Pumpálási lemma reguláris nyelvekre

Állítás (Ez a lemma állítása)
Ha L reguláris nyelv, akkor $\exists n \geq 0$ egész szám, hogy $\forall w \in L$, ahol $|w| \geq n$ és w minden w_1 részszavára, ahol $w = uw_1v$ és $|w_1| \geq n$, létezik w_1-nek egy $w_1 = xyz$ felosztása, amelyben $y \neq \epsilon$ és amely felbontásra $uxy^izv \in L$ minden $i \geq 0$ esetén.

Azaz: Ha L egy reguláris nyelv, akkor létezik egy olyan hossz, aminél hosszabb tetszõleges L-beli szóban tetszõlegesen kiválasztva egy ennél a hossznál nagyobb részszót igaz lesz a következõ: ezen részszóban van egy olyan rész valahol, melyet tetszõlegesen sokszor megismételve (vagy elhagyva) még mindig L-beli szavakat kapunk.

Ennek a haszna:
Segítségével be lehet látni nyelvekrõl, hogy nem regulárisak. Ez úgy megy, hogy ha megmutatjuk egy nyelvrõl, hogy van olyan szava, amiben nem lehet a fentiek szerint pumpálni, akkor az tuti nem reguláris.
Vigyázat! Lehetnek olyan nyelvek, amiket lehet pumpálni, de mégsem regulárisak. Vagyis ez a lemma csak a nem-reguláris nyelvek egy részét buktatja le.

Például az $\{a^kb^k \mid k \geq 1\}$ nyelv nem reguláris, mert:

Tegyük fel, hogy az. Vegyük az a^nb^n szót, ahol az n az az n, aminek létezését a lemma garantálja és jelöljük ki w_1-nek az a^n részszót! Ha ebben pumpálunk, akkor az a-k száma megnõ, de közben a b-k száma nem változik, azaz a kapott szó nem lesz eleme a nyelvnek.
Ezzel beláttuk, hogy a nyelv nem reguláris.

A lemma bizonyítása:
Ha az L nyelv reguláris, akkor van egy olyan véges automata, ami elfogadja õt. Legyen $n := |Q|$, azaz n legyen az automata állapotainak a száma. Nézzük, hogy az automata milyen állapotokat érint, onnantól kezdve, hogy rálép a $w_1 = x_1 \cdots x_t$ ($t \geq n$) szó elsõ betûjére, egészen odáig, hogy lelép az utolsóról is. Ez összesen $t+1$ állapot: $q_{i_1}, \cdots, q_{i_{t+1}}$. Mivel ez több, mint amennyi állapota az automatának van, biztosan van közöttük két egyforma. Ha $q_{i_k} = q_{i_l} = q$, akkor ez éppen azt jelenti, hogy a szó olvasása során az automata az $y = x_{i_k} \cdots x_{i_{l-1}}$ részszó olvasását a q állapotban kezdi el, elolvassa részszót és ekkor is q-ban van. Ha ekkor megint az y részszó jönne (akárhányszor), akkor megint csak q-ba kerülne. Mivel az eredeti teljes w szó elolvasása után elfogadóban áll meg az automata (azaz a q-ból tovább olvasva a maradék szót elfogadóba jutunk), a sok y-t tartalmazó futásokkor is lesz vele, azaz ezeket a szavakat is elfogadja.
Az $i = 0$ eset úgy jön ki, hogy ha kihagyjuk az eredeti szóból az y-t, a maradék szót akkor is a q állapotból kezdjük el elolvasni, úgy meg elfogadunk.