

hyväksymispäivä arvosana

arvostelija

## **COBOL ja Python: Tyypitys ja laskennan kapselointi**

Erkki Heino, Tero Huomo, Eeva Terkki

Helsinki 13.2.2013

HELSINGIN YLIOPISTO

Tietojenkäsittelytieteen laitos

# Sisältö

<b>1</b>	<b>Tyypijärjestelmät</b>	<b>1</b>
1.1	COBOL . . . . .	1
1.1.1	Alkeis- ja perustyytit . . . . .	1
1.1.2	Tyypitys . . . . .	1
1.2	Python . . . . .	1
1.2.1	Alkeis- ja perustyytit . . . . .	2
1.2.2	Tyypitys . . . . .	2
<b>2</b>	<b>Laskennan kapselointi</b>	<b>5</b>
2.1	COBOL . . . . .	5
2.1.1	Parametrivälitys . . . . .	5
2.1.2	Poikkeuskäsittely ja rinnakkaisuus . . . . .	6
2.2	Python . . . . .	7
2.2.1	Parametrivälitys . . . . .	7
2.2.2	Poikkeuskäsittely ja rinnakkaisuus . . . . .	9
<b>3</b>	<b>Etuja ja haittoja</b>	<b>9</b>
	<b>Lähteet</b>	<b>11</b>

# 1 Tyyppijärjestelmät

Ohjelmointikielen tyypitys on hyvin olennaista ohjelmointikielen käytössä. Vahvalla ja heikolla tyypityksellä on omat vahvuutensa ja heikkoutensa.

## 1.1 COBOL

Vaikka COBOL näyttää ulkoapäin vahvasti tyypitetyltä ja Pythonin voisi kuvitella olevan heikosti tyypitetty — totuus on kuitenkin päinvastainen.

### 1.1.1 Alkeis- ja perustyypit

COBOLissa muuttujat ovat tyypiltään aakkosellisia (alphabetic), numeerisia (numeric), muokattu numeerinen (numeric-edited), aakkosnumeerinen (alphanumeric), kanji ja muokattu kanji (kanji-edited) [Cou99a].

Kielen eri toteutuksille yhtenäisten perustyyppien lisäksi COBOLin kääntäjien toteuttajat ja levittäjät tarjoavat usein lisää valmiita perustyypppejä [Wik13a]. Suurin osa tarjoaa esimerkiksi osoittimet ja bittikentät (bit field).

### 1.1.2 Tyypitys

Michael Coughlanin COBOL Tutorialin [Cou99a] mukaan kääntäjä ei tee tyypitarkistuksia, joten COBOL on siis heikosti tyypitetty kieli. Esimerkiksi ei-numeerinen arvo on mahdollista sijoittaa numeerista tietoa sisältävään muuttujaan. Kun väärää tietotyyppiä sisältävällä muuttujalla yritetään tehdä laskutoimituksia, COBOL-ohjelma kaatuu suorituksenaikaiseen virheeseen. Muuttujat määritelläänkin siten, että niistä annetaan ikään kuin kuva, miltä niiden pitäisi näyttää — tätä kuvaa määrittelyssä käytetty PICTURE-lauseke, joka suomentuu kuvaksi.

COBOLin oliolaajennos, Object-Oriented COBOL, sisältää sekä vahvan että heikon tyypityksen piirteitä [AC96].

## 1.2 Python

Pythonin tyypit eroavat huomattavasti COBOLin tyypeistä.

### 1.2.1 Alkeis- ja perustyytit

Pythonissa ei ole erikseen alkeistyypppejä, vaan Pythonissa kaikki on ilmaistu olioina tai olioiden välisinä suhteina [Pyt13d]. Jokaisella oliolla on *identiteetti* (identity), tyyppi ja arvo. Identiteetti on olion luonnin jälkeen muuttumaton. Identiteettiä voidaan ajatella olion osoitteena muistissa. Myös olion tyyppi on muuttumaton. Tietyissä kontrolloiduissa tilanteissa tyyppiä on mahdollista muuttaa, mutta usein tyyppimuunnoksilla voi olla odottamattomia seurauksia.

Kun kaikki on ilmaistu olioina, Pythonissa myös esimerkiksi None on olio. None on vastaavuus Java-kielen null-arvolle. Toisin kuin Pythonin None, Javan null ei kuitenkaan ole olio, eikä sillä ole tyyppiä.

