Tutorial 7

1. Following is the table built using dynamic programming algorithm.

i/j	1	2	3	4	5	6	7	8
1	CD	A	CSB	A	CSB	A	CSB	A
2		CD	A	CB	A	CSB	A	CSB
3			BCD	A	CB	A	CSB	A
4				CD	A	CSB	A	CSB
5					CD	A	CSB	A
6						BCD	A	CB
7							BCD	A
8								CD

Table 1:

In above, entry in row i, column j indicates $X_{i,j}$. Thus, $aabaabba$ does not belong to the language (as S is not in the $X_{1,8}$).

2. (We call the language corresponding to the question as L in the proofs below).

(a) $\{a^i b^j : i = j^2\}$.
No. Suppose by way of contradiction otherwise. Let n be as in pumping lemma.
Let $z = a^{n^2} b^n$.
Let $z = uvwxy$ as in the pumping lemma.
Case 1: v contains both a and b or x contains both a and b.
In this case uv^2wx^2y has a appearing after a b, thus $uv^2wx^2y \notin L$.
So for the following cases assume that v and x are members of $a^* + b^*$.
Case 2: v and x are both members of a^*.
In this case $uv^2wx^2y = a^{n^2+|vx|}b^n \notin L$ (as $vx \neq \epsilon$).
Case 3: v and x are both members of b^*.
In this case $uv^2wx^2y = a^{n^2}b^{n+|vx|} \notin L$ (as $vx \neq \epsilon$).
Case 4: $v \in a^+$ and $x \in b^+$.
Suppose $|v| = r$ and $|x| = s$. Then, $uv^2wx^2y = a^{n^2+r}b^{n+s}$. For this string to be in L, $(n^2 + r) = (n + s)^2 = n^2 + s^2 + 2ns$, or $r = s^2 + 2ns$, which is not possible as $1 \leq r \leq n$ and $1 \leq s \leq n$.
Alternative Proof: Suppose by way of contradiction that L is context free. Let n be as in the pumping lemma. Without loss of generality assume $n > 2$.
Let $z = a^{n^2} b^n$.
Let $z = uvwxy$ as in pumping lemma.
Then, $uwy = a^{n^2-k}b^{n-s}$, for some $k, s \geq 0$ such that $1 \leq k + s \leq n$.
For $uwy \in L$, we must have $n^2 - k = n^2 + s^2 - 2ns$, or $k = 2ns - s^2 = s(2n - s)$.
Thus, if $s = 0$, then $k = 0$; if $0 < s \leq n$, then $k \geq n$. Both violate the condition $0 < k + s \leq n$.
Thus, L is not context free.

(b) $\{a^n b^m c^q \mid n \neq m \text{ or } m \neq q\}$.

It is context free. Following is the grammar for it, where S is the starting symbol.

$S \rightarrow TC|AX$.

$T \rightarrow aTb|bB|aA$

$X \rightarrow bXc|bB|cC$

$A \rightarrow aA|\epsilon$

$B \rightarrow bB|\epsilon$

$C \rightarrow cC|\epsilon$

Intuitively, A, B, C respectively generate arbitrary number of a, b, c's.

T generates equal number of a and b, plus extra a's or b's (at least one extra).

X generates equal number of b and c, plus extra b's or c's (at least one extra).

(c) $\{a^m \mid m \text{ is a prime }\}$.

No. Similar proof as for regular languages works here also.

Suppose by way of contradiction that the language is context free. Then let n be as in the pumping lemma.

Let $z = a^p$, where p is a prime number greater than n.

Then, let $z = uvwxy$ as in the pumping lemma.

Then, $uv^{p+1}wx^{p+1}y = a^{p+p*|vx|} \notin L$, as $|vx| \geq 1$, and thus, $p + p * |vx|$ is not a prime number.

(d) $\{a^i b^j c^k d^l \mid i = 0 \text{ or } j = k = l\}$.

No. Suppose by way of contradiction, otherwise. Consider the intersection of the language in the question with $ab^*c^*d^*$. Then, the resulting langauge $\{ab^j c^k d^l \mid j = k = l\}$ must be context free. Now, consider the substitution, $s(a) = \epsilon$, $s(b) = a, s(c) = b, s(d) = c$. Then, the resulting language $\{a^j b^k c^l \mid j = k = l\}$ must be context free. But this is not true as proved in class.

