Name:_____

1. What are the two main problems with probabilistic context free grammars?

Answer: Lack of sensitivity to lexical information
Lack of sensitivity to structural frequency

2. Given the two sentences below assume the most likely parse trees are marked with square brackets. Name one problem that a PCFG would have when finding the most likely parse trees (as indicated) corresponding to each of the sentences:

   Workers dumped [NP sacks] [PP into a bin]
   Fishermen caught [NP tons [PP of herring]]

   Answer: Both of these sentences have a PP attachment ambiguity. The correct parse tree for the first sentence should attach the PP to the verb, while the correct parse for the second sentence should attach PP to the noun "tons". Because a PCFG does not take into account lexical information, the choice between the two parse trees would depend on the probabilities of the rules VP -> VP NP, and VP -> VP PP , so that for the both sentences any PCFG would select the same attachment type (therefore any PCFG is bound to make a mistake either on the first sentence or on the second sentence).

3. We have a lexicalized grammar with some valid rules and some invalid rules. For each of the following rules, indicate whether it is valid or invalid:
   1. DT(the)-> a
   2. NN(a)-> a
   3. SBAR(that)→ 1 COMP(that) S(was)
   4. SBAR(was) → 2 COMP(that) S(was)
   5. PP(in) → 1 IN(of) NP(company)

   Answer:
   1. Invalid
   2. Valid
   3. Valid
   4. Valid
   5. Invalid

4. Given the following PCFG grammar, compute the probabilities of the two derivations that are possible for the sentence *"He saw the man with the telescope".* Draw a tree that corresponds to each of the derivations.

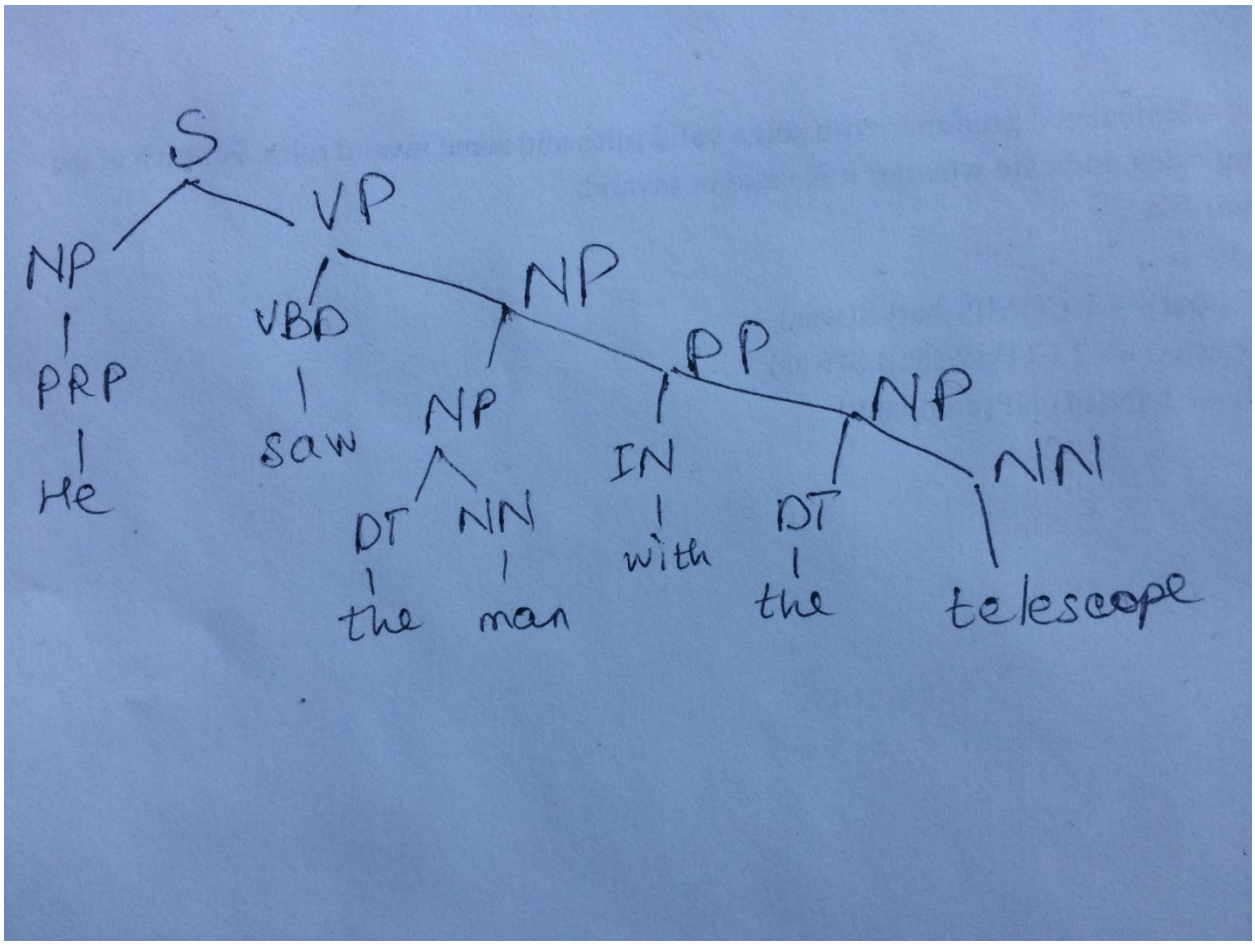
   q(S -> NP VP) = 0.9
   q(S -> NP) = 0.1
   q(NP -> DT NN) = 0.3
   q(NP -> NP PP) = 0.5
   q(NP->PRP)=0.2
   q(VP -> V NP) = 0.3
   q(VP -> V NP PP)=0.7
   q(PP -> IN NP)=1.0
   q(NN-> man)=0.7
   q(NN-> telescope)=0.3
   q(IN->with)=1.0
   q(PRP->he)=1.0
   q(V-> saw)=1.0
   q(DT -> the) = 0.8

