Devoir Non Surveillé 2

À rendre le 6 janvier 2025

INFORMATIQUE MPI/MPI*

Tournez la page S.V.P.

Vue d'ensemble du sujet

Dans ce sujet, on se propose d'étudier le problème du voyageur de commerce (TSP pour Travelling Salesman Problem)

Chemins hamiltoniens, et problème du voyageur de commerce (TSP)

Soit $G = (S, A)$ un graphe non orienté. Un **cycle hamiltonien** de G est un cycle passant une fois et une seule par chaque sommet.
On définit le problème d'optimisation du voyageur de commerce (TSP) :

Instance : $G = (S, A, w)$ un graphe pondéré non orienté, avec $|S| = n$.

Solution : Un cycle hamiltonien $\mathcal{C} = (s_0, s_1, \ldots, s_{n-1}, s_n = s_0)$ minimisant la fonction de cout :

$$w(\mathcal{C}) = \sum_{i=0}^{n-1} w(s_i, s_{i+1})$$

Ce sujet est composé de 5 parties indépendantes, utilisant les langages de programmation C et OCaml.

Les différentes parties sont indépendantes et peuvent être traités dans n'importe quel ordre. Il n'est pas nécessaire d'avoir répondu aux questions d'une partie pour répondre aux questions d'une autre partie.

- La partie I étudie une résolution du problème par recherche exhaustive naïve.
- La partie II s'intéresse à l'algorithme de programmation dynamique proposé par Held et Karp (et aussi indépendamment par Bellman).
- La partie III s'intéresse à une résolution gloutonne, et en montre les limites.
- La partie IV traite la difficulté du problème de seuil associé.
- La partie V montre la difficulté d'approximer une solution à un facteur constant près.
- Enfin, la partie VI montre que dans certains cas, il est possible de fournir une 2-approximation de la solution.

Pour répondre à une question, il est permis de réutiliser le résultat d'une question antérieure, même sans avoir réussi à établir ce résultat. En langage C, il est inutile de rappeler que les entêtes `<assert.h>`, `<stdbool.h>`, etc. doivent être inclus.

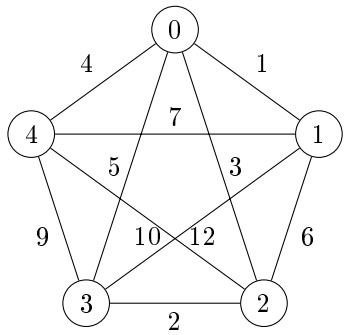
Quand l'énoncé demande de coder une fonction, sauf indication explicite de l'énoncé, il n'est pas nécessaire de justifier que celle-ci est correcte ou de tester que des préconditions sont satisfaites.

Le barème tient compte de la clarté des programmes : nous recommandons de choisir des noms de variables intelligibles ou encore de structurer de longs codes par des blocs ou par des fonctions auxiliaires dont on décrit le rôle. Lorsqu'une réponse en pseudo-code est permise, seule la logique de programmation est évaluée, même dans le cas où un code en C a été fourni en guise de réponse.

Partie I. Approche naïve (C) – 12 pts

☐ 1 – On suppose que G est un graphe complet d'ordre n. Déterminer le nombre de cycles hamiltoniens différents dans G. On supposera que deux cycles hamiltoniens sont différents s'ils ne sont pas constitués des mêmes arêtes.

Pour la suite de cette partie, on suppose qu'un graphe $G = (S, A, w)$ pondéré, non orienté et complet est représenté par sa matrice d'adjacence implémentée en C dans un tableau unidimensionnel d'entiers, les lignes de la matrice étant consécutives dans le tableau. Le graphe étant supposé complet, la valeur ∞ n'apparaitra pas dans la matrice. Par exemple, le graphe G_0 de la figure pourra être représenté par le code suivant :



$$M_0 = \begin{pmatrix} 0 & 1 & 3 & 5 & 4 \\ 1 & 0 & 6 & 12 & 7 \\ 3 & 6 & 0 & 2 & 10 \\ 5 & 12 & 2 & 0 & 9 \\ 4 & 7 & 10 & 9 & 0 \end{pmatrix}$$

C
```c
int G0[25] =
{0, 1, 3, 5, 4,
1, 0, 6, 12, 7,
3, 6, 0, 2, 10,
5, 12, 2, 0, 9,
4, 7, 10, 9, 0};
```

FIGURE 1 – Le graphe G_0, sa matrice d'adjacence et sa représentation en C

☐ 2 – Écrire une fonction `int w(int *G, int n, int s, int t)` qui prend en argument un tableau correspondant à un graphe $G = (S, A, w)$, un entier $n = |S|$, et deux entiers $(s, t) \in S^2$ et renvoie la valeur $w(s, t)$ correspondant au poids de l'arête (s, t).

