

Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Marcin Potkański

Nr albumu: 209402

**Analiza porównawcza bajtkodów i
maszyn wirtualnych świata języków
obiektowych: Java vs Python.**

**Praca magisterska
na kierunku INFORMATYKA**

Praca wykonana pod kierunkiem
dr. Jacka Chrzęszcza
Instytut Informatyki

Warszawa, Kwiecień 2025

Streszczenie

W pracy porównano bajtkody Javy i Pythona. Omówiono również maszyny wirtualne, na których wykonywany jest kod pośredni tych języków. Szacuje się czas wykonania bajtkodów tych samych programów napisanych w badanych językach programowania na maszynach wirtualnych przy różnych ustawieniach optymalizacji. Programy te pochodzą ze strony [geeksforgeeks \[6\]](#) a także CLBG [\[7\]](#). Wskazano co sprawia, że Java jest szybsza od Pythona. Omówiono JIT wykorzystywany przez JVM jak i jego odpowiednik w Pythonie, który pojawił się dopiero w wersji 3.13. Wspomniano o optymalizacjach. W pracy skupiono się na najpopularniejszej (napisanej w C) implementacji języka Python jaką jest CPython.

Słowa kluczowe

Java, Python, JVM, PVM, maszyna wirtualna, bajtkod

Dziedzina pracy (kody wg programu Socrates-Erasmus)

11.3 Informatyka

Klasyfikacja tematyczna

D. Software

Tytuł pracy w języku angielskim

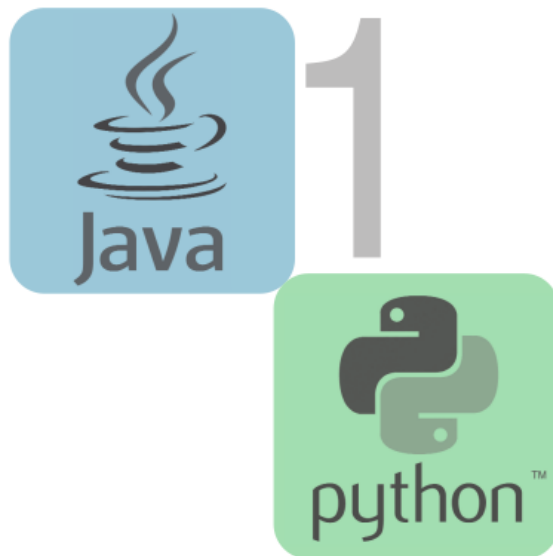
Comparative analysis of bytecodes and virtual machines in the world of object-oriented languages: Java vs Python.

Spis treści

| | |
|---|----|
| 1. Wprowadzenie | 7 |
| 1.1. Historia Javy | 7 |
| 1.2. Historia Pythona | 9 |
| 1.3. Popularność i znaczenie | 10 |
| 2. Podstawowe pojęcia | 12 |
| 2.1. Maszyna wirtualna | 12 |
| 2.1.1. Rodzaje maszyn wirtualnych | 12 |
| 2.2. Bajtkod | 12 |
| 3. Omówienie instrukcji bajtkodów | 14 |
| 3.1. Instrukcje bajtkodu Javy | 14 |
| 3.2. Struktura bajtkodu Javy | 15 |
| 3.3. Instrukcje bajtkodu Pythona | 16 |
| 3.4. Struktura bajtkodu Pythona | 17 |
| 3.5. Porównanie bajtkodów | 17 |
| 4. Wstępne porównanie JVM i PVM | 22 |
| 4.1. Języki programowania: | 22 |
| 4.2. Statyczne vs dynamiczne typowanie: | 22 |
| 4.3. Zarządzanie pamięcią: | 22 |
| 4.4. Niezależność od platformy: | 23 |
| 4.5. Wydajność: | 23 |
| 5. Architektura | 25 |
| 5.1. Architektura maszyny wirtualnej Javy | 25 |
| 5.2. Class Loader Subsystem | 25 |
| 5.2.1. Loading | 25 |
| 5.2.2. Linking | 26 |
| 5.2.3. Initialization | 26 |
| 5.3. Runtime Data Area | 26 |
| 5.3.1. Method Area | 26 |
| 5.3.2. Heap Area | 27 |
| 5.3.3. Stack Area | 27 |
| 5.3.4. PC Register | 27 |
| 5.3.5. Native Method stacks | 27 |
| 5.4. Execution Engine | 27 |
| 5.4.1. Interpreter | 27 |
| 5.4.2. JIT Compiler | 27 |

| | | |
|------------|--|-----------|
| 5.4.3. | Garbage Collector: | 28 |
| 5.4.4. | Java Native Interface (JNI): | 28 |
| 5.4.5. | Native Method Libraries: | 28 |
| 5.5. | Architektura maszyny wirtualnej Pythona | 28 |
| 5.5.1. | Code object | 29 |
| 5.5.2. | Function object | 29 |
| 5.5.3. | Frame object | 30 |
| 5.5.4. | Thread state | 31 |
| 5.5.5. | Interpreter state | 31 |
| 5.5.6. | Runtime state | 32 |
| 5.5.7. | Evaluation loop | 32 |
| 6. | Porównanie JIT Python'a i Javy. | 35 |
| 6.1. | JIT Javy | 35 |
| 6.1.1. | Co to jest kompilacja Just-In-Time (JIT)? | 35 |
| 6.1.2. | Jak działa JIT: | 35 |
| 6.1.3. | Optymalizacje stosowane przez kompilatory JIT: | 36 |
| 6.1.4. | Zalety kompilatora JIT: | 37 |
| 6.1.5. | Wady kompilatora JIT: | 37 |
| 6.2. | JIT Pythona | 38 |
| 7. | Zarządzanie pamięcią | 40 |
| 7.1. | Memory Allocation and Object Creation: | 40 |
| 7.1.1. | JVM | 40 |
| 7.1.2. | PVM | 40 |
| 7.2. | Garbage Collection: | 41 |
| 7.2.1. | JVM | 41 |
| 7.2.2. | PVM | 41 |
| 8. | Środowisko uruchomieniowe | 43 |
| 9. | Porównanie czasu działania prostych fragmentów kodu | 45 |
| 9.1. | Pusta pętla | 45 |
| 9.2. | Dodawanie w pętli jednej liczby | 47 |
| 9.3. | Duże liczby | 48 |
| 9.4. | Kolekcje | 50 |
| 9.4.1. | Java Array vs Python List, NumPy Array | 50 |
| 9.4.2. | Dodawanie String'ów do kolekcji | 51 |
| 9.4.3. | Sprawdzanie czy kolekcja zawiera Stringi | 57 |
| 9.4.4. | Usuwanie String'ów z kolekcji | 57 |
| 10. | Porównanie czasu działania programów | 60 |
| 10.1. | Narzędzia | 60 |
| 10.2. | The Computer Language Benchmarks Game | 60 |
| 10.2.1. | PIDIGITS | 60 |
| 10.2.2. | binary-trees | 62 |
| 10.2.3. | fannkuch-redux | 65 |
| 10.2.4. | N-body | 67 |
| 10.2.5. | fasta | 69 |

| | |
|--|-----------|
| 10.2.6. reverse-complement | 71 |
| 10.2.7. mandelbrot | 73 |
| 10.2.8. spectral-norm | 75 |
| 10.3. Przerobione programy ze strony geeksforgeeks.org | 77 |
| 10.3.1. Mergesort | 77 |
| 10.3.2. Problem n hetmanów | 82 |
| 11.Podsumowanie | 88 |
| Bibliografia | 89 |



Rozdział 1

Wprowadzenie

Java i Python to bardzo popularne obecnie języki programowania. Na rynku pracy jest duże zapotrzebowanie na programistów zarówno jednego, jak i drugiego języka. W niniejszej pracy postanowiłem porównać bajtkody i maszyny wirtualne owych języków w celu wskazania szczegółów, którymi się różnią. Porównałem je też pod względem zużycia pamięci.

Praca składa się z dziesięciu rozdziałów. W rozdziale 2 przypomniano podstawowe pojęcia związane z językami Java i Python takie jak bajtkod czy maszyny wirtualne. Omówienie instrukcji bajtkodów zawarto w rozdziale 3. W rozdziale 4 dokonano wstępnego porównania JVM i PVM. W rozdziale 5 omówiono architekturę maszyn wirtualnych Javy i Pythona. W rozdziale 6 porównano JIT Javy i Pythona. W rozdziale 7 przedstawiono zarządzanie pamięcią. W rozdziale 8 wspomniano o środowisku uruchomieniowym. Rozdział 9 zawiera porównanie czasu działania prostych fragmentów kodu. W rozdziale 10 porównano czas działania większych programów. Rozdział 11 zawiera podsumowanie.

1.1. Historia Javy

James Gosling, Mike Sheridan i Patrick Naughton zainicjowali projekt języka Java w czerwcu 1991 r. Java została pierwotnie zaprojektowana dla telewizji interaktywnej, ale była wówczas zbyt zaawansowana dla branży cyfrowej telewizji kablowej. Gosling zaprojektował Javę ze składnią w stylu C/C++, którą programiści systemów i aplikacji mogliby uznać za znajomą.

Firma Sun Microsystems wydała pierwszą publiczną implementację jako Java 1.0 w 1996 roku. Zapewniała funkcjonalność WORA (ang. Write Once, Run Anywhere), udostępniała darmowe środowisko uruchomieniowe na popularnych platformach. Dość bezpieczna i wyposażona w konfigurowalne zabezpieczenia pozwalała na ograniczenia dostępu do sieci i plików. Popularne przeglądarki internetowe wkrótce włączyły możliwość uruchamiania apletów Javy na stronach internetowych, a Java szybko stała się popularna. Kompilator Java 1.0 został ponownie napisany w Javie przez Arthura van Hoffa w celu ścisłej zgodności ze specyfikacją języka Java 1.0.

Wraz z pojawieniem się desktopowej wersji Java 2 (wydanej początkowo jako J2SE 1.2 w grudniu 1998 - 1999, przemianowanej w 2004 na Java SE), nowe wersje miały wiele konfiguracji zbudowanych dla różnych typów platform. J2EE zawierała technologie i interfejsy API dla aplikacji korporacyjnych, które zwykle działają w środowiskach serwerowych, natomiast J2ME zawierało interfejsy API zoptymalizowane pod kątem aplikacji mobilnych.

| Rok | Wersja | Zmiany |
|------|------------|--|
| 1995 | | Java przeznaczona dla telewizji interaktywnej |
| 1995 | | The Green Team inicjuje rozwój Javy |
| 1995 | | „Oak” - początkowa nazwa Javy |
| 1995 | | Zmieniono nazwę na „Java” ze względu na problemy ze znakiem towarowym |
| 1996 | JDK 1.0 | Pierwsza oficjalna wersja, podstawowe funkcje języka (OOP), JVM, podstawowe biblioteki |
| 1997 | JDK 1.1 | Istotne usprawnienia, włączając inner classes, JavaBeans, JDBC, and RMI. |
| 1998 | J2SE 1.2 | Przemianowano na "Java 2 Platform, Standard Edition"(J2SE). Dodano Swing GUI toolkit, Collections framework, JIT compiler. |
| 2000 | J2SE 1.3 | Poprawiona wydajność, Java Naming i Directory Interface (JNDI). |
| 2002 | J2SE 1.4 | Dalsza poprawa wydajności. Wprowadzono regular expressions, XML parsing, Non-blocking I/O, Logging API. |
| 2004 | Java SE 5 | Przemianowano na „Java Platform, Standard Edition” (Java SE) Główna aktualizacja: Generics, metadata annotations, ulepszona for loop. |
| 2006 | Java SE 6 | Skupienie na poprawie wydajności, scripting support (via JSR 223). |
| 2011 | Java SE 7 | Dodano try-with-resources, NIO 2.0, i the diamond operator. |
| 2014 | Java SE 8 | Lambda expressions, Stream API, new Date/Time API, default methods. |
| 2017 | Java SE 9 | Wprowadzono modularność (Project Jigsaw), JShell, HTTP/2 support. |
| 2018 | Java SE 10 | Local-variable type inference (var), JEP 286. |
| 2018 | Java SE 11 | LTS release, new HTTP client, usunięto moduły Java EE. |
| 2019 | Java SE 12 | Shenandoah garbage collector, JVM constants API. |
| 2019 | Java SE 13 | Text Blocks (preview), dynamic CDS archives. |
| 2020 | Java SE 14 | Records (preview), helpful NPE, pattern matching dla instanceof. |
| 2020 | Java SE 15 | Sealed classes (preview), Hidden classes, Foreign Function API (incubator). |
| 2021 | Java SE 16 | Records (finalized), Pattern Matching for switch (preview). |
| 2021 | Java SE 17 | Wersja LTS, Sealed classes (finalized), silna enkapsulacja wnętrza JDK. |
| 2022 | Java SE 18 | Simple Web Server, vector API (incubator). |
| 2022 | Java SE 19 | Virtual threads (preview), Foreign function & memory API (preview). |
| 2023 | Java SE 20 | Record patterns (preview), Virtual threads (preview). |
| 2023 | Java SE 21 | Virtual threads (finalized), sequenced collections, structured concurrency (preview). |
| 2024 | Java SE 22 | Usprawnienia do projektu Loom, nowe uaktualnienia API. |
| 2024 | Java SE 23 | Usprawniony pattern matching, Foreign Function & Memory API (finalized), poprawiona skalowalność. |

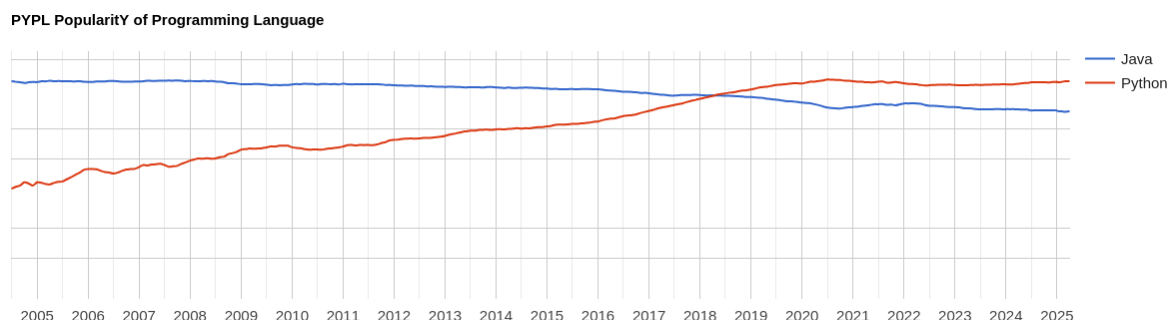
1.2. Historia Pythona

Pythona stworzył we wczesnych latach 90. Guido van Rossum – jako następcę języka ABC, stworzonego w Centrum voor Wiskunde en Informatica (CWI – Centrum Matematyki i Informatyki w Amsterdamzie). Van Rossum jest głównym twórcą Pythona, choć spory wkład w jego rozwój pochodzi od innych osób. Z racji kluczowej roli, jaką van Rossum pełnił przy podejmowaniu ważnych decyzji projektowych, często określano go przydomkiem „Benevolent Dictator for Life” (BDFL).

Nazwa języka nie pochodzi od zwierzęcia, lecz od serialu komediowego emitowanego w latach siedemdziesiątych przez BBC – „Monty Python’s Flying Circus” (Latający cyrk Monty Pythona). Projektant, będąc fanem serialu i poszukując nazwy krótkiej, unikalnej i nieco tajemniczej, uznał tę za świetną.

| Rok | Wersja | Zmiany |
|------|---------------|---|
| 1991 | Python 0.9.0 | Wersja początkowa z podstawowymi typami danych jak list, dict, string |
| 1994 | Python 1.0 | Wprowadzono lambda, map, filter, reduce |
| 2000 | Python 2.0 | Dodano list comprehensions, wsparcie dla Unicode, garbage collection |
| 2008 | Python 3.0 | Gruntowna przebudowa, lepsze wsparcie dla Unicode, bardziej spójna składnia |
| 2018 | Python 3.7 | Data classes, async/await, context variables |
| 2020 | Python 3.8 | Walrus operator, positional-only parameters, ulepszenia f-stringów |
| 2021 | Python 3.9 | Type hinting generics, nowy parser, moduł zoneinfo |
| 2022 | Python 3.10 | Strukturalny pattern matching, precyzyjne lokalizowanie błędów |
| 2023 | Python 3.11 | Poprawa wydajności, exception groups |
| 2025 | Python 3.13.2 | Free-threaded mode (experimental), Just-In-Time (JIT) compiler, usprawniony, interaktywny interpreter, colorized tracebacks |
| 2025 | Python 3.13.3 | Nowy interaktywny shell, free-threading, JIT preview, mimalloc, aktualizacja platformy |

1.3. Popularność i znaczenie



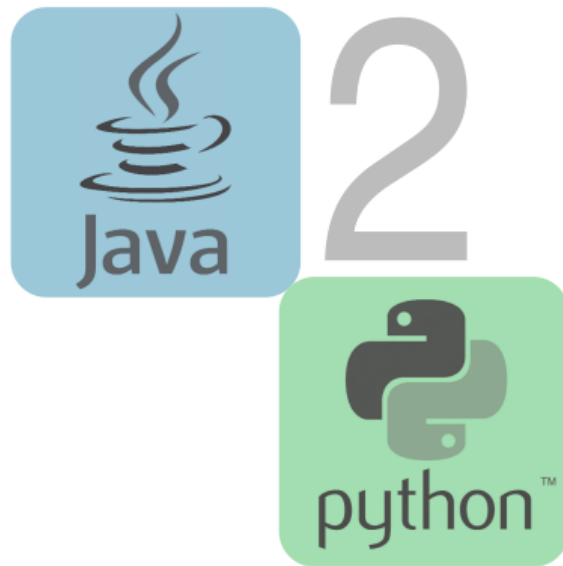
Rysunek 1.1: źródło: <https://pypl.github.io/PYPL.html>

Patrząc na załączoną grafikę widać, że na przestrzeni ostatnich dwóch dekad Java traciła na popularności a Python zyskiwał. Gdzieś w połowie 2018 roku, Python prześcignął w popularności Javę.

| Worldwide, Apr 2025 : | | | | |
|-----------------------|--------|------------|---------|--------------|
| Rank | Change | Language | Share | 1-year trend |
| 1 | | Python | 30.27 % | +1.4 % |
| 2 | | Java | 15.04 % | -0.6 % |
| 3 | | JavaScript | 7.93 % | -0.7 % |
| 4 | ↑ | C/C++ | 6.99 % | +0.6 % |
| 5 | ↓ | C# | 6.2 % | -0.6 % |

Rysunek 1.2: źródło: <https://pypl.github.io/PYPL.html>

Jak widać na załączonej grafice, Python i Java zajmują dwa pierwsze miejsca na liście popularności języków. Sprawia to, że obydwa te języki mają obecnie bardzo duże znaczenie w programowaniu. Ich łączny udział w rynku wynosi około 50% przy czym Python jest 2 razy popularniejszy od Javy (stan na kwiecień 2025);



Rozdział 2

Podstawowe pojęcia

2.1. Maszyna wirtualna

Maszyna wirtualna (ang. virtual machine, VM) – przyjęta nazwa środowiska uruchomieniowego programów.

Maszyna wirtualna kontroluje i obsługuje wszystkie odwołania uruchamianego programu bezpośrednio do sprzętu lub systemu operacyjnego. Dzięki temu program uruchomiony na maszynie wirtualnej działa jak na rzeczywistym sprzęcie.

Wykonywanym programem może być pojedyncza aplikacja, jak i cały system operacyjny lub nawet kolejna maszyna wirtualna. Odizolowanie ich przez VM od maszyny fizycznej odróżnia ją od klasycznego systemu operacyjnego.

2.1.1. Rodzaje maszyn wirtualnych

Maszyny wirtualne to m.in.:

- interpretery, szczególnie interpretery kodu bajtowego
- kompilatory JIT
- emulatory rzeczywiście istniejącego sprzętu, np. emulatory konsol.

Różnice między poszczególnymi typami takich maszyn są płynne. Np. wirtualna maszyna Javy jest powszechnie znana jako samodzielny interpreter, ale ponieważ istniały komputery, które potrafiły wykonywać programy w kodzie bajtowym Javy bezpośrednio, można ją także traktować jako emulator tych maszyn. Ponadto kompilator JIT również jest rozwiązaniem wykorzystywanym podczas interpretacji kodu bajtowego Javy.

2.2. Bajtkod

Kod bajtowy (ang. bytecode) – nazwa reprezentacji kodu używanej przez maszyny wirtualne oraz przez niektóre kompilatory. Kod składa się z ciągu instrukcji (których kody operacji mają zwykle długość jednego bajtu), które nie odpowiadają bezpośrednio instrukcjom procesora i mogą zawierać instrukcje wysokiego poziomu. W przeciwieństwie do kodu źródłowego wymagają jednak analizy tylko poszczególnych operacji.



Rozdział 3

Omówienie instrukcji bajtkodów

3.1. Instrukcje bajtkodu Javy

Każdy bajtkod składa się z jednego bajtu reprezentującego kod operacji (opcode) oraz zera lub większej liczby bajtów reprezentujących operandy.

Spośród 256 możliwych kodów operacji, w 2015 r. 202 są w użyciu ($\sim 79\%$), 51 jest zarezerwowanych do wykorzystania w przyszłości ($\sim 20\%$), a 3 instrukcje ($\sim 1\%$) są trwale zarezerwowane do wykorzystania przez implementacje JVM. Dwie z nich (**impdep1** i **impdep2**) mają na celu zapewnienie pułapek dla oprogramowania i sprzętu specyficznego dla implementacji. Trzecia jest używana przez debugery do implementowania punktów przerwania.

