**Module Title: Informatics 2A**
**Exam Diet (Dec/April/Aug): August 2013 2012–13**
**Brief notes on answers:**

## PART A

1. (a) Equations:

$$
\begin{aligned}
X_1 &= aX_3 + bX_2 + \epsilon \\
X_2 &= aX_1 + bX_2 \\
X_3 &= aX_3 + bX_1
\end{aligned}
$$

(b) Applying Arden's rule to $X_2$ and $X_3$ we get respectively

$$X_2 = b^* a X_1$$
$$X_3 = a^* b X_1$$

Substituting in to equation for $X_1$, we get:

$$
\begin{aligned}
X_1 &= aa^*bX_1 + bb^*aX_1 + \epsilon \\
&= (aa^*b + bb^*a)X_1 + \epsilon \\
&= (aa^*b + bb^*a)^*
\end{aligned}
$$

with the last equation again by Arden's rule. Thus the regular expression for the language is

$$(aa^*b + bb^*a)^*$$

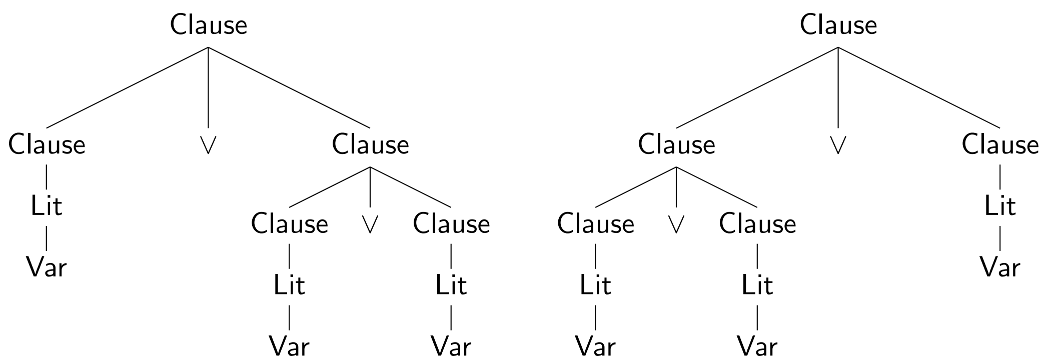
2. (a)

| transition | control state | unread input | stack |
|---|---|---|---|
| | $q1$ | $aab$ | $\bot$ |
| $q1 \xrightarrow{a,\bot:\bot\bot} q1$ | $q1$ | $ab$ | $\bot\bot$ |
| $q1 \xrightarrow{a,\bot:\bot\bot} q1$ | $q1$ | $b$ | $\bot\bot\bot$ |
| $q1 \xrightarrow{\epsilon,\bot:\epsilon} q2$ | $q2$ | $b$ | $\bot\bot$ |
| $q2 \xrightarrow{b,\bot:\epsilon} q2$ | $q2$ | $\epsilon$ | $\bot$ |
| $q2 \xrightarrow{\epsilon,\bot:\epsilon} q2$ | $q2$ | $\epsilon$ | $\epsilon$ |

(b) The language is:

$$\{a^n b^m \mid n \geq m \geq 0\}$$

3. (a) *Structurally ambiguous:* there is a phrase that has more than one parse tree. An example: Var ∨ Var ∨ Var.
This has the two parse trees below.



(b) An LL(1) grammar:

$$\begin{aligned} \mathsf{Lit} &\rightarrow \mathsf{Var} \mid \neg\, \mathsf{Var} \\ \mathsf{Clause} &\rightarrow \mathsf{Lit\ ClauseTail} \\ \mathsf{ClauseTail} &\rightarrow \epsilon \mid \vee\ \mathsf{Clause} \end{aligned}$$

4. The Earley parsing chart is:

| | | | | |
|---|---|---|---|---|
| S | $\rightarrow$ | $\bullet$ NP V | [0,0] | P |
| NP | $\rightarrow$ | $\bullet$ N | [0,0] | P |
| NP | $\rightarrow$ | $\bullet$ N N | [0,0] | P |
| NP | $\rightarrow$ | N $\bullet$ | [0,1] | S |
| NP | $\rightarrow$ | N $\bullet$ N | [0,1] | S |
| S | $\rightarrow$ | NP $\bullet$ V | [0,1] | C |
| NP | $\rightarrow$ | N N $\bullet$ | [0,2] | S |
| S | $\rightarrow$ | NP $\bullet$ V | [0,2] | C |
| S | $\rightarrow$ | NP V $\bullet$ | [0,3] | S |

(Or they may write bread, price, rises in place of N,V to the left of the $\bullet$.)

