Concours commun Centrale - Supelec 2014

Corrigé de l'épreuve d'informatique

MP - option informatique

Partie I: préliminaires

Il n'y a aucun problème pour écrire les trois premières fonctions:

En notant (i, j) l'indice de la case de bloc (b, r) et $b = 3b_1 + b_2$ et $r = 3r_1 + r_2$ les divisions euclidiennes de b et de r par 3, nous avons $i = 3b_1 + r_1$ et $j = 3b_2 + r_2$, ce qui donne:

```
let indice (b,r) = 3*(b/3)+(r/3)3, 3*(b mod 3)+(r mod 3);
```

Partie II: codage de la formule initiale

II.A -Formule logique décrivant la règle du jeu

A.1) La clause s'écrit $\bigvee_{k=1}^{9} x_{i,j}^{k}$. La règle (K1) s'écrit mathématiquement: $\forall (i,j) \in [0,8]^{2}$, $\exists k \in [1,9]$, $x_{i,j}^{k}$, ce qui correspond à la forme normale conjonctive $K1 = \bigwedge_{i=0}^{8} \bigwedge_{j=0}^{8} \left(\bigvee_{k=1}^{9} x_{i,j}^{k}\right)$ qui contient 81 clauses.

On construit facilement K1 à l'aide d'une triple boucle: chaque boucle sur k construit une clause qui est ajoutée à la formule F (on utilise des référence sur des listes pour faire évoluer dynamiquement la clause cet la formule f):

```
let case1 () =
 let f = ref [] in
 for i = 0 to 8 do for j = 0 to 8 do
   let c = ref [] in
   for k = 1 to 9 do
     c := X(i,j,k) :: !c
   done;
    f := !c :: !f
 done; done;
 !f;;
```

A.2) (L1) s'écrit:
$$\forall (i,k) \in [0,8] \times [1,9], \ \exists j \in [0,8], \ x_{i,j}^k$$
, ce qui donne $L1 = \bigwedge_{i=0}^8 \bigwedge_{k=1}^9 \left(\bigvee_{j=0}^8 x_{i,j}^k\right)$.

$$\textbf{A.3)} \qquad \text{On obtient de la même façon } C1 = \bigwedge_{j=0}^8 \bigwedge_{k=1}^9 \left(\bigvee_{i=0}^8 x_{i,j}^k\right) \text{ et } B1 = \bigwedge_{b=0}^8 \bigwedge_{k=1}^9 \left(\bigvee_{r=0}^8 x_{\mathtt{indice}(b,r)}^k\right).$$

On obtient ainsi la fonction bloc1, presque identique à la fonction ligne1:

```
let bloc1 () =
 let f = ref [] in
 for b = 0 to 8 do for k = 1 to 9 do
   let c = ref [] in
   for r = 0 to 8 do
     let i, j = indice(b,r) in
     c := X(i,j,k) :: !c
   done;
    f := !c :: !f
 done; done;
 !f;;
```

La ligne i ne contient pas deux fois la valeur k si et seulement: A.4)

$$\forall j_1, j_2 \in [0, 8], \ j_1 < j_2 \Longrightarrow (\neg x_{i, j_1}^k \lor \neg x_{i, j_2}^k)$$

ce qui se traduit par la forme normale conjonctive $\bigwedge_{j_1=0}^7 \bigwedge_{j_2=j_1+1}^8 \left(\neg\, x_{i,j_1}^k \lor \neg\, x_{i,j_2}^k\right).$

(L2) s'écrit donc $\forall i \in Ii, \ \forall k \in \llbracket 1, 9 \rrbracket, \ \forall j_1, j_2 \in \llbracket 0, 8 \rrbracket, \ j_1 \neq j_2 \Longrightarrow \left(\neg \, x_{i,j_1}^k \lor \neg \, x_{i,j_2}^k \right)$ et la formule:

$$L2 = \bigwedge_{i=0}^{8} \bigwedge_{k=1}^{9} \bigwedge_{j_{1}=0}^{8} \bigwedge_{j_{2}=0 \atop j_{2} \neq j_{1}}^{8} \left(\neg x_{i,j_{1}}^{k} \lor \neg x_{i,j_{2}}^{k} \right).$$

Cette formule contient $9 \times 9 \times \frac{9 \times 8}{2}$, soit 2916 clauses et est calculée par la fonction:

```
let lig2 () =
  let f = ref [] in
  for i = 0 to 8 do for k= 1 to 9 do for j1= 0 to 7 do for j2= j1 to 8 do
    f := [NonX(i,j1,k); NonX(i,j2,k)] :: !f
  done; done; done;
!f;;
```