Kielessä on valmiina useita kymmeniä sisäänrakennettuja tyypppejä, jotka on lueteltu Pythonin tyyppihierarkiassa [Pyt13d]. Valmiita numeerisia tyypppejä ovat esimerkiksi totuusarvot, kompleksiluvut, kokonaisluvut ja reaaliluvut. Pythonin versiossa 2 kokonaislukuja kuvaavat kiinteän mittainen `int` ja rajattoman pituisen kokonaisluvun mahdollistava `long`. Pythonin versiossa 3 ei ole kiinteän pituisia kokonaislukuja lainkaan, vaan siinä `int`-tyyppi vastaa version 2 `long`-tyyppiä [Pyt13e]. Reaalilukujen ja kompleksilukujen arvoalueet voivat riippua käytettävästä konearkkitehtuurista, virtuaalimuistin määrästä sekä kääntäjän asetuksista. Pythonissa ei ole valmiina olemassa erillistä character-tyyppiä, vaan yksittäiset merkit ovat merkkijonoja. Erilaiset laajennusmoduulit lisäävät tyypppejä Javan kirjastojen tapaan.

### 1.2.2 Tyypitys

Python on vahvasti tyypitetty kieli [Stat09]. Tietyn tyyppiselle muuttujalle ei voida tehdä toisen tyyppin operaatioita ennen eksplisiittistä tyyppimuunnosta.

```
a = 5
b = "9"
c = a + int(b)
```

Esimerkissä `b` sisältää merkkijonon "9", mutta yhteenlaskussa merkkijonosta jäsenetään kokonaisluku. Jos kokonaislukujäsennyksen jättää tekemättä, antaa ohjelma kyseisellä rivillä poikkeuksen.

Python on dynaamisesti tyypitetty kieli. Suoritusaikana muuttujan tyyppi ei ole sidottu, vaan muuttujaan voi dynaamisesti sitoa eri vaiheessa eri tyyppisiä olioita.

ta. Seuraavassa esimerkissä muuttuja `a` saa ensin kokonaislukuarvon 5. Sen jälkeen muuttujan `a` arvoksi muutetaan merkkijono "Hei".

```
a = 5  
a = "Hei"
```

Pythonissa käytetään *duck typing* -ohjelmointiparadigmaa [Wik13b]. Olioiden tyyppiä tärkeämpää ovat niiden toteuttamat ominaisuudet ja metodit. Käytännössä tämä tarkoittaa sitä, että funktioiden parametreina annettujen olioiden tyyppiä ei tarkisteta suoritettavassa funktiossa. Jos parametrina ollut olio ei toteuta funktion parametrilta kutsuttavia metodeja, funktio signaloi suorituksenaikaisen virhetilan-

teen.

```
class Ankka:
    def vaaku(self):
        print("Vaaaak!")
    def ui(self):
        print("Ankka pysyy pinnalla.")

class Ihminen:
    def vaaku(self):
        print("Ihminen imitoi ankkaa.")
    def ui(self):
        print("Ihminen uppoaa.")
    def name(self):
        print("Arto Wikla")

def ankkalampi(ankka):
    ankka.vaaku()
    ankka.ui()

aku = Ankka()
arto = Ihminen()
ankkalampi(aku)
ankkalampi(arto)
```

Esimerkissä määritellään kaksi luokkaa, **Ankka** ja **Ihminen** sekä funktio **ankkalampi**. Funktio ei välitä siitä, onko sen parametrin tyyppinä **Ankka**, vaan sitä voi kutsua myös luokan **Ihminen** ilmentymällä, sillä luokassa toteutetaan funktion vaatimat metodit. Javassa samantapainen käyttäytyminen toteutettaisiin perinnän tai rajapintojen avulla.

Duck typingin lisäksi Pythonissa käytetään tyyppipäätelyä ennen versiota 3 [Pyt13d]. Tyyppipäätelyn säännöt vaihtelevat Pythonin varhaisempien versioiden välillä suuresti, jonka vuoksi ohjelmointikielen referenssi ei tarjoa tarkkaa tyyppipäätelysäännöstöä, vaan ainoastaan epävirallisia ohjenuoria. Pythonin versiosta 3 tyyppipäätely on kokonaan karsittu.

## 2 Laskennan kapselointi

Ohjelmiston jaottelu eri funktioihin ja aliohjelmiin on tärkeä osa ohjelmointia. Kapseloimalla laskentaa voidaan koodissa tuoda esiin erilaisia abstraktion tasoja.

