Programmation linéaire en nombres entiers

Résolution : coupes et séparation - évaluation

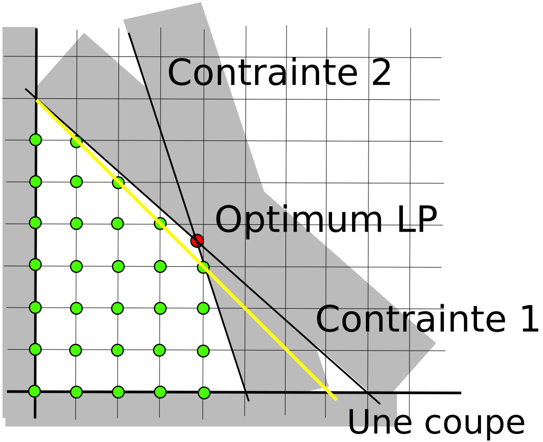
Hugues Talbot

Laboratoire A2SI

26 mars 2009

Plan

Principe des coupes



Formulation initiale

- Soit une “relaxation” linéaire d’un problème IP, c’est-à-dire le problème de PL correspondant, dont la solution (par le simplexe par exemple) optimale est X^*.
  - soit X^* est entière, auquel cas c’est la solution optimale du problème IP.
  - soit elle ne l’est pas, X^* est fractionnaire.
- Considérons la base correspondant à la solution optimale de PL

$$X_b = B^{-1}[b - EX_e]; \bar{b} = B^{-1}b.$$

avec $\bar{b}_i = B^{-1}b$ fractionnaire.

- Nous avons

$$X_{b_i} = e_i^t B^{-1}[b - EX_e]$$

où $e_i^t = [000 \ldots \underset{\text{pos}\,i}{1} \ldots 00]$

Formulation (suite)

- Ainsi :

$$\begin{aligned} X_{b_i} &= e_i^t B^{-1} b - e_i^t B^{-1} E X_e \\ &= \bar{b}_i - \alpha^i X_e \end{aligned}$$

c-à-d :

$$\alpha^i = e_i^t B^{-1} E \text{ (ligne } i \text{ de } B^{-1}E)$$

- En réordonnant :

$$\begin{aligned} X_{b_i} + \alpha^i X_e &= \bar{b}_i \\ X_{b_i} + \sum_j \alpha_j^i X_{e_j} &= \bar{b}_i \qquad (1) \end{aligned}$$

Partie entière / partie fractionnaire

- Définition : soit a un non-entier,
  - ent(a) = plus grand entier $< a$.
  - frac(a) = a – ent(a)($donc$) > 0.
- de (1) on a :

$$\begin{aligned} X_{b_i} + \sum_j [\text{ent}(\alpha_j^i) + \text{frac}(\alpha_j^i)] X_{e_j} &= \text{ent}(\bar{b}_i) + \text{frac}(\bar{b}_i) \\ X_{b_i} + \sum_j \text{ent}(\alpha_j^i) X_{e_j} + \sum_j \text{frac}(\alpha_j^i) X_{e_j} &= \text{ent}(\bar{b}_i) + \text{frac}(\bar{b}_i) \qquad (2) \end{aligned}$$

  avec $\text{frac}(\alpha_j^i) \geq 0$ et $X_{e_j} \geq 0$.
- donc

$$X_{b_i} + \sum_j \text{ent}(\alpha_j^i) X_{e_j} \leq \text{ent}(\bar{b}_i) + \text{frac}(\bar{b}_i)$$

Coupe de Gomory

- Mais X_{b_i} est entier et X_{e_j} aussi.
- Donc

$$X_{b_i} \sum_j \text{ent}(\alpha_j^i) X_{e_j} \leq \text{ent}(\bar{b}_i) \qquad (3)$$

- de (3)–(2) on déduit

$$\begin{aligned} -\sum_j \text{frac}(\alpha_j^i) X_{e_j} &\leq -\text{frac}(\bar{b}_i) \\ \sum_j \text{frac}(\alpha_j^i) X_{e_j} &\geq \text{frac}(\bar{b}_i) \qquad (4) \end{aligned}$$

(4) est une coupe de Gomory.