Thus, the language given in the question is not context free.

(e) Suppose by way of contradiction that L is context free. Then, let $L' = L \cap c\{a,b\}^* cdc\{a,b\}^* c$. Then, $L' = \{cwcdcwc : w \in \{a,b\}^*\}$, must be context free. Now, considering the substitution, $s(c) = s(d) = \epsilon$, $s(a) = a, s(b) = b$, we get that $s(L') = \{ww : w \in \{a,b\}^*\}$, which is shown in class not to be context free. A contradiction.

3. Prove or Disprove:

(a) True.

Suppose (V, Σ, S, P) is the grammar for L without any useless symbols.

The new grammar for Prefix(L) is $(V \cup \{A^p : A \in V\}, \Sigma, S^p, P')$, where P' is defined as follows.

Intuitively, A^p generates prefixes of all strings which are generated by A.

For each production $A \rightarrow \alpha$ in P, we have the following productions in P':

(i) production $A \rightarrow \alpha$

(ii) all productions of form $A^p \to \gamma$, where γ is a prefix of α.

(iii) all productions of form $A^p \to \gamma B^p$, where γB is a prefix of α and $B \in V$.

(b) True.

Suppose $(Q, \Sigma, \delta, q_0, F)$ is a DFA for L.

Then, let $F' = \{q \in Q : q$ is not a dead state, i.e., there exists a w such that $\hat{\delta}(q, w) \in F\}$.

Then, $(Q, \Sigma, \delta, q_0, F')$ is a DFA for $Prefix(L)$.

This holds, as for any $w \in L$, for any prefix u of w, $\hat{\delta}(q_0, u)$ is a not a dead state and thus in F'. Furthermore, for any u such that $\hat{\delta}(q_0, u) \in F'$, there exists a $y \in \Sigma^*$ such that $uy \in L$, and thus $u \in Prefix(L)$.

4. (sketch): Very similar to the proof done in class for pumping lemma except that one only counts the distinguished positions for various parts in the proof.

   Let m = number of non-terminals in the CNF grammar for L. Let $n = 2^m$. Consider any string $z \in L$ with $\geq n$ distinguished positions.

   Consider the derivation tree of z using the CNF grammar. In the derivation tree, mark an internal node if and only if either it has distinguished positions on both left and right subtrees, or if it has only one child which is a distinghished position.

   For a marked internal node s, let $level(s)$ = number of marked ancestors the node has.

   Let s be the internal node with largest level. Let s' be an ancestor of s with level $level(s) - m$. Now, in the path between s' and s there must be at least two marked nodes with same non-terminal: say these nodes are n_1, n_2 with associated non-terminal A (where n_1 is an ancestor of n_2). Thus, we have (in a way similar to that done for pumping lemma)

   $S \Rightarrow^* uAy \Rightarrow^* uvAxy \Rightarrow^* uvwxy = z$. Thus, $S \Rightarrow uv^iwx^iy$ for all i. Now, vwx has at most 2^m distinguished positions (as the subtree rooted at s' has at most 2^m distinguished positions). Furthermore vx has at least one distinguished position as n_1 is a marked internal node and thus both its left and right children have at least one marked descendant (and A at n_2 is on one side).

   Ogden's lemma follows.

5. Consider string $z = 0^n1^n2^{n!+n}$. Let 0's be distinsuished positions. Let $z = uvwxy$. Note that if v or x contain two distinct symbols from $\{0, 1, 2\}$, then clearly $uv^2wx^2y \notin L$. Thus, $v \in 0^*$. Furthermore, if $x \notin 1^+$ or if $|x| \neq |v|$, then, uv^2wx^2y contains different number of 0's than 1's. Thus, we have that $|v| = |x| > 0$, and $v \in 0^+, x \in 1^+$. Now, let $m = |v|$. Then, $uv^{1+n!/m}wx^{1+n!/m}y = 0^{n!+n}1^{n!+n}2^{n!+n} \notin L$. A contradiction.