### 2.1 COBOL

COBOLissa ohjelmiston struktuuri on hyvin yksinkertainen, eikä se tarjoa samantaisia moderneja toiminnallisuuksia kuin Python.

#### 2.1.1 Parametrivälitys

COBOLissa parametreja voidaan välittää joko arvo- tai viiteparametreina. Kutsuja määrittää, välitetäänkö parametri arvo- vai viitesemantiikalla. Kuvassa 1 nähdään esimerkki aliohjelmakutsuista.

IDENTIFICATION DIVISION.	IDENTIFICATION DIVISION.
PROGRAM-ID. ENSIMMAINEN.	PROGRAM-ID. TOINEN.
ENVIRONMENT DIVISION.	ENVIRONMENT DIVISION.
DATA DIVISION.	DATA DIVISION.
WORKING-STORAGE SECTION.	WORKING-STORAGE SECTION.
77 OMA PIC X(20) VALUE "EKAN ARVO".	LINKAGE SECTION.
PROCEDURE DIVISION.	77 PARAMETRI PIC X(20).
* KUTSUTAAN TOISTA (ARVOPARAMETRI)	PROCEDURE DIVISION USING PARAMETRI.
CALL "TOINEN" USING BY VALUE OMA.	MOVE "TOKAN ARVO" TO PARAMETRI.
* ENSIMMAISEN MUUTTUJA EI MUUTTUNUT	EXIT PROGRAM.
DISPLAY OMA.	
* KUTSUTAAN TOISTA (VIITEPARAMETRI)	
CALL "TOINEN" USING BY REFERENCE OMA.	
* ENSIMMAISEN MUUTTUJA MUUTTUI	
DISPLAY OMA.	
END RUN.	

Kuva 1: Parametrivälitys COBOLissa

Ohjelmassa ENSIMMAINEN kutsutaan kaksi kertaa ohjelmaa TOINEN: ensimmäisellä kerralla välitetään muuttujan OMA arvo ja toisella viite samaiseen muuttujaan. Aliohjelma TOINEN muuttaa parametrina saadun muuttujan arvoa. Ohjelma tulostaa:

EKAN ARVO  
 TOKAN ARVO

Kutsu

CALL "TOINEN" USING BY VALUE OMA

tarkoittaa, että kutsutaan ohjelmaa TOINEN ja välitetään parametrina muuttujan OMA arvo.

CALL "TOINEN" USING BY REFERENCE OMA

sen sijaan tarkoittaa, että samaiselle aliohjelmalle välitetään parametrina viite muuttujaan OMA.

### 2.1.2 Poikkeuskäsittely ja rinnakkaisuus

COBOLissa ei ole yleistä virheiden käsittelyyn käytettävää rakennetta, kuten `try...catch`-lausetta. Virheet on siis käsiteltävä toisin — esimerkiksi näin:

```
IDENTIFICATION DIVISION.
PROGRAM-ID. VIRHE.
ENVIRONMENT DIVISION.
DATA DIVISION.
WORKING-STORAGE SECTION.
77 NUMERO    PIC S9(3).
PROCEDURE DIVISION.
    ACCEPT NUMERO FROM CONSOLE.
    IF NUMERO IS NUMERIC
        DISPLAY "ANNOIT NUMERON"
    ELSE
        DISPLAY "VIRHE: ET ANTANUT NUMEROA"
    STOP RUN
END-IF.
STOP RUN.
```

Esimerkissä hyväksytään käyttäjältä syöte, minkä jälkeen tarkistetaan, onko syöte numeerinen ja tulostetaan virheviesti, jos mikäli syöte ei ole numeerinen.

COBOLista puuttuu luontainen tuki rinnakkaislaskennalle [Law97]. Osa COBOLin välittäjistä tarjoaa uudempiin COBOLin versioihin apuvälineitä rinnakkaisuuden



toteuttamista varten. Yksi tällainen on Micro Focus [Mic99].

## 2.2 Python

Pythonissa funktioiden parametrit ovat arvoparametreja. Lähestulkoon kaikki esitetään kielessä olioina, ja arvoparametrit ovat viitteitä olioihin; ne siis toimivat samalla tavalla kuin parametrina annetut oliot Javassa.

### 2.2.1 Parametrivälitys

Funktioiden parametrit ovat tavallisesti ns. *positional-or-keyword*-parametreja [Pyt13b]. Tämä tarkoittaa sitä, että funktiokutsussa voi välittää argumentin joko sen sijainnin perusteella tai käyttämällä avainsanaa. Parametrille voi määritellä oletusarvon. Tällöin vastaavan argumentin voi jättää funktiokutsusta pois ja parametrin arvoksi annetaan määrätty oletusarvo. Ilman oletusarvoa olevat parametrit on listattava funktion määritelmässä ensimmäisenä, ja niiden jälkeen tulevat oletusarvoiset parametrit.