Instrukcje można podzielić na następujące grupy:

- ładujące i zapisujące (np. **aload_0**, **istore**),
- arytmetyczne i logiczne (np. **ladd**, **fcmpl**),
- zmieniające typ (np. **i2b**, **d2i**),
- przenoszące sterowanie (np. **ifeq**, **goto**),
- tworzące obiekt i manipulujące nim (np. **new**, **putfield**),
- wywołujące metodę i wracające z niej (np. **invokespecial**, **areturn**).

Są również instrukcje służące do wyspecjalizowanych czynności, takich jak zwracanie wyjątków, synchronizacja itd. Wiele instrukcji ma też prefiksy lub sufiksy związane z typami zmiennych, na których pracują. Oto ich lista:

| Prefiks/ Sufiks | Typ w języku Java | Opis | Przyjmowane wartości |
|--------------------|-------------------|--------------------------------------|----------------------|
| i | int | 32-bitowy typ całkowity ze znakiem | $-2^{31}..2^{31}-1$ |
| l | long | 64-bitowy typ całkowity ze znakiem | $-2^{63}..2^{63}-1$ |
| s | short | 16-bitowy typ całkowity ze znakiem | $-2^{15}..2^{15}-1$ |
| b | byte | 8-bitowy typ całkowity ze znakiem | -128..127 |
| c | char | 16-bitowy typ całkowity bez znaku | 0..65535 |
| f | float | 32-bitowy typ zmiennoprzecinkowy | N/A |
| d | double | 64-bitowy typ zmiennoprzecinkowy | N/A |
| z | boolean | 1-bitowy typ Boolowski | 0 lub 1 |
| a | Object i pochodne | Referencja do instancji klasy Object | N/A |

3.2. Struktura bajtkodu Javy

Każda instrukcja ma następującą postać:

```
1 <opcode> [ <operand1> [ <operand2>... ]]
```

W przypadku korzystania z narzędzia javap otrzymuje się plik z wpisami o następującej postaci:

```
1 <index> <opcode> [ <operand1> [ <operand2>... ]] [<comment>]
```

Przykład wywołania narzędzia javap:

```
1 class Simple {
2     public static void main(String argv[]) {
3         for (int i=1;i<=10;i++){
4             System.out.println(i);
5         }
6     }
7 }
8 $ javac Simple.java
9 $ javap -c Simple

1 Method Simple()
2     0 aload_0
3     1 invokespecial #1 <Method java.lang.Object()>
4     4 return
5 Method void main(java.lang.String[])
6     0 iconst_1
7     1 istore_1
8     2 goto 15
9     5 getstatic #2 <Field java.io.PrintStream out>
10    8 iload_1
11    9 invokevirtual #3 <Method void println(int)>
12   12 iinc 1 1
13   15 iload_1
14   16 bipush 10
15   18 if_icmple 5
16   21 return
```

Pole <index> jest indeksem kodu operacji instrukcji w tablicy zawierającej bajty kodu wirtualnej maszyny Java dla tej metody. Alternatywnie, <indeks> można traktować jako offset bajtowy od początku metody. Pole <opcode> jest mnemonikiem kodu operacji instrukcji, a zero lub więcej <operandN> to operandy instrukcji. Opcjonalny <komentarz> podaje się w składni komentarza na końcu wiersza:

```
1 8    bipush 100        // Push int constant 100
```

Pole <indeks> poprzedzające każdą instrukcję może być celem instrukcji przekazania sterowania. Na przykład instrukcja goto 8 przekazuje sterowanie do instrukcji o indeksie 8. Rzeczywiste argumenty instrukcji przesyłania sterowania wirtualnej maszyny Java są przesunięte w stosunku do adresów rozkazów (<opcode>) tych instrukcji; te operandy są wyświetlane przez javap, ponieważ łatwiej jest odczytać przesunięcia w ich metodach.

Poprzedza się operand reprezentujący indeks w run-time constant pool znakiem hash i postępuje zgodnie z komentarzem identyfikującym element run-time constant pool, do którego się odwołuje, na przykład:

```
1 10   ldc #1           // Push float constant 100.0
```

albo:

```
1 9    invokevirtual #4   // Method Example.addTwo(II)I
```

3.3. Instrukcje bajtkodu Pythona

Dokładny zestaw instrukcji zależy od wersji CPythona. W CPython'ie 3.6 wszystkie instrukcje bajtkodu zajmują dokładnie 2 bajty i mają format: <INSTRUCTION> <ARGUMENT> (po 1 bajcie). Istnieje pewien magiczny numer który określa, czy instrukcja wymaga argumentu. Jeśli liczbową reprezentacją instrukcji jest mniejsza od tego numeru to nie ma ona argumentu. W CPython'ie 3.6 tym magicznym numerem jest 90. **BINARY_POWER** nie ma argumentu, jako że jego bajtowa reprezentacja wynosi 19.

Instrukcje można podzielić na następujące grupy:

- Operacje unarne, które pobierają element ze szczytu stosu, dokonują operację a następnie umieszczają wynik na szczyście stosu (np. **UNARY_POSITIVE**, **UNARY_NOT**).
- Operacje binarne, które pobierają szczyt stosu (TOS) oraz drugi element ze szczytu stosu (TOS1), dokonują operację a następnie umieszczają wynik na szczyście stosu (np. **BINARY_POWER**, **BINARY_ADD**).
- Operacje w miejscu, które dają taki sam efekt jak operacje binarne tzn. pobierają element ze szczytu stosu (TOS) oraz drugi element ze szczytu stosu (TOS1), dokonują operację a następnie umieszczają wynik na szczyście stosu, ale są wykonywane w miejscu o ile TOS1 to obsługuje a wynikowy TOS może być oryginalnym TOS1. Przykłady takich operacji to (np. **INPLACE_POWER**, **INPLACE_ADD**)
- Operacje na listach wymagające do 3 argumentów (np. **SLICE+0**).
- Operacje na listach wymagające nawet dodatkowego argumentu. Nie umieszczają one niczego na stosie (np. **STORE_SLICE**).

```
1 end = STACK.pop()
2 start = STACK.pop()
3 container = STACK.pop()
4 values = STACK.pop()
5 container[start:end] = value
```

- Rozmaite inne operacje, niekiedy wymagające argumentów (np. **PRINT_NEWLINE**, **LOAD_CONST**).

3.4. Struktura bajtkodu Pythona

Analizuję strukturę bajtkodu Pythona na przykładzie funkcji myfunc:

```
1 def myfunc(alist):
2     return len(alist)
```

i jej kod bajtowy:

```
1 >>> dis.dis(myfunc)
2      2          0 RESUME          0
3
4      3          2 LOAD_GLOBAL      1 (NULL + len)
5          12 LOAD_FAST          0 (alist)
6          14 CALL              1
7          22 RETURN_VALUE
```

Numer w pierwszej kolumnie oznacza numer linii w kodzie źródowym Pythona. Wiele instrukcji może być przypisanych do tego samego numeru linii. Wartość ta jest obliczana na podstawie informacji z pola `co_lnotab` obiektu **code object**. Druga kolumna to offset danej instrukcji od początku bajtkodu. Zakładam, że ciąg kodu bajtowego jest zawarty w tablicy, więc wartość ta jest indeksem instrukcji w tablicy. Trzecia kolumna to rzeczywisty, czytelny kod instrukcji. Czwarta kolumna to argument instrukcji.

Instrukcja `LOAD_FAST` przyjmuje argument 0. Ta wartość jest indeksem tablicy `co_varnames`. Ostatnia kolumna to wartość argumentu - podana przez funkcję `dis`. Niektóre opcode'y nie wymagają jawnych argumentów. Zwraca się uwagę, że instrukcja `RETURN_VALUE` nie przyjmuje jawnego argumentu. Jak również, że maszyna wirtualna Pythona jest oparta na stosie, więc ta instrukcja odczytuje wartość z góry stosu.

Argumenty są przechowywane przy użyciu kodowania Little-Endian (LE) (najpierw najmniej znaczący bajt). Jeśli wartość argumentu jest mniejsza niż 65536 (0xFFFF), jest ona zapisywana jako wartość dwubajtowa, nawet jeśli jest po prostu wartością zerową.

```
1 <opcode> <2 bajty w kodowaniu LE>
```

Jeżeli wartość argumentu jest większa niż 65535, jest ona kodowana w sposób specjalny: Wstawiany jest kod bajtu `EXTENDED_ARG` (145), po którym następują cztery bajty wartości argumentu w kodowaniu LE.

```
1 <opcode> <EXTENDED_ARG> <niższe 2 bajty argumentu> <wyższe 2 bajty argumentu>
```

Takie rozszerzone argumenty zdarzają się niezwykle rzadko i mogą wystąpić wyłącznie w przypadku ogromnej funkcji zawierającej więcej niż 65535 kodów bajtowych lub w przypadku tworzenia ogromnej, wstępnie wypełnionej krotki, listy, zestawu, mapy lub wycinka (w takim przypadku kod bajtowy takiej funkcji również będzie ogromny i może wystąpić sytuacja podobna do tej w pierwszym przypadku).

Zatem rozmiar pojedynczego polecenia kodu bajtowego może zajmować 1, 3 lub 6 bajtów.

3.5. Porównanie bajtkodów

Maszyna wirtualna Javy obsługuje ponad 200 różnych bajtkodów, podczas gdy PVM rozróżnia 118 bajtkodów (CPython 3.6). Implementacje bajtkodów zarówno jednego jak i drugiego języka wykorzystują maszyny stosowe do wykonywania obliczeń. W przypadku Javy jest to **operand stack**. W Pythonie jest to **value stack** (zwany też **data stack** lub **evaluation**

stack).

Już na poziomie bajtkodów można dostrzec istotne różnice między statyczną naturą Javy a dynamiczną naturą Pythona.

Niech za przykład posłużą instrukcje **iadd** Javy i **BINARY_ADD** Pythona.

W przypadku Javy zasadne wydaje się rozpatrzenie kodu emitowanego przez tzw. CppInterpreter i TemplateInterpreter. CppInterpreter jest niezależny od platformy, natomiast TemplateInterpreter używa kodu specyficznego dla platformy.

Poniżej pokazuje się kod implementacji operacji **iadd** dla CppInterpretera (plik *bytecodeInterpreter.cpp*).

```
1 #define OPC_INT_BINARY(opcname, opname, test)
2     CASE(_i##opcname):
3         if (test && (STACK_INT(-1) == 0)) {
4             VM_JAVA_ERROR(vmSymbols::java_lang_ArithmeticException(),
5                           "/ by zero");
6         }
7         SET_STACK_INT(VMint##opname(STACK_INT(-2),
8                                     STACK_INT(-1)),
9                       -2);
10        UPDATE_PC_AND_TOS_AND_CONTINUE(1, -1);

1 OPC_INT_BINARY(add, Add, 0);
```

W pliku *bytecodeInterprete_zero.inline.hpp* znajduje się kod operacji dodawania.

```
1 inline jint BytecodeInterpreter::VMintAdd(jint op1, jint op2) {
2     return op1 + op2;
3 }
```

Poniżej pokazuje się z kolei kod implementacji operacji **iadd** dla TemplateInterpretera (plik *templateTable_x86.cpp*).

```
1 void TemplateTable::iop2(Operation op) {
2     transition(itos, itos);
3     switch (op) {
4         case add : -- pop_i(rdx); -- addl (rax, rdx); break;
5         case sub : -- movl(rdx, rax); -- pop_i(rax); -- subl (rax, rdx); break;
6         case mul : -- pop_i(rdx); -- imull(rax, rdx); break;
7         case _and : -- pop_i(rdx); -- andl (rax, rdx); break;
8         case _or : -- pop_i(rdx); -- orl (rax, rdx); break;
9         case _xor : -- pop_i(rdx); -- xorl (rax, rdx); break;
10        case shl : -- movl(rcx, rax); -- pop_i(rax); -- shll (rax); break;
11        case shr : -- movl(rcx, rax); -- pop_i(rax); -- sarl (rax); break;
12        case ushr : -- movl(rcx, rax); -- pop_i(rax); -- shr1 (rax); break;
13        default : ShouldNotReachHere();
14    }
15 }
```

Na tych przykładach widać że maszyna wirtualna Javy wymaga by liczby, na których operuje były 32-bitowymi zmiennymi a maszyna wirtualna Pythona wywołuje zaś dynamicznie odpowiedni kod aby dodać 2 obiekty na stosie bazując na ich typach czasu wykonania. Pokazano to na poniższym kodzie:

```
1 TARGET(BINARY_ADD) {
2     PyObject *right = POP();
```

```

3 PyObject *left = TOP();
4 PyObject *sum;
5
6 if (PyUnicode_CheckExact(left) &&
7     PyUnicode_CheckExact(right)) {
8     sum = unicode_concatenate(left, right, f, next_instr);
9 }
10 else {
11     sum = PyNumber_Add(left, right);
12     Py_DECREF(left);
13 }
14 Py_DECREF(right);
15 SET_TOP(sum);
16 if (sum == NULL)
17     goto error;
18 DISPATCH();
19 }

```

Python sprawdza czy lewy i prawy operand są instancjami Unicode np. stringami. Doko-
nuje tego sprawdzając typ ich obiektu. Jeśli obydwa operandy są stringami to je łączy. W
przeciwnym przypadku wywoływane jest `PyNumber_Add()`.

```

1 PyObject *
2 PyNumber_Add(PyObject *v, PyObject *w)
3 {
4     // NB_SLOT(nb_add) expands to "offsetof(PyNumberMethods, nb_add)"
5     PyObject *result = binary_op1(v, w, NB_SLOT(nb_add));
6     if (result == Py_NotImplemented) {
7         PySequenceMethods *m = Py_TYPE(v)->tp_as_sequence;
8         Py_DECREF(result);
9         if (m && m->sq_concat) {
10             return (*m->sq_concat)(v, w);
11         }
12         result = binop_type_error(v, w, "+");
13     }
14     return result;
15 }

```

`PyNumberAdd()` najpierw próbuje przeprowadzić operację dodawania na operandach `v` i
`w` (dwa wskaźniki do `PyObject`) wywołując `binary_op1(v, w, NB_SLOT(nb_add))`. Jeśli
rezultatem wywołania jest `Py_NotImplemented`, próbuje połączyć operandy jako sekwencję:

```

1 static PyObject *
2 binary_op1(PyObject *v, PyObject *w, const int op_slot)
3 {
4     PyObject *x;
5     binaryfunc slotv = NULL;
6     binaryfunc slotw = NULL;
7
8     if (Py_TYPE(v)->tp_as_number != NULL)
9         slotv = NB_BINOP(Py_TYPE(v)->tp_as_number, op_slot);
10    if (!Py_IS_TYPE(w, Py_TYPE(v)) &&
11        Py_TYPE(w)->tp_as_number != NULL) {
12        slotw = NB_BINOP(Py_TYPE(w)->tp_as_number, op_slot);
13        if (slotw == slotv)
14            slotw = NULL;
15    }
16    if (slotv) {
17        if (slotw && PyType_IsSubtype(Py_TYPE(w), Py_TYPE(v))) {
18            x = slotw(v, w);
19            if (x != Py_NotImplemented)

```

```

20         return x;
21         Py_DECREF(x); /* can't do it */
22         slotw = NULL;
23     }
24     x = slotv(v, w);
25     if (x != Py_NotImplemented)
26         return x;
27     Py_DECREF(x); /* can't do it */
28 }
29 if (slotw) {
30     x = slotw(v, w);
31     if (x != Py_NotImplemented)
32         return x;
33     Py_DECREF(x); /* can't do it */
34 }
35 Py_RETURN_NOTIMPLEMENTED;
36 }

```

Funkcja `binary_op1()` pobiera trzy parametry: lewy operand, prawy operand i offset, który identyfikuje slot. Typy obydwu operandów mogą implementować slot. Z tego powodu, `binary_op1()` sprawdza obydwie implementacje. Aby uzyskać wynik, woła jedną implementację lub drugą opierając się na następującej logice:

1. Jeśli typ jednego operandu jest podtypem drugiego, wywołaj slot podtypu.
2. Jeśli lewy operand nie ma slotu, wywołaj slot prawego operandu.
3. W przeciwnym przypadku, wywołaj slot lewego operandu.



Rozdział 4

Wstępne porównanie JVM i PVM

Do określenia maszyn wirtualnych Javy i Pythona stosujemy odpowiednio skróty: JVM (Java Virtual Machine) i PVM (Python Virtual Machine). JVM jest częścią Java Virtual Machine specification. PVM jest z kolei używana przez CPythona, standardową implementację Pythona. Maszyny te służą podobnym celom, ale jednocześnie zasadniczo się różnią.

4.1. Języki programowania:

- **JVM:** JVM został zaprojektowany specjalnie do uruchamiania kodu bajtowego Javy. Chociaż istnieją inne języki, które można skompilować do kodu bajtowego Javy (np. Kotlin, Scala), JVM obiera za cel przede wszystkim programy Javy.
- **PVM:** PVM jest przeznaczony do uruchamiania kodu bajtowego Pythona wygenerowanego przez CPythona.

4.2. Statyczne vs dynamiczne typowanie:

- **JVM:** Java jest językiem statycznie typowanym, co oznacza, że typy zmiennych są znane w czasie kompilacji. JVM wymusza ściśle sprawdzanie typów w czasie kompilacji.
- **PVM:** Python jest typowany dynamicznie, co oznacza, że typy zmiennych są określane w czasie wykonywania, co sprawia, że kodowanie jest ułatwione i bardziej elastyczne. Może jednak prowadzić do błędów w czasie wykonywania, jeśli problemy z typami nie będą wcześniej odpowiednio rozwiązane.

4.3. Zarządzanie pamięcią:

- **JVM:** JVM korzysta z automatycznego zarządzania pamięcią poprzez garbage collection. Automatycznie zwalnia pamięć dla obiektów, do których nie ma już odniesień.
- **PVM:** PVM wykorzystuje również garbage collection do zarządzania pamięcią. Zarządzanie pamięcią w Pythonie opiera się na zliczaniu odwołań, które automatycznie zwalnia pamięć, gdy liczba odwołań do obiektu spadnie do zera. Garbage Collector posiada dodatkowe algorytmy wykrywające cykle i jeśli takowe wykryje – to je przerywa. To natomiast powoduje spadek liczby referencji a w konsekwencji sprzątnięcie obiektu.

4.4. Niezależność od platformy:

- **JVM:** Maksyma Java „Write Once, Run Anywhere” jest podstawową zasadą dla języka Java, a kod bajtowy jest niezależny od platformy i jeśli na platformie istnieje implementacja JVM, można na niej uruchomić kod Java. JVM nie zmienia się co wersję Javy.
- **PVM:** Python ma również na celu niezależność platformy, ale dostępność PVM dla konkretnej platformy może nie być tak powszechna, jak implementacje JVM. Ponadto PVM zmienia się co wersję Pythona. Uważa się, że przenośność Pythona jest wystarczająco dobra.

4.5. Wydajność:

- **JVM:** Programy Java są zazwyczaj znane ze swojej wydajności i efektywności. Kompilacja Just-In-Time (JIT) maszyny JVM może zoptymalizować kod bajtowy by zwiększyć szybkość wykonywania.
- **PVM:**
Python jest ogólnie uważany za wolniejszy niż Java, szczególnie w przypadku zadań związanych z procesorem, ze względu na jego dynamiczne typowanie i interpretacyjną naturę. W wielu aplikacjach łatwość użycia Pythona i dostępne biblioteki dają przewagę nad wadami wydajnościowymi.

Podsumowując, JVM i PVM to maszyny wirtualne, które mają różne filozofie projektowania i nadają się do różnych przypadków użycia. Koncentracja Java na silnym typowaniu i wydajności sprawia, że jest popularna w aplikacjach korporacyjnych, podczas gdy prostota i wszechstronność Pythona sprawia, że jest ulubionym narzędziem do tworzenia skryptów, tworzenia stron internetowych, analizy danych.



Rozdział 5

Architektura

5.1. Architektura maszyny wirtualnej Javy

Programiści Javy wiedzą, że kod bajtowy będzie wykonywany przez JRE (Java Runtime Environment). JRE to z kolei implementacja wirtualnej maszyny Java (JVM), która analizuje kod bajtowy, interpretuje i wykonuje go.

Java została opracowana w oparciu o koncepcję WORA (Write Once Run Anywhere), która działa na maszynie wirtualnej. Kompilator kompiluje plik Java do pliku .class Javy, następnie ten plik .class jest wprowadzany do maszyny JVM, która ładuje i wykonuje plik klasy.

Jak działa JVM?

JVM jest podzielona na trzy główne podsystemy:

- Class Loader Subsystem
- Runtime Data Area
- Execution Engine

5.2. Class Loader Subsystem

Funkcja dynamicznego ładowania klas w Javie jest obsługiwana przez podsystem ładujący klasy, który ładuje, linkuje i inicjuje plik klasy przy pierwszym odwołaniu.

5.2.1. Loading

Klasy będą załadowane przez ten komponent. Bootstrap ClassLoader, Extension ClassLoader i Application ClassLoader to trzy klasy modułu ładującego, które pomogą w osiągnięciu tego celu.

- **Bootstrap ClassLoader (Primordial ClassLoader)** – Bootstrap ClassLoader to kod maszynowy odpowiedzialny za inicjowanie operacji JVM. W wersjach Java do 8 ładował podstawowe pliki Java z rt.jar. Jednak począwszy od Javy 9 ładuje podstawowe

pliki Java z Java Runtime Image (JRT). Bootstrap ClassLoader działa niezależnie, bez żadnych nadrzędnych ClassLoader'ów.

