Typage fort et typage souple des collections topologiques et des transformations

Julien Cohen

LaMI UMR 8042,
CNRS - Université d'Évry Val d'Essonne
523 Place des Terrasses de l'Agora
91025 Évry, France
`jcohen@lami.univ-evry.fr`

Résumé

Les collections topologiques permettent de considérer uniformément de nombreuses structures de données dans un langage de programmation et sont manipulées par des fonctions définies par filtrage appelées des transformations.

Nous présentons dans cet article deux systèmes de types pour des langages intégrant les collections topologiques et les transformations. Le premier est un système à typage fort à la Hindley/Milner qui peut être entièrement typé à la compilation. Le second est un système à typage mixte statique/dynamique permettant de gérer des collections hétérogènes, c'est-à-dire qui contiennent des valeurs de types distincts. Dans les deux cas l'inférence de types automatique est possible.

1. Introduction

Les *collections topologiques* sont une famille de structures de données que l'on peut voir comme des fonctions d'un ensemble de positions vers un ensemble de valeurs et sur lesquelles on dispose d'une relation de voisinage entre les positions. De nombreuses structures de données usuelles peuvent être vues comme des collections topologiques : ensembles, séquences, tableaux généralisés, graphes, *etc.* On peut programmer par filtrage sur ces collections grâce à des fonctions particulières appelées *transformations*. Les transformations permettent d'écrire des programmes opérant uniformément sur diverses structures de données. De tels programmes sont dits *polytypiques* [JJ96].

Nous montrons dans cet article que les collections topologiques et les transformations peuvent s'intégrer dans un langage fortement typé. De plus elles y apportent des caractéristiques importantes comme le polytypisme et le filtrage sur des structures non-algébriques sans perdre le polymorphisme paramétrique ou l'inférence automatique de types.

Dans un second temps, nous proposons un système de typage plus souple permettant de manipuler des collections hétérogènes, c'est à dire des collections contenant des valeurs de types différents. En effet, la programmation par transformations trouve un champ d'application important en simulation biologique où les collections manipulées sont souvent hétérogènes. Le cadre hétérogène permet aux transformations de gagner en expressivité par rapport au langage fortement typé. Des exemples de tels programmes [BT02, BdRTG03] ont été codés dans le langage déclaratif MGS [Gia03] qui intègre les collections et les transformations dans un contexte dynamiquement typé. Les langages dynamiquement typés ont fait l'objet de nombreux travaux visant à leur donner un système de types proche d'un typage statique avec inférence de types automatique [ALW94, CF91, Dam94b, Fur02]. Le langage que nous présentons qui permet la manipulation de collections hétérogènes est un sous-ensemble du langage

MGS et nous lui associons un système de types souple en partie basé sur les travaux de Aiken *et al.* [ALW94]. Ce système est doté d'une procédure de typage automatisée permettant de construire un type précis pour les transformations. Ce système de types est une étape importante dans l'élaboration d'un compilateur efficace pour le langage MGS et peut s'adapter à d'autres langages à base de règles de réécriture.

La section suivante donne une intuition du fonctionnement des collections topologiques et des transformations et introduit leur typage. La section 3 présente notre langage fortement typé et son système de types. La section 4 présente le langage à typage souple et énonce la correction de ce typage puis donne un aperçu de l'intérêt du typage pour la compilation des transformations. La dernière section conclut cet article en proposant des extensions directes de nos travaux et en présentant les travaux proches des nôtres et les perspectives ouvertes par notre travail.

2. Collections topologiques et transformations

Dans cette section nous introduisons les collections topologiques puis les transformations de manière informelle. Nous donnons également des éléments pour comprendre comment un système de types pour un langage fonctionnel peut les intégrer.

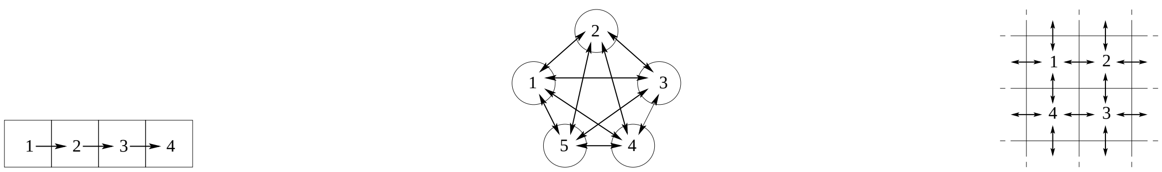
Collections topologiques

Une collection topologique est une structure de données sur laquelle il existe une relation de voisinage notée *Vois* entre les éléments. Lorsqu'on a $Vois(e_1, e_2)$ on dira que e_2 est un voisin de e_1. Par exemple, une *séquence* est une collection topologique telle que :

- chaque élément possède au plus un voisin ;
- chaque élément ne peut être le voisin que d'un élément au plus ;
- il n'existe pas de cycle dans la relation de voisinage.

Un ensemble ou un multi-ensemble peuvent être vus comme une collection dont tout élément est voisin de tous les autres éléments.

La *grille* est un autre exemple de collection topologique qui est similaire à une matrice. Chaque élément contenu dans une grille peut avoir quatre voisins se trouvant respectivement au nord, au sud, à l'est et à l'ouest. Contrairement aux tableaux, la grille est une structure de données partielle.



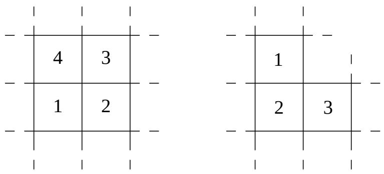
(a) séquence (b) ensemble (c) grille

FIG. 1 – exemples de collections topologiques

De nombreuses structures de données peuvent être considérées uniformément comme des collections topologiques, aussi bien des structures usuelles comme celles que nous venons de décrire que des structures plus spécialisées comme les graphes de Delaunay.

Dans le langage que nous proposons les collections peuvent être construites à partir de collections vides et d'opérateurs de construction. Par exemple 1 :: *empty_set* produit l'ensemble contenant l'entier 1. L'opérateur :: peut être utilisé pour construire n'importe quelle sorte de collection[1], ainsi on peut l'utiliser pour produire une séquence comme dans 1 :: 2 :: *empty_seq*.

Quatre opérateurs sont réservés à la construction de grilles : *nord*, $-$*nord*, *est* et $-$*est*. Ceux-ci permettent de spécifier l'organisation entre les éléments lors de la construction de la collection. Ainsi on peut construire une grille carrée : 1 *est* 2 *nord* 3 $-$*est* 4 :: *empty_grid* ou une grille triangulaire : 1 $-$*nord* 2 *est* 3 :: *empty_grid* par exemple.



Les systèmes de types que nous proposons dans cet article contiennent des types particuliers de la forme $[\tau]\rho$ pour les collections où τ est le type des éléments contenus dans la collection aussi appelé le *type contenu* de la collection et ρ est sa *topologie*. Une topologie peut être soit un symbole de l'ensemble $\{set,\ bag,\ seq,\ grid,\ \ldots\}$ qui représente les topologies possibles des collections, soit une variable de topologie que l'on pourra dénoter par la lettre grecque θ. Notons qu'une topologie n'est pas un type et qu'une variable de topologie ne peut être utilisée à la place d'une variable de type et *vice versa*. Les deux grilles ci-dessus par exemple ont le type $[int]grid$. Le constructeur *nord* et les trois autres constructeurs spécifiques aux grilles ont le type $\alpha \to [\alpha]grid \to [\alpha]grid$ où α est une variable de type. Le constructeur générique :: a quant à lui le type $\alpha \to [\alpha]\theta \to [\alpha]\theta$ car il peut être utilisé avec toute collection, quelle que soit sa topologie.

On parle de collections *hétérogènes* lorsque les valeurs contenues dans les collections peuvent être de types différents. L'ensemble $\{1,\ true\}$ contient deux valeurs de types respectifs *int* et *bool* et est donc un exemple de collection hétérogène. Le langage fortement typé ne permettra pas de manipuler des collections hétérogènes.

Pour rendre compte de l'hétérogénéité des collections dans notre système de types souple nous utilisons des types *unions*, déjà utilisés par d'autres auteurs [AW93, PS94, Dam94a, FCB02]. Une valeur du type union $\tau_1 \cup \tau_2$ est soit du type τ_1 soit du type τ_2. Savoir qu'une valeur est du type $\tau_1 \cup \tau_2$ ne permet pas de déduire qu'elle est du type τ_1. L'entier 1 du type *int* est aussi du type $int \cup bool$.

