VI – Modelo Categórico de Traduções entre Linguagens

6.1. Introdução

Traduções entre linguagens de programação estão presentes em diversas áreas da ciência da computação; como, por exemplo, na tradução fonte-a-fonte, na geração de código em um compilador e na otimização de programas. Apesar da relevância do assunto, existem muito poucos estudos teóricos sobre ele. O texto padrão sobre tradução entre linguagens é "The Theory of Parsing, Translating and Compiling" [Aho e Ullman, 1972] [Aho e Ullman, 1972b].

Iniciamos o capítulo apresentando o modelo de tradução de Aho e Ullman, após o qual, apresentamos um modelo categórico de tradução entre linguagens, baseado nas idéias do capítulo III. Em seguida, estudamos a relação entre o modelo categórico introduzido e o modelo usual e finalmente definimos formalmente o que é uma tradução que preserve a semântica de árvores (introduzida no capítulo IV).

6.2. Noção usual de tradução entre linguagens

O nosso conceito intuitivo de "tradução" é uma correspondência entre sentenças de duas linguagens que mantenha o significado dessas sentenças. Assim como no caso da formalização usual de linguagens, o aspecto do significado é deixado de lado das definições iniciais, podendo ser introduzido posteriormente.

Definição 6.1. Dada uma linguagem L_1 sobre um alfabeto Σ, e uma linguagem L_2 sobre um alfabeto Δ, definimos uma **tradução** entre L_1 e L_2 como uma relação $\tau \subseteq L_1 \times L_2$.

Além do aspecto semântico, também o aspecto sintático da linguagem é deixado de fora do conceito de tradução. Aho e Ullman abordam essa questão através do conceito de tradução dirigida por sintaxe, que apresentamos a seguir.

Definição 6.2. Um **esquema de tradução dirigido por sintaxe** (ETDS) é uma tupla $T = (N, \Sigma, \Delta, R, S)$, onde:

- N é um conjunto finito de símbolos não-terminais.
- Σ é o "alfabeto de entrada".
- Δ é o "alfabeto de saída".
- R é um conjunto finito de regras da forma $A \to \alpha,\beta,p$, onde $\alpha \in (N\cup\Sigma)^*$, $\beta \in (N\cup\Delta)^*$ e p é uma permutação entre os não-terminais de α e os não-terminais de β. Essa permutação é indicada através da colocação de índices nos não-terminais quando necessário. Convencionamos que $R_1 = \{ A \to \alpha \mid A \to \alpha,\beta,p \in R \}$ e $R_2 = \{ A \to \beta \mid A \to \alpha,\beta,p \in R \}$
- S é um não-terminal, chamado de símbolo inicial.

A idéia dessa definição é que $G_E(T)=(N, \Sigma, R_1 ,S)$ (**gramática de entrada** de T) e $G_S(T)=(N, \Delta, R_2 ,S)$ (**gramática de saída** de T) são duas gramáticas que funcionam de forma acoplada, onde um passo em uma derivação em G_E está associado a um passo em uma derivação de G_S.

Definição 6.3. Um ETDS é simples se toda regra $A \to \alpha,\beta,p \in R$, p é a identidade.

Exemplo 6.4. Neste exemplo, $N = \{E, T, F\}$, $\Sigma = \Delta = \{+, *, (,), Id\}$, $S = E$ e R é dado pelas regras abaixo. Mais adiante, veremos que este ETDS simples traduz uma expressão na notação infixa para a notação pósfixa.

$$
\begin{array}{lll}
E\to & E+T, & ET+ \\
E\to & T, & T \\
T\to & T*F, & TF* \\
T\to & F, & F \\
F\to & Id, & Id \\
F\to & (E), & E
\end{array}
$$

Figura 6.1 ETDS que mapeia expressões infixas em pósfixas.

