# IT3322E Xây dựng chương trình dịch

# Tên môn tiếng Anh – COMPILER CONSTRUCTION

Version: 2019.05.30

1. **THÔNG TIN CHUNG**

**GENERAL INFORMATION**

|  |  |
| --- | --- |
| **Tên học phần**  **Course name:** | *Compiler construction* |
| **Mã học phần**  **Code:** | IT3322E  IT3322E |
| **Khối lượng**  **Credit:** | 2(2-1-0-6)   * Lý thuyết - Lecture: 30 hours * Bài tập - Exercise: 30 hours   (If capstone project is used, please indicate clearly)   * Thí nghiệm - Experiments: 0 hours |
| **Học phần tiên quyết**  **Prerequisite:** |  |
| **Học phần học trước**  **Prior course:** | * IT3312E: Data structure and algỏithm   IT3320E: Data structure and algỏithm lab |
| **Học phần song hành**  **Paralell course:** | No |

1. **MÔ TẢ HỌC PHẦN - COURSE DESCRIPTION**

Môn học này nhằm cung cấp cho sinh viên kiến thức cơ bản về hoạt động của một trình biên dịch của ngôn ngữ lập trình. Ngoài ra sinh viên cũng có được những kiến thức cơ bản về lý thuyết ngôn ngữ, đặc biệt là cách mô tả cú pháp và ngữ nghĩa của ngôn ngữ. từ đó sinh viên hiểu được nguyên lý làm việc của ngôn ngữ lập trình. Với những hiểu biết về lĩnh vực này này, sinh viên sẽ viết chương trình một cách hiệu quả hơn, dễ dàng học tập các ngôn ngữ mới. Sinh viên cũng hiểu được các mô hình xử lý ngôn ngữ để nghiên cứu sang các lĩnh vực khác như xử lý ngôn ngữ tự nhiên, sinh tin, nhận dạng cấu trúc...

Bài tập của môn học yêu cầu sinh viên hoàn thiện một chương trình dịch cho một ngôn ngữ lập trình đơn giản, từ đó cho phép sinh viên làm việc với các cấu trúc dữ liệu, các phương pháp xử lý và các giải thuật quan trọng, liên kết chúng để tạo ra một phần mềm hoàn chỉnh, từ đó rèn luyện kỹ năng làm việc trong những dự án lớn sau này.

This course is intended to provide a basic understanding of how a compiler of a programming language works. In addition, after learning the course, students also have basic knowledge of language theory, especially the method for syntax and semantics description. From there, students understand the principles of programming languages. With knowledge of this field, students will write programs in a more efficient way, and make it easier to learn new languages. Students also understand models language processing to apply to many other fields such as natural language processing, bioinformatics, structural recognition ...

The assignments require students to complete a compiler for a simple programming language, thereby allowing students to work with important data structures and algorithms, to link them together to create a complete software, thereby training skills to work in large projects in the future.

**MỤC TIÊU VÀ CHUẨN ĐẦU RA CỦA HỌC PHẦN - GOAL AND OUTPUT REQUIREMENT**

Sinh viên hoàn thành học phần này có khả năng

After this course the student will obtain the followings:

| **Mục tiêu/CĐR**  **Goal** | **Mô tả mục tiêu/Chuẩn đầu ra của học phần**  **Description of the goal or output requirement** | **CĐR được phân bổ cho HP/ Mức độ (I/T/U)**  **Output division/ Level (I/T/U)** |
| --- | --- | --- |
| **[1]** | **[2]** | **[3]** |
| **M1** | **Hiểu các đặc trưng của ngôn ngữ lập trình và có khả năng sử dụng các ngôn ngữ lập trình mới**  **Understand features of programming languages and be able to use new programming languages** | 1.1.1; 1.2.5; 1.1.4; 1.5.2; 2.3.3; 3.1.4 |
| M1.1 | Nhận diện và hiểu rõ các thành phần của chương trình dịch  Identify and understand components of a compiler | [1.1.1; 1.2.5; 1.1.4] (I) |
| M1.2 | Nhận diện, so sánh và phân loại được các dạng khác nhau của ngôn ngữ lập trình và chương trình dịch  Identify, compare and categorize different types of programming languages ​​and translation programs | [1.1.1; 1.2.5; 1.1.4] (T) |
| M1.3 | Có khả năng thiết một ngôn ngữ về cú pháp và ngữ nghĩa, viết chương trình xử lý ngôn ngữ đơn giản  Be able to design a language in syntax and semantics, and write a simple language processor | [1.5.2; 2.3.3; 3.1.4] (TU) |
| **M2** | **Nhận diện và làm chủ được các cơ hội trên thị trường do công nghệ thông tin đem lại để phát triển tổ chức sẵn có và tạo ra các tổ chức mới**  **Identify and master opportunitiesin of the market brought by information technology to develop existing organizations and create new ones** | 1.1.4; 1.3.3; 1.1.4; 1.3.3; 1.3.4; 1.6.3; |
| M2.1 | Hiểu và vận dụng được các ứng dụng công nghệ thông tin đương đại nhằm hỗ trợ các hoạt động trong tổ chức  Understand and use current applications of information technology to support activities in the organization | [1.1.4; 1.3.3; 1.6.3] (T) |
| M2.2 | Nhận diện được các tác động của công nghệ thông tin đối với tổ chức và môi trường hoạt động của tổ chức  Identify the impacts of information technology on the organization and its operating environment | [1.3.4] (U) |
| **M3** | **Nhận diện các xu hướng phát triển của các kỹ thuật xử lý ngôn ngữ.**  **Identify development trends of language processing techniques** | 4.1.1; 4.1.2; 4.1.3 |
| M3.1 | Chủ động tìm hiểu các ngôn ngữ lập trình mới  Learn new programming languages proactively | [4.1.1; 4.1.2] (T) |
| M3.2 | Xác định được các cơ hội mà sự phát triển của công cụ mới đem lại để phát triển tổ chức sẵn có  Identify the opportunities the development of new tools brings to development of the organization. | [4.1.3] (U) |

1. **TÀI LIỆU HỌC TẬP**

**Reference**

**Reference book**

|  |  |
| --- | --- |
| [1] | Aho.A.V, Sethi.R., Lam M., Ullman.J.D.  *Compiler : Principles, Techniques and Tools.*  Addison Wesley. 2007. |
| [2] | Andrew.W.Appel  Modern Compiler Implementation in Java (C)  *Princeton University. 2002* |

**Online open courseware**

|  |
| --- |
| [1] |
| [2] |
| [3] |
| [4] |

1. **CÁCH ĐÁNH GIÁ HỌC PHẦN - EVALUATION**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Điểm thành phần**  **Module** | **Phương pháp đánh giá cụ thể**  **Evaluation method** | **Mô tả**  **Detail** | **CĐR được đánh giá**  **Output** | **Tỷ trọng**  **Percent** |
| **[1]** | **[2]** | **[3]** | **[4]** | **[5]** |
| **A1. Điểm quá trình**  **Mid-term (\*)** | **Đánh giá quá trình Progress** |  |  | **50%** |
| A1.1. Bài tập về nhà  Homework | Tự luận  Written |  | 20% |
| A1.2. Bài tập lớn  Capstone project | Lập trình  Programming |  | 30% |
| **A2. Điểm cuối kỳ**  **Final term** | **A2.1. Thi cuối kỳ**  **Final exam** | Thi viết  Written exam |  | **50%** |

*\* Điểm quá trình sẽ được điều chỉnh bằng cách cộng thêm điểm chuyên cần. Điểm chuyên cần có giá trị từ –2 đến +1, theo Quy chế Đào tạo đại học hệ chính quy của Trường ĐH Bách khoa Hà Nội.*