A.5) De la même manière, on obtient:

$$\begin{cases} C2 = \bigwedge_{j=0}^{8} \bigwedge_{k=1}^{9} \bigwedge_{i_{1}=0}^{7} \bigwedge_{i_{2}=i_{1}+1}^{8} \left(\neg x_{i_{1},j}^{k} \lor \neg x_{i_{2},j}^{k}\right) \\ K2 = \bigwedge_{i=0}^{8} \bigwedge_{j=0}^{8} \bigwedge_{k_{1}=1}^{8} \bigwedge_{k_{2}=k_{1}+1}^{9} \left(\neg x_{i,j}^{k_{1}} \lor \neg x_{i,j}^{k_{2}}\right) \\ B2 = \bigwedge_{b=0}^{8} \bigwedge_{k=1}^{9} \bigwedge_{r_{1}=0}^{7} \bigwedge_{r_{2}=r_{1}+1}^{8} \left(\neg x_{\mathrm{indice}(b,r_{1})}^{k} \lor \neg x_{\mathrm{indice}(b,r_{1})}^{k}\right) \end{cases}$$

Les formules K1, L1, C1, B1 (resp. K2, L2, C2, B2) contiennent chacunes 81 (resp. 2916) clauses, donc F_{regle} en contient bien $4 \times (81 + 2916) = 11988$.

II.B -Formule logique décrivant la grille initiale

B.1) On initialise une pile vide de clause puis on parcourt le tableau. Si l'on trouve une valeur $k \neq 0$ dans la case (i, j), on ajoute à la pile les clauses $x_{i,j}^k$ et $\neg x_{i,j}^q$ pour tout $q \neq k$ (attention: une clause unitaire est une liste de longueur 1):

B.2) La fonction indice est égale à son inverse, donc b = 3(i/3) + j/3 où / désigne la division entière (comme dans Caml).

Pour éviter d'écrire plusieurs fois la même clause unitaire négative, nous utilisons la fonction ajoute de la partie I.

```
let interdites_ij t i j = let f = ref [] in
  let b = 3*(i/3)+(j/3) in
  for l = 0 to 8 do
    if t.(1).(j) <> 0 then
        f := ajoute NonX(i,j,t.(1).(j)) !f;
    if t.(i).(1) <> 0 then
        f := ajoute NonX(i,j,t.(i).(1)) !f;
    let i1,j1 = indice(b,l) in
    if t.(i1).(j1) <> 0 then
        f := ajoute NonX(i,j,t.(i1).(j1)) !f;
    done;
  !f ;;
```

En remarquant que l'utilisation de la concaténation de liste n'est pas trop couteuse, puisque les listes interdites_ij T i j est de longueur au plus 8, cela donne la fonction:

```
let interdites t =
  let f = ref [] in
  for i = 0 to 8 do for j = 0 to 8 do
    if t.(i).(j)=0 then
      f := (interdites_ij t i j) !f;
  done; done;
!f;;
```

B.3) Chaque case remplie (i, j) donne 9 clauses et chaque case non remplie donne au maximum 8 clauses (comme le problème a une solution, il y a au moins une valeur non interdite pour chaque case). La formule F contient donc au maximum 9N + 8(81 - N) clauses, où N est le nombre de cases pleines de la grille initiale. Dans le pire des cas (toutes les cases sont remplies), F contient $9^3 = 729$ clauses.

Partie III: résolution d'une grille de sudoku

III.A - Règle de propagation unitaire

- **A.1)** Comme la grille initiale est associée à une et une seule grille finale, il existe une unique valuation satisfaisant F_{initiale} .
- **A.2)** Il y a $9^3 = 729$ variables, donc la table de vérité de F contient 2^{729} lignes (soit plus de 10^{219}); il est bien illusoire de tester toutes les lignes à la recherche de l'unique valuation solution du problème.
- **A.3)** Si σ satisfait F', prolongeons σ en une valuation $\overline{\sigma}$ définie sur \mathcal{V} en posant $\overline{\sigma}(p) = \mathtt{true}$ si $\ell = p$ et $\overline{\sigma}(p) = \mathtt{false}$ si $\ell = \neg p$. Nous avons alors

$$\operatorname{Ev}_{\overline{\sigma}}(F) = \operatorname{Ev}_{\overline{\sigma}}(\ell) \wedge \operatorname{Ev}_{\overline{\sigma}}(F_1) \wedge \operatorname{Ev}_{\overline{\sigma}}(F_2) \wedge \operatorname{Ev}_{\overline{\sigma}}(F_3).$$