5. (a) The raw lambda expression arising from the definition is
$(\lambda x.\ (\lambda z.\mathit{child}(z))(x) \wedge\ (\lambda y.(\lambda uv.\mathit{likes}(u,v)(\mathit{Anna},y)(x)))\ \ (\mathit{Bill})$

(b) This reduces via four $\beta$-reductions (which can be done in various orders) to
$\mathit{child}(\mathit{Bill}) \wedge\ \mathit{likes}(\mathit{Anna},\mathit{Bill})$

**PART B**

6. (a) *Not regular.*

   We show ¬ P (the negation of the pumping property).

   Suppose $k \geq 0$.

   Consider $x = \epsilon$, $y =$ <a>$^k$ and $z =$ </a>$^k$. Then $xyz =$ <a>$^k$</a>$^k \in L$ and clearly $|y| \geq k$.

   Suppose $y = uvw$ where $|v| \geq 1$.

   Then $uv^0w = uw =$ <a>$^m$ for some $m < k$. Whence $xyv^0wz =$ <a>$^m$</a>$^k \notin L$ since $m < k$.

   Thus the pumping property fails for $i = 0$.

   (b)

   *First*(Doc) = { $\epsilon$, text, <a>, <b> }
   *Follow*(Doc) = { \$, </a>, </b> }

   (c)

| | text | <a> | <b> | </a> | </b> | \$ |
|---|---|---|---|---|---|---|
| Doc | text | <a>Doc</a>Doc | <b>Doc</b>Doc | $\epsilon$ | $\epsilon$ | $\epsilon$ |

   (d) (i). End of input reached but </a> left on stack.
   Error message: "</a> expected at end of document."

   (ii). </b> in input does not match </a> on stack.
   Error message: "</b> occurs where </a> expected."

   (iii). Stack empties before end of input reached.
   Error message: "completed document before occurrence of </b>."

7. (a) Bookwork. The probability assigned to a rule X → $\alpha$ is the number of applications of this rule in the corpus divided by the total number of appearances of the non-terminal X in the corpus.

   (b) The grammar with probabilities added is:

   S → Name VP (1.0)
   VP → V NP (0.6) | V NP PP (0.4)
   NP → the N (0.875) | the N PP (0.125)
   PP → with NP (1.0)

   (Note that VP occurs 5 times in the corpus, and NP occurs 8 times.)

   (c) There are two parse trees:

   (S (Name John)(VP (V saw)(NP the (N man)(PP with (NP the (N telescope))))))
   (S (Name John)(VP (V saw)(NP the (N man))(PP with (NP the (N telescope)))))

   The first of these (where the man has the telescope) has probability
   1.0 x 0.6 x 0.125 x 1.0 x 0.875 = 0.065625.

   The second (where John has the telescope) has probability
   1.0 x 0.4 x 0.875 x 1.0 x 0.875 = 0.28125.