*The evaluation about the progress can be adjusted with some bonus. The bonus should belong to [-2, +1], according to the policy of Hanoi University of Science and Technology.*

1. **KẾ HOẠCH GIẢNG DẠY – SCHEDULE**

| **Tuần**  **Week** | **Nội dung**  **Content** | **CĐR học phần**  **Output** | **Hoạt động dạy và học**  **Activities** | **Bài đánh giá**  **Evaluation** |
| --- | --- | --- | --- | --- |
| **[1]** | **[2]** | **[3]** | **[4]** | **[5]** |
| 1 | **Bài 1: Bộ xử lý ngôn ngữ**  1.1 Đặc trưng của ngôn ngữ lập trình  1.2 Bộ xử lý ngôn ngứ  1.3 Compiler và Interpreter  1.4 Tính khả chuyển của trình biên dịch  1.5. Các công cụ liên quan đến trình biên dịch.  **Unit 1: Language processors**  1.1 Features of programming languages  1.2 Languages processors  1.3 Compiler vs Interpreter   * 1. Portability of compilers   1.5. Cousins of compiler | M1.1  M2.2  M3.2 | Giảng bài  Lecturing | A2.1 |
| 2 | **Bài 2: Cấu trúc của chương trình dịch**  2.1. Các giai đoạn của trình biên dịch  2.2. Quá trinh dịch một câu lệnh  2.3. Giới thiệu các pha trong trình biên dịch  **Unit 2: Structure of a typical compiler**  2.1. Phases of a compiler  2.2. Translation of a statement  2.3. Phases of compiler | M1.1 | Đọc trước tài liệu;  Giảng bài  Notes reading  Lecturing | A2.1 |
| 3 | **Bài 3 : Văn phạm sản sinh**  3.1. Kiến thức cơ bản về lý thuyết ngôn ngữ: Ngôn ngữ và xâu, phân cấp Chomsky  3.2. Văn phạm phi ngữ cảnh  3.3. Suy dẫn  3.4. Cây phân tích cú pháp  3.5. Vấn đề nhập nhằng và khử nhập nhằng  **Unit 3. Generative Grammas**  3.1. Fundamentals of languages theory: strings and languages, Chomsky hierarchy.  3.2. Context free grammar  3.3. Derivation  3.4.Ambiguity. Disambiguation | M1.1 | Đọc trước tài liệu;  Giảng bài;  Notes reading  Lecturing | A2.1 |
| 4 | **Bài 4: BNF và sơ đồ cú pháp**  4.1. Ký pháp BNF  4.2. Ký pháp EBNF  4.3. Văn phạm KPL biểu diễn bằng BNF  4.4. Sơ đồ cú pháp  4.5. Văn phạm KPL biểu diễn bằng sơ đồ cũ pháp  **Unit 4. BNF and syntax diagram**  4.1. BNF notation  4.2. EBNF notation  4.3. KPL grammar in BNF  4.5. Syntax diagrams | M1.2; M1.3; M2.1; M2.2 | Đọc trước tài liệu;  Giảng bài;  Bài tập về nhà  Notes reading  Lecturing  Home works | A1.1  A2.1 |
| 5 | **Bài 5: Phân tích từ vựng**  5.1. Nhiệm vụ của bộ phân tích từ vựng  5.2. Ôtômat hữu hạn làm nhiệm vụ phân tích từ vựng  5.3. Cài đặt bộ phân tích từ vựng  *Thực tập với máy tính*  **Unit 5. Lexical analysis**   * 1. Tasks of lexical analyzer (scanner)   2. Finite automata and lexical analysis   3. Scanner implementation | M1.2; M1.3; M2.1; M2.2 | Đọc trước tài liệu;  Giảng bài  Notes reading  Lecturing | A2.1 |
