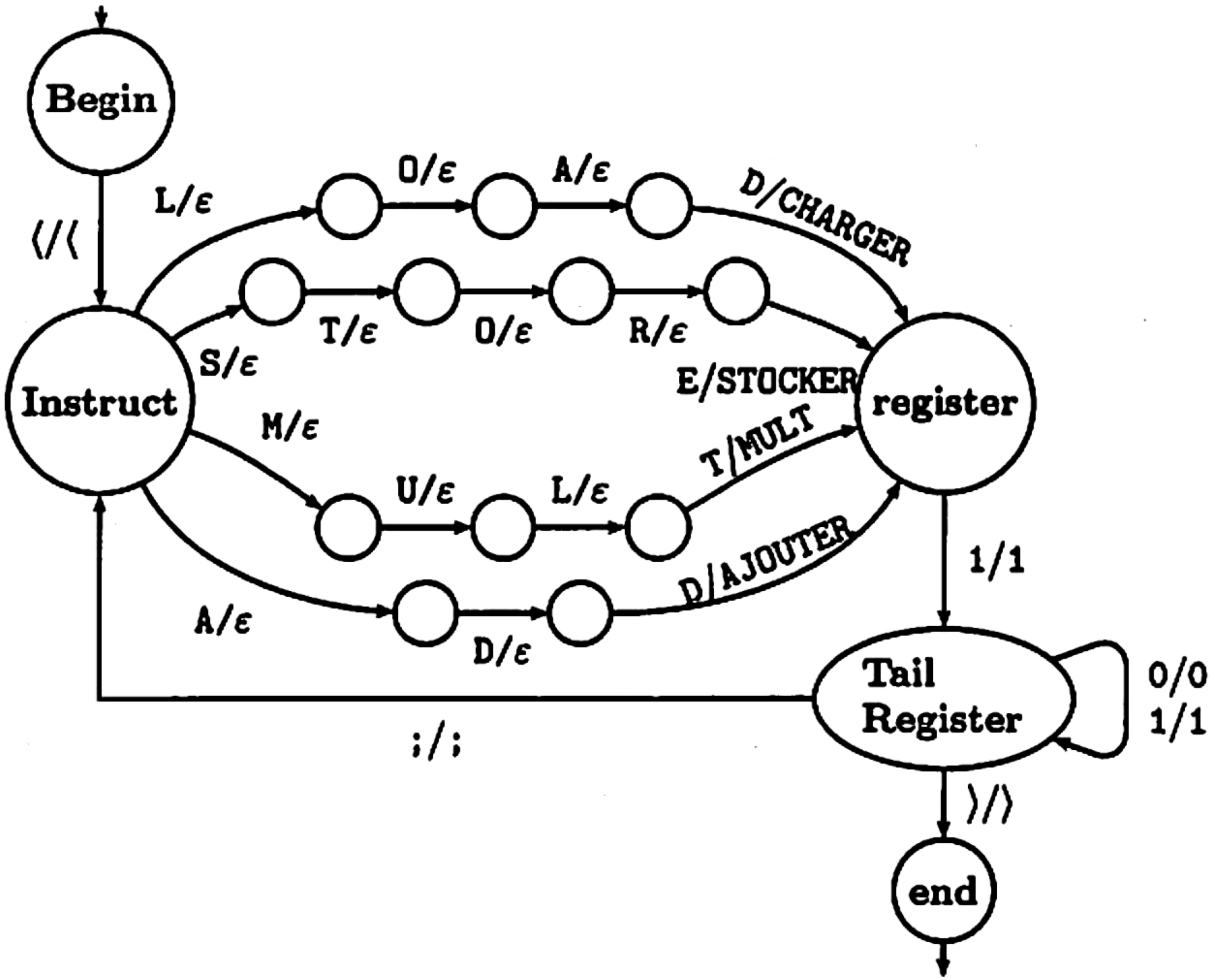
木変換系 (Tree Transducers)