Nous pouvons à présent donner un type à l'ensemble $\{1,\ true\}$, ce type est $[int \cup bool]set$. On peut lire ce type de la manière suivante : «collection de topologie *set* qui contient des valeurs de type *int* et des valeurs de type *bool*» mais il est plus juste de le comprendre ainsi : «collection de topologie *set* dont les éléments ont le type $int \cup bool$».

L'utilisation des collections hétérogènes nécessite la possibilité de tester le type des valeurs à l'exécution. Ceci explique que le langage les manipulant ne puisse être entièrement typé à la compilation.

Transformations

Une transformation est une fonction opérant sur les collections définie par un ensemble de règles de réécriture de la forme $m \Rightarrow e$ appelées simplement *règles*. La partie gauche d'un règle est appelée *motif* et la partie droite *expression de remplacement*. On note une transformation par l'énumération de ses règles entre accolades : $\{m_1 \Rightarrow e_1; \ldots; m_n \Rightarrow e_n\}$.

[1] La sémantique de l'opérateur :: dépend de la topologie de la collection à laquelle il est appliqué. Cette forme de surcharge est de la même nature que la surcharge du = de ML.

L'application d'une transformation à une collection se fait en appliquant les règles de la transformation de la manière suivante : des instances disjointes du motif de la première règle sont recherchées puis, lorsqu'on ne peut plus trouver de nouvelle instance du motif on recherche des instances du motif de la seconde règle parmi les éléments qui n'ont pas déjà été filtrés, et ainsi de suite. Lorsque ce processus de filtrage est terminé on substitue les parties filtrées par les parties remplaçantes correspondantes et la nouvelle collection ainsi créée est retournée.

Le motif x filtre une valeur quelconque ; le motif $x : int$ filtre une valeur de type *int* (uniquement dans le langage à typage souple) ; le motif gardé $x/p(x)$ filtre une valeur v telle que $p(v)$ s'évalue à *true* ; le motif x, y filtre deux valeurs voisines quelconques ; enfin le motif $x : int,\ y : int\ /\ x > y$ filtre deux valeurs entières voisines telles que la première est plus grande que la seconde. Le processus de filtrage a été décrit formellement dans [GMC02].

Les transformations permettent d'exprimer simplement des programmes classiques comme le tri d'une séquence par exemple. Pour trier une séquence on peut chercher des couples d'éléments voisins mal ordonnés dans la séquence et les placer dans le bon ordre. Lorsqu'il n'existe plus de couples d'éléments voisins mal ordonnés la séquence est triée. Cette sorte de tri à bulles est obtenue en itérant l'application de la transformation suivante jusqu'à atteindre un point fixe : $\{x, y/x > y \Rightarrow [y, x]\}$.

Dans cet exemple, l'expression $[y, x]$ en partie droite de la règle dénote une séquence à deux éléments. En effet, lors du processus de filtrage, une instance du motif x, y est représentée par une suite de deux valeurs se trouvant en des positions voisines dans la collection. Cette suite de valeurs positionnées est appelée un *chemin* dans la collection. Ici le chemin est de longueur 2. Les valeurs dénotées par la séquence $[y, x]$ viendront remplacer les valeurs du chemin filtré par la partie gauche de la règle.

L'utilisation d'une séquence comme expression de remplacement convient quelle que soit la topologie de la collection à laquelle la transformation est appliquée. En effet, un motif filtre un chemin qui peut être vu comme une séquence de valeurs positionnées. Une séquence de valeurs est donc suffisante en partie droite pour spécifier le remplacement point à point des éléments filtrés. C'est pourquoi nous forçons les parties droites de règles à être des séquences.

Dans certaines collections comme les grilles la longueur de la séquence remplaçante doit être égale à la longueur du chemin filtré et la substitution se fera point à point. Dans le cas contraire, la topologie de la collection ne serait pas préservée : une valeur ne peut être remplacée par plusieurs valeurs dans une grille car il faudrait pour cela insérer de nouvelles positions et la collection ne serait plus une grille. Ces collections dont l'ensemble de positions n'est pas modifiable sont dites *newtoniennes*[2]. En revanche dans une séquence, un ensemble ou un multi-ensemble le chemin filtré peut être remplacé par un nombre arbitraire d'éléments car on peut toujours insérer une position entre deux positions dans ces collections. Ces collections dont l'ensemble de positions peut varier sont dites *leibnitziennes*.

La fonction *map* qui applique une fonction à tout élément d'une collection est un autre exemple de programme simple à écrire : $\lambda f.\lambda c.(let\ t = \{x \Rightarrow [f\ x]\}\ in\ (t\ c))$ ou de manière équivalente $\lambda f.\{x \Rightarrow [f\ x]\}$. Ceci implémente bien un map car chaque élément e de la collection sera filtré par le motif x et sera remplacé par $f(e)$. Cette fonction peut s'appliquer à toute collection, indépendamment de sa topologie. De telles fonctions sont dites *polytypiques* [JJ96]. Le polytypisme est l'un des avantages à considérer les structures de données dans un cadre unificateur.

L'identité sur les collections peut s'écrire $\{x \Rightarrow [x]\}$ et a le type $[\alpha]\theta \to [\alpha]\theta$. En effet cette transformation s'applique à toute collection topologique et ne change ni sa topologie, ni son type contenu. De manière générale une transformation ne change pas la topologie de la collection à laquelle elle est appliquée. La transformation $\{x : int \Rightarrow [x + 1]\}$ où $+$ est l'addition entière a également le type $[\alpha]\theta \to [\alpha]\theta$ (dans le typage souple). En revanche la transformation $\{x \Rightarrow [x + 1]\}$ a le type $[int]\theta \to [int]\theta$ dans les deux systèmes car une erreur de type aura lieu si la règle est appliquée à une

[2]Cette appellation vient de la vision différente de la notion d'espace selon Leibnitz ou Newton [Gia03].

valeur qui n'est pas du type *int*.

Considérons à présent la transformation $\{x : int \ / \ (x \ mod \ 2 = 0) \Rightarrow [true]\}$ dans le langage à typage souple. Celle-ci est du type $[\alpha]\theta \rightarrow [\alpha \cup bool]\theta$ mais ce type manque de précision. Il ne porte pas l'information que des booléens apparaissent dans la collection renvoyée uniquement si la collection en argument contient des entiers. Pour traduire cette information nous utiliserons des types *conditionnels* de la forme $\tau_1?\tau_2$. Le type $\tau_1?\tau_2$ se lit «τ_1 *if* τ_2» et vaut τ_1 lorsque τ_2 n'est pas égal au type nul noté 0 et il vaudra 0 sinon. Ainsi on peut donner le type $[\alpha]\theta \rightarrow [\alpha \cup (bool?(\alpha \cap int))]\theta$ à la transformation mentionnée ci-dessus. Ce type signifie que si la collection en argument contient des valeurs de type *int* alors la collection renvoyée pourra contenir des valeurs de type *bool* car dans ce cas $\alpha \cap int$ n'est pas nul et vaut *int*. En revanche si le type de la collection en argument indique qu'elle ne contient pas de valeurs du type *int* alors le type de la collection renvoyée est le même que le type de la collection en argument. En effet dans ce cas $\alpha \cap int$ sera nul et donc $bool?(\alpha \cap int)$ sera également nul et $\alpha \cup bool?(\alpha \cap int)$ vaudra α.

Le typage fin des transformations avec des types conditionnels peut être vu comme de l'analyse de flots. On peut noter que les types conditionnels ont déjà été utilisés à cette fin (voir [FA97]).

3. Typage fort

Le premier langage que nous présentons, noté $\mathcal{L}_=$ permet de manipuler des collections homogènes et des transformations. Il a pour objet de montrer que les collections topologiques et les transformations s'intègrent bien dans un langage fortement typé comme ML. Nous présentons donc un système de types pour $\mathcal{L}_=$ qui permet l'inférence automatique des types par une extension de l'algorithme de Damas/Milner.

3.1. Langage $\mathcal{L}_=$

Le langage $\mathcal{L}_=$ est un λ-calcul avec constantes et *let* auquel on ajoute les transformations.

$$e \quad ::= \quad x \mid c \mid \lambda x.e \mid e \ e \mid let \ x \ = \ e \ in \ e \mid \{m/e \Rightarrow e; \ldots; m/e \Rightarrow e; x \Rightarrow e\}$$

$$m \quad ::= \quad x, \ \ldots \ , x$$

Une transformation est composée d'une suite de règles dont la dernière dispose d'un motif qui se réduit à une variable. Une telle règle de la forme $x \Rightarrow e$ est appelée *attrape-tout*. Cette dernière règle permet de garantir que toutes les valeurs d'une collection seront filtrées par la transformation. Ainsi si les règles d'une transformation remplacent des entiers par des flottants par exemple on est sûr que tous les entiers seront remplacés et on peut garantir que la transformation préserve l'homogénéité des collections.