Definição 6.5. Dado um ETDS T = (N, Σ, Δ, R, S), definimos a relação $(u_1Au_2, v_1Av_2, p) \Rightarrow_T (u_1\alpha u_2, v_1\beta v_2, p')$ se e somente se:

- $A \in N$; $u_1Au_2, u_1\alpha u_2 \in (\Sigma \cup N)^*$; $v_1Av_2, v_1\beta v_2 \in (\Delta \cup N)^*$
- u_1Au_2 e v_1Av_2 possuem os mesmos não-terminais, permutados de acordo com p, que mapeia A em A.
- $u_1\alpha u_2$ e $v_1\beta v_2$ possuem os mesmos não-terminais, permutados de acordo com p'.
- Existe $A \rightarrow \alpha,\beta,q \in R$, tal que $p'(B) = q(B)$ se B está em β, ou $p'(B) = p(B)$, caso contrário.

Definição 6.6. A relação $\Rightarrow_T^*$ é o fecho transitivo-reflexivo da relação $\Rightarrow_T$.

Definição 6.7. A tradução definida por um ETDS T = (N, Σ, Δ, R, S) é dada por:

$$trad(T) = \{ (u,v) \mid \exists p \text{ tq } (S,S,id) \Rightarrow_T^* (u,v,p),\ u \in \Sigma^*,\ v \in \Delta^* \}$$

Exemplo 6.8. O SDTS do exemplo 6.4 acima traduz a expressão (Id+Id)*Id para Id Id + Id *, como mostrado na "derivação" abaixo:

$$(E,E) \Rightarrow_T (T,T) \Rightarrow_T (T{*}F, TF{*}) \Rightarrow_T (F_1{*}F_2, F_1F_2{*}) \Rightarrow_T ((E){*}F, EF{*}) \Rightarrow_T ((E+T){*}F, ET+F{*}) \Rightarrow_T$$

$$((T_1+T_2) * F,\ T_1T_2+F{*}) \Rightarrow_T ((F_1+T){*}F_2, F_1T+F_2{*}) \Rightarrow_T ((Id+T){*}F,\ IdT+F{*}) \Rightarrow_T$$

$$((Id+F_1){*}F_2,\ Id\ F_1+F_2{*}) \Rightarrow_T ((Id+Id){*}F,\ Id\ Id+F{*}) \Rightarrow_T ((Id+Id){*}\ Id,\ Id\ Id + Id\ {*})$$

Definição 6.9. Uma tradução τ é uma tradução dirigida por sintaxe se existe ETDS T tal que τ = trad(T).

6.3 Noção categórica de tradução entre linguagens

Nesta seção buscamos um modelo categórico para um esquema de tradução. A tentativa mais natural é definir tradução como um morfismo em uma das categorias apresentadas no capítulo III ($\mathbf{Gram}^{\Sigma}$ e $\mathbf{Cat}_{\square}^{\Sigma}$). Pode-se definir um morfismo de domínio $(F_1,\Sigma,\phi,\mathbf{C_1}) = L_{\square}(G_1,\Sigma,\phi)$ e codomínio $(F_2,\Sigma,\phi,\mathbf{C_2}) = L_{\square}(G_2,\Sigma,\phi)$ através do seu comportamento nas arestas de $F_1(M)$ que são regras de $G_1(i)$, seu comportamento nas demais arestas será obtido pela definição de $L_{\square}$. Assim, um morfismo de $\mathbf{Cat}_{\square}$ leva regras de derivação de G_1 em derivações em F_2. No entanto, existem algumas situações, que correspondem a nossa intuição de tradução composicional, que não podem ser modeladas por tais morfismos, por exemplo:

$$\tau_M(E_0 \rightarrow E_1\, M\, T) \quad = \quad E_0 \overset{*}{\Rightarrow} T\, M\, E_1$$

$$\tau_M(E_0 \rightarrow Sqr\, E_1) \quad = \quad E_0 \overset{*}{\Rightarrow} E_1\, P\, E_1$$

No primeiro caso, a tradução da regra $E_0 \rightarrow E_1\, M\, T$ é uma derivação na qual a ordem dos não-terminais está trocada; portanto, τ_M não seria a componente-morfismo de um funtor entre as categorias $\mathbf{C_1}$ e $\mathbf{C_2}$. No segundo caso, τ_M também não seria a componente-morfismo de um funtor entre $\mathbf{C_1}$ e $\mathbf{C_2}$, pois E_1 está repetido na tradução. É fácil construir um modelo destas traduções através de uma pequena modificação da nossa construção no capítulo III; basta tomarmos as categorias com soma no lugar das categorias estritamente monoidais. As definições a seguir formalizam esta construção.

