

Course Syllabus
CS 410
 Compiler Construction
 fall semester
 August 24, 2019

| | | | | | | | |
|-----------|---------------------------|-------------------|--------------|----------------------------------|----------------------------------|------------------------|-------------|
| class | 3 | | 2 | 1 | | 0 | |
| language | regular expression | linear | context free | context sensitive | recursive, decidable | recursively enumerable | undecidable |
| grammar | right linear, left linear | linear | context free | context sensitive | (none) | phrase structure | (none)- |
| automaton | finite state | one-turn pushdown | pushdown | 2 stack pushdown, linear bounded | Turing machine that always halts | Turing machine | (none) |

Course Introduction

Credit Hours: 3

Prerequisite Courses: CS 220 and 310

Method of Instruction: lecture

Instructor: Frances Van Scoy, Ph.D., associate professor of computer science

Class Meets: 2:00-3:15. spring semester, 2019

Class Meets in 135 Advanced Engineering Research

Course Description:

"Compiler construction" is one of the earliest success stories for computer science. That is, there was a real problem to solve--how to translate notation easily written and understood by humans into strings of binary digits understood by a computer--for which a rather elegant body of theory provided much of the solution.

There are two organizing themes for this course.

the phases of a compiler:

scanner, parser, symbol table manager, intermediate code generator, code generator

the Chomsky hierarchy:

regular expressions, context free, context sensitive, phrase structure

This course is a mix of formal language theory, construction of a software system using lex and yacc, pragmatic programming details, comparative programming languages, and natural language

processing.

| class | 3 | | 2 | 1 | | 0 | |
|-----------|---------------------------|-------------------|--------------|----------------------------------|----------------------------------|------------------------|-------------|
| language | regular expression | linear | context free | context sensitive | recursive, decidable | recursively enumerable | undecidable |
| grammar | right linear, left linear | linear | context free | context sensitive | (none) | phrase structure | (none)- |
| automaton | finite state | one-turn pushdown | pushdown | 2 stack pushdown, linear bounded | Turing machine that always halts | Turing machine | (none) |

There is a similar table at https://en.wikipedia.org/wiki/Recursively_enumerable_language

Learner Support

Instructor Office Location: 343 Advanced Engineering Research Building

Office Hours: Tuesdays and Thursdays 9:30-10:45, ESB ground floor hallway (starting in September); 12:30-1:45 AER atrium

Instructor Email: frvanscoy@mail.wvu.edu

Use frvanscoy@mix.wvu.edu *only* for submitting course assignments.

Otherwise (including to ask questions about course assignments) use frvnscoy@mail.wvu.edu

Phone: 304.293.0960

Method of Making Appointment:

email to frvanscoy@mail.wvu.edu;

phone call to 304.293.0960 (leave voice mail if no answer; will be sent to me as email attachment)

ITS contact: ?

Other: ?

Instructional Materials

Required Instructional Materials:

Alfred V. Aho, Monica S. Lam, Ravi Sethi, Jeffrey D. Ullman
Compilers Principles, Techniques, and Tools, 2nd edition
Addison-Wesley, 2007 (other editions are probably acceptable)

Doug Brown, *John Levine, Tony Mason*

lex & yacc, second edition

O'Reilly Media, 1992

available for on-line reading via WVU Libraries' subscription to Safari Books Online

Smalltalk-80: The Language and its Implementation

by Adele Goldberg and David Robson;

Xerox Palo Alto Research Center ISBN 0-201-11371-6. 344 pp. 1983

available for legal free download (32 MB) at

<http://stephane.ducasse.free.fr/FreeBooks/index.html>

Steven Bird, Ewan Klein, and Edward Loper

Natural Language Processing with Python

O'Reilly Media, 2016

available for free download at

<http://www.nltk.org/book/>

Some software packages to install on your computer are:

| | |
|----------------------------------|---|
| Python 3 | https://www.python.org/ |
| NLTK (Natural Language Tool Kit) | http://www.nltk.org/ |
| lex or flex or similar | http://flex.sourceforge.net/ |
| yacc or bison or similar | https://www.gnu.org/software/bison/ |

Optional Instructional Materials:

Inform 7 implementation and documenttion found at www.inform7.com

Course Learning Objectives

1. Understand Chomsky hierarchy and its relationship to compiler construction
2. Construct compiler components
3. Apply compiler theory, techniques, tools to problems in natural language processing