On choisit de représenter un cycle hamiltonien $\mathcal{C} = (s_0, s_1, \ldots, s_{n-1}, s_n = s_0)$ d'un graphe G d'ordre n par un tableau de taille n contenant les sommets du cycle.

☐ 3 – Écrire une fonction `int poids_cycle(int *G, int* c, int n)` qui prend en argument un graphe $G = (S, A, w)$, un cycle $\mathcal{C}$ de G et un entier $n = |S|$ et renvoie $w(\mathcal{C})$ tel que défini précédemment. Par exemple, si `int *c = {0, 2, 4, 3, 1};` , alors `poids_cycle(G0, c, 5)` renverrai `35`.

On remarque que toute permutation de $[\![0, n-1]\!]$ forme un cycle hamiltonien de G. On propose de parcourir toutes les permutations par ordre lexicographique pour conserver celle de poids minimal. Par exemple, la permutation qui suit $(0, 2, 4, 3, 1)$ dans l'ordre lexicographique est $(0, 3, 1, 2, 4)$.
Si $p = (p_0, p_1, \ldots, p_{n-1})$ est une permutation de $[\![0, n-1]\!]$, on définit les indices suivants :
- $j = \max\{i \in [\![0, n-2]\!] \mid p_i < p_{i+1}\}$ et $j = -1$ si cet ensemble est vide.
- $k = \max\{i \in [\![j+1, n-1]\!] \mid p_j < p_i\}$ et $k = n$ si $j = -1$.

☐ 4 – En utilisant les indices j et k, décrire en français ou en pseudo-code un algorithme permettant de modifier p pour obtenir la permutation suivante dans l'ordre lexicographique et renvoyer «faux» si p était la dernière permutation et «vrai» sinon. On supposera pour la suite qu'une telle fonction est implémentée par une fonction `bool permut_suivante(int *p, int n)`.

☐ 5 – Écrire une fonction `int *PVC_naif(int *G, int n)` qui prend en argument un graphe $G = (S, A, w)$ et un entier $n = |S|$ et renvoie un pointeur vers un tableau contenant un cycle hamiltonien de G de poids minimal.

☐ 6 – Déterminer la complexité temporelle de `PVC_naif` en fonction de $n = |S|$. On admettra que `permut_suivante` a une complexité amortie en $\mathcal{O}(1)$.

Partie II. Algorithme de Held-Karp (OCaml) – 8 pts

L'approche naïve précédente effectue beaucoup de calculs superflus. En effet, on peut remarquer que si $(0, s_1, s_2, \ldots, s_k)$, $k < n-1$, est le début d'un cycle hamiltonien de poids minimal, alors la manière de le compléter ne dépend pas de l'ordre des sommets $s_1, s_2, \ldots, s_{k-1}$ mais uniquement de s_k et de $S \setminus \{0, s_1, \ldots, s_k\}$.
L'algorithme de Held-Karp est un algorithme de programmation dynamique, qui utilise une table de hachage pour mémoïser les résultats.
On choisit de représenter un ensemble de sommets qui a déjà été ajouté au cycle par un tableau de booléens `vus`, valant `true` si le sommet appartient à l'ensemble, et `false` sinon.
Pour pouvoir stocker les résultats déjà calculés dans la table de hachage, il est nécéssaire de transformer les valeurs en des objets non mutables, comme par exemple des entiers ou des tuples d'objets non mutables.

☐ 7 – Écrire une fonction `clef : int -> bool array -> int * int` qui prend en argument un entier s_k et un tableau de booléens représentant un sous-ensemble $S' \subset S$, $(S = [\![0, n-1]\!])$ et renvoie un couple d'entiers représentant de manière unique s_k et S' et pouvant servir de clef dans une table de hachage.
On pourra assimiler le tableau de booléens à la décomposition binaire d'un entier compris entre 0 et 2^{n-1}.