Exemple

- Soit le problème suivant :

$$\begin{array}{lccccc} \max & x_1 & + & x_2 & & \\ & 3x_1 & + & 4x_2 & \leq & 15 \\ & x_1 & - & 4x_2 & \leq & 0 \end{array}$$

- Sous forme standard :

$$\begin{array}{cccccl} 3x_1 & +4x_2 & +x_3 & & = & 15 \\ x_1 & -4x_2 & & +x_4 & = & 0 \end{array}$$

Coupe 1

- Solution optimale : base $X_b = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 4 \\ 1 & -4 \end{bmatrix}$
- Inverse : $B^{-1} = -\frac{1}{16}\begin{bmatrix} -4 & -4 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/4 \\ 1/16 & -3/16 \end{bmatrix}$
- Solution :
  $\bar{b} = B^{-1}b = \begin{bmatrix} 1/4 & 1/4 \\ 1/16 & -3/16 \end{bmatrix}\begin{bmatrix} 15 \\ 0 \end{bmatrix} = \begin{bmatrix} 15/4 \\ 15/16 \end{bmatrix}$
- Coupe (1) : $\alpha^1 = [\frac{1}{4}\frac{1}{4}], \bar{b}_1 = \frac{15}{4}$
  $\frac{1}{4}x_3 + \frac{1}{4}x_4 \geq \frac{3}{4}$

$$x_3 + x_4 \geq 3 \qquad \text{coupe (1)}$$

Coupe 2

- Coupe (2) :
  $\alpha^2 = [\frac{1}{16} - \frac{3}{16}], \bar{b}_2 = \frac{15}{16}$, Attention ! $\text{frac}(\alpha^2) = [\frac{1}{16} \frac{13}{16}]$
  $\frac{1}{16}x_3 + \frac{13}{16}x_4 \geq \frac{15}{16}$

$$x_3 + 13x_4 \geq 15 \qquad \text{coupe (2)}$$

- En soit pas très intéressantes : on exprime les variables hors-base en fonction des variables de base :

$$\begin{array}{rcrrr} x_3 & = & 15 & -3x_1 & -4x_2 \\ x_4 & = & & -x_1 & +4x_2 \end{array}$$

Coupes finales

- En substituant dans coupe 1 :

$$\begin{aligned} 15 - 3x_1 - 4x_2 - x_1 + 4x_2 &\geq 3 \\ 12 &\geq 4x_1 \\ x_1 &\leq 3 \end{aligned}$$

- En substituant dans coupe 2 :

$$\begin{aligned} 15 - 3x_1 - 4x_2 - 13x_1 + 52x_2 &\geq 15 \\ -16x_1 + 48x_2 &\geq 0 \\ x_1 - 3x_2 &\leq 0 \end{aligned}$$

- On ajoute les nouvelles coupes au problème de départ, et cette fois-ci on trouve :
  $\bar{b} = [3, \frac{3}{2}]$, on doit faire une nouvelle itération.

Solution par séparation-évaluation

- On a déjà vu une méthode de solution des problèmes de IP en utilisant des coupes (de Gomory).
- Dans le cas où les contraintes et variables sont rationnelles (a fortiori entières), les coupes de Gomory convergent vers la solution optimale en temps fini.
- Les coupes attaquent le problème par l'extérieur en gardant le domaine réalisable convexe, ce qui peut être inefficace dans certaines configurations
- On peut aussi résoudre les problèmes de IP/MP en fractionnant le domaine en régions faisables bornées par des frontières alignées sur des entiers : on *sépare* ainsi les domaines, et on *évalue* quelle région explorer en premier : on appelle cette méthode la séparation-évaluation, ou *branch-and-bound* en anglais.