- 木変換系：木構造を利用した変換系
  例：LaTeX から HTML への変換
  - ボトムアップ木変換系
  - トップダウン木変換系
  - 準同型写像による木変換系

文字列変換系

- **Rational Transducers** (ミーリー機械)



- **Rational Transducers:** $R = (Q, \mathcal{F}, \mathcal{F}', Q_i, Q_f, \Delta)$
  - $\mathcal{F}$ ($\mathcal{F}'$): 入力(出力)の記号の集合
  - Q_i (Q_f): 開始(終了)状態の集合
  - Δ : $q \stackrel{f/m}{\rightarrow} q'$ の形式の規則集合

    ここで、$f \in \mathcal{F} \cup \{\varepsilon\}$, $m \in \mathcal{F}'^*$, $q, q' \in Q$
- 遷移関係$\rightarrow_R$ ($\subseteq \mathcal{F}^* \times Q \times \mathcal{F}'^*$)

  $(ft, q, u) \rightarrow_R (t, q', um)$ for $q \stackrel{f/m}{\rightarrow} q' \in \Delta$
- Rが受理する関係T_R:

  $T_R = \{(t, u) \mid (t, q, \varepsilon) \rightarrow_R^* (\varepsilon, q', u), q \in Q_i, q' \in Q_f\}$

- 二つの準同型写像による定義：
  - $B = (\Phi, L, \Psi)$ ここで、

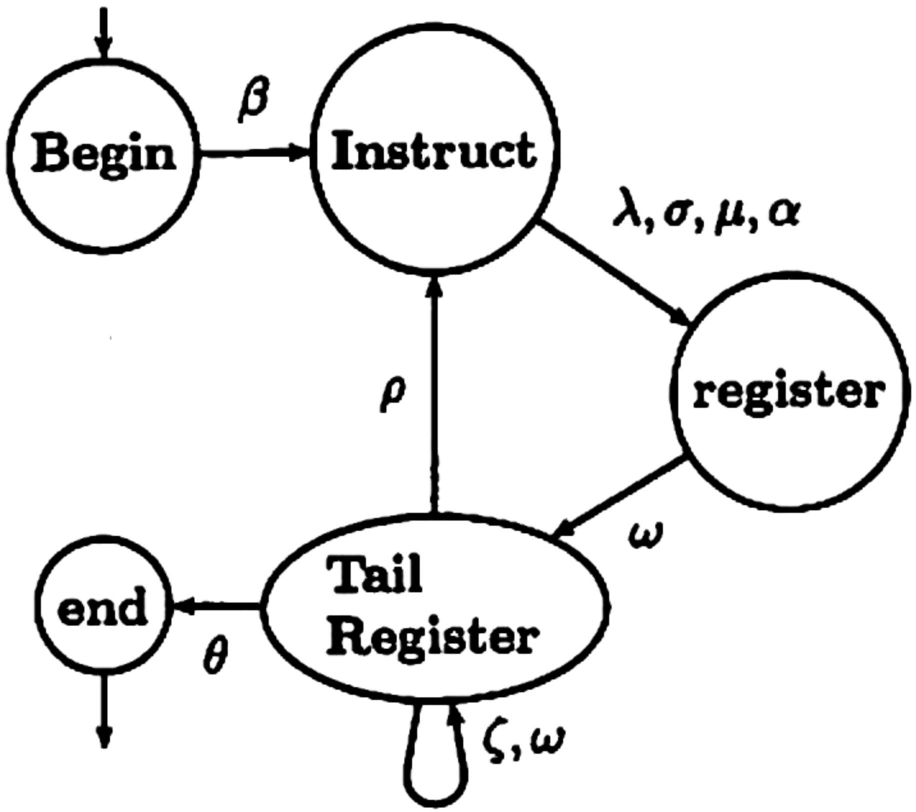
    $\Phi : \mathcal{F}''^* \to \mathcal{F}^*$

