Ouv (MU4IN511) : Programmation COURS 2 - suite

Ouverture (STL)

Emmanuel Chailloux

Valeurs : représentation uniforme

Nécessité de parcourir les valeurs :

- Fonctions primitives génériques : égalité, sérialisation, etc.
- Gestion mémoire (cf. suite du cours)
- Introspection, affichage générique, etc.

Solution logique : **uniformiser la structure des valeurs**

Question centrale : distinction entre

- Valeurs immédiates (entiers, caractères, etc.)
- Valeurs allouées (tableaux, structures, etc.)
- Différentes sortes de valeurs allouées.

En machine : **un pointeur = un entier = un mot machine**

Valeurs : solution plus avancées

Bit(s) discrimant(s) :

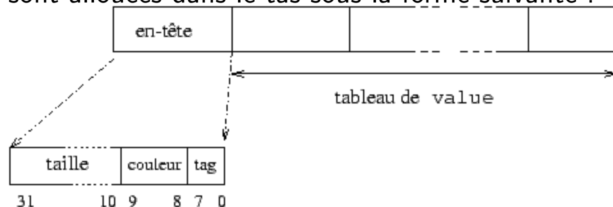
- On mange un bit sur le mot machine pour discriminer entre entier et pointeur
- Éventuellement plus de bits pour plusieurs types d'immédiats
- On utilise un système de tags comme précédemment pour les valeurs allouées
- On limite l'étendue des immédiats

En OCaml tout ce qui peut être représenté par un entier : entier, caractère, booléen, constructeur constant, est codé dans un entier et tout ce qui est plus gros qu'un entier est alloué dans le tas .

Le pointeur de la zone allouée occupe un mot mémoire, comme les entiers.

Représentation des valeurs allouées en OCaml (1)

Les constructeurs paramétrés, les enregistrements et les fermetures sont allouées dans le tas sous la forme suivante :



le tableau de `value` correspond aux champs pour un enregistrement ou un n-uplet, ou l'environnement et le code d'une fermeture. L'entête contient la taille du tableau de `value`, 2 bits pour le gestionnaire mémoire, et un `tag` correspondant au constructeur qui a alloué cette valeur.

Représentation des valeurs allouées en OCaml (2)

créer ou utiliser une fonction d'exploration : bibliothèque `Obj`

```
# Obj.reachable_words;;
- : Obj.t -> int = <fun>
# Obj.repr;;
- : 'a -> Obj.t = <fun>
```

```
# let calcule_taille f x =
   let u = f x in
    u,Obj.reachable_words (Obj.repr u) ;;
val mesure2 : ('a -> 'b) -> 'a -> 'b * int = <fun>

# let _,r=calcule_taille (i 1) 1000;;
val r : int = 3000
```

Représentation des valeurs allouées en OCaml (3)

On reprend la fonction `calcule_taille` pour explorer les valeurs :

```
1  # let calcule_taille o = Obj.reachable_words (Obj.repr o)  ;;
2  val calcule_taille : 'a -> int = <fun>
```

On commence par des valeurs allouées plates que l'on définit spécialement :

```
1  # type ty2 = Ctor2 of int * ↩
       int;;
2  type ty2 = Ctor2 of int * int
3  # let v2 = Ctor2 (3,4);;
4  val v2 : ty2 = Ctor2 (3, 4)
5  # calcule_taille v2;;
6  - : int = 3
```

```
1  # type ty5 = Ctor5 of int * int * int * ↩
        int * int ;;
2  type ty5 = Ctor5 of int * int * int * int↩
         * int
3  # let v5 = Ctor5 (3,6,9,12,15);;
4  val v5 : ty5 = Ctor5 (3, 6, 9, 12, 15)
5  # calcule_taille v5;;
6  - : int = 6
```

- v2 : a 2 valeurs : 3 et 4 dans le tableau de valeur, son entête contient cette longueur de 2 est dansl'entête ainsi que le numéro de constructeur `Ctor2` dans l'octet de `tag` , au total : 2 valeur + 1 entête = 3 mots
- v5 : c'est pareil mais avec un tableau de 5 valeurs = 6 au total

Représentation des valeurs allouées en OCaml (4)

Calcul sur une structure dynamique:

- Le constructeur `::` des listes peut être considéré comme le `Ctor2` de l'exemple précédent.
- Le constructeur constant `[]`, comme tous les constructeurs constants, est codé par un entier.

```
1  # let c = 4::[];;
2  val c : int list = [4]                    (* taille 3 *)
3  # let d = 22::c;;
4  val d : int list = [22; 4]                (* taille 6 *)
5  # let e = 77::d;;
6  val e : int list = [77; 22; 4]            (* taille 9 *)
7  # List.map calcule_taille [c;d;e];;
8  - : int list = [3; 6; 9]
```

- c a 2 valeurs 4 et `[]` + l'entête
- d prend 3 mots + la taille de c
- e prend 3 Mots plus la taille de `d`

Pile ou tas

- size1 : utilise la pile des appels de fonction
- size2 : récursive terminale mais utilise une pile à la main

```
1   # let rec size1 (bt:'a btree) : int =
2      match bt with
3     | Empty -> 0
4     | Node(_, bt1, bt2) -> 1 + (size1 bt1) + (size1 bt2) ;;
5   val size1 : 'a btree -> int = <fun>
6
7   # let rec size_aux (bt:'a btree) (bts:('a btree) list) (r:int) =
8      match bt with
9      | Node (_,bt1,bt2) -> size_aux bt1 (bt2::bts) (r+1)
10     | Empty -> ( match bts with
11                  | [] -> r
12                  | bt::bts -> size_aux bt bts r
13                ) ;;
14  val size_aux : 'a btree -> 'a btree list -> int -> int = <fun>
15  # let size2 (bt:'a btree) : int = size_aux bt [] 0 ;;
16  val size2 : 'a btree -> int = <fun>
```

Quid pour height ? ⇒ lire
“Mesurer la hauteur d’un arbre” - JC Filliâtre - JFLA 2020
https ://hal.inria.fr/hal-02427360/document

GC en Objective Caml

à générations: 2 générations (ancienne et nouvelle)

- `Stop&Copy` sur la nouvelle (GC mineur)
- `Mark&Sweep` incémental sur l'ancienne génération (GC majeur)

caractéristiques:

- Un objet jeune qui survit à 1 GC change de zone.
- Conservation des pointeurs de zone ancienne vers zone jeune
- Si le `Mark&Sweep` échoue, alors un GC compactant est déclenché pour la génération ancienne.

