Programmation dynamique

INFORMATIQUE COMMUNE - TP n° 3.3 - Olivier Reynet

À la fin de ce chapitre, je sais :

- ☞ cerner les limitations des algorithmes gloutons dans certaines situations
- ☞ justifier l'optimalité d'une sous-structure d'un problème en programmation dynamique
- ☞ coder de bas en haut une problème en programmation dynamique
- ☞ utiliser la mémoïsation et un dictionnaire pour pallier l'inefficacité de l'approche récursive

A Le sac à dos est de retour

On cherche à remplir un sac à dos. Chaque objet que l'on peut insérer dans le sac est **insécable** [1] et possède une valeur et un poids connus. On cherche à maximiser la valeur totale emportée dans la sac à dos tout en limitant [2] le poids à π.

Soit un ensemble $\mathcal{O}_n = \{o_1, o_2, \ldots, o_n\}$ de n objets de valeurs $v_1, v_2, \ldots, v_n$ et de poids respectifs $p_1, p_2, \ldots, p_n$. Soit un sac à dos n'admettant pas un poids emporté supérieur à π. On note également qu'on peut mettre au plus n objets dans le sac.

Les objets sont rangés dans une liste et dans un ordre quelconque. Ils sont indicés par i variant de 1 à n. Un objet o_i possède une valeur v_i et un pèse p_i.

Avec ces notations, on peut formuler le problème du sac à dos comme suit.

■ **Définition 1 — Problème du sac à dos.** Comment remplir un sac à dos en maximisant la valeur totale emportée $V = \sum_{i=1}^{n} v_i$ tout en ne dépassant pas le poids maximal admissible par le sac à dos, c'est à dire en respectant la contrainte $\sum_{i=1}^{n} p_i \leqslant \pi$? On note [a] le problème du sac à dos $\text{KP}(n, \pi)$ et une solution optimale à ce problème $S(n, \pi)$.

a. en anglais, ce problème est nommé Knapsack Problem, d'où le KP.

On dispose d'une collection d'objets dont les valeurs et les poids sont les suivants :

- `valeurs = [100, 700, 500, 400, 300, 200]`,
- `poids = [40, 15, 2, 9, 18, 2]`.

Un objet `i` possède une valeur `valeurs[i]` et un poids `poids[i]`.

A1. Pour une poids maximal admissible de 11 kg, quel peut-être un chargement optimal du sac à dos ?

A2. Pour le problème $\text{KP}(n, \pi)$, implémenter un algorithme de résolution glouton.

1. Soit on le met dans le sac, soit on ne le met pas. Mais on ne peut pas en mettre qu'une partie.
2. On accepte un poids total inférieur ou égal à π.

A3. Le problème du sac à dos est-il à sous-structure optimale ?

A4. En utilisant la programmation dynamique de bas en haut, coder la résolution du problème KP(n,π).

A5. En comparant les deux stratégies précédentes, identifier les cas pour lesquels l'algorithme glouton n'est pas optimal.

A6. Programmer une version gloutonne de KP(n,π) qui choisit le meilleur objet d'après le ratio valeur / poids.

A7. En comparant les trois stratégies, identifier les cas pour lesquels les algorithmes gloutons ne sont pas optimaux.

B Rendu de monnaie

Un commerçant doit à rendre la monnaie à un client[3]. La somme à rendre est un somme entière P et le commerçant cherche à utiliser le moins de pièces possibles. On considère qu'il dispose d'autant de pièces qu'il le souhaite parmi un système monétaire $M = \{m_1, m_2, \dots, m_n\}$ qui possède n valeurs différentes.

On nomme ce problème le rendu de monnaie[4] et on note CCP(M,i,P) le problème où il s'agit de rendre la monnaie P à l'aide des i premières pièces du système M. On note $S(i,P)$ une solution optimale au problème CCP(M,i,P), c'est à dire une solution qui nécessite **le moins de pièces possibles**.

