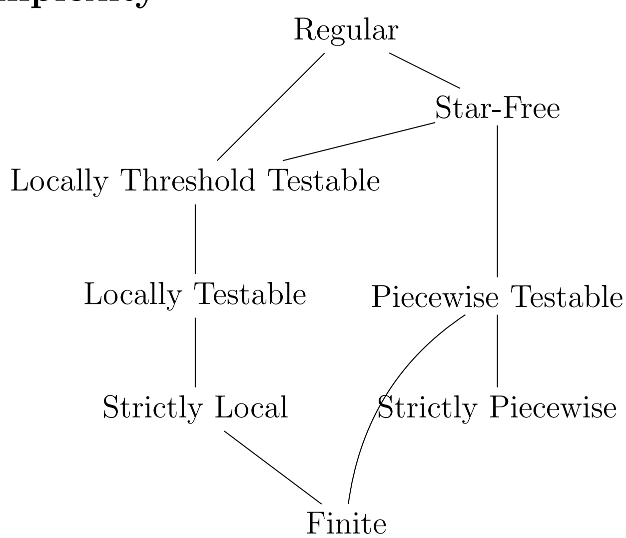
## Subregular Complexity




## Examples

1. Strings end with a `b`. (2-SL)
2. The second to last symbol in all strings is `b`. (3-SL)
3. The third to last symbol in all strings is `b`. (4-SL)
4. Strings contain at least one `b`. (1-LT = 1-PT)
5. Strings contain at most one `b`. (2-SP)
6. Strings contain exactly one `b`. (2-PT)
7. Strings contain at least one `bb` substring. (2-LT)
8. Strings contain two `b`s. (2-PT)
9. Strings contain two `bb` substrings. (2,2-LTT)
10. Strings contain an `a` between every pair of `b`s. (SF)
11. Strings contain an odd number of `b`s. (REG)

## Grammar-Independent Characterizations

SL Suffix Substitution Closure:

A language $L$ is Strictly Local iff there is a $k$ such that for all strings $u_1, v_1, u_2, v_2 \in \Sigma^*$ and for all strings $x$ of length $k$ whenever $u_1xv_1, u_2xv_2 \in L$ then $u_1xv_2 \in L$.

LT Local Testability:

A language $L$ is Locally Testable iff there exists $k$ such that for all $u, v \in \Sigma^*$, if $u$ and $v$ have the same $k-1$ prefix, $k$-long substrings, and $k-1$ suffix then either $u, v \in L$ or $u, v \notin L$.

LTT Local Theshold Testabilty:

A language $L$ is Locally Testable iff there exists $k$ such that for all $u, v \in \Sigma^*$, if $u$ and $v$ have the same $k-1$ prefix, $k-1$ suffix, and the same number of occurrences of the same $k$-long substrings, counting up to some threshold $t$, then either $u, v \in L$ or $u, v \notin L$.

SP Subsequence Closure:

A language $L$ is Strictly Piecewise iff whenever $w \in L$ every subsequence of $w$ also belongs to $L$.

PT Piecewise Testability: A language $L$ is Piecewise Testable iff there exists $k$ such that for all $u, v \in \Sigma^*$, if $u$ and $v$ have the same $k$-long subsequences then either $u, v \in L$ or $u, v \notin L$.

SF Aperiodicity:

A language $L$ is Star-Free iff there exists $n$ such that for all $x, y, z \in \Sigma^*$ and $m > n$ if $xy^nz \in L$ then $xy^mz \in L$.

Reg MyHill/Nerode Theorem:

$L$ is regular iff $|\{T_L(u) \mid u \in \Sigma^*\}|$ is finite, where for all $L \subseteq \Sigma^*, u \in \Sigma^*$, $T_L(u) = \{v \mid uv \in L\}$.

# References

McNaughton, Robert, and Seymour Papert. 1971. *Counter-Free Automata*. MIT Press.

Rogers, James, Jeffrey Heinz, Gil Bailey, Matt Edlefsen, Molly Visscher, David Wellcome, and Sean Wibel. 2010. On languages piecewise testable in the strict sense. In *The Mathematics of Language*, edited by Christian Ebert, Gerhard Jäger, and Jens Michaelis, vol. 6149 of *Lecture Notes in Artifical Intelligence*, 255–265. Springer.

Rogers, James, Jeffrey Heinz, Margaret Fero, Jeremy Hurst, Dakotah Lambert, and Sean Wibel. 2013. Cognitive and sub-regular complexity. In *Formal Grammar*, edited by Glyn Morrill and Mark-Jan Nederhof, vol. 8036 of *Lecture Notes in Computer Science*, 90–108. Springer.

Rogers, James, and Geoffrey Pullum. 2011. Aural pattern recognition experiments and the subregular hierarchy. *Journal of Logic, Language and Information* 20:329–342.

Thomas, Wolfgang. 1997. Languages, automata, and logic. In *Handbook of Formal Languages*, vol. 3, chap. 7. Springer.