Inférence de Types

Soit $\mathcal{A} = \{A, B, C, \dots\}$ un ensemble infini de *variables de types*.
On considère $\mathcal{T}$ l'ensemble des *types* définis inductivement par :

- $\mathbb{N} \in \mathcal{T}$. ($\mathbb{N}$ est un type)
- $\mathcal{A} \subseteq \mathcal{T}$. (une variable de types est un type)
- si $T_1 \in \mathcal{T}$ et $T_2 \in \mathcal{T}$ alors $T_1 \to T_2 \in \mathcal{T}$. (la "flèche" de deux types est un type)

Une *substitution de types* $\sigma : \mathcal{A} \longrightarrow \mathcal{T}$ est une fonction qui vaut l'identité presque partout, c'est-à-dire telle que $\{A \in \mathcal{A} \mid \sigma(A) \neq A\}$ est fini.
Si σ est une substitution de types et $T \in \mathcal{T}$ un type, on note $\sigma_\uparrow(T)$ (par abus de notation, on s'autorisera à écrire $\sigma(T)$) le type obtenu inductivement par :

- $\sigma_\uparrow(\mathbb{N}) = \mathbb{N}$.
- pour $\mathcal{A}$, $\sigma_\uparrow(A) = \sigma(A)$.
- pour tout $T_1, T_2 \in \mathcal{T}^2$, $\sigma_\uparrow(T_1 \to T_2) = \sigma_\uparrow(T_1) \to \sigma_\uparrow(T_2)$

Un *système d'équations de types* (ou juste *système*) est un ensemble fini $\{(T_k, T'_k)\}_{1 \leq k \leq n}$ de couples de types.
Un *unificateur* d'un système $\{(T_k, T'_k)\}_{1 \leq k \leq n}$ est une substitution de types σ telle que $\forall k, \sigma_\uparrow(T_k) = \sigma_\uparrow(T'_k)$.

Par exemple, on pose $S_0 = \{(B \to A, C),\ (\mathbb{N}, B)\}$

et σ_0 définie par $\begin{cases} \sigma_0(B) &= \mathbb{N} \\ \sigma_0(C) &= \mathbb{N} \to A \\ \sigma_0(X) &= X \text{ pour tout} X \notin \{B, C\} \end{cases}$

Alors σ_0 est un unificateur de S_0 car :

1. $\sigma_0(B \to A) = \mathbb{N} \to A$ et $\sigma_0(C) = \mathbb{N} \to A$
2. $\sigma_0(\mathbb{N}) = \mathbb{N}$ et $\sigma_0(B) = \mathbb{N}$

Question 1. Trouver si possible un unificateur des systèmes suivants :

1. $\{(\mathbb{N} \to A, B \to \mathbb{N})\}$
2. $\{(\mathbb{N} \to A, B \to (C \to \mathbb{N})),\ (\mathbb{N} \to \mathbb{N}, C)\}$
3. $\{(A \to B), (B \to A)\}$
4. $\{(\mathbb{N} \to (A \to \mathbb{N}), \mathbb{N} \to B),\ (B, A)\}$

Question 2. Un système S admet-il toujours un unificateur ?

Quand un unificateur pour S est-il unique ?

Question 3. En considérant les différentes possibilités pour les formes des types T_1 et T'_1, exprimer l'existence d'un unificateur pour $\{(T_k, T'_k)\}_{1 \leq k \leq n}$ en fonction de l'existence d'un unificateur pour un autre système, bien choisi.

Question 4. Rédiger l'algorithme induit par la question précédente, étudier sa terminaison.

Soit $\mathcal{X} = \{x, y, z, \dots\}$ un ensemble de *variables de termes*.
On considère un langage d'expressions fonctionnelles $\mathcal{E}$ défini inductivement par :

- $\mathbb{N} \subseteq \mathcal{E}$, (les entiers sont des expressions)
- $\mathcal{X} \subseteq \mathcal{E}$ (les variables de termes sont des expressions)
- Si $(E_1, E_2) \in \mathcal{E}^2$, $E_1 + E_2 \in \mathcal{E}$ (la somme de deux expressions est une expression)
- Si $(E_1, E_2) \in \mathcal{E}^2$, $E_1(E_2) \in \mathcal{E}$ (l'application d'une expression à une expression est une expression)
- Si $x \in \mathcal{X}$ et $E \in \mathcal{E}$ alors $(x \mapsto E) \in \mathcal{E}$ (si E est une expression et x une variable, l'expression fonctionnelle $x \mapsto E$ est une expression)

On supposera que les variables x apparaissant dans des sous-expressions $x \mapsto E'$ d'une expression E sont toutes distinctes et distinctes deux à deux des variables de E qui n'apparaissent jamais directement à gauche d'un $\mapsto$. On note $FV(E)$ ce dernier ensemble de variables.

Par exemple, $f_0 = (x \mapsto (y \mapsto x(y) + 1))$ est un élément de $\mathcal{E}$.

Les contextes de types sont des fonctions de $D \subseteq \mathcal{X}$ dans T.
On note $\emptyset$ le contexte de domaine vide et $(\Gamma, x : T)$, le contexte Γ' défini par $\Gamma'(x) = T$ et pour tout y dans le domaine de Γ, $y \neq x \Rightarrow \Gamma'(y) = \Gamma(y)$.

Dans la suite on s'intéresse au problème de donner un type de $\mathcal{T}$ à une expression de $E \in \mathcal{E}$ avec un contexte Γ qui sert d'oracle donnant le type des variables de $FV(E)$, quand c'est possible. On dit qu'on *infère un type de E dans le contexte* Γ.

Question 5. Donner deux types différents qu'il semble légitime de donner à f_0 dans le contexte $\emptyset$.

Question 6. Donner des règles qui formalisent quand il est légitime de donner un type T à E dans le contexte Γ.

Question 7. En déduire un algorithme qui infère un type d'une expression E.

Question 8. Utiliser cet algorithme pour trouver un type de $\mathbf{S} = x \mapsto (\ y \mapsto (\ z \mapsto x(z)(y(z))\ )\ )$ dans le contexte $\emptyset$.