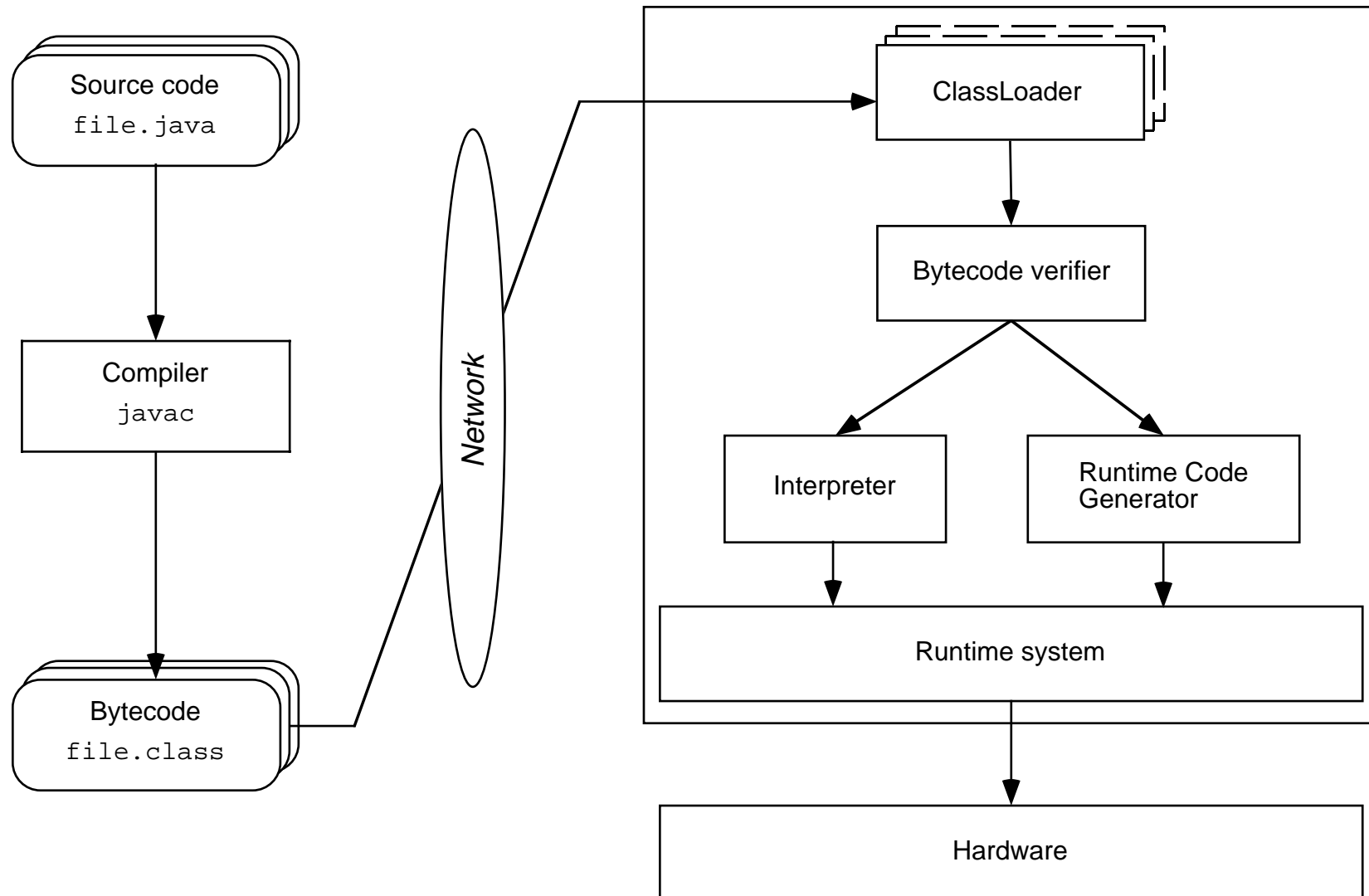


ANALYSIS OF THE JAVA CLASS FILE

DENIS N. ANTONIOLI, MARKUS PILZ
{antonio,pilz}@ifi.unizh.ch

**DEPARTMENT OF COMPUTER SCIENCE,
UNIVERSITY OF ZÜRICH**
<<http://www.ifi.unizh.ch>>



ifi THE IMPORTANCE OF THE CLASS FILE

