

B. ROY

Procédure d'exploration par séparation et évaluation

Revue française d'informatique et de recherche opérationnelle. Série verte, tome 3, n° V1 (1969), p. 61-90

http://www.numdam.org/item?id=RO_1969__3_1_61_0

© AFCET, 1969, tous droits réservés.

L'accès aux archives de la revue « Revue française d'informatique et de recherche opérationnelle. Série verte » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

PROCEDURE D'EXPLORATION PAR SEPARATION ET EVALUATION

(P. S. E. P. et P. S. E. S.)

par B. ROY (1)

Résumé. — *A côté des procédures traditionnelles de recherche directe, rigoureuse et exclusive d'un optimum, la recherche opérationnelle voit depuis quelques années se développer des procédures d'un genre nouveau. Il s'agit des procédures d'exploration. Comme les précédentes elles visent à guider le choix au sein d'un ensemble de décisions réalisables A. Toutefois, la responsabilité de ce choix n'est plus nécessairement confiée à une unique fonction, la souplesse de l'approche permet de s'adapter à une réalité complexe et changeante, enfin les divers éléments de A constituant le résultat sont le fruit d'une exploration de cet ensemble.*

Après avoir situé les procédures d'exploration par rapport aux procédures de pure optimisation (§ 1), nous nous efforçons (§ 2) de mieux cerner une classe importante de procédures d'exploration : celles reposant sur un principe de séparation (« branching » principe) et une application d'évaluation (« bounding » fonction). Le paragraphe 3 traite des Procédures par Séparation et Évaluation Séquentielle (P.S.E.S.) qui constituent un premier type de telles procédures. Le reste de l'article est consacré à un second type : les Procédures par Séparation et Évaluation Progressive (P.S.E.P.).

Les paragraphes 4 et 5 jettent les bases axiomatiques sur lesquelles fonder les P.S.E.P. Le mécanisme de ce type de procédure est ensuite décrit au paragraphe 6 où il est illustré sur un exemple. Enfin le dernier paragraphe indique comment s'y prendre en pratique pour concevoir et adapter une P.S.E.P. en présence de tel ou tel problème particulier.

1. PROCEDURES D'OPTIMISATION ET PROCEDURES D'EXPLORATION

On a pris l'habitude à propos de problèmes qui s'insèrent dans des contextes fort variés de la vie économique et sociale, de définir et de formaliser ce qu'il est maintenant convenu d'appeler « l'ensemble des décisions réalisables » eu égard aux problèmes considérés. D'une façon générale, les éléments d'un tel ensemble se rapportent aussi bien à des actions envisageables qu'à des situations possibles.

Du fait même de leur définition, ces décisions se situent dans un certain univers \mathcal{U} lequel peut être par exemple :

- l'ensemble R^n ,
- l'ensemble Z^n (problèmes en nombres entiers),
- l'ensemble $\mathcal{p}(E)$ des parties d'un ensemble,
- l'ensemble $\pi(E)$ des permutations d'un ensemble E ,
- ...

(1) S.E.M.A.

Dans ce qui suit A désignera le sous-ensemble de \mathcal{U} des *décisions* auxquelles on s'intéresse, il pourra s'agir de la totalité des décisions *réalisables*, ou plus généralement seulement de celles d'entre elles qui présentent un intérêt certain et que l'on qualifiera d'*admissibles*.

Supposons maintenant que l'on se préoccupe de guider un choix devant s'opérer au sein de A . Soit alors γ , l'application :

$$\gamma : A \rightarrow \mathcal{C}$$

par laquelle chaque élément $a \in A$ se voit associer une gamme de conséquences $\gamma(a) = C \in \mathcal{C}$, sur la base de laquelle il doit être apprécié et comparé aux autres.

EXEMPLE 1

Dans un problème d'ordonnancement du type T (cf. B. Roy, B. Sussman [17]), l'inconnue est un élément $(t_0, t_1, \dots, t_{n+1}) \in R^{n+2}$ (qui constitue l'univers \mathcal{U}). L'ensemble A des ordonnancements auxquels on s'intéresse revêt alors des formes variées selon la nature des contraintes que l'on veut introduire (cf. B. Roy [14]). Enfin, les conséquences à prendre en compte peuvent être une durée globale, la tenue de délais partiels, des coûts, des effectifs de main-d'œuvre, des taux d'occupation de matériels...

EXEMPLE 2

Considérons l'univers \mathcal{U} défini comme suit : $\mathcal{U} = \mathcal{P}(Y)$ avec Y ensemble de parties d'un ensemble X , en vue de nous intéresser au sous-ensemble A des éléments de \mathcal{U} qui définissent un recouvrement de X . Les conséquences susceptibles d'intervenir pour guider le choix pourront être, le nombre de parties appartenant à Y entrant dans le recouvrement, le nombre minimum, moyen ou maximum d'éléments de X composant ces parties, un indicateur de chevauchement (cf. B. Roy [15]), ...

EXEMPLE 3

Dans un problème de distribution ou de collecte (cf. Bertier P. [5]) devant conduire à visiter n centres formant un ensemble X , on sera généralement amené à définir un certain nombre de propriétés que devra satisfaire une suite d'éléments de X (avec ou sans répétitions) pour définir un programme de visites admissible. La durée du parcours, son kilométrage, la facilité avec laquelle il se décompose en sous-parcours journaliers, ... sont autant de conséquences susceptibles de guider un choix.

La nature des ensembles \mathcal{U} , A , \mathcal{C} qui joueront un rôle central dans tout ce dernier chapitre, pourrait encore être illustrée grâce à bien

d'autres exemples ayant trait notamment à la recherche de sous-ensembles de sommets remarquables d'un graphe à la conception de réseaux de communication, à des problèmes de circulation, à des problèmes d'investissements (cf. J. M. Audiber, J. C. Holl, J. P. Plas [2]), de localisation (cf. R. Benayoun et B. Roy [4]), à la programmation linéaire continue, entière ou mixte (cf. G. B. Dantzig [8], P. Bertier, Nghiem, B. Roy [7]...).

Souvent, l'introduction d'une fonction économique Φ permet de ramener le problème du choix à celui de la recherche d'un élément de A qui confère à Φ (définie sur A) sa valeur extrémum sur A . Le problème sera théoriquement résolu si l'on connaît une *procédure d'optimisation* permettant d'isoler un tel élément privilégié de A . De telles procédures ont été élaborées en très grand nombre à propos de problèmes particuliers fort divers (cf. B. Roy [15]). Il en existe qui sont relatives à des problèmes plus généraux (voir par exemple M. Simonnard [18], H. P. Kunzi [13], J. Abadie [1], J. Wilde [19]), que nous n'aborderons pas ici.

Les procédures de recherche directe rigoureuses et exclusives d'un optimum apparaissent cependant fréquemment insuffisantes, et cela pour les raisons suivantes :

1° L'optimum n'étant pas *a priori* quelconque dans A (c'est un sommet du polyèdre en programmation linéaire...), la procédure en tire profit, mais elle est inapte en revanche à isoler d'autres éléments (des points intérieurs au polyèdre par exemple) ;

2° La validité de la procédure étant étroitement liée à la nature mathématique de A et Φ , cela interdit toutes sortes (non linéarités, discontinuités, alternatives,...) d'adaptations de l'ensemble et de la fonction économique à une réalité quelque peu complexe et changeante ;

3° L'optimum étant trouvé, l'étude de son voisinage, tenant compte de ce que l'appartenance à A de certains éléments de \mathcal{U} peut être mise en cause, ainsi que de la présence des critères secondaires, n'est pas un prolongement naturel de la procédure.

