

Formal Languages and Compilers

Prof. Angelo Morzenti
Academic year 2019-2020

Introduction /1

- Formal language:
 - a language with precisely defined syntax and semantics
 - it is possible to define procedures to
 - analyze the syntactic structure and check the grammatical correctness of the phrases
 - compute its meaning (for FLC semantics \equiv translation into a different language)
- The theory of formal languages:
 - studies the form, or ***syntax*** of the phrases
 - defines analysis methods and algorithms
- Compilation
 - applying theory of formal languages
 - to the compilation (and, at large, to the automatic processing)
 - of *programming* languages
 - in general, to all languages of informatics and to «technical languages»

Introduction /2

■ Historical notes

- 1950's – Noam Chomsky: mathematical model of grammars in connection with the study of natural languages
- 1960's – (a lot of research in progr. lang.) ... ALGOL and CFG's
 - syntax-directed compilation, meta-compilers
 - deep link between *formal language theory* and *automata theory*
 - grammars: regular, context free, context-sensitive
- 1970's-1980's: Formal languages, automata and Compilation become fundamental disciplines of Computer science
- Afterwards: these theories continuously foster innovative research and applications

Program/1

■ Syntax

- Formal language Theory
- Regular expressions and languages
- Context-free grammars (briefly: free grammars)
- Grammars for regular languages

■ Finite automata and regular language recognition

- Deterministic finite automata
- Nondeterministic finite automata
- Relations between regular languages and finite automata

Program/2

- Phrase recognition and parsing
 - Recognizing free languages through pushdown automata
 - Deterministic syntax analysis (bottom-up / top-down)
- Semantic translation and static analysis
 - translation relation and function
 - transliterations
 - regular translations
 - purely syntactic translations
 - semantic translations (attribute grammars)
 - static analysis of programs (based on finite automata and Data Flow equations)

Laboratory/design lessons

- Some practice for the theoretical part of the course
- Tools for compiler design and a «toy compiler»
 - **Lexical analysis (scanner)**
 - token identification, vocabulary search, attribute extraction
 - **flex** is a scanner generator
 - GNU General Public License (GPL)
 - **Syntax Analysis (parser)**
 - identify the phrase structure according to a given grammar
 - **bison** is a parser generator
 - GNU GPL
- where: in classroom
- When: 18 hours, around november/december
- Who: Ph.D. Marcello Bersani

Teaching Material

- Text

- Stefano Crespi Reghizzi, Luca Breveglieri, Angelo Morzenti, Formal Languages and Compilation, Springer Verlag, Texts in Computer Science, 2nd Ed. 2013 or 3rd Ed. 2019
- Italian Version (equivalent to the 2nd English one): Stefano Crespi Reghizzi, Luca Breveglieri, Angelo Morzenti, Linguaggi Formali e Compilazione, Esculapio Ed., 2015

- on the Beep platform

- Slides
- Laboratory lessons
- Exam exercises with solutions

Teaching and exams

- «Theory» Lessons (~32 hrs)
 - Includes examples useful for written exam
- Laboratory Lessons (~18 hrs)
 - On the tools for compiler design, plus some on the theory part
- Written Exams
 - No midterm test
 - Exams calls at course end according to School rules
 - Exam consists of two “open book” tests that can be taken separately:
 - Theory and Practice (modify a fragment of a given compiler)
 - Final mark: weighted average of the Theory (80%) and Practice (20%) parts (both ≥ 15)
 - Usually Practice part delivered first (≈ 60 min.), then Theory (≈ 2 hrs. 15 min.)
 - Validity of a passed test: \approx one year (NB: 5 calls, including the one where the mark was obtained)
 - If you redo a part you lose the corresponding previous mark, if you refuse an overall mark (2 partial sufficient marks) you must redo at least one part (with the usual constraint that any partial mark is valid for 4 calls after it is obtained); no sufficient mark can be «frozen» (if not registered it is lost).
- Oral exam
 - no oral exam, only upon request by the teacher

References

- Prof. Angelo Morzenti
 - angelo.morzenti@polimi.it
 - Tel 02/23993634
 - Office: 3rd floor Ed. 22, Via Golgi 42
 - Meetings with students on appointment via email
- Ph.D. Marcello Bersani marcellomaria.bersani@polimi.it