Les tables et les chaises

- Une entreprise fabrique des armoires et des tables ;
- Une armoire nécessite 1h de travail et 9 m^2 de bois ;
- Une table nécessite 1h de travail et 5m^2 de bois ;
- On dispose de 6h de travail et de 45 m^2 de bois ;
- Chaque armoire génère un profit de 8 €, et chaque table €;
- Formuler et résoudre en entiers.

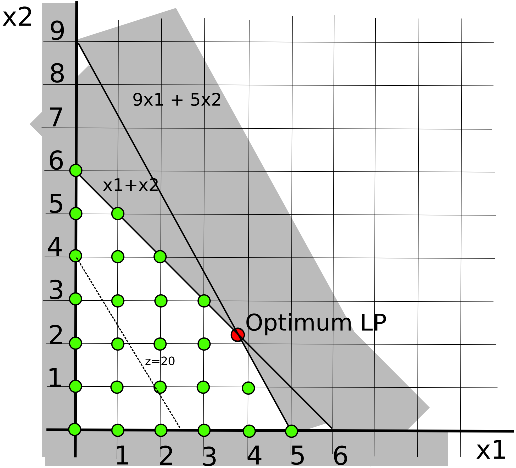
Formulation

On formule de la façon suivante :

$$\begin{array}{rcccl} \max z = & 8x_1 & + & 5x_2 & \\ & x_1 & + & x_2 & \leq 6 \\ & 9x_1 & + & 5x_2 & \leq 45 \end{array}$$

$x_i \geq 0, x_i \in \mathbb{N}$.

Représentation



Probleme initial (1)

Méthode générale

- On commence par résoudre la formulation de relaxation du problème en entiers, c-à-d le problème en variables continues (LP). Ici, on obtient :

$$\begin{aligned} z &= \frac{165}{4} \\ x_1 &= \frac{15}{4} \\ x_2 &= \frac{9}{4} \end{aligned}$$

- Si la solution est en entiers, on s'arrête, on a trouvé l'optimal. Ici ce n'est pas le cas, mais la valeur z obtenue est une borne supérieure pour l'optimum en entiers.

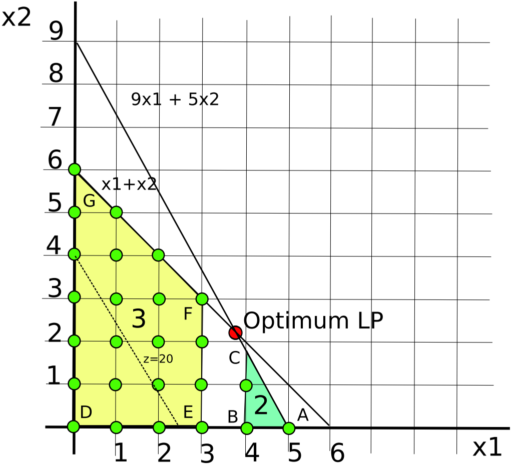
Methode générale (suite)

- Si la solution n'est pas entière, on partitionne le domaine : on choisit arbitrairement une variable qui est fractionnaire dans la solution optimale relaxée, par exemple ici x_1.
- On applique des contraintes supplémentaires dues à la nature de la variable, ici par exemple on impose soit $x_1 \leq 3$, soit $x_1 \geq 4$ (Question : pourquoi pouvons nous éliminer les solutions $3 < x_1 < 4$?)

Séparation

- On a donc maintenant deux sous-problèmes :
  2. Problème initial + contrainte $x_1 \geq 4$
  3. Problème initial + contrainte $x_2 \leq 3$.