Course Activities

Major Learning Activities:

12 homework assignments at the New Gradience site (24% course grade)

2 programming assignments (40% course grade)

1 term project (36% course grade)

Expected Timeline:

Six homework assignments are due weekly before midterm, and six are due just before the end of the semester.

| w | # | day | date | class topic | due 11:59 Tuesday | reading assignment |
|----|----|--------|----------|---|----------------------|---------------------------|
| 1 | 1 | Thurs. | 08/16/18 | Introduction to Compilers and Automata Theory | | Aho 1 |
| 2 | 2 | Tues. | 08/21/18 | Inform 7 programming language | | Aho 2 |
| | 3 | Thurs. | 08/23/18 | Regular Expressions | | Aho 3 |
| 3 | 4 | Tues. | 08/28/18 | * Lex (using lex) | | Brown 1,2,6 |
| | 5 | Thurs. | 08/30/18 | left linear and right linear grammars, deterministic and nondeterministic finite automata | | |
| 4 | 6 | Tues. | 09/04/18 | * Natural Language Processing (using NLTK) 1 | | Bird 0-4 |
| | 7 | Thurs. | 09/06/18 | the lexical analyzer (or scanner) module | | |
| 5 | 8 | Tues. | 09/11/18 | * Natural Language Processing (using NLTK) 2 | | |
| | 9 | Thurs. | 09/13/18 | | | |
| 6 | 10 | Tues. | 09/18/18 | context-free grammars | proj 1 | Aho 4 |
| | 11 | Thurs. | 09/20/18 | * Yacc (using yacc) 1 | | Brown 3, 4, 5, 7, 8, 9 |
| 7 | 12 | Tues. | 09/25/18 | LL parsing | | |
| | 13 | Thurs. | 09/27/18 | LR parsing | | |
| 8 | 14 | Tues. | 10/02/18 | pushdown automata | proj 2 | |
| | 15 | Thurs. | 10/04/18 | Symbol table | | |
| 9 | 16 | Tues. | 10/09/18 | context sensitive grammars and 2-stack pushdown automata | | |
| | 17 | Thurs. | 10/11/18 | (time built in to catch up) | | |
| 10 | 18 | Tues. | 10/16/18 | * intermediate code generation using yacc | proj 3 | |
| | 19 | Thurs. | 10/19/18 | intermediate code generation for common language structures | | |
| 11 | 20 | Tues. | 10/23/18 | * Smalltalk bytecode generation | | Goldberg 26-30 |
| | 21 | Thurs. | 10/25/18 | type checking; Ada example | | |
| 12 | 22 | Tues. | 10/30/18 | * Unicode | | |

| | | | | | | |
|----|----|--------|----------|--|--------|----------|
| | 23 | Thurs. | 11/01/18 | Run-time environments | | Aho 7 |
| 13 | | Tues. | 11/06/18 | (time built in to catch up) | | |
| | 24 | Thurs. | 11/08/18 | | | |
| 14 | 25 | Tues. | 11/13/18 | Optimization (1) | proj 4 | Aho 8 |
| | 26 | Thurs. | 11/15/18 | Optimization (2) | | Aho 9-12 |
| 15 | | Tues. | 11/20/18 | WVU CLOSED - Thanksgiving Break | | |
| | | Thurs. | 11/22/18 | WVU CLOSED - Thanksgiving Day | proj 5 | |
| 16 | 27 | Tues. | 11/27/18 | Implementation: Curveship implementation example | | |
| | 28 | Thurs. | 11/29/18 | application of Curveship | | |
| 17 | 29 | Tues. | 12/04/18 | TBD | | |
| | 30 | Thurs. | 12/06/18 | TBD (last day of classes) | | |

Assessment

Major Assignments/Assessments:

Homework assignments, typically 5 questions per set (12 sets out of 15), will be constructed from a pool of questions provided by the text book author. Students may attempt each assignment multiple times before the due date, with the recorded grade being that of the most recent attempts. Students will receive feedback on incorrect answers with explanations of why this answer is wrong and often suggestions on how to discover the correct answer.