- **Platform Class Loader (Extension ClassLoader)** – W wersjach Javy przed Javą 9 istniał moduł Extension ClassLoader, ale od Javy 9 jest on nazywany modulem Platform Class Loader. Ładuje on rozszerzenia specyficzne dla platformy z systemu modułów JDK. Platform Class Loader ładuje pliki z obrazu środowiska wykonawczego Java lub z dowolnego innego modułu określonego przez właściwość systemową `java.platform` lub `-module-path`.
- **System ClassLoader (Application ClassLoader)** – Znany również jako Application ClassLoader, ładuje klasy z classpath aplikacji. Jest potomkiem Platform Class Loader. Klasy są ładowane z katalogów określonych przez zmienną środowiskową `CLASSPATH`, opcję wiersza poleceń `-classpath` lub `-cp`.

5.2.2. Linking

Gdy klasa zostanie załadowana do pamięci, nadal nie jest gotowa do wykonania. Zanim jakiegokolwiek metody będą mogły zostać uruchomione, JVM musi przetworzyć klasę dalej. W tym miejscu pojawia się linkowanie.

Linkowanie następuje po załadowaniu klasy i przygotowaniu jej do wykonania. Proces ten składa się z trzech głównych kroków: weryfikacji, przygotowania i rozwiązywania.

- **Verification** – Pierwszą podfazą fazy Linking jest weryfikacja (ang. *verifying*). Podczas tej fazy sprawdzana jest poprawność kodu bajtowego. Proces ten wykonywany jest przez weryfikator kodu bajtowego. Sprawdzane są tam między innymi informacje o tym czy nie wywołujemy metod prywatnych z innych klas, czy nie nadpisujemy finalnych klas. Jeśli podczas weryfikacji wystąpi błąd to otrzymamy `VerifyError`.

Proces weryfikacji jest procesem kosztownym w kontekście ładowania klas. Natomiast wykonanie tego sprawdzania pozwala na szybsze działanie naszej aplikacji w czasie wykonania, ponieważ żadna weryfikacja kodu bajtowego nie musi już zachodzić.

- **Preparation** – Wszystkim zmiennym statycznym przydzielana jest pamięć i wartości domyślne.
- **Resolving** – Wszystkie symboliczne odniesienia do pamięci są zastępowane początkowymi odniesieniami z `Method Area`.

5.2.3. Initialization

Jest to ostatnia faza ładowania klas, podczas której wszystkim zmiennym statycznym przypisane są początkowe wartości i zostaje wykonany blok statyczny.

5.3. Runtime Data Area

The Runtime Data Area jest podzielona na 5 głównych komponentów:

5.3.1. Method Area

Wszystkie dane na poziomie klasy są tutaj przechowywane, łącznie ze zmiennymi statycznymi. Na każdą maszynę JVM przypada tylko jeden Method Area i jest to zasób współdzielony.

JC: na-
zwy po
angiel-
sku?

5.3.2. Heap Area

Wszystkie obiekty i odpowiadające im zmienne instancji oraz tablice są tutaj przechowywane. Istnieje również jeden obszar sterty na każdą maszynę JVM. Ponieważ Method Area i Heap Area są wspólne dla wszystkich wątków, dostęp do danych musi być synchronizowany pomiędzy wątkami.

5.3.3. Stack Area

Dla każdego wątku zostaje utworzony oddzielny Runtime Stack. Dla każdego wywołania metody zostaje utworzony jeden wpis w pamięci stosu, nazywany Stack Frame. Wszystkie zmienne lokalne zostają utworzone w pamięci stosu. Obszar stosu jest bezpieczny dla wątków, ponieważ nie jest zasobem współdzielonym. Stack frame jest podzielony na trzy podjednostki:

- **Local Variable Array** – Tablica, w której przechowywane są wartości zmiennych lokalnych danej metody.
- **Operand Stack** – Jeśli do wykonania wymagana jest jakakolwiek operacja pośrednia, operand stack pełni rolę obszaru roboczego środowiska wykonawczego dla wykonania operacji.
- **Frame Data** – Tutaj przechowywane są wszystkie symbole odpowiadające tej metodzie. W przypadku jakiegokolwiek wyjątku, informacja o bloku catch zachowana jest we frame data.

5.3.4. PC Register

Każdy wątek ma osobny rejestr PC, aby przechowywać adres aktualnie wykonywanej instrukcji. Po jej wykonaniu rejestr PC zostaje zaktualizowany i wskazuje następną instrukcję.

5.3.5. Native Method stacks

– Native Method Stack przechowuje informacje o metodach natywnych. Dla każdego wątku tworzony jest oddzielny stos metod natywnych.

5.4. Execution Engine

Kod bajtowy przypisany do Runtime Data Area jest wykonywany przez Execution Engine, który odczytuje kod bajtowy i wykonuje go fragmentami.

5.4.1. Interpreter

Interpretuje bajtkod linia po linii odpowiednio uaktualniając struktury czasu wykonania, a następnie wykonuje. Działa dość wolno. Dodatkową wadą jest to, że gdy jedna metoda jest wywoływana wiele razy, za każdym razem wymagana jest interpretacja.

5.4.2. JIT Compiler

Służy do zwiększenia wydajności interpretera. Kompiluje kod bajtowy i zmienia go na kod natywny, więc kiedykolwiek interpreter widzi powtarzające się wywołania metod, JIT dostarcza bezpośredni kod natywny dla tej części, więc reinterpretacja nie jest wymagana, a tym samym wydajność jest zwiększona.

JC: Ale co ten interpreter robi? Bezpośrednio wykonuje każdą instrukcję odpowiednio uaktualniając struktury czasu

- **Intermediate Code generator** – tworzy kod pośredni
- **Code Optimizer** – odpowiedzialny za optymalizację kodu pośredniego wygenerowanego powyżej
- **Target Code Generator** – odpowiedzialny za generowanie kody maszynowego lub natywnego
- **Profiler** – Specjalny komponent odpowiedzialny za wyszukiwanie hotspotów, czyli czy metoda jest wywoływana wielokrotnie czy nie.

5.4.3. Garbage Collector:

Zbieracz śmieci jest odpowiedzialny za zwalnianie pamięci w stercie. Robi to, patrząc na pamięć sterty, identyfikując, które obiekty są używane, a które nie, i usuwając nieużywane obiekty. Proces ten odbywa się w trzech krokach, mianowicie:

- **Oznaczanie:** GC identyfikuje, które części pamięci sterty są używane, a które nie. Odbywa się to poprzez skanowanie wszystkich obiektów w celu zidentyfikowania obiektów odwoływanych i nieodwoływanych.
- **Zwykłe usuwanie:** Zwykłe usuwanie usuwa obiekty nieodwoływane, pozostawiając obiekty odwoływane i wskaźniki do wolnej przestrzeni.
- **Usuwanie i kompaktowanie:** Obiekty nieodwoływane są usuwane, a obiekty odwoływane są kompaktowane razem, co ułatwia i przyspiesza przydzielanie nowej pamięci.

5.4.4. Java Native Interface (JNI):

Dzięki temu Java może używać metod natywnych (C, C++) podczas pisania programów Java. JNI osiąga to poprzez interakcję z bibliotekami metod natywnych, które są implementowane w C, C++ i będą udostępniać te biblioteki silnikowi wykonawczemu. Pozwala to na bezpośredni dostęp do kodu asemblera. JVM będzie mieć dwa różne typy kodów, czyli Java i Native, a JNI utworzy „dobrze zdefiniowane” łącze między nimi.

5.4.5. Native Method Libraries:

Jest to zbiór bibliotek natywnych, które są wymagane przez silnik wykonawczy.

5.5. Architektura maszyny wirtualnej Pythona

W pracy skupiono się na najpopularniejszej (napisanej w C) implementacji języka Python jaką jest CPython.

Wywołanie programu Pythona składa się z trzech etapów:

- **Inicjalizacja**
- **Kompilacja**
- **Interpretacja**

Podczas fazy inicjalizacji, CPython inicjalizuje struktury danych wymagane do działania Pythona. Przygotowuje również typy wbudowane, konfiguruje moduły wbudowane, konfiguruje system importu i wykonuje wiele innych czynności.

Teraz następuje faza translacji. CPython jest interpreterem, a nie kompilatorem w tym sensie, że nie tworzy kodu maszynowego. Zamiast tego, jak wiele innych interpreterów, przed wykonaniem tłumaczy kod źródłowy do postaci pośredniej, aby zmniejszyć potem narzut choćby na parsowanie źródeł. Etap translacji wykonuje te same czynności jakie wykonuje typowy kompilator: parsuje kod źródłowy, buduje AST (Abstract Syntax Tree), generuje bajtkod z AST i wykonuje nawet pewne optymalizacje bajtkodu.

Podczas fazy interpretacji następuje wykonanie bajtkodu przez maszynę wirtualną. Wykonanie to ma miejsce w ogromnej pętli, która działa póki są instrukcje do wykonania. Zatrzymuje się ona by pobrać wartość lub gdy wystąpił błąd.

5.5.1. Code object

Fragmenty kodu, które wykonywane są jako pojedyncza jednostka takie jak moduł lub treść funkcji zwane są blokami kodu. CPython przechowuje informacje na temat tego co kod bloku robi w strukturze zwanej obiektem kodu. Zawiera ona bajtkod i dane takie jak lista nazw zmiennych używanych w bloku. Uruchomienie modułu lub wywołanie funkcji oznacza start obliczeń odpowiadającego obiektu kodu.

Poniżej przedstawiam definicję struktury obiektu kodu.

```
1 struct PyCodeObject {
2     PyObject_HEAD
3     int co_argcount;           /* #arguments, except *args */
4     int co_posonlyargcount;    /* #positional only arguments */
5     int co_kwonlyargcount;    /* #keyword only arguments */
6     int co_nlocals;           /* #local variables */
7     int co_stacksize;         /* #entries needed for evaluation stack */
8     int co_flags;             /* CO_..., see below */
9     int co_firstlineno;       /* first source line number */
10    PyObject *co_code;         /* instruction opcodes */
11    PyObject *co_consts;       /* list (constants used) */
12    PyObject *co_names;        /* list of strings (names used) */
13    PyObject *co_varnames;     /* tuple of strings (local variable names) */
14    PyObject *co_freevars;     /* tuple of strings (free variable names) */
15    PyObject *co_cellvars;     /* tuple of strings (cell variable names) */
16
17    Py_ssize_t *co_cell2arg;    /* Maps cell vars which are arguments. */
18    PyObject *co_filename;     /* unicode (where it was loaded from) */
19    PyObject *co_name;         /* unicode (name, for reference) */
20    /* ... more members ... */
21 };
```

5.5.2. Function object

Funkcje są nie tylko obiektami kodu. Muszą zawierać dodatkowe informacje takie jak nazwa funkcji, domyślne argumenty i wartości zmiennych zdefiniowanych w ich zasięgu. Te informacje, razem z obiektem są przechowywane w obiekcie funkcji.

Poniżej przedstawiam definicję struktury obiektu funkcji.

```

1 typedef struct {
2     PyObject_HEAD
3     PyObject *func_code;           /* A code object, the __code__ attribute */
4     PyObject *func_globals;        /* A dictionary (other mappings won't do) */
5     PyObject *func_defaults;       /* NULL or a tuple */
6     PyObject *func_kwdefaults;     /* NULL or a dict */
7     PyObject *func_closure;        /* NULL or a tuple of cell objects */
8     PyObject *func_doc;            /* The __doc__ attribute, can be anything */
9     PyObject *func_name;           /* The __name__ attribute, a string object */
10    PyObject *func_dict;            /* The __dict__ attribute, a dict or NULL */
11    PyObject *func_weakreflist;     /* List of weak references */
12    PyObject *func_module;          /* The __module__ attribute, can be anything */
13    PyObject *func_annotations;     /* Annotations, a dict or NULL */
14    PyObject *func_qualname;        /* The qualified name */
15    vectorcallfunc vectorcall;
16 } PyFunctionObject;

```

5.5.3. Frame object

Kiedy maszyna wirtualna wykonuje obiekt kodu, musi śledzić wartości zmiennych i stale zmieniający się stos wartości. Musi także pamiętać, gdzie zatrzymała wykonywanie bieżącego obiektu kodu, aby wykonać inny i dokąd ma się udać po powrocie. CPython przechowuje te informacje wewnątrz obiektu ramki lub po prostu w ramce. Ramka zapewnia stan, w którym można wykonać obiekt kodu.

Poniżej przedstawiam definicję obiektu ramki.

```

1 struct _frame {
2     PyObject_VAR_HEAD
3     struct _frame *f_back;         /* previous frame, or NULL */
4     PyCodeObject *f_code;          /* code segment */
5     PyObject *f_builtins;          /* builtin symbol table (PyDictObject) */
6     PyObject *f_globals;           /* global symbol table (PyDictObject) */
7     PyObject *f_locals;            /* local symbol table (any mapping) */
8     PyObject **f_valstack;         /* points after the last local */
9
10    PyObject **f_stacktop;          /* Next free slot in f_valstack. ... */
11    PyObject *f_trace;              /* Trace function */
12    char f_trace_lines;             /* Emit per-line trace events? */
13    char f_trace_opcodes;          /* Emit per-opcode trace events? */
14
15    /* Borrowed reference to a generator, or NULL */
16    PyObject *f_gen;
17
18    int f_lasti;                   /* Last instruction if called */
19    /* ... */
20    int f_lineno;                  /* Current line number */
21    int f_iblock;                  /* index in f_blockstack */
22    char f_executing;              /* whether the frame is still executing */
23    PyTryBlock f_blockstack[CO_MAXBLOCKS]; /* for try and loop blocks */
24    PyObject *f_localsplus[1];    /* locals+stack, dynamically sized */
25 };

```

Pierwsza ramka jest tworzona w celu wykonania obiektu kodu modułu. CPython tworzy nową ramkę za każdym razem, gdy musi wykonać inny obiekt kodu. Każda ramka ma odniesienie

do poprzedniej ramki. W ten sposób ramki tworzą stos ramek, zwany także stosem wywołań, przy czym bieżąca ramka znajduje się na górze. Po wywołaniu funkcji nowa ramka jest umieszczana na stosie. Po powrocie z aktualnie wykonywanej ramki CPython kontynuuje wykonywanie poprzedniej ramki, zapamiętując jej ostatnią przetworzoną instrukcję. W pewnym sensie maszyna wirtualna CPython nie robi nic innego, jak konstruuje i wykonuje ramki.

5.5.4. Thread state

Stan wątku to struktura danych zawierająca dane specyficzne dla wątku, w tym stos wywołań, stan wyjątku i ustawienia debugowania.

Poniżej znajduje się definicja typu struktury thread state.

```
1  typedef struct _ts {
2      struct _ts *prev;
3      struct _ts *next;
4      PyInterpreterState *interp;
5
6      struct _frame *frame;
7      int recursion_depth;
8      char overflowed;
9      char recursion_critical;
10     int tracing;
11     int use_tracing;
12
13     Py_tracefunc c_profilefunc;
14     Py_tracefunc c_tracefunc;
15     PyObject *c_profileobj;
16     PyObject *c_traceobj;
17
18     PyObject *curexc_type;
19     PyObject *curexc_value;
20     PyObject *curexc_traceback;
21
22     PyObject *exc_type;
23     PyObject *exc_value;
24     PyObject *exc_traceback;
25
26     ...
27
28 } PyThreadState;
```

5.5.5. Interpreter state

Stan interpretera to grupa wątków wraz z danymi specyficznymi dla tej grupy. Wątki współdzielą takie rzeczy, jak załadowane moduły (*sys.modules*), wbudowane elementy (*builtins.__dict__*) i system importowania (*importlib*).

Przedstawiam poniżej znajduje definicję typu struktury interpreter state.

```
1  typedef struct _is {
2
3      struct _is *next;
4      struct _ts *tstate_head;
5
6      PyObject *modules;
7      PyObject *modules_by_index;
```

```

8      PyObject *sysdict;
9      PyObject *builtins;
10     PyObject *importlib;
11
12     PyObject *codec_search_path;
13     PyObject *codec_search_cache;
14     PyObject *codec_error_registry;
15     int codecs_initialized;
16     int fscodec_initialized;
17
18     ...
19
20     PyObject *builtins_copy;
21     PyObject *import_func;
22 } PyInterpreterState

```

5.5.6. Runtime state

Stan środowiska wykonawczego jest zmienną globalną. Przechowuje dane specyficzne dla procesu. Obejmuje to stan CPythona (np. czy został zainicjowany czy nie) i mechanizm GIL.

5.5.7. Evaluation loop

Wykonanie kodu bajtowego Pythona przez maszynę wirtualną wydaje się skomplikowane. W rzeczywistości jedyne co musi reobić to iterować po instrukcjach i działać zgodnie z nimi. I to robi `_PyEval_EvalFrameDefault()`. Zawiera nieskończoną pętlę `for (;;)` , którą nazywamy pętlą obliczeń. Wewnątrz tej pętli znajduje się ogromna instrukcja `switch` obejmująca wszystkie możliwe kody operacji. Do wykonania każdego kodu operacji potrzebny jest odpowiedni blok przypadków z kodem. Kod bajtowy jest reprezentowany przez tablicę 16-bitowych liczb całkowitych bez znaku, po jednej liczbie całkowitej na instrukcję. Maszyna wirtualna śledzi następną instrukcję do wykonania za pomocą zmiennej `next_instr`, która jest wskaźnikiem do tablicy instrukcji. Na początku każdej iteracji pętli ewaluacyjnej maszyna wirtualna oblicza następny kod operacji i jego argument. Bierze odpowiednio najmniej znaczący i najbardziej znaczący bajt następnej instrukcji i zwiększa wartość `next_instr`. Funkcja `_PyEval_EvalFrameDefault()` ma prawie 3000 linii, ale jej można przedstawić ją w uproszczonej wersji:

```

1 PyObject*
2 _PyEval_EvalFrameDefault(PyThreadState *tstate, PyFrameObject *f, int
   throwflag)
3 {
4     // ... declarations and initialization of local variables
5     // ... macros definitions
6     // ... call depth handling
7     // ... code for tracing and profiling
8
9     for (;;) {
10         // ... check if the bytecode execution must be suspended,
11         // e.g. other thread requested the GIL
12

```

```

13     // NEXTOPARG() macro
14     _Py_CODEUNIT word = *next_instr; // _Py_CODEUNIT is a typedef for
uint16_t
15     opcode = _Py_OPCODE(word);
16     oparg = _Py_OPARG(word);
17     next_instr++;
18
19     switch (opcode) {
20         case TARGET(NOP) {
21             FAST_DISPATCH(); // more on this later
22         }
23
24         case TARGET(LOAD_FAST) {
25             // ... code for loading local variable
26         }
27
28         // ... 117 more cases for every possible opcode
29     }
30
31     // ... error handling
32 }
33
34 // ... termination
35 }

```



Rozdział 6

Porównanie JIT Python'a i Javy.

6.1. JIT Javy

6.1.1. Co to jest kompilacja Just-In-Time (JIT)?

Kompilacja JIT polega na tłumaczeniu kodu bajtowego Java na natywny kod maszynowy w locie tuż przed wykonaniem. Jest zasadniczym elementem modelu wykonawczego Java. W przeciwieństwie do AOT (Ahead Of Time), który kompiluje kod przed uruchomieniem programu, oferując szybkie wykonanie pętli, ale wolniejsze uruchamianie, JIT tłumaczy kod na bieżąco podczas działania programu, zapewnia szybsze uruchamianie i dostosowuje się do aktywnych ścieżek kodu.

JVM obsługuje aplikacje Java i współpracuje z JIT. Podczas wykonywania JVM komunikuje się z kompilatorem JIT. JIT identyfikuje często używane ścieżki kodu i tłumaczy kod bajtowy na wydajny kod natywny, przyspieszając wykonanie. JVM i JIT razem zapewniają, że aplikacje Java działają płynnie i dobrze.

6.1.2. Jak działa JIT:

Kompilator Just-In-Time (JIT) jest ważnym składnikiem środowiska Java Runtime Environment (JRE). Znacznie zwiększa wydajność i szybkość aplikacji Java w czasie wykonywania. Zrozumienie działania JIT zapewnia wgląd w szybkość i wydajność Java. Przedstawię w punktach, jak dochodzi do jego zastosowania:

- Java Source to Bytecode:

Wszystko zaczyna się od kodu źródłowego Java. Tworzymy ten kod skrupulatnie. Ale zanim zacznie działać, musi zostać przekształcony. Kompilator Java, często nazywany „javac”, kompiluje kod do kodu bajtowego. Ten kod bajtowy jest neutralną dla platformy wersją naszego kodu źródłowego. Można nazwać go uniwersalnym przepisem naszego programu.

- Loading and Interpretation:

Po naciśnięciu przycisku „Uruchom” w aplikacji Java zaczyna pracę wirtualna maszyna Java (JVM). Ładuje skompilowane klasy i metody, gotowe do wykonania. Początkowo JVM interpretuje kod bajtowy. To traktujemy jak czytanie tego uniwersalnego przepisu, punkt po punkcie. Działa, ale czasami może działać nieco wolno, szczególnie w przypadku często używanych części kodu.

- JIT Compilation Activation

Kiedy aplikacja wywołuje metodę, JVM identyfikuje często używane metody i oznacza je jako „gorące” ścieżki kodu. W przypadku tych „gorących” metod rozpoczyna się aktywacja JIT. Jest to dynamiczna kompilacja zachodząca w czasie rzeczywistym podczas działania programu.