Parmi les constantes du langage on trouve notamment des collections vides (*empty_seq*, *empty_grid*, ...) et des opérateurs comme le constructeur générique de collections noté :: ou le constructeur spécialisé *nord*.

Nous ferons un grand usage de sucre syntaxique, notamment :
- les opérateurs binaires seront écrits en position infixe,
- on pourra écrire une séquence en énumérant ses éléments entre crochets au lieu d'utiliser le constructeur standard comme le montre l'exemple suivant : $[1, 2, 3]$ pour $1 :: 2 :: 3 :: empty\_seq$,
- on pourra omettre la garde d'un motif lorsque celle-ci est la constante *true*.

Nous ne donnons pas ici la sémantique formelle de $\mathcal{L}_=$. Le lecteur pourra se référer à [GM02] ou à [GM01] pour le modèle des collections topologiques et à [Coh03a] pour la sémantique des transformations. Deux valeurs particulières dénotent des erreurs : *wrong* lorsqu'une erreur de type survient et *shape_err* lorsque l'application d'une règle viole la topologie d'une collection newtonienne.

3.2. Types

Nous associons à $\mathcal{L}_=$ un système de types à la Hindley/Milner augmenté des types collections. L'ensemble des topologies possibles pour une collection topologique est noté R et contient au moins *set*, *seq*, *bag* et *grid*. On note B l'ensemble des types de base $\{int,\ bool,\ float,\ string\}$.

$$\tau ::= \alpha \mid B \mid \tau \to \tau \mid [\tau]\rho \qquad\qquad \rho ::= \theta \mid R$$

La lettre θ sera utilisée pour désigner une variable de topologie.

Un schéma de types σ est un type quantifié sur des variables de type et des variables de topologie de la forme $\forall\alpha_1,\ldots,\alpha_n,\theta_1,\ldots,\theta_m.\tau$. On dit qu'un type τ est une instance d'un schéma de types $\sigma = \forall\alpha_1,\ldots,\alpha_n,\theta_1,\ldots,\theta_m.\tau'$ s'il existe une instanciation s des variables quantifiées de σ telle que $s(\tau') = \tau$ et on le note $\sigma \leq \tau$.

La fonction TC donne le schéma de type associé aux constantes du langage. Par exemple $TC(::)$ vaut $\forall\alpha,\theta.\alpha \to [\alpha]\theta \to [\alpha]\theta$.

Un *contexte* de typage Γ est une fonction d'un ensemble de variables du langage vers l'ensemble des schémas de types.

Règles de typage

Les règles d'inférence sont celles d'Hindley/Milner augmentées d'une règle pour les transformations.

$$\frac{\Gamma(x) \leq \tau}{\Gamma \vdash x : \tau}\,(var - inst) \qquad \frac{TC(c) \leq \tau}{\Gamma \vdash c : \tau}\,(const - inst)$$

$$\frac{\Gamma \cup \{x : \tau_1\} \vdash e : \tau_2}{\Gamma \vdash (\lambda x.e) : \tau_1 \to \tau_2}\,(fun) \qquad \frac{\Gamma \vdash e_1 : \tau' \to \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1\ e_2 : \tau}\,(app)$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma \cup \{x : Gen(\tau_1, \Gamma)\} \vdash e_2 : \tau_2}{\Gamma \vdash (let\ x = e_1\ in\ e_2) : \tau_2}\,(let)$$

$$\frac{\Gamma_i \vdash e_i : [\tau']seq \quad \Gamma_i \vdash g_i : bool \quad (1 \leq i \leq n)}{\Gamma \vdash \{m_1/g_1 \Rightarrow e_1; \ldots; m_n/g_n \Rightarrow e_n\} : [\tau]\rho \to [\tau']\rho}(trans)$$

où $\Gamma_i = \Gamma \cup \{self : [\tau]\rho\} \cup \gamma(m_i, \tau)$ et avec γ définie par $\gamma((x_1, \ldots, x_k), \tau) = \{x_1 : \tau, \ldots, x_k : \tau\}$.

Comme usuellement la fonction *Gen* généralise un type τ en schéma de type $\forall\alpha_1,\ldots,\alpha_n,\theta_1,\ldots,\theta_m.\tau$ où les variables de type et de topologie quantifiées sont les variables du type qui ne sont pas liées dans le contexte de typage.

Dans la règle *(trans)* on type toutes les règles comme si elles avaient la même forme bien que la dernière règle n'ait pas de garde. Ceci est naturel car le motif x est équivalent au motif $x/true$. Intuitivement la règle *(trans)* exprime le fait suivant : si la partie droite de chaque règle est une séquence de type contenu τ' lorsqu'on suppose que les variables liées dans son motif ont le type τ alors la transformation renvoie une collection de type contenu τ' lorsqu'elle est appliquée à une collection de type τ. Ceci est vrai grâce à la règle attrape-tout rendue obligatoire par la syntaxe des transformations. La collection renvoyée a la même topologie que la collection en argument.

À l'intérieur d'une transformation l'identificateur *self* est lié à la collection à laquelle la transformation est appliquée. On peut remarquer que si *self* n'est pas utilisé dans le corps de la transformation celle-ci pourra toujours être polytypique.

Si $\emptyset \vdash e : \tau$ (noté aussi $\vdash e : \tau$) alors l'évaluation de e ne provoquera pas d'erreur de type *wrong*. En revanche, l'absence d'erreur de structure newtonienne n'est pas assurée.

Exemple

On peut montrer avec les règles de typage que la transformation suivante a le type $[int]\theta \rightarrow [int]\theta$ pour toute topologie θ : $\{x, y/x > y \Rightarrow [x, y, (x - y)]; x \Rightarrow [x]\}$

La preuve est donnée ci-dessous avec $\Gamma_1 = \{x : int;\ y : int;\ self : [int]\theta\}$ et $\Gamma_2 = \{x : int;\ self : [int]\theta\}$.

$$\dfrac{\dfrac{\cdots}{\Gamma_1 \vdash x > y : bool} \quad \dfrac{\dfrac{\Gamma_0(x) \leq int}{\Gamma_1 \vdash x : int} \quad \dfrac{\cdots}{\Gamma_1 \vdash [y, (x-y)] : [int]seq}}{\Gamma_1 \vdash [x, y, (x-y)] : [int]seq} \quad \dfrac{\dfrac{\Gamma_2(x) \leq int}{\Gamma_2 \vdash x : int}}{\Gamma_2 \vdash [x] : [int]seq}}{\vdash \{x, y/x > y \Rightarrow [x, y, (x-y)]; x \Rightarrow [x]\} : [int]\theta \rightarrow [int]\theta}$$

3.3. Inférence automatique

L'algorithme d'inférence de type automatique $\mathcal{W}$ de Damas/Milner s'étend simplement au langage $\mathcal{L}_=$ et aux règles de typage correspondantes. Il suffit pour cela d'étendre la procédure d'unification de Robinson aux types collections et aux topologies ainsi que d'ajouter un cas dans $\mathcal{W}$ pour le typage des transformations. Cet algorithme est donné dans [Coh03b] et a été implémenté afin d'être intégré à un compilateur pour une version fortement typée du langage MGS. Comme $\mathcal{W}$, il calcule le type le plus général d'un programme.

4. Typage souple

Le langage $\mathcal{L}_=$ ne permet la manipulation des collections que dans un cadre homogène. Nous présentons à présent le langage $\mathcal{L}_\subseteq$ qui est presque identique à $\mathcal{L}_=$ mais qui permettra des tests dynamiques de type afin de manipuler des collections hétérogènes. Nous associons à ce langage un système de types approprié qui effectue un typage statique tout en laissant certains tests de types à l'exécution. Ce système de types plus avancé que le premier utilise des types union et du sous-typage non-structurel. La procédure d'inférence automatique est plus complexe que celle basée sur $\mathcal{W}$.

4.1. Langage $\mathcal{L}_\subseteq$

Au niveau syntaxique $\mathcal{L}_\subseteq$ diffère de $\mathcal{L}_=$ par la possibilité de tester dynamiquement le type d'une valeur durant le filtrage et la liberté d'avoir une règle attrape-tout dans la transformation ou non. Les tests de type sont spécifiés par annotation des variables des motifs comme le montre la syntaxe ci-dessous. Ces conditions de types sont restreintes aux types de base. La construction μ est appelée *motif élémentaire* et b désigne un type de base de $B = \{int,\ bool,\ float,\ string\}$.

$$e \quad ::= \quad x \mid \lambda x.e \mid e\ e \mid c \mid \text{let } x = e \text{ in } e \mid \{m/e \Rightarrow e; \ldots; m/e \Rightarrow e\}$$

$$m \quad ::= \quad \mu, \ldots, \mu \qquad\qquad \mu \quad ::= \quad x \mid x : b$$

Le langage $\mathcal{L}_\subseteq$ s'évalue dans un domaine D qui contient les collections topologiques en plus des valeurs usuelles (voir [GS90] pour une introduction aux domaines sémantiques et [GM01] ou [GM02] pour le modèle des collections topologiques). Les transformations sont représentées dans D par des fonctions continues[3]. Leur sémantique est donnée dans [Coh03a]. D contient les valeurs spéciales *wrong* et *shape_err* ainsi que $\perp$ qui dénote un calcul qui ne termine pas.