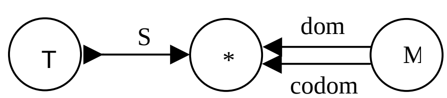

Figura 6.2 Categoria **A**.

Definição 6.10. Uma **floresta de DAGs** sobre um alfabeto Σ é uma tupla $\mathbf{D}$ = $(D, \Sigma, \phi, \mathbf{C})$, onde:

- D é funtor D: $\mathbf{A} \rightarrow \mathbf{SETS}$, tal que:
  - ❑ D(T) é um objeto terminal em **Sets**.
  - ❑ D(*) e D(M) são respectivamente os objetos e os morfismos da categoria pequena $\mathbf{C}$ = (C, +, $\mathbf{e}$) com somas finitas, onde a soma é estritamente associativa e tem um elemento neutro $\mathbf{e}$.

- D(dom) e D(codom) são as funções que retornam, respectivamente, o domínio e o codomínio de um dado morfismo em C.

- ϕ é uma função dos "morfismos 0-ários" de C em Σ^*. Isto é seja Z = { m $\in$ D(M) | m: **e** $\rightarrow$A, A $\neq$ **e** } então ϕ é uma função ϕ:Z$\rightarrow\Sigma^*$.

Definição 6.11. $\mathbf{Cat}_+^{\Sigma}$ é a categoria cujos objetos são florestas de DAGs e cujos morfismos entre $(D_1,\Sigma,\phi_1,\mathbf{C_1})$ e $(D_2,\Sigma,\phi_2,\mathbf{C_2})$ são transformações naturais $\eta:D_1 \xrightarrow{\bullet} D_2$, tais que $\eta(*)$ e $\eta(M)$ são respectivamente as componentes-objeto e morfismo de um funtor entre $\mathbf{C_1}$ e $\mathbf{C_2}$, que preserva +.

Exemplo 6.12. Considere a gramática G_{EXP}, do exemplo 3.8. Existe um morfismo entre ID e E, correspondente ao DAG da figura 6.3, i.e., ao trecho de derivação E $\Rightarrow$ E P E $\Rightarrow$ ID P E $\Rightarrow$ ID $\times$ E $\Rightarrow$ ID $\times$ ID, onde as ocorrências de ID são obrigatoriamente iguais. A figura 6.4 ilustra a construção deste morfismo.

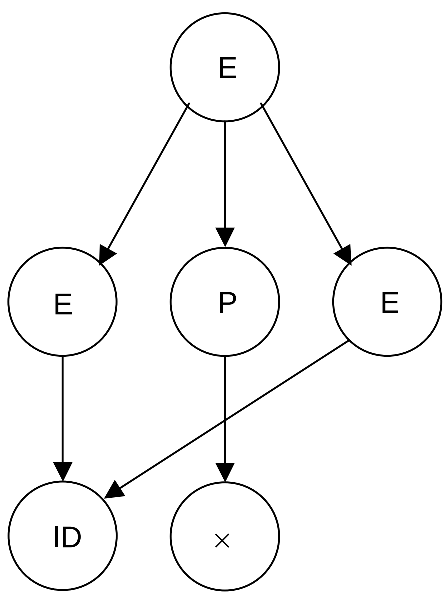

Figura 6.3 DAG correspondente a uma derivação na qual as ocorrências de ID são igualadas.

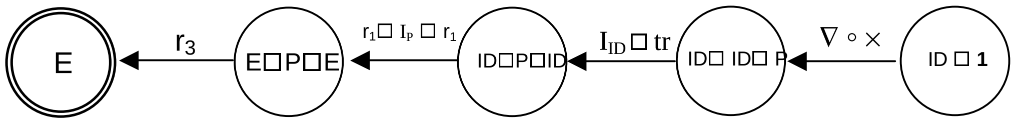

Figura 6.4 Construção do morfismo correspondente ao DAG.

