Large Hand-Coded Grammars

• Characteristics
  broad-coverage
  grounded in linguistic theory
  reusable across domains

• Three generations
  First: mixed procedural and declarative knowledge
  Second: separation of declarative and procedural parts
  Third: formalization, elimination of redundancy
First Generation of Broad-Coverage Grammars

• Characteristics
  large inventory of rules, simple lexical entries
  use of procedures to encode some grammatical properties
  atomic categories
  single-site development of system

• Examples
  Fidditch deterministic parser [Hindle 1983]
  Linguistic String Project [Sager 1981]
  Systran: machine translation (P. Toma, starting in 1968)
Second Generation of Broad-Coverage Grammars

• Characteristics
  grammatical properties encoded declaratively
  central role of lexicon in encoded constraints
  rich use of features on categories
  still mostly single-site development of system

• Examples
  SRI Core Language Engine [Alshawi 1992]
  IBM PLNLP; Microsoft WinNLP [Jensen et al. 1993]
  ATR Fifth Generation [Nagao et al. 1985]
  Penn English XTAG [Joshi et al, 1975]
  Hewlett-Packard NL system [Flickinger et al, 1985]
Third Generation of Broad-Coverage Grammars

• Characteristics
  hierarchical organization of constraints
  support for reversibility (parsing and generation)
  focus on efficient processing
  multi-site collaboration

• Examples
  ParGram (PARC, Xerox Grenoble, Stuttgart, Bergen) [Butt et al. 1999]
  LinGO (Stanford, Saarbrücken, Tokyo, Cambridge, Sussex, Groningen, Trondheim, Oslo, Torino) [Oepen et al. 2003]
  F/XTAG (U. Penn, Paris, Korea) [Doran et al. 1994]
  ALE/ConTroll (Tübingen, Ohio State) [Götz and Meurers 1997]