1. (18 pts) These should be readily found in the text.
   (a) Generates/represents logical form.
   (b) System-generated var for discourse variables.
   (c) Agreement of any sort, mostly number.
   (d) Verb feature: tense.
   (e) Word class.
   (f) Word subclass (like TO-LOC).
   (g) A WH-Q. This is a Boolean.
   (h) Indicates whether a PP is part of predicate. This is a Boolean.
   (i) Root form of a word.

2. (6 pts) Local movement is where a part of speech is relocated in the sentence to a fixed position within the sentence. For example, yes-no questions exhibit subject-verb inversion. Unbounded means that parts can move to many positions. For example in WH-questions, the WH component can represent/take the place of the subject, direct object, indirect object, etc.

3. (2 pts) Will, can, would, could, should, etc.

4. (6 pts) A gap is literally a ”hole” in a sentence where a component is missing; the filler is a sentence part/constituent that plugs up the hole. In WH-Qs, when the questions is converted to a declarative sentence, there is a hole to be filled by the answer to the question (the who, what, where, etc.).

   Where did he put the book? ⇒ He put the put GAP.

5. (2 pts) With variables and values, specifically the HOLD list which stores categories that are moved. The hold action places constituants in the hold list. VIR arcs pull things out of the hold list.

6. (2 pts) Plain old variables and values, namely the fillers-in and fillers-out lists, which pass potential fillers in and out of constituents. The process called gap threading.

7. (4 pts) Determine the probability that a given word has a particular sense based on its occurrence in a corpus.

8. (2 pts) $\text{Prob}(N|flies) = 20/80$ and $\text{Prob}(V|flies) = 60/80$.

9. (2 pts) See the previous question, which calculates the MLI.

10. (2 pts) A word may not occur in the corpus, so its MLI = 0.

11. (4 pts) It assigns a small frequency to every word, thus no word will ever have MLI = 0. But other words will have their MLI slightly different than it is in actuality because of this add’l amount added to its frequency.

12. (2 pts) It uses local context of $n$ words to determine the probability of a word’s sense.
13. (2 pts) An ngram of two words.

14. (3 pts) \( \text{Prob}(N|P) = 25/100. \)

15. (2 pts) Essentially an ATN with probabilities associated with the arcs.

16. (2 pts) Follow the arcs that correspond to the sequence, multiplying the probabilities on the arcs.

17. (4 pts) Given a sequence of words, it finds the most likely sequence of word senses on a constituent-by-constituent basis. For a given pair in the sequence, it finds the most likely senses. Only the most likely is kept. Then the next constituent is considered, and its most likely interpretation is determined. The algorithm moves "layer-by-layer", retaining only a single path to any layer.

18. (3 pts) See discussion pp 213 - 219. Basically most likely constituent and/or rules are used first. The fly in the ointment is that now are not parsing strictly left-to-right and so have to modify the rule-replacement steps accordingly.

19. (2 pts) Meaning of a sentence (context independent, except for local context).

20. (2 pts) Logical generally generated during parsing, so a roughly 1:1 correspondence. No context info maintained/used. Internal represents everything, including aspects related to discourse context (pronoun referents, WH referents, etc.)

21. (3 pts) Ambiguity based on meaning given 2 syntactically similar structures (same parse trees). "He drove the cattle at the fair."

22. (3 pts) Differences due to structural differences (different parse trees). "He saw the girl with a telescope." where the PP can modify "girl" or "saw".

23. (2 pts) For example, if \( A \) believes \( X \), and \( X \equiv Y \), it doesn’t mean that \( A \) believes \( Y \).

24. (2 pts) In a group of people, every person in the group did not vote. OR in a group of people, some of them did not vote.

25. (4 pts) There is so much variation in the way Allen handle sthis that anything close to anything he did was acceptable. Ultimately, thematic-based roles are the easiest and the ones we’ll concentrate on.

\[
(\text{EARN1 e1 (PAST e1)} \: <\text{EVERY b1 BOY}> \: (\text{RESPECT1 r1 (OF r1 < \text{THE t1 TEACHER1}>))))
\]

26. (12 pts) Explain what each of these thematic roles represents.
   
   (a) Perpetrator of an action.

   (b) Result of an action.

   (c) Animate object that senses/undergoes something.

   (d) Recipient of an action.

   (e) Tool of an action.

   (f) Location from which an action occurs.

27. (6 pts)
\[
(\text{GIVE1 g1 [AGENT < \text{THE d1 DOCTOR1}>] [THEME < \text{THE m1 MEDICINE1}>] [BENEFICIARY h1 HER1]})
\]
\[
(\text{WATCH1 w1 [EXPERIENCER < \text{THE b1 BOY1}>] [THEME < \text{THE g1 GAME1}>] [INSTRUMENT b1 BINOCULAR1 (PLURAL b1)]})
\]
28. (2 pts) Structure built piece-by-piece. Components can be added in.

29. (3 pts) Uniform structure that can can be used anywhere with no special rules, etc. for special forms.

30. (3 pts) \( \lambda x \ (\text{HAMMER} \ x < \ \text{THE} \ d1 \ \text{DRUMs} >) \)

31. (3 pts) \( (\lambda x \ (\text{LIKE1} \ l1 \ x \ (\text{CANDY1} \ c1)) \ (\text{PRO} \ h1 \ \text{HE})) \)

32. (3 pts) \( \lambda x \ (\text{INTO} \ x) \)

33. (2 pts) In the grammar, can have rule for each sense of a word and so will be able to represent nuances well. And for all possible roles that may (or may not) be included in a sentence; i.e., instrument may or may not appear with the verb, theme may or may not appear with the verb, and so would need rules for each possibility. Will require lots of rules with lots of duplication for words that have similar structure. If move semantics into the lexicon, makes rules simpler, but now have more complexity at the lexical level, with possible duplication here for words with same structure.

34. (2 pts) Place thematic structure in the hierarchy and have root words inherit from only those that share the structure. Eliminates duplication as mentioned above as now common aspects attached to higher nodes in the hierarchy.