| | | | |
|----|---|--|--------|
| 5 | Computer programs | 1st @ 4%, 2nd and 3rd @5%, 4th and 5th @ 10% | 34.00% |
| 2 | Term project | 36% each | 30.00% |
| 12 | Homework assignments (at New Gradiance) | 2% each | 36.00% |

Homework (via NewGradiance)

<http://www.newgradiance.com/>

class token **D37FF6A0** (last character is zero)

(due via New Gradiance web site at 11:59 pm of indicated date)

| | Title/topic | Due date | chapter |
|----|------------------------------------|--------------------------------------|---------|
| 1 | Programming language concepts | 11:59 pm, Monday, October 7, 2019 | 1 |
| 2 | Parse trees and derivations | | 2, 4 |
| 3 | Code generation for expressions | | 2, 6 |
| 4 | Lexical analysis | | 3 |
| 5 | Design of grammars | | 4 |
| 6 | FIRST and FOLLOW | | 4 |
| 7 | LR(0) Sets of items | 11:59 pm, December 13, 2019 | 4 |
| 8 | LR(1) parsing | | 4 |
| 9 | Types | | 6 |
| 10 | Basic garbage collection | | 7 |
| 11 | Heap allocation and run-time stack | | 7 |
| 12 | Loop optimization and | | 11 |

Programming assignments

| | | |
|---|--------------------|--------|
| | Project assignment | 0% |
| 1 | lex assignment 1 | 20.00% |
| 2 | yacc assignment | 20.00% |

Expected Timeline:

I anticipate that the first six homework assignments and the lex / flex assignment will provide the midterm grade.

Response Time and Feedback Plan:

University policy is now that all submissions be graded within two weeks of submission.

Final Grading Scale:

Most homework assignments will consist of 5 questions, equally weighted.

Final course grades will be letter grades assigned based on the percentage of points received.

| | |
|---|---------|
| A | 90-100% |
| B | 80-89% |
| C | 70-79% |

| | |
|---|--------|
| D | 60-69% |
| F | 0-59% |

Course and Institutional Policies

Attendance Policy:

Attendance is expected but not recorded or graded.

Participation Policy:

Participation is expected but not recorded or graded.

Late Assignment and Missed Exam Policy:

Institutional Policies:

Students are responsible for reviewing policies on inclusivity, academic integrity, incompletes, sale of course materials, sexual misconduct, adverse weather, as well as student evaluation of instruction, and days of special concern/religious holiday statements.

Many policies at the university and college level have been updated. Here are some that focus on student success.

University sanctioned absences now require instructors to provide students with an opportunity to make up missed work or activities and not penalize students for those absences. More details are in the Catalog at <http://catalog.wvu.edu/undergraduate/enrollmentandregistration/#enrollmenttext>.

The university's new policy on smoking, vaping, and tobacco has gone into effect and is now enforceable. The university has several programs to assist smoking cessation.

<https://tobaccoandsmokefree.wvu.edu/home>

The Statler College Smart Device Policy: "The use of programmable calculators or smart devices (including smart-phones, smart watches, tablets, cameras, wearable devices, etc.) on exams and quizzes prohibited

<https://outlook.office365.com/mail/inbox/id/AAQkADQ2NmY5OT...>

unless specifically indicated by the instructor."

The Statler College Sanction Policy for Academic Integrity Offenses is in the Catalog online (along with other college policies) at

<http://catalog.wvu.edu/undergraduate/collegeofengineeringandmineralresources/#policiestextand>

<http://catalog.wvu.edu/graduate/collegeofengineeringandmineralresources/#masterstext>. If you

discover what you think is an academic integrity violation, faculty are to report it online and communicate the charge to the student using the template letter. These, along with the WVU policy and other resources on academic integrity, can be found online at <https://provost.wvu.edu/governance/academic-standards-resources>. The faculty resources are at the bottom of the page. Please consult the Statler College sanction policy when you prepare your letter to insure consistency across the college.

One significant change this year is the university will automatically look for repeat offenses. D/F Repeats can now be done at any point in the student's academic career, not limited to the first 60 hours attempted.

Students are limited in attempts to successfully pass a course. Statler students on the 2018-19 and newer Catalogs have only two attempts to pass a course, including W grades; they may appeal for a third attempt. All students at WVU are limited to three attempts to successfully pass a course, with some leeway granted in 2019-20.