Seuraavan funktion parametreista kahdella on oletusarvo ja yhdellä ei:

```
def tervehdi(nimi, tervehdys="Hei", viesti=None):
    print tervehdys, nimi
    if viesti:
        print viesti
```

Parametrilla `tervehdys` on oletusarvo `hei` ja parametrilla `viesti` oletusarvo `None`. Funktio tulostaa parametrit `tervehdys` ja `nimi`, ja tutkii parametrin `viesti` arvoa. Jos se on jotakin muuta kuin tyhjä merkkijono tai `None`, arvo tulostetaan.

Pythonissa funktiokutsussa funktiolle välitettävät argumentit voivat olla avainsana-argumentteja (keyword argument) tai "tavallisia" argumentteja [Pyt13b]. Avainsana-argumentissa argumenttia edeltää tunniste, joka kertoo, mistä funktion parametrissa on kyse. Avainsana-argumentit voivat olla funktiokutsussa keskenään eri järjestyksessä kuin vastaavat parametrit funktion määritelmässä. Argumentit, joissa ei käytetä parametrin tunnistetta, on annettava funktion määritelmän parametrilistauksen mukaisessa järjestyksessä.

Yllä määriteltyä funktiota `tervehdi` kutsuttaessa vain yksi argumentti on pakollinen. Funktiokutsut erilaisilla argumenteilla tuottavat seuraavanlaisia tulostuksia:

```
>>> tervehdi("Pekka", "Moi")
```

```
Moi Pekka
```

```
>>> tervehdi("Pekka", viesti="Kvaak kvaak")
```

```
Hei Pekka
```

```
Kvaak kvaak
```

```
>>> tervehdi(viesti="Opettele Pythonia", nimi="Pekka", tervehdys="Terve")
```

```
Terve Pekka
```

```
Opettele Pythonia
```

Funktio voidaan määrittää myös siten, että sitä voidaan kutsua mielivaltaisilla argumenteilla:

```
def mielivaltainen(*parametrit, **avainsanaparametrit):
```

```
    for parametri in parametrit:
```

```
        print parametri
```

```
    for avain, arvo in avainsanaparametrit.items():
```

```
        print avain, arvo
```

Funktio `mielivaltainen` tulostaa sille annetut argumentit ja avainsana-argumenttien nimet. Esimerkiksi kutsulla `mielivaltainen("eka", "toka", kolmas="kolmonen")` funktio tulostaa:

```
eka
```

```
toka
```

```
kolmas kolmonen
```

Avainsanattomat argumentit ovat saatavilla monikossa (tuple) `parametrit` ja avainsana-argumentit assosiaatiotaulussa (dictionary) `avainsanaparametrit`. Asteriskit määrittävät muodollisten parametrien tarkoituksen: yhdellä asteriskilla alkava parametri kuvaa avainsanattomia argumentteja ja kahdella asteriskilla alkava avainsana-argumentteja. Funktion määrittelyssä voidaan määrittää myös nimettyjä muodollisia parametreja, kunhan ne ovat ennen asteriskeilla varustettuja parametreja. \*-alkuinen parametri on oltava aina ennen \*\*-alkuista parametria funktion määrittelyssä [Pyt13c].

## 2.2.2 Poikkeuskäsittely ja rinnakkaisuus

Edellisessä luvussa määritellyn `tervehdi`-funktion kutsuminen esimerkiksi seuraavilla argumenteilla johtaisi suoritusaikaisiin virheisiin:

```
>>> tervehdi()
TypeError: tervehdi() takes at least 1 argument (0 given)

>>> tervehdi("Pekka", "Moi", tervehdys="Terve")
TypeError: tervehdi() got multiple values for keyword argument 'tervehdys'
```

Suoritusaikaisen virheen sattuessa Python-tulkki nostaa poikkeuksen [Pyt13a]. Poikkeus voidaan nostaa myös ohjelmakoodissa käyttämällä `raise`-lausetta. Poikkeuskäsittelijä määritellään `try...except`-lauseella, joka muistuttaa Javan `try...catch`-rakennetta. Poikkeuskäsittelijällä voi olla useita `except`-haaroja, ja Javan tavoin tilanteeseen sopiva haara valitaan poikkeuksen luokan mukaan.