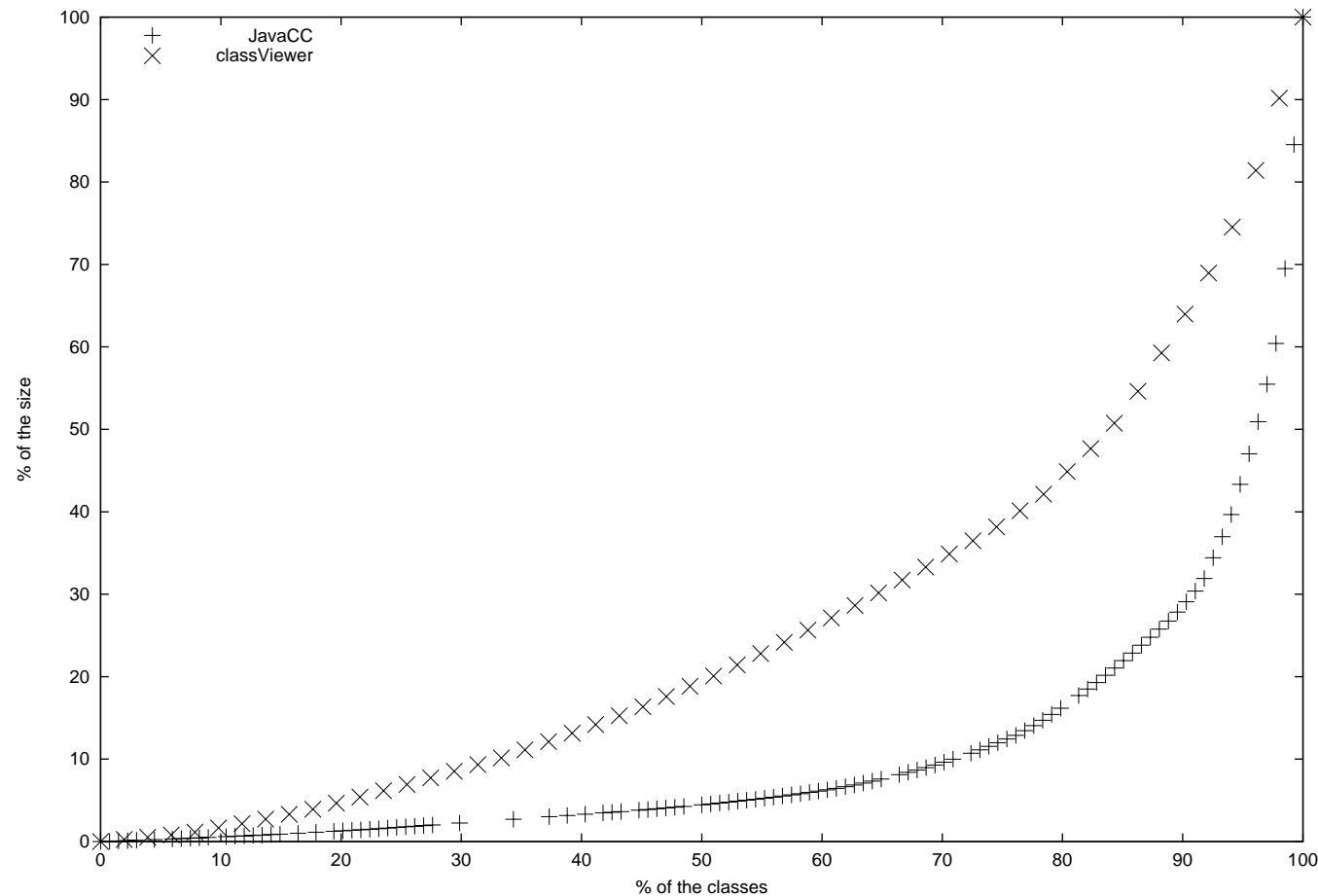
The choice of an intermediary format has consequences for:

- **distribution**
 - ◆ transmission time
 - ◆ secondary storage efficiency
- **execution**
 - ◆ security checks
 - ◆ run time linking
 - ◆ run time code generation

