

ĐẠI HỌC QUỐC GIA - THÀNH PHỐ HỒ CHÍ MINH

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÀNH PHỐ HỒ CHÍ MINH

KHOA KHOA HỌC VÀ KỸ THUẬT MÁY TÍNH



NGUYÊN TẮC LẬP TRÌNH - C03005

---

# BÀI VIẾT 1

## Nhận dạng

---

THÀNH PHỐ HỒ CHÍ MINH, 12/2023



## BÀI VIẾT 1

### Nhận dạng

Sau khi hoàn thành nhiệm vụ này, bạn sẽ có thể

- xác định từ vựng chính thức của ngôn ngữ lập trình.
- sử dụng ANTLR để triển khai từ vựng cho ngôn ngữ lập trình.
- xác định ngữ pháp chính thức của một ngôn ngữ lập trình.
- sử dụng ANTLR để triển khai trình nhận dạng ngôn ngữ lập trình.

## 1 Đặc điểm kỹ thuật

Trong bài tập này, bạn được yêu cầu viết một từ vựng và một bộ nhận dạng cho một chương trình được viết bằng ZCode. Để hoàn thành nhiệm vụ này, bạn cần:

- Cài đặt Python 3 nếu bạn chưa cài đặt.
- Tải mã ban đầu của Bài tập 1 (zcode-initial.zip) và giải nén.
- Tải xuống antlr-4.9.2-complete.jar từ <https://www.antlr.org/download.html> (phần Phiên bản trước), đặt biến môi trường ANTLR\_JAR cho tệp này; cài đặt antlr4-python3-runtime phiên bản 4.9.2 (xem hướng dẫn tại phần Python Targets của trang web trên hoặc gõ `pip3 install antlr4-python3-runtime==4.9.2`).
- Thay đổi thư mục ban đầu thành bài tập 1

Để hoàn thành nhiệm vụ này, bạn cần:

- đọc kỹ các đặc điểm kỹ thuật của ngôn ngữ.
- sửa đổi ZCode.g4 trong mã ban đầu để mô tả ngôn ngữ ZCode chính thức. Vui lòng điền id của bạn vào tiêu đề của tệp tin này.

Nhiệm vụ này được chia thành hai giai đoạn: giai đoạn từ vựng và giai đoạn nhận dạng. Các giai đoạn này được đánh giá độc lập.



## 1.1 Giai đoạn 1: Lexer

Trong giai đoạn này, bạn được yêu cầu viết từ vựng cho chương trình viết bằng ANTLR. Để hoàn thành giai đoạn này, bạn cần:

- Sửa đổi `ZCode.g4` để phát hiện mã thông báo bằng ngôn ngữ `ZCode`.
- Tạo 100 trường hợp thử nghiệm cho `LexerSuite` để kiểm tra mã của bạn.
- Đối với các lỗi từ vựng, vui lòng trả lại các mã thông báo sau cùng với các từ vựng cụ thể:
  - `ERROR_TOKEN` với từ vựng `<unrecognized char>`: khi lexer phát hiện một ký tự không được nhận dạng.
  - `UNCLOSE_STRING` với từ vựng `<unclosed string>`: khi lexer phát hiện một chuỗi không kết thúc. Từ vựng `<unclosed string>` không bao gồm dấu ngoặc kép mở đầu.
  - `ILLEGAL_ESCAPE` với từ vựng `<wrong string>`: khi lexer phát hiện một lỗi thoát bất hợp pháp trong chuỗi. Sai chuỗi từ đầu chuỗi (không có dấu ngoặc mở đầu) đến chỗ thoát trái phép.
- Bạn có thể giả sử rằng chỉ có một lỗi trong mỗi ca kiểm thử.

## 1.2 Giai đoạn 2: Bộ nhận dạng

Trong giai đoạn này, bạn được yêu cầu viết trình nhận dạng cho chương trình được viết bằng `ZCode`. Để hoàn thành giai đoạn này, bạn cần:

- Sửa đổi `ZCode.g4`.
- Tạo 100 trường hợp thử nghiệm cho `ParserSuite` để kiểm tra mã của bạn.
- Bạn có thể giả định rằng có nhiều nhất một lỗi trong mỗi ca kiểm thử.

## 2 Yêu cầu

Chạy `python run.py gen` để tạo các tệp sau trong thư mục đích hoặc thư mục con của nó. Gửi các tệp này cùng với `ZCode.g4`, `LexerSuite.py` và `ParserSuite.py` (10 tệp).

`ZCode.interp`

`ZCodeVisitor.py`

`ZCodeParser.py`

`ZCodeLexer.tokens`

`ZCodeLexer.py`

`ZCodeLexer.interp``ZCode.tokens`

KHÔNG NÉN CHÚNG khi gửi. Sau đó bấm vào nút Kiểm tra để kiểm tra bài gửi của bạn. Có một số trường hợp thử nghiệm đơn giản để kiểm tra nội dung gửi của bạn và sẽ được đánh dấu sau. Học sinh có thể nộp bài bao nhiêu lần tùy thích nhưng chỉ bài nộp cuối cùng mới được chấm điểm. Vì hệ thống không thể chịu tải khi có quá nhiều học sinh nộp bài cùng một lúc nên học sinh nên nộp bài càng sớm càng tốt. Sinh viên sẽ chịu mọi trách nhiệm về rủi ro của mình nếu nộp bài gần thời hạn. Khi hết thời hạn nộp bài, hệ thống sẽ đóng nên học viên không thể nộp bài được nữa. Các phương pháp gửi khác sẽ không được chấp nhận.

Thời hạn của cả hai giai đoạn của bài tập 1 được thông báo trên trang web của lớp. Bạn phải tự mình hoàn thành bài tập và không được để người khác xem bài làm của mình, nếu không, bạn sẽ bị quy định của trường đại học trừng phạt vì tội đạo văn.

### 3 Các quy định khác

- Học sinh phải tự mình hoàn thành bài tập này và phải ngăn chặn người khác lấy cắp kết quả của mình. Nếu không, học sinh coi là gian lận theo quy định của nhà trường về gian lận.
- Mọi quyết định của giáo viên phụ trách nhiệm vụ này là quyết định cuối cùng.
- Học sinh sẽ không được phát testcase sau khi chấm điểm mà chỉ được cung cấp thông tin về chiến lược thiết kế testcase và phân bổ số lượng học sinh chính xác theo từng testcase.

### 4 Nhật ký thay đổi

- Thêm (phần Phiên bản trước)
- Thêm pip3 cài đặt antlr4-python3-runtime==4.9.2