- JIT Compilation:

Kompilator JIT, zawsze gotowy w tle, zaczyna działać. Pobiera kod bajtowy „gorących” metod i przekształca go w wydajny natywny kod maszynowy. Ten natywny kod jest dostosowany do konkretnego kontekstu wykonania, jest wyjątkowo zoptymalizowany i bardzo szybki.

- Optimized Execution:

Dzięki nowemu kodowi natywnemu JVM nie musi już interpretować kodu bajtowego. Zamiast tego bezpośrednio uruchamia wysoko zoptymalizowany kod natywny „gorących” metod. To zoptymalizowane wykonanie jest znacznie szybsze dzięki wydajności natywnego kodu maszynowego. Kompilatory JIT mogą wykonywać różne optymalizacje, takie jak upraszczanie wyrażeń, ograniczanie dostępu do pamięci i używanie wydajnych operacji na rejestrach zamiast bardziej złożonych operacji na stosie.

- Performance Enhancement: Chociaż kompilacja JIT potrzebuje trochę czasu i pamięci procesora przy uruchamianiu maszyny i wywoływaniu wielu metod, opłaca się, ponieważ aplikacja działa płynniej i wydajniej. Podczas uruchamiania aplikacji Java kompilator JIT dynamicznie konwertuje „gorące” ścieżki kodu z kodu bajtowego na bardzo wydajny natywny kod maszynowy. Te wspaniałe cechy JIT zapewniają wysoką jakość i niezależność od platformy. Jeszcze raz podkreślę, że dzięki kompilatorze JIT środowisko Java jest tak wyjątkowe, a aplikacje Java zachwycają szybkością i responsywnością.

6.1.3. Optymalizacje stosowane przez kompilatory JIT:

Optymalizacja kodu jest siłą kompilacji JIT. Przedstawię kolejne punkty kompilacji JIT.

- Inlining:

Inlining to proces łączenia mniejszych metod z ich obiektami wywołującymi, zmniejszający obciążenie wywołania metod. Technika ta przyspiesza często wykonywane wywołania metod. Obejmuje optymalizacje, takie jak trivial inlining, call graph inlining, tail recursion elimination i virtual call guard optimizations.

- Local Optimizations:

Lokalne optymalizacje ulepszają małe fragmenty kodu na raz, wykorzystując techniki takie jak local data flow analyses, register usage optimization i simplifications of Java idioms.

- Control Flow Optimizations:

Control flow optimizations analizują i reorganizują ścieżki kodu poprawiając wydajność. Obejmują one code reordering, loop-related enhancements (reduction, inversion, unrolling), i doskonalsze exception handling.

- **Global Optimizations:**

Globalne optymalizacje działają na całą metodę, wymagają więcej czasu kompilacji, ale oferują znaczny wzrost wydajności. Obejmują one global data flow analyses, partial redundancy elimination, escape analysis i optymalizacje związane z garbage collection i memory allocation.

- **Native Code Generation:**

W końcowej fazie drzewa są tłumaczone na instrukcje kodu maszynowego i stosowane są pewne optymalizacje specyficzne dla platformy. Skompilowany kod jest przechowywany w pamięci podręcznej kodu JVM, gotowy do użycia w przyszłości, zapewniając szybkość wykonania.

Można powiedzieć, że kompilacja JIT jest jak magiczna transformacja, która zamienia zwykły kod Java w superwydajny, błyskawiczny kod maszynowy.

6.1.4. Zalety kompilatora JIT:

- **Efficient Memory Usage:**

Kompilatory JIT zużywają mniej pamięci, ponieważ generują natywny kod maszynowy tylko dla wykonywanych ścieżek kodu, zamiast kompilować cały program od razu.

- **Dynamic Code Optimization:**

Kompilatory JIT optymalizują kod w czasie wykonywania, umożliwiając im dostosowanie się do kontekstu wykonania, co prowadzi do poprawy wydajności.

- **Multiple Optimization Levels:**

Kompilatory JIT wykorzystują różne poziomy optymalizacji, dopasowując optymalizacje do konkretnego wykonywanego kodu, co może znacznie zwiększyć szybkość wykonywania.

- **Reduced Page Faults:**

Generując w razie potrzeby natywny kod maszynowy, kompilatory JIT zmniejszają prawdopodobieństwo błędów stron i operacji we/wy dysku, zwiększając czas reakcji aplikacji.

6.1.5. Wady kompilatora JIT:

- **Increased Program Complexity:**

Kompilacja JIT powoduje że proces wykonywania programu jest bardziej złożony, co może utrudniać debugowanie i profilowanie.

- **Limited Benefit for Short Code:**

Kompilacja JIT może nie zapewniać znaczących korzyści w przypadku małych fragmentów kodu lub programów o minimalnym czasie wykonania, ponieważ narzut związany z kompilacją może przeważać nad korzyściami.

- **Cache Memory Usage:**

Kompilatory JIT zużywają znaczną ilość pamięci podręcznej, potencjalnie wpływając na ogólną wydajność systemu, szczególnie w środowiskach o ograniczonych zasobach..

Wnioski:

Zalety JIT obejmują efektywne wykorzystanie pamięci, dynamiczną optymalizację kodu, wiele poziomów optymalizacji i zmniejszoną liczbę błędów stron, a wszystko to przyczynia się do zwiększonej wydajności. Jednak JIT wprowadza pewną złożoność i może nie zapewniać znacznych korzyści w przypadku bardzo krótkich fragmentów kodu. Warto dokonać takiego kompromisu ze względu na szybkość i możliwości adaptacji, jakie wnosi do ekosystemu Java.

6.2. JIT Pythona

CPython od wersji 3.13 obsługuje JIT. Wykorzystuje on technikę copy-and-patch. Copy-and-patch to prosta technika kompilatora przeznaczona do kompilacji just-in-time (kompilacja JIT), która wykorzystuje pattern matching w celu dopasowania wstępnie wygenerowanych szablonów do części abstrakcyjnego drzewa składni (AST) lub strumienia kodu bajtowego. Emituje odpowiednio wcześniej napisane fragmenty kodu maszynowego, poprawia je i wstawia adresy pamięci, adresy rejestrów, stałych i innych parametrów potrzebnych do wytworzenia kodu wykonywalnego. Kod niepasujący do szablonów można albo zinterpretować w normalny sposób, albo utworzyć kod w celu bezpośredniego wywołania kodu interpretera.

Benchmarki wskazują na 2-9 procentowy przyrost szybkości w stosunku do Pythona bez JIT.



Rozdział 7

Zarządzanie pamięcią

Tylko odpowiednie zarządzanie pamięcią maszyn wirtualnych doprowadzi do efektywnego wykorzystania zasobów. Postaram się przedstawić mechanizmy alokacji pamięci, tworzenia obiektów oraz mechanizmy usuwania śmieci stosowane przez wirtualną maszynę Java i wirtualną maszynę Python.

7.1. Memory Allocation and Object Creation:

7.1.1. JVM

JVM wykorzystuje trzy główne obszary do zarządzania pamięcią: obszar metod, stos i stertę. Kiedy obiekty są tworzone przy użyciu słowa kluczowego `new`, pamięć jest alokowana na stercie, a konstruktor służy do inicjowania obiektu. Sterta jest podzielona na trzy pokolenia:

- **Young Generation:** To tu powstają nowe obiekty
- **Old Generation:** Obiekty, które przetrwały kilka rund zbiórki śmieci w ramach Young Generation, ostatecznie awansują do Old Generation, wykorzystywanego do obiektów długowiecznych
- **Metaspace:** Jest to natywny obszar pamięci, który dynamicznie dostosowuje swój rozmiar

7.1.2. PVM

Python wykorzystuje prywatną stertę do zarządzania pamięcią. Globalna blokada interpretera (GIL) zapewnia, że tylko jeden wątek może jednocześnie wykonywać kod bajtowy w czasie zarządzania pamięcią. Python jest typowany dynamicznie, a obiekty tworzone są przy użyciu prostej składni z automatycznie zarządzaną pamięcią.

Python obsługuje różne typy obiektów, takie jak `int`, `dict` sklasyfikowane jako obiekty wysokiego poziomu. Podczas schodzenia o jeden poziom obiekty te są definiowane jako typy obiektów Pythona. Alokator obiektów Pythona wywołuje interfejs API alokatora pamięci `raw` w celu przydzielenia pamięci dla obiektów w systemie operacyjnym. Ten interfejs API wykorzystuje alokatory ogólnego przeznaczenia, wykorzystując funkcję `malloc` języka C do przypisywania pamięci dla zmiennych

7.2. Garbage Collection:

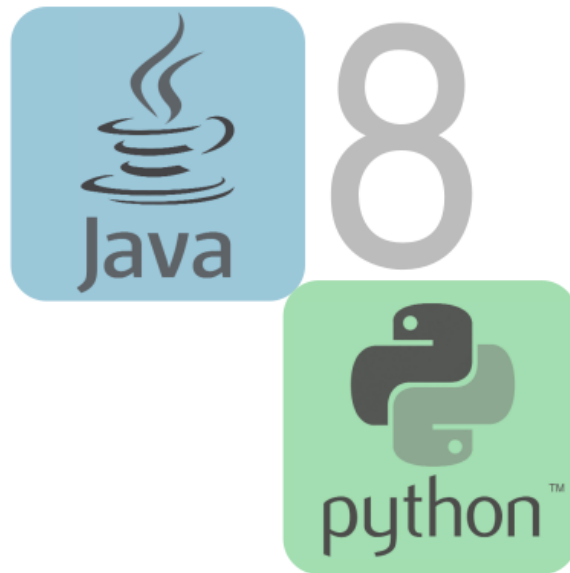
7.2.1. JVM

Garbage Collection w maszynie JVM działa tylko na obszarze sterty. Pamięć przydzielona na stosie jest automatycznie usuwana po zdjęciu ramki. Istnieją różne algorytmy dla garbage collection.

- **Garbage-First (G1) Collector:** Działa na Young Generation, identyfikując i kolekcjonując obiekty krótkotrwałe. Ocalałe obiekty podlegają dalszej obróbce i mogą zostać przeniesione w inne miejsce w ramach Young Generation.
- **Mark and Sweep (Old Generation):** Identyfikuje i oznacza osiągalne obiekty w Old Generation. Nieoznaczone obiekty są dealokowane, zwalniając pamięć.
- Nie ma konkretnych algorytmów dla Metaspace

7.2.2. PVM

PVM wykorzystuje przede wszystkim mechanizm zliczania referencji do zarządzania pamięcią. Obiekty mają liczbę odwołań, a pamięć jest zwalniana, gdy liczba spada do zera. Wbudowany garbage collector służy do obsługi scenariuszy takich jak cykliczne odniesienia, identyfikowanie i zbieranie obiektów, które nie są już osiągalne.



Rozdział 8

Środowisko uruchomieniowe

Do wykonania pomiarów wydajnościowych zastosowaliśmy komputer o następującej konfiguracji:

- Procesor AMD Ryzen 5 3600
- 16 GB pamięci RAM
- karta graficzna GeForce GTX 1660
- Płyta główna PRIME B450M-A II

Wykorzystaliśmy następujące oprogramowanie:

- System operacyjny Debian 11
- Java w wersji "23.0.1" z 2024-10-15
- Python 3.13.0 z 2024-10-07



Rozdział 9

Porównanie czasu działania prostych fragmentów kodu

W celu porównania czasu działania prostych operacji takich jak dodawanie czy mnożenie stosujemy narzędzie `perf_counter` w przypadku Pythona i polecenie `System.nanoTime()` w pętli `for` w przypadku Javy.

9.1. Pusta pętla

W przypadku Javy do pomiaru czasu wykonania pustej pętli stosuję następujący kod [\[Java\]](#):

```
1 import java.io.*;
2
3 public class Empty {
4     // main function
5     public static void main(String[] args)
6     {
7         // Start measuring execution time
8         long startTime = System.nanoTime();
9
10        count_function(100000000);
11
12        // Stop measuring execution time
13        long endTime = System.nanoTime();
14
15        // Calculate the execution time in milliseconds
16        long executionTime
17            = (endTime - startTime);
18
19        double seconds = (double)executionTime / 1_000_000_000.0;
20        System.out.println(seconds + " s");
21    }
22
23    public static void count_function(long x)
24    {
25        for (long i = 0; i < x; i++)
26            ;
27    }
28 }
29 }
```

Wykonanie powyższego kodu zajmuje około 0,0416 s. W przypadku wyłączenia JIT (`java -Xint Empty`) czas wykonania wzrasta do około 0,618 s.

Do pomiaru czasu wykonania pustej pętli stosuję następujący kod w przypadku Pythona [Python]:

```
1 from time import perf_counter_ns
2
3 def my_function():
4     for i in range(100000000):
5         pass
6
7 # Start the stopwatch / counter
8 t1_start = perf_counter_ns()
9
10 my_function()
11
12 # Stop the stopwatch / counter
13 t1_stop = perf_counter_ns()
14
15 print((t1_stop-t1_start) / 1000000000, 's')
```

Wykonanie powyższego kodu zajmuje około 1,62 s.

Jest to więc $1,62/0,0416 = 3,89$ raza wolniej niż w przypadku Javy. Po wyłączeniu JIT, Java jest już jedynie $1,62/0,618=2,62$ razy szybsza niż Python.

Warto w tym miejscu zaznaczyć, że gdyby kod funkcji `my_function` znajdował się poza nią, tak jak na poniższym listingu [Python]:

```
1 from time import perf_counter_ns
2
3 # Start the stopwatch / counter
4 t1_start = perf_counter_ns()
5
6 for i in range(100000000):
7     pass
8
9 # Stop the stopwatch / counter
10 t1_stop = perf_counter_ns()
11
12 print((t1_stop-t1_start) / 1000000000, 's')
```

to program działałby zdecydowanie wolniej. W tym przypadku byłoby to 3,68 s zamiast 1,62 s.

Generalnie przyjmuje się, że szybsze działanie kodu w funkcjach niż poza nimi spowodowane jest przez różnicę prędkości między dostępem do zmiennych globalnych i lokalnych, optymalizację kodu bajtowego, wydajność pamięci podręcznej instrukcji i zarządzanie przestrzenią nazw.

Szczególnie istotny wydaje się czas dostępu do zmiennych. Widać to dobrze na poziomie kodu bajtowego. Na poniższym listingu widzimy kod, który korzysta z bajtkodu `STORE_FAST` do zachowania aktualnej wartości licznika pętli w zmiennej lokalnej.

| | | | |
|---|-----|---------------|-----------|
| 1 | L1: | FOR_ITER | 3 (to L2) |
| 2 | | STORE_FAST | 0 (i) |
| 3 | | | |
| 4 | 6 | JUMP_BACKWARD | 5 (to L1) |

Natomiast na poniższym listingu widzimy kod, który korzysta z bajtkodu `STORE_NAME` do zachowania aktualnej wartości licznika pętli w zmiennej globalnej.

| | | | |
|---|-----|---------------|-----------|
| 1 | L1: | FOR_ITER | 3 (to L2) |
| 2 | | STORE_NAME | 4 (i) |
| 3 | | | |
| 4 | 7 | JUMP_BACKWARD | 5 (to L1) |

Jak wynika z pomiarów (3,68 s vs 1,62 s), wykorzystanie zmiennych globalnych w pętli znacznie spowalnia program.

9.2. Dodawanie w pętli jednej liczby

W przypadku Javy do pomiaru czasu wykonania dodawania jednej liczby w pętli stosuję następujący kod [\[Java\]](#):

```

1 import java.io.*;
2
3 public class Addition {
4     // main function
5     public static void main(String[] args)
6     {
7         // Start measuring execution time
8         long startTime = System.nanoTime();
9
10        count_function(100000000);
11
12        // Stop measuring execution time
13        long endTime = System.nanoTime();
14
15        // Calculate the execution time in milliseconds
16        long executionTime
17            = (endTime - startTime);
18
19        double seconds = (double)executionTime / 1_000_000_000.0;
20        System.out.println(seconds + " s");
21    }
22
23    public static void count_function(long x)
24    {
25        long j = 0;
26        for (long i = 0; i < x; i++)
27            j = j + 1;
28    }
29 }

```

Wykonanie powyższego kodu zajmuje około 0,0339 s. W przypadku wyłączenia JIT (java -Xint Addition) czas wykonania wzrasta do około 0,918 s.

Do pomiaru czasu wykonania dodawania jednej liczby w pętli stosuję następujący kod w przypadku Pythona [\[Python\]](#):

```

1 from time import perf_counter_ns
2
3 def my_function():
4     sum = 0
5     for i in range(100000000):
6         sum = sum + 1
7
8 # Start the stopwatch / counter
9 t1_start = perf_counter_ns()
10

```

```

11
12 my_function()
13
14 # Stop the stopwatch / counter
15 t1_stop = perf_counter_ns()
16
17 print((t1_stop-t1_start) / 1000000000, 's')

```

Wykonanie powyższego kodu zajmuje około 3,27 s.

Jest to więc $3,27/0,0339=96,5$ raza wolniej niż w przypadku Javy. Po wyłączeniu JIT, Java jest już jedynie $10,9/0,918=11,9$ szybsza niż Python.

9.3. Duże liczby

W przypadku korzystania z bardzo dużych liczb Python okazuje się być szybszy od Javy. Weźmy pod uwagę poniższy kod w Javie [\[Java\]](#):

```

1 import java.io.*;
2 import java.math.BigInteger;
3
4 public class Addition4 {
5     // main function
6     public static void main(String[] args)
7     {
8         // Start measuring execution time
9         long startTime = System.nanoTime();
10
11         count_function(1000000000);
12
13         // Stop measuring execution time
14         long endTime = System.nanoTime();
15
16         // Calculate the execution time in milliseconds
17         long executionTime
18             = (endTime - startTime);
19
20         double seconds = (double)executionTime / 1_000_000_000.0;
21         System.out.println(seconds + " s");
22     }
23
24     public static void count_function(long x)
25     {
26         BigInteger b1, b2;
27
28         b1 = new BigInteger("100");
29         int exponent = 100;
30
31         BigInteger result = b1.pow(exponent);
32         BigInteger result2 = result.add(BigInteger.valueOf(x));
33
34         b2 = new BigInteger("2000");
35         int exponent2 = 2000;
36
37         BigInteger result3 = b2.pow(exponent2);
38
39         BigInteger sum = BigInteger.ZERO;
40

```

```

41     for (BigInteger bi = result; bi.compareTo(result2) < 0; bi = bi.add(
42         BigInteger.ONE))
43         sum = sum.add(result3);
44 }

```

Czas jego wykonania wynosi 83 s.

Odpowiednik tego kodu w Pythonie wykonuje się w czasie 59,4 s [\[Python\]](#):

```

1 from time import perf_counter_ns
2 import sys
3
4 def my_function():
5     sum = 0
6     r = 2000**2000
7     for i in range(100**100, 100**100 + 1000000000):
8         sum = sum + r
9
10 # Start the stopwatch / counter
11 t1_start = perf_counter_ns()
12
13 my_function()
14
15 # Stop the stopwatch / counter
16 t1_stop = perf_counter_ns()
17
18 print((t1_stop-t1_start) / 1000000000, 's')

```

Widać, że implementacja dużych liczb w Pythonie jest lepsza od implementacji BigInteger w Javie.

Dobrze widać różnicę w czasie działania dodawania między int, long a BigInteger analizując kod bajtowy dla odpowiednio long:

```

1 public static void count_function(long);
2 Code:
3     0: lconst_0
4     1: lstore_2
5     2: lload_2
6     3: lload_0
7     4: lcmp
8     5: ifge             15
9     8: lload_2
10    9: lconst_1
11   10: ladd
12   11: lstore_2
13   12: goto             2
14   15: return

```

i dla BigInteger:

```

1 public static void count_function(long);
2 Code:
3     0: getstatic          #35             // Field java/math/BigInteger.
4     ZERO:Ljava/math/BigInteger;
5     3: astore_2
6     4: aload_2
7     5: lload_0
8     6: invokestatic        #41             // Method java/math/BigInteger.
9     valueOf:(J)Ljava/math/BigInteger;

```

```

8      9: invokevirtual #45                // Method java/math/BigInteger.
compareTo:(Ljava/math/BigInteger;)I
9      12: ifge                26
10     15: aload_2
11     16: getstatic    #49                // Field java/math/BigInteger.ONE
:Ljava/math/BigInteger;
12     19: invokevirtual #52                // Method java/math/BigInteger.
add:(Ljava/math/BigInteger;)Ljava/math/BigInteger;
13     22: astore_2
14     23: goto         4
15     26: return

```

Ten drugi jest jak widać bardziej skomplikowany.

9.4. Kolekcje

W następnych podpunktach mierzymy czas działania kolekcji.

9.4.1. Java Array vs Python List, NumPy Array

Weźmy pod uwagę kod dodawania elementów dwóch tablic [\[Java\]](#).

```

1 import java.util.Arrays;
2
3 public class Concat2 {
4
5     public static void main(String[] args) {
6         int[] array1 = new int[10000000]; //000000
7         int[] array2 = new int[10000000];
8
9         for (int i=0; i < 10000000;i++) {
10             array1[i] = i;
11             array2[i] = i + 10000000;
12         }
13
14         int aLen = array1.length;
15         int bLen = array2.length;
16         int[] result = new int[aLen];
17
18         // Start measuring execution time
19         long startTime = System.nanoTime();
20
21         for (int i = 0; i < aLen; i++) {
22             result[i] = array1[i] + array2[i];
23         }
24
25         // Stop measuring execution time
26         long endTime = System.nanoTime();
27
28         // Calculate the execution time in milliseconds
29         long executionTime
30             = (endTime - startTime);
31
32         double seconds = (double)executionTime / 1_000_000_000.0;
33         System.out.println(seconds + " s");
34     }
35 }

```

Czas działania w Javie wynosi 0,0132 s.