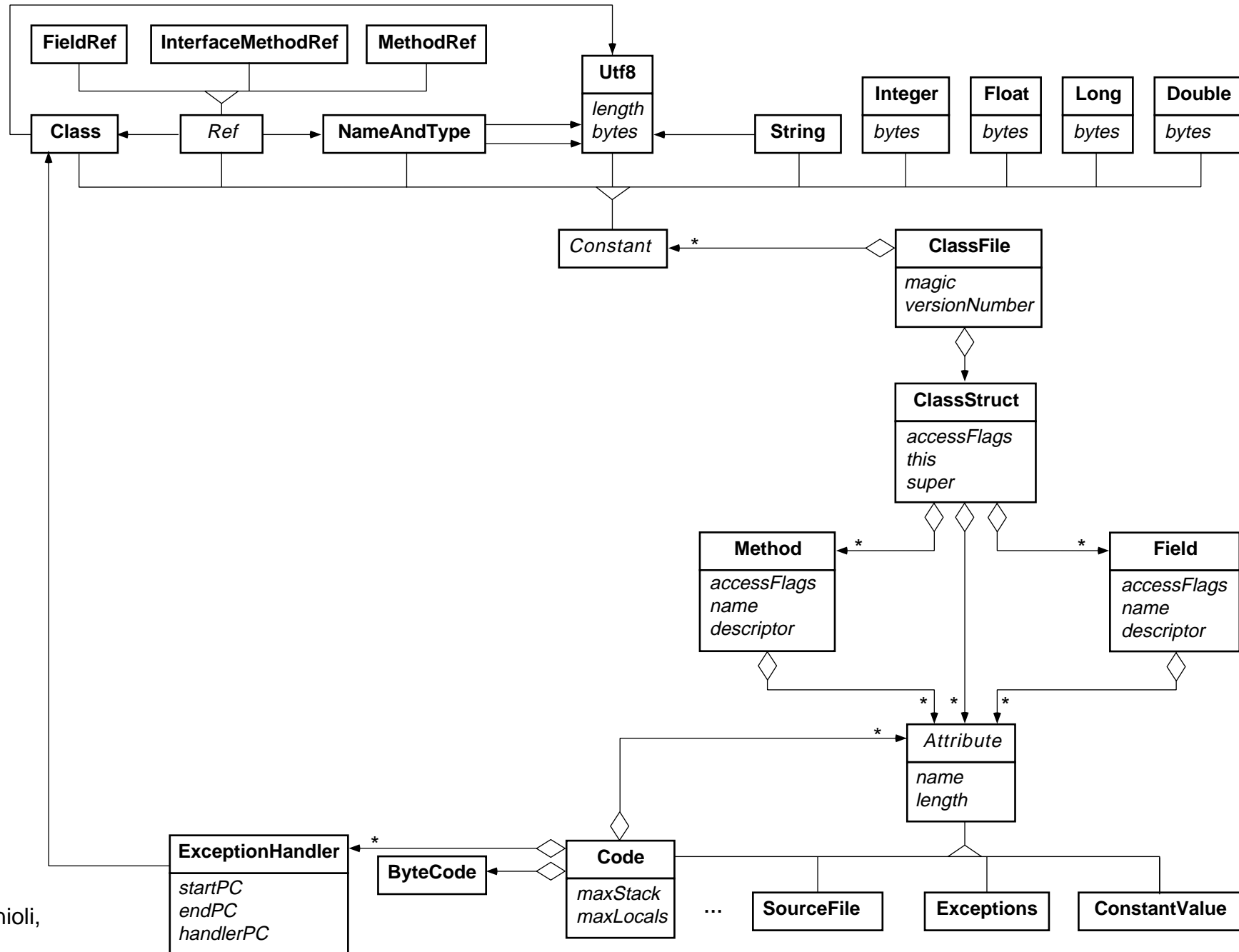
Fortunately, this format is *not* mandatory.

CUMULATIVE FREQUENCIES OF THE SIZE

Average size: 4'541 \pm 13'017 bytes, Median: 1'746 bytes.