On choisit en OCaml de représenter un graphe par matrice d'adjacence. Ainsi, le graphe G_0 sera représenté par :

```
let g0 = [| [|0; 1; 3; 5; 4|];
            [|1; 0; 6; 12; 7|];
            [|3; 6; 0; 2; 10|];
            [|5; 12; 2; 0; 9|];
            [|4; 7; 10; 9; 0|] |];;
```

On rappelle que les tables de hachages peuvent être manipulées en OCaml avec les fonctions suivantes :
- `Hashtbl.create : int -> ('a, 'b) Hashtbl.t`
  prend en argument un entier m et crée une table vide occupant un espace mémoire de taille proportionnelle à m (on pourra choisir m = 1 en pratique) ;
- `Hashtbl.add : ('a, 'b) Hashtbl.t -> 'a -> 'b -> unit`
  prend en argument une table, une clé et une valeur et rajoute une association ;
- `Hashtbl.mem : ('a, 'b) Hashtbl.t -> 'a -> bool`
  teste si une table contient une clé donnée ;
- `Hashtbl.find : ('a, 'b) Hashtbl.t -> 'a -> 'b`
  prend en argument une table et une clé et renvoie la valeur associée.

On suppose créées en variables globales :
- une matrice d'entiers correspondants à un graphe $G = (S, A, w)$;
- une table de hachage par la commande `let tabh = Hashtbl.create 1;;`.

☐ 8 – Écrire une fonction

`completer_cycle : int array -> bool array -> int -> unit`

qui prend en argument un tableau d'entiers `c`, un tableau de booléens `vus` et un entier k tels que :
- `c` et `vus` sont deux tableaux de même taille $n = |S| > k$;
- les éléments `c.(i)` tels que $i \leq k$ sont tous distincts et compris entre 0 et $n-1$, et `c.(0) = 0` ;
- les éléments s tels que `vus.(s)` vaut `true` sont exactement `c.(0)`,..., `c.(k)`

La fonction doit modifier `c` en un cycle hamiltonien de G sans modifier les éléments d'indices inférieurs ou égaux à k, en minimisant le poids du cycle parmi ceux qui commencent par `c.(0)`,..., `c.(k)`.

☐ 9 – En déduire une fonction `PVC_dynamique : unit -> int array` qui renvoie un tableau correspondant à un cycle hamiltonien de poids minimal du graphe G, calculé selon l'algorithme de Held-Karp.

☐ 10 – Déterminer la complexité de `PVC_dynamique` en fonction de $n = |S|$.

Partie III. Heuristique du plus proche voisin (C) – 8 pts

Pour éviter la recherche exhaustive, on propose un algorithme glouton qui ajoute les sommets un par un en choisissant toujours celui qui coûte le moins cher, c'est-à-dire le plus proche voisin.
Pour l'implémenter, on utilise un tableau `c` correspondant au cycle en cours de construction et un tableau `vus` de booléens, permettant de savoir quels sommets ont déjà été vus et rajoutés au cycle.

☐ 11 – Écrire une fonction

```
int plus_proche(int *G, bool *vus, int n, int s)
```

qui prend en argument un graphe $G = (S, A, w)$, un tableau `vus` de booléens, un entier $n = |S|$ et un sommet $s \in S$ et renvoie un sommet t vérifiant :

- `vus[t]` vaut `false` ;
- $w(s,t)$ est minimal parmi les sommets t non vus. Par exemple, si `vus` est défini par

```
bool vus[5] = {true, false, true, true, false};
```

alors l'appel à `plus_proche(G0, vus, 5, 2)` renverra `1` car les seuls sommets non vus sont 1 et 4, et $w(2,1) = 6 < w(2,4) = 10$.

☐ 12 – En déduire une fonction

```
int *PVC_glouton(int *G, int n)
```

qui prend en argument un graphe $G = (S, A, w)$ et un entier $n = |S|$ et renvoie un pointeur vers un tableau contenant un cycle hamiltonien de G construit selon l'heuristique du plus proche voisin. On commencera par le sommet 0 comme sommet initial.

☐ 13 – Déterminer la complexité temporelle de `PVC_glouton` en fonction de $n = |S|$.

☐ 14 – Déterminer et représenter graphiquement un graphe pondéré complet d'ordre 5 tel que l'heuristique du plus proche voisin ne renvoie jamais de cycle hamiltonien de poids minimal en partant du sommet 0, quelles que soient les manières de traiter les cas d'égalité.

