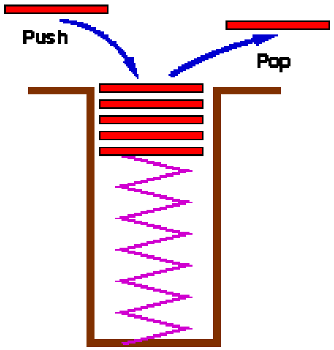
Pushdown Automata (PDA)

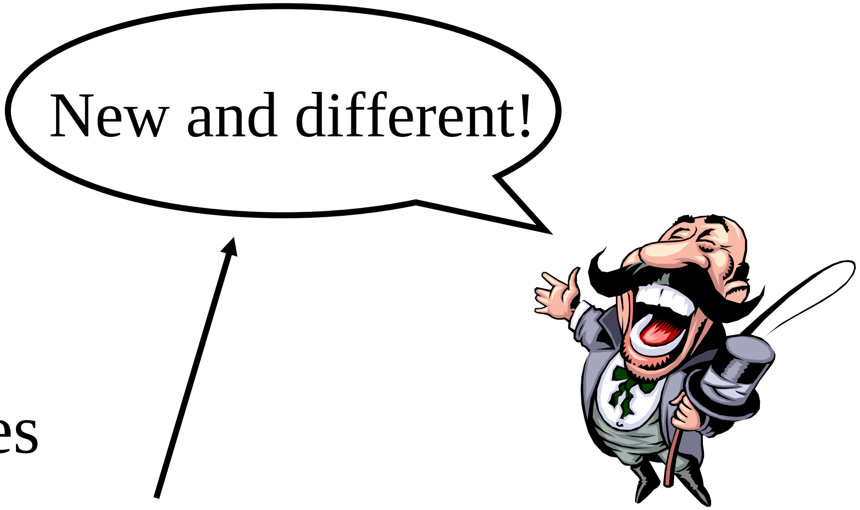
- Main Idea: Add a stack to an NFA
  - Stack provides potentially unlimited memory to an otherwise finite memory machine (finite memory = finite no. of states)
  - PDA = NFA +



  - Stack is LIFO ("Last In, First Out")
  - Two operations:
    - <u>"Push"</u> symbol onto top of stack
    - <u>"Pop"</u> symbol from top of stack

6 Components of a PDA = $(Q, \Sigma, \Gamma, \delta, q_0, F)$

- Q = set of states
- Σ = input alphabet
- **Γ = stack alphabet**
- q_0 = start state
- $F \subseteq Q$ = set of accept states
- **Transition function $\delta: Q \times \Sigma_\varepsilon \times \Gamma_\varepsilon \rightarrow Pow(Q \times \Gamma_\varepsilon)$**
  - **(current state, next input symbol, popped symbol) $\rightarrow$ {set of (next state, pushed symbol)}**
  - **Input/popped/pushed symbol can be ε**



When does a PDA accept a string?

- A PDA M accepts string $w = w_1 w_2 \ldots w_m$ if and only if there exists <u>at least one accepting computational path</u> i.e. a sequence of states $r_0, r_1, \ldots, r_m$ and strings $s_0, s_1, \ldots, s_m$ (denoting stack contents) such that:

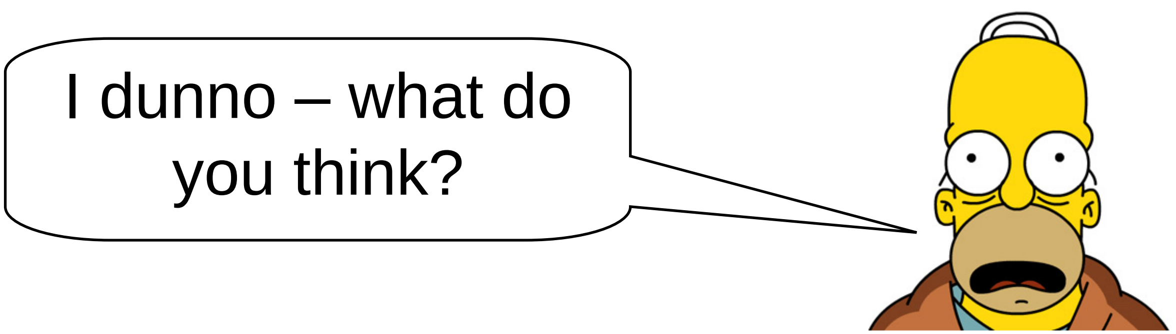
1. $r_0 = q_0$ and $s_0 = \varepsilon$ *(M starts in q_0 with empty stack)*
2. $(r_{i+1}, b) \in \delta(r_i, w_{i+1}, a)$ for some $a, b \in \Gamma_\varepsilon$ *(States follow transition rules)*
3. $s_i = at$ and $s_{i+1} = bt$ for some $t \in \Gamma^*$
   (M pops "a" from top of stack and pushes "b" onto stack)
4. $r_m \in F$ *(Last state in the sequence is an accept state)*

On-Board Examples

- PDA for $L = \{w\#w^R \mid w \in \{0,1\}^*\}$ (# acts as a "delimiter")
  - E.g. 0#0, 1#1, 10#01, 01#10, 1011#1101 $\in$ L
  - L is a CFL (what is a CFG for it?)
  - Recognizing L using a PDA:
    - Push each symbol of w onto stack
    - On reaching # (middle of the input), pop the stack – this yields symbols in w^R – and compare to rest of input
- PDA for $L_1 = \{ww^R \mid w \in \{0,1\}^*\}$
  - Set of all even length palindromes over {0,1}
- Recognizing L_1 using a PDA:
  - <u>Problem</u>: Don't know the middle of input string
  - <u>Solution</u>: Use nondeterminism (ε-transition) to guess!
  - See lecture notes on class website

Are context free grammars equivalent to PDAs?

