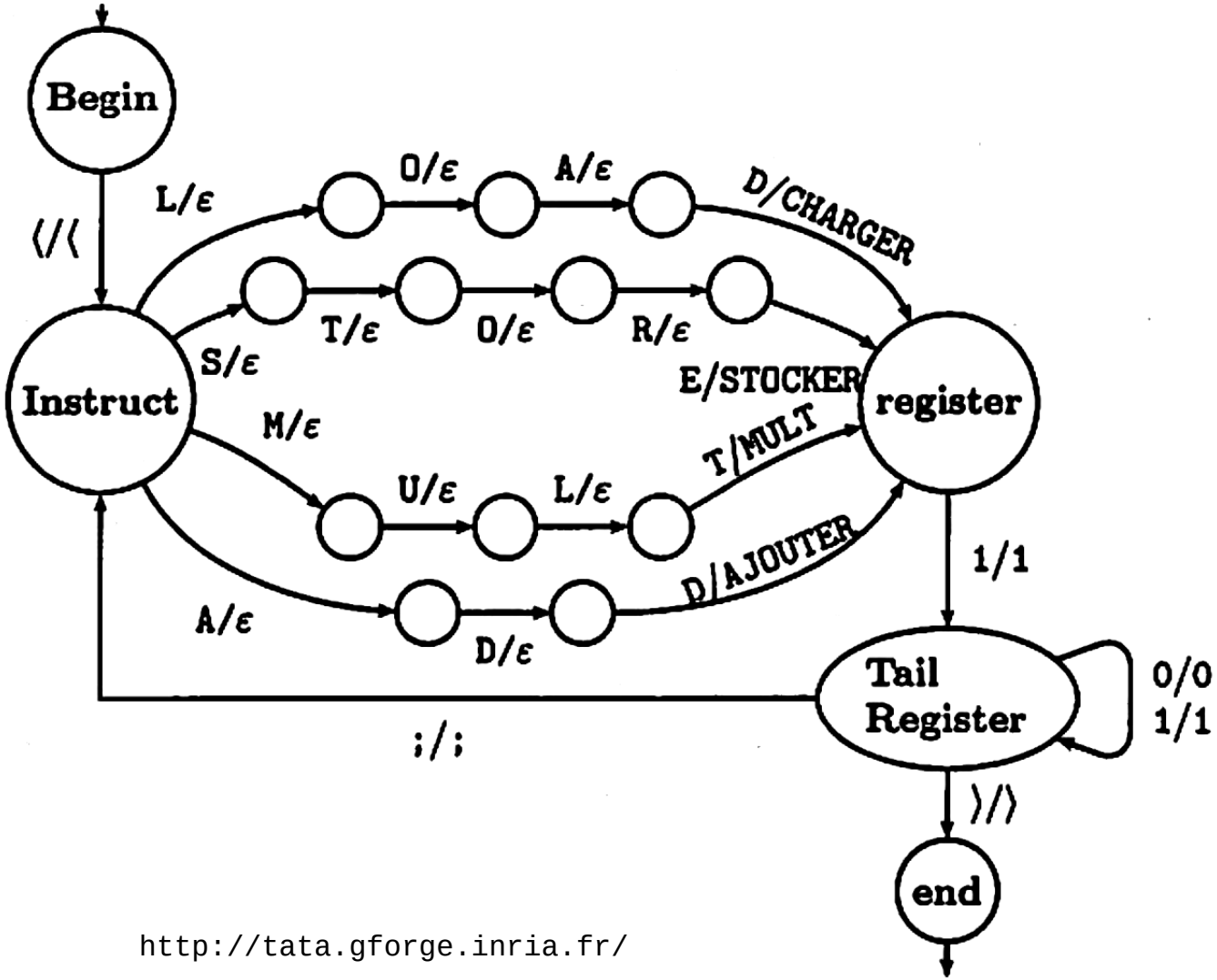
木変換系（Tree Transducers)

- 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)**



http://tata.gforge.inria.fr/

- **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 : $q \xrightarrow{f/m} q'$ 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 $\rightarrow_R$ **(**$\subseteq (\mathcal{F} \cup Q)^* \times \mathcal{F}'^*$**)**

  $(ft, q, u) \rightarrow_R (t, q', um)$ **for** $q \xrightarrow{f/m} q' \in \Delta$
- 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、