A côté de ces procédures de pure optimisation, se sont développées d'autres procédures, visant elles aussi à guider le choix au sein de A , mais sans que la responsabilité en soit exclusivement confiée à une fonction unique devant aboutir à la sélection d'un élément théoriquement le meilleur. Ces procédures que l'on peut caractériser par leur aptitude à :

- α) isoler *a priori* n'importe quel élément de A ,
- β) s'adapter à des réalités complexes et changeantes,

γ) permettre l'exploration de certains voisinages, seront englobées sous le terme général de *procédures d'exploration*. Insistons au préalable sur le fait que les aptitudes (α) (β) (γ) qui caractérisent les procédures d'exploration, ne conduisent en aucun cas à les opposer aux recherches d'optimum, qu'elles utilisent en fait largement.

La littérature (cf. bibliographie en fin d'article) incite à distinguer trois grands styles de procédures d'exploration (cf. B. Roy [15], chapitre X) :

- exploration aléatoire,
- exploration dirigée,
- exploration par séparation.

Dans cet article, c'est le dernier qui retiendra seul notre attention. Il s'oppose aux précédents en ce sens qu'il ne conduit pas à engendrer une succession d'éléments de A mais une succession de parties de A que l'on subdivise progressivement : autrement dit, au lieu de procéder de façon « ponctuelle », on procède par « grappes ».

2. GENERALITES SUR LES PROCEDURES D'EXPLORATION PAR SEPARATION ET EVALUATION

Les procédures de ce style reposent sur :

— *un principe de séparation* qui, du fait de la structure de l'univers \mathcal{U} permet d'isoler des sous-ensembles de A de plus en plus « réduits » en vue d'aboutir à des problèmes simples que l'on sait résoudre (cette réduction pouvant conduire par exemple, à un ensemble qui ne contient plus qu'un seul élément) ;

— *des règles d'examen et de sélection* qui, en fonction du ou des critères, servent, après chaque séparation, à éliminer certains des sous-ensembles de A obtenus, et à sélectionner parmi ceux qui restent, celui auquel le principe de séparation doit à nouveau être appliqué.

La relation de filiation qui lie un sous-ensemble à ceux auxquels il donne naissance par séparation est le plus souvent représenté par une arborescence (voir fig. 1).

Il s'agit évidemment de concevoir ces mécanismes d'examen, de sélection, et de séparation de telle sorte que leur fonctionnement itéré aboutisse à une mise en évidence de quelques décisions satisfaisantes, voire optimales. Il est clair que ceci est parfaitement compatible avec les conditions α , β , γ du paragraphe 1.

Voici, afin d'illustrer les généralités qui précèdent, une description très sommaire d'une procédure proposée par HUARD [12] — sous le nom de *méthode des centres* — pour rechercher le minimum d'une fonction ν sur une partie A de R^n .

Cette méthode suppose que, à tout sous-ensemble $B \subset A$, on ait le moyen d'associer un élément $c_B \in B$ occupant, en un certain sens, une position centrale dans B , centre de gravité d'une distribution de masses,

milieu d'un diamètre... Considérons alors la surface d'indifférence définie par les éléments $A \in A$ solution de :

$$v(a) = v(c_B),$$

elle forme une frontière grâce à laquelle B peut en général être séparé en deux sous-ensembles (n'ayant que cette frontière en commun). Soit

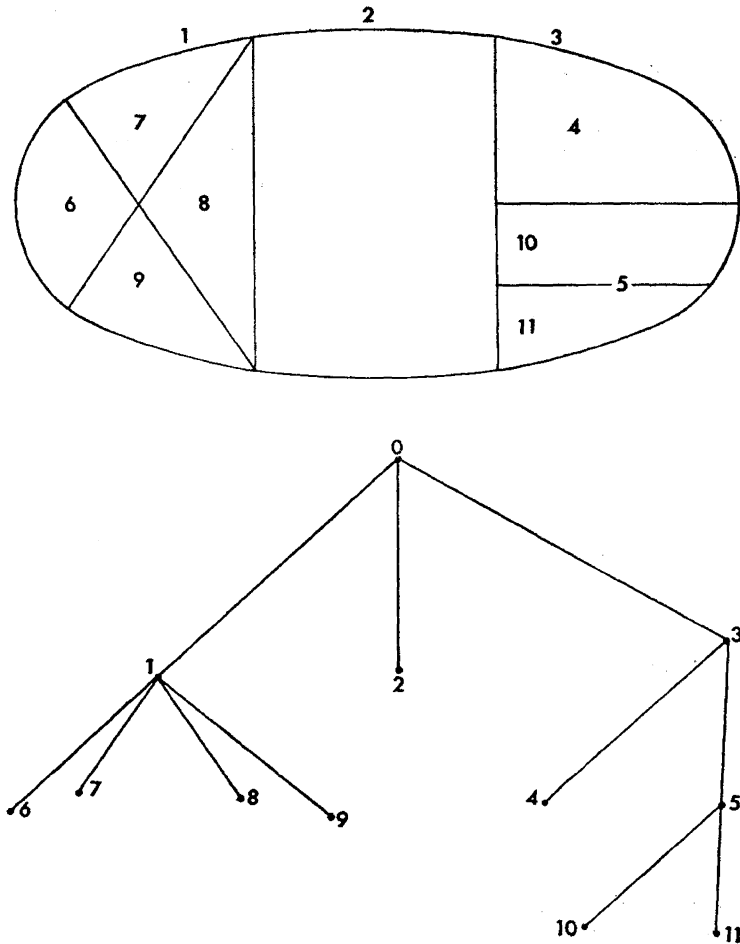


Figure 1

Exemple d'arborescence associée à une exploration par séparation. L'ensemble A , numéroté 0, est ici séparé en trois sous-ensembles numérotés 1, 2, 3. L'ensemble 3, sélectionné, donne naissance aux 2 sous-ensembles 4 et 5... Sachant que les sous-ensembles sont numérotés conformément à l'ordre dans lequel ils apparaissent, il est facile de reconstituer les sélections successives.

en s'appuyant sur le gradient de la fonction ν au centre c_B , soit en calculant la valeur de cette fonction aux centres de chacun des deux sous-ensembles ainsi engendrés, on peut éliminer l'un et par conséquent sélectionner l'autre. L'arborescence associée sera dans ce cas bien particulière (voir fig. 2). Il est clair qu'en procédant de la sorte, moyennant quelques précautions élémentaires, on aboutira à un minimum local de ν sur A .

L'un des principaux mérites de ce style de procédure est de fournir un cadre rigoureux pour aborder des problèmes complexes tels que ceux posés par la recherche d'une typologie (cf. B. Roy [15], chap. X), d'un

ordonnement (cf. B. Roy et B. Sussmann [17]), d'un plan d'investissement (cf. J. M. Audibert, J. C. Holl, J. P. Plas [2]), de circuits de distributions (cf. P. Bertier [6]), ..., cadre grâce auquel on pourra concevoir des procédures heuristiques efficaces sur le plan concret. C'est là un aspect qui sera illustré et précisé tout au long de cet article.

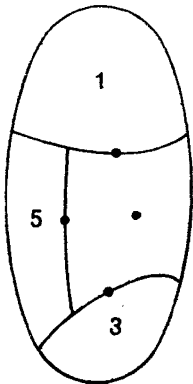


Illustration de la
méthode des centres.

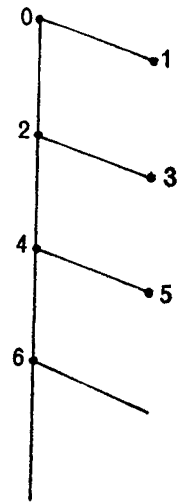


Figure 2

C'est probablement cette aptitude à s'adapter à des réalités complexes qui explique le foisonnement des procédures relevant de ce style. Très nombreuses (voir bibliographie) sont celles rangées sous l'étiquette *Branch and bound* laquelle peut être comprise comme précisant que la sélection et l'élimination repose sur l'introduction de bornes (inférieures ou supérieures). Le terme « branch » se réfère à l'organisation arborescente qui traduit la séparation. Cette acception conduit notamment à ranger

sous cette étiquette ⁽¹⁾ les procédures par séparation et évaluation progressive, lesquelles correspondent à un type bien précis auquel nous consacrerons les quatre derniers paragraphes de cet article.

À côté des P.S.E.P., et bien que très proche, il convient de distinguer un autre type précis de procédure d'exploration par séparation et évaluation et se rangeant lui aussi sous l'étiquette Branch and Bound. Il s'agit du type que nous proposons d'appeler P.S.E.S. : Procédure par Séparation et Evaluation Séquentielle. Nous lui consacrerons le paragraphe suivant. Celui-ci ne constitue toutefois qu'une présentation très sommaire du type P.S.E.S. que le lecteur pourra approfondir en cherchant à transposer et à adapter l'approche beaucoup plus détaillée qu'apportent pour le type P.S.E.P. les paragraphes 4, 5 et 6.

3. ESQUISSE DES PROCEDURES P.S.E.S.

Ce type de procédure repose sur l'introduction :

— d'un principe de séparation auquel on associe une relation d'ordre définie sur tout ensemble de sous-ensembles de A engendrés par une séparation ; il s'ensuit que dans l'arborescence liée à un tel principe de séparation les suivants d'un sommet quelconque se trouvent toujours complètement ordonnés par cet ordre auquel nous donnerons le nom d'ordre transverse ;

— d'une fonction d'évaluation ν définie pour tout sous-ensemble résultant d'une séparation, et conçue pour apprécier par défaut ou par excès — selon que l'on recherche un minimum ou un maximum — ce que valent (compte tenu du ou des critères) les meilleures solutions du sous-ensemble considéré.

Les règles d'examen et de sélection qui définissent alors la procédure sont les suivantes :

— en chaque sommet de l'arborescence on sélectionne le premier suivant dans l'ordre transverse, non encore sélectionné ;

— pour chaque sous-ensemble représenté par un tel sommet sélectionné, on calcule la valeur de l'évaluation ν , et, lorsqu'on en a la possibilité, on exhibe une ou plusieurs des meilleures solutions que renferme ce sous-ensemble, à moins qu'on ne prouve qu'il est vide ; de telles solutions doivent nécessairement être révélées par tous ceux des sommets terminaux de l'arborescence qui correspondent à des sous-ensembles non vides ;

— si l'examen précédent permet de prouver que le sous-ensemble considéré ne renferme pas de solutions optimales ou « satisfaisantes » (soit parce qu'il est vide, soit parce que l'évaluation prouve que les meilleures solutions qu'ils renferment sont moins bonnes que d'autres

(1) La méthode des centres qui vient d'être présentée fournit un exemple de procédure de séparation ne se rangeant pas sous cette étiquette : l'élimination et la sélection, reposent sur une fonction d'évaluation mais celle-ci ne conduit pas forcément à une borne supérieure ou inférieure.

déjà exhibées), il est inutile de le séparer et la sous-arborescence ayant pour racine le sommet qui le représente n'a pas à être construite ; il en est de même lorsque le problème relatif au sous-ensemble considéré a pu être résolu, dans ce cas, il convient toutefois de conserver les solutions exhibées ; dans ces deux cas, on retourne au dernier sommet de l'arborescence qui possède encore au moins un suivant non sélectionné (le recours à une pile ⁽¹⁾ facilite cette mémorisation ;

— dans le cas contraire on procède à la séparation du sommet considéré et l'on poursuit avec le premier, dans l'ordre transverse, des sous-ensembles ainsi engendrés (conformément à ce qui a été dit plus haut).

Ces règles d'examen et de sélection conduisent à décrire l'arborescence (aux sous-arborescences non décrites près) conformément à une séquence préétablie entièrement déterminée par l'ordre transverse. Ce type de procédure présente donc un caractère séquentiel : à des lacunes près, les sommets se succèdent conformément à un ordre intrinsèquement lié au principe de séparation. C'est pourquoi nous nous référons à ce type de procédure sous le sigle P.S.E.S. (Procédure par Séparation et Évaluation Séquentielle). C'est seulement par ce caractère séquentiel que P.S.E.S. se différencie de P.S.E.P. étudié ci-après. Dans ce dernier type de procédure en effet, l'arborescence est décrite du fait d'une progression gouvernée à la fois par la séparation et par l'évaluation..

Cette différence d'enchaînement mise à part, P.S.E.S. et P.S.E.P. présentent une grande similitude, et bon nombre des considérations des paragraphes suivants pourront être aisément transposés à P.S.E.S. par le lecteur. Précisons simplement ici, que, d'une façon générale :

— P.S.E.S., du fait de son caractère séquentiel, est assez facile (plus facile souvent que P.S.E.P.) à programmer sur calculateur électronique ;

— P.S.E.S., pour être efficace, doit reposer sur un ordre transverse qui conduira à exhiber assez tôt les solutions les meilleures, et sur une fonction d'évaluation qui interrompra rapidement l'exploration une fois les meilleures solutions exhibées ; on retrouvera à propos de P.S.E.P. des exigences similaires ;

— P.S.E.S., du fait de l'exigence précédente relative à l'ordre transverse, paraît mieux adaptée à des problèmes clairement et fortement structurés sur le plan théorique (tels ceux étudiés dans B. Roy, [15], chapitre VI, où l'on trouvera plusieurs exemples de P.S.E.S.) qu'à des problèmes résultant d'un empilement de conditions assez disparates (tels ceux résultant de réalités complexes : investissement, localisation, conception de réseau, ordonnancement, distribution...), auxquels P.S.E.S. sera souvent mieux adaptée parce que plus souple.

Mentionnons enfin qu'un type de procédure d'exploration par séparation, extrêmement général, a été proposé par P. Hervé dans [11].

(1) Ce qui explique l'expression LIFO — Last In First Out — qu'emploient les Anglo-Saxons pour désigner cette règle.

4. LES PROCEDURES P.S.E.P. : CADRE GENERAL

Soient toujours \mathcal{U} un univers et A un sous-ensemble dont un élément quelconque a s'interprète comme une décision admissible. Supposons maintenant que pour comparer de telles décisions, on dispose :

— d'une application Φ dite critère principal :

$$\Phi : A \rightarrow F$$

F étant un ensemble doté d'un ordre complet tel que par définition :

$\Phi(a) < \Phi(a')$ signifie que le critère principal Φ conduit à préférer a à a' ;

$\Phi(a) < \Phi(a')$ signifie que, du point de vue du critère principal a et a' sont des décisions équivalentes.

— éventuellement d'autres applications ψ_1, ψ_2, \dots de A dans le même ensemble ordonné F , celles-ci s'interprétant comme des critères secondaires relatifs chacun à un point de vue particulier que l'on ne peut, ou ne veut, pas égréger au critère principal Φ ; nous supposerons que l'ordre défini sur F traduit la préférence pour les critères secondaires avec les mêmes conventions de sens que pour le critère principal.

REMARQUES

1° Les applications Φ et ψ ne sont pas nécessairement des fonctions (économiques). Cela implique en particulier que même si F est une partie de l'ensemble R des nombres réels, ordonné comme le sont usuellement ces nombres, on ne suppose pas pour autant que les opérations d'addition et de multiplication soient susceptibles d'une interprétation quelconque. C'est ainsi que F peut être une simple échelle.

2° Toutefois la procédure peut être améliorée lorsque, a étant, d'un certain point de vue Φ , préféré à a' , cela a un sens de dire que cette préférence dépasse significativement ou non un certain seuil. Théoriquement ceci revient à admettre que l'on peut définir une fonction δ :

$$\delta : F \times F \rightarrow B = \{0, 1\}$$

avec :

$\delta(f, f') = 0$ si $f = f'$ ou si $f < f'$ avec un « écart de préférence » jugé inférieur au seuil ;

$\delta(f, f') = 1$ si non.

Il convient également de faire observer qu'aucune perte de généralité n'est entraînée par l'identité des ensembles cibles pour les divers critères : rien ne s'oppose en effet à ce que $\Phi(A), \psi_1(A), \dots$ soient des parties disjointes d'un même ensemble F fabriqué artificiellement pour la seule commodité du symbolisme.

Notre propos est ici de définir une classe de procédure d'exploration par séparation qui soit applicable à de nombreux modèles faisant intervenir un ensemble de décisions admissibles A plongé dans un univers \mathcal{U} , un critère principal Φ et éventuellement des critères secondaires ψ_1, ψ_2, \dots

Un tel modèle $(\mathcal{U}, A, \Phi, \psi)$ étant défini, le type de problème que concerne la procédure que nous allons étudier peut revêtir l'une des deux versions suivantes :

Version α : Isoler une ou plusieurs décisions de A qui soit *optimale* pour le critère principal Φ .

Version β : Isoler une ou plusieurs décisions de A qui soit *satisfaisante* pour le ou les critères considérés. Une décision a sera qualifiée de *satisfaisante pour un critère Φ* si $\Phi(a)$ est jugé suffisamment proche de l'optimum Φ^* de Φ sur A ; autrement dit, si δ étant une fonction seuil définie on a :

$$\delta[\Phi^*, \Phi(a)] = 0$$

Une décision a sera dite *satisfaisante pour un critère principal Φ et des critères secondaires ψ_1, ψ_2, \dots* si les valeurs $\psi_1(a), \psi_2(a), \dots$ sont jugées suffisamment bonnes et si $\Phi(a)$ est suffisamment proche de l'optimum Φ^* qui peut atteindre Φ sur A sous les contraintes additionnelles qui imposent à chacun des critères secondaires d'atteindre respectivement les niveaux $\psi_1(a), \psi_2(a), \dots$

Les procédures de cette classe se reconnaissent en ce que, d'une part, le principe de séparation satisfait aux trois axiomes énoncés ci-après en 5a, d'autre part, les règles d'examen et de sélection reposent sur une application d'évaluation satisfaisant à son tour deux nouveaux axiomes 5b (cf. b). Étant donné un problème $(\mathcal{U}, A, \Phi, \psi)$, ces axiomes laissent une grande liberté pour choisir le principe de séparation et l'application d'évaluation. Ce choix étant fait, les consignes opératoires décrites au 6a ci-après achèvent de définir la procédure : il suffit en effet de s'y conformer pour atteindre l'un ou l'autre des deux types de solutions que l'on a qualifiées d'optimales et de satisfaisantes. Nous terminerons ce paragraphe en indiquant comment une procédure conçue en conformité avec ce schéma rigoureux, peut être aménagée en une procédure heuristique efficace pour des problèmes complexes.

Achevons ce paragraphe par quelques définitions et notations essentielles pour la suite.

Les sous-ensembles de A engendrés par les séparations progressives seront notées V_q : il est commode de convenir que l'indice q traduit l'ordre dans lequel ces sous-ensembles apparaissent (voir fig. 1). Il s'ensuit que :

$$V_0 = A.$$

Soient V_0, V_1, \dots, V_p une famille de sous-ensembles engendrés en séparant V_0 , puis l'un quelconque des sous-ensembles ainsi obtenus, puis un autre, et cela après un nombre fini de séparations.

Considérons un entier q compris entre 0 et p , et tel que V_q n'ait pas encore donné lieu à séparation (voir fig. 3). Supposons que V_q soit, à ce stade, sélectionné pour être séparé. Le résultat de cette opération sera :

— soit une suite finie, non vide et non réduite à un seul élément, définie de façon unique et sans ambiguïté, de sous-ensembles de V_q , ceux-ci étant notés V_{p+1}, V_{p+2}, \dots ; nous dirons qu'ils résultent de la séparation, ou encore que ce sont les suivants de V_q ; nous noterons $S(q)$ le groupement d'indices qui sert à les repérer ($|S(q)| \geq 2$) ;

— soit V_q lui-même si le principe de séparation conduit à ne pas séparer un tel ensemble : celui-ci sera alors qualifié de terminal et l'on posera $S(q) = \emptyset$;

précisons que ce résultat ne dépend que de V_q et qu'il ne saurait en aucune manière être influencé par les autres ensembles déjà engendrés.

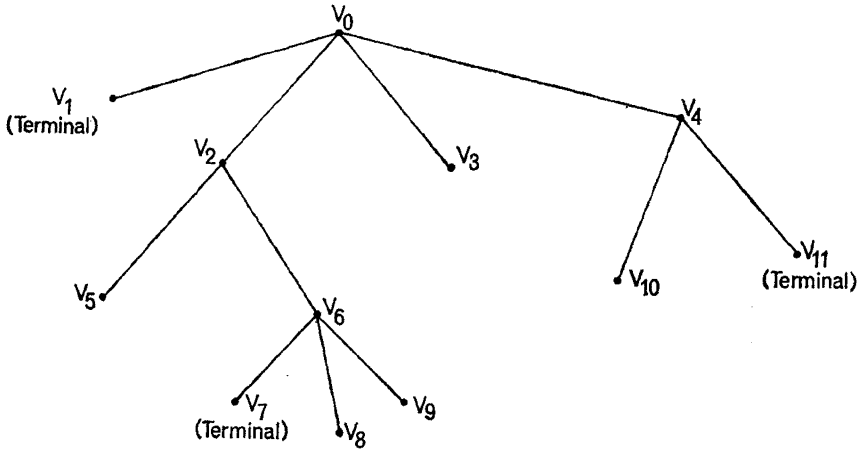


Figure 3

Sous-arborescence relative $\alpha : V_0, V_1, \dots, V_{11}$ susceptible d'être poursuivie seulement à partir de : $V_3, V_5, V_8, V_9, V_{10}$.

Si l'on poursuit ainsi la séparation jusqu'à ce que tous les ensemble^s non séparés soient terminaux, on aboutit à une suite de parties de A : (non nécessairement toutes distinctes)

$$V_0, V_1, \dots, V_q, \dots$$

sur laquelle est définie une relation de filiation S dont le graphe est une arborescence de racine $V_0 = A$. Il est clair que cette arborescence est indépendante de l'ordre dans lequel les choix successifs ont introduit les différentes séparations (cet ordre n'influençant que la numérotation des sous-ensembles).

5. AXIOMES DE BASE DES P.S.E.P.

a) Axiomes relatifs au principe de séparation

Axiome de finitude :

La suite maximale de parties $V_0, V_1, \dots, V_q, \dots$ qu'il est possible d'engendrer par séparation, est finie : Q désignera la valeur maximum de q .

Axiome de conservation :

Quel que soit

$$q = 0, 1, \dots, Q : \bigcup_{r \in S(q)} V_r = V_q$$

Mentionnons qu'il est souvent préférable mais non indispensable que $\{V_r / r \in S(q)\}$ constitue une partition de V_q ; il est toujours intéressant d'examiner si l'on peut ou non avoir $V_r = V_q (r \in S(q))$. Si une telle égalité est impossible, et si A est fini, alors l'axiome de finitude découle de l'axiome de conservation.

Axiome d'arrêt :

Quel que soit le sous-ensemble terminal V_q , on sait :

— soit prouver que $V_q = \emptyset$ (l'existence de tels sous-ensembles terminaux vides ne pouvant, en général, pas être exclue) ;

— soit trouver une solution $a_0(q) \in V_q$ optimale (resp. satisfaisante) pour le problème posé modifié en substituant V_q à A ; dans la version β $a_0(q)$ pourra ne pas être optimale, mais seulement suffisante relativement à V_q .

Voici tout d'abord quelques idées très générales sur lesquelles fonder un tel principe de séparation selon la structure de \mathcal{U} :

- le signe ou la valeur prise par une certaine fonction si $\mathcal{U} = R^n$;
- les diverses valeurs possibles pour une variable discrète si $\mathcal{U} = Z^m \times R^n$;
- la présence ou l'absence de tel élément si $\mathcal{U} = \mathcal{E}(E)$;
- l'ordre relatif de deux éléments si $\mathcal{U} = \pi(E)$.

Illustrons maintenant sur un exemple précis — lequel nous servira de référence tout au long de ce paragraphe et du suivant — comment on peut concevoir un principe de séparation conforme à ces trois axiomes.

EXEMPLE 4. — PROBLEME DE LA PLUS COURTE CHAINE HAMILTONIENNE

Considérons un graphe $G = (X, U)$ à chaque arête u duquel est associé un nombre positif $c(u)$ représentatif d'une longueur (voir à titre

d'exemple le graphe de la figure 5) (1). La longueur d'une chaîne étant définie comme la somme des longueurs des arêtes qui la compose, on cherche à élaborer une procédure permettant de construire la ou les plus courtes chaînes hamiltoniennes de G .

Rappelons qu'une chaîne est une suite de sommets dans laquelle deux sommets consécutifs sont toujours reliés par une arête. Pour être hamiltonienne, la chaîne doit passer une fois et une seule par chacun des sommets du graphe. Une telle chaîne peut également être définie comme une suite d'arêtes (une de moins qu'il y a de sommets dans le graphe).

L'univers \mathcal{U} peut être défini ici comme étant formé de toutes les parties de U comprenant $n - 1$ arêtes (n désignant le nombre des sommets de X); quant à l'ensemble A des « décisions admissibles » il est constitué par celles de ces parties qui définissent une chaîne hamiltonienne. Enfin, le critère principal est tout simplement défini par :

$$\Phi(a) = \sum_{u \in a} c(u) \quad (a \in A)$$

(F étant ici l'ensemble \mathbb{R}^+ des réels positifs).

Précisons encore que, tel qu'il vient d'être formulé, le problème correspond à la version α du problème général, et ne fait pas intervenir de critère secondaire.

Supposons les arêtes de U rangées selon un certain ordre (que l'on précisera plus loin) traduit par leur numérotage u_1, u_2, \dots, u_m . L'univers \mathcal{U} est alors doté d'une structure naturelle qui conduit à distinguer :

- celles des parties de U (que sont les éléments de \mathcal{U}) qui contiennent l'arête u_1 , et celles qui ne la contiennent pas ;
- chacun des deux sous-ensembles ainsi défini peut à son tour être séparé en deux autres à partir de la seconde arête u_2 ;
- et ainsi de suite jusqu'à la prise en considération de l'arête u_m .

Il y a là un principe de séparation extrêmement sommaire, valable non seulement pour \mathcal{U} mais aussi pour A . Il est clair qu'un tel principe satisfait aux deux axiomes de finitude et de conservation. Il mérite cependant d'être affiné afin d'éviter les nombreuses séparations inutiles auquel il conduit (dans A plus encore que dans \mathcal{U} , ce qui aura pour effet de réduire le nombre Q).

Remarquons à cet effet qu'une suite de $n - 1$ arêtes constitue un élément $a \in A$ si et seulement si (2) le sous-graphe $G(a)$ (3) qu'elle définit possède les deux propriétés suivantes :

(1) Rappelons que les sommets du graphe dont l'ensemble est noté X correspondent aux points marqués x_1, x_2, \dots, x_6 sur la figure ; quant aux arêtes dont l'ensemble est noté U , ce sont les lignes qui joignent ces points (dans l'exemple considéré, deux sommets quelconques sont toujours reliés par une arête).

(2) Graphe déduit de G par suppression de toutes les arêtes qui n'appartiennent pas à a .

(3) Il découle en effet de (P_1) que $G(a)$ est un arbre partiel de G , et du fait de (P_2) cet arbre ne peut être qu'une chaîne, laquelle est forcément hamiltonienne.

(P_1) : $G(a)$ sans cycle (un cycle est une chaîne dont le dernier sommet coïncide avec le premier).

(P_2) : n'avoir aucun sommet de degré supérieur à 2 partiel dans $G(a)$ (le degré d'un sommet est le nombre d'arêtes qu'il admet pour extrémité).

Remarquons encore qu'étant donné un sous-ensemble a^h de U formé de h arêtes bien définies, il n'existe des éléments de A contenant a^h que si le graphe $G(a^h)$ défini par ces h arêtes possède les propriétés P_1 et P_2 . Il est de ce fait très facile de mettre en évidence l'inutilité de certaines des dichotomies envisagées plus haut lorsque l'acceptation d'une arête s'avère incompatible avec les acceptations antérieures.

Désignons par k un entier quelconque au plus égal à $m - 1$, ($m = |U|$), et par a^h un sous-ensemble de h ($h < n - 1$) arêtes prises parmi les k premières. Soit alors $V(k, a^h)$ le sous-ensemble de A défini par :

a) ($V(k \cdot a^h)$) si et seulement si :

$$a \in A, \quad a \supset a^h, \quad a \cap (\{u_1, \dots, r_k\} - a^h) = \Phi.$$

Il est clair que le principe de séparation envisagé plus haut, fondé sur une décision d'acceptation ou de rejet de chaque arête, décisions prises successivement conformément à l'ordre de rangement des arêtes, engendre nécessairement des sous-ensembles V de la forme $V(k, a^h)$: k est le nombre des arêtes qui ont donné lieu à décision — ou comme nous dirons désormais pour éviter toute confusion, à *arbitrage*, a^h correspond à celles des arêtes acceptées, — ou comme nous dirons encore, *arbitrées positivement*.

Nous sommes maintenant en mesure de progresser dans notre définition d'un principe de séparation. L'arête u_{k+1} ne devra jouer un rôle dans la séparation d'un ensemble $V(k, a^h)$ que si elle est compatible avec les choix déjà faits, c'est-à-dire, si son adjonction à a^h n'est pas *a priori* incompatible avec les propriétés P_1 et P_2 (sinon, on est en droit de rejeter u_{k+1} sans pour autant compromettre l'axiome de conservation). Soit donc $u_{k+\varepsilon}$ la première arête postérieure à u_k telle que :

$$a^{h+1} = a^h \cup \{u_{k+\varepsilon}\}$$

définisse un graphe $G(a^{h+1})$ jouissant des propriétés P_1 et P_2 . On pourra adopter le principe de séparation suivant (toujours conforme aux deux premiers axiomes) :

tout ensemble $V(k, a_k)$ qui ne sera pas décrété terminal (voir ci-après) donnera naissance, par séparation, aux deux sous-ensembles $V(k + \varepsilon, a^h)$ et $V(k + \varepsilon, a^{h+1})$ correspondant respectivement au rejet et à l'acceptation de l'arête $u_{k+\varepsilon}$.

Il ne reste plus maintenant qu'à définir les conditions dans lesquelles un ensemble $V(k, a^h)$ pourra être décrété terminal. Il nous faut pour cela prendre appui sur l'axiome d'arrêt.

Observons tout d'abord qu'un ensemble $V(k, a^h)$ tel que :

$$h + m - (k + \varepsilon - 1) < n - 1$$

est nécessairement vide (puisque même en acceptant toutes les arêtes $u_{k+\varepsilon}, \dots, u_m$ il sera impossible d'en accepter en tout $n - 1$). Un tel ensemble peut donc toujours être décrété terminal vide.

Observons ensuite que, si l'ordre des arêtes a été choisi conforme aux valeurs croissantes de leurs longueurs, alors, chaque fois que les arêtes $u_{k+\varepsilon}, u_{k+\varepsilon+1}, \dots, u_{k+\varepsilon+n-2-h}$ formeront avec a_h une chaîne hamiltonienne, celle-ci sera de longueur minimum parmi celles que renferme $V(k, a^h)$. Un tel ensemble pourra donc être décrété terminal : la solution a_0 dont l'axiome postule l'existence étant précisément définie par a^h et les $n - 1 - h$ premières arêtes suivant $u_{k+\varepsilon-1}$.

Le lecteur vérifiera aisément que tout ensemble $V(k, a^h)$ qui n'est pas terminal, en ce sens qu'il ne satisfait aucune des deux conditions qui précèdent, peut être séparé sans ambiguïté, et l'on a bien défini un principe de séparation conforme aux trois axiomes.

Comme le raisonnement qui nous a conduit à ce principe l'a laissé entrevoir à plusieurs reprises, bien d'autres variantes pourraient être imaginées à partir de la même idée de base que constitue la structure de \mathcal{U} décrite au début de l'exemple.

b) Axiomes relatifs à l'application d'évaluation

Cette application ν a pour objet d'associer, en connexion avec le critère principal Φ , un élément $\nu(q) \in F$ à chacun des V_q engendrés au cours de la séparation, et cela, précisément en vue de guider celle-ci. Rappelons qu'optimum signifie ici minimum du fait des conventions relatives au sens de la préférence, posées au début du paragraphe 4. Pour pouvoir jouer ce rôle (cf. b.6 ci-après), l'application ν doit satisfaire aux deux axiomes suivants.

Axiome de minoration

$$\nu : \{ V_q / q = 1, \dots, Q \} \rightarrow F$$

vérifie :

$$\nu(q) \leq \text{Min}_{a \in V_q} \Phi(a) \quad (1)$$

quel que soit $q = 1, \dots, Q$.

Axiome de coïncidence

Si V_q est un sous-ensemble terminal non vide, alors :

- $\nu(q) = \Phi[a_0(q)]$ dans la version α (ou β) ;
- $\delta[\nu(q), \Phi[a_0(q)]] = 0$ dans la version β ;

(1) Ce qui suppose que $\{ \Phi(a) / a \in V_q \}$ est une partie de F qui admet un élément minimum c'est là une hypothèse nullement gênante dans les problèmes qui nous occupent ; on pourrait l'affaiblir en substituant Inf à Min.

seulement $(\delta[f, f'] = 0$ signifiant que « l'écart » entre f et f' est jugé négligeable) ⁽¹⁾.

Le plus souvent, cette application d'évaluation résultera de la résolution d'un problème déduit de celui de la minimation de Φ sur V_q , soit en simplifiant Φ , soit en plongeant V_q dans un certain sous-ensemble de \mathcal{U} (cf. E. Balas [3]).

Tel que le principe de séparation a été défini, l'application — ici on peut dire la fonction — d'évaluation s'impose de façon à peu près évidente. Posons :

$$V_q = V(k, a^h)$$

et continuons à désigner par $u_{k+\varepsilon}$ la première arête susceptible de donner lieu à arbitrage en vue de séparer $V(k, a^h)$. Le lecteur vérifiera sans peine que :

$$v(q) = c(a^h) + c(u_{k+\varepsilon}) + c(u_{k+\varepsilon+1}) + \dots + c(u_{k+\varepsilon+n-2-h})$$

(ou $c(a^h)$ représente la somme des longueurs des arêtes formant a^h) satisfait au deux axiomes de minoration et de coïncidence.

Dans ce qui suit, nous noterons $\omega_1, \omega_2, \dots$ des applications d'évaluation répondant aux deux axiomes précédents, mais relativement aux critères secondaires ψ_1, ψ_2, \dots

6. CONSIGNES OPERATOIRES

Supposons maintenant choisis le principe de séparation et l'application d'évaluation, en conformité avec les axiomes. La procédure consiste en une suite d'itérations, constituées chacune de trois phases (décrites ci-après en 1^o, 2^o, 3^o).

La première d'entre elles consiste dans le cadre de la $p^{\text{ième}}$ itération à sélectionner, parmi un groupe Z^p de sous-ensembles dits candidats, celui qui doit donner lieu à la $(p+1)^{\text{ième}}$ séparation, laquelle conduit à des calculs faisant l'objet de la seconde phase, les sous-ensembles ainsi engendrés sont enfin examinés dans la troisième phase ; à l'issue de laquelle un nouvel ensemble de candidats Z^{p+1} se trouve défini. Cette répétition s'opère aussi longtemps que les conditions d'arrêt décrites en 4^o ne sont pas remplies.

Après cette description de la procédure, on en développera une application numérique. Nous n'établirons pas ici la validité de cette procédure ; le lecteur la trouvera dans B. Roy [15], chapitre X.

a) Description de la procédure

Elle débute en posant :

$$Z^0 = V_0 = A$$

(1) Voir seconde remarque au début de la section.

ce qui définit l'ensemble Z^0 des candidats au début de l'itération initiale $P_1 = 0$. L'application successive des trois phases ci-après conduira à définir Z^1 point de départ de l'itération suivante $P = 1$ et ainsi de suite.

1° PHASE DE SÉLECTION

Pour $p = 0$, on sélectionne l'unique élément V_0 de Z^0 , et l'on passe à la phase de séparation et évaluation en posant :

$$z_0^0 = V_0.$$

Pour $p = 1, 2, \dots$, à chaque sous-ensemble $V_q \in Z^p$ se trouve associé l'élément $\nu(q) \in F$ calculé antérieurement au cours d'une phase de séparation et évaluation. De cette application de Z^p dans F , découle, par transfert de l'ordre complet de F , un préordre complet sur Z^p .

Désignons par :

— z_1^p le (ou l'un quelconque des) premier(s) élément(s) de Z^p , ainsi par définition :

$$\nu[z_1^p] \leq \nu[z] \quad \text{quel que soit} \quad z \in Z^p$$

(si $z = V_q$, il est commode de poser $\nu(q) = \nu[z]$;

— z_0^{p-1} l'élément de Z^{p-1} qui a été sélectionné au cours de l'itération $p - 1$ et qui a par conséquent donné lieu à la $p^{\text{ème}}$ séparation ; il découle de la phase d'examen que cet élément ne peut appartenir à Z^p , où se trouvent par contre certains de ses suivants (s'il en a) ;

— z_1^{*p} le (ou l'un quelconque des) premier(s) suivant(s) de z_0^{p-1} appartenant à Z^p , il vient par définition (si z_1^{*p} est défini) :

$$\nu[z_1^{*p}] \leq \nu[z] \quad \text{quel que soit} \quad z \in S[z_0^{p-1}] \cap Z^p \quad (1)$$

(avec la même commodité d'écriture pour S que pour ν).

Il se peut naturellement que l'on ait $z_1^p = z_1^{*p}$.

Afin d'inciter à définir z_0^p parmi les suivants de z_0^{p-1} , on aura souvent intérêt (cf. § 7 ci-après) à introduire une fonction seuil δ^* permettant d'apprécier s'il convient ou non de tenir compte de l'écart qui existe entre les évaluations de deux éléments quelconques de Z^p ; cette fonction seuil δ^* peut ou non être confondue avec la fonction δ introduite dans l'axiome de coïncidences.

On pourra alors sélectionner z_0^p conformément aux consignes suivantes :

$$\begin{array}{ll} z_0^p = z_1^p & \text{si} \quad \delta^*(\nu[z_1^p], \nu[z_1^{*p}]) = 1 \quad \text{ou si } z_1^{*p} \text{ n'est pas défini} \\ z_0^p = z_1^{*p} & \text{si} \quad \delta^*(\nu[z_1^p], \nu[z_1^{*p}]) = 0. \end{array}$$

(1) La limitation aux seuls éléments de Z^p sera justifiée plus loin par les possibilités d'élimination à la phase d'examen.

Il sera souvent plus simple, mais parfois aussi moins efficace de sélectionner z_0^p en posant tout simplement :

$$z_0^p = z_1^p.$$

Cette sélection étant faite (d'une manière ou d'une autre), on passera à la phase de calcul et d'évaluation

2° PHASE DE SÉPARATION ET ÉVALUATION

Appliquer le principe de séparation à z_0^p afin de définir ses suivants éventuels ainsi que leur évaluation dans F : on se trouve nécessairement dans l'un des trois cas décrits ci-après (du fait de l'axiome d'arrêt).

— z_0^p est un sous-ensemble terminal dont on prouve qu'il est vide : revenir alors à la phase de sélection avec :

$$Z^{p+1} = Z^p - \{z_0^p\};$$

— z_0^p est un sous-ensemble terminal non vide au sein duquel on sait déterminer une solution $a_0[z_0^p] = a_0^p$ optimale ou satisfaisante pour le problème modifié en substituant z_0^p à A : selon le but que l'on s'est proposé, la procédure pourra alors être arrêtée, la solution z_0^p étant solution du problème initial, ou au contraire poursuivie en vue d'engendrer d'autres solutions ; des consignes précises à ce sujet sont données plus loin en 4° ;

— a_0^p n'est pas terminal et l'on sait définir les éléments de $S(z_0^p)$, ainsi que pour chacun d'eux son image dans F par l'application ν : passer dans ce cas à la phase d'examen.

3° PHASE D'EXAMEN

Examiner un à un les suivants de z_0^p afin d'éliminer ceux qui correspondent à des sous-ensembles de A dont on peut prouver qu'ils ne renferment aucune solution susceptible d'être retenue comme satisfaisante ou optimale. On est en particulier en droit d'éliminer un sous-ensemble dès l'instant où :

- on a la preuve qu'il est vide ;
- des applications d'évaluation w_1, w_2, \dots relatives aux critères secondaires ψ_1, ψ_2, \dots font apparaître pour ces derniers des bornes inférieures rendant inacceptables les décisions que renferme le sous-ensemble considéré ;
- l'on connaît une solution a_0 jugée acceptable et telle que $\Phi(a_0)$ soit inférieure à l'évaluation par ν du sous-ensemble considéré, toutefois, cette règle peut devoir être nuancée.

$\hat{S}[z_0^p]$ désignant le sous-ensemble des suivants de z_0^p non éliminés par ces règles (ou par d'autres), revenir à la phase de sélection, en posant :

$$Z^{p+1} = Z^p - (z_0^p) \cup \hat{S}[z_0^p].$$

Lorsque la phase de séparation et évaluation précédente a (comme sous-produit de l'évaluation par exemple) permis de mettre en évidence un élément a_0 de A , jugé préférable à tous ceux antérieurement exhibés, on aura généralement intérêt à définir Z^{p+1} en éliminant du second nombre ci-dessus, tous les candidats antérieurs dont l'évaluation dépasse $\Phi(a_0)$; toutefois, quelques précautions sont à prendre dans le cas où l'on recherche non pas une seule, mais plusieurs solutions satisfaisantes.

4° CONDITIONS D'ARRÊT

Dans la classe de procédures décrite ci-dessus, dont l'organisation des consignes opératoires est schématisée (fig. 6), l'arrêt ne peut être que la conséquence de la sélection d'un sous-ensemble terminal non vide. Une telle sélection ne peut manquer de se produire au bout d'un nombre fini d'itérations en raison de l'axiome de finitude.

Soit donc z_0^p un tel sous-ensemble terminal non vide a_0^p et la solution optimale (resp. satisfaisante) qui a été obtenue à propos d'un problème déduit du problème initial par réduction de A à z_0^p .

Au cours de $p - 1$ itérations qui précèdent cette sélection, un certain nombre de sous-ensemble de A ont pu être reconnu comme sans intérêt et éliminés à l'issue d'une phase d'examen : A^p désignera désormais ce qui reste de A au début de l'itération P du fait des éliminations justifiées par les itérations antérieures. Il est clair que toute solution optimale (resp. satisfaisante) pour le problème déduit du problème initial par réduction de A à A^p est également optimale (resp. satisfaisante) pour ce problème initial. Or, on peut démontrer que :

— si $z_0^p = z_1^p$, alors a_0^p est optimale (resp. satisfaisante) relativement à A^p , donc aussi relativement à A ;

— si $z_0^p = z_1^{*p}$ alors a_0^p présente un écart que l'on sait borner supérieurement (du point de vue du critère principal) vis-à-vis d'une solution optimale relativement à A^p , donc aussi relativement à A .

ORGANISATION DES CONSIGNES OPÉRATOIRES D'UNE P. S. E. P.

Dans la mesure où l'on souhaite exhiber non pas une solution mais plusieurs (notamment en vue de les diversifier d'après les critères secondaires), il suffit de poursuivre en éliminant cette fois a_0^p de Z_p tout en conservant trace de z_0^p et en adaptant convenablement, pour les itérations ultérieures, la phase d'examen au but qu'on se propose (voir l'exemple ci-après ainsi que le paragraphe 7).

b) Exemple d'application de la procédure

Revenons à l'exemple 4, et examinons pas à pas, à propos du graphe de la figure 5, le déroulement des calculs qui permettent d'obtenir une plus courte chaîne hamiltonienne, le principe de séparation et la fonction d'évaluation étant ceux définis aux 5 a et 5 b (dont on conservera les notations).

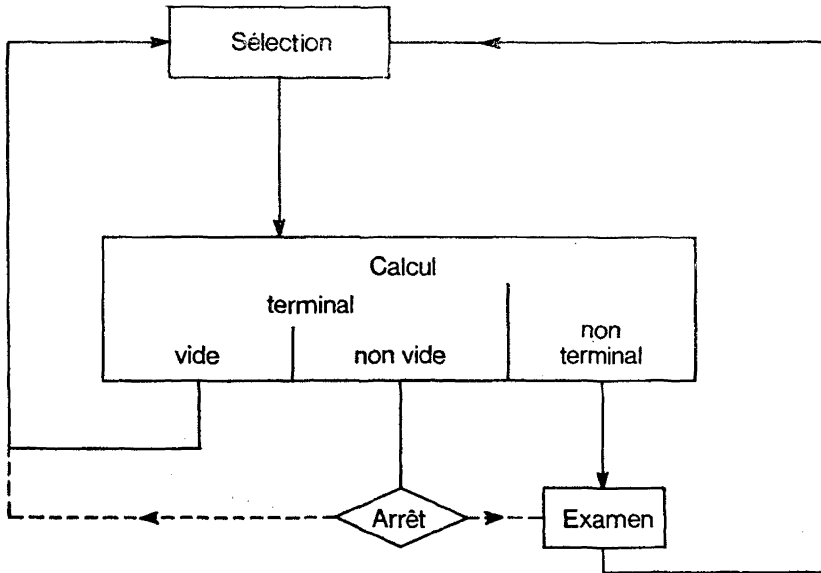


Figure 4

Les 2 flèches en pointillés signifient que dans le cas où l'on veut poursuivre, il faut modifier certaines des consignes des phases de sélection et d'examen.

Dans ce second cas ($z^{p_0} = z^{*1^p}$) on pourra toujours, en vue d'améliorer a^{p_0} , poursuivre la procédure en supprimant, à partir de l'itération p , la possibilité de sélectionner d'après δ^* (on aura intérêt à tenir compte, dans les phases d'examen des itérations ultérieures, de la connaissance de a^{p_0}).

La procédure débute par la sélection de :

$$V_0 = V(0, \emptyset)$$

que la phase de séparation et évaluation conduit à séparer en deux sous-ensembles sur la base de l'acceptation et du rejet de l'arête (4, 5) (première arête dans l'ordre croissant des longueurs : voir tableau situé à droite de la figure 6). Les deux sous-ensembles ainsi engendrés sont respectivement :

$$V_1 = V(1, \{(4, 5)\})$$

$$V_2 = V(1, \emptyset);$$

ils sont représentés (figure 6) par les sommets numérotés 1 et 2, les indications (4, 5) et $\bar{(4, 5)}$ qui accompagnent respectivement l'un et l'autre des arcs aboutissant à ces sommets symbolisent le contenu et la nature de l'arbitrage (positif et négatif) qui caractérise le dit sommet.

Pour achever cette phase de séparation et évaluation, il faut chercher pour chacun des deux sommets quelle est la première arête après (4, 5) qui peut donner lieu à arbitrage. Dans l'un et l'autre cas, il est évident que c'est l'arête (1, 5). Il s'ensuit que :

$$\nu(1) = 6 + 7 + 8 + 9 + 11 = 41$$

$$\nu(2) = 7 + 8 + 9 + 11 + 13 = 48$$

(ces valeurs de la fonction d'évaluation sont portées sur la figure à côté des sommets correspondants). Ni l'un ni l'autre des deux sous-ensembles d'arcs sur lesquels reposent ces évaluations, ne constitue une solution ; ils ne sont donc pas terminaux.

La phase d'examen ne permet aucune élimination, et cette partie préliminaire du calcul se termine en posant :

$$Z^1 = [V_0 - V_0] \cup [V_1, V_3] = [V_1, V_3].$$

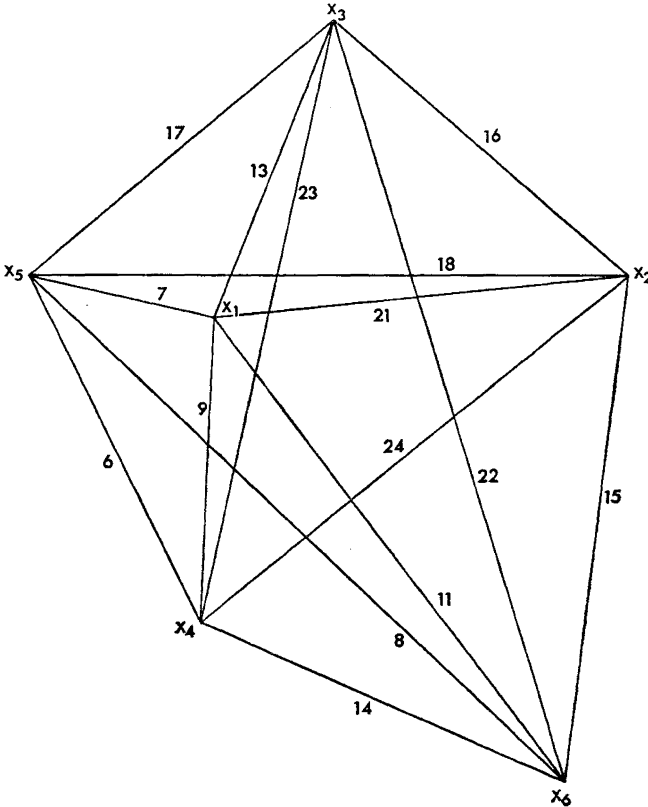


Figure 5

Graphe $G = (X, U)$ formé de 6 sommets et 15 arêtes ; le chiffre qui accompagne chaque arête représente sa longueur.

L'itération 1 débute ensuite avec la sélection de V_1 , lequel engendre (phase de séparation et évaluation) deux nouveaux suivants V_3 et V_4 (voir fig. 6).

Le sous-ensemble $V_3 = V(2, \{ (4, 5), (1, 5) \})$ fait apparaître les arêtes (5, 6) et (1, 4) (respectivement 3^e et 4^e dans le rangement par longueur croissante) comme incompatibles avec la portion de chaîne

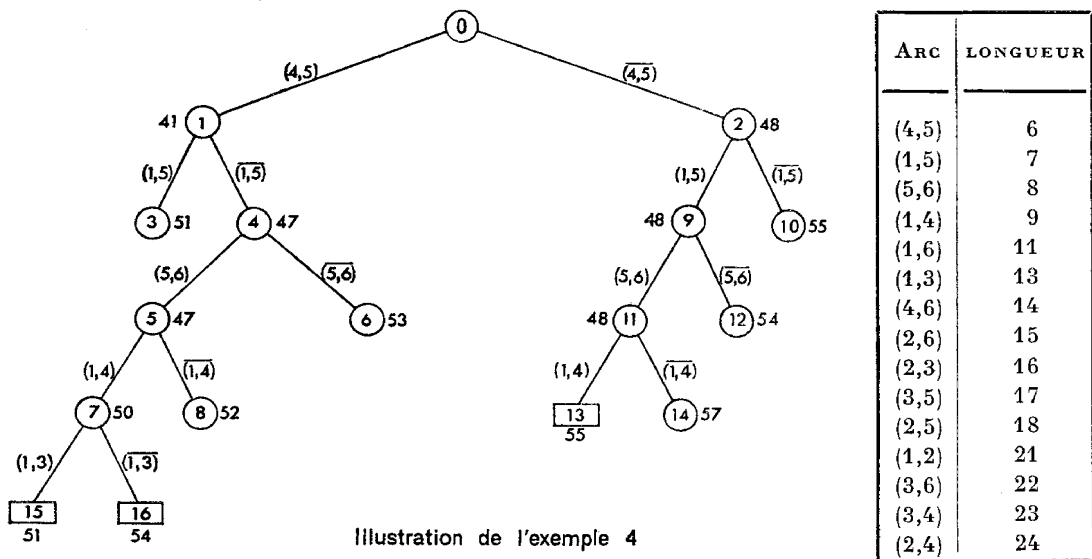


Illustration de l'exemple 4

Figure 6

acceptée : (5, 6) du fait de la propriété P_2 (voir 5 a, exemple 4) et (1,4) du fait de P_1 . On est ainsi conduit à poser pour cet ensemble V_3 :

$$u_{2+\epsilon} = (1, 6)$$

et :

$$\nu(3) = 6 + 7 + 11 + 13 + 14 = 51.$$

Le lecteur vérifiera