Représentation (2)



Sous-problemes 2+3

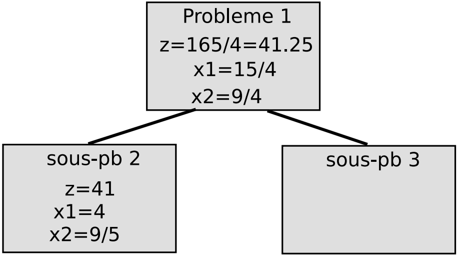
Récursion

- On constate qu'on a, ce faisant, éliminé la région contenant l'optimum de LP.
- On a maintenant deux problèmes qu'on peut de nouveau résoudre en LP.
- On en choisi un arbitrairement parmi ceux non résolus, par exemple ici le problème 2.
- La solution de relaxation (LP) pour la région 2 est

$$\begin{aligned} z &= 41 \\ x_1 &= 4 \\ x_2 &= \frac{9}{5} \end{aligned}$$

  Correspondant au point C.
- On a créé la première branche d'un *arbre*.
- Comme x_2 est toujours fractionnaire, on décide de séparer sur cette variable, on sépare donc la région 2 entre deux zones : celle pour $x_2 \geq 2$ et celle pour $x_2 \leq 1$.

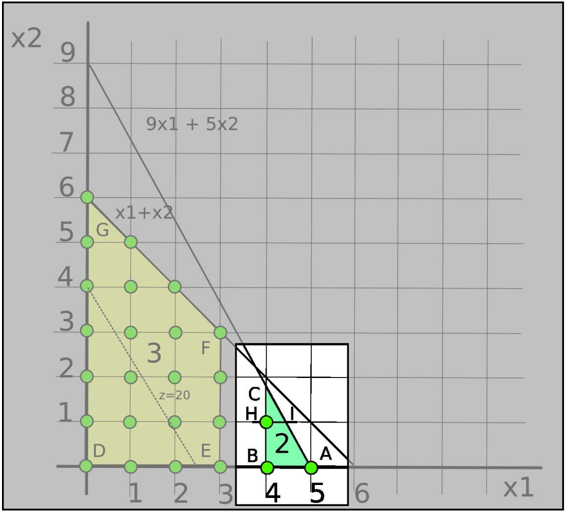
Arbre (1+2)


Probleme 1
z=165/4=41.25
x1=15/4
x2=9/4
sous-pb 2
z=41
x1=4
x2=9/5
sous-pb 3

Nouvelle séparation (4 + 5)

- Nous voici avec deux nouveaux sous-problèmes
  4. Problème 2 + contrainte $x_2 \geq 2$
  5. Problème 2 + contrainte $x_2 \leq 1$.
- Il est préférable d'explorer l'arbre en profondeur d'abord (on verra plus tard pourquoi), on choisi donc de regarder l'un ou l'autre des nouveaux sous-problèmes, par exemple le sous-problème 4. Or ce problème n'est pas réalisable en entiers.

Représentation (4 et 5)



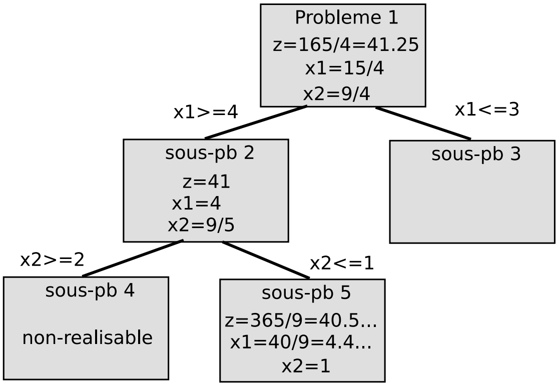
Évaluation (4 + 5)

- On constate que le problème 4 (région $x_2 \geq 2$) n'est pas réalisable.
- En résolvant le problème de relaxation LP lié au sous-problème 5, on trouve l'optimum I avec

$$
\begin{aligned}
z &= 365/9 = 40.555\ldots \\
x_1 &= \frac{40}{9} = 4.444\ldots \\
x_2 &= 1
\end{aligned}
$$

- il faut donc de nouveau séparer sur x_1, avec les contraintes
  6. Problème 5 + contrainte $x_1 \geq 5$
  7. Problème 5 + contrainte $x_2 \leq 4$.

