322 Midterm Review

- ## Formal Languages
  - Alphabet (Σ)
  - String (Σ^*)
  - Length ($|x|$)
  - Empty String (ε)
  - Empty Language ($\varnothing$)

- ## Language/String Operations
  - "Regular" Operations:
    - Union ($\cup$)
    - Concatenation ($\bullet$)
    - (Kleene) Star (*)
  - Other:
    - Intersection
    - Complement
    - Reversal
    - ...

Finite Defns of Infinite Languages

- English, mathematical
- DFAs
  - States
  - Start states
  - Accept states
  - Transitions (δ function)
  - M accepts w $\in \Sigma^*$
  - M recognizes L $\subseteq \Sigma^*$

- Nondeterminism
- NFAs
  - Transitions (δ relation)
    - Missing out-edges
    - ε-moves
    - Multiple out-edges
  - N accepts w $\in \Sigma^*$
  - N recognizes L $\subseteq \Sigma^*$

- Regular Expressions
  - $\varnothing$, a$\in \Sigma$, $\cup$, $\bullet$, $*$, ( )
- GNFAs

Key Results, Constructions, Methods

- L is regular iff it is:
  - Recognized by a DFA
  - Recognized by a NFA
  - Recognized by a GNFA
  - Defined by a Regular Expr

Proofs:

GNFA → Reg Expr

(Kleene/Floyd/Warshall: $R_{ij} R_{jj}{}^* R_{jk}$)

Reg Expr → NFA

(join NFAs w/ ε-moves)

NFA → DFA

(subset construction)

- The class of regular languages is closed under:
  - Regular ops: union, concatenation, star
  - Also: intersection, complementation, (& reversal, prefix, no-prefix, …)

- NOT closed under $\subseteq$, $\supseteq$

- Also: Cross-product construction (union, …)

Non-Regular Languages

- Key idea: once M is in some state q, it doesn't remember how it got there.

  E.g. "hybrids":
      if $xy \in L(M)$ and
      x, x' both go to q, then
      $x'y \in L(M)$ too.

  E.g. "loops":
      if $xyz \in L(M)$ and
      x, xy both go to q, then
      $xy^iz \in L(M)$ for all $i \geq 0$.

- Cor: Pumping Lemma
- Important examples:

  $L_1 = \{ a^nb^n \mid n > 0 \}$
  $L_2 = \{ w \mid \#_a(w) = \#_b(w) \}$
  $L_3 = \{ ww \mid w \in \Sigma^* \}$
  $L_4 = \{ ww^R \mid w \in \Sigma^* \}$
  $L_5 = \{ \text{balanced parens} \}$

- Also: closure under $\cap$, complementation sometimes useful:
  - $L_1 = L_2 \cap a^*b^*$

- PS: don't say "Irregular"

Applications

- "globbing"
  - lpr *.txt

- pattern-match searching:
  - grep "Ruzzo.*terrific" *.txt

- Compilers:
  - Id ::= letter (letter|digit)*
  - Int ::= digit digit*
  - Float ::=
    d d* . d* (ε | E d d*)