(i.e. Are the languages generated by CFGs the same as the languages recognized by PDAs?)



From CFGs to PDAs

- L is context free $\Rightarrow$ there exists a PDA that accepts it
- Proof idea:
  PDA "simulates" context-free grammar (CFG) for L by:
  1. Nondeterministically generating strings (in parallel) using rules of the CFG starting from the start symbol,
  2. Using the stack to store each intermediate string,
  3. Checking the generated part of each string with the input string in an "on-line" manner, and
  4. Going to the accept state if and only if all characters of the generated string match the input string.

From CFGs to PDAs: Details

- L is a CFL $\Rightarrow$ L = L(M) for some PDA M
- Proof Summary:
  - L is a CFL means L = L(G) for some CFG G = (V, Σ, R, S)
  - Construct PDA M = (Q, Σ, Γ, δ, q_0, $\{q_{acc}\}$)
    M has only 4 main states (plus a few more for pushing strings)
    Q = $\{q_0, q_1, q_2, q_{acc}\} \cup E$ where E are states used in 2 below
  - δ has 4 components:

  1. **Init. Stack**: $\delta(q_0, \varepsilon, \varepsilon) = \{(q_1, \$)\}$ and $\delta(q_1, \varepsilon, \varepsilon) = \{(q_2, S)\}$
  2. **Push & generate strings**: $\delta(q_2, \varepsilon, A) = \{(q_2, w)\}$ for $A \rightarrow w$ in R
  3. **Pop & match to input**: $\delta(q_2, a, a) = \{(q_2, \varepsilon)\}$ for all a in Σ
  4. **Accept if stack empty**: $\delta(q_2, \varepsilon, \$) = \{(q_{acc}, \varepsilon)\}$
- Can prove by induction: $w \in L$ iff $w \in L(M)$

Relationship to Compilers and Parsing

- The PDA in the previous proof is doing what a *compiler* does to your Java/C program: **parsing** an input string based on a grammar G
- This type of parsing is called "top-down" or LL parsing (Left to right parse, Leftmost derivation)
- For details and an example implementation, see: http://en.wikipedia.org/wiki/LL_parser
(they even use $ to represent their end of stack!)

Can PDAs be converted to CFGs??

From PDAs to CFGs

- L = L(M) for some PDA M $\Rightarrow$ L = L(G) for some CFG G
- Proof Summary: Simulate M's computation using a CFG
  - First, simplify M: 1. Only 1 accept state, 2. M empties stack before accepting, 3. Each transition either Push or Pop, not both or neither.
  - Let this M = $(Q, \Sigma, \Gamma, \delta, q_0, \{q_{acc}\})$
  - Construct grammar G = (V, Σ, R, S)

From PDAs to CFGs

- Construct grammar $G = (V, \Sigma, R, S)$

Basic Idea: Define variables A_{pq} for simulating M

A_{pq} generates all strings w such that w takes M from state p with empty stack to state q with empty stack

Then, A_{q0qacc} generates all strings w accepted by M

Review: From PDAs to CFGs (cont.)

- $L = L(M)$ for some PDA $M \Rightarrow L = L(G)$ for some CFG G
- Proof (cont.)
  - Construct grammar $G = (V, \Sigma, R, S)$ where

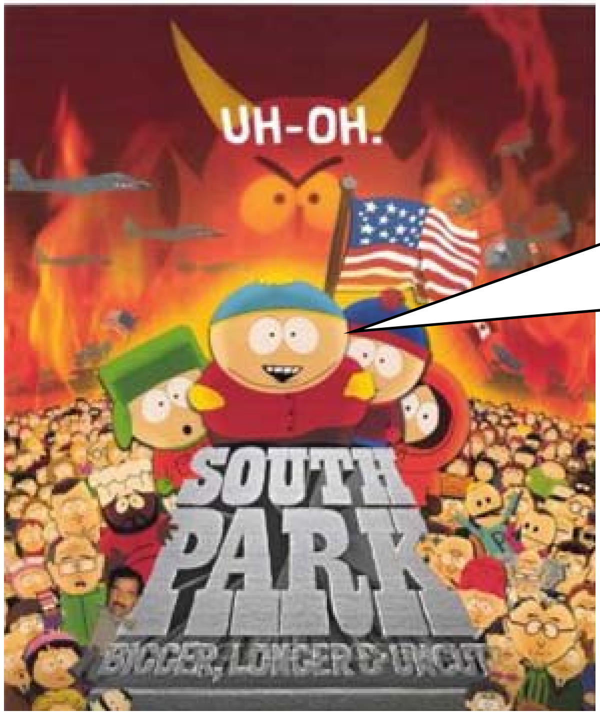
    $V = \{A_{pq} \mid p, q \in Q)$

    $S = A_{q0qacc}$

    $R = \{A_{pq} \to aA_{rs}b \mid \; p \xrightarrow{a,\, \varepsilon \to c} r \overset{A_{rs}}{\dashrightarrow} s \xrightarrow{b,\, c \to \varepsilon} q\}$

    $\cup \{A_{pq} \to A_{pr} A_{rq} \mid p, q, r \in Q\}$

    $\cup \{A_{qq} \to \varepsilon \mid q \in Q\}$
- See textbook for details and proof: $w \in L(M)$ iff $w \in L(G)$
- Try to get G from M where $L(M) = \{0^n1^n \mid n \geq 1\}$



Next class: Return of the Pumping Lemma (bigger and better)