Arbre (4+5)


Probleme 1
z=165/4=41.25
x1=15/4
x2=9/4
x1>=4
x1<=3
sous-pb 2
z=41
x1=4
x2=9/5
sous-pb 3
x2>=2
x2<=1
sous-pb 4
non-realisable
sous-pb 5
z=365/9=40.5...
x1=40/9=4.4...
x2=1

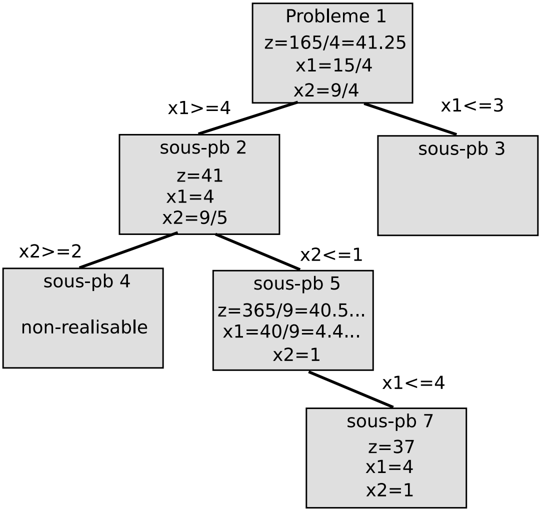
Séparation (6+7)

- Nous voici avec deux nouveaux problèmes, comme d'habitude, on va éliminer la solution fractionnaire en séparant de part et d'autre.
- Nous procédons en profondeur d'abord parmi les nouveaux sous-problèmes non-résolus (6 et 7). Nous choisissons arbitrairement 7.
- La solution de relaxation LP est maintenant :

$$\begin{aligned} z &= 37 \\ x_1 &= 4 \\ x_2 &= 1 \end{aligned}$$

  Cette solution est réalisable et correspond au point H, c'est une *solution candidate* qui nous donne une *borne inférieure* sur notre résultat final.
- Il est inutile de continuer à séparer sur cette branche.

Arbre (7)


Probleme 1
z=165/4=41.25
x1=15/4
x2=9/4
x1>=4
x1<=3
sous-pb 2
z=41
x1=4
x2=9/5
sous-pb 3
x2>=2
x2<=1
sous-pb 4
non-realisable
sous-pb 5
z=365/9=40.5...
x1=40/9=4.4...
x2=1
x1<=4
sous-pb 7
z=37
x1=4
x2=1

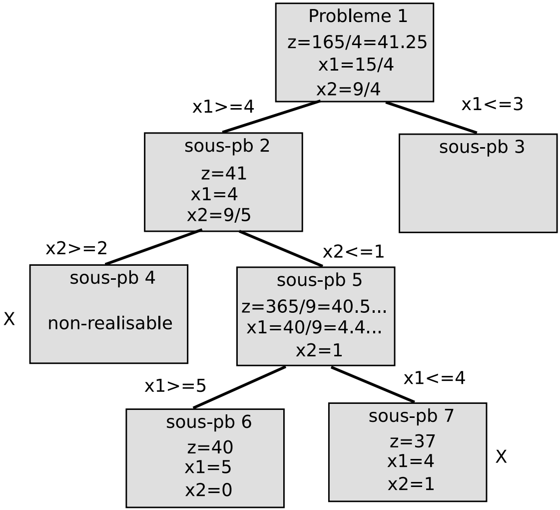
Évaluation (6)

- On continue à évaluer en profondeur d'abord, soit maintenant le sous-problème 6.
- On trouve la solution correspondant au point A :

$$\begin{aligned} z &= 40 \\ x_1 &= 5 \\ x_2 &= 0 \end{aligned}$$

  Il s'agit également d'une solution réalisable candidate. La valeur de la borne inférieure de notre problème est donc maintenant 40.
- La solution 7 n'est donc pas optimale.
- Il est inutile de séparer davantage sur 6.

