Compiler to JVM

Jiawei Chen, Zhetuo Qi

Zhejiang University

May 29, 2019

Language Specification

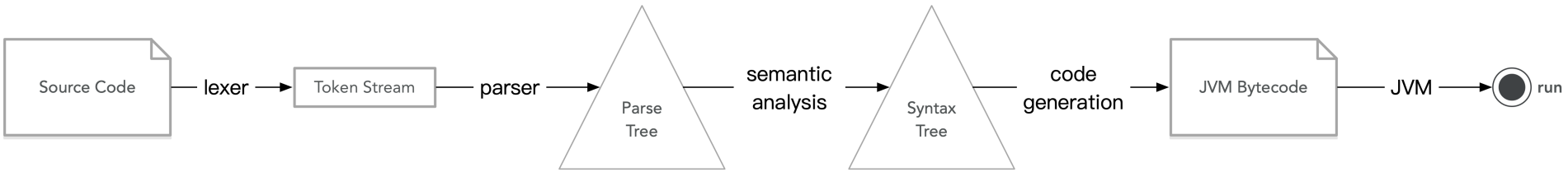
```
namespace Test {
    struct MyStruct {
        int a = f();
    }

    double d = 4;

    def void main(String[] args) {
        print("Hello world!\n");
    }

    def int f() {
        return 3;
    }
}
```

Compiler Architecture



Third Party Library

haha

Type I and Type II Errors

False Negative Errors

A false negative is an error in which a test/study result is negative but in reality it is positive.

True Positive Rate, TPR

$$TPR = \frac{TP}{TP + FN} \tag{1}$$

False Positive Rate, FPR

$$FPR = \frac{FP}{FP + TN} \tag{2}$$

ROC Space

True Positive Rate, TPR

$$TPR = \frac{TP}{TP + FN} \tag{3}$$

False Positive Rate, FPR

$$FPR = \frac{FP}{FP + TN} \tag{4}$$

ROC Space

An ROC space is defined by **FPR** and **TPR** as x and y axes, respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs)

- Each prediction result represents one point in the ROC space.

- What if a result point locates in the diagonal line?

Accuracy

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{5}$$

- Suppose the number of positive results equals to the number of negative results in reality.

$$TP + FN = FP + TN \tag{6}$$

- Then a result point located in the diagonal line means:

$$\begin{aligned} \frac{TP}{TP + FN} &= \frac{FP}{FP + TN} \\ \Rightarrow Accuracy &= \frac{TP + TN}{TP + FP + FN + TN} = 50\% \end{aligned} \tag{7}$$

- So the diagonal line are called line of no-discrimination (line of random guess)

- The result located in the diagonal line has 50% accuracy

- The closer a result is to the upper left corner, the better it predicts.
- The closer a result is to the lower right corner, the worse it predicts.

ROC Curves

ROC Curve

Fix binary classifier model and change the threshold, the result points will generate the ROC curve.

Example

AUC

AUC

Area under the curve of ROC curve.

Example

- $AUC \leqslant 1$

Example

- $AUC = 1$, perfect classifier
- $0.5 < AUC < 1$, better than random guess
- $AUC = 0.5$, same as random guess
- $AUC < 0.5$, worse than random guess

Balance of FPR and FNR

- With the decrease of FPR, FNR increases.

Example

Balancing according to concrete situations

- **Prediction of earthquakes:** we prefer to predict every earthquake successfully, which means that we can tolerate the increase of FPR compared to FNR.

Balancing according to concrete situations

- **Conviction:** By the presumption of innocence, we prefer to be cautious when involving criminal justice, so we have to tolerate larger FNR compared to FPR.