| 6 | **Bài 6: Phân tích cú pháp trên xuống có quay lui**  6.1. Bài toán phân tích cú pháp  6.2. Phân tích trái và phân tích phải  6.3. Mô hình ôtômât đẩy xuống  6.4. Mô tả giải thuật phân tích cú pháp quay lui  6.5. Cài đặt giải thuật phân tích cú pháp quay lui  **Unit 6: Topdown parser with backtracking**  6.1. Parsing problem  6.2. Left parse and right parse  6.3. Model of Push down automata  6.4. Parsing algorithm with backtracking  6.5. Implementation of parsing algorithm | M1.2; M1.3 | Đọc trước tài liệu;  Giảng bài  Notes reading  Lecturing  Homework | A1.1  A1.2  A1.1  A2.1 |
| 7 | **Bài 7: Phân tích cú pháp trên xuống tiền định**  7.1. Vấn đề tiền định trong phân tích cú pháp trên xuống  7.2. Điều kiện để phân tích cú pháp trên xuống tiền định  7.3. Giải thuật phân tích cú pháp tiền định   * Xây dựng bảng phân tích cú pháp * Thực thi giải thuật   **Unit 7: Predictive top down parsing**  7.1. How to avoid bactraking?  7.2. Conditions for predictive parsing  7.3. Algorithm of predictive parsing   * Parsing table * Parsing algorithm | M1.2; M1.3 | Giảng bài  Lecturing | A1.1  A2.1 |
| 8 | **Bài 8: Văn phạm LL(k)**  8.1. Tương quan giữa các lớp con của lớp ngôn ngữ phi ngữ cảnh  8.2. Khái niệm FIRST, FOLLOW  8.3. Văn phạm LL(k)  8.4. Kiểm tra điều kiện LL(1) qua văn phạm và sơ đồ cú pháp  **Unit 8. LL(k) grammars**  8.1. LL(k) grammar vs classes of Context Free Grammars  8.2. FIRST and FOLLOW  8.3. LL(k) grammars  8.4. LL(1) condition for BNF and syntax diagrams | M1.2; M2.2 | Đọc trước tài liệu;  Giảng bài;  Bài tập tại phòng lab: Hoàn thiện bộ phân tích từ vựng của ngôn ngữ KPL  Notes reading  Lecturing  Labwork: Complete the scanner of KPL | A1.2  A2.1 |
| 9 | **Bài 9. Phương pháp đệ quy trên xuống**  9.1. Giải thuật phân tích cú pháp đệ quy trên xuống  9.2. Triển khai thủ tục xử lý trong bộ phân tích cú pháp đệ quy trên xuống  9.3. Cài đặt bộ phân tích cú pháp cho ngôn ngữ KPL  **Unit 9. Recursive descent parsing**  9.1. Recursive descent parsing algorithm  9.2. Implementation of the algorithm  9.3. Recursive descent parser for KPL | M1.2; M2.2 | Đọc trước tài liệu;  Giảng bài  Bài tập tại phòng lab: Hoàn thiện bộ phân tích cú pháp của ngôn ngữ KPL: phàn khai báo  Notes reading  Lecturing  Labwork: Complete the first part of KPL parser: declarations | A1.2  A2.1 |
| 10 | **Bài 10 Phân tích ngữ nghĩa**  10.1. Những vấn đề ngữ nghĩa trong ngôn ngữ lập trình  10.2. Bảng ký hiệu  10.3. Quản lý phạm vi  **Unit 10: Semantic analysis**  10.1. Semantics of programming languages  10.2. Symbol table  10.3. Scope management | M1.2  M2.1 | Đọc trước tài liệu;  Giảng bài;  Bài tập tại phòng lab: Hoàn thiện toàn bộ bộ phân tích cú pháp của ngôn ngữ KPL  Notes reading  Lecturing  Labwork: Complete the whole KPL parser. | A1.2  A2.1 |