W przypadku Pythona skorzystamy z biblioteki NumPy. Gdy korzystamy z danych liczbowych, biblioteka ta pozwala na znaczne przyspieszenie działań. Dla porządku wywołamy również łączenie elementów dwóch list, aby wskazać, że jest nieefektywne [\[Python\]](#).

```
1 import numpy as np
2 from time import perf_counter_ns
3
4 # Define two Python lists
5 py_list1 = list(range(10000000))
6 py_list2 = list(range(10000000, 20000000))
7
8 # Define two Numpy arrays
9 np_array1 = np.arange(10000000)
10 np_array2 = np.arange(10000000, 20000000)
11
12 # Adding Python lists
13 t1_start = perf_counter_ns()
14 result_list = [a + b for a, b in zip(py_list1, py_list2)]
15 t1_stop = perf_counter_ns()
16
17 print((t1_stop-t1_start) / 1000000000, 's')
18
19 # Adding Numpy arrays
20 t2_start = perf_counter_ns()
21 result_array = np_array1 + np_array2
22 t2_stop = perf_counter_ns()
23
24 print((t2_stop-t2_start) / 1000000000, 's')
```

Czas działania wynosi 0,416 s dla łączenia list i 0,0178 s dla łącznia dwóch tablic z biblioteki NumPy. Dzięki bibliotece NumPy kod w Pythonie wykonuje się w zbliżonym czasie do kodu w Javie.

9.4.2. Dodawanie String'ów do kolekcji

Dla poniższego programu [\[Java\]](#) mierzącego czas dodawania String'ów do HashSet, TreeSet, HashMap i TreeMap:

```
1 package com.avenuecode.snippet;
2
3 import org.openjdk.jmh.annotations.*;
4
5 import java.util.concurrent.TimeUnit;
6
7 import java.io.*;
8
9 import java.util.*;
10
11 @State(Scope.Thread)
12 public class MyBenchmark {
13
14     int x = 10000000;
15     HashSet<String> counts = new HashSet<String>();
16     TreeSet<String> counts2 = new TreeSet<String>();
17     HashMap<String, String> map = new HashMap<>();
18     TreeMap<String, String> map2 = new TreeMap<>();
19     ArrayList<String> mylist = new ArrayList<String>();
```

```

20 String[] s = new String[x];
21 int y = -1;
22
23 /**
24  * Setup method to initialize data for the benchmark.
25  */
26 @Setup(Level.Trial)
27 public void setup() {
28     for (int i=0; i < x; i++)
29         mylist.add(String.valueOf(i));
30
31     Collections.shuffle(mylist);
32
33     for (int i=0; i < x; i++)
34         s[i] = mylist.get(i);
35 }
36
37 @Fork(value = 1, warmups = 1)
38 @Warmup(iterations = 1)
39 @Benchmark
40 @BenchmarkMode(Mode.AverageTime)
41 public long _0_HashSetAdd() {
42     for (int i=0; i < x; i++)
43         counts.add(s[i]);
44     y = 0;
45     return 0;
46 }
47
48 @Fork(value = 1, warmups = 1)
49 @Warmup(iterations = 1)
50 @Benchmark
51 @BenchmarkMode(Mode.AverageTime)
52 public long _1_TreeSetAdd() {
53     for (int i=0; i < x; i++)
54         counts2.add(s[i]);
55     y = 1;
56     return 0;
57 }
58
59 @Fork(value = 1, warmups = 1)
60 @Warmup(iterations = 1)
61 @Benchmark
62 @BenchmarkMode(Mode.AverageTime)
63 public long _2_HashMapPut() {
64     for (int i=0; i < x; i++)
65         map.put(s[i], s[i]);
66     y = 2;
67     return 0;
68 }
69
70 @Fork(value = 1, warmups = 1)
71 @Warmup(iterations = 1)
72 @Benchmark
73 @BenchmarkMode(Mode.AverageTime)
74 public long _3_TreeMapPut() {
75     for (int i=0; i < x; i++)
76         map2.put(s[i], s[i]);
77     y = 3;
78     return 0;
79 }

```

```

80     @TearDown(Level.Invocation)
81     public void doTearDown() {
82         switch (y) {
83             case 0:
84                 counts.clear();
85                 break;
86             case 1:
87                 counts2.clear();
88                 break;
89             case 2:
90                 map.clear();
91                 break;
92             case 3:
93                 map2.clear();
94                 break;
95         }
96     }
97
98     /**
99      * Main method to run the benchmark.
100     *
101     * @param args Command line arguments
102     * @throws Exception If an error occurs during benchmark execution
103     */
104     public static void main(String[] args) throws Exception {
105         org.openjdk.jmh.Main.main(args);
106     }
107
108 }

```

uzyskujemy następujące czasy dodawania:

HashSet:

```

1 # Benchmark: com.avenuecode.snippet.MyBenchmark._0_HashSetAdd
2
3 # Run progress: 0,00% complete, ETA 00:08:00
4 # Warmup Fork: 1 of 1
5 # Warmup Iteration 1: 1,333 s/op
6 Iteration 1: 0,952 s/op
7 Iteration 2: 0,906 s/op
8 Iteration 3: 1,027 s/op
9 Iteration 4: 1,045 s/op
10 Iteration 5: 1,048 s/op
11
12 # Run progress: 12,50% complete, ETA 00:07:29
13 # Fork: 1 of 1
14 # Warmup Iteration 1: 1,496 s/op
15 Iteration 1: 1,084 s/op
16 Iteration 2: 1,049 s/op
17 Iteration 3: 1,011 s/op
18 Iteration 4: 0,949 s/op
19 Iteration 5: 0,942 s/op

```

TreeSet:

```

1 # Benchmark: com.avenuecode.snippet.MyBenchmark._1_TreeSetAdd
2
3 # Run progress: 25,00% complete, ETA 00:06:25
4 # Warmup Fork: 1 of 1
5 # Warmup Iteration 1: 13,906 s/op

```

```

6 Iteration 1: 12,645 s/op
7 Iteration 2: 12,154 s/op
8 Iteration 3: 11,993 s/op
9 Iteration 4: 12,054 s/op
10 Iteration 5: 11,838 s/op
11
12 # Run progress: 37,50% complete, ETA 00:05:40
13 # Fork: 1 of 1
14 # Warmup Iteration 1: 15,035 s/op
15 Iteration 1: 12,804 s/op
16 Iteration 2: 12,588 s/op
17 Iteration 3: 12,710 s/op
18 Iteration 4: 12,403 s/op
19 Iteration 5: 12,679 s/op

```

HashMap:

```

1 # Benchmark: com.avenuecode.snippet.MyBenchmark._2_HashMapPut
2
3 # Run progress: 50,00% complete, ETA 00:04:43
4 # Warmup Fork: 1 of 1
5 # Warmup Iteration 1: 1,438 s/op
6 Iteration 1: 0,958 s/op
7 Iteration 2: 0,889 s/op
8 Iteration 3: 0,914 s/op
9 Iteration 4: 0,947 s/op
10 Iteration 5: 0,988 s/op
11
12 # Run progress: 62,50% complete, ETA 00:03:28
13 # Fork: 1 of 1
14 # Warmup Iteration 1: 1,312 s/op
15 Iteration 1: 0,927 s/op
16 Iteration 2: 0,945 s/op
17 Iteration 3: 0,916 s/op
18 Iteration 4: 0,931 s/op
19 Iteration 5: 0,897 s/op

```

TreeMap:

```

1 # Benchmark: com.avenuecode.snippet.MyBenchmark._3_TreeMapPut
2
3 # Run progress: 75,00% complete, ETA 00:02:17
4 # Warmup Fork: 1 of 1
5 # Warmup Iteration 1: 15,210 s/op
6 Iteration 1: 13,199 s/op
7 Iteration 2: 13,195 s/op
8 Iteration 3: 12,313 s/op
9 Iteration 4: 12,285 s/op
10 Iteration 5: 12,118 s/op
11
12 # Run progress: 87,50% complete, ETA 00:01:10
13 # Fork: 1 of 1
14 # Warmup Iteration 1: 15,400 s/op
15 Iteration 1: 12,677 s/op
16 Iteration 2: 12,756 s/op
17 Iteration 3: 12,658 s/op
18 Iteration 4: 12,662 s/op
19 Iteration 5: 12,568 s/op

```

W przypadku Pythona dla programu [\[Python\]](#):

```

1 from time import perf_counter

```



```

2 import random
3 import sys
4 from sortedcontainers import SortedSet
5 from sortedcontainers import SortedDict
6
7 def my_function():
8     s = list(range(0,10000000))
9
10    for i in range(0, 10000000):
11        s[i] = str(s[i])
12
13    random.shuffle(s)
14
15    counts = set()
16    sorted_set = SortedSet()
17    dict1 = {}
18    dict2 = SortedDict()
19
20    for j in range(0,5):
21
22        counts.clear()
23
24        # Start the stopwatch / counter
25        t1_start = perf_counter()
26
27        for i in range(0, 10000000):
28            counts.add(s[i])
29
30        # Stop the stopwatch / counter
31        t1_stop = perf_counter()
32
33        print("set: ", t1_stop-t1_start)
34
35
36    for j in range(0,5):
37
38        sorted_set.clear()
39
40        # Start the stopwatch / counter
41        t1_start = perf_counter()
42
43        for i in range(0, 10000000):
44            sorted_set.add(s[i])
45
46        # Stop the stopwatch / counter
47        t1_stop = perf_counter()
48
49        print("SortedSet: ", t1_stop-t1_start)
50
51    for j in range(0,5):
52
53        dict1.clear()
54
55        # Start the stopwatch / counter
56        t1_start = perf_counter()
57
58        for i in range(0, 10000000):
59            dict1[s[i]] = s[i]
60
61        # Stop the stopwatch / counter

```

```

62     t1_stop = perf_counter()
63
64     print("Dictionary: ", t1_stop-t1_start)
65
66     for j in range(0,5):
67
68         dict2.clear()
69
70         # Start the stopwatch / counter
71         t1_start = perf_counter()
72
73         for i in range(0, 10000000):
74             dict2[s[i]] = s[i]
75
76         # Stop the stopwatch / counter
77         t1_stop = perf_counter()
78
79         print("SortedDict: ", t1_stop-t1_start)
80
81 my_function()

```

uzyskujemy następujące czasy wstawiania do kolekcji:

```

1 set:      2.207043123999938
2 set:      2.200626075999935
3 set:      2.1995758820003175
4 set:      2.197942086000239
5 SortedSet: 24.410430537999673
6 SortedSet: 25.729037049
7 SortedSet: 24.10598873800018
8 SortedSet: 24.190924873000313
9 SortedSet: 24.490848085999914
10 Dictionary: 4.109121431999938
11 Dictionary: 4.080353903999821
12 Dictionary: 4.059659220999947
13 Dictionary: 4.15668830300001
14 Dictionary: 4.085921346999839
15 SortedDict: 28.95898643199962
16 SortedDict: 28.782305533
17 SortedDict: 28.483486593999714
18 SortedDict: 28.313258532999953
19 SortedDict: 28.271739520999745

```

Przjrzyjmy się jeszcze raz wynikom dla Javy:

| Benchmark | Mode | Cnt | Score | Error | Units |
|---------------------------|------|-----|--------|-----------|-------|
| MyBenchmark._0_HashSetAdd | avgt | 5 | 1,007 | +/- 0,238 | s/op |
| MyBenchmark._1_TreeSetAdd | avgt | 5 | 12,637 | +/- 0,584 | s/op |
| MyBenchmark._2_HashMapPut | avgt | 5 | 0,923 | +/- 0,070 | s/op |
| MyBenchmark._3_TreeMapPut | avgt | 5 | 12,664 | +/- 0,257 | s/op |

Powyższe czasy dodawania elementów do kolekcji zgodne są ze złożonością czasową tych operacji.

| collection | operation | complexity |
|------------|-----------|-----------------------|
| HashSet | add | $O(1)$ |
| TreeSet | add | $O(\log N)$ |
| HashMap | put | $O(1)$ |
| TreeMap | put | $O(\log N)$ |
| set | add | $O(1)$ (average case) |
| SortedSet | add | $O(\log N)$ |

```

9 Dictionary    setitem    O(1)
10 SortedDict   setitem    O(log N)

```

9.4.3. Sprawdzanie czy kolekcja zawiera Stringi

Korzystając z programu podobnego do używanego do szacowania czasu dodawania String'ów do kolekcji [\[Java\]](#) [\[Python\]](#), mierzę czas sprawdzania czy String należy do kolekcji.

Wyniki dla Javy:

| Benchmark | Mode | Cnt | Score | Error | Units |
|-----------------------------------|------|-----|--------|-----------|-------|
| MyBenchmark._0_HashSetContains | avgt | 5 | 0,594 | +/- 0,024 | s/op |
| MyBenchmark._1_TreeSetContains | avgt | 5 | 11,100 | +/- 0,154 | s/op |
| MyBenchmark._2_HashMapContainsKey | avgt | 5 | 0,536 | +/- 0,017 | s/op |
| MyBenchmark._3_TreeMapContainsKey | avgt | 5 | 11,472 | +/- 0,297 | s/op |

Wyniki dla Pythona:

```

1 set: 2.9993871530000433
2 set: 2.9143904570000814
3 set: 2.880181250000078
4 set: 2.879620507999789
5 set: 2.8789269859998967
6 SortedSet: 3.3288162799999554
7 SortedSet: 3.3365371639997647
8 SortedSet: 3.3687667489998603
9 SortedSet: 3.368098517999897
10 SortedSet: 3.353463620000184
11 Dictionary: 4.0150994229998105
12 Dictionary: 4.100978073999613
13 Dictionary: 4.031463173999782
14 Dictionary: 4.091581018000397
15 Dictionary: 4.087747910999951
16 SortedDict: 4.311155410000083
17 SortedDict: 4.318198027000108
18 SortedDict: 4.194814053999835
19 SortedDict: 4.197658074999708
20 SortedDict: 4.210030474999712

```

Powyższe czasy sprawdzania czy elementy należą do kolekcji zgodne są ze złożonością czasową tych operacji.

| collection | operation | complexity |
|------------|-------------|---------------------|
| HashSet | contains | O(1) |
| TreeSet | contains | O(log N) |
| HashMap | containsKey | O(1) |
| TreeMap | containsKey | O(log N) |
| set | in | O(1) (average case) |
| SortedSet | in | O(1) |
| Dictionary | in | O(1) |
| SortedDict | in | O(1) |

9.4.4. Usuwanie String'ów z kolekcji

Korzystając z programu podobnego do używanego do szacowania czasu dodawania String'ów do kolekcji [\[Java\]](#) [\[Python\]](#), mierzę czas usuwania String'ów z kolekcji.

Wyniki dla Javy:

| Benchmark | Mode | Cnt | Score | Error | Units |
|------------------------------|------|-----|--------|-----------|-------|
| MyBenchmark._0_HashSetRemove | avgt | 5 | 1,288 | +/- 0,153 | s/op |
| MyBenchmark._1_TreeSetRemove | avgt | 5 | 11,823 | +/- 2,925 | s/op |
| MyBenchmark._2_HashMapRemove | avgt | 5 | 1,244 | +/- 0,052 | s/op |
| MyBenchmark._3_TreeMapRemove | avgt | 5 | 11,110 | +/- 1,195 | s/op |

Wyniki dla Pythona:

```

1 set: 2.962287095999727
2 set: 2.93804236699998
3 set: 2.975594524999906
4 set: 2.9539603660000466
5 set: 2.9535686730000634
6 SortedSet: 14.213103547999708
7 SortedSet: 14.19680624900002
8 SortedSet: 14.200566351999896
9 SortedSet: 14.25538980600004
10 SortedSet: 14.142081440999846
11 Dictionary: 4.246978401999968
12 Dictionary: 4.244396434999999
13 Dictionary: 4.245342547999826
14 Dictionary: 4.243117217999952
15 Dictionary: 4.247132589999637
16 SortedDict: 16.82463326800007
17 SortedDict: 16.839339327999824
18 SortedDict: 16.89218398200046
19 SortedDict: 17.11353439100003
20 SortedDict: 16.469016367999757

```

Powyższe czasy usuwania elementów z kolekcji zgodne są ze złożonością czasową tych operacji.

| collection | operation | complexity |
|------------|-----------|-----------------------|
| HashSet | remove | $O(1)$ |
| TreeSet | remove | $O(\log N)$ |
| HashMap | remove | $O(1)$ |
| TreeMap | remove | $O(\log N)$ |
| | | |
| set | discard | $O(1)$ (average case) |
| SortedSet | discard | $O(\log N)$ |
| Dictionary | pop | $O(1)$ |
| SortedDict | pop | $O(\log N)$ |



Rozdział 10

Porównanie czasu działania programów

10.1. Narzędzia

- `runexec` - program dostarczany wraz z frameworkiem `BenchExec`, który pozwala na mierzenie zużycia zasobów, podobnie jak polecenie `time`, ale z dokładniejszym pomiarem czasu działania i z pomiarem zużycia pamięci

10.2. The Computer Language Benchmarks Game

W tej sekcji porównuję czas działania wybranych programów w Javie i Pythonie pochodzących ze strony [\[CLBG\]](#).

10.2.1. PIDIGITS

Program `PIDIGITS` oblicza `N` pierwszych cyfr liczby `Pi`. Następnie wypisuje 10 kolejnych cyfr liczby `Pi` w każdej linii. Dla pomiarów za bazowe `N` przyjąłem liczbę 10000. Dokładniejszy opis: [\[8\]](#)

Java

W przypadku Javy walltime programu [\[Java\]](#) przy wykorzystaniu polecenia `runexec` wyniósł 5,93 s. Komplet wyników:

```
starttime=2024-12-20T18:01:42.080860+01:00
returnvalue=0
walltime=5.927161733000048s
cputime=6.890361028s
cputime-cpu0=0.079110021s
cputime-cpu1=0.032669473s
cputime-cpu10=0.063455014s
cputime-cpu11=5.658510060s
cputime-cpu2=0.260540423s
cputime-cpu3=0.042436937s
cputime-cpu4=0.050863610s
cputime-cpu5=0.151805614s
cputime-cpu6=0.057804142s
cputime-cpu7=0.403615176s
cputime-cpu8=0.037124300s
```

```
cputime-cpu9=0.052426258s  
memory=383864832B  
blkio-read=0B  
blkio-write=0B
```

Python

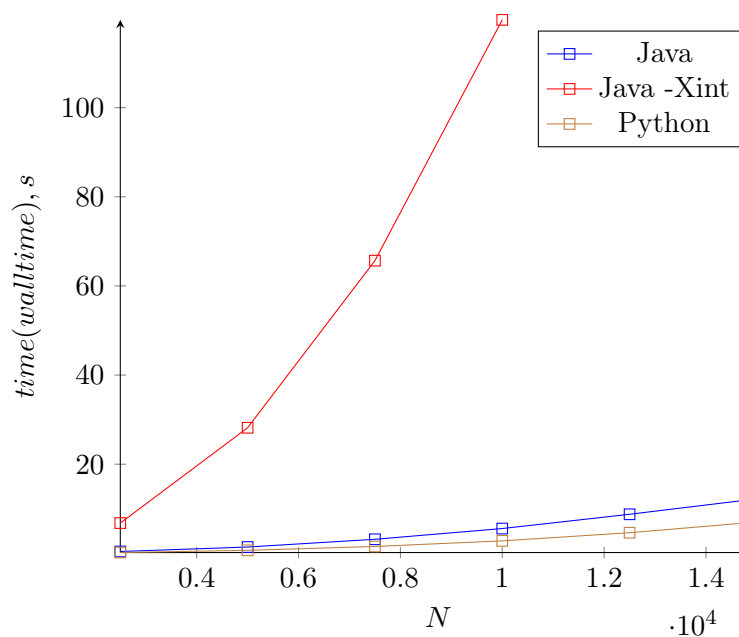
W przypadku Pythona walltime programu [\[Python 3 #4\]](#) przy wykorzystaniu polecenia `runex` wyniósł 2,86 s. Komplet wyników:

```
starttime=2024-12-20T18:03:53.441213+01:00  
returnvalue=0  
walltime=2.8576857300001848s  
cputime=2.857317687s  
cputime-cpu3=2.102434451s  
cputime-cpu9=0.754883236s  
memory=4349952B  
blkio-read=0B  
blkio-write=0B
```

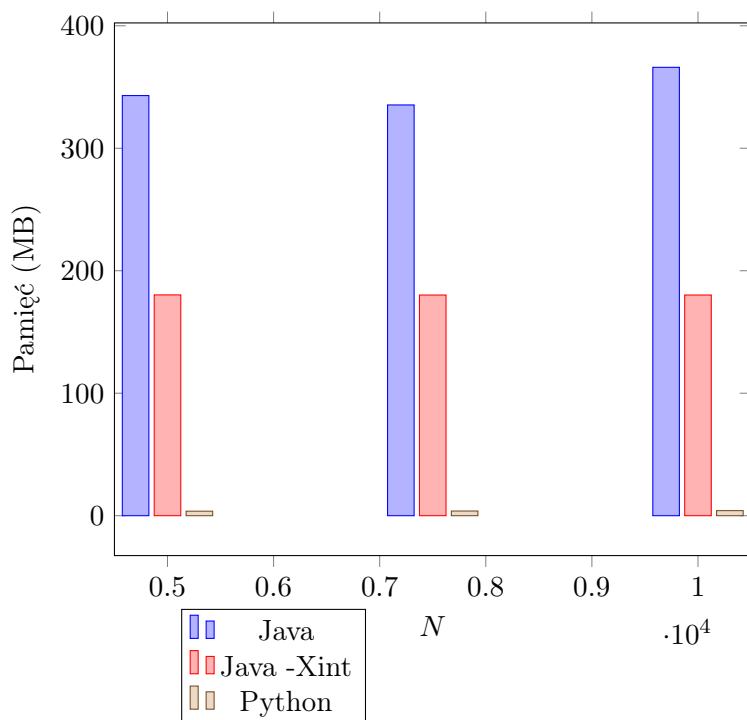
Jak widać w przypadku tego programu Java jest wolniejsza od Pythona. Może to wynikać z wykorzystania do obliczeń liczb typu `BigInteger`, które jak pokazaliśmy działają wolniej niż duże liczby w Pythonie.

Rozwiązaniem tego problemu może być wykorzystanie biblioteki GMP, która znacząco przyspiesza obliczenia arytmetyczne [\[Java #3\]](#).

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .



Jak widać Java potrzebuje do wykonania programu znacznie większej ilości pamięci niż Python niezależnie od wejściowej wartości N .

10.2.2. binary-trees

Ten program alokuje wpierw głębokie drzewo binarne aby rozciągnąć pamięć. Następnie alokuje drzewo binarne, które istnieje aż do końca działania programu. Pozostała część wykona-

nia polega na alokowaniu i dealokowaniu wielu drzew przy jednoczesnym zliczaniu liczby ich wierzchołków.

Podstawową liczbę dla pomiarów stanowi 21. Dokładniejszy opis: [9].

Java

W przypadku Javy walltime programu [Java #7] przy wykorzystaniu polecenia runexec wyniósł 2,28 s. Komplet wyników:

```
starttime=2024-12-20T19:58:59.059904+01:00
returnvalue=0
walltime=2.2797591759999705s
cputime=16.777071931s
cputime-cpu0=1.853680797s
cputime-cpu1=1.892253950s
cputime-cpu10=1.848870294s
cputime-cpu11=2.032843361s
cputime-cpu2=0.928579037s
cputime-cpu3=0.677434905s
cputime-cpu4=1.476493191s
cputime-cpu5=1.654540192s
cputime-cpu6=0.403547661s
cputime-cpu7=0.814800178s
cputime-cpu8=1.478957829s
cputime-cpu9=1.715070536s
memory=2565365760B
blkio-read=94208B
blkio-write=0B
```

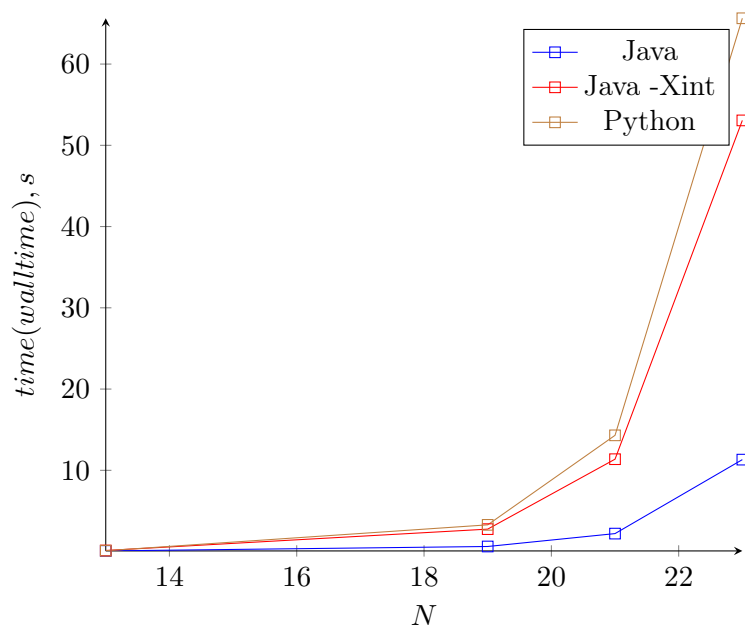
Python

W przypadku Pythona walltime programu [Python 3 #4] przy wykorzystaniu polecenia runexec wyniósł 15,6 s. Komplet wyników:

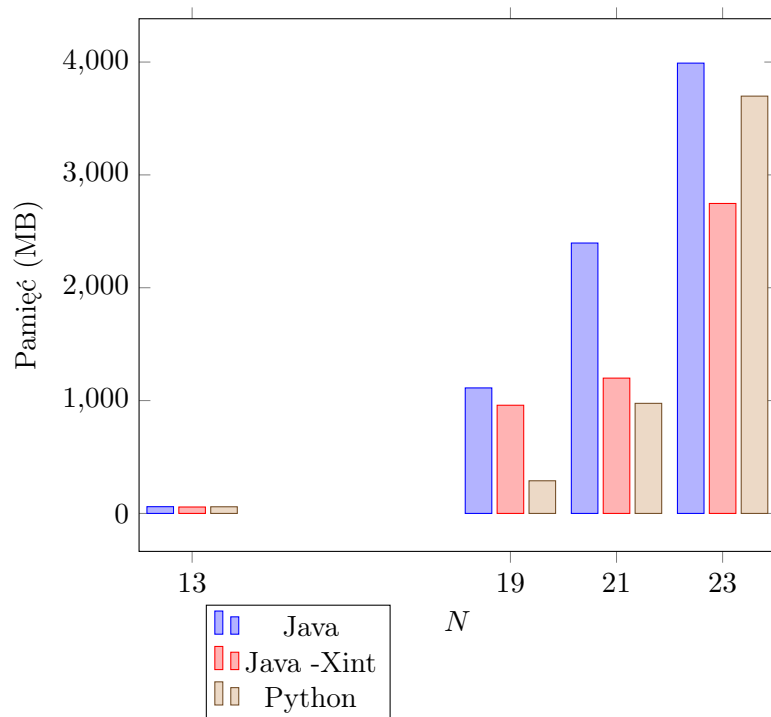
```
starttime=2024-12-20T19:59:42.184918+01:00
returnvalue=0
walltime=15.58560308199958s
cputime=142.584223976s
cputime-cpu0=11.524329568s
cputime-cpu1=11.418168383s
cputime-cpu10=11.543249594s
cputime-cpu11=11.336778209s
cputime-cpu2=11.776000717s
cputime-cpu3=11.450827422s
cputime-cpu4=14.548266545s
cputime-cpu5=11.834461877s
cputime-cpu6=11.801861248s
cputime-cpu7=11.703029944s
cputime-cpu8=11.800562881s
cputime-cpu9=11.846687588s
memory=1025269760B
blkio-read=4157440B
blkio-write=0B
```

Jak widać w przypadku tego programu Java jest wyraźnie szybsza od Pythona. Dla $N = 21$ jest to $15,6 \text{ s} / 2,28 \text{ s} = 6,84$ raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .



10.2.3. fannkuch-redux

W przypadku tego programu rozpatrywane są wszystkie permutacje liczb 1,...,n dla danego n. Dla każdej permutacji rozpatrywany jest jej pierwszy element, nazwijmy go x. Następnie porządek x pierwszych elementów jest odwracany. Procedura ta jest powtarzana aż pierwszym elementem stanie się 1. Zapamiętywana jest maksymalna liczba odwróceń spośród wszystkich permutacji. Obliczana jest również wartość checksum aby zagwarantować poprawność implementacji.

Dla pomiarów za bazowe N przyjąłem liczbę 12. Dokładniejszy opis: [10].

Java

W przypadku Javy walltime programu [Java] przy wykorzystaniu polecenia runexec wyniósł 3,56 s. Komplet wyników:

```
starttime=2024-12-24T12:38:48.923297+01:00
returnvalue=0
walltime=3.559818760999974s
cputime=40.68653424s
cputime-cpu0=3.422547344s
cputime-cpu1=3.466155625s
cputime-cpu10=3.316182076s
cputime-cpu11=3.447459427s
cputime-cpu2=3.412224327s
cputime-cpu3=3.411831401s
cputime-cpu4=3.314061044s
cputime-cpu5=3.302599848s
cputime-cpu6=3.352211472s
cputime-cpu7=3.460125386s
cputime-cpu8=3.293854854s
cputime-cpu9=3.487281436s
memory=31907840B
blkio-read=0B
blkio-write=0B
```

Python

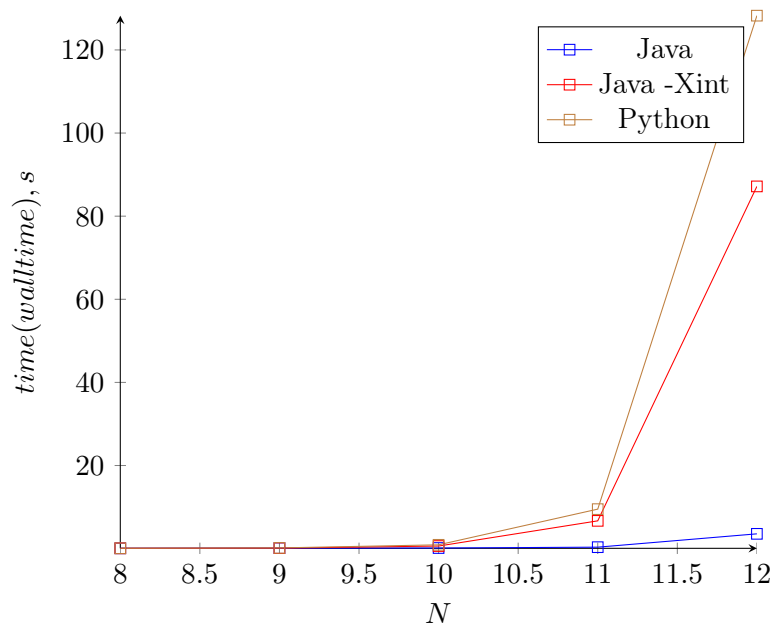
W przypadku Pythona walltime programu [Python 3 #4] przy wykorzystaniu polecenia runexec wyniósł 128 s. Komplet wyników:

```
starttime=2024-12-24T12:41:12.741818+01:00
returnvalue=0
walltime=128.24714689400002s
cputime=1487.220224683s
cputime-cpu0=120.576718880s
cputime-cpu1=125.762417764s
cputime-cpu10=122.406405365s
cputime-cpu11=124.644218967s
cputime-cpu2=124.072275975s
cputime-cpu3=124.788649866s
cputime-cpu4=122.481056674s
cputime-cpu5=124.992362780s
cputime-cpu6=121.394777709s
cputime-cpu7=126.898999067s
cputime-cpu8=123.979731513s
cputime-cpu9=125.222610123s
```

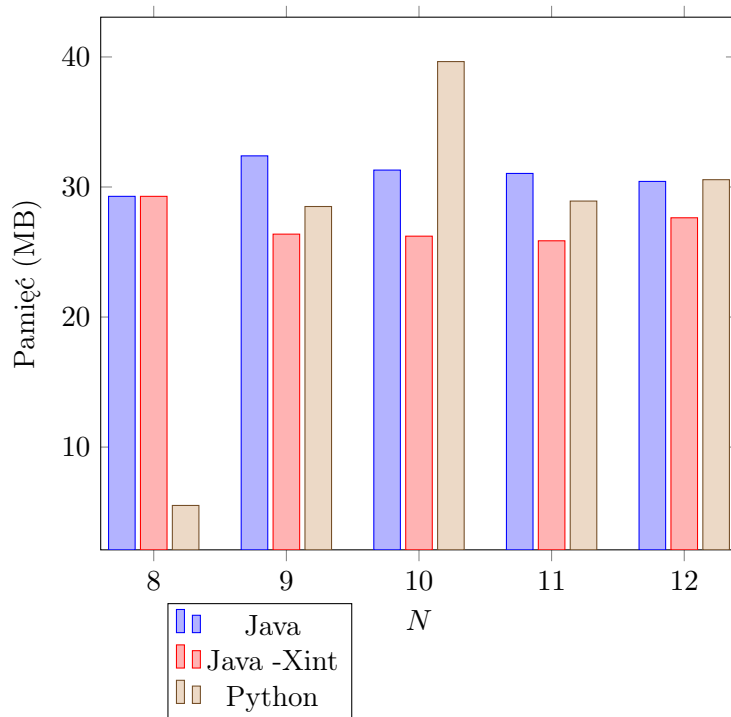
```
memory=32047104B  
blkio-read=0B  
blkio-write=0B
```

Jak widać w przypadku tego programu Java jest wyraźnie szybsza od Pythona. Dla $N = 12$ jest to $128\text{ s} / 3,56\text{ s} = 3,60$ raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .



10.2.4. N-body

Modeluje orbity planet z grupy Jowian (Jowisz, Saturn, Uran, and Neptun), używając tego samego prostego integratora symplektycznego.

Dla pomiarów za bazowe n przyjąłem liczbę 50000000. Dokładniejszy opis: [11].

Java

W przypadku Javy walltime programu [Java #4] przy wykorzystaniu polecenia runexec wyniósł 4,22 s. Komplet wyników:

```
starttime=2024-12-25T11:53:57.697005+01:00
returnvalue=0
walltime=4.217370437999989s
cputime=4.218599014s
cputime-cpu10=0.008972176s
cputime-cpu11=0.005400429s
cputime-cpu2=0.022193033s
cputime-cpu3=4.148934257s
cputime-cpu4=0.003043800s
cputime-cpu5=0.016407312s
cputime-cpu6=0.009972404s
cputime-cpu7=0.000178814s
cputime-cpu8=0.000066114s
cputime-cpu9=0.003430675s
memory=25059328B
blkio-read=544768B
blkio-write=0B
```

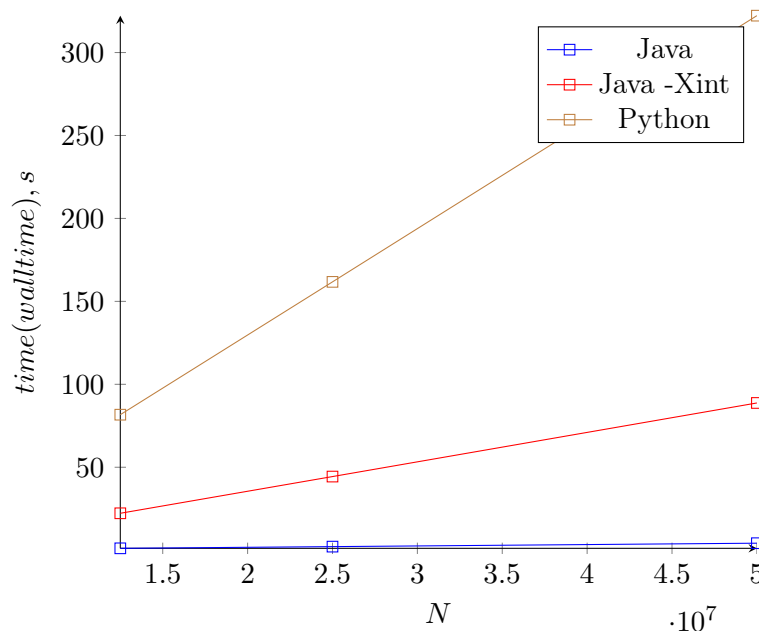
Python

W przypadku Pythona walltime programu [Python 3] przy wykorzystaniu polecenia runexec wyniósł 322 s. Komplet wyników:

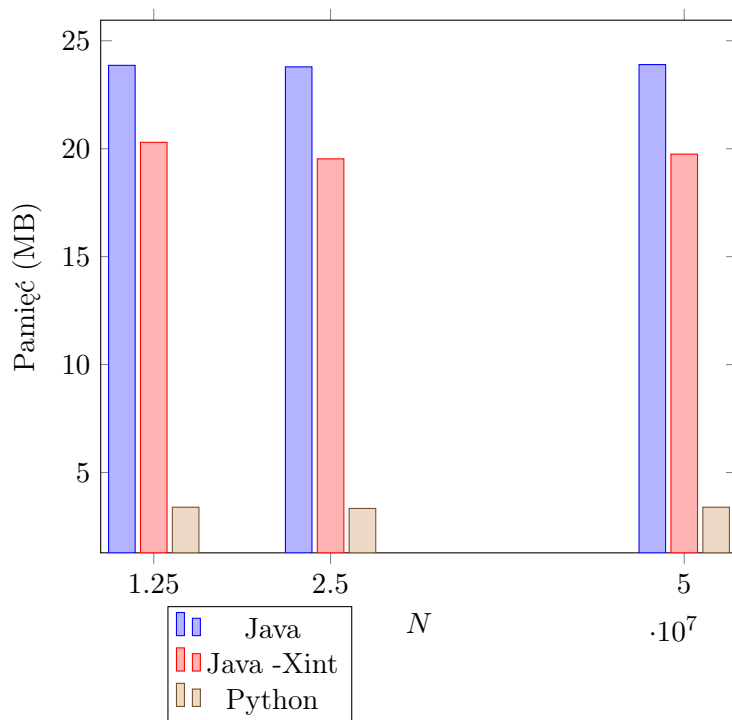
```
starttime=2024-12-25T11:57:15.773324+01:00
returnvalue=0
walltime=322.2353108489999s
cputime=322.226826032s
cputime-cpu0=97.789804078s
cputime-cpu10=0.012211654s
cputime-cpu11=72.973624307s
cputime-cpu3=1.512004426s
cputime-cpu5=22.195560776s
cputime-cpu6=127.743620791s
memory=3559424B
blkio-read=53248B
blkio-write=0B
```

Jak widać w przypadku tego programu Java jest wyraźnie szybsza od Pythona. Dla $n = 50000000$ jest to $322 \text{ s} / 4,22 \text{ s} = 76,3$ raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .



10.2.5. fasta

- generuje sekwencje DNA, kopiując z danej sekwencji
- generuje sekwencje DNA poprzez ważony losowy wybór z 2 alfabetów

Dla pomiarów za argument wejściowy przyjąłem 25000000. Dokładniejszy opis: [12].

Java

W przypadku Javy walltime programu [Java #6] przy wykorzystaniu polecenia runexec wyniósł 1.16 s. Komplet wyników:

```
starttime=2024-12-25T19:50:16.432118+01:00
returnvalue=0
walltime=1.1648875109999608s
cputime=4.734007739s
cputime-cpu0=0.547533909s
cputime-cpu1=0.477176959s
cputime-cpu10=0.005104631s
cputime-cpu11=0.128299400s
cputime-cpu2=0.377512691s
cputime-cpu3=0.578542008s
cputime-cpu4=0.240216090s
cputime-cpu5=0.852080439s
cputime-cpu6=0.441369131s
cputime-cpu7=0.478429609s
cputime-cpu8=0.533055693s
cputime-cpu9=0.074687179s
memory=298942464B
blkio-read=0B
blkio-write=0B
```

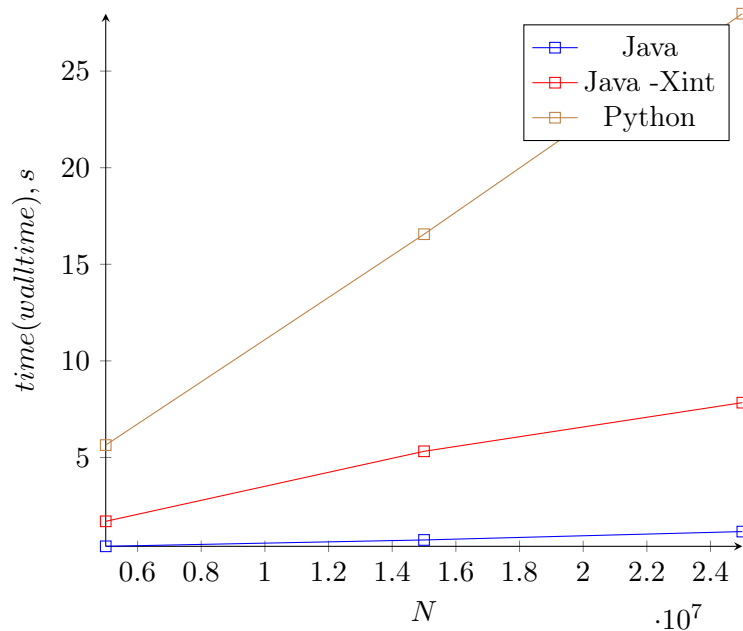
Python

W przypadku Pythona walltime programu [Python 3 #5] przy wykorzystaniu polecenia runnec wyniósł 28 s. Komplet wyników:

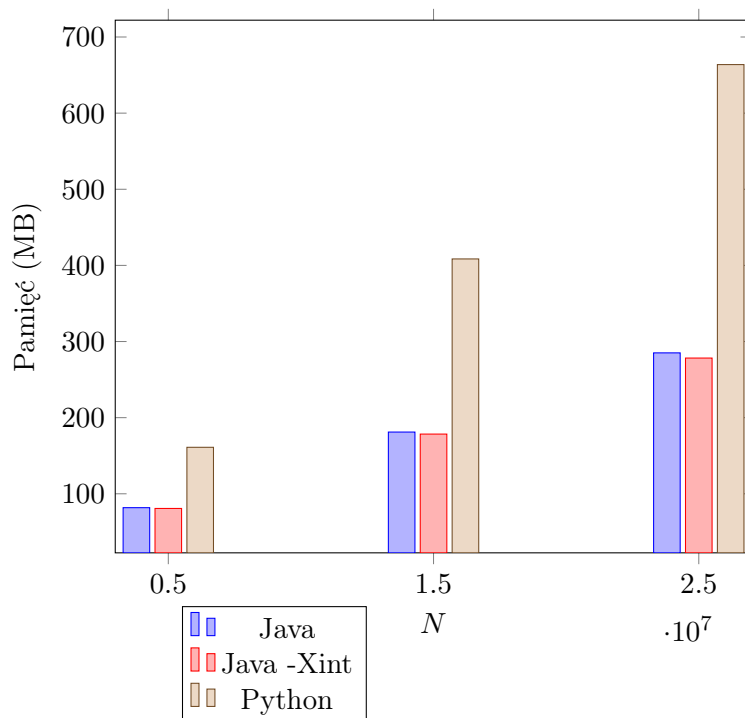
```
starttime=2024-12-25T19:56:00.767456+01:00
returnvalue=0
walltime=27.973746784000014s
cputime=49.00618006s
cputime-cpu0=2.295515155s
cputime-cpu1=11.375523946s
cputime-cpu10=5.328212206s
cputime-cpu11=13.194317161s
cputime-cpu2=2.877192826s
cputime-cpu3=1.743911116s
cputime-cpu4=0.995536218s
cputime-cpu5=4.384794556s
cputime-cpu6=2.489412690s
cputime-cpu7=2.882490619s
cputime-cpu8=1.439273567s
memory=696020992B
blkio-read=0B
blkio-write=0B
```

Jak widać w przypadku tego programu Java jest wyraźnie szybsza od Pythona. Dla argumentu 100000001 jest to $28 \text{ s} / 1,16 \text{ s} = 2,41$ raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .



10.2.6. reverse-complement

Znając sekwencję zasad jednej nici DNA, od razu znamy sekwencję nici DNA, która się z nią łączy, nie ta nazywa się odwrotnym dopełnieniem.

Program czyta ze standardowego wejścia plik w formacie FASTA i wypisuje na standardowe wyjście id, opis i odwrotne dopełnienie w formacie FASTA.

Dane wejściowe zostały wygenerowane za pomocą programu fasta. Długość sekwencji wynosi 100000001. Dokładniejszy opis: [13].

Java

W przypadku Javy walltime programu [Java #6] przy wykorzystaniu polecenia runexec wyniósł 2,37 s. Komplet wyników:

```
starttime=2024-12-25T13:36:16.654139+01:00
returnvalue=0
walltime=2.366201705000094s
cputime=2.680236005s
cputime-cpu0=0.000092554s
cputime-cpu1=0.000087434s
cputime-cpu10=0.001456222s
cputime-cpu11=0.001197352s
cputime-cpu2=0.000078177s
cputime-cpu3=2.040724949s
cputime-cpu4=0.016653725s
cputime-cpu5=0.013650926s
cputime-cpu6=0.000079108s
cputime-cpu7=0.601405798s
```

```
cputime-cpu8=0.000336001s
cputime-cpu9=0.004473759s
memory=3115700224B
blkio-read=0B
blkio-write=0B
```

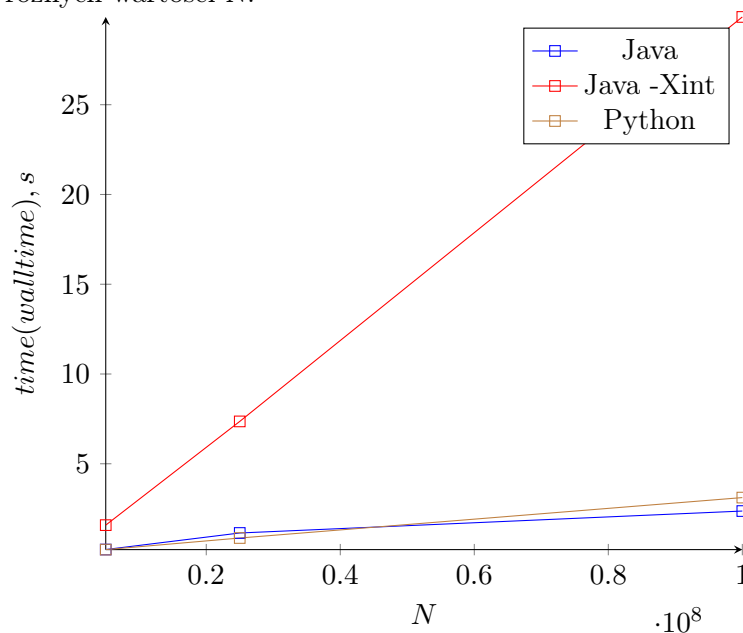
Python

W przypadku Pythona walltime programu [Python 3 #5] przy wykorzystaniu polecenia runnecx wyniósł 3,11 s. Komplet wyników:

```
starttime=2024-12-25T13:38:31.845847+01:00
returnvalue=0
walltime=3.1158195779999005s
cputime=5.790129265s
cputime-cpu0=1.019473178s
cputime-cpu1=0.567983424s
cputime-cpu10=0.009591560s
cputime-cpu2=0.000162765s
cputime-cpu3=0.007990056s
cputime-cpu5=0.053840911s
cputime-cpu6=1.809747270s
cputime-cpu7=0.583753191s
cputime-cpu8=0.001700389s
cputime-cpu9=1.735886521s
memory=2274304000B
blkio-read=0B
blkio-write=0B
```

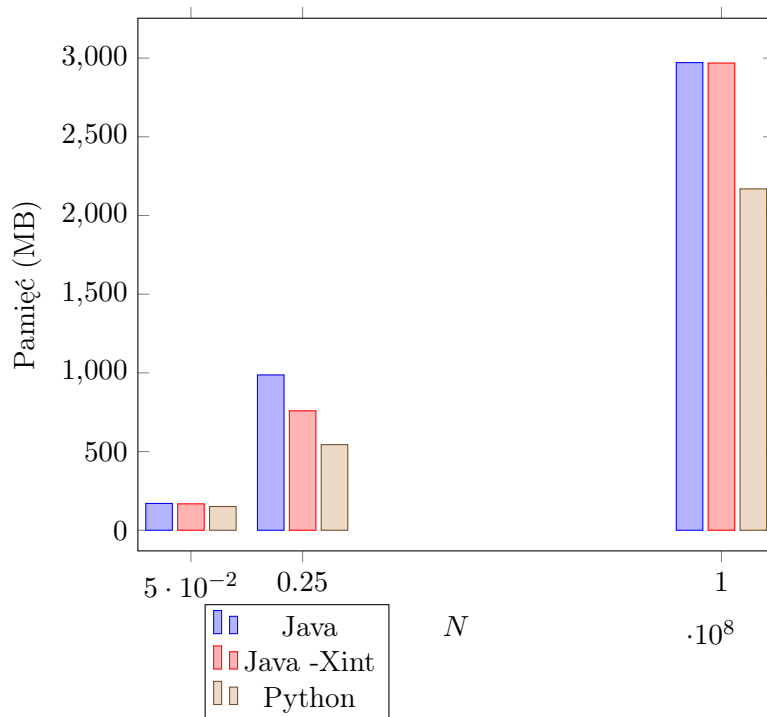
Jak widać w przypadku tego programu Java jest stosunkowo niewiele szybsza od Pythona. Dla $N = 100000001$ jest to $3,11 \text{ s} / 2,37 \text{ s} = 1,31$ raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla

różnych wartości N .



10.2.7. mandelbrot

Kreśli zbiór Mandelbrota $[-1.5-i, 0.5+i]$ na bitmapie N na N . Zapisuje wynik bajt po bajcie w pliku PBM.

Dla pomiarów za argument wejściowy przyjąłem 16000. Dokładniejszy opis: [14].

Java

W przypadku Javy walltime programu [Java #2] przy wykorzystaniu polecenia runexec wyniósł 1,13 s. Komplet wyników:

```
walltime=1.1342887529999643s
cputime=11.93891101s
cputime-cpu0=1.019560255s
cputime-cpu1=0.989544058s
cputime-cpu10=0.975318721s
cputime-cpu11=1.044600389s
cputime-cpu2=0.983189828s
cputime-cpu3=0.978190282s
cputime-cpu4=0.990735071s
cputime-cpu5=0.997777040s
cputime-cpu6=0.987225069s
cputime-cpu7=1.012641294s
cputime-cpu8=0.981878383s
cputime-cpu9=0.978250620s
memory=112623616B
blkio-read=0B
```

```
blkio-write=0B
```

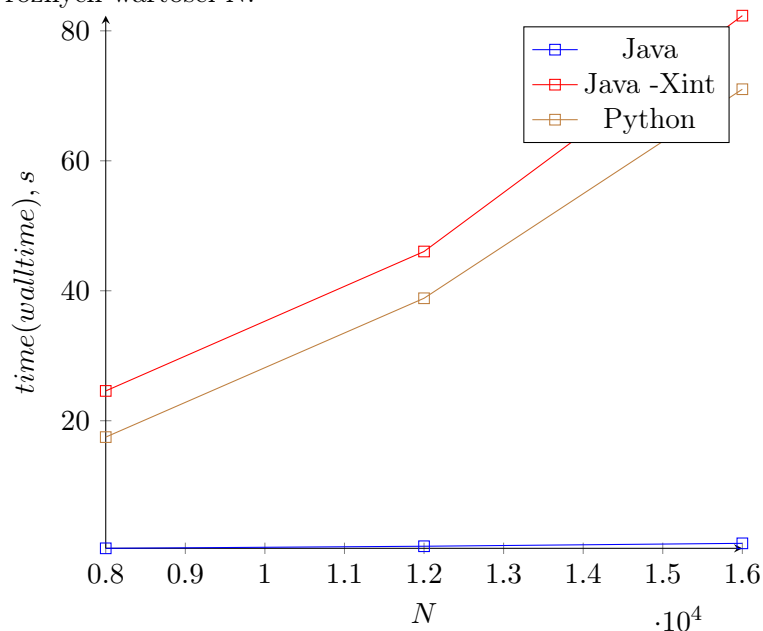
Python

W przypadku Pythona walltime programu [Python 3 #7] przy wykorzystaniu polecenia runnexec wyniósł 71 s. Komplet wyników:

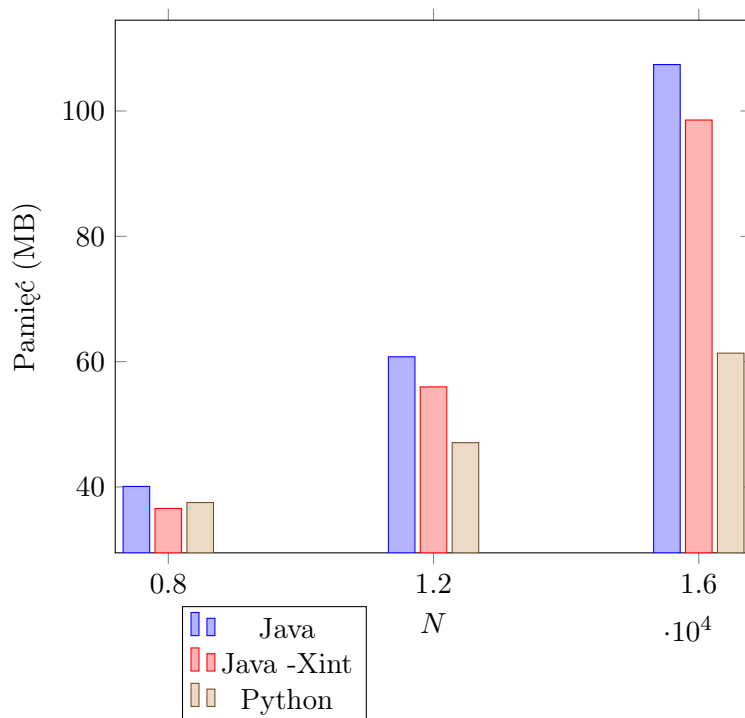
```
starttime=2024-12-25T20:12:15.167569+01:00
returnvalue=0
walltime=71.016525645999999s
cputime=843.977451867s
cputime-cpu0=70.294594650s
cputime-cpu1=70.500735403s
cputime-cpu10=69.902568873s
cputime-cpu11=70.183822083s
cputime-cpu2=70.520069691s
cputime-cpu3=70.074994223s
cputime-cpu4=70.495964229s
cputime-cpu5=70.273095224s
cputime-cpu6=70.291708382s
cputime-cpu7=70.591711177s
cputime-cpu8=70.442063069s
cputime-cpu9=70.406124863s
memory=64335872B
blkio-read=151552B
blkio-write=0B
```

Jak widać w przypadku tego programu Java jest wyraźnie szybsza od Pythona. Dla argumentu 16000 jest to $71\text{ s} / 1,13\text{ s} = 62,8$ raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .



10.2.8. spectral-norm

Oblicza normę widmową nieskończonej macierzy A z wpisami $a_{11} = 1, a_{12} = 1/2, a_{21} = 1/3, a_{13} = 1/4, a_{22} = 1/5, a_{31} = 1/6$, itd.

Dla pomiarów za argument wejściowy przyjąłem 5500. Dokładniejszy opis: [15].

Java

W przypadku Javy walltime programu [Java #3] przy wykorzystaniu polecenia runexec wyniósł 0,618 s. Komplet wyników:

```
starttime=2024-12-25T20:19:10.698662+01:00
returnvalue=0
walltime=0.6107691810002507s
cputime=4.24365897s
cputime-cpu0=0.502503172s
cputime-cpu1=0.299692284s
cputime-cpu10=0.303250014s
cputime-cpu11=0.292081007s
cputime-cpu2=0.415043783s
cputime-cpu3=0.294426456s
cputime-cpu4=0.338604593s
cputime-cpu5=0.284607222s
cputime-cpu6=0.511081268s
cputime-cpu7=0.284779923s
cputime-cpu8=0.411520472s
cputime-cpu9=0.306068776s
memory=27381760B
blkio-read=0B
blkio-write=0B
```

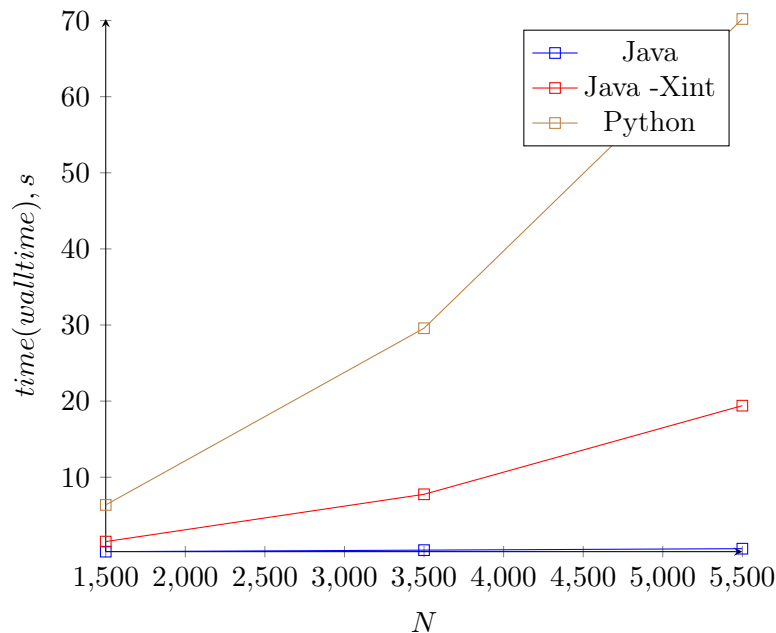
Python

W przypadku Pythona walltime programu [Python 3 #4] przy wykorzystaniu polecenia runexec wyniósł 70,2 s. Komplet wyników:

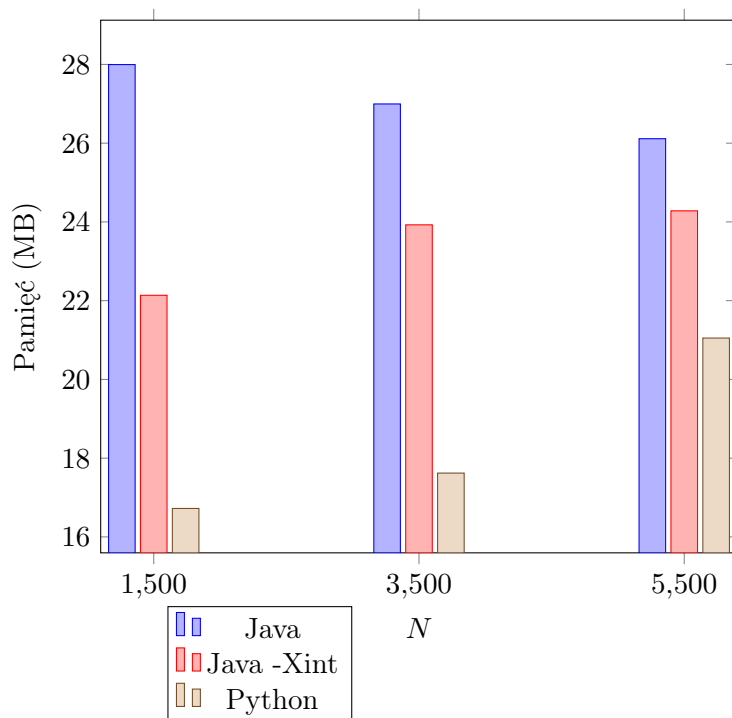
```
starttime=2024-12-25T20:24:30.374322+01:00
returnvalue=0
walltime=70.20847665300016s
cputime=277.682246331s
cputime-cpu0=23.436026862s
cputime-cpu1=33.131612600s
cputime-cpu10=14.584188744s
cputime-cpu11=12.294572871s
cputime-cpu2=16.686515146s
cputime-cpu3=39.313063969s
cputime-cpu4=46.249951552s
cputime-cpu5=8.991012702s
cputime-cpu6=8.721421147s
cputime-cpu7=31.267817521s
cputime-cpu8=27.581012635s
cputime-cpu9=15.425050582s
memory=22110208B
blkio-read=0B
blkio-write=0B
```

Jak widać w przypadku tego programu Java jest wyraźnie szybsza od Pythona. Dla argumentu 5500 jest to 70.2 s/0.618 s = 114 raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .



10.3. Przerobione programy ze strony geeksforgeeks.org

W tej sekcji porównuję czas działania przerobionych programów testowych pochodzących ze strony geeksforgeeks.org.

10.3.1. Mergesort

Program Mergesort sortuje przy pomocy algorytmu sortowania mergesort listę/tablicę, w której znajdują się liczby uporządkowane malejąco: od n do 1. W pomiarach za bazowe n przyjąłem 1000000.

