3 Koodin kääntäminen ajon aikana

Pythonissa on sisäänrakennetut funktiot `compile`, `eval` ja `exec`, joiden avulla voidaan kääntää ja ajaa Python-koodia ohjelman ajon aikana [Cro14, Mar06].

Esimerkki ajonaikaisesta koodin kääntämisestä on listauksessa 4. Esimerkin rivillä 1 luodaan muuttuja, joka sisältää käännettävän koodin merkkijo-

nona. Rivillä 2 käännetään koodi merkkijonosta tavukoodiksi käyttämällä funktiota `compile`, jonka parametrit ovat käännettävä koodi, koodin sisältävän tiedoston nimi, jossa käytetään merkkijonoa '`<string>`' tarkoittamaan, että koodia ei luettu tiedostosta ja käännöstila (mode), jossa '`single`' kertoo, että käännetään yksi lause (statement). Rivi 4 tulostaa `code_obj`-muuttujan merkkijonoesityksen. Rivillä 7 ajetaan tavukoodi muuttujasta `code_obj`.

```
1 code_str = 'print("Hello world!")'
2 code_obj = compile(code_str, '<string>', 'single')
3
4 print(code_obj)
5 # <code object <module> at 0x7f06341f2cb0, file "<string>", line 1>
6
7 exec(code_obj)
8 # Hello world!
```

Listaus 4: Esimerkki Python-komennon kääntämisestä tavukoodiksi ohjelman ajon aikana ja käännetyin koodin ajamisesta [Cro14].

3.4 Tavukoodin tarkastelu ja muokkaus

Standardikirjaston `dis`-moduulilla voidaan tutkia Python-tavukoodia. Listauksessa 5 eräs esimerkkifunktio sekä sen tavukoodin muodostaminen `dis`-moduulin `dis`-metodilla sekä tulostettu tavukoodi. Tavukoodi on tulostettu Python 2:lla ja se on erinäköinen Python 3:lla.

Moduuli `parser` antaa rajapinnan Python-kääntäjän sisäiseen jäsennyshuuhun ja mahdollistaa sen muokkaamisen. Tämän moduulin lisäämisen jälkeen on kuitenkin standardikirjastoon lisätty `ast`-moduuli, joka mahdollistaa kääntäjän abstraktin syntaksipuun (abstract syntax tree) luomisen annetun ohjelmakoodin perusteella ja sen muokkaamisen. Funktiolle `compile` voidaan antaa parametrina muokatun AST-puun sisältävä AST-olio ja kääntää se tavukoodiksi.

```

1  # -*- coding: utf-8 -*-
2
3  import dis
4
5  def neliot(iteroitava):
6      '''Palauttaa listan parametrin alkioden neliöistä'''
7      return [x**2 for x in iteroitava]
8
9  dis.dis(neliot)
10 # 4          0 BUILD_LIST          0
11 #           3 LOAD_FAST           0 (iteroitava)
12 #           6 GET_ITER
13 #      >>    7 FOR_ITER           16 (to 26)
14 #           10 STORE_FAST         1 (x)
15 #           13 LOAD_FAST         1 (x)
16 #           16 LOAD_CONST         1 (2)
17 #           19 BINARY_POWER
18 #           20 LIST_APPEND        2
19 #           23 JUMP_ABSOLUTE      7
20 #      >>    26 RETURN_VALUE

```

Lista 5: Python-tavukoodin tarkastelu dis-moduulilla.

3.5 Käännösaikainen metaohjelmointi

Käännösaikaista metaohjelmointia ei ole suoraan tuettu Pythonissa. Template-metaohjelmointi on mahdollista esimerkiksi Jinja2 -template-moottorilla (template engine) [Ron14].

Käännösaikainen metaohjelmointi on myös lisätty suoraan osaksi kahta Pythonista jatkokehitettyä kieltä, Mythoniin[Rie08] ja Convergeen[Tra05]. Mython kääntää ohjelmakoodia suoraan Python-tavukoodiksi, mutta Convergen tuottamaa koodia ajetaan kielen omalla virtuaalikoneella.

Lähteet

- [Cro14] Crosta, D.: *Exploring Python Code Objects*, 2014. <http://late.am/post/2012/03/26/exploring-python-code-objects> [17.11.2014].
- [Guo14] Guo, P.: *Python is Now the Most Popular Introductory Teaching Language at Top U.S. Universities*. Communications of The ACM Blog, heinäkuu 2014. <http://cacm.acm.org/blogs/blog-cacm/176450/fulltext> [06.11.2014].
- [HB12] Hazzard, K. ja Bock, J.: *Metaprogramming in. NET*. Manning Publications, 2012.
- [Mar06] Martelli, A.: *Python in a Nutshell*. O'Reilly Media, Inc., 2006.
- [Pil04] Pilgrim, M.: *Dive Into Python*, toukokuu 2004. <http://www.diveintopython.net/> [06.11.2014].
- [pyt14] *Metaprogramming – Python 3 Patterns, Recipes and Idioms*, 2014. <http://python-3-patterns-idioms-test.readthedocs.org/en/latest/Metaprogramming.html> [20.11.2014].
- [Rie08] Riehl, J.: *The Mython Programming Language*, 2008. <http://mython.org/> [16.11.2014].
- [Ron14] Ronacher, A.: *Jinja2 (The Python Template Engine)*, 2014. <http://jinja.pocoo.org/> [20.11.2014].
- [Tra05] Tratt, L.: *Compile-time meta-programming in a dynamically typed OO language*. Teoksessa *Proceedings Dynamic Languages Symposium*, sivut 49–64, October 2005.
- [VWC13] Van Rossum, G., Warsaw, B. ja Coghlan, N.: *PEP 8 – Style guide for python code*. 2013. <http://www.python.org/dev/peps/pep-0008> [11.11.2014].