Arbre (6)


Probleme 1
z=165/4=41.25
x1=15/4
x2=9/4
x1>=4
x1<=3
sous-pb 2
z=41
x1=4
x2=9/5
sous-pb 3
x2>=2
x2<=1
sous-pb 4
X
non-realisable
sous-pb 5
z=365/9=40.5...
x1=40/9=4.4...
x2=1
x1>=5
x1<=4
sous-pb 6
z=40
x1=5
x2=0
sous-pb 7
z=37
x1=4
x2=1
X

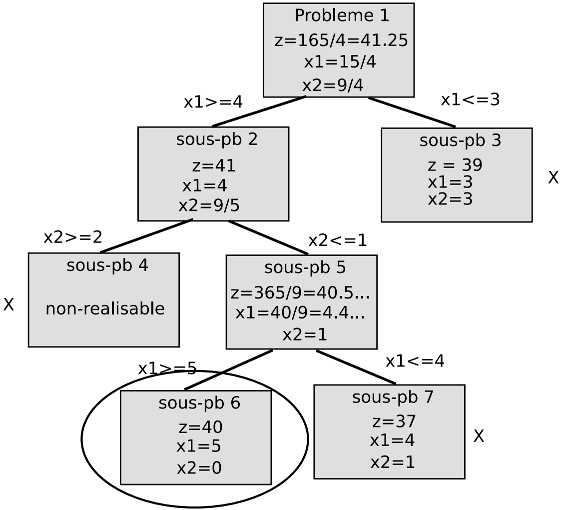
Évaluation (3)

- Il reste à évaluer la solution pour (3). On trouve la solution correspondant au point F :

$$\begin{aligned} x_1 &= 3 \\ x_2 &= 3 \\ z &= 39 \end{aligned}$$

- Ce résultat est inférieur à 40, notre borne inf, donc cette branche de l'arbre de peut pas produire un meilleur résultat que celui déjà connu correspondant au point A.
- Il ne reste plus de nœud de l'arbre à explorer, on a donc trouvé notre optimum en nombre entier : fabriquer 5 armoires et 0 table pour 40€.

Arbre (3)


Probleme 1
z=165/4=41.25
x1=15/4
x2=9/4
x1>=4
x1<=3
sous-pb 2
z=41
x1=4
x2=9/5
sous-pb 3
z = 39
x1=3
x2=3
X
x2>=2
x2<=1
sous-pb 4
non-realisable
X
sous-pb 5
z=365/9=40.5...
x1=40/9=4.4...
x2=1
x1>=5
x1<=4
sous-pb 6
z=40
x1=5
x2=0
sous-pb 7
z=37
x1=4
x2=1
X

Label des points

- L'algorithme de séparation/évaluation labellise tous les points de la région réalisable, certains de manière explicite (les nœuds de l'arbre), et d'autres implicitement.
- Par exemple, le point ($x_1 = 2, x_2 = 3$) est contenu dans le sous-problème 3, dont l'optimum est en $(3, 3)$, il est donc inutile de l'évaluer.
- L'algorithme divise pour régner, et converge nécessairement du au fait qu'il élimine toujours des points à chaque étape de séparation, alors que ceux-ci sont en nombre fini.

Règles de stérilisation de sommets

1. On dit qu'un sommet de l'arbre qu'il est inutile de séparer est *stérilisé*. Il y a plusieurs possibilités pour cela :
   1.1 Le sommet est associé à une solution réalisable candidate (séparer ne l'améliorera pas)
   1.2 Le sommet est associé à un sous-problème non-réalisable.
   1.3 La valeur z optimal du sous-problème associé est inférieure (pour un problème max, et vice-versa pour un problème min) à la borne inférieure courante.
2. Un sous-problème peut-être éliminé directement si
   2.1 Le sous-problème est non-réalisable
   2.2 La borne inférieure courante est supérieure à la valeur z du sous-problème.