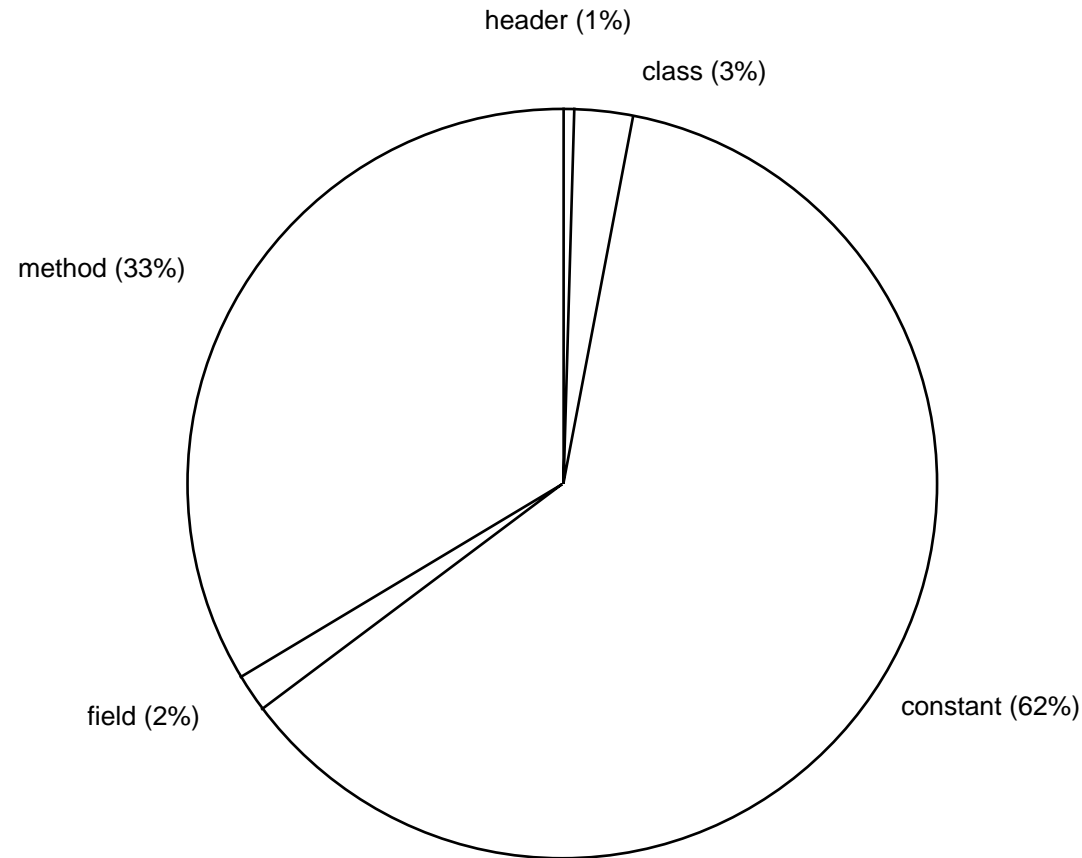


FORMAT OF THE CLASS FILE



ConstantPool: 62%

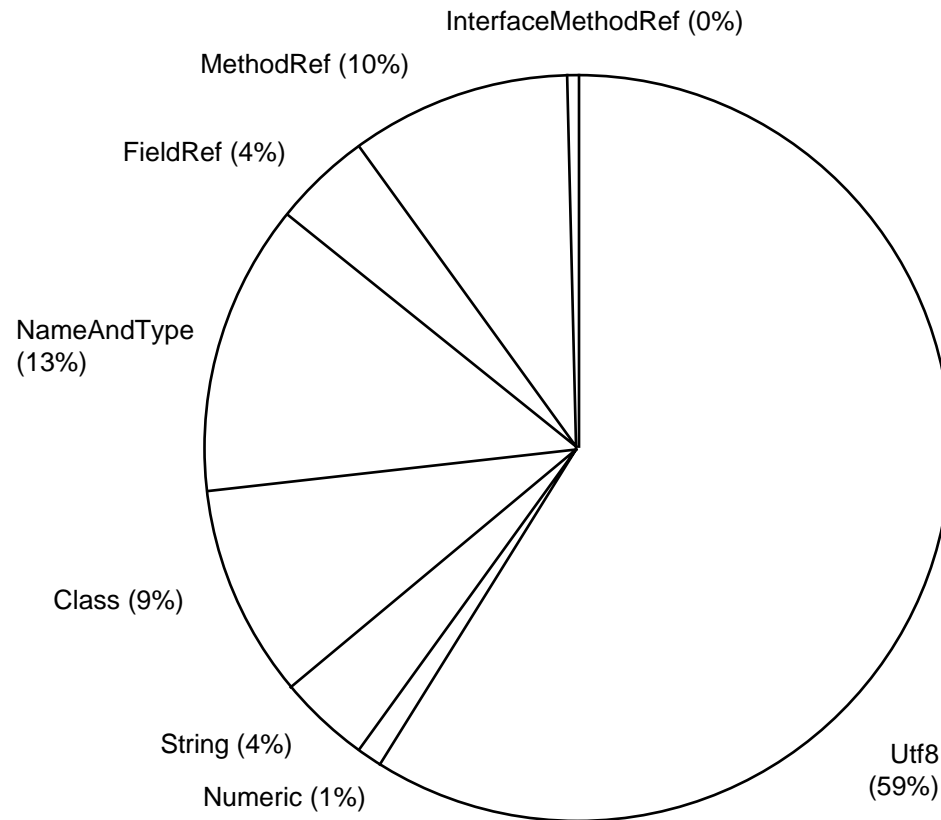
Methods: 33%



Type & Link: 36%

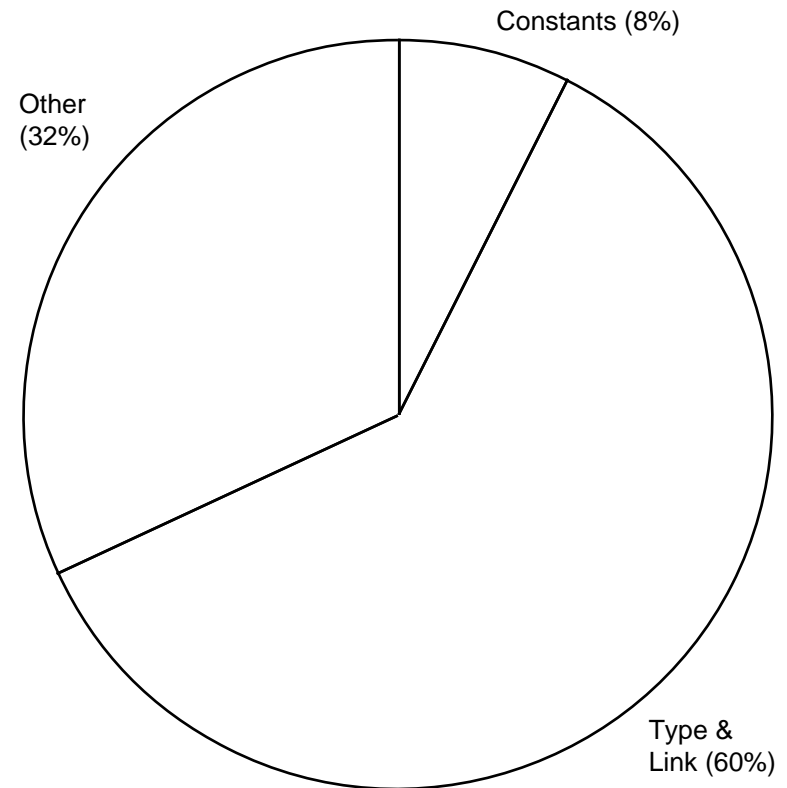
Constant: 5%

Utf8: 59%