Le module `Gc` permet de contrôler les paramètres du GC.

Module Gc (1)

- statistiques (type *stat* :
  - `Gc.stat : unit -> Gc.stat`
- contrôler les paramètres du tas (type *control*)
  - Gc.get : unit -> Gc.control+
    `Gc.set : Gc.control -> unit`
- forcer le GC :
  - `Gc.minor() : unit -> unit`
  - `Gc.major() : unit -> unit`
  - `Gc.compact() : unit -> unit`

Module Gc (2)

Statistiques:

```
# Gc.stat();;
- : Gc.stat =
{Gc.minor_words = 112065.; promoted_words = 0.;
 major_words = 60074.; minor_collections = 0;
 major_collections = 0; heap_words = 126976;
 heap_chunks = 1; live_words = 60074;
 live_blocks = 11226; free_words = 66902;
 free_blocks = 1; largest_free = 66902;
 fragments = 0; compactions = 0;
 top_heap_words = 126976; stack_size = 53}
```

et fonction d'affichage :

`print_stat : out_channel -> unit`

Module Gc (3)

```
# Gc.stat();;
- : Gc.stat =
{Gc.minor_words = 680793.; promoted_words = 7360.;
 major_words = 117587.; minor_collections = 2;
 major_collections = 4; heap_words = 126976;
 heap_chunks = 1; live_words = 106362;
 live_blocks = 23795; free_words = 20614;
 free_blocks = 1; largest_free = 20614;
 fragments = 0; compactions = 2;
 top_heap_words = 126976; stack_size = 53}
# Gc.major();;
- : unit = ()
# Gc.stat() ;;
- : Gc.stat =
{Gc.minor_words = 702675.; promoted_words = 7360.;
 major_words = 118290.; minor_collections = 2;
 major_collections = 5; heap_words = 126976;
 heap_chunks = 1; live_words = 106671;
 live_blocks = 23880; free_words = 20305;
 free_blocks = 18; largest_free = 19911;
 fragments = 0; compactions = 2;
 top_heap_words = 126976; stack_size = 53}
```

Module Gc (4)

Contrôle:

```
# let c = Gc.get () ;;
val c : Gc.control =
  {Gc.minor_heap_size = 262144;
   major_heap_increment = 126976;
   space_overhead = 80;
   verbose = 0;
   max_overhead = 500;
   stack_limit = 1048576;
   allocation_policy = 0}
```

Module Gc (5)

Par exemple, le champ `verbose` peut prendre des valeurs de 0 à 127 activant 7 indicateurs différents.

```
# c.Gc.verbose <- 127 ;;
- : unit = ()
# Gc.set c ;;
- : unit = ()
# Gc.compact () ;;
```

affichage:

```
Heap compaction requested
<>Sweeping 9223372036854775807 words
Starting new major GC cycle
Marking 9223372036854775807 words
Subphase = 10
Sweeping 9223372036854775807 words
Compacting heap...
done.
```

Les différentes phases du GC sont indiquées ainsi que le nombre d'objets traités.

Mesure de taille de valeurs OCaml (1)

- calcul sur les valeurs vivantes entre 2 Gc encadrant un calcul (si calcul pur)
- exploration d'une valeur en tenant compte du partage (`calcul_taille`

fonction de test : intervalle

```
let rec i a b =  if  a > b then [] else a :: (i (a+1) b);;
```

calcul sur les objets vivants :

```
let mesure f x =
  Gc.compact ();
  let s1 = Gc.stat() in
    let u = f x in
    Gc.compact();
    let s2= Gc.stat() in
    (u,s1,s2) ;;

let _,b,c = mesure (i 1) 1000;;
```

Mesure de taille de valeurs OCaml (2)

résultats :

```
1  #  let _,b,c = mesure (i 1) 1000;;
2  val b : Gc.stat =
3    {Gc.minor_words = 397708.; promoted_words = 142768.; major_words = 302854.;
4     minor_collections = 2; major_collections = 2; heap_words = 311296;
5     heap_chunks = 4; live_words = 275375; live_blocks = 64498;
6     free_words = 35918; free_blocks = 1; largest_free = 35918; fragments = 3;
7     compactions = 1; top_heap_words = 311296; stack_size = 97;
8     forced_major_collections = 1}
9  val c : Gc.stat =
10   {Gc.minor_words = 400732.; promoted_words = 145792.; major_words = 305878.;
11    minor_collections = 3; major_collections = 4; heap_words = 311296;
12    heap_chunks = 4; live_words = 278399; live_blocks = 65502;
13    free_words = 32894; free_blocks = 1; largest_free = 32894; fragments = 3;
14    compactions = 2; top_heap_words = 311296; stack_size = 99;
15    forced_major_collections = 2}
16 # c.promoted_words -. b.promoted_words;;
17 - : float = 3024.
18 # c.live_words - b.live_words;;
19 - : int = 3024
```