Os morfismos em $\mathbf{Cat}^{\Sigma}_{+}$ são morfismos entre "DAGs de derivação livres de ordem", i.e., esses morfismos devem levar um DAG em outro DAG com a mesma "assinatura", a menos da ordem de apresentação dos não-terminais.

Definição 6.13. $\mathbf{U}_{+/\square}$: $\mathbf{Cat}^{\Sigma}_{+} \to \mathbf{Cat}^{\Sigma}_{\square}$ é um funtor definido como:

- A componente-objeto de $U_{+/\square}$ leva uma floresta de DAGs (D,Σ,ϕ_D) a uma floresta (F,Σ,ϕ_F) tal que:
  - F = D. Observe que, pela definição, uma floresta de DAGs é uma floresta de árvores, pois uma categoria que possui soma estritamente associativa e elemento neutro é estritamente monoidal.
  - $\phi_F = \phi_D$.
- A componente-morfismo de $U_{+/\square}$ leva um morfismo η_+ de $\mathbf{Cat}^{\Sigma}_{+}$ em um morfismo $\eta_{\square}$ em $\mathbf{Cat}^{\Sigma}_{\square}$ tal que:
  - $\eta_{\square} = \eta_+$. Esta igualdade está justificada no apêndice.

Teorema 6.14. $U_{+/\square}$ tem um adjunto à esquerda $L_{+/\square}$: $\mathbf{Cat}^{\Sigma}_{\square} \to \mathbf{Cat}^{\Sigma}_{+}$ tal que:

- A componente-objeto de $L_{+/\square}$ leva uma floresta (F, Σ, ϕ_F, **C**) na floresta de DAGs (D, Σ, ϕ_D, **C'**) tal que:
  - D(S) = F(S), portanto D(T) = F(T) e D(*) = F(*) .
  - D(M) = Morf(C'). Onde **C'** é a menor categoria tal que Obj(C') = F(*), Morf(C') ⊇ F(M) e que possui uma soma que é uma extensão de □.

- Para $m_1, m_2 \in F(M)$, $D(codom)(m_1) = F(codom)(m_1)$ e $D(dom)(m_1) = F(dom)(m_1)$ e $m_1+m_2 = m_1 \square m_2$.
- $\phi_D = \phi_F$.

- A componente-morfismo de $L_{+/\square}$ leva um morfismo $\eta_\square$ de $\mathbf{Cat}^{\Sigma}_{\square}$ em um morfismo η_+ em $\mathbf{Cat}^{\Sigma}_{+}$ tal que:
  - $\eta_{+T} = \eta_{\square T}$, $\eta_{+*} = \eta_{\square *}$.
  - η_{+M} estende $\eta_{\square M}$ preservando a soma.

Podemos ver $U_{+/\square}$ como uma inclusão de $\mathbf{Cat}^{\Sigma}_{+}$ em $\mathbf{Cat}^{\Sigma}_{\square}$, e $L_{+/\square}$ é um funtor que estende o produto tensorial para uma soma, introduzindo os morfismos que forem necessários.

Definição 6.15. $\mathbf{U_+}$: $\mathbf{Cat}^{\Sigma}_{+} \rightarrow \mathbf{Gram}^{\Sigma}$ é o funtor $\mathbf{U_\square} \circ \mathbf{U_{+/\square}}$. $\mathbf{L_+}$:$\mathbf{Gram}^{\Sigma} \rightarrow \mathbf{Cat}^{\Sigma}_{+}$ é o funtor $\mathbf{L_{+/\square}} \circ \mathbf{L_\square}$.

Neste ponto, estamos aptos a apresentar nossa definição de esquema categórico de tradução.

Definição 6.16. Um **esquema categórico de tradução** entre (G_1,Σ,ϕ_1) e (G_2,Σ,ϕ_2) é um morfismo em $\mathbf{Cat}^{\Sigma}_{+}$, τ_+:$L_+(G_1,\Sigma,\phi_1) \rightarrow L_+(G_1,\Sigma,\phi_1)$.

