Natural Language Processing 2001

EJB

ambiguities require a) general knowledge (telescopes are instruments of seeing etc) or b) statistical preferences that may be partly structurally mediated but are largely lexical (see+telescope is more common than man+telescope)

2)

Problems: ran up = vb+prt or vb+prep; byPP is adjectival/adverbial, definite description with superlative picks out a unique path on the hill via a bridging inference; had cues a discourse explanation rather than narrative sequence; her corefers with Kim (not Sandy) by no reference within clause rule for non-refl pronouns and because Sandy is discourse topic; egged on is an idiomatic phrasal verb discontinuous because of particle movement.

Stages: morph, syn, sem, prag/discourse and descrip of what's involved and potentially resolved at each stage

Eval: morph/PoS tagging shld identify PoS reliably but might treat ran+up incorrectly as vb+prt as it's common (ran up the bill) and might get egg+on wrong as discontinuous; syn might get byPP attachment wrong cos (hill by the lake, location by location, plausible) given knowledge-based or statistical interp.; bridging inference needs deeper world knowledge (hill has inclined paths etc); egg+on is an idiomatic non-compositional predicate, but egg alone is an unlikely verbal predicate, so good semantic lexical info shld resolve this; had is a good simple lexical cue for explanation role / reversed temporal sequencing, so don't need to know about the causal relations between egg+on and run; her is resolved by syn and simple discourse structural rules, so no need for complex world knowledge.