[3]On peut considérer que les transformations sont déterministes en supposant que la stratégie d'application d'une règle est fixée mais non spécifiée.

À tout type de base b on associe un sous-ensemble de $D - \{wrong, shape\_err\}$ noté $[\![b]\!]$ et contenant $\perp$. L'intersection de ces sous-ensembles deux à deux doit toujours valoir $\{\perp\}$.

Un *environnement* est une fonction d'un ensemble d'identifiants vers l'ensemble des valeurs D. On note $Eval(e, E)$ la sémantique de l'expression e dans l'environnement E.

4.2. Types

La syntaxe des types est la suivante :

$$\tau \quad ::= \quad B \mid \alpha \mid \tau_1 \to \tau_2 \mid [\tau]\rho \mid \tau_1 \cup \tau_2 \mid \tau_1 \cap \tau_2 \mid 0 \mid 1 \mid \tau_1?\tau_2$$

$$\rho \quad ::= \quad R \mid \theta$$

L'ensemble R est le même que pour $\mathcal{L}_=$. La sémantique des types est basée sur la notion d'*idéal* [MPS86] : un type correspond à un sous ensemble particulier du domaine des valeurs D. La relation de sous-typage est notée $\subseteq$ et correspond à l'inclusion ensembliste sur D. Un type ne peut contenir ni $wrong$ ni $shape\_err$.

Voici une interprétation intuitive des types avant que nous ne donnions leur sémantique formelle.

- Le type *flèche*, les types de base et les variables de type sont interprétés comme usuellement dans les langages fonctionnels.
- Dans le type collection $[\tau]\rho$, le type contenu est τ et la topologie est ρ comme pour $\mathcal{L}_=$.
- Les types $\tau_1 \cup \tau_2$ et $\tau_1 \cap \tau_2$ correspondent à l'union et l'intersection ensembliste des types.
- Le type 0 contient uniquement la valeur $\perp$, qui représente la non-terminaison. Le type 0 est inclus dans tous les autres types car $\perp$ appartient à tous les types. Le type 1 contient toutes les valeurs sauf *wrong* et *shape_err*. Le type 1 inclut tous les autres types. Le type $0 \to 1$ convient à toute fonction, y-compris aux transformations.
- Le type conditionnel $\tau_1?\tau_2$ vaut τ_1 lorsque τ_2 est différent de 0 et il vaut 0 sinon. Par exemple le type $int?(\tau \cap float)$ vaut int si τ contient $float$ et 0 sinon.

Un schéma de type σ est de la forme $\forall \alpha_1, \ldots, \alpha_n, \theta_1, \ldots, \theta_m.\tau \ where \ S$ où S est un ensemble de contraintes de types de la forme $\tau_1 \subseteq \tau_2$. On utilisera par la suite le symbole χ pour désigner indifféremment une variable de type ou une variable de topologie. On note $Sol(S)$ l'ensemble des solutions de S.

Interprétation sémantique des types

Étant donné une instanciation s des variables de type et de topologie, la sémantique d'un type et d'un schéma de types sont définis dans la figure 2 par la fonction $[\![.]\!]_s$.

La sémantique d'un type de base b est l'ensemble $[\![b]\!]$ défini en section 4.1. La sémantique du type $\tau_1 \to \tau_2$ est l'ensemble des fonctions continues de D vers D telles que $f(v)$ est dans $[\![\tau_2]\!]_s$ (ou provoque une erreur différente de *wrong*) si v est dans $[\![\tau_1]\!]_s$. La sémantique de $\tau_1 \cup \tau_2$ est l'union de la sémantique de τ_1 et de la sémantique de τ_2. La sémantique du type $[\tau]\rho$ est l'ensemble des collections de D dont la topologie correspond à ρ et dont les éléments sont dans $[\![\tau]\!]_s$. La sémantique du schéma de types $\forall \chi_1, \ldots, \chi_n.\tau \ where \ S$ est l'ensemble des valeurs qui sont dans $[\![\tau]\!]_{s'}$ pour toute instanciation s' des χ_i solution de S (s' doit être compatible avec s sur les variables non quantifiées).

À chaque constante c du langage on associe un schéma de types $TC(c)$. On suppose que TC est correct par rapport à la sémantique : pour toute constante c du langage et pour toute instanciation s des variables de type et de topologie, $Eval(c, \emptyset) \in [\![TC(c)]\!]_s$.

$$
\begin{array}{rcl}
[\![\alpha]\!]_s & = & [\![s(\alpha)]\!]_\emptyset \\
[\![b]\!]_s & = & [\![b]\!] \\
[\![\tau_1 \to \tau_2]\!]_s & = & \{f \in D \to D \mid f([\![\tau_1]\!]_s - \{\bot\}) \subseteq [\![\tau_2]\!]_s \cup \{shape\_err\}\} \cup \{\bot\} \\
[\![\tau_1 \cup \tau_2]\!]_s & = & [\![\tau_1]\!]_s \cup [\![\tau_2]\!]_s \\
[\![\tau_1 \cap \tau_2]\!]_s & = & [\![\tau_1]\!]_s \cap [\![\tau_2]\!]_s \\
[\![\tau_1 ? \tau_2]\!]_s & = & \begin{cases} [\![\tau_1]\!]_s & \text{si } [\![\tau_2]\!]_s \neq \{\bot\} \\ \{\bot\} & \text{sinon} \end{cases} \\
[\![0]\!]_s & = & \{\bot\} \\
[\![1]\!]_s & = & D - \{wrong, shape\_err\} \\
[\![[\tau]\rho]\!]_s & = & \{c \in D \mid s(\rho) \text{ est la topologie de } c \text{ et } \forall e \in c.e \in [\![\tau]\!]_s\} \\
 & & \\
[\![\forall \alpha_1, \dots, \alpha_n, \theta_1, \dots, \theta_m.\tau \ where \ S]\!]_s & = & \bigcap_{s' \in X} [\![\tau]\!]_{s'} \\
 & & \text{où } X = Sol(S) \cap \\
 & & \qquad \{s' | s'(\chi) = s(\chi) \text{ si } \chi \notin \{\alpha_1, \dots, \alpha_n, \theta_1, \dots, \theta_m\}\}
\end{array}
$$

FIG. 2 – Sémantique des types et des schémas de types

4.3. Règles de typage

La figure 3 donne notre système de règles de typage pour $\mathcal{L}_\subseteq$. La relation de typage comporte un contexte Γ et un ensemble S de contraintes de types. Le jugement $\Gamma, S \vdash e : \tau$ peut se lire «dans le contexte de typage Γ, l'expression e a le type $s(\tau)$ pour toute solution s de S». La présence d'un ensemble de contraintes dans la relation de typage est standard dans les systèmes de types en présence de sous-typage. Toutes les règles sauf *(const)* et *(trans)* sont similaires à celles de Aiken *et al.*

Voici une description des règles de la figure 3 :

(var) : Cette règle correspond à la règle standard de Hindley/Milner.

(const) : TC donne les schémas de types des constantes.

(fun) et (app) : La règle *(fun)* correspond à celle de Hindley/Milner. Dans *(app)* les contraintes expriment qu'une fonction du type $\tau_3 \to \tau_4$ ne peut être appliquée à une valeur du type τ_2 que si τ_2 est un sous-type de τ_3.

(gen) et (inst) : La règle *(gen)* sert à introduire le polymorphisme paramétrique dans les types. En effet elle exprime que si une expression a le type τ sous les contraintes S alors elle a aussi le schéma de type $\forall(\chi_i).\tau \ where \ S$ où les χ_i sont des variables libres de τ. La règle *(inst)* sert à instancier les schémas de types en types. Pour utiliser cette règle, les contraintes du schéma de types doivent avoir une solution. Les τ_i et ρ_j sont libres dans cette règle.

(let) : Le *let-polymorphisme* est obtenu en appliquant la règle *(gen)* juste après la règle *(let)*.