Partie IV. Difficulté du problème de seuil – 12 pts

Dans cette partie, on se propose de montrer que le problème de seuil associé au problème du voyageur de commerce est NP-complet. Pour se faire, on va montrer une suite de réductions depuis le problème 3-SAT dont on sait qu'il est NP-complet vers le problème de seuil associé au problème du voyageur de commerce, en passant par 3 intermédiaires : l'existences de chemins ou de cycles hamiltoniens dans des graphes orientés ou non-orientés.

Différents problèmes de décision

Problème du voyageur de commerce (TSP) - seuil

Instance : $G = (S, A, w)$ un graphe pondéré non orienté, avec $|S| = n$, et un seuil k.
Solution : V si il existe un cycle hamiltonien $\mathcal{C} = (s_0, s_1, \ldots, s_{n-1}, s_n = s_0)$ tel que :

$$w(\mathcal{C}) \leq k$$

F sinon.

uHAM-CYCLE

Instance : $G = (S, A)$ un graphe non orienté, avec $|S| = n$.
Solution : V si il existe un cycle hamiltonien $\mathcal{C} = (s_0, s_1, \ldots, s_{n-1}, s_n = s_0)$, F sinon.

uHAM-PATH

Instance : $G = (S, A)$ un graphe non orienté, avec $|S| = n$, s et t deux sommets de S.
Solution : V si il existe un chemin hamiltonien $\mathcal{C} = (s = s_0, s_1, \ldots, s_{n-1} = t)$, F sinon.

dHAM-PATH

Instance : $G = (S, A)$ un graphe orienté, avec $|S| = n$, s et t deux sommets de S.
Solution : V si il existe un chemin hamiltonien $\mathcal{C} = (s = s_0, s_1, \ldots, s_{n-1} = t)$, F sinon.

3-SAT

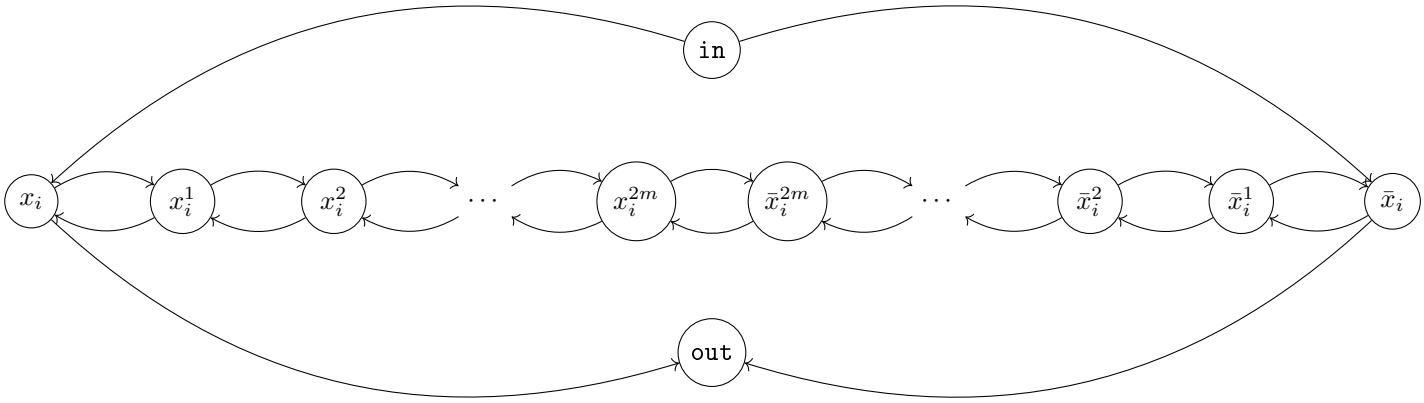
Instance : $\varphi = \wedge_{j=1}^{m} C_j$ avec $C_j = \ell_{j,1} \vee \ell_{j,2} \vee \ell_{j,3}$ où $\ell_{j,k} \in \cup_{i=1}^{n} \{x_i, \bar{x}_i\}$, une formule booléenne de m clauses sur n variables dont chaque clause contient 3 littéraux.
Solution : V si il existe une valuation des x_i qui satisfait φ, F sinon.

☐ 15 – Donner une réduction polynomiale de uHAM-CYCLE à la version seuil du problème du voyageur du commerce : à partir d'un graphe non orienté $G = (S, A)$, construire un graphe pondéré non orienté $G' = (S', A', w)$ et un seuil k tel que G' admet un cycle hamiltonien de poids inférieur à k si et seulement si G admet un cycle hamiltonien.

