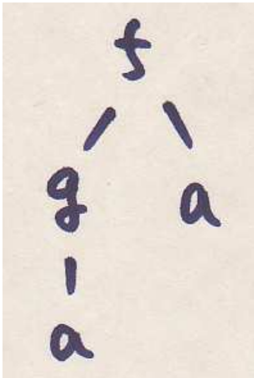
Formal Languages and Automata Theory

Terms and Trees

- **Set** $\mathcal{F}$ **of function symbols with rank:**
  Number arity **of arity:** $\mathcal{F} \to N$
  $\mathcal{F} = \{f(,), g(), a\}$
- **Set** $\mathcal{X}$ **of variables Especially** $\mathcal{X}_n = \{x_1, \ldots, x_n\}$
- **Set** $\mathsf{T}(\mathcal{F}, \mathcal{X})$ **of terms: Minimal set satisfying**
  $x \in \mathsf{T}(\mathcal{F}, \mathcal{X})$ **...** $x \in \mathcal{X}$
  $a \in \mathsf{T}(\mathcal{F}, \mathcal{X})$ **...** $a \in \mathcal{F}$ **and** $\text{arity}(a) = 0$
  $f(t_1, \ldots, t_n) \in \mathsf{T}(\mathcal{F}, \mathcal{X})$
  ... $f \in \mathcal{F}$**,** $\text{arity}(f) = n$**, each** $t_i \in \mathsf{T}(\mathcal{F}, \mathcal{X})$
  Ex. of terms: $f(g(a), a)$
- **Set** $\mathsf{T}(\mathcal{F})$ **of ground terms:** $\mathsf{T}(\mathcal{F}, \emptyset)$

- **Labelled tree**:
  Partial function $t : N^* \to \mathcal{F} \cup \mathcal{X}$
  $\mathsf{Pos}(t)$ **denotes the domain of** t
- **Tree** t **that represents** $f(g(a), a)$
  $\mathsf{Pos}(t) = \{\varepsilon, 1, 11, 2\}$

p	ε	1	11	2
$t(p)$	f	g	a	a



- **Height of tree** t**:** $|t| = \max\{|p| \mid p \in \mathsf{Pos}(t)\}$

- **Substitution**: **Function** $\sigma : \mathcal{X} \to \mathrm{T}(\mathcal{F}, \mathcal{X})$
  - **Domain of** σ: $\mathrm{Dom}(\sigma) = \{x \in \mathcal{X} \mid \sigma(x) \neq x\}$
  - **Notation**: $\{x_1 \leftarrow t_1, \ldots, x_n \leftarrow t_n\}$ **for** σ **where** $\mathrm{Dom}(\sigma) = \{x_1, \ldots, x_n\}$, $\sigma(x_i) = t_i$
  - **Naturally extended on terms**
    $\sigma(f(t_1, \ldots, t_n)) = f(\sigma(t_1), \ldots, \sigma(t_n)) \quad (n \geq 0)$
- **Context**: **Linear term** $C \in \mathrm{T}(\mathcal{F}, \mathcal{X})$
  - **For a context with** n **variables, assume variables are** $x_1, \ldots, x_n$ **from left to right**
    For $C = f(g(x_1), x_2)$, $C[a, g(b)] = f(g(a), g(b))$
  - **For a context** C **with** n **variables,** $C[t_1, \ldots, t_n]$ **represents** $C\{x_1 \leftarrow t_1, \ldots, x_n \leftarrow t_n\}$

Finite Tree Automata

- **(Bottom-up) non-deterministic finite tree automata (NFTA)** $A = (Q, \mathcal{F}, Q^f, \Delta)$
  - Q: Finite set of states
  - $\mathcal{F}$: Finite set of function symbols
  - Q^f ($\subseteq Q$): Set of finial states
  - Δ: Set of transition rules

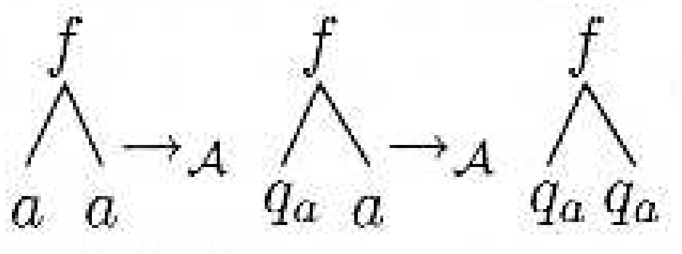
    Form of transition rules:

    $f(q_1, \ldots, q_n) \to q$

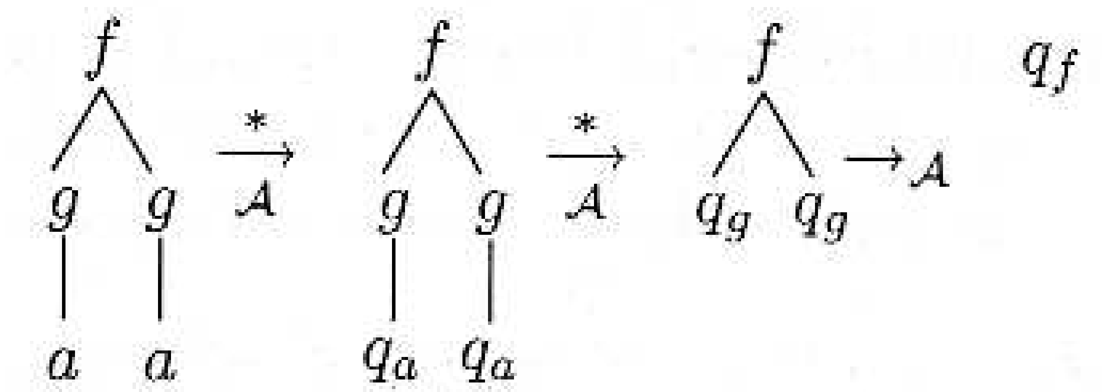
    where $f \in \mathcal{F}$, arity$(f) = n$, $q, q_i \in Q$

- **Ex.:** $A = (\{q_a, q_g, q_f\}, \{a, g(), f(,)\}, \Delta, \{q_f\})$
  where Δ consists of $a \to q_a$, $g(q_a) \to q_g$, $g(q_g) \to q_g$, $f(q_g, q_g) \to q_f$

  - Run for $f(a, a)$



  - Run for $f(g(a), g(a))$



- Transition relation $\rightarrow_A$: Minimal set that satisfies
  - $\Delta \subseteq \rightarrow_A$
  - $s \rightarrow_A t$ implies $f(\cdots s \cdots) \rightarrow_A f(\cdots t \cdots)$
- $\rightarrow_A^*$: Reflexive and transitive closure of $\rightarrow_A$
- A accepts t:

  $t \rightarrow_A^* q \in Q^f$
- Language recognized by A:

  $L(A) = \{t \mid t \rightarrow_A^* q \in Q^f\}$
- Regular tree languages: Languages recognizable by NFTA

- **Ex.:** $A = (\{q, q_g, q_f\}, \{a, g(), f(,)\}, \Delta, \{q_f\})$

  $a \to q$, $g(q) \to q$, $g(q) \to q_g$,

  $g(q_g) \to q_f$, $f(q, q) \to q$

  - **Run for** $g(g(f(g(a), a)))$

    $g(g(f(g(a), a))) \to_A^* g(g(f(q_g, q)))$

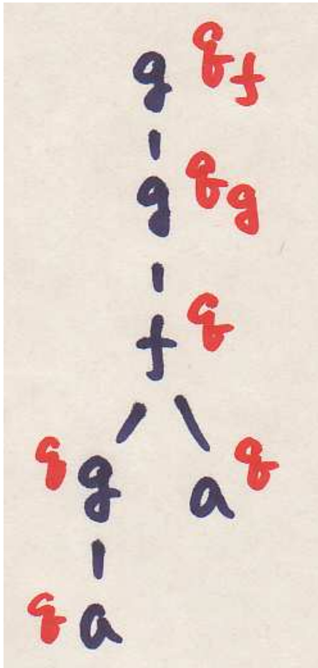
    $g(g(f(g(a), a))) \to_A^* g(g(q)) \to_A^* q$

    $g(g(f(g(a), a))) \to_A^* g(g(q)) \to_A^* q_f$

  - $L(A) = \{g(g(t)) \mid t \in \mathsf{T}(\mathcal{F})\}$

- **Ex.(cont.):**
  - **Alternative representation of run**

    Δ: $a \to q$, $g(q) \to q$, $g(q) \to q_g$, $g(q_g) \to q_f$, $f(q, q) \to q$



- **Q1-1: Let** $\mathcal{F} = \{f(,), g(), a\}$**. Show NFTA that recognizes** $L_1 = \{g(f(a, s)) \mid s \in \mathsf{T}(\mathcal{F})\}$
- **Q1-2: Let** $\mathcal{F} = \{f(,), a, b\}$**. Show NFTA that recognizes** L_2 **defined below:**
  - $a \in L_2$
  - $s \in L_2$ **implies** $f(f(a, s), b) \in L_2$

- **NFTA with ε (ε-NFTA):**
  NFTA which also contains ε-rule in forms of $q \to q'$
- **Deterministic FTA (DFTA):**
  NFTA in which each pair of rules has different lefthand-side (Moreover no ε-rules)
- **Complete NFTA:**
  For any ground term t, $\exists q.\, t \to_A^* q$

- **Ex. of complete DFTA:**

  $A = (\{q_0, q_1\}, \{0, 1, not(), and(,), or(,)\}, \Delta, \{q_1\})$

  $0 \rightarrow q_0$, $1 \rightarrow q_1$, $not(q_0) \rightarrow q_1$, $not(q_1) \rightarrow q_0$,
  $and(q_0, q_0) \rightarrow q_0$, $and(q_0, q_1) \rightarrow q_0$,
  $and(q_1, q_0) \rightarrow q_0$, $and(q_1, q_1) \rightarrow q_1$,
  $or(q_0, q_0) \rightarrow q_0$, $or(q_0, q_1) \rightarrow q_1$,
  $or(q_1, q_0) \rightarrow q_1$, $or(q_1, q_1) \rightarrow q_1$

  - **Run for** $and(not(or(0,1)), or(1, not(0)))$