Java

Kod źródłowy programu w Javie [\[Java\]](#):

```

1 // Java program for Merge Sort
2 import java.io.*;
3
4 class GfG {
5
6     // Merges two subarrays of arr[].
7     // First subarray is arr[l..m]
8     // Second subarray is arr[m+1..r]
9     static void merge(int arr[], int l, int m, int r)
10    {
11        // Find sizes of two subarrays to be merged
12        int n1 = m - l + 1;
13        int n2 = r - m;
14    }

```

```

15 // Create temp arrays
16 int L[] = new int[n1];
17 int R[] = new int[n2];
18
19 // Copy data to temp arrays
20 for (int i = 0; i < n1; ++i)
21     L[i] = arr[l + i];
22 for (int j = 0; j < n2; ++j)
23     R[j] = arr[m + 1 + j];
24
25 // Merge the temp arrays
26
27 // Initial indices of first and second subarrays
28 int i = 0, j = 0;
29
30 // Initial index of merged subarray array
31 int k = l;
32 while (i < n1 && j < n2) {
33     if (L[i] <= R[j]) {
34         arr[k] = L[i];
35         i++;
36     }
37     else {
38         arr[k] = R[j];
39         j++;
40     }
41     k++;
42 }
43
44 // Copy remaining elements of L[] if any
45 while (i < n1) {
46     arr[k] = L[i];
47     i++;
48     k++;
49 }
50
51 // Copy remaining elements of R[] if any
52 while (j < n2) {
53     arr[k] = R[j];
54     j++;
55     k++;
56 }
57 }
58
59 // Main function that sorts arr[l..r] using
60 // merge()
61 static void sort(int arr[], int l, int r)
62 {
63     if (l < r) {
64
65         // Find the middle point
66         int m = l + (r - l) / 2;
67
68         // Sort first and second halves
69         sort(arr, l, m);
70         sort(arr, m + 1, r);
71
72         // Merge the sorted halves
73         merge(arr, l, m, r);
74     }

```



```

75     }
76
77     // Driver code
78     public static void main(String args[])
79     {
80         int n = Integer.parseInt(args[0]);
81
82         int arr[] = new int[n];
83         for (int i = 0; i < n; i++)
84             arr[i] = n - i;
85
86         sort(arr, 0, n - 1);
87     }
88 }

```

W przypadku Javy walltime programu dla $n=1000000$ przy wykorzystaniu polecenia `runexec` wyniósł 0,111 s. Komplet wyników:

```

starttime=2024-12-23T18:23:52.973959+01:00
returnvalue=0
walltime=0.11144095099962215s
cputime=0.11972095s
cputime-cpu0=0.020661110s
cputime-cpu1=0.003487471s
cputime-cpu10=0.005665164s
cputime-cpu11=0.001796297s
cputime-cpu2=0.000269566s
cputime-cpu3=0.001961156s
cputime-cpu4=0.004080414s
cputime-cpu5=0.001882721s
cputime-cpu6=0.000039865s
cputime-cpu8=0.078099802s
cputime-cpu9=0.001777384s
memory=90558464B
blkio-read=0B
blkio-write=0B

```

Python

Kod źródłowy programu w Pythonie [\[Python\]](#):

```

1 from sys import argv
2
3 def merge(arr, left, mid, right):
4     n1 = mid - left + 1
5     n2 = right - mid:
6         arr[k] = R[j]
7         j += 1
8         k += 1
9
10 def merge_sort(arr, left, right):
11     if left < right:
12         mid = (left + right) // 2
13
14         merge_sort(arr, left, mid)
15         merge_sort(arr, mid + 1, right)
16         merge(arr, left, mid, right)
17
18 def main():
19     n=int(argv[1])

```

```

20     arr = [0] * n
21     for i in range(0, n):
22         arr[i] = n - i
23
24     merge_sort(arr, 0, n - 1)
25
26 main() right - mid
27
28     # Create temp arrays
29     L = [0] * n1
30     R = [0] * n2
31
32     # Copy data to temp arrays L[] and R[]
33     for i in range(n1):
34         L[i] = arr[left + i]
35     for j in range(n2):
36         R[j] = arr[mid + 1 + j]
37
38     i = 0 # Initial index of first subarray
39     j = 0 # Initial index of second subarray
40     k = left # Initial index of merged subarray
41
42     # Merge the temp arrays back
43     # into arr[left..right]
44     while i < n1 and j < n2:
45         if L[i] <= R[j]:
46             arr[k] = L[i]
47             i += 1
48         else:
49             arr[k] = R[j]
50             j += 1
51         k += 1
52
53     # Copy the remaining elements of L[],
54     # if there are any
55     while i < n1:
56         arr[k] = L[i]
57         i += 1
58         k += 1
59
60     # Copy the remaining elements of R[],
61     # if there are any
62     while j < n2:
63         arr[k] = R[j]
64         j += 1
65         k += 1
66
67 def merge_sort(arr, left, right):
68     if left < right:
69         mid = (left + right) // 2
70
71         merge_sort(arr, left, mid)
72         merge_sort(arr, mid + 1, right)
73         merge(arr, left, mid, right)
74
75 def main():
76     n=int(argv[1])
77     arr = [0] * n
78     for i in range(0, n):
79         arr[i] = n - i

```

```

80
81     merge_sort(arr, 0, n - 1)
82
83 main()

```

W przypadku Pythona walltime programu dla $n=1000000$ przy wykorzystaniu polecenia `ru-nexec` wyniósł 3,55 s. Komplet wyników:

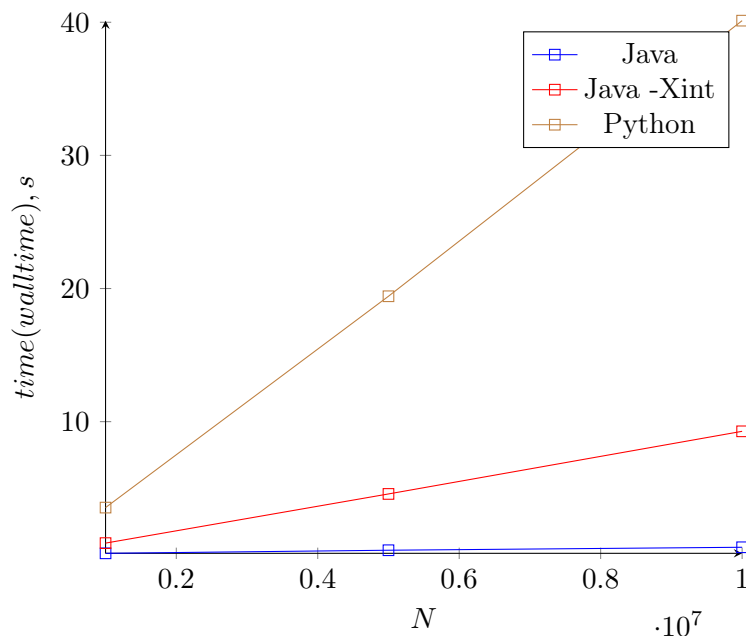
```

starttime=2024-12-23T18:26:32.246044+01:00
returnvalue=0
walltime=3.5459886850003386s
cputime=3.545602191s
cputime-cpu11=3.540892358s
cputime-cpu5=0.004709833s
memory=51707904B
blkio-read=0B
blkio-write=0B

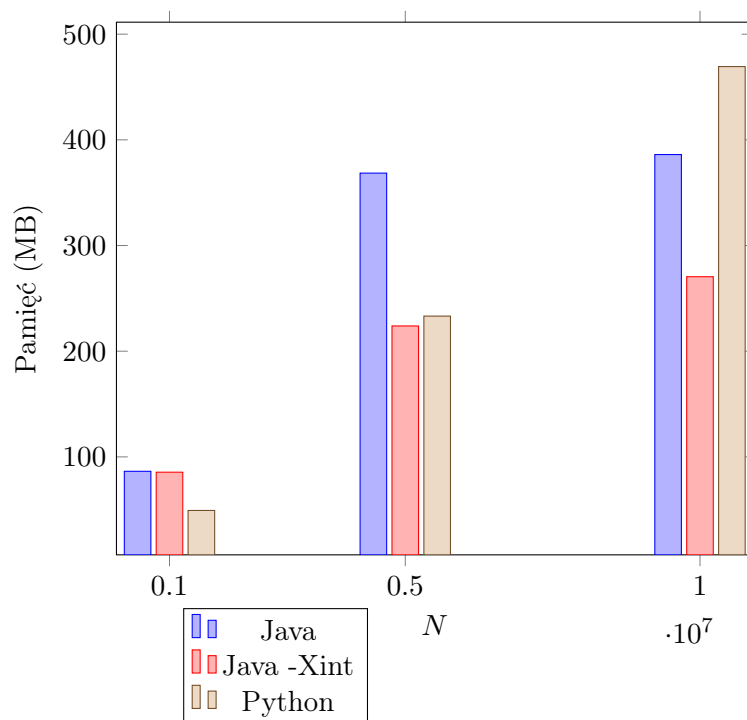
```

Jak widać w przypadku tego programu Java jest wyraźnie szybsza od Pythona. Dla $n=1000000$ jest to $3,55 \text{ s} / 0,111 \text{ s} = 3,20$ raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .



10.3.2. Problem n hetmanów

Mając daną liczbę całkowitą n , zadanie polega na znalezieniu wszystkich różnych rozwiązań problemu n -królowych, w którym n hetmanów jest umieszczonych na szachownicy $n \times n$ w taki sposób, że żadne dwie hetmany nie mogą się atakować.

Każde rozwiązanie jest unikalną konfiguracją n królowych, reprezentowaną jako permutacja $[1, 2, 3, \dots, n]$. Liczba na i -tym miejscu oznacza rząd królowej w i -tej kolumnie. Na przykład $[3, 1, 4, 2]$ reprezentuje rozwiązanie dla $n=4$. W pomiarach za bazowe n przyjąłem 13.

Java

Kod źródłowy programu w Javie [\[Java\]](#):

```

1 import java.util.ArrayList;
2 import java.util.List;
3
4 // Utility class for N-Queens
5 class GfG {
6
7     // Utility function for solving the N-Queens
8     // problem using backtracking.
9     static void nQueenUtil(int j, int n,
10         List<Integer> board,
11         List<List<Integer>> result,
12         boolean[] rows, boolean[] diag1,
13         boolean[] diag2){
14
15         if (j > n) {
16
17             // A solution is found
18         }
19     }
20 }

```

```

17         result.add(new ArrayList<>(board));
18         return;
19     }
20     for (int i = 1; i <= n; i++) {
21         if (!rows[i] && !diag1[i + j]
22             && !diag2[i - j + n]) {
23
24             // Place queen
25             rows[i] = diag1[i + j] = diag2[i - j + n]
26                 = true;
27             board.add(i);
28
29             // Recurse to next column
30             nQueenUtil(j + 1, n, board, result, rows,
31                 diag1, diag2);
32
33             // Remove queen (backtrack)
34             board.remove(board.size() - 1);
35             rows[i] = diag1[i + j] = diag2[i - j + n]
36                 = false;
37         }
38     }
39 }
40
41 // Solves the N-Queens problem and returns
42 // all valid configurations.
43 static List<List<Integer> > nQueen(int n){
44
45     List<List<Integer> > result = new ArrayList<>();
46     List<Integer> board = new ArrayList<>();
47     boolean[] rows = new boolean[n + 1];
48     boolean[] diag1 = new boolean[2 * n + 1];
49     boolean[] diag2 = new boolean[2 * n + 1];
50
51     // Start solving from first column
52     nQueenUtil(1, n, board, result, rows, diag1, diag2);
53     return result;
54 }
55
56 public static void main(String[] args){
57
58     int n = Integer.parseInt(args[0]);
59     List<List<Integer> > result = nQueen(n);
60     //for (List<Integer> res : result) {
61     //    System.out.print("[");
62     //    for (int i = 0; i < res.size(); i++) {
63     //        System.out.print(res.get(i));
64     //        if (i != res.size() - 1)
65     //            System.out.print(", ");
66     //    }
67     //    System.out.println("]");
68     //}
69 }
70 }

```

W przypadku Javy walltime programu dla n=13 przy wykorzystaniu polecenia runexec wyniósł 0,459 s. Komplet wyników:

```

starttime=2024-12-23T18:44:30.970948+01:00
returnvalue=0
walltime=0.4590756580000743s

```

```
cputime=0.506433078s
cputime-cpu0=0.005824122s
cputime-cpu1=0.059389618s
cputime-cpu10=0.005215852s
cputime-cpu11=0.002777278s
cputime-cpu2=0.003773366s
cputime-cpu6=0.000202028s
cputime-cpu7=0.000066765s
cputime-cpu8=0.429184049s
memory=33517568B
blkio-read=0B
blkio-write=0B
```

Python

Kod źródłowy programu w Pythonie [\[Python\]](#):

```
1 from sys import argv
2
3 # Utility function for solving the N-Queens
4 # problem using backtracking.
5 def nQueenUtil(j, n, board, result,
6                 rows, diag1, diag2):
7     if j > n:
8
9         # A solution is found
10        result.append(board.copy())
11        return
12    for i in range(1, n + 1):
13        if not rows[i] and not diag1[i + j] and \
14            not diag2[i - j + n]:
15
16            # Place queen
17            rows[i] = diag1[i + j] = \
18                diag2[i - j + n] = True
19            board.append(i)
20
21            # Recurse to next column
22            nQueenUtil(j + 1, n, board,
23                      result, rows, diag1, diag2)
24
25            # Remove queen (backtrack)
26            board.pop()
27            rows[i] = diag1[i + j] = \
28                diag2[i - j + n] = False
29
30 # Solves the N-Queens problem and returns
31 # all valid configurations.
32 def nQueen(n):
33     result = []
34     board = []
35     rows = [False] * (n + 1)
36     diag1 = [False] * (2 * n + 1)
37     diag2 = [False] * (2 * n + 1)
38
39     # Start solving from first column
40     nQueenUtil(1, n, board, result,
41               rows, diag1, diag2)
42     return result
43
```

```

44 def main():
45     n=int(argv[1])
46     result = nQueen(n)
47     # for res in result:
48     #     print("[", end="")
49     #     for i in range(len(res)):
50     #         print(res[i], end="")
51     #         if i != len(res)-1:
52     #             print(", ", end="")
53     #     print("]")
54
55 main()

```

W przypadku Pythona walltime programu dla $n=13$ przy wykorzystaniu polecenia runexec wyniósł 3,44 s. Komplet wyników:

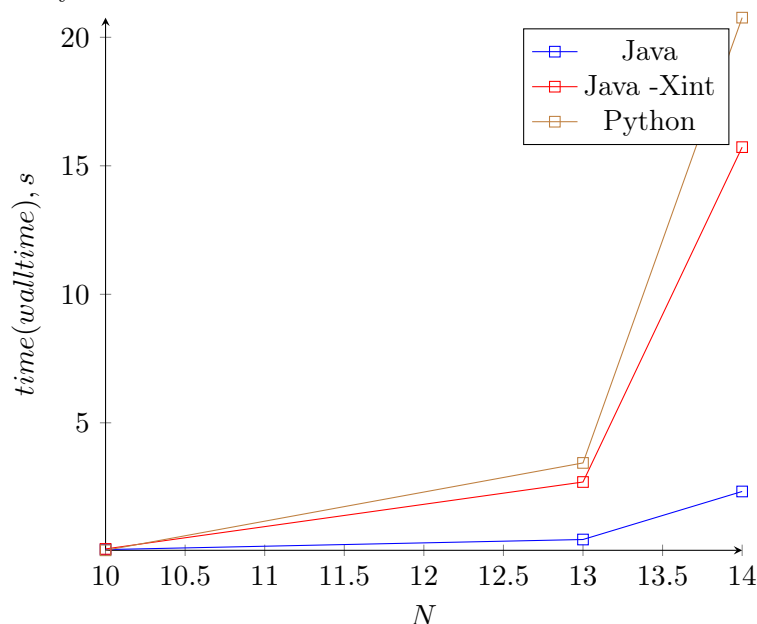
```

starttime=2024-12-23T18:45:24.910125+01:00
returnvalue=0
walltime=3.439188954999736s
cputime=3.401129201s
cputime-cpu0=1.877275741s
cputime-cpu6=1.523853460s
memory=17215488B
blkio-read=0B
blkio-write=0B

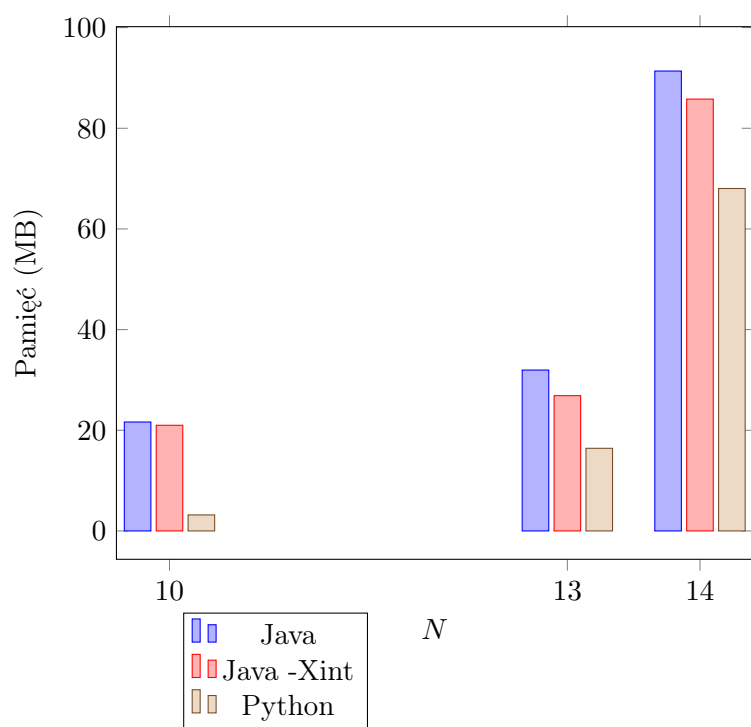
```

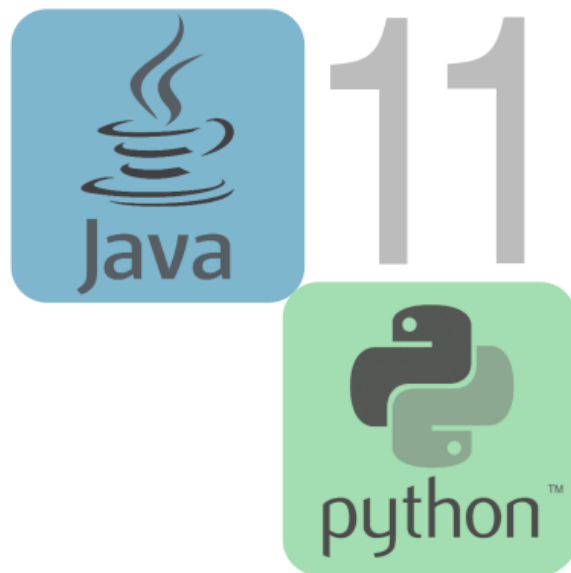
Jak widać w przypadku tego programu Java jest wyraźnie szybsza od Pythona. Dla $n=13$ jest to 3,44 s/0,459 s = 7,49 raza szybciej.

Poniżej znajduje się wykres czasu działania podstawowego algorytmu w Javie i Pythonie dla różnych wartości N .



Poniżej znajduje się wykres zapotrzebowania na pamięć w przypadku Javy i Pythona dla różnych wartości N .





Rozdział 11

Podsumowanie

W pracy porównałem bajtkody i maszyny wirtualne Javy i Pythona. Omówiłem architekturę JVM i PVM, JIT Javy i stosowane przez niego optymalizacje oraz dopiero co dodany JIT do Pythona 3.13. Dokonałem analizy porównawczej kolekcji Javy i Pythona. Porównałem czas działania i zużycie pamięci dwóch programów (Mergesort, problem n hetmanów) ze strony [geeksforgeeks \[6\]](#) w Javie i Pythonie. Zbadałem też czas działania oraz zużycie pamięci programów rozwiązujących te same problemy w Javie i Pythonie ze strony CLBG [\[7\]](#): binary-trees, fannkuch-redux, n-body, fasta, reverse-complment, mandelbrot, spectral-norm.

Doszedłem do następujących wniosków:

- Piętą achillesową Javy jest obsługa dużych liczb. Klasa BigInteger Javy działa wolniej niż duże liczby w Pythonie co sprawia, że programy na niej oparte są również wolniejsze od analogicznych programów w Pythonie. Z pomocą przychodzą tu Javie zewnętrzne biblioteki takie jak GMP, ale komplikują one z kolei kod programu.
- Z pomocą Pythonowi przychodzą zewnętrzne biblioteki takie jak np. NumPy do obliczeń numerycznych, ale nie zawsze ich zastosowanie przyspiesza działanie programów.
- Okazuje się, że w niektórych przypadkach kolekcje Pythona są szybsze od ich odpowiedników w Javie. Ma to miejsce np. w przypadku metody containsKey klasy TreeSet (Java) i in klasy SortedSet (Python). Jest to związane z tym, że złożoność tej metody w Javie wynosi $O(\log N)$ a w Pythonie $O(1)$. W przypadku, gdy złożoność metody jest taka sama w obu językach, Java okazuje się szybsza.
- Kod Pythonowy działa szybciej w funkcji niż poza nią. Jest to odczuwalne, gdy elementów, po których iterujemy jest dużo.

Bibliografia

- [1] *The Java® Virtual Machine Specification*, <https://docs.oracle.com/javase/specs/jvms/se13/html/index.html>, 2019-08-21
- [2] Sanket Bhosale, *Memory Management in JVM and PVM: A Comparative Overview*, https://www.linkedin.com/posts/sanket-bhosale-3108751a7_java-python-core2web-activity-7134199599744843776-rjUK/, 2024
- [3] Sakshee Agrawal, *Understanding Just-In-Time (JIT) Compilation in Java*, https://medium.com/@sakshee_agrawal/understanding-just-in-time-jit-compilation-in-java-ae2a6b9fa931, 2023-10-17
- [4] *JVM vs PVM*, <http://simplealgo.com/jvm-vs.pvm/>, 2024
- [5] Obi Ike-Nwosu, *Inside The Python Virtual Machine*, <https://leanpub.com/insidethepythonvirtualmachine/read>, 2020-08-07
- [6] Gouy, Isaac. *geeksforgeeks*, <https://www.geeksforgeeks.org/> (odwiedzone 31/12/2024).
- [7] Gouy, Isaac. *The Computer Language Benchmarks Game*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/index.html> (odwiedzone 31/12/2024).
- [8] Gouy, Isaac. *The Computer Language Benchmarks Game: pidigits description.*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/description/pidigits.html#pidigits> (odwiedzone 31/12/2024).
- [9] Gouy, Isaac. *The Computer Language Benchmarks Game: binarytrees description.*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/description/binarytrees.html#binarytrees> (odwiedzone 31/12/2024).
- [10] Gouy, Isaac. *The Computer Language Benchmarks Game: fannkuchredux description.*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/description/fannkuchredux.html#fannkuchredux> (odwiedzone 31/12/2024).
- [11] Gouy, Isaac. *The Computer Language Benchmarks Game: nbody description.*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/description/nbody.html#nbody> (odwiedzone 31/12/2024).
- [12] Gouy, Isaac. *The Computer Language Benchmarks Game: fasta description.*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/description/fasta.html#fasta> (odwiedzone 31/12/2024).
- [13] Gouy, Isaac. *The Computer Language Benchmarks Game: revcomp description.*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/description/revcomp.html#revcomp> (odwiedzone 31/12/2024).

- [14] Gouy, Isaac. *The Computer Language Benchmarks Game: mandelbrot description.*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/description/mandelbrot.html#mandelbrot> (odwiedzone 31/12/2024).
- [15] Gouy, Isaac. *The Computer Language Benchmarks Game: spectralnorm description.*, <https://benchmarksgame-team.pages.debian.net/benchmarksgame/description/spectralnorm.html#spectralnorm> (odwiedzone 31/12/2024).
- [16] Liquid Web *Python version history*, <https://www.liquidweb.com/blog/latest-python-version/>.
- [17] Hitesh Umaletiya *History of various Java versions*, <https://www.brilworks.com/blog/java-versions-new-features-and-deprecation/>.
- [18] Admin 3 *Java Version History*, <https://manaschool.in/history-of-java/>.
- [19] *Maszyna wirtualna*, https://pl.wikipedia.org/wiki/Maszyna_wirtualna.
- [20] *Bajtkod*, https://pl.wikipedia.org/wiki/Kod_bajtowy.