

The Chomsky Hierarchy

Software Evolution – L2P1

Dr. Vadim Zaytsev aka @grammarware, February 2021



Unrestricted grammars

Context-sensitive

$\alpha \rightarrow \beta$

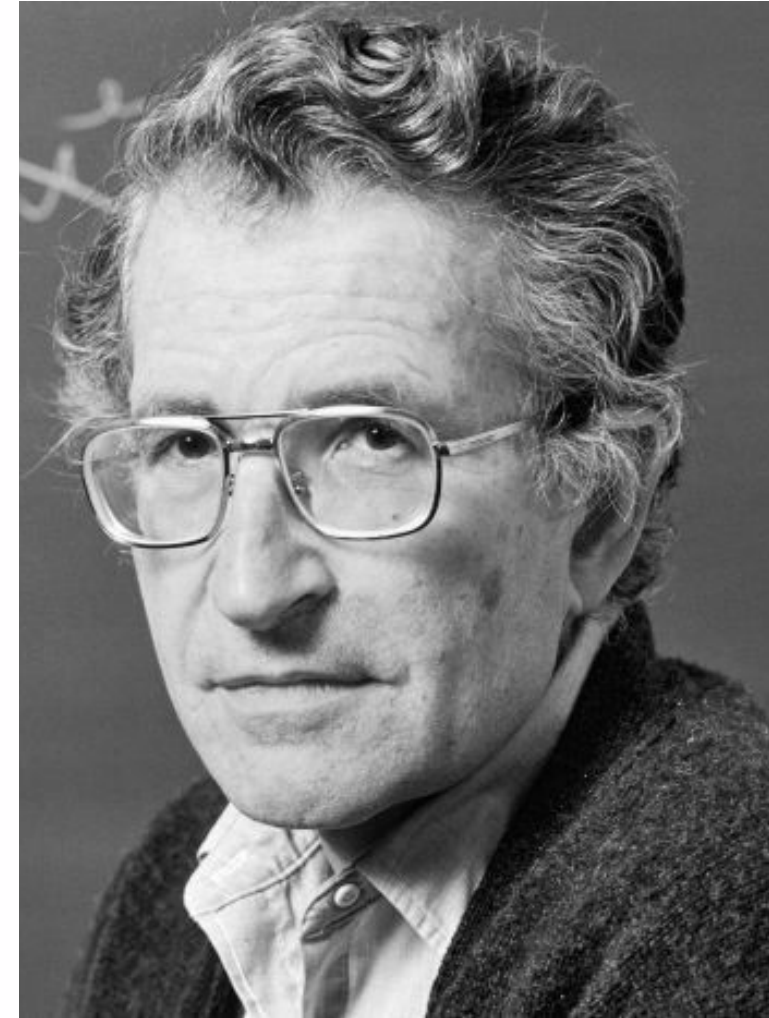
Context-free grammars

$\alpha X \beta \rightarrow \alpha \gamma \beta$

Regular grammars

$X \rightarrow \gamma$

$X \rightarrow a$
 $X \rightarrow aB$



Noam Chomsky. On Certain Formal Properties of Languages. *Journal of Information & Control* 2(2):137-167, 1959.

Unrestricted grammars

Decidable grammars

Context-sensitive

Indexed grammars

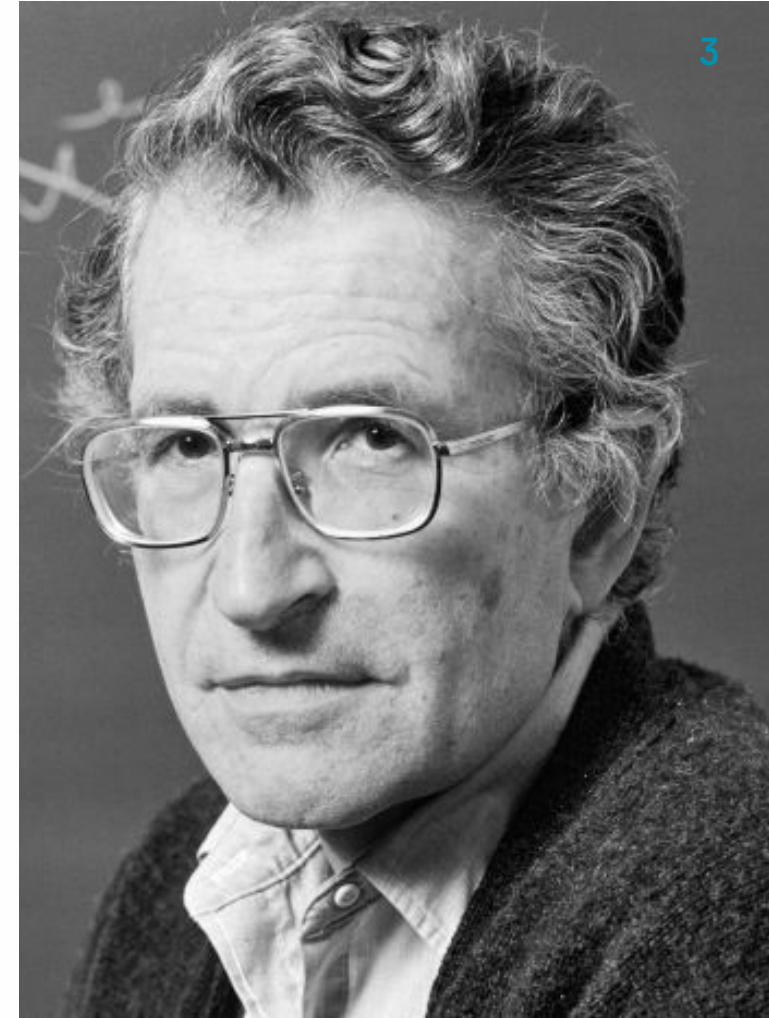
Context-free grammars

Deterministic CFG

Nested word

Regular grammars

Non-recursive grammars



Noam Chomsky. On Certain Formal Properties of Grammars,
Information & Control 2(2):137-167, 1959.