Type & Link: 60%

Other: 32%



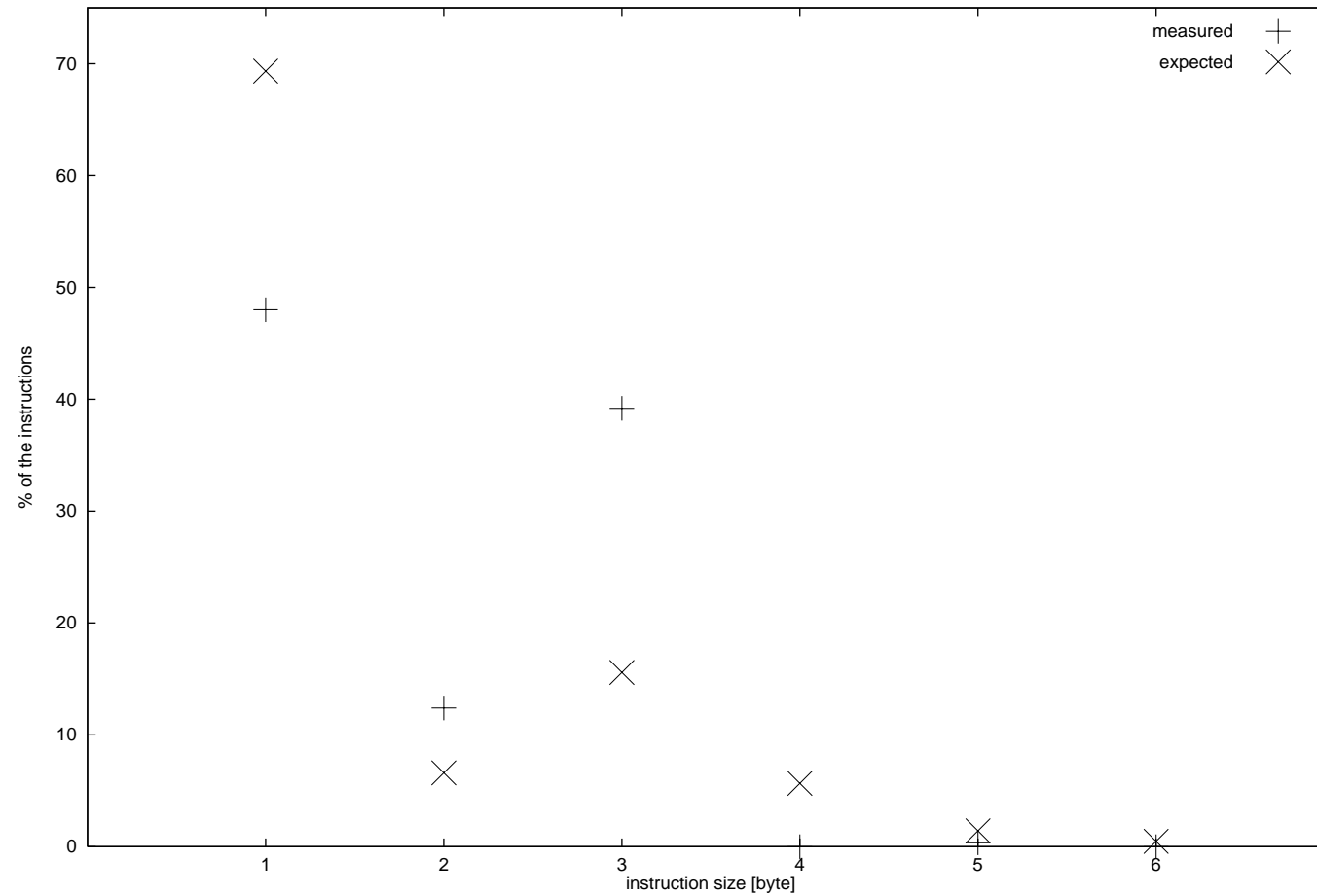
The content of the class file is customizable through its *attribute* and *constant pools*.

“Of the predefined attributes, the `Code`, `ConstantValue`, and `Exceptions` attributes must be recognized and correctly read by a class file reader for correct interpretation of the class file by a Java Virtual Machine” [Lind96].

On the average, 33% of the class file is *dispensable*.

