Verifying commenting conventions

Bachelor’s Thesis – First Presentation, 30/03/2021

By Michael Dooley

Supervised by Pooja Rani and Dr. Nataliia Stulova

University of Bern

Motivation

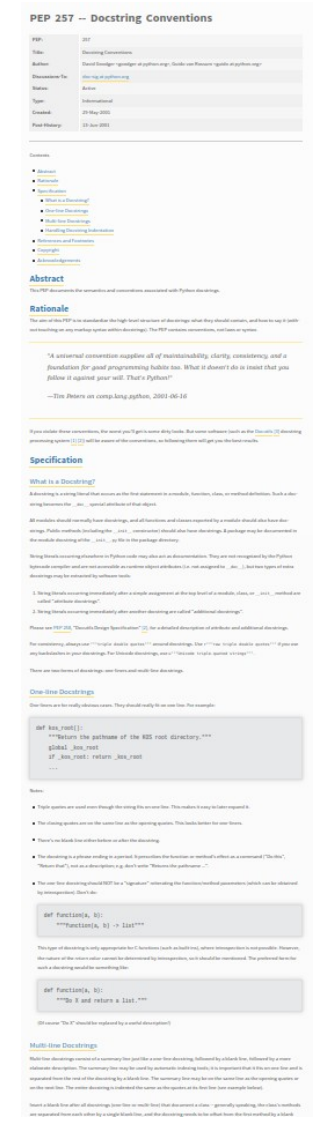
- Style Guidelines cover comments
- Style Checkers check comments

...but how?

Style Guidelines

- Codification of best stylistic practices

- For comments: Cover both content and style



PEP 257 -- Docstring Conventions

Abstract

This PEP documents the semantics and conventions associated with Python docstrings.

Rationale

The aim of this PEP is to standardize the high-level structure of docstrings: what they should contain, and how to say it (without touching on any markup syntax within docstrings). The PEP contains conventions, not laws or syntax.

> "A universal convention supplies all of maintainability, clarity, consistency, and a foundation for good programming habits too. What it doesn't do is insist that you follow it against your will. That's Python!"
>
> —Tim Peters on comp.lang.python, 2001-06-16

Specification

What is a Docstring?

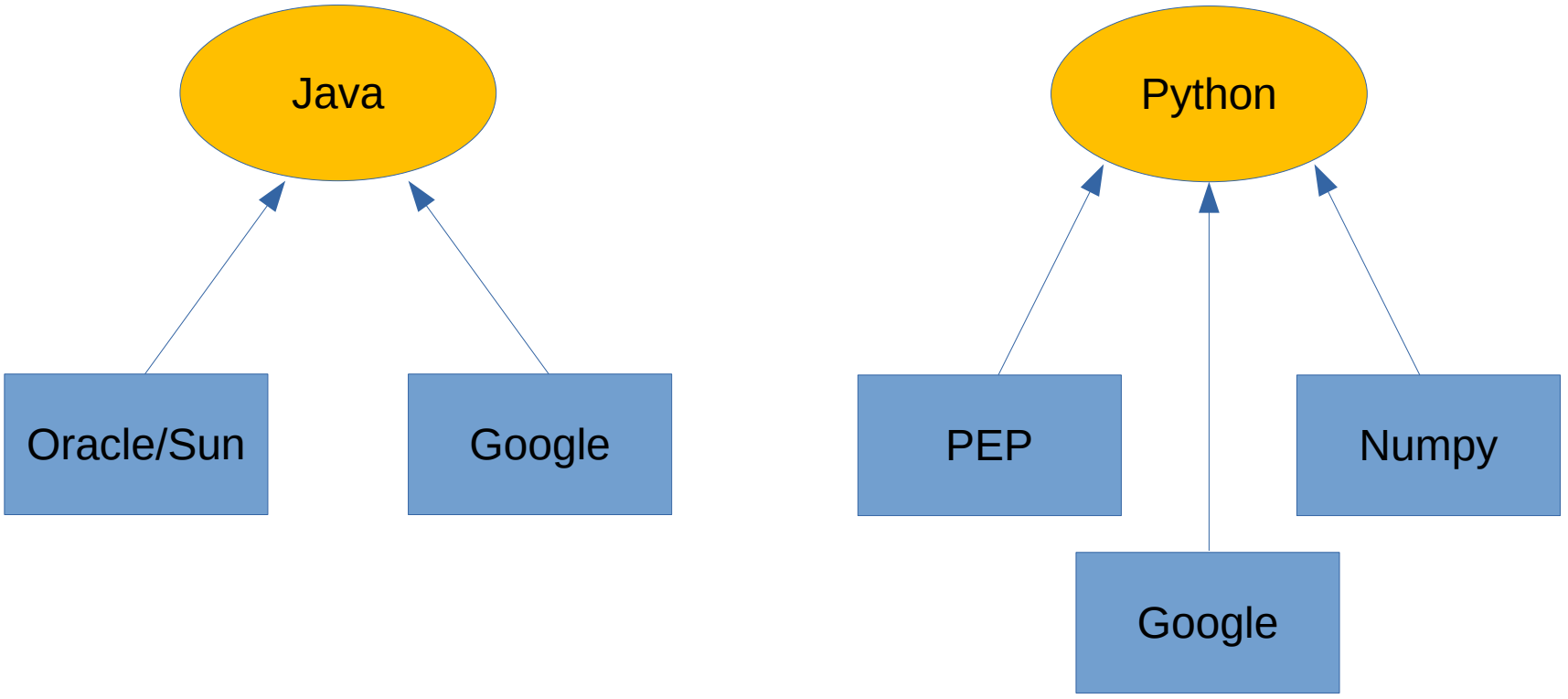
One-line Docstrings

```
def kos_root():
    """Return the pathname of the KOS root directory."""
    global _kos_root
    if _kos_root: return _kos_root
    ...
```

```
def function(a, b):
    """function(a, b) -> list"""
```

```
def function(a, b):
    """Do X and return a list."""
```