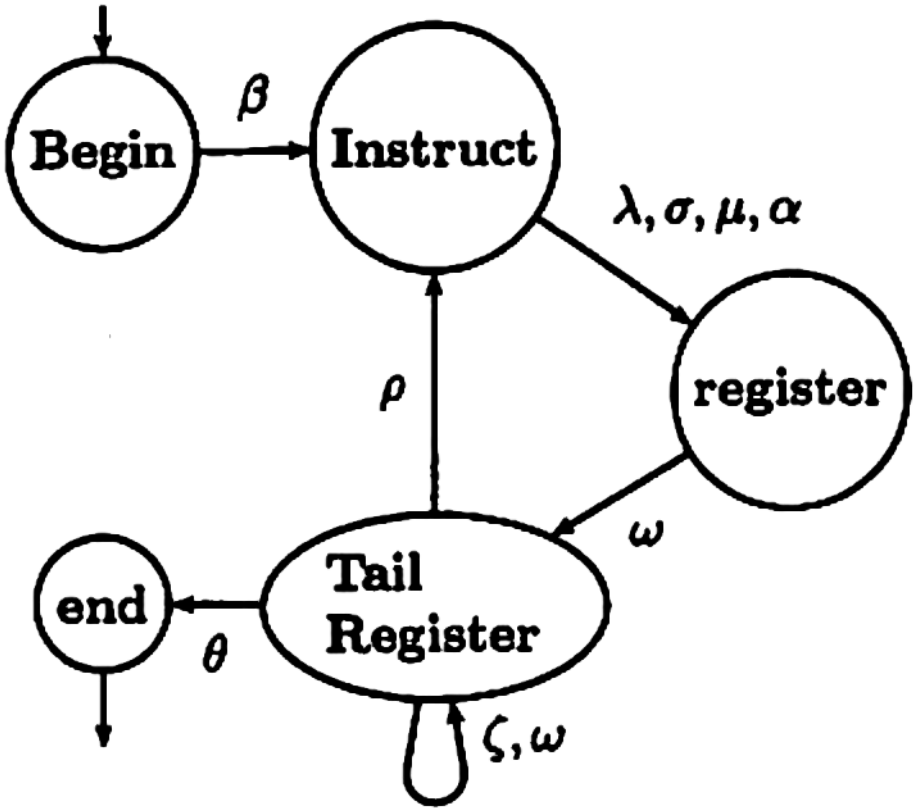
    $\Phi : \mathcal{F}''^* \to \mathcal{F}^*$

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

    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：

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



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

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

- 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：

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

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

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

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

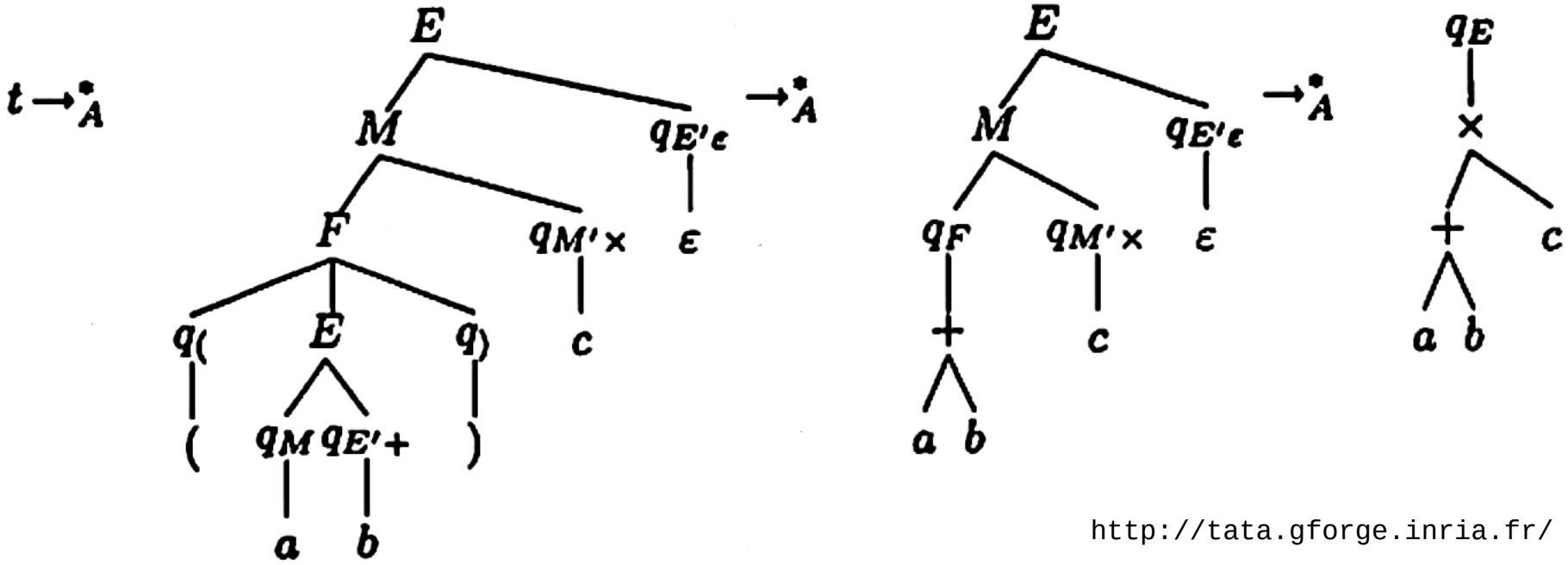
  - 線形**(linear)**：右辺の変数が線形**(**コピーなし**)**
  - 非消去**(non-erasing)**：右辺に$\mathcal{F}'$の記号が必ず出現
  - 非削除**(non-deleting)**：左辺の変数は右辺にも出現