    $\Psi : \mathcal{F}''^* \to \mathcal{F}'^*$

    L **(**$\subseteq \mathcal{F}''^*$**):** 正規言語
  - $T_B = \{(\Phi(w), \Psi(w)) \mid w \in L\}$
- εなし：$\Phi(a) = \varepsilon$となる$a \in \mathcal{F}''$が存在しない

- 二つの準同型写像による**Transducer**の例：

$$
\begin{array}{llll}
\Phi(\beta) = \langle & \Phi(\lambda) = \texttt{LOAD} & \Phi(\sigma) = \texttt{STORE} & \Phi(\mu) = \texttt{MULT} \\
\Phi(\alpha) = \texttt{ADD} & \Phi(\rho) =; & \Phi(\omega) = 1 & \Phi(\zeta) = 0 \\
\Phi(\theta) =\rangle & & & \\
\Psi(\beta) = \langle & \Psi(\lambda) = \texttt{CHARGER} & \Psi(\sigma) = \texttt{STOCKER} & \Psi(\mu) = \texttt{MULT} \\
\Psi(\alpha) = \texttt{ADD} & \Psi(\rho) =; & \Psi(\omega) = 1 & \Psi(\zeta) = 0 \\
\Psi(\theta) =\rangle & & &
\end{array}
$$



木変換系

- 例題

  $(a+b)\times c$ の G の構文木を $\times(+(a,b),c)$ のように変換

- 算術式の一般的な文法 (構文解析には向かない)

$$
\begin{array}{lcl}
E & \rightarrow & M \mid M+E \\
M & \rightarrow & F \mid F\times M \\
F & \rightarrow & I \mid (E) \\
I & \rightarrow & a \mid b \mid \cdots \mid z
\end{array}
$$

- **LL(1)** 構文解析可能な文法 G

$$
\begin{array}{lcl}
E & \rightarrow & ME' \\
E' & \rightarrow & +E \mid \varepsilon \\
M & \rightarrow & FM' \\
M' & \rightarrow & \times M \mid \varepsilon \\
F & \rightarrow & I \mid (E) \\
I & \rightarrow & a \mid b \mid \cdots \mid z
\end{array}
$$

- ボトムアップ木変換系: **NUTT**

  $A = (Q, \mathcal{F}, \mathcal{F}', Q^f, \Delta)$、ここで$\Delta$は以下の形式の規則

  $$f(q_1(x_1), \ldots, q_n(x_n)) \to q(u)$$

  $$f \in \mathcal{F}, u \in \mathsf{T}(\mathcal{F}', \mathcal{X}_n)$$

  - 遷移関数Aの定義：

    $$s \to_A t \quad \overset{\text{def}}{\Longleftrightarrow}$$

    $$\exists C, \sigma, l \to r \in \Delta.\ s = C[l\sigma] \wedge t = C[r\sigma]$$

  - 変換の定義： $T \subseteq \mathsf{T}(\mathcal{F})$ に対して、

    $$A(T) = \{s \in \mathsf{T}(\mathcal{F}') \mid t \in T,\ t \to_A^* q(s),\ q \in Q^f\}$$