B1. On considère les systèmes monétaire $M_c = [5,1,2]$ et $M = [4,1,3]$. Rendre la monnaie de manière optimale sur 10 et 6 avec chaque système.

B2. Résoudre le problème CCP(M,n,P) en implémentant un algorithme glouton. Tester l'algorithme sur les systèmes $M_c = [5,1,2]$ et $M = [4,1,3]$. Que constatez-vous ?

B3. Le problème du rendu de monnaie est-il à sous-structure optimale ? Pourquoi ?

B4. En utilisant la programmation dynamique et un tableau, coder la résolution du problème CCP(M,n,P) de bas en haut. On utilisera un dictionnaire pour mémoriser la liste des pièces nécessaires.

B5. En comparant les deux stratégies précédentes, identifier les cas pour lesquels l'algorithme glouton n'est pas optimal pour les systèmes monétaires M et M_c définis plus haut.

C Distance d'édition

Les séquences de caractères peuvent encoder de nombreuses informations de nature différente, par exemple du texte, de la voix ou des séquences ADN. L'alignement de deux chaînes des caractères consiste à comparer deux séquences de caractères afin d'évaluer la similarité entre les deux.

La distance d'édition ou distance de Levenshtein est une mesure de la similarité entre deux chaînes de caractères. Cette distance est le nombre minimal de caractères qu'il faut supprimer, insérer ou substituer pour passer d'une chaîne à l'autre.

■ **Définition 2 — Distance d'édition.** Soit a et b deux chaînes de caractères. On note $|a|$ le cardinal de a, c'est à dire le nombre de caractères de la chaîne. $a[0]$ désigne le premier caractère de la chaîne a. On dénote par $a[:-1]$ la chaîne a tronquée de son premier caractère.

3. Mais on pourrait considérer d'autres problèmes qui se résoudraient de la même manière. Par exemple, le remplissage d'un conteneur dont le volume total est V à l'aide d'objets de volume $v_1, v_2, \dots, v_n$. On dispose d'autant d'objets que l'on veut pour compléter le conteneur mais on souhaite en charger le moins possible.

4. En anglais Coin Change Problem.

On suppose que :

- supprimer un caractère,
- insérer un caractère,
- substituer un caractère,

sont des opérations qui ont toute un coût **unitaire (1)**. Si le caractère est identique, la substitution en coûte rien (0).

La distance d'édition est définition par induction de la manière suivante :

$$d_e(a,b) = \begin{cases} \max(|\,a\,|,|\,b\,|) & \text{si } \min(|\,a\,|,|\,b\,|) = 0 \\ d(a[1:],b[1:]) & \text{si } a[0] = b[0] \\ 1 + \min \begin{cases} d(a[1:],b) \\ d(a,b[1:]) \\ d(a[1:],b[1:]) \end{cases} & \text{sinon} \end{cases} \tag{3}$$

Si on souhaite programmer dynamiquement par le bas en utilisant un tableau S contenant les distances, il est utile d'exprimer le résultat d'une case en fonction de celles dont elle dépend dans le schéma dynamique :

$$S[i,j] = \begin{cases} \max(i,j) & \text{si } \min(i,j) = 0 \\ S(i-1,j-1) & \text{si } a[i] = b[j] \\ 1 + \min \begin{cases} S[i-1,j] \\ S[i,j-1] \\ S[i-1,j-1] \end{cases} & \text{sinon} \end{cases} \tag{4}$$

C1. La distance d'édition de `"chien"` à `"niche"` vaut 4. Expliquer pourquoi.

C2. La distance d'édition représente-t-elle un problème à sous-structure optimale ? Pourquoi ?

C3. On souhaite utiliser la programmation dynamique. Compléter à la main et de bas en haut le tableau associé à la distance d'édition de `"chien"` à `"niche"`.