(trans) : Dans cette règle nous utilisons deux fonctions définies inductivement sur les motifs : *Comp* et γ. La première, *Comp* s'applique à un motif m et à un type τ et renvoie un type qui vaudra toujours $\{\bot\}$ si le motif ne peut s'appliquer dans une collection de type contenu τ et qui vaudra un type différent de $\{\bot\}$ sinon. On dira que cette fonction calcule la ***compatibilité*** d'un motif avec un type. Par exemple $Comp\ ((x_1 : int,\ x_2 : float),\ \tau) = (\tau \cap int)\ ?\ (\tau \cap float)$. Ainsi si τ ne contient pas *int* et *float* ce type vaut $\{\bot\}$. La seconde fonction, γ, calcule le contexte induit par un motif m sachant que ce motif est appliqué à une collection de type contenu τ.

Pour une règle $m_i/g_i \Rightarrow e_i$ dans *(trans)*, τ_i est le type contenu de la séquence e_i sachant que la transformation s'applique à une collection de type contenu τ et en considérant le contexte induit par m_i. Le type $\tau_i ? Comp(m_i, \tau)$ vaudra $\{\bot\}$ si les conditions de types de m_i font que le motif n'a jamais d'instances dans une collection de type contenu τ (incompatibilité), il vaudra τ_i si la règle peut s'appliquer (compatibilité).

Le type contenu de la collection renvoyée doit être un sur-type de $\tau_i?Comp(m_i, \tau)$ pour chaque i, d'où les contraintes $\tau_i?Comp(m_i, \tau) \subseteq \tau'$ car si la règle peut s'appliquer, la collection renvoyée pourra contenir des éléments de type τ_i. Par ailleurs lors de l'application d'une transformation des valeurs peuvent ne pas être filtrées et resteront dans la collection renvoyée, d'où la contrainte[4] $\tau \subseteq \tau'$.

$$\frac{}{\Gamma \cup \{x : \sigma\}, S \vdash x : \sigma}\ (var) \qquad \frac{\Gamma \cup \{x : \tau_1\}, S \vdash e : \tau_2}{\Gamma, S \vdash \lambda x.e : \tau_1 \to \tau_2}\ (fun)$$

$$\frac{}{\Gamma, S \vdash c : TC(c)}\ (const) \qquad \frac{\Gamma, S \vdash e_1 : \tau_1,\ e_2 : \tau_2}{\Gamma, S \cup \{\tau_2 \subseteq \tau_3, \tau_1 \subseteq \tau_3 \to \tau_4\} \vdash e_1\ e_2 : \tau_4}\ (app)$$

$$\frac{\Gamma, S \vdash e : \tau}{\Gamma, \emptyset \vdash e : \forall \chi_1, \ldots, \chi_n.\tau\ where\ S}\ (gen) \text{ si } Sol(S) \neq \emptyset \text{ et } \chi_1, \ldots, \chi_n \text{ non libres dans } \Gamma$$

$$\frac{\Gamma, S \vdash e : \forall \alpha_1, \ldots, \alpha_n, \theta_1, \ldots, \theta_m.\tau\ where\ S'}{\Gamma, S \cup S'[\tau_i/\alpha_i, \rho_j/\theta_j] \vdash e : \tau[\tau_i/\alpha_i, \rho_j/\theta_j]}\ (inst) \qquad \frac{\Gamma, S \vdash e_1 : \sigma \quad \Gamma \cup \{x : \sigma\}, S \vdash e_2 : \tau}{\Gamma, S \vdash \text{let } x = e_1 \text{ in } e_2 : \tau}\ (let)$$

$$\frac{\Gamma_i, S \vdash g_i : bool \qquad \Gamma_i, S \vdash e_i : [\tau_i]seq \qquad (1 \leq i \leq n)}{\Gamma, S \cup S' \vdash \{m_1/g_1 \Rightarrow e_1; \ldots; m_n/g_n \Rightarrow e_n\} : [\tau]\rho \to [\tau']\rho}\ (trans)$$

où $S' = \{\tau \subseteq \tau'\} \cup \bigcup_{1 \leq i \leq n} \{\tau_i?Comp(m_i, \tau) \subseteq \tau'\}$ et $\Gamma_i = \Gamma \cup \{self : [\tau]\rho\} \cup \gamma(m_i, \tau)$.

$$\begin{array}{llll}
Comp\ (x,\ \tau) & = \tau \cap 1 = \tau & \gamma(x,\ \tau) & = \{x : \tau\} \\
Comp\ (x : b,\ \tau) & = \tau \cap b & \gamma(x : b,\ \tau) & = \{x : \tau \cap b\} \\
Comp\ ((x, m'),\ \tau) & = Comp\ (m', \tau) & \gamma((x, m'),\ \tau) & = \{x : \tau\} \cup \gamma(m',\ \tau) \\
Comp\ ((x : b, m'),\ \tau) & = (\tau \cap b)\ ?\ Comp\ (m', \tau) & \gamma((x : b, m'),\ \tau) & = \{x : \tau \cap b\} \cup \gamma(m',\ \tau)
\end{array}$$

FIG. 3 – Règles de typage

Exemples

Voici les preuves de typage de deux transformations simples.

$$\cfrac{\cfrac{\cfrac{\ldots}{\{self : [\tau]\theta; x : \alpha \cap int\}, \emptyset \vdash [x; 1] : [int]seq}}{\emptyset, \{\alpha \subseteq \alpha, (int?(\alpha \cap int)) \subseteq \alpha\} \vdash \{x : int \Rightarrow [x; 1]\} : [\alpha]\theta \to [\alpha]\theta}}{\emptyset, \emptyset \vdash \{x : int \Rightarrow [x; 1]\} : \forall \alpha, \theta.[\alpha]\theta \to [\alpha]\theta\ where\ \{\alpha \subseteq \alpha, (int?(\alpha \cap int)) \subseteq \alpha\}}$$

Dans cet exemple l'ensemble de contraintes du schéma se réduit à $\emptyset$. En effet on peut montrer que $(int?(\alpha \cap int)) \subseteq \alpha$ est toujours vrai. Donc n'importe quelle instanciation de α et θ convient.

[4]On peut ne pas considérer cette contrainte lorsque l'on sait détecter que toutes les valeurs seront filtrées, afin d'avoir un type plus précis. Par exemple dans $\mathcal{L}_=$ la règle attrape-tout assure que tous les éléments sont filtrés.

$$\dfrac{\dfrac{\dfrac{\cdots}{\{self : [\tau]\theta; x : \alpha \cap int\}, \emptyset \vdash [true] : [bool]seq}}{\emptyset, \{\alpha \subseteq \beta, (bool?(\alpha \cap int)) \subseteq \beta\} \vdash \{x : int \Rightarrow [true]\} : [\alpha]\theta \to [\beta]\theta}}{\emptyset, \emptyset \vdash \{x : int \Rightarrow [true]\} : \forall \alpha, \beta, \theta.[\alpha]\theta \to [\beta]\theta \ where \ \{\alpha \subseteq \beta, (bool?(\alpha \cap int)) \subseteq \beta\}}$$

Ici le type $\alpha \cup (int?(\alpha \cap int))$ est la plus petite instanciation de β vérifiant les contraintes du schéma. Le type suivant est donc valable pour cette transformation : $[\alpha]\theta \to [\alpha \cup (bool?(\alpha \cap int))]\theta$.

Sans l'utilisation de types conditionnels, le type le plus précis pour cette transformation aurait été $[\alpha]\theta \to [\alpha \cup bool]\theta$ qui porte moins d'informations que le type précédent.

4.4. Propriétés

On dira qu'un environnement E est correct par rapport à un contexte Γ et une instanciation s des variables de type et de topologie lorsque $E(x) \in [\![\Gamma(x)]\!]_s$ pour tout x lié dans Γ et E.

Lemme 1 (Correction) Soit un typage $\Gamma, S \vdash e : \sigma$, une solution s de S, un environnement E correct par rapport à Γ et s portant sur les variables libres de e. Alors $Eval(e, E) \in [\![\sigma]\!]_s \cup \{shape\_err\}$.

La preuve est donnée dans [Coh03a]. Le corollaire suivant découle de ce lemme : si e est une expression sans variables libres et si $\emptyset, \emptyset \vdash e : \sigma$ alors $Eval(e, \emptyset) \neq wrong$. Ceci est vrai car $wrong$ n'appartient à aucun type. On dit qu'un programme e est *bien typé* si il existe un schéma de type σ tel que $\emptyset, \emptyset \vdash e : \sigma$. Notons que le lemme ne garantit rien sur les erreurs de structure newtonienne.