Multi-line Docstrings


Java
Oracle/Sun
Google
Python
PEP
Numpy
Google

Style Guideline differences

PEP	Google
Use 4 spaces per indentation level.	Indent your code blocks with *4 spaces*.

Style Guideline differences

PEP	Google
Unless the entire docstring fits on a line, place the closing quotes on a line by themselves.	-

Style Guideline differences

PEP	Google
Limit all lines to a maximum of 79 characters. For flowing long blocks of text with fewer structural restrictions (docstrings or comments), the line length should be limited to 72 characters.	Maximum line length is 80 characters.

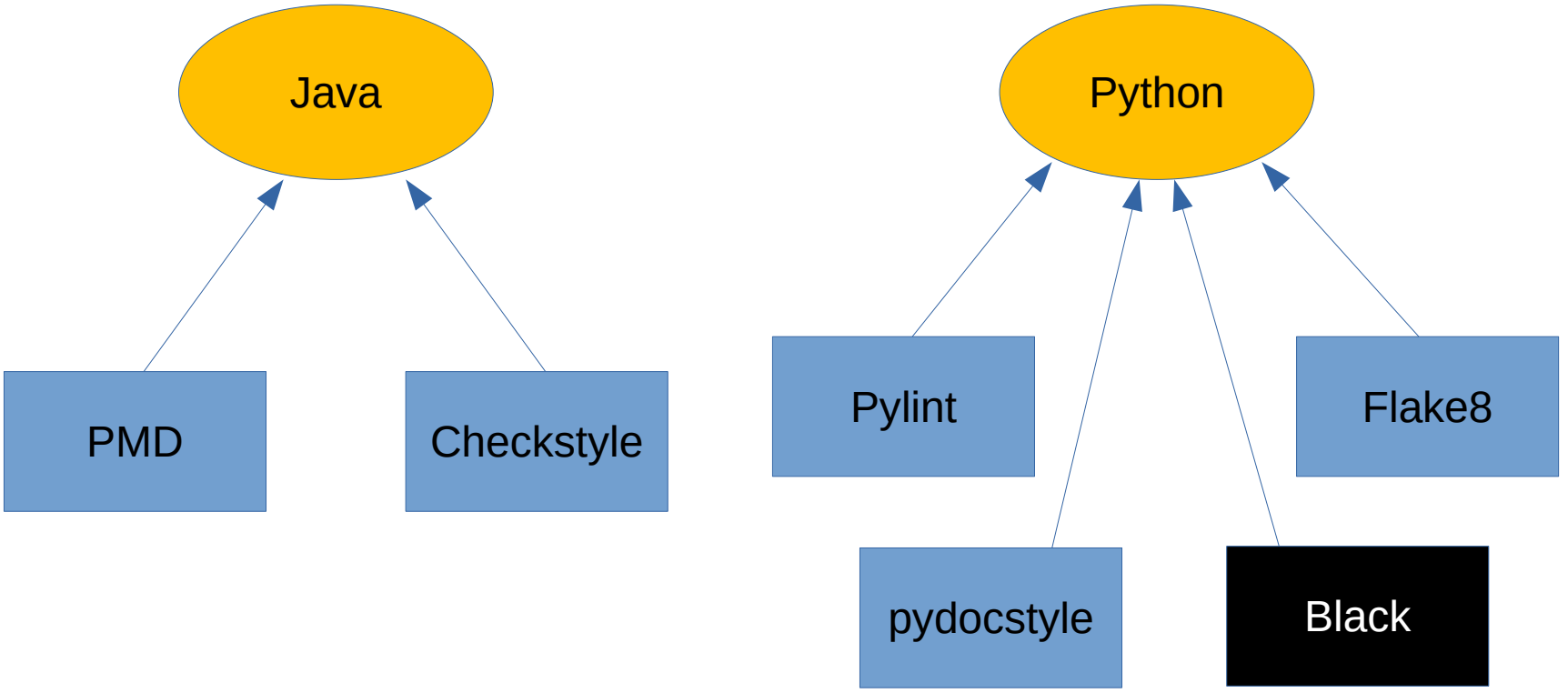
Style Checkers

Programs, that check code for consistent style

Style Checker Rules

Example (Pylint)
empty-docstring (C0112) Used when a module, function, class or method has an empty docstring (it would be too easy ;).