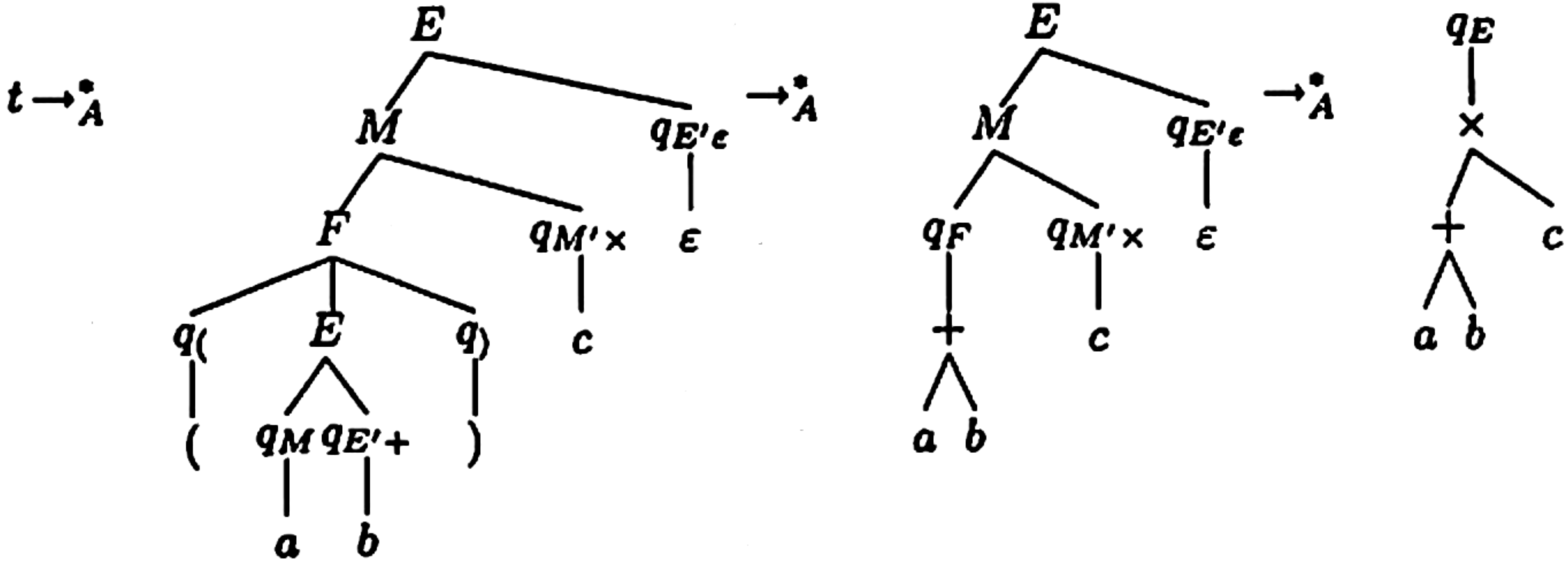
  - 線形(**linear**)：右辺の変数が線形(コピーなし)
  - 非消去(**non-erasing**)：右辺に$\mathcal{F}'$の記号が必ず出現
  - 非削除(**non-deleting**)：左辺の変数は右辺にも出現

- 例：受理状態q_E

$$
\begin{aligned}
a &\longrightarrow q(a) \\
c &\longrightarrow q(c) \\
) &\longrightarrow q_{)}()) \\
+ &\longrightarrow q_{+}(+) \\
I(q(x)) &\longrightarrow q_I(x) \\
M'(q_\varepsilon(x)) &\longrightarrow q_{M'\varepsilon}(x) \\
M(q_F(x), q_{M'\varepsilon(y)}) &\longrightarrow q_M(x) \\
M'(q_\times(x), q_M(y)) &\longrightarrow q_{M'\times}(y) \\
E'(q_+(x), q_E(y)) &\longrightarrow q_{E'+}(y) \\
F(q_{(}(x), q_E(y), q_{)}(z)) &\longrightarrow q_F(y)
\end{aligned}
$$

$$
\begin{aligned}
b &\longrightarrow q(b) \\
\varepsilon &\longrightarrow q_\varepsilon(\varepsilon) \\
(&\longrightarrow q_{(}(() \\
\times &\longrightarrow q_\times(\times) \\
F(q_I(x)) &\longrightarrow q_F(x) \\
E'(q_\varepsilon(x)) &\longrightarrow q_{E'\varepsilon}(x) \\
E(q_M(x), q_{E'\varepsilon}(y)) &\longrightarrow q_E(x) \\
M(q_F(x), q_{M'\times}(y)) &\longrightarrow q_M(\times(x,y)) \\
E(q_M(x), q_{E'+}(y)) &\longrightarrow q_E(+(x,y))
\end{aligned}
$$