- 例：受理状態q_E

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

http://tata.gforge.inria.fr/

- ボトムアップ木変換系 (続き)



http://tata.gforge.inria.fr/

- 例 U_1: 完全

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

$$a \to q(a) \qquad f(q(x)) \to q(f(x))$$
$$f(q(x)) \to q(f'(x)) \quad f(q(x)) \to q'(g(x,x))$$

$U_1(\{f(f(f(a)))\})$は以下の項からなる集合

$$g(f(f(a)), f(f(a))),\ g(f(f'(a)), f(f'(a))),$$
$$g(f'(f(a)), f'(f(a))),\ g(f'(f'(a)), f'(f'(a)))$$

- 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、

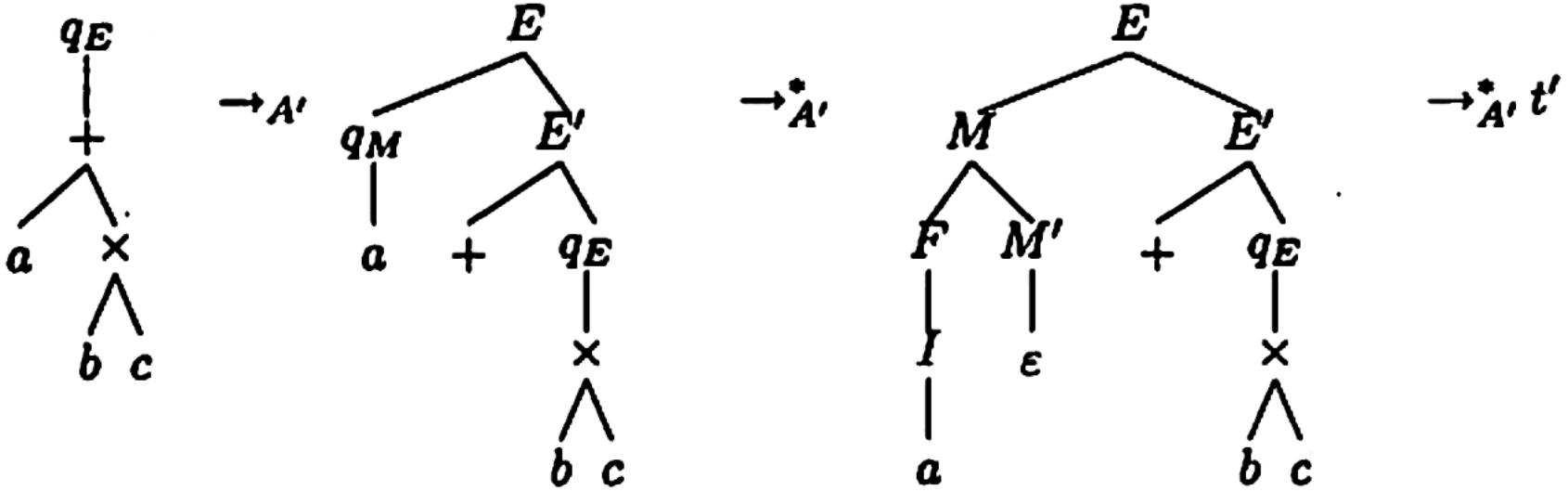
$$A(T) = \{s \in \mathsf{T}(\mathcal{F}') \mid t \in T,\ q \in Q^i,\ q(t) \to_A^* s\}$$

- 例：初期状態q_E

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

http://tata.gforge.inria.fr/

- トップダウン木変換系 (続き)



http://tata.gforge.inria.fr/

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

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

- 例 U_1:

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

$$q(f(x)) \to g(q'(x), q'(x)) \quad q'(f(x)) \to f(q'(x))$$
$$q'(a) \to a \quad q'(f(x)) \to f'(q'(x))$$

$U_1(\{f(f(f(a)))\})$は以下の**16**個の項からなる集合

$$g(f(f(a)), f(f(a))), \quad g(f(f(a)), f(f'(a))),$$
$$g(f(f(a)), f'(f(a))), \quad g(f(f(a)), f'(f'(a))),$$
$$\vdots$$
$$g(f'(f'(a)), f'(f(a))), \quad g(f'(f'(a)), f'(f'(a)))$$

- 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 $\Longleftrightarrow$ Ψ が線形
  - U 非削除 $\Longleftrightarrow$ Ψ が非削除
  - U が ε なし $\Longleftrightarrow$ Ψ が ε なし