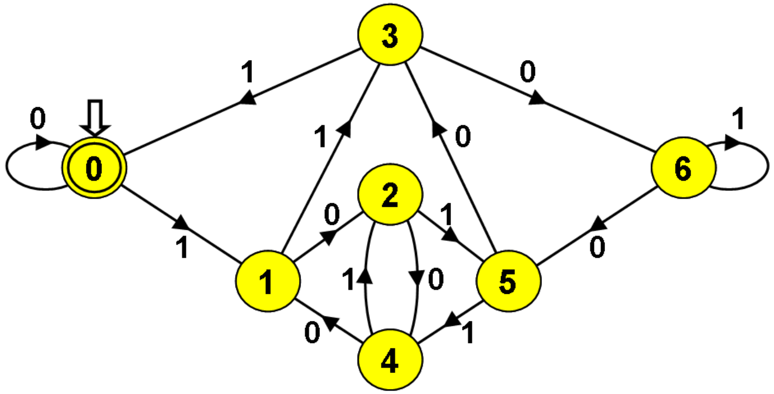
THEOREM OF THE DAY

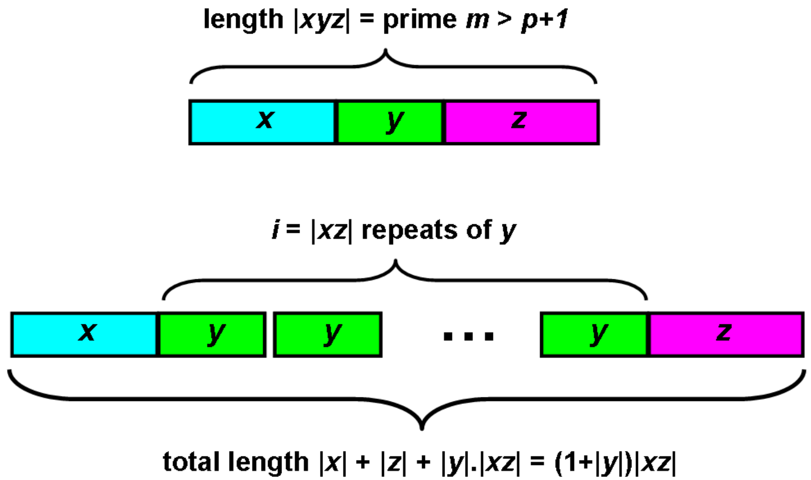
The Pumping Lemma *Let $\mathcal{L}$ be a regular language. Then there is a positive integer p such that any word $w \in \mathcal{L}$ of length exceeding p can be expressed as $w = xyz$, $|y| > 0$, $|xy| \leq p$, such that, for all $i \geq 0$, xy^iz is also a word of $\mathcal{L}$.*



repeat

input	0	1	0	1	0	0	1	1	0	0	1
state	0	1	2	5	3	6	6	6	5	3	0
accept	✓										✓

The set {x : x = $7k$, x in base 2 } is a regular language



The set {a^n : n prime} is not a regular language

Regular languages over an alphabet Σ (e.g. $\{0, 1\}$) are precisely those strings of letters which are 'recognised' by some *deterministic finite automaton* (DFA) whose edges are labelled from Σ. Above left, such a DFA is shown, which recognises the language consisting of all positive multiples of 7, written in base two. The number $95 \times 7 = 665 = 2^9 + 2^7 + 2^4 + 2^3 + 2^0$ is expressed in base 2 as 1010011001. Together with any leading zeros, these digits, read left to right, will cause the edges of the DFA to be traversed from the initial state (heavy vertical arrow) to an accepting state (coincidentally the same state, marked with a double circle), as shown in the table below the DFA. Notice that the bracketed part of the table corresponds to a cycle in the DFA and this may occur zero or more times without affecting the string's recognition. This is the idea behind the pumping lemma, in which p, the 'pumping length', may be taken to be the number of states of the DFA.

So a DFA can be smart enough to recognise multiples of a particular prime number. But it cannot be smart enough recognise all prime numbers, even expressed in *unary* notation ($2 = aa$, $3 = aaa$, $5 = aaaaa$, etc). The proof, above right, typifies the application of the pumping lemma in disproofs of regularity : assume a recognising DFA exists and exhibit a word which, when 'pumped' must fall outside the recognised language.

This lemma, which generalises to context-free languages, is due to Yehoshua Bar-Hillel (1915–1975), Micha Perles and Eli Shamir.

Web link: www.seas.upenn.edu/~cit596/notes/dave/pumping0.html (and don't miss www.cs.brandeis.edu/~mairson/poems/node1.html!)

Further reading: *Models of Computation and Formal Languages* by R Gregory Taylor, Oxford University Press Inc, USA, 1997.