☐ 16 – Donner une réduction polynomiale de uHAM-PATH à uHAM-CYCLE.
On pourra ajouter un sommet pour clore le cycle.

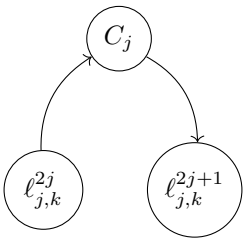
☐ 17 – Donner une réduction polynomiale de dHAM-PATH à uHAM-PATH.
On pourra par exemple découper chaque sommet en 3.

Pour chaque variable x_i, on considère le gadget G_i suivant, que l'on va ensuite mettre bout à bout :



☐ 18 – Combien y-a-t'il de chemins hamiltoniens de `in` à `out` dans ce graphe ?

Pour chaque littéral $\ell_{j,k}$ dans la clause C_j, on ajoute aussi la construction suivante :



☐ 19 – Montrer qu'un chemin hamiltonien dans G_i peut alors être modifié pour passer par tous les sommets C_j dans lesquels x_i figure mais aucun sommet C_j où $\bar{x}_i$ ne figure, ou vice-versa par tous les sommets C_j dans lesquels $\bar{x}_i$ figure sans passer par les sommets C_j dans lesquels x_i figure.

☐ 20 – Donner une réduction polynomiale de 3-SAT à dHAM-PATH, et conclure quand à la difficulté du problème de seuil associé au problème du voyageur du commerce.

Partie V. Difficulté d'approximer le problème d'optimisation – 4 pts

On peut aussi montrer qu'il est difficile d'approximer la solution avec un facteur constant d'approximation garanti.

On suppose dans cette partie qu'il existe un algorithme de complexité polynomiale qui, en prenant un graphe $G = (S, A, w)$ pondéré, non orienté et complet, renvoie un cycle hamiltonien $\mathcal{C}$ de G tel que si $\mathcal{C}^*$ est un cycle hamiltonien de poids minimal, alors $w(\mathcal{C}) \leq \alpha w(\mathcal{C}^*)$, avec $\alpha \geq 1$ est une constante fixée.
Pour $G = (S, A)$ un graphe non orienté, on considère le graphe pondéré non orienté complet $K_G = (S, S^2, w)$ défini par :

- pour $a \in A$, $w(a) = 1$;
- pour $a \notin A$, $w(a) = \alpha|S|$.

☐ 21 – Montrer que $G = (S, A)$ possède un cycle hamiltonien si et seulement si K_G possède un cycle hamiltonien de poids inférieur ou égal à $\alpha|S|$.

☐ 22 – En déduire qu'on peut résoudre uHAM-PATH en temps polynomial, et conclure sur la difficulté d'approximer la solution du problème du voyageur de commerce.

Partie VI. Approximation dans le cas métrique en utilisant un arbre couvrant minimal (OCaml) – 6 pts

Le problème du voyageur de commerce est difficile à approximer dans le cas général, mais sous certaines hypothèse, on peut trouver des algorithmes efficaces qui s'approchent du résultat. On s'intéresse ici à un graphe pondéré complet $G = (S, A, w)$ dont la fonction de poids w vérifie l'inégalité triangulaire :

$$\forall s, t, u \in S^3, w(s, u) \leq w(s, t) + w(t, u).$$

C'est le cas par exemple si on prend pour sommet des points du plan et pour poids la distance euclidienne entre ces points.
Pour la suite, on note $\mathcal{C}^*$ un cycle hamiltonien de poids minimal de G et T un arbre couvrant minimal. On étend de manière naturelle la fonction w aux sous-graphes et aux ensembles d'arêtes de G comme la somme des poids des arêtes qui les composent.

☐ 23 – Montrer que $w(T) \leq w(\mathcal{C}^*)$

☐ 24 – En considérant un parcours en profondeur préfixe de T, en déduire l'existence d'un cycle hamiltonien $\mathcal{C}_T$, calculable en temps polynomial en la taille de G, tel que $w(\mathcal{C}_T) \leq 2w(\mathcal{C}^*)$

☐ 25 – Écrire une fonction `PVC_approx : int array array -> int array` qui prend en argument un graphe $G = (S, A, w)$ représenté par une matrice d'adjacence et renvoie un tableau correspondant à un cycle hamiltonien de G dont le poids est au plus le double du poids minimal d'un cycle hamiltonien de G.

On pourra suppose l'existence d'une fonction `kruskal` qui implémente l'algorithme de Kruskal et dont vous préciserez le type.

Fin de l'épreuve