Answer:

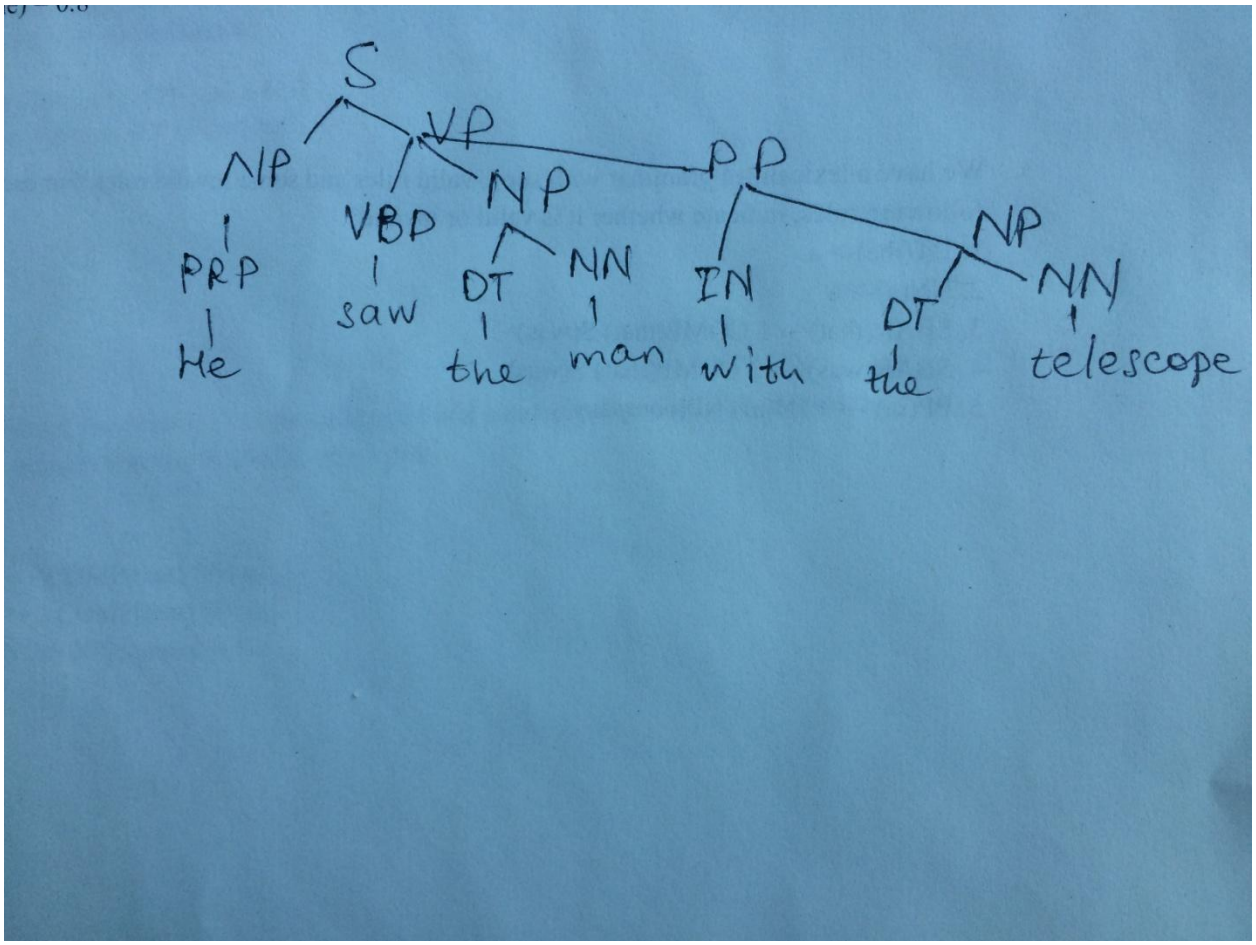
Derivation 1	**Derivation 2**
S -> NP VP 0.9	S -> NP VP 0.9
NP -> PRP 0.2	NP -> PRP 0.2
PRP -> he 1.0	PRP -> he 1.0
VP -> V NP 0.3	VP -> V NP PP 0.7
V -> saw 1.0	V -> saw 1.0
NP -> NP PP 0.5	NP -> DT NN 0.3
NP -> DT NN 0.3	DT -> the 0.8
DT -> the 0.8	NN -> man 0.7
NN -> man 0.7	PP -> IN NP 1.0
PP -> IN NP 1.0	IN -> with 1.0
IN -> with 1.0	NP -> DT NN 0.3
NP -> DT NN 0.3	DT -> the 0.8
DT -> the 0.8	NN -> telescope 0.3
NN -> telescope 0.3	

Derivation 1: 0.9*0.2*1.0*0.3*1.0*0.5*0.3*0.8*0.7*1.0*1.0*0.3*0.8*0.3
Derivation 2: 0.9*0.2*1.0*0.7*1.0*0.3*0.8*0.7*1.0*1.0*0.3*0.8*0.3

Tree 1


S
VP
NP
VBD
NP
PP
NP
PRP
saw
NP
IN
DT
NN
He
DT
NN
with
the
man
the
telescope

Tree 2


S
NP
VP
PP
PRP
VBD
NP
NP
DT
NN
IN
DT
NN
He
saw
the
man
with
the
telescope