Retour en arrière

- On peut résoudre les problèmes de séparation-évaluation soit en largeur d'abord (on évalue tous les sous-problèmes à un niveau avant de séparer au niveau inférieur : dans l'exemple on commencerait à évaluer (3) avant (4) ou (5)).
- Ou alors comme dans l'exemple en profondeur d'abord comme dans l'exemple.
- La pratique montre que l'exploration en profondeur d'abord fonctionne mieux, en particulier avec une bonne fonction d'estimation comme la relaxation LP (voir plus loin) : on divise pour régner plutôt que d'explorer systématiquement.
- ATTENTION : séparation-évaluation n'est pas glouton : on effectue des retours en arrière (backtrack).

Résoudre les problèmes mixtes

- On ne sépare que sur les variables entières !

Généralisation de sep/eval

- La relaxation LP n'est qu'une heuristique !
- Elle peut être remplacée par une plus performante
- Exemple : algorithme du knapsack (voir en TD).
- Lien avec l'algorithme $A^\star$ d'exploration combinatoire.

Knapsack

- Rappel : un knapsack (problème de sac à dos) est un problème à une seule contrainte.

$$\begin{aligned} \max z &= \textstyle\sum_i c_i x_i \\ & \textstyle\sum_i a_i x_i \quad \leq b \end{aligned}$$

  avec $x_i \in \{0, 1\}$
- Séparer sur x_i donne une branche $x_i = 0$ et une branche $x_i = 1$.
- On peut résoudre la relaxation LP par inspection en remarquant qu'on peut ordonner les variables par rapport $\frac{c_i}{a_i}$ décroissant. Les variables avec le meilleur rapport sont préférables.
- Voir TD associé.
- On peut aussi résoudre un knapsack par programmation dynamique.

Planification

- Les problèmes de planification avec dates butoir sont très courants, par exemple : optimisation de ressource, gestion temps-réel, etc.
- Illustration :

	durée	date butoir
tâche 1	6	8
tâche 2	4	4
tâche 3	5	12
tâche 4	8	16

- On lit la table de la façon suivante : la tâche 1 requiert 6 jours et doit être complétée à la fin du 8ème jour.
- Chaque jour de retard résulte en des pénalités. On doit minimiser les jours de retard.

Solution par S/E

- On propose d'utiliser les variables suivantes :

$$x_{ij} = \begin{cases} 1 & \text{si la tâche } i \text{ est en position } j \\ 0 & \text{sinon} \end{cases}$$

- On peut ici proposer de séparer en assignant d'abord la *dernière* tâche, et en évaluant le délai minimal total en faisant la somme des durées pour tâches assignées et restantes moins le délai pour la dernière tâche
- Ici, si la tâche 3 est la dernière assignée ($x_{34} = 1$), on a un délai minimal estimé de $6 + 4 + 8 + 5 - 12 = 23 - 12 = 11$.

Solution par S/E, suite

- Par la suite dans l'arbre on sépare en assignant les tâches en commençant vers la fin (dernière tâche, puis avant-dernière, etc), et on évalue en faisant la somme des délais évalués. Par exemple si la tâche 3 est choisie en dernier, puis la tâche 4 en avant-dernier, le délai total ne pourra pas être inférieur à $11 + (6 + 4 + 8 - 16) = 11 + 2 = 13$.
- On pourra vérifier que la solution optimale dans l'exemple est 2-1-3-4 avec un délai de 12 jours.