| 11 | **Bài 10 Phân tích ngữ nghĩa (tiếp theo)**  10.4. Định nghĩa tựa cú pháp  10.5. Hệ thống kiếu  10.6. Tương ứng kiểu trong các lệnh cơ bản  **Unit 10: Semantic analysis (continued)**  10.4. Syntax Directed Definition  10.5. Type system  10.5. Type checking of statements | M1.2  M2.1 | Giảng bài;  Bài tập tại phòng lab: Hoàn thiện bộ phân tích ngữ nghĩa của ngôn ngữ KPL: Bảng lý hiệu và quản lý phạm vi.  Notes reading  Lecturing  Labwork: Complete the first part of KPL semantic analyser: Symbol table and scope management | A1.2  A2.1 |
| 12 | **Bài 11: Sinh mã trung gian**  11.1. Mã ba địa chỉ  11.2. Sinh mã cho lệnh gán  11.3. Sinh mã cho các biểu thức logic  11.4. Sinh mã cho các cấu trúc lập trình  **Unit 11: Intermediate code generation**  11.1. Three address code  11.2. Three addresss code for assignments  11.3.Three address code for logic expressions  11.4. Three address code for flow-of-control statements | M1.2  M2.1 | Giảng bài;  Bài tập tại phòng lab: Hoàn thiện bộ phân tích ngữ nghĩa của ngôn ngữ KPL: Kiểm tra kiểu.  Lecturing  Labwork: Complete KPL semantic analyser: Type checking | A1.2  A2.1 |
| 13 | **Bài 12: Tối ưu mã**  12.1.Tối ưu cục bộ  12.2. Tối ưu trong khối cơ bản  12.3. Tồi ưu trên đồ thị  **Unit 12: Code optimization**  12.1. Peephole optimization  12.2. Optimization for basic blocks  12.3. Optimization on control flow graphs | M3.1; M3.2 | Đọc trước tài liệu;  Giảng bài;  Bài tập tại phòng lab: Hoàn thiện bộ sinh mã đích cho ngôn ngữ KPL phần phân phối bộ nhớ trên stack  Notes reading  Lecturing  Labwork: Complete target code generator of KPL: Memory allocation on stack | A1.2  A2.1 |
| 14 | **Bài 13: Sinh mã đích**  13.1. Máy đích: Stack Calculator  13.2. Tập lệnh của ngôn ngữ đích  13.3. Sinh mã đích cho các cấu trúc cơ bản.  **Unit 13: Target code generation**  13.1. Target machine: Stack Calculator  13.2. Instruction set  13.3. Code generation for basic structures | M3.1; M3.2 | Đọc trước tài liệu;  Giảng bài;  Bài tập tại phòng lab: Hoàn thiện bộ sinh mã đích cho ngôn ngữ KPL phần phân phối bộ nhớ trên stack  Notes reading  Lecturing  Labwork: Complete the whole target code generator of KPL. | A1.2  A2.1 |
| 15 | ***Tổng kết và ôn tập***  ***Summary*** |  | Trao đổi  Đánh giá toàn bộ project compiler  Discussion  Evaluate the whole KPL compiler project |  |

1. **QUY ĐỊNH CỦA HỌC PHẦN - COURSE REQUIREMENT**

(The specific requirements if any)

1. **NGÀY PHÊ DUYỆT - DATE: …………………..**

|  |  |
| --- | --- |
| **Chủ tịch hội đồng**  **Committee chair** | **Nhóm xây dựng đề cương**  **Course preparation group** |

1. **QUÁ TRÌNH CẬP NHẬT - UPDATE INFORMATION**

| **STT**  **No** | **Nội dung điều chỉnh**  **Content of the update** | **Ngày tháng được phê duyệt**  **Date accepted** | **Áp dụng từ kỳ/ khóa**  **A**  **pplicable from** | **Ghi chú**  **Note** |
| --- | --- | --- | --- | --- |
| 1 | …………… |  |  |  |
| 2 | …………… |  |  |  |