C4. Écrire un code qui calcule la distance d'édition de deux chaînes de caractères par programmation dynamique de bas en haut. On pourra tester sur `"AGTTC"` et `"AGCTC"`, sur `"chien"` et `"niche"` ou `"sunday"` et `"saturday"`.

C5. Écrire un code similaire en programmation dynamique avec memoïsation et comparer les résultats.

D Plus longue sous-chaîne commune

La distance d'édition permet de mesurer le degré de similarité de deux chaînes. Elle ne donne pas d'information quant aux séquences maximales communes aux deux chaînes. Hors, en génétique par exemple, il peut s'avérer très important de savoir quels sont les points communs de deux génomes. Le problème de la plus longue sous-chaîne permet d'apporter une réponse à cette question.

■ **Définition 3 — Sous-chaîne.** On appelle sous-chaine d'une chaîne de caractères $a = a_1 ... a_n$ toute chaine de caractères s extraite de a telle que $s = a_i ... a_k$ où $i \leqslant \ldots \leqslant k < n$ et $\forall p \in \{i, \ldots, k\}, a_p \in a$. Les caractères de s n'apparaissent pas nécessairement de manière consécutive dans la chaîne a.

■ **Définition 4 — Plus longue sous-chaîne commune.** Soit $a = a_1 \dots a_q$ et $B = b_1 \dots b_p$ deux chaines de caractères non vides. On appelle plus longue sous-chaine commune à a et b toute sous-chaîne commune à a et b de longueur maximale.

Si l'une des chaînes a ou b est vide ou si a et b nont aucune sous-chaîne commune, la chaine vide est alors l'unique plus longue sous-chaîne commune à a et b.

■ **Exemple 1 — Plus longue sous-chaîne commune.** Par exemple, les chaines de caractères `"AAA"` et `"TAA"` sont les plus longues sous-chaînes communes aux chaines de caractères `"ATAGA"` et `"TAACA"`. La chaine de caractères `"ATGC"` est une plus longue sous-chaine commune aux chaines de caractères `"AATGCG"` et `"TATTAGC"`.

Le problème de la plus longue sous-chaîne commune entre a et b est noté $\mathcal{L}(a,b)$, son résultat est la longueur maximale d'une sous-chaîne commune à a et b.

D1. On considère les chaines `a="AATGCG"` et `b="TATTAGC"` ? Donner les solutions de $\mathcal{L}(a,b)$.

D2. Écrire une fonction de prototype `is_ss(ch,sch)` où les paramètres sont deux chaînes de caractères et qui renvoie `True` si `sch` est une sous-chaîne de `ch` et `False` sinon.

D3. Écrire une fonction de prototype `is_common_ss(a,b,sch)` où les paramètres sont des chaînes de caractères et qui renvoie `True` si `sch` est une sous-chaîne commune à `a` et `b`.

D4. Formuler le problème $\mathcal{L}(a,b)$ récursivement afin de pourvoir justifier de sa sous-structure optimale.

D5. Écrire un code qui résout $\mathcal{L}(a,b)$ avec la programmation dynamique du haut vers le bas.

D6. Résoudre $\mathcal{L}(a,b)$ récursivement avec mémoïsation.

D7. Faire en sorte que les codes qui résolvent $\mathcal{L}(a,b)$ produisent également les sous-chaînes solutions.

E Algorithme de Floyd-Warshall

L'algorithme de Floyd-Warshall est l'application de la programmation dynamique à la recherche du plus court chemin entre deux sommets d'un graphe orienté et valué. Le plus court chemin n'est pas celui qui comporte le moins de sommets mais celui dont la somme des poids de chaque arc est la plus faible [5]. Les valuations peuvent être négatives mais on exclue tout circuit de poids strictement négatif.

Soit un graphe orienté et valué $G = (S, A, C)$ décrit par ses $n = |\,S\,|$ sommets $S = \{s_1, s_2, \dots, s_n\}$, ses arcs A et les coûts associés aux arcs C.