Travelling Salesperson Problem

- Notez la définition politiquement correcte...
- Rappel : on a un ensemble de villes qu'on veut toutes visiter, au coût moindre.
- Prenons la matrice de coût suivante, où c_{ij} représente le coût d'aller de la ville i à la ville j :

	ville 1	ville 2	ville 3	ville 4	ville 5
ville 1	0	132	217	164	58
ville 2	132	0	290	201	79
ville 3	217	290	0	113	303
ville 4	164	201	113	0	196
ville 5	58	79	303	196	0

TSP (II)

- Le voyage de coût minimum est nécessairement un circuit, qui est un cycle Hamiltonien du graphe associé.
- On peut le résoudre par E&S, pour cela, il faut un moyen d'évaluer et un moyen de séparer. On peut par exemple transformer TSP en problème d'affectation avec la matrice de coût c : soit x_{ij} un ensemble de variables binaires, indiquant que si $x_{ij} = 1$ alors le voyage de la ville i à la ville j est effectué, et 0 sinon.
- Dans ce cas, si on les variables $x_{12} = x_{24} = x_{45} = x_{53} = x_{31} = 1$, alors on a réalisé le circuit 1-2-4-5-3-1.
- Effectivement, si le problème d'affectation génère un circuit, alors il est optimal (pourquoi ?), mais l'affectation peut ne pas générer de circuits, par exemple on pourrait obtenir la solution $x_{15} = x_{21} = x_{34} = x_{43} = x_{52} = 1$, qui est une solution avec 2 sous-circuits non connexes.

TSP (III)

- On doit aussi imposer $i \neq j$, ce qui peut se faire en imposant un coût $c_{ii} = M$ élevé.
- On peut par exemple partir de la solution d'affectation, puis séparer en éliminant des sous-circuits. Pour éliminer un sous-circuit, il suffit d'imposer un coût élevé M à une des arêtes d'un des sous-circuits.
- On peut choisir, pour une arête joignant la ville i à la ville j, d'imposer un coût élevé dans un sens ou dans l'autre : on sépare sur ce choix.
- Par exemple, avec la table donnée (modifiée pour un coût c_{ii} élevé), la première solution d'affectation est $x_{15} = x_{21} = x_{34} = x_{43} = x_{52} = 1$.
- On a deux sous-circuits : 1-5-2-1 et 3-4-3. On peut décider d'éliminer le sous-circuit 3-4-3, soit en imposant $c_{34} = M$, soit $c_{43} = M$. (Q : pourquoi est-ce légitime ?)

TSP (IV)

- Par S/E on vérifie que l'on converge vers la solution $x_{15} = x_{24} = x_{31} = x_{43} = x_{52} = 1$ soit le tour 1-5-2-4-3-1 pour un coût $z = 668$.

Résolution des problèmes binaires par énumération implicite

- Les problèmes binaires (BP) forment une très grande classe de problèmes en IP.
- Remarque : on peut exprimer tout IP en BP (par décomposition sur puissances de 2)
- Il existe des méthodes de S/E pour BP particulières, par exemple la méthode d'énumération implicite : en chaque branche les branches supérieures représentent les variables fixes. On branche sur la variable qui améliore la fonction de coût le plus (comme dans le knapsack), tant que la solution est faisable.
- Même si la solution n'est pas faisable, elle génère un coût d'évaluation.

Sudoku

Comme promis.

Conclusion

- Optimisation par PL associée avec une PI ou PM par S/E donne une méthodologie génerale pour résoudre les PI/PM.
- Un intérêt majeur de S/E est que la convergence est garantie en temps fini, et que toute solution réalisable intermédiaire est associée avec une borne inférieure et supérieure pour le coût optimal.
- Dans certains cas on a intérêt à exhiber une meilleure fonction d'évaluation et de séparation, mais la méthode générale reste valide.
- De la qualité de ces fonctions et de la formulation dépend le résultat.
- Lien avec heuristiques, $A^{\star}$, algorithmes sur les graphes, etc.
- En général les problèmes solubles par S/E sont difficiles, mais parfois ce n'est pas le cas.