Rules are often toggle-able or configurable


Java
PMD
Checkstyle
Python
Pylint
pydocstyle
Black
Flake8

Question 1

What parts of style guidelines do style checkers cover with regards to documentation comments?

Related: How do style checkers differ compared to each other?

Example (PEP)

The docstring is a phrase ending in a period.

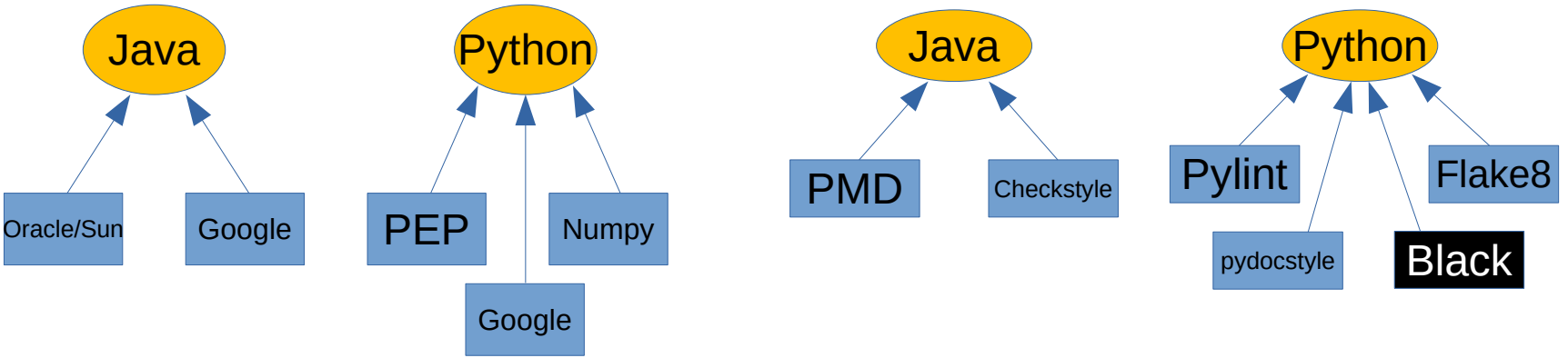
Question 2

What style checkers are used in open-source projects and how are they configured for comments?

Pipeline

1. Q1: Select guidelines and style checkers
2. Q1: Create a rule mapping
3. Q2: Mine projects and see if they use style checkers
4. Q2: Look at config files and compare them to defaults
5. Document and interpret results

Summary



What parts of style guidelines do style checkers cover with regards to documentation comments?

What style checkers are used and how are they configured for comments?