# CFL closure under complement?

- Show that $\overline{L_{\text{double}}} = \{y \in \{a,b\}^* \mid \forall x \in \{a,b\}^*, y \neq x.x\}$ is a CFL

- Note: $\overline{L_{\text{double}}} = L_{\text{odd-length}} \cup L'$ where $L'$ is generated by this CFG:

$$S \rightarrow AB \mid BA$$
$$A \rightarrow aAa \mid aAb \mid bAa \mid bAb \mid a$$
$$B \rightarrow aBa \mid aBb \mid bBa \mid bBb \mid b$$

# CFL closure under intersection?

- **Claim 1**: $L_1 = \{a^n.b^n.a^m \mid m,n \geq 0\}$ is a CFL
- **Claim 2**: $L_2 = \{a^m.b^n.a^n \mid m,n \geq 0\}$ is a CFL
- **Claim 3**: $L_1 \cap L_2 = \{a^n.b^n.a^n \mid n \geq 0\}$ is NOT a CFL
  - Proof using Pumping Lemma
- **Claim 4**: If $L_1$ is a CFL and $L_2$ is regular, then $L_1 \cap L_2$ is a CFL
- **Proof sketch**: Let $G = (N, A, S, P)$ be a CFG in Chomsky Normal Form for $L_1$ and let $M = (Q, s, \delta, F)$ be a DFA for $L_2$
- Let $G' = (N \times Q \times Q \cup \{S_0\}, A, S_0, P')$ ; $P' = \{S_0 \rightarrow (S,s,q) | q \in F\} \cup$

$$\{(X,q,q') \rightarrow (Y,q,p)(Z,p,q') | X \rightarrow YZ \in P, p,q,q' \in Q\}$$
$$\cup \{(X,q,q') \rightarrow a | X \rightarrow a \in P, \delta(q,a) = q'\}$$

Handle ε separately

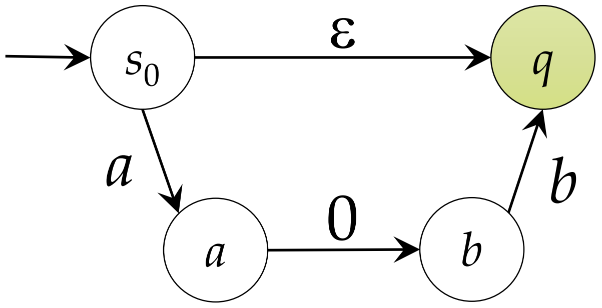
# Automata for CFLs?

- Automata model control flow in programs naturally, so it is useful to have automata models for CFLs
- Recursive Automata (RA)
  - NFAs which can call each other (potentially recursively, implicit call stack)
  - RA $\leftrightarrow$ CFG conversion is trivial
- Pushdown Automata (PDA)
  - Classical model, explicit stack
- Visibly Pushdown Automata (VPA)
  - Restricted PDA + nice properties (e.g., DVPL = VPL $\subset$ DCFL)

Deterministic versions are strictly less powerful
DCFL $\subset$ CFL

# Recursive Automata

- Models programs with finite memory + unlimited recursion
- A recursive automaton over an alphabet $A$ is a non-empty set of NFA $\{N_i = (Q_i, s_i, \delta_i, F_i) \mid i = 0, \dots, k\}$ and each $\delta_i \subseteq Q_i \times \hat{A} \times Q_i$ where
$$\hat{A} = A \cup \{\varepsilon\} \cup \{0,1,\dots,k\}$$
- The RA starts from the initial state of NFA $N_0$ (the **`main`** NFA)
- *Example*:




This RA accepts the language
$\{a^n . b^n | n \geq 0\}$

# Formal definition of acceptance

- A stack $t \in Q^*$ lists the return states for incomplete recursive calls (where $Q = \bigcup_{i=0}^{k} Q_i$)
- A configuration $c$ is a pair $(q, t) \in Q \times Q^*$
  - Initial configuration is $c_0 = (s_0, \varepsilon)$
- We say that configuration $c$ can go to $c'$ on reading $e \in A \cup \{\varepsilon\}$ ($c \xrightarrow{e} c'$) if:
  - **Internal**: $c = (q, t)$ ; $c' = (q', t)$ ; $q, q' \in Q_i$ ; $(q, e, q') \in \delta_i$
  - **Call**: $e = \varepsilon$ ; $c = (q, t)$ ; $c' = \left(q_j, t.q'\right)$ ; $q, q' \in Q_i$ ; $(q, j, q') \in \delta_i$
  - **Return**: $e = \varepsilon$ ; $c = (q, t.q')$ ; $c' = (q', t)$ ; $q \in F_i$

  Extend $\xrightarrow{e}$ to strings
- The RA accepts the language $\left\{w \in A^* \middle| c_0 \xrightarrow{w} (q, \varepsilon) \text{ for some } q \in F_0\right\}$

# RA ↔ CFG equivalence (sketch)

- Let $G = (N, A, X_0, P)$ be a CFG where $N = \{X_i | i = 0, \dots, k\}$
- The equivalent RA is the set of NFAs $\{N_i | i = 0, \dots, k\}$ where each $N_i$ accepts the regular language $L_i = \left\{\hat{\alpha} \in \hat{A}^* \middle| X_i \to \alpha \in P\right\}$ where $\hat{\alpha}$ is obtained from $\alpha$ by replacing each $X_i$ with $i$

- Given an RA $R = \{N_i | i = 0, \dots, k\}$ where $Q = \bigcup_i Q_i$, the equivalent CFG is $G = \left(\left\{X_q \middle| q \in Q\right\}, A, X_{q_0}, P\right)$ where $P$ is generated as follows:

**Internal**: if $(p, e, q) \in \delta_i$ then $X_p \to e.X_q \in P$

**Call**: if $(p, j, q) \in \delta_i$ then $X_p \to X_{q_j}.X_q \in P$

**Return**: if $p \in F_i$ then $X_p \to \varepsilon \in P$

# Example

- Construct an RA for the CFG with these rules:

$$S \to AB \mid BA$$
$$A \to aAa \mid aAb \mid bAa \mid bAb \mid a$$
$$B \to aBa \mid aBb \mid bBa \mid bBb \mid b$$