Definição 6.17. A **tradução gerada** por esquema de tradução categórico $\tau_+ = \eta_+$ entre (G_1,Σ_1,ϕ_1) e (G_2,Σ_2,ϕ_2) é definida como:

$$trad(\tau_+) = \{(\mathbf{T}(m), \mathbf{T}(n)) \mid m \in Hom_{\mathbf{C_1}}(\mathbf{1},F(T)F(S)) \text{ e } n = \eta_{+M}(m) \}$$

onde $(F_1,\Sigma,\phi_1,\mathbf{C_1}) = L_\square(G_1,\Sigma,\phi_1)$ e $\mathbf{T}(\ \_\ )$ é a função definida no capítulo 3 que retorna os terminais do morfismo m e $S = F_1(S)(F_1(T))$. Usaremos informalmente $trad(\tau_\square)$, ou "tradução gerada por $\tau_\square$", para designar $trad(L_{+/\square}\tau_\square)$.

6.4 Relação entre a noção usual de tradução e a categórica

Esta seção investiga a relação entre o conceito categórico de esquema de tradução e o usual. Iniciamos pelo tipo mais simples de esquema categórico de tradução, aquele que é da forma $L_{+/\square}\tau_{\square}$. Esses ETCs possuem o mesmo poder dos ETDSs simples. Essa relação é formalizada no teorema abaixo, cuja demonstração encontra-se no apêndice A.

Teorema 6.18. Dado um ETDS simples T, existe um morfismo $\tau_{\square}$ de $\mathbf{Cat}_{\square}^{\Sigma}$, para algum Σ, tal que trad($\tau_{\square}$) = trad(T). Reciprocamente, dado um morfismo $\tau_{\square}:L_{\square}\mathbf{G_1}\rightarrow L_{\square}\mathbf{G_2}$ de $\mathbf{Cat}_{\square}^{\Sigma}$, existe um ETDS simples T, tal que trad(T) = trad($\tau_{\square}$).

Os morfismos τ_{+} de $\mathbf{Cat}_{+}^{\Sigma}$ possuem um poder de expressividade maior, pois permitem a duplicação e a permutação de não-terminais, i.e., são morfismos entre DAGs não-ordenados e não morfismos entre árvores. Aho e Ullman também sugerem uma generalização nesse sentido quando definem um *pushdown processor* que mapeia *strings* em grafos [Aho e Ullman, 1972b, pág. 737]; no entanto eles permitem a manipulação algorítmica destes grafos (bem como de outros tipos de dados) e, devido à complexidade introduzida pela generalização, não chegam a formalizar o conceito.

Na próxima seção, apresentamos um modelo categórico de tradução intermediário entre $\tau_{\square}$ e τ_{+}, cujo poder de expressividade é equivalente aos ETDS. Este modelo evita precisamente as traduções que introduzem redundância no código traduzido, através da duplicação de não-terminais, o que prejudica a manutenibilidade. Portanto, o modelo apresentado é um primeiro passo na busca de uma definição de tradução que preserva a manutenibilidade.

6.5 Modelo categórico de um ETDS

Na seção anterior, vimos que os morfismos de $\mathbf{Cat}_{\Box}^{\Sigma}$ da forma $\tau_{\Box}:L_{\Box}\mathbf{G_1}\rightarrow L_{\Box}\mathbf{G_2}$ são um modelo para esquemas de tradução dirigidos por sintaxe simples. Se procuramos um modelo para esquemas de tradução dirigidos por sintaxe, devemos procurar morfismos com maior poder de expressão. Esse é o caso dos morfismos de $\mathbf{Cat}_{+}^{\Sigma}$. No entanto, esses morfismos, além de permitir a permutação, permitem também a duplicação de informação. Nós vamos procurar os modelos de STDS em uma categoria intermediária entre $\mathbf{Cat}_{\Box}^{\Sigma}$ e $\mathbf{Cat}_{+}^{\Sigma}$, que definimos a seguir.