- si C est une clause contenant ℓ , elle est satisfaite par $\overline{\sigma}$: on en déduit que $\operatorname{Ev}_{\overline{\sigma}}(\ell) = \operatorname{Ev}_{\overline{\sigma}}(F_1) = \operatorname{true}$;
- si C est une clause contenant $\neg \ell$ et ne contenant pas ℓ , $\overline{\sigma}(C) = \overline{\sigma}(C')$ où C' est la clause obtenue en supprimant de C les occurences du littéral $\neg \ell$: on en en déduit que $\operatorname{Ev}_{\overline{\sigma}}(F_2) = \operatorname{Ev}_{\sigma}(F_2')$;
- si C est une clause ne contenant ni ℓ , ni $\neg \ell$, la variable p n'intervient pas dans C et $\operatorname{Ev}_{\overline{\sigma}}(C) = \operatorname{Ev}_{\sigma}(C)$: on en déduit que $\operatorname{Ev}_{\overline{\sigma}}(F_3) = \operatorname{Ev}_{\sigma}(F_3)$.

Nous obtenons donc:

$$\operatorname{Ev}_{\overline{\sigma}}(F) = \operatorname{Ev}_{\sigma}(F_2') \wedge \operatorname{Ev}_{\sigma}(F_3) = \operatorname{Ev}_{\sigma}(F') = \operatorname{true}.$$

Supposons réciproquement qu'il existe une valuation $\overline{\sigma}$ qui prolonge σ et qui satisfait F. Comme ℓ est une clause de F, σ satisfait ℓ et on obtient comme ci-dessus $\operatorname{Ev}_{\sigma}(F') = \operatorname{Ev}_{\overline{\sigma}}(F) = \operatorname{true}$.

- **A.4** Une première étape permet de remplacer F par $\left(x_{(2,2)}^4 \lor x_{(3,6)}^6 \lor x_{(7,7)}^7\right) \land \left(\neg x_{(3,6)}^6\right)$, qui devient ensuite $\left(x_{(2,2)}^4 \lor x_{(7,7)}^7\right)$. Ces transformations doivent être accompagnées de l'affectation $T.(0).(0) \leftarrow 1$.
- A.5 Seules les valeurs 1, 3, 4 et 7 ne sont pas déjà présentes sur la ligne, la colonne ou le bloc de la case (0,7). Les cases vides de la ligne 0 sont les cases (0,0), (0,2), (0,4), (0,5), (0,7). Comme le chiffre 7 apparaît déjà dans les blocs 0 et 1, (dans les cases (2,1) et (1,5)), il ne peut être placé ni en (0,0), (0,2), (0,4) ou (0,5): on en déduit que le 7 de la première ligne est nécessairement dans la case (0,7).

En terme de formule, nous pouvons écrire:

$$F_{\mathtt{initiale}} = F' \wedge F'_1 \wedge F'_2 \wedge F_3$$

$$\hat{\mathbf{e}} \left\{ \begin{array}{l} F_1' = \neg x_{(0,1)}^7 \wedge \neg x_{(0,3)}^7 \wedge \neg x_{(0,6)}^7 \wedge \neg x_{(0,8)}^7 \\ F_2' = \neg x_{(0,0)}^7 \wedge \neg x_{(0,2)}^7 \wedge \neg x_{(0,4)}^7 \wedge \neg x_{(0,5)}^7 \wedge \neg x_{(0,5)}^7 \\ F_3' = x_{(0,0)}^7 \vee x_{(0,1)}^7 \vee x_{(0,2)}^7 \vee x_{(0,3)}^7 \vee x_{(0,4)}^7 \vee x_{(0,5)}^7 \vee x_{(0,6)}^7 \vee x_{(0,7)}^7 \vee x_{(0,8)}^7 \vee x_{(0,8)}^7 \end{array} \right.$$

En effet, F_1' est une partie de F_1 , F_2' une partie de F_2 et F_3' une partie de L1. En appliquant l'algorithme de propagation unitaire successivement avec les littéraux isolés $\neg x_{(0,0)}^7, \neg x_{(0,1)}^7, \neg x_{(0,2)}^7, \neg x_{(0,3)}^7, \neg x_{(0,4)}^7, \neg x_{(0,5)}^7, \neg x_{(0,6)}^7, \neg x_{(0,6)}^7, \neg x_{(0,8)}^7, \text{ nous obtenons } F_{\text{initiale}} \equiv F'' \wedge x_{(0,7)}^7$ qui donne enfin $x_{(0,7)}^7 = \text{true}$ en choisissant le littéral isolé $x_{(0,7)}^7$.