4.5. Algorithme d'inférence automatique

L'inférence automatique des types d'un programme consiste en deux étapes. En premier lieu, le schéma de type le plus général du programme est calculé en appliquant les règles de typage selon une stratégie appropriée. Ensuite on calcule les solutions des contraintes de types générées. Nous décrivons ces deux étapes dans cette section.

4.5.1. Production du type et des contraintes

Nous suivons la stratégie d'application des règles proposée par [AW93] qui définit la dérivation la plus générale modulo renommage des variables de type et de topologie :
- On utilise des variables fraîches partout où cela est possible.
- On applique la règle *(gen)* immédiatement après la règle *(let)*.
- On applique *(inst)* après avoir appliqué la règle *(var)* ou la règle *(const)*.
- La dérivation se termine par une application de la règle *(gen)*.
- Les règles *(gen)* et *(inst)* ne sont appliquées nulle part ailleurs.

4.5.2. Résolution des contraintes

Une procédure de résolution de systèmes de contraintes ensemblistes est donnée dans [AW93] et [ALW94]. Elle est basée sur un système de réécriture qui met les contraintes sous une forme où leurs solutions peuvent être directement extraites. Cette procédure s'étend aux types collections en considérant l'équivalence suivante que l'on orientera de gauche à droite et en résolvant les égalités entre topologies par unification : $\{[\tau_1]\rho_1 \subseteq [\tau_2]\rho_2\} \equiv \{\rho_1 = \rho_2 \ ; \ \tau_1 \subseteq \tau_2\}$.

La présence d'opérateurs typés dans notre langage impose également de modifier l'une des règles de réécriture de la procédure d'Aiken *et al.* : $\{\tau_1 \to \tau'_1 \subseteq \tau_2 \to \tau'_2\}$ se réécrit en $\{\tau_1 \subseteq \tau'2 \; ; \; \tau_2 \subseteq \tau'1\}$.

La procédure de résolution proposée n'est correcte que sous une condition sur la forme du système à résoudre, laquelle est détaillée dans [Coh03a]. Nous montrons dans ce même document que notre procédure de production de contraintes ne produit que des systèmes solvables par notre procédure de résolution. On peut noter que cette condition empêche de donner d'utiliser des types intersections pour dénoter la surcharge d'opérateurs. Cependant nous verrons en section 5.2 que la surcharge peut s'exprimer autrement dans notre système de types.

4.6. Compilation

En plus du gain en performances attendu lorsqu'on passe d'un langage dynamiquement typé à un langage statiquement typé, le typage de $\mathcal{L}_\subseteq$ apporte une information pouvant s'apparenter à de l'analyse de flot de contrôle, permettant des optimisations dans le processus d'application des transformations. Dans cette section, nous esquissons certaines de ces optimisations.

Considérons la règle $x : int => [x + 1]$ et une collection c de type contenu τ. Deux cas particuliers peuvent se présenter :
- Si d'après le type τ la collection c ne contient pas d'entiers alors la règle ne peut s'appliquer. De manière générale, une règle $m/g \Rightarrow e$ ne peut s'appliquer si $Comp(m, \tau)$ vaut $\{\perp\}$. On sait donc dès la compilation qu'il est inutile d'essayer d'appliquer cette règle et on peut donc la sauter (élimination des règles inutiles).
- Si $\tau = int$ alors on sait que les valeurs de la collection vérifieront la condition de type du motif. Les tests de type sont donc inutiles à l'exécution (élimination des conditions de types inutiles).

Supposons à présent que les collections soient implémentées de manière à optimiser la recherche d'éléments lorsque leur type est connu. Par exemple les ensembles peuvent être implémentés par des sous-ensembles homogènes. Alors une optimisation est possible et elle généralise les deux premières : on peut chercher les instances d'un motif élémentaire dans la partie appropriée de la collection.

Les deux premières optimisations sont simples à implémenter mais des études sont nécessaires pour savoir si elles s'appliquent souvent dans les programmes réels. La dernière s'applique plus souvent mais il peut être difficile d'implémenter les collections de manière à favoriser à la fois la recherche en fonction du type et en fonction de la topologie de la collection. En effet le processus de filtrage est profondément lié à la topologie des collections puisqu'un motif filtre des valeurs voisines.

5. Conclusion

5.1. Comparaison des deux approches

Comme le montre [Wan87], l'algorithme $\mathcal{W}$ de Damas/Milner est équivalent à un algorithme procédant par production de contraintes d'égalités entre types suivie de résolution du système d'équations par unification de Robinson. Par conséquent les moteurs d'inférence automatique pour $\mathcal{L}_=$ et pour $\mathcal{L}_\subseteq$ peuvent être implémentés en suivant un même schéma production/résolution. Pour passer de la production de contraintes pour $\mathcal{L}_\subseteq$ à la production de contraintes pour $\mathcal{L}_=$ il s'agit essentiellement de remplacer les inclusions par des égalités dans les règles *(trans)* et *(app)*. Par ailleurs l'équivalence $\{X = Y\} \equiv \{X \subseteq Y, \; Y \subseteq X\}$ montre que l'on peut mêler dans un même système de contraintes des inclusions et des égalités. Une stratégie d'application de l'unification et de la résolution de Aiken *et al.* permet de résoudre de tels systèmes. Ces deux observations nous on conduit naturellement à implémenter les deux systèmes de types dans un seul moteur d'inférence automatique, en cours d'intégration dans un compilateur MGS expérimental.

5.2. Extensions

Nous présentons ici des extensions directes de nos travaux. Certaines font déjà partie de notre implémentation du système alors que d'autres sont des voies pour des travaux futurs.

Constructions usuelles. Les constructions usuelles telles que le produit ou le *let-rec* n'ont pas été considérées pour ne pas alourdir la présentation mais leur ajout au langage et au typage se fait de manière habituelle et ils sont présents dans notre implémentation.

Gardes dans le motif. Pour la simplicité de la présentation nous avons restreint l'usage de la garde dans un motif mais on peut étendre aisément le langage des motifs afin d'avoir une garde associée à chaque motif élémentaire, comme fait dans [Coh03b] et garder les propriétés du typage fort ou du typage souple. Mettre des gardes au plus tôt dans le motif permet d'optimiser le processus de filtrage. Par exemple le filtrage du motif $(x/x>0), y$ est plus rapide que pour $x, y/x>0$. Cette extension fait partie de notre implémentation.

Des types plus précis. Nous avons vu que le système de types de $\mathcal{L}_{\subseteq}$ permettait d'inférer des types assez précis. Pourtant dans au moins deux cas que nous montrons ici des types plus précis existent.

- Le type inféré pour la transformation $\{x : \mathit{int} \Rightarrow [\mathit{true}]\}$ est $[\alpha]\theta \to [\alpha \cup (\mathit{bool}?(\alpha \cap \mathit{int}))]\theta$. Ce type ne porte pas l'information que la collection renvoyée ne contient plus d'entiers.
- Le type inféré pour la transformation $\{x : \mathit{int} \Rightarrow [x]; x : \mathit{int} \Rightarrow [\mathit{true}]\}$ est $[\alpha]\theta \to [\alpha \cup (\mathit{bool}?(\alpha \cap \mathit{int}))]\theta$. Or la deuxième règle ne s'applique jamais car tous les entiers sont consommés par la première donc le type $[\alpha]\theta \to [\alpha]\theta$ qui est plus précis convient.

La règle *(trans)* de $\mathcal{L}_{\subseteq}$ analyse sommairement les conditions dans lesquelles les règles peuvent s'appliquer (par le biais des types conditionnels et de la fonction *Comp*). Les deux exemples ci-dessus montrent que l'analyse des motifs et des transformations doit être plus subtile pour inférer le type le plus précis d'un programme. Par exemple l'amélioration proposée dans la note de bas de page numéro 4 permet d'obtenir le typage le plus précis pour la fonction *map* qui est $(\alpha \to \beta) \to [\alpha]\theta \to [\beta]\theta$ au lieu de $(\alpha \to \beta) \to [\alpha]\theta \to [\alpha \cup \beta]\theta$.

Étoile dans un motif. Les motifs tels que nous les avons présentés permettent uniquement de filtrer des parties de taille prédéterminée. Il existe toutefois des cas où le programmeur aimerait filtrer des parties de la collection de taille arbitraire, comme montré dans [GMC02]. Afin de permettre ceci, nous introduisons une nouvelle construction dans les motifs appelée l'*étoile* et notée $*$ qui exprime le filtrage d'un nombre d'éléments arbitraire. La grammaire des motifs élémentaires est modifiée comme suit :

$$\mu \ ::= \ x \mid x : b \mid \ * \ \mathit{as}\ x \mid b*\ \mathit{as}\ x$$

Dans le motif élémentaire $*\ \mathit{as}\ x$, l'identificateur x dénote la séquence des valeurs filtrées par l'étoile et peut être utilisé dans la garde du motif et dans l'expression de remplacement de la règle. Le fait que x dénote une séquence plutôt qu'une partie de la collection est en accord avec la contrainte d'avoir une séquence comme expression de remplacement.