Definição 6.19. Uma **floresta de árvores não ordenadas** sobre um alfabeto Σ é uma tupla $(T, \Sigma, \phi, \mathbf{C})$, onde

- T é funtor $T:\mathbf{A} \rightarrow \mathbf{SETS}$, tal que:
  - $T(\top)$ é um objeto terminal em **Sets**.
  - $T(*)$ e $T(M)$ são, respectivamente, os objetos e os morfismos de uma categoria pequena estritamente monoidal simétrica **C**. Categorias monoidais simétricas são definidas em [Barr e Wells 1995, pp. 290].
  - $T(dom)$ e $T(codom)$ são as funções que retornam, respectivamente, o domínio e o codomínio de um dado morfismo em C.
- ϕ é uma função dos “morfismos 0-ários” de C em Σ^{*}. Isto é, seja $Z = \{ m \in T(M) \mid m: \mathbf{1} \rightarrow A, A \neq \mathbf{1} \}$ então ϕ é uma função $\phi:Z\rightarrow\Sigma^{*}$.

Definição 6.20. $\mathbf{Cat}_{\mathrm{T}}^{\Sigma}$ (o índice T sugere “troca”, ou permutação de posições) é a categoria cujos objetos são florestas de árvores não-ordenadas e cujos morfismos entre $(T_1, \Sigma, \phi_1, \mathbf{C_1})$ e $(T_2, \Sigma, \phi_2, \mathbf{C_2})$ são transformações naturais $\eta:T_1 \overset{\bullet}{\rightarrow} T_2$ tal que $\eta(*)$ e $\eta(M)$ são respectivamente as componentes objeto e morfismo de um funtor F entre $\mathbf{C_1}$ e $\mathbf{C_2}$, que preserva $\square$ e $F(s_{a,b}) = s_{F(a),F(b)}$. Onde $s_{a,b}$ é a transformação natural $s_{a,b}: A\square B \xrightarrow{\bullet} B\square A$.

Definição 6.21. $U_{T/\square}$: $\mathbf{Cat}^{\Sigma}_{T} \rightarrow \mathbf{Cat}^{\Sigma}_{\square}$ é um funtor definido como:

- A componente-objeto de $U_{T/\square}$ leva uma floresta de árvores não-ordenadas (T, Σ, ϕ_T, $\mathbf{C_1}$) a uma floresta (F, Σ, ϕ_F, $\mathbf{C_2}$) tal que:
  - T = F. Observe que, pela definição, uma floresta de árvores não-ordenadas é uma floresta de árvores.
  - $\phi_T = \phi_F$.
- A componente-morfismo de $U_{T/\square}$ leva um morfismo η_T de $\mathbf{Cat}_T$ em um morfismo $\eta_\square$ em $\mathbf{Cat}_\square$ tal que:
  - $\eta_\square = \eta_T$. Observe que $\eta_T(*)$ e $\eta_T(M)$ são, por definição, as componentes de um funtor que preserva $\square$.

Teorema 6.22. $U_{T/\square}$ tem um adjunto à esquerda, $L_{T/\square}$: $\mathbf{Cat}^{\Sigma}_{\square} \rightarrow \mathbf{Cat}^{\Sigma}_{T}$ tal que:

- A componente-objeto de $L_{T/\square}$ leva uma floresta (F, Σ, ϕ_F, **C**) na floresta de árvores não ordenadas (T, Σ, ϕ_T, **C’**) tal que:
  - T(S) = F(S), portanto T(T) = F(T) e T(*) = F(*).
  - T(M) = Morf(C’). Onde C’ é a menor categoria tal que Obj(C’) = Obj(C), Morf(C’) $\supseteq$ Morf(C) e que é simétrica.
  - Para m_1, $m_2 \in$ F(M), T(codom)(m_1) = F(codom)(m_1), T(dom)(m_1) = F(dom)(m_1).
  - $\phi_T = \phi_F$.
- A componente-morfismo de $L_{T/\square}$ leva um morfismo $\eta_\square$ de $\mathbf{Cat}^{\Sigma}_{\square}$ em um morfismo η_T em $\mathbf{Cat}^{\Sigma}_{T}$ tal que:

- $\eta_{T_T} = \eta_{\square_T}$, $\eta_{T_*} = \eta_{\square_*}$.
- η_{T_M} estende $\eta_{\square_M}$ preservando $s_{a,b}$.