A.6) Il suffit de parcourir récursivement la formule à la recherche d'une clause de longueur 1 (sans faire appel à la fonction list_length!), ce qui donne:

```
let rec nouveau_lit_isole f = match f with
| [] -> X(-1,-1,-1)
| [lit] :: f1 -> lit
| _ :: f1 -> nouveau_lit_isole f1;;
```

A.7) On peut traduire directement la transformation proposée:

```
let rec simplification 1 f = match f with | [] -> []
| C :: f1 when appartient 1 C -> simplification 1 f1
| C :: f1 -> supprime (neg 1) C :: simplification 1 f1;;
```

A.8) Il suffit d'appliquer la méthode proposée:

```
let rec propagation t f = match nouveau_lit_isole f with | X(-1,-1,-1) -> f  | X(i,j,k) -> t.(i).(j) <- k; propagation t (simplification (X(i,j,k)) f) | 1 ->  propagation t (simplification 1 f) ;;
```

- A.9) À la fin du calcul, si la méthode fonctionne, le tableau T contient la solution au problème posé et la formule renvoyée F' est équivalente à $F_{initiale}$ et ne contient aucun littéral isolé. Supposons que $F \neq []$: il existe alors une variable $x_{i,j}^k$ qui est encore présente dans F', ce qui impose que les littéraux isolés X(i,j,k) et NonX(i,j,k) ne sont jamais apparus dans la formule (sinon, une simplification aurait supprimé tous les $x_{(i,j)}^k$). Quatre cas sont alors possibles:
 - 1. la case (i,j) était remplie au début du calcul et T.(i).(j) = k; comme le littéral isolé $x_{(i,j)}^k$ était présent dans la formule initiale, il n'a pu être supprimé que lors d'une simplification faite avec le littéral isolé X(i,j,k) et $x_{(i,j)}^k$ n'apparaît plus dans F': c'est absurde;
 - 2. la case (i,j) était remplie au début du calcul et $T.(i).(j) = k' \neq k$; comme la clause $C = \neg x_{(i,j)}^k \lor \neg x_{(i,j)}^{k'}$ était dans la formule initiale et que $x_{(i,j)}^{k'}$ n'apparaît pas dans F' (cas 1.), F' contient le littéral isolé $\neg x_{(i,j)}^k$: c'est absurde;
 - 3. la case (i,j) n'était pas remplie au début du calcul et T.(i).(j) = k; il y a donc un moment où, dans la propagation, on a effectué l'affectation $T.(i).(j) \leftarrow k$, qui a été suivie de la simplification par $x_{(i,j)}^k$: c'est absurde, puisque cette simplification supprime tous les $x_{(i,j)}^k$;
 - 4. la case (i,j) n'était pas remplie au début du calcul et $T.(i).(j) = k' \neq k$; comme précédemment, F' ne contient pas la variable $x_{(i,j)}^{k'}$ et la clause $C = \neg x_{(i,j)}^k \lor \neg x_{(i,j)}^{k'}$ présente dans F est devenue $\neg x_{(i,j)}^k$ au moment de la simplification par le littéral isolé X(i,j,k'): c'est une nouvelle fois absurde.

Si la méthode ne fonctionne pas, le tableau T n'est pas entièrement rempli et la formule renvoyée n'est pas vide et ne contient aucun littéral isolé.

A.10) La fonction nouveau_lit_isole parcourt, dans le pire des cas, toute la liste F en effectuant une opération de coût constant sur chaque clause: son coût est donc dans le pire des cas de l'ordre de la longueur de F, ce qui donne (grossièrement) un O(n).

La fonction simplification demande, dans le pire des cas, le parcours de toutes les clauses au plus deux fois (une seule fois si le littéral est présent dans la clause et une seconde fois s'il n'est pas présent, lors du calcul supprime (neg 1) C1): le temps de calcul est une nouvelle fois un O(n).

Enfin, dans le pire des cas (pour le temps de calcul, puisque ce pire des cas est celui où l'algorithme permet de résoudre le sudoku), chaque variable est supprimée de la formule initiale et la simplification est effectuée autant de fois qu'il y a de variables. Comme on demande un résultat ne dépendant que de n, on fait l'hypothèse que toutes les variables apparaissent dans la formule (ce qui est le cas avec F_{initiale}): le nombre de variable est donc inférieur à n et le coût total de propagation est un $O(n^2)$.

