Formal Grammars
Reading Group 4

Marco Kuhlmann
Programming Systems Lab
Saarland University, Saarbrücken

2007-02-21

Regular string languages and finite string automata

- **overview**: one of the simplest classes of string languages with an equivalent characterization in terms of automata

- **motivation**: fundamental importance for formal language theory; finite-state methods abound

- **focus of the talk**: elementary techniques; closure properties; relation to MSO on strings

Betreuer: Gert Smolka

Regular tree languages and finite tree automata

- **overview**: natural correspondent of regular string languages with similar concepts and techniques

- **motivation**: ubiquitous applications in computer science; model of derivational processes in grammar formalisms

- **focus of the talk**: finite tree automata; regular tree grammars; relation to derivations in CFGs; MSO on trees

Betreuer: Gert Smolka

Mildly context-sensitive grammar formalisms

- **overview**: CFGs have been extended into 'mildly' context-sensitive grammar formalisms (TAG, CCG, LIG, MCFG)

- **motivation**: several syntactic phenomena in natural language cannot be modelled by CFGs, or at least not easily

- **focus of this talk**: linguistic motivations with examples; overview of the formalisms and results

Betreuer: Ralph Debusmann

Combinatory categorial grammar (CCG)

- **overview**: mildly context-sensitive grammar formalism that combines categorial grammar and combinatory logic

- **motivation**: proof-theoretic approach towards syntactic modelling; semantics via Curry-Howard isomorphism

- **focus of the talk:** presentation of the formalism; linguistic modelling; parsing algorithm

Betreuer: Ralph Debusmann

Multiple CFGs

- **overview**: generalization of CFGs where not one, but several non-terminals are substituted in parallel

- **motivation**: increased expressivity of the resulting string languages; polynomial membership problem

- **focus of the talk:** presentation of the formalism; results on expressivity and processing complexity

Betreuer: Mathias Möhl

Tree transducers

- **overview**: class of automata that maps trees into trees or strings, and provides a formal model of translation

- **motivation**: alternative characterization of mildly context-sensitive formalisms; practical applications (XML)

- **focus of the talk**: overview of the formalism; relation between tree transducers and grammar formalisms

Betreuer: Marco Kuhlmann

CFG parsing as Boolean matrix multiplication

- **overview:** CFG parsing can be reduced to Boolean matrix multiplication and vice versa

- **motivation:** establish better bounds for CFG parsing; explain why sub-cubic algorithms are not in use

- **focus of the talk:** motivation of the relevance of the result; technical aspects of the reduction

Betreuer: Guido Tack

Parsing as deduction

- **overview**: methods and techniques for proving parsing algorithms correct

- **motivation**: formal reasoning about parsing algorithms; separation of concerns (parsing schema, implementation)

- **focus of this talk**: overview of the concepts; complexity analyses; correctness proofs

Betreuer: Guido Tack

Lexicalized parsing

- **overview**: variants of CFGs and TAGs where rules are specialized for individual words

- **motivation**: adequate representation of linguistic information; faster parsing; interesting formal analyses

- **focus of the talk:** benefits and problems of lexicalized formalisms; parsing algorithms

Betreuer: Ralph Debusmann

Transformations of parsing schemata

- **overview**: systematic manipulation of parsing schemata to influence the operational behaviour of the parsers

- **motivation**: parsing schemata as logic programs; improvements in worst-case runtime

- **focus of this talk**: presentation of the program transformations with examples

Betreuer: Mathias Möhl

Semiring parsing

- **overview**: extension of the parsing-as-deduction framework to weighted deduction

- **motivation**: integration of probabilistic information; statistical natural language processing

- **focus of the talk**: general formal setup; training of probabilistic context-free grammars

Betreuer: Marco Kuhlmann

PTIME Languages

- **overview**: characterization of languages in terms of the complexity of their membership problems

- **motivation**: declarative characterization of tractability; more expressive formalisms

- **focus of this talk**: presentation of the LFP formalism; proof of the equivalence between LFP and PTIME

Betreuer: Marco Kuhlmann

Zeitplan

- **bis Freitag, 09:00 Uhr:**
  E-Mail mit 3 Themenwünschen an kuhlmann@ps.uni-sb.de

- **bis Freitag, 18:00 Uhr:**
  Verteilung der Themen wird per E-Mail mitgeteilt

- **bis Anfang nächster Woche:**
  Kontakt mit dem Betreuer aufnehmen

- **Blockseminar vom 22.–23.3.**

Grundlage der Bewertung

- **mündliche Präsentation** (Deutsch oder Englisch)

  - elektronische Folien (Abgabe als PDF)

  - 25 Minuten + Diskussion

- **schriftliche Ausarbeitung** (Englisch)

  - erweiterte Zusammenfassung des Vortrags

  - 10–15 Seiten in vorgegebenem Layout (Abgabe als PDF)