Definição 6.23. $\mathbf{U}_T$: $\mathbf{Cat}^{\Sigma}_{T} \rightarrow \mathbf{Gram}^{\Sigma}$ é o funtor $\mathbf{U}_{\square} \circ \mathbf{U}_{T/\square}$. $\mathbf{L}_T$: $\mathbf{Gram}^{\Sigma} \rightarrow \mathbf{Cat}^{\Sigma}_{T}$ é o funtor $\mathbf{L}_{T/\square} \circ \mathbf{L}_{\square}$.

Teorema 6.24. Dado um ETDS T, existe um morfismo τ_T de $\mathbf{Cat}^{\Sigma}_{T}$, tal que trad($\tau_T$) = trad(T). Reciprocamente, dado um morfismo τ_T:$L_T\mathbf{G_1} \rightarrow L_T\mathbf{G_2}$ de $\mathbf{Cat}^{\Sigma}_{T}$, existe um ETDS T, tal que trad(τ_T) = trad(T).

No teorema acima trad(τ_T) significa trad($L_{+/T}\tau_T$). Portanto precisamos definir mais uma adjunção, desta vez entre $\mathbf{Cat}^{\Sigma}_{T}$ e $\mathbf{Cat}^{\Sigma}_{+}$. Esta nova adjunção também será útil para enunciarmos a condição para a preservação da semântica de árvores na próxima seção.

Definição 6.25. $\mathbf{U}_{+/T}$: $\mathbf{Cat}^{\Sigma}_{+} \rightarrow \mathbf{Cat}^{\Sigma}_{T}$ é um funtor definido como:

- A componente-objeto de $U_{+/T}$ leva uma floresta de DAGs (D, Σ, ϕ_D, $\mathbf{C_1}$) a uma floresta de árvores não-ordenadas (T, Σ, ϕ_T, $\mathbf{C_2}$) tal que:
  - T = D. Observe que, pela definição, uma floresta de DAGs é uma floresta de árvores não ordenadas, basta tomar a soma como o produto $\square$ e $\langle i_b, i_a \rangle$ como $s_{a,b}$.
  - $\phi_T = \phi_D$.
- A componente-morfismo de $U_{+/T}$ leva um morfismo η_+ de $\mathbf{Cat}^{\Sigma}_{+}$ em um morfismo η_T em $\mathbf{Cat}^{\Sigma}_{T}$ tal que:
  - $\eta_T = \eta_+$. Tomando a soma como o produto $\square$ e $\langle i_b, i_a \rangle$ como $s_{a,b}$, um funtor que preserve a soma preservará também $\square$ e $s_{a,b}$.

Teorema 6.26. $U_{+/T}$ tem um adjunto à esquerda, $L_{+/T}$: $\mathbf{Cat}^{\Sigma}_{T} \rightarrow \mathbf{Cat}^{\Sigma}_{+}$, tal que:

- A componente-objeto de $L_{+/T}$ leva uma floresta de árvores não-ordenadas (T, Σ, ϕ_T, **C**) à floresta de DAGs (D ,Σ, ϕ_D, **C'**) tal que:
    - D(S) = T(S), portanto D(T) = T(T) e D(*) = T(*).
    - D(M) = Morf(C'). Onde **C'** é a menor categoria tal que Obj(**C'**) = Obj(**C**), Morf(**C'**) $\supseteq$ Morf(**C**), e que possui uma soma que é uma extensão do produto $\square$ em **C** e tal que $s_{a,b} = \langle i_b, i_a \rangle$.
    - Para m_1, $m_2 \in$ T(M), D(codom)(m_1) = T(codom)(m_1) e D(dom)(m_1) = T(dom)(m_1) e $m_1+m_2 = m_1 \square m_2$.
    - $\phi_D = \phi_T$.
- A componente-morfismo de $L_{+/T}$ leva um morfismo η_T de $\mathbf{Cat}_T^{\Sigma}$ em um morfismo η_+ em $\mathbf{Cat}_+^{\Sigma}$ tal que:
    - $\eta_{+T} = \eta_{T_T}$, $\eta_{+*} = \eta_{T*}$.
    - η_{+M} estende η_{T_M} preservando a soma.