Unrestricted grammars	Recursively enumerable	Turing machine
Decidable grammars	Recursive languages	Terminating automata
Context-sensitive	Context-sensitive	Linear-bounded automata
Indexed grammars	Languages with macros	Nested stack automata
Context-free grammars	Context-free languages	Pushdown automata
Deterministic CFG	Deterministic CFL	Deterministic PDA
Nested word	Nested word	Visibly PDA
Regular grammars	Regular languages	FSMs
Non-recursive grammars	Finite languages	FSMs without cycles

The Finite Language World

- Examples:
 - Boolean values
 - languages
 - countries
 - cities
 - postcodes

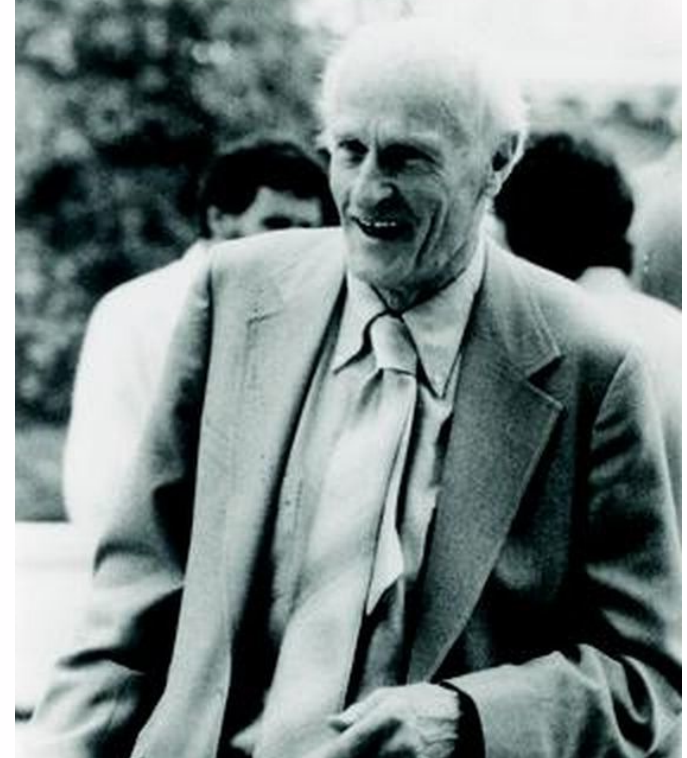


The Finite Language World

- Typical actions
 - union
 - difference
 - intersection
 - element of
 - range

The Regular Language World

- Regular sets by Stephen Kleene in 1956
- \emptyset , ε , letters from Σ
- concatenation
- iteration
- alternation
- Precisely fit the regular class



S. C. Kleene, Representation of Events in Nerve Nets and Finite Automata. In Automata Studies, pp. 3-42, 1956.
photo from: Konrad Jacobs, S. C. Kleene, 1978, MF0.

The Regular Language World

- **PCRE**

- "Perl-compatible regular expressions"
- (not compatible with Perl)
- (not regular)
- C library
- (backrefs, recursion, assertions...)

The Regular Language World

- Typical actions
 - substring search
 - substring replace
 - counting
 - repetition

The Context-Free World

- FSM + memory (stack)
- Modular composition
 - $A ::= "[B]" ;$
 - $B ::= A? ;$
- Forget intersection & diff
- Closed under substitution



on photo: John Backus (1924-2007)

The Context-Free World

- Typical actions
 - recognising structure
 - parsing
 - hierarchy
 - traversal

The Context-Sensitive World

- Explainable only in context
 - Sentence \rightarrow List End
 - List \rightarrow Name;
 - List \rightarrow List “,” Name;
 - “,” Name End \rightarrow “and” Name
- Parsing in exponential time

The Context-Sensitive World

- Typical actions
 - disambiguation based on context

The Unbounded World

- (almost) anything
- recognising is impossible
- parsing is impossible

The Harry Potter Effect

- Gets progressively gloomier
- $\{a^i b^n \dots\}$
- Rule of thumb
 - 0 counters
 - 1 counter
 - n counters
 - ∞ counters



on photo: Walther von Dyck (1856–1934)
by [Zeitlupe](#), CC-BY-SA, 2012

Conclusion

- Finite: profit is minimal
- Regular: linear. Beware of impostors!
- Context-free: hierarchy. Near-linear to cubic.
- Context-sensitive: exponential.
- Many other useful classes! PEG, AG, TGG, TAG, DDG, ...
- Q&A Sessions @ Canvas
 - \Rightarrow v.zaytsev@utwente.nl
 - \Rightarrow <https://discord.gg/n7VQAPNBPD>