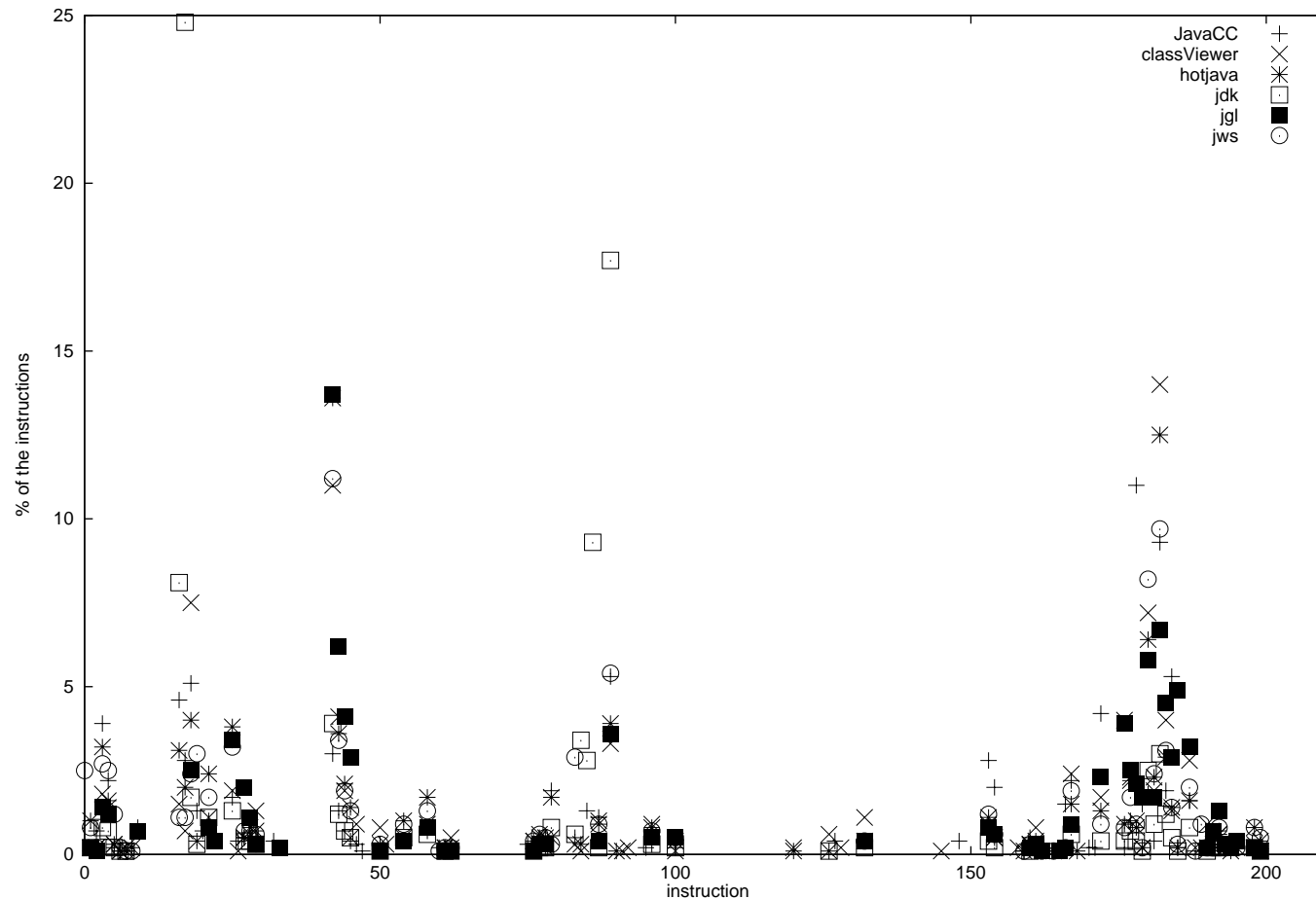
ifi DISTRIBUTION OF THE INSTRUCTIONS

Average instruction size: 1.9 byte



ifi UNEQUAL USE OF THE INSTRUCTIONS

5 instructions account for more than 5% of the occurrences in at least 2 programs.



For n instructions, with measured frequencies

$$p_i = \frac{c_i}{\sum c}$$

the entropy H is:

$$Hs = -\sum p_i \times \log p_i$$

The average entropy is 3.46.

- 50% of the files are smaller than 1'746 bytes, 80% of the files are smaller than 6'000 bytes.
- The *constant pool* makes 62% of the class file, of which the *Linking information* represents 60%.
- 33% of the class file is dispensable.
- The average size of an instruction is 1.9 bytes.
- Classes use 25 instructions, and 5 instructions are preferred.
- An optimal encoding needs just 4 bits.

It is a capital mistake to theorize before one has data.

C. Doyle, Scandal in Bohemia

Program	Company	Version	Size [classes]
Java Developer's Kit (JDK)	JavaSoft	1.1.5	1622
HotJava	Sun	1.1	539
Java Workshop (JWS)	Sun	2.0	1408
JavaCC	SunTest	0.7.1	134
Java Generic Library (JGL)	ObjectSpace	3.0	262
classViewer	ifi		51
			4016