- ボトムアップ木変換系(続き)



- 例 U_1: 非消去、非削除

  $\mathcal{F} = \{f(), a\}$, $\mathcal{F}' = \{g(,), f(), f'(), a\}$, $q' \in Q^f$

$$
\begin{array}{ll}
a \to q(a) & f(q(x)) \to q(f(x)) \\
f(q(x)) \to q(f'(x)) & f(q(x)) \to q'(g(x,x))
\end{array}
$$

  $U_1(\{f(f(f(a)))\})$は以下の項からなる集合

$$
\begin{array}{l}
g(f(f(a)), f(f(a))),\ \ g(f(f'(a)), f(f'(a))), \\
g(f'(f(a)), f'(f(a))),\ \ g(f'(f'(a)), f'(f'(a)))
\end{array}
$$

- トップダウン木変換系: **NDTT** $A = (Q, \mathcal{F}, \mathcal{F}', Q^i, \Delta)$、ここで Δ は以下の形式の規則

$$q(f(x_1, \ldots, x_n)) \to C[q_1(x_{i_1}), \ldots, q_p(x_{i_p})]$$
$$f \in \mathcal{F}, C \in \mathsf{T}(\mathcal{F}', \mathcal{X}_p), x_{i_1}, \ldots, x_{i_p} \in \mathcal{X}_n$$

- 遷移関数 A の定義：

$$s \to_A t \quad \stackrel{\text{def}}{\Longleftrightarrow}$$
$$\exists C, \sigma, l \to r \in \Delta.\ s = C[l\sigma] \wedge t = C[r\sigma]$$

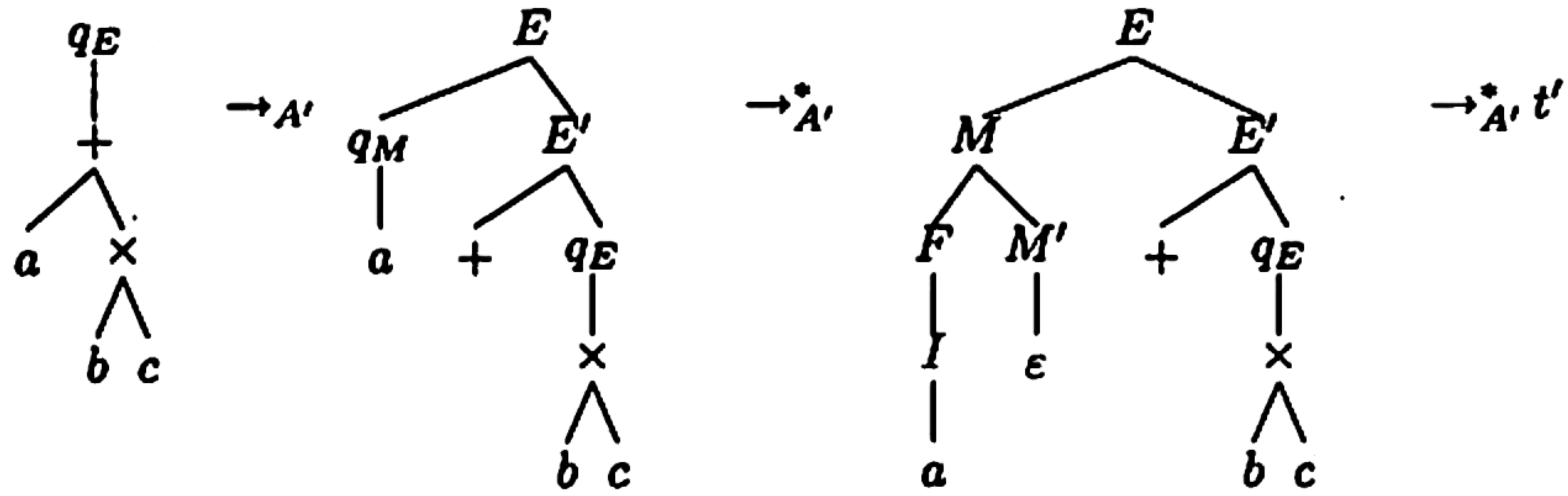
- 変換の定義： $T \subseteq \mathsf{T}(\mathcal{F})$ に対して、