Le typage de ces nouveaux motifs nécessite uniquement la modification des fonctions γ et *Comp* de la manière suivante :

$$\begin{array}{llll} \gamma(*\ \mathit{as}\ x, \tau) & = \{x : [\tau]\mathit{seq}\} & \mathit{Comp}(*\ \mathit{as}\ x, \tau) & = 1 \\ \gamma(b*\ \mathit{as}\ x, \tau) & = \{x : [\tau \cap b]\mathit{seq}\} & \mathit{Comp}(b*\ \mathit{as}\ x, \tau) & = \tau \cap b \end{array}$$

Les répétitions arbitraires sont intégrées à notre implémentation.

Direction à la place de la virgule dans un motif. Dans une structure de données comme la grille, on peut vouloir filtrer des éléments voisins selon une direction donnée. Ceci peut ce faire simplement comme dans le motif suivant : $x, y/(\mathit{nord\_nb}\ \mathit{self}\ x\ y)$ où $\mathit{nord\_nb}$ est la relation de voisinage correspondant au constructeur *nord*. Toutefois il est avantageux de placer cette information de voisinage au niveau syntaxique afin de l'utiliser efficacement durant le filtrage. Ainsi dans le langage MGS on écrit directement $x\ |\mathit{nord}>y$. La virgule a été remplacée par le constructeur *nord* encadré des symboles | et >. Cette nouvelle construction dans le langage ne nécessite pas de modification dans le typage puisqu'elle peut être vue comme du sucre syntaxique.

Lisibilité des types. Un schéma de types de $\mathcal{L}_\subseteq$ contient un ensemble de contraintes qui peuvent se révéler obscures pour le programmeur. F. Pottier a formalisé une notion de simplification de systèmes de contraintes dans [Pot98] dans le but de rendre les schémas de types plus lisibles par le programmeur. Une telle simplification nous semble essentielle dans un but d'aide au développement.

Surcharge. Comme le font remarquer Pantel et Sallé dans [PS94] les types conditionnels introduits par Aiken *et al.* peuvent servir à dénoter la surcharge de la manière suivante. Si $\{U_i \to V_i\}$ est l'ensemble des types d'une fonction surchargée alors on peut lui donner le type suivant dans notre système :

$$\forall\alpha.\alpha \to \bigcup_i (V_i?\alpha \cap U_i) \ \mathit{where}\ \{\alpha \subseteq \bigcup_i U_i\}$$

Exemple. On peut donner le type suivant à l'addition sur les entiers et les flottants :

$$\forall\alpha.\alpha \to \alpha \to (\mathit{int}?\alpha \cap \mathit{int}) \cup (\mathit{float}?\alpha \cap \mathit{float}) \ \mathit{where}\ \{\alpha \subseteq \mathit{int} \cup \mathit{float}\}$$

Types récursifs. Les types récursifs ne sont pas intégrés directement à la grammaire des types mais on peut définir des types récursifs à l'aide de contraintes de la manière suivante : le schéma de types $\forall\alpha.[\alpha]\mathit{bag}\ \mathit{where}\ \{\alpha = [\alpha]\mathit{bag} \cup \mathit{int}\}$ dénote les multi-ensembles pouvant contenir des entiers et des multi-ensembles contenant à leur tour des entiers et des multi-ensembles et ainsi de suite. Dans cet exemple on a utilisé l'égalité entre types définie par $\{\tau_1 = \tau_2\} \equiv \{\tau_1 \subseteq \tau_2, \tau_2 \subseteq \tau_1\}$.

5.3. Discussions

Erreurs de structure newtonienne. La faiblesse de nos systèmes de types est qu'ils ne permettent pas de détecter les violations de structures newtoniennes à la compilation. Pour le faire il faudrait garantir que le chemin filtré et la séquence remplaçante ont la même taille. Or ceci n'est pas toujours possible car :

- la taille de la séquence remplaçante ne peut pas toujours être calculée à la compilation ;
- la taille du motif n'est pas connue à la compilation si une répétition arbitraire $*$ y est utilisée.

Toutefois il existe des cas où l'on peut détecter que *shape_err* sera renvoyé ou ne sera pas renvoyé. Déterminer l'ensemble des programmes pour lesquels l'apparition d'une *shape_err* est décidable est souhaitable mais notre système de types est mal adapté à ceci. Il existe également des systèmes de types prévus pour l'inférence automatique de tailles [Gia92, CK01].

Types sommes ou types unions ? Les types unions et les tests de type à l'exécution peuvent sembler superflus lorsque l'on sait que les types sommes avec constructeurs permettent de programmer de manière similaire dans un contexte statiquement typé. Les collections hétérogènes et les types unions sont bien plus flexibles que les types sommes mais permettent moins de vérifications automatiques, lesquelles sont essentielles à la construction de logiciels sûrs. Par exemple ils ne permettent pas de vérifier l'exhaustivité du filtrage, contrairement aux type sommes (voir [Mar03]).

Toutefois, dans certains domaines d'application, la complexité des processus modélisés est telle que les types sommes deviennent trop lourds. D'autres auteurs partagent cette idée. Par exemple J. Garrigue étend les types sommes aux types sommes polymorphes dont les types ne sont pas nécessairement déclarés et dont les constructeurs peuvent être utilisés dans différents types [Gar98]. Il les utilise intensivement pour interfacer des bibliothèques C (comme OpenGL ou GTK) au langage O'CAML. Les types sommes polymorphes apparaissent dans la version 3 de O'CAML [LDGV02]. Par ailleurs A. Frisch *et al.* utilisent des types unions pour typer le langage CDUCE dédié à la manipulation de données XML [FCB02]. Il existe d'autres exemples où la manipulation de données venant de l'extérieur nécessite un langage supportant l'hétérogénéité mais nous sommes particulièrement motivés par des programmes issus de la biologie tels que la simulation de cellules. Dans ce type de simulation on doit gérer des entités qui migrent entre compartiments et des réactions pouvant avoir lieu sans que l'ensemble des entités pouvant être rencontrées ne soit connu *a priori*. La définition d'un type somme pour représenter l'ensemble des entités possibles est très lourde dans ce cas. Les collections hétérogènes sont ici une solution appropriée. De telles simulations ont été programmées à l'aide du langage MGS (voir [BT02]).

On peut également noter que l'approche orientée objet se prête à l'implémentation de collections hétérogènes. Toutefois, que ce soit avec des *Vector* en JAVA ou des listes en O'CAML l'utilisateur est obligé à un moment ou à un autre de coercer explicitement les éléments des collections. Par ailleurs cette approche contraint à considérer toujours des objets alors que bien souvent des types de base suffisent. Enfin, la définition d'opérations polytypiques par le programmeur reste assez lourde dans ce cadre.

5.4. Travaux apparentés et perspectives

Les travaux de Aiken *et al.* [AW93, ALW94] fournissent une méthode d'inférence de types en présence de types unions qui sert de base à notre système de types souple, en effet les types union sont la clé de notre représentation de l'hétérogénéité. Le filtrage apparaissant dans [ALW94] est en revanche prévu pour des termes et n'est pas adapté à notre filtrage. Toutefois les types conditionnels apparaissent déjà dans ces travaux pour exprimer une analyse de flots lors du filtrage. Notre travail se démarque fortement du leur par le point de vue original sur le typage des structures de données et la puissance du processus de filtrage que nous typons.

Les travaux de A. Frisch *et al.* [FCB02] sont assez proches des nôtres puisqu'ils proposent un système de types basé sur un modèle ensembliste pour un langage avec filtrage sur le type des valeurs où l'hétérogénéité des données est représentée par des types unions. Toutefois la puissance de leurs motifs les contraint à imposer une déclaration du type des fonctions dans [FCB02].

D'autres systèmes de types existent pour des langages à base de règles. Par exemple les travaux de P. Fradet *et al.* [FLM98] munissent le langage Gamma d'un système de types dédié à la vérification de propriétés des programmes. En effet Gamma est à l'origine un langage dédié à la spécification formelle de haut niveau. Dans une optique d'implémentation efficace de $\mathcal{L}_{\subseteq}$ notre système semble plus approprié. Le système de types de Gamma permet en revanche d'envisager des vérifications de préservations de topologie (sans résoudre les difficultés discutées en section 5.3).