(d) The lexicalized rules and head selection rules have probabilities as follows:

$$
\begin{array}{rcl}
\mathsf{VP}[\text{saw}] & \rightarrow & \text{saw } \mathsf{NP}\ (0.667) \ \mid\ \text{saw } \mathsf{NP}\ \mathsf{PP}\ (0.333) \\
\mathsf{NP}[\text{man}] & \rightarrow & \text{the man } (0.5) \ \mid\ \text{the man } \mathsf{PP}\ (0.5) \\
\mathsf{NP}[\text{telescope}] & \rightarrow & \text{the telescope } (1.0) \ \mid\ \text{the telescope } \mathsf{PP}\ (0.0) \\
\mathsf{VP} & \rightarrow & \mathsf{VP}[\text{saw}]\ (0.6) \\
\mathsf{NP} & \rightarrow & \mathsf{NP}[\text{man}]\ (0.25) \\
\mathsf{NP} & \rightarrow & \mathsf{NP}[\text{telescope}]\ (0.125)
\end{array}
$$

(e) The first of the above parse trees now has probability
1.0 x 0.2 x 0.4 x 0.667 x 0.25 x 0.5 x 0.125 x 1.0 = 0.0008333.
The second has probability
1.0 x 0.2 x 0.4 x 0.333 x 0.25 x 0.5 x 0.125 x 1.0 = 0.0004167.

8. (a) This is an example of a noncontracting grammar. This is because the right side of every production is at least as long as the left side. Since noncontracting grammars are equivalent in power to context-sensitive grammars, we can say that the language is (at worst) context sensitive.

(b) (i). Underlining the sequence to be expanded on the next line:

$$
\begin{array}{rcl}
\underline{\mathsf{S}} & \Rightarrow & \underline{\mathsf{E}\ \mathsf{B}_{\mathsf{apples}}} \\
 & \Rightarrow & \underline{\mathsf{G}}\ \mathsf{apples}\ . \\
 & \Rightarrow & \mathsf{Granny\ went\ to\ market\ and\ she\ bought\ some\ apples}\ .
\end{array}
$$

(ii).

$$
\begin{array}{rcl}
\underline{\mathsf{S}} & \Rightarrow & \mathsf{E}\ \underline{\mathsf{S}}\ \mathsf{B}_{\mathsf{apples}} \\
 & \Rightarrow & \mathsf{E}\ \underline{\mathsf{E}\ \mathsf{B}_{\mathsf{bananas}}}\ \mathsf{B}_{\mathsf{apples}} \\
 & \Rightarrow & \mathsf{E}\ \mathsf{G}\ \mathsf{bananas}\ \underline{.\ \mathsf{B}_{\mathsf{apples}}} \\
 & \Rightarrow & \mathsf{E}\ \mathsf{G}\ \underline{\mathsf{bananas}\ \mathsf{B}_{\mathsf{apples}}}\ . \\
 & \Rightarrow & \mathsf{E}\ \underline{\mathsf{G}\ \mathsf{B}_{\mathsf{apples}}}\ \mathsf{bananas}\ . \\
 & \Rightarrow & \underline{\mathsf{E}\ \mathsf{B}_{\mathsf{apples}}}\ \mathsf{G\ apples\ and\ some\ bananas}\ . \\
 & \Rightarrow & \mathsf{G\ apples}\ .\ \mathsf{G\ apples\ and\ some\ bananas}\ .
\end{array}
$$

From this the desired phrase is obtained by two G expansions.

(c) $L_2$ is context free. A grammar for it:

$$
\begin{array}{rcl}
\mathsf{S} & \rightarrow & \mathsf{T}\ . \\
\mathsf{T} & \rightarrow & \mathsf{G\ i\ .\ G\ i} \ \mid\ \mathsf{G\ i\ .\ T\ and\ some\ i} \\
\mathsf{G} & \rightarrow & \mathsf{Granny\ went\ to\ market\ and\ she\ bought\ some}
\end{array}
$$

(d) A possible grammar:

$$
\begin{array}{rcll}
\mathsf{S} & \rightarrow & \mathsf{G\ i\ .\ G\ i\ .} \ \mid\ \mathsf{G\ i\ .\ S\ B_i}. & \\
\mathsf{t\ B_i} & \rightarrow & \mathsf{B_i\ t} & \mathsf{t} \in \Sigma - \{\mathsf{bought}\} \\
\mathsf{G\ B_i} & \rightarrow & \mathsf{G\ i\ and\ some} & \\
\mathsf{G} & \rightarrow & \mathsf{Granny\ went\ to\ market\ and\ she\ bought\ some} &
\end{array}
$$