$$A(T) = \{s \in \mathsf{T}(\mathcal{F}') \mid t \in T,\ q \in Q^i,\ q(t) \to_A^* s\}$$

- 例：初期状態q_E

$$
\begin{aligned}
q_E(x) &\to E(q_M(x), E'(\varepsilon)) & q_E(+(x,y)) &\to E(q_M(x), E'(+, q_E(y))) \\
q_M(x) &\to M(q_F(x), M'(\varepsilon)) & q_M(\times(x,y)) &\to M(q_F(x), M'(\times, q_M(y))) \\
q_F(x) &\to F((, q_E(x),)) & q_F(a) &\to F(I(a)) \\
q_F(b) &\to F(I(b)) & q_F(c) &\to F(I(c))
\end{aligned}
$$

- トップダウン木変換系(続き)



変換前の木 $s \in \mathrm{T}(\mathcal{F}')$，変換して得られる木 $t' \in \mathrm{T}(\mathcal{F})$

$$q_E(s) \to^*_{A'} t'$$

- 例 D_1:

  $\mathcal{F} = \{f(), a\}$, $\mathcal{F}' = \{g(,), f(), f'(), a\}$, $q \in Q^i$

$$\begin{array}{ll} q(f(x)) \to g(q'(x), q'(x)) & q'(f(x)) \to f(q'(x)) \\ q'(a) \to a & q'(f(x)) \to f'(q'(x)) \end{array}$$

  $D_1(\{f(f(f(a)))\})$は以下の**16**個の項からなる集合

$$\begin{array}{l} g(f(f(a)), f(f(a))),\ g(f(f(a)), f(f'(a))), \\ g(f(f(a)), f'(f(a))),\ g(f(f(a)), f'(f'(a))), \\ \vdots \\ g(f'(f'(a)), f'(f(a))),\ g(f'(f'(a)), f'(f'(a))) \end{array}$$

- クラス間の性質
  - **NUTT**のクラスと**NDTT**のクラスは異なる**(**U_1とD_1**)**
  - 線形**NDTT**のクラスは、線形**NUTT**に含まれる
  - 線形かつ非削除なら、**NDTT**と**NUTT**は等価
- 合成に関する性質
  - **NUTT**、**NDTT**共に、閉じていない
  - 線形**NUTT**は、閉じている
  - 決定的な**NUTT**は、閉じている
  - 決定的な**NDTT**の合成クラスは、非削除**NDTT**と線形準同型写像の合成クラスと等しい
- 木変換系の定義域は正規木言語
- 正規木言語の線形木変換による像は正規木言語

- 準同型写像による定義：文字列変換系と同じ
  - $B = (\Phi, L, \Psi)$ ここで、

    $\Phi : \mathsf{T}(\mathcal{F}'') \to \mathsf{T}(\mathcal{F})$

    $\Psi : \mathsf{T}(\mathcal{F}'') \to \mathsf{T}(\mathcal{F}')$

    L **(**$\subseteq \mathsf{T}(\mathcal{F}'')$**):** 正規木言語
- **NUTT** U のクラスは、準同型写像による木変換

  $B = (\Phi,\ L,\ \Psi)$ のクラスと等価
  - U が線形 $\iff$ Ψ が線形
  - U 非削除 $\iff$ Ψ が非削除
  - U が ε なし $\iff$ Ψ が ε なし