G peut être modélisé par une matrice d'adjacence M :

$$\forall i,j \in [\![1, |\,S\,|]\!], M = \begin{cases} C(s_i,s_j) & \text{si } (s_i,s_j) \in A \\ +\infty & \text{si } (s_i,s_j) \notin A \\ 0 & \text{si } i = j \end{cases} \tag{6}$$

Un exemple de graphe associé à sa matrice d'adjacence est donné sur la figure 1.

Pour trouver le plus court chemin entre deux sommets, on essaye tous les chemins de toutes les longueurs possibles et on ne garde que les plus courts. Chaque étape p de l'algorithme de Floyd-Warshall est donc constitué d'un allongement **éventuel** du chemin par le sommet s_p. À l'étape p, on associe une matrice M_p qui contient la longueur des chemins les plus courts d'un sommet à un autre passant par

5. Dans un réseau de télécommunications, il s'agit bien du chemin le plus court si les poids des arcs sont les débits en Gbits/s des liens.

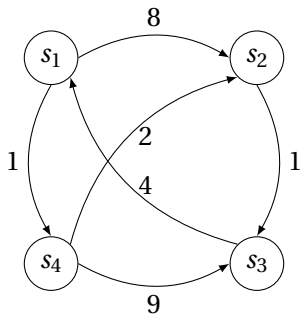

FIGURE 1 – Exemple de graphe orienté et valué. La matrice d'adjacence de ce graphe $G = (S, A, V)$ vaut : $M = \begin{pmatrix} 0 & 8 & +\infty & 1 \\ +\infty & 0 & 1 & +\infty \\ 4 & +\infty & 0 & +\infty \\ +\infty & 2 & 9 & 0 \end{pmatrix}$. Sur cet exemple, le chemin le plus court de s_4 à s_3 vaut 3 et passe par s_2.

des sommets de l'ensemble $\{s_1, s_2, \ldots, s_p\}$. On construit ainsi une suite de matrice finie $\left(M_p\right)_{p\in[\![0,n]\!]}$ avec $M_0 = M$.

Supposons qu'on dispose de M_p. Considérons un chemin $\mathcal{C}$ entre s_i et s_j dont la longueur est minimale et dont les sommets intermédiaires sont dans $\{s_1, s_2, \ldots, s_{p+1}\}$, $p \leqslant n$. Pour un tel chemin :

- soit $\mathcal{C}$ passe par s_{p+1}. Dans ce cas, $\mathcal{C}$ est la réunion de deux chemins dont les sommets sont dans $\{s_1, s_2, \ldots, s_{p+1}\}$: celui de s_i à s_{p+1} et celui de s_{p+1} à s_j.
- soit $\mathcal{C}$ ne passe pas par s_{p+1}.

Entre ces deux chemins, on choisira le chemin le plus court.

Disposer d'une formule de récurrence entre M_{p+1} et M_p permettrait de montrer que le problème du plus court chemin entre deux sommets d'un graphe orienté et valué est à sous-structure optimale. On pourrait alors utiliser la programmation dynamique pour résoudre le problème.

E1. Formuler le problème du plus court chemin entre deux sommets d'un graphe orienté afin de montrer que ce problème est à sous-structure optimale.

E2. Coder une fonction qui implémente l'algorithme de Floyd-Warshall et la programmation dynamique de bas en haut. Tester ce code sur l'exemple de la figure 1. On pourra utiliser un tableau numpy à trois dimensions.

E3. Quelle est la complexité temporelle de cet algorithme ?

E4. Quelle est la complexité spatiale de cet algorithme ? Pourrait-on l'améliorer ? Comment ?

R Pour ceux qui font l'option informatique, cet algorithme est celui qui permet de trouver la forme d'une expression régulière dénotant le langage rationnel défini par un automate fini (algorithme de McNaughton et Yamada).