Nous avons utilisé nos deux systèmes de types dans un compilateur MGS pour éliminer l'étiquetage des valeurs par leur type lorsque possible. Ceci permet de grandes améliorations des performances, ce qui nous incite à intégrer d'autres optimisations basées sur les types telles que l'élimination des règles inutiles (voir section 4.6). D'autres langages à base de règles de réécriture pourraient tirer bénéfice de ces systèmes de types.

De nombreux travaux s'attellent à rendre le filtrage plus expressif. Citons TOM [MRV03] un compilateur de filtrage adapté à plusieurs langages (C, JAVA, Eiffel) et permettant le filtrage associatif ; ELAN [MK98], un langage fondé sur la réécriture et proposant une implémentation très efficace du filtrage associatif et commutatif ; enfin CDUCE [FCB02] et G'CAML [Fur02] qui permettent de filtrer les valeurs en fonction de leur type. Les langages $\mathcal{L}_{=}$ et $\mathcal{L}_{\subseteq}$ qui sont des sous-ensembles du langage MGS proposent du filtrage associatif, commutatif, sur des structures non algébriques et également sur le type des valeurs dans le cas de $\mathcal{L}_{\subseteq}$.

Enfin, notre expérience de la programmation en MGS montre que les enregistrements extensibles [Rém93] sont particulièrement utiles et adaptés à une utilisation dans un cadre hétérogène (voir de nombreux exemples de programmes dans [Gfx]). Nous souhaitons ajouter les enregistrements à nos systèmes de types et permettre les tests de types d'enregistrements lors du filtrage. Nous pensons également étudier comment augmenter les tests de types dans le filtrage sans perdre l'inférence automatique.

5.5. Remerciements

Je remercie Jacques Garrigue, Pascal Fradet, Catherine Dubois, Giuseppe Castagna et Alain Frisch pour le temps qu'ils m'ont consacré. Je remercie également Olivier Michel et Jean-Louis Giavitto pour leur aide et l'équipe SPÉCIF du LaMI pour leurs encouragements constants.

Page web du projet MGS : `http://mgs.lami.univ-evry.fr`. Une page est notamment dédiée aux travaux présentés dans cette article et aux implémentations correspondantes : `http://mgs.lami.univ-evry.fr/TypeSystem/`.

Références

[ALW94] Alexander Aiken, T.K. Lakshman, and E. Wimmers. Soft typing with conditional types. In *Proceedings of the Twenty-First Annual ACM Symposium on Principles of Programming Languages*, pages 163–173, January 1994.

[AW93] Alexander Aiken and Edward L. Wimmers. Type inclusion constraints and type inference. In *Proceedings of the Seventh ACM Conference on Functional Programming Languages and Computer Architecture*, pages 31–41, June 1993.

[BdRTG03] Pierre Barbier de Reuille, Jan Traas, and Christophe Godin. Towards a dynamic model of the arabidopsis meristem. Modélisation et simulation de processus biologiques dans le contexte de la génomique, Dieppe, France, May 2003. (Poster).

[BT02] Clément Boin and Nicolas Thibault. Hyperstructure et modélisation de chimie artificielle en MGS. Master's thesis, Université d'Évry Val d'Essonne, 2002. `http ://mgs.lami.univ-evry.fr/ImageGallery/EXEMPLES/Bugrim`.

[CF91] Robert Cartwright and Mike Fagan. Soft typing. In *Proceedings of the ACM SIGPLAN'91 Conference on Programming Language Design and Implementation (PLDI'91)*, pages 278–292. ACM Press, 1991.

[CK01] Wei-Ngan Chin and Siau-Cheng Khoo. Calculating sized types. *Higher-Order and Symbolic Computing*, 14(2/3) :261–300, 2001.

[Coh03a] Julien Cohen. Typage des collections topologiques hétérogènes et des transformations. Technical Report LaMI-89-2003, LaMI, October 2003.

[Coh03b] Julien Cohen. Typing rule-based transformations over topological collections. In Jean-Louis Giavitto and Pierre-Etienne Moreau, editors, *Electronic Notes in Theoretical Computer Science*, volume 86. Elsevier, 2003.

[Dam94a] Flemming Damm. Subtyping with union types, intersection types and recursive types. In *Symposium on Theoretical Aspects of Computer Software*, volume 789 of *Lecture Notes in Computer Science*, pages 687–706. Springer-Verlag, 1994.

[Dam94b] Flemming Damm. Type inference with set theoretic type operators. Technical Report 838, IRISA, 1994.

[FA97] Manuel Fähndrich and Alexander Aiken. Refined type inference for ML. In *Proceedings of the 1st Workshop on Types in Compilation*, 1997.

[FCB02] Alain Frisch, Giuseppe Castagna, and Véronique Benzaken. Semantic Subtyping. In *Proceedings of the Seventeenth Annual IEEE Symposium on Logic in Computer Science*, pages 137–146. IEEE Computer Society Press, 2002.

[FLM98] Pascal Fradet and Daniel Le Métayer. Structured gamma. *Science of Computer Programming*, 31(2-3) :263–289, 1998.

[Fur02] Jun P. Furuse. *Extensional Polymorphism : Theory and Application*. PhD thesis, Université Paris 7, 2002.

[Gar98] Jacques Garrigue. Programming with polymorphic variants. In *ML Workshop*, 1998.

[Gfx] The MGS graphic gallery. Web page. `http ://mgs.lami.univ-evry.fr/ImageGallery/mgs_gallery.html`.

[Gia92] Jean-Louis Giavitto. Typing geometries of homogeneous collections. In *2nd Int. workshop on array manipulation (ATABLE)*, Montréal, 1992.

[Gia03] Jean-Louis Giavitto. Topological collections, transformations and their application to the modeling and the simulation of dynamical systems. In Robert Nieuwenhuis, editor, *Rewriting Techniques and Applications, 14th International Conference*, volume 2706 of *Lecture Notes in Computer Science*, pages 208–233. Springer, 2003.

[GM01] Jean-Louis Giavitto and Olivier Michel. MGS : a programming language for the transformations of topological collections. Technical Report LaMI-61-2001, LaMI, 2001.

[GM02] Jean-Louis Giavitto and Olivier Michel. The topological structures of membrane computing. *Fundamenta Informaticae*, 49 :107–129, 2002.

[GMC02] Jean-Louis Giavitto, Olivier Michel, and Julien Cohen. Pattern-matching and rewriting rules for group indexed data structures. *ACM SIGPLAN Notices*, 37(12) :76–87, December 2002.

[GS90] Carl A. Gunter and Dana S. Scott. *Handbook of Theoretical Computer Science*, volume B, chapter Semantic domains, pages 633–674. J. van Leeuween, 1990.

[JJ96] J. Jeuring and P. Jansson. Polytypic programming. In J. Launchbury, E. Meijer, and T. Sheard, editors, *Advanced Functional Programming, Second International School*, pages 68–114. Springer-Verlag, 1996. LNCS 1129.

[LDGV02] Xavier Leroy, Damien Doligez, Jacques Garrigue, and Jérôme Vouillon. *The Objective Caml system release 3.06, Documentation and user's manual.* Projet Cristal, INRIA, 2002.

[Mar03] Luc Maranget. Les avertissements du filtrage. In *Journées Francophones des Langages Applicatifs*, pages 3–19. INRIA, 2003.

[MK98] Pierre-Etienne Moreau and Hélène Kirchner. A compiler for rewrite programs in associative-commutative theories. In *Proceedings of Algebraic and Logic Programming - Programming Language Implementation and Logic Programming, ALP/PLI LP'98*, Pisa, September 1998. Lecture Notes in Computer Science.

[MPS86] David MacQueen, Gordon Plotkin, and Ravi Sethi. An ideal model for recursive polymorphic types. *Information and Control*, 71 :95–130, 1986.

[MRV03] Pierre-Etienne Moreau, Christophe Ringeissen, and Marian Vittek. A pattern matching compiler for multiple target languages. In G. Hedin, editor, *12th Conference on Compiler Construction, Warsaw (Poland)*, volume 2622 of *LNCS*, pages 61–76. Springer-Verlag, May 2003.

[Pot98] François Pottier. *Synthèse de types en présence de sous-typage : de la théorie à la pratique.* PhD thesis, Université Paris 7, July 1998.

[PS94] Marc Pantel and Patrick Sallé. Typage souple pour le langage FOL. In *Journées francophones des langages applicatifs*, pages 21–51. INRIA, 1994.

[Rém93] Didier Rémy. Type inference for records in a natural extension of ML. In Carl A. Gunter and John C. Mitchell, editors, *Theoretical Aspects Of Object-Oriented Programming. Types, Semantics and Language Design.* MIT Press, 1993.

[Wan87] Mitchell Wand. A simple algorithm and proof for type inference. *Fundamenta Informaticae*, 10 :115–122, 1987.