Pythonissa on tuki rinnakkaiselle ohjelmoinnille. Kielen standardikirjaston `threading`-moduuli tarjoaa tähän esimerkiksi säikeet, lukot ja semaforit [Pyt13g], ja `Queue`-moduuli tarjoaa turvallisen tavan säikeiden käsittelyyn [Pyt13f]. CPythonissa, joka on laajimmin käytetty Pythonin toteutus, säikeet eivät kuitenkaan voi olla suoritettavina samanaikaisesti, sillä CPythonin muistinhallinta ei ole säieturvallinen [PyW13].

## 3 Etuja ja haittoja

Pythonissa laajasti hyödynnetty duck typing mahdollistaa samojen funktioiden käyttämisen useille olioille ilman rajapintoja ja perintää. Tyypitarkistusten puuttuessa duck typing kuitenkin turvautuu siihen, että sitä hyödyntävät funktiot ja ohjelmakoodi ovat luotettavasti dokumentoituja ja testattuja. Esimerkiksi Javan rajapinnat ja perintä toimivat osaltaan selkeytyksenä ja dokumentaationa sille, mitä eri metodit vaativat parametreiltaan.

Pythonin parametrien oletusarvot mahdollistavat usein lyhyemmän ja selkeämmän ilmaisutavan kuin esimerkiksi Java, jossa oletusarvojen puutetta korvataan kuormittamalla metodeja. Jos samasta metodista tarvitsee neljä eri vaihtoehtoa, joutuu Java-ohjelmoija luomaan neljä samannimistä metodia eri parametrivaihtoehtoilta. Pythonissa riittää yksi, jossa on ilmaistu vapaaehtoisten parametrien oletusarvot.

COBOLin rajattu tuki rinnakkaisohjelmoinnille hankaloittaa tehokkaiden rinnakkaisuuteen perustuvien ohjelmien ohjelmoimista kielellä. COBOL-ohjelmia käytetään usein tietokantojen kanssa. Tällöin ohjelmasta voi käynnistää useita instansseja, ja relaatiotietokannan toteutus huolehtii yhteiskäyttöisen tietokannan eheydestä ja tiedon muokkaamisen rinnakkaisuudesta.

## Lähteet

- AC96 Arranga, E. C. ja Coyle, F. P., *Object-Oriented COBOL*. Cambridge University Press, New York, NY, USA, 1983. painos, 1996.
- Cou99a Cobol tutorial: Selection, <http://www.csis.ul.ie/cobol/course/DataDeclaration.htm>. [12.2.2013].
- Wik13b Duck typing, [http://en.wikipedia.org/wiki/Duck\\_typing](http://en.wikipedia.org/wiki/Duck_typing). [11.2.2013].
- Pyt13a Python language reference – execution model, <http://docs.python.org/2/reference/executionmodel.html#exceptions>. [12.2.2013].
- PyW13 Python wiki – globalinterpreterlock, <http://wiki.python.org/moin/GlobalInterpreterLock>.
- Pyt13b Glossary, <http://docs.python.org/2/glossary.html>. [12.2.2013].
- Law97 LAWLIS, P. K., Guidelines for choosing a computer language: Support for the visionary organization, <http://archive.adaic.com/docs/reports/lawlis/m.htm>, 1997. [13.2.2013].
- Mic99 Programmer's guide to writing programs, <http://supportline.microfocus.com/documentation/books/oc41books/prconc.htm>. [13.2.2013].
- Pyt13c The python tutorial - 4. more control flow tools, <http://docs.python.org/2/tutorial/controlflow.html>. [12.2.2013].
- Pyt13d Python v2.7.3 documentation language reference - data model, <http://docs.python.org/2/reference/datamodel.html>. [12.2.2013].
- Pyt13e Python v3.3 documentation language reference - data model, <http://docs.python.org/3.3/reference/datamodel.html>. [12.2.2013].
- Pyt13f The python standard library – queue – a synchronized queue class, <http://docs.python.org/2/library/queue.html>.
- Stat09 Static vs. dynamic typing of programming languages, <http://pythonconquerstheuniverse.wordpress.com/2009/10/03/static-vs-dynamic-typing-of-programming-languages/>, 2009. [11.2.2013].

- Pyt13g      The python standard library – threading, <http://docs.python.org/2/library/threading.html>. [13.2.2013].
- Wik13a      Cobol, <http://en.wikipedia.org/wiki/COBOL>. [11.2.2013].