The Pumping Lemma

See section 4.1

In mathematics a *lemma* is a minor result that is proved as one step in proving something more important. Sometimes the lemmas turn out to be more important than the theorems. There are several Pumping Lemmas in automata theory that are used to show that specific languages *don't* have particular properties. The pumping lemma for regular languages is used to show that some languages are not regular.

The Pumping Lemma For Regular Languages (Dana Scott and Michael Rabin, 1959): Let $\mathcal{L}$ be a regular language. Then there is a constant n (depending on $\mathcal{L}$) such that for every string w in $\mathcal{L}$ with $|w| > n$ we can partition w into 3 substrings, $w = xyz$, such that

1. $|y| > 0$
2. $|xy| <= n$
3. For all $k >= 0$ the string xy^kz is also in $\mathcal{L}$ (i.e. we can "pump" y any number of times and stay in the language.

This shouldn't be a huge surprise. Intuitively it says that if a DFA accepts a very long string then that string must go around a loop in the DFA and we could go around that loop any number of times.

Proof of the pumping lemma: Let (Σ, Q, δ, s, F) be a DFA that accepts language $\mathcal{L}$. Let n be the number of states in Q. Let w=a_0...a_{k-1} be a string in $\mathcal{L}$ of length k >= n. Let P_0,...P_k be the sequence of states w takes the automaton takes the automaton through. There are k+1>n states here and the DFA has only n states so there must be a repeated state, ie there are i and j such that i < j <= n with $P_i = P_j$. So let x=a_0..a_{i-1}, y=a_i..a_{j-1}, and z=a_j..a_{k-1}. The automaton will accept strings xz. xyz, xy^2z, etc. Note that |xy| = j <= n and |y| = j-i > 0.

Example: Show that the set of strings of 0's and 1's with the same number of 0's and 1's is not regular.

Proof: Suppose it is regular. Let n be its pumping constant. Consider the string $w=0^n1^n$, which is a string in this language with length greater than n. There must be a pumping decomposition of w: $w=xyz$ where $|xy|<=n$ and $|y|>0$, so y consists only of a positive number of 0's. However, xy^2z has more 0's than 1's, so it is not in the language. This violates the pumping lemma.

Is $\{0^{2n}1^{2m} | n>=0, m>=0\}$ regular? Sure; it is (00)*(11)*

Is $\{0^n1^m | n>=m\}$ regular? No. Let n be the pumping constant and $w=0^n1^n$. The pumping lemma says there will be a decomposition w=xyz, where |xy|<=n. Just as before y has a positive number of 0's. However, $xy^0z = xz$ is not in the language since it has fewer 0's than 1's.

Is $\{ss^R | s \in (0+1)^*\}$ (i.e., even-length palindromes) regular? No; use $w=0^n110^n$ where n is the pumping constant.