6.6 Tradução que preserva a semântica de árvores

Finalmente, estamos aptos a definir quando uma tradução preserva a semântica de árvores. Estamos interessados em traduções que mapeam "estruturas" na linguagem de entrada em "estruturas" na linguagem de saída com a mesma intenção, isto é, a mesma semântica de árvores.

Pelos argumentos apresentados no capítulo 5, as "estruturas" às quais nos referimos não são árvores de derivação (pois estas são muito restritivas ao dependerem da ordem dos não-terminais), nem tampouco DAGs de derivação (pois estes são excessivamente flexíveis, permitindo repetição de não terminais). As estruturas que procuramos são árvores não-ordenadas.

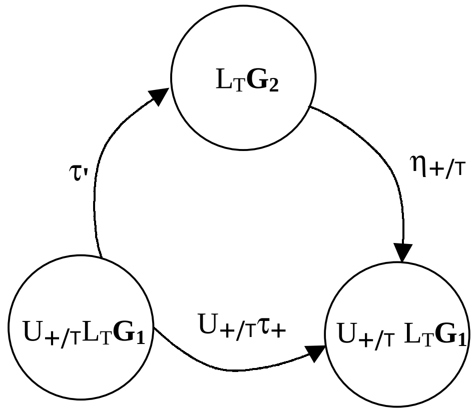

Figura 6.5 Condição para τ_+ mapear árvores não-ordenadas em árvores não-ordenadas.

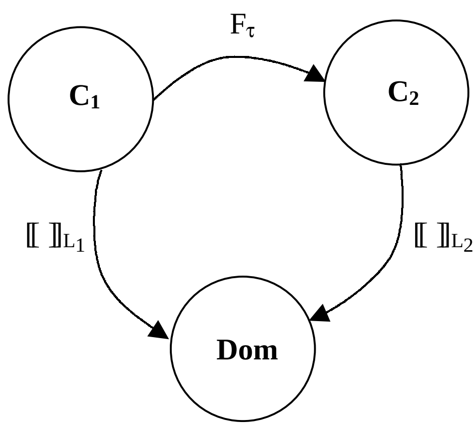

Figura 6.6 Condição para τ_+ preservar a semântica das árvores traduzidas.

Definição 6.27. Sejam $\mathbf{G_1} = (G_1,\Sigma_1,\phi_1,\mathbf{C_1})$ e $\mathbf{G_2} = (G_2,\Sigma_2,\phi_2,\mathbf{C_1})$ GLCCs. Um esquema categórico de tradução $\tau_+:L_+\mathbf{G_1}\rightarrow L_+\mathbf{G_2}$ preserva a semântica de árvores se e somente satisfaz as seguintes condições.

- τ_+ mapeia árvores não-ordenadas em árvores não-ordenadas, ou seja, existe τ' tal que $U_{+/T}\tau_+$ fatora como indicado no diagrama em $\mathbf{Cat}_T^\Sigma$ da figura 6.5, onde $\eta_{+/T}$ é a unidade da adjunção entre $\mathbf{Cat}_T^\Sigma$ e $\mathbf{Cat}_+^\Sigma$.
- A semântica de árvores da árvore traduzida é a mesma, ou seja, o diagrama na categoria das categorias pequenas estritamente monoidais indicado na figura 6.6

comuta. Neste diagrama F_τ é o funtor entre $\mathbf{C_1}$ e $\mathbf{C_2}$ cujas componentes objeto e morfismo são respectivamente as componentes * e M da transformação natural $U_{+/T}\tau_+$.