On pourrait faire un calcul un peu plus précis en exprimant le temps de calcul (toujours grossièrement) comme un O(g(N)) où N est la taille du sudoku (ici, N=9 et n est de l'ordre de N^4).

III.B - Règle du littéral infructueux

B.1) Si l'on peut déduire la clause vide de la formule $F \wedge \neg x$, $F \wedge \neg x$ est une antilogie, ce qui donne:

$$F \equiv F \land (x \lor \neg x) \equiv (F \land x) \lor (F \land \neg x) \equiv (F \land x) \lor \bot \equiv F \land x$$

B.2) Je ne vois pas l'intérêt d'utiliser la fonction flatten qui va nécessiter de concaténer les clauses. Nous obtenons la fonction variables qui utilise les fonctions auxiliaires ajouter_clause et ajouter_formule qui ajoutent (sans doublons) les variables rencontrées à la pile Pile:

B.3 Nous allons appliquer la fonction propagation à des formules qui ne seront plus satisfiables (on ajoute un littéral quelconque que l'on espère être infructueux): en appliquant la fonction simplification, une instruction supprime (neg 1) C peut donc conduire à la liste vide \bot : il est alors maladroit de laisser cette liste vide dans la formule finale, puisqu'une formule contenant la clause \bot n'est pas satisfiable (c'est une antilogie). Autrement-dit, dès que l'on obtient une clause vide, il serait plus intéressant de faire renvoyer par la fonction simplification la formule L = [[]] qui représente une antilogie (et non pas la formule F = [] qui représente une tautologie), plutôt qu'une longue formule contenant une clause vide (cela évitera de chercher à chaque étape si la formule contient une clause vide, alors que la présence d'une telle clause peut être détectée à l'étape précédente). Nous allons donc modifier la fonction simplification de sorte qu'elle renvoie [[]] dès qu'une clause vide a été trouvée:

La fonction deduction s'écrit alors facilement: on applique la fonction propagation à la formule $[\neg x] :: F$; si l'on obtient l'antilogie [[]], on renvoie 1; sinon, on applique la fonction propagation à la formule [x] :: F; si l'on obtient l'antilogie [[]], on renvoie -1; sinon, on renvoie 0. On utilise des copies de la matrice T car la fonction propagation modifie la matrice à laquelle elle est appliquée.

```
let deduction t x f = let t1 = copier_matrice t in match x with
    | X(i,j,k) when propagation t1 ([NonX(i,j,k)] :: f) = [[]] -> 1
    | _ -> let t2 = copier_matrice t in
    if propagation t2 ([x] :: f) = [[]] then -1 else 0;;
```

B.4 Nous allons utiliser l'analyse suivante: un booléen fini initialisé à la valeur false permet de savoir si le calcul est terminé. Il prend la valeur true si après une propagation, la formule obtenue est [] (le sudoku est alors résolu), ou bien si on ne trouve pas de littéral infructueux. Cela donne:

```
let deduction2 t f = let fini = ref false and f1 = ref f in
while not !fini do
  f1 := propagation t !f1;
  if !f1=[] then
    fini:= true
  else
    begin
    match litteral_infructueux t (variables !f1) !f1 with
    | X(-1,-1,-1) -> fini := true
    | x -> f1 := [x] :: !f1
    end
done;;
```

B.4 Il suffit de construire la formule $F_{\tt initiale}$ en concaténant les résultats des huit fonctions construites dans la partie II (pour accélérer le temps de calcul de nouveau_lit_isole, nous mettons en tête de liste les littéraux isolés, puis les clauses de longueur 2). Cela donne:

Il est enfin possible de tester la fonction sur l'exemple donné en début d'énoncé, qui donne la solution du problème en 5 secondes:

qui donne la table complétée:

		_	_		_		_	
8	9	1	2	3	4	6	7	5
3	2	5	6	8	7	4	1	9
4	7	6	9	1	5	3	2	8
9	1	4	7	5	2	8	6	3
2	5	7	3	6	8	1	9	4
6	8	3	1	4	9	7	5	2
1	4	8	5	2	6	9	3	7
7	3	2	4	9	1	5	8	6
5	6	9	8	7	3	2	4	1

On pourra tester cette fonction sur d'autres sudokus, en particulier sur les deux exemples ci-dessous, qualifiés respectivement de "très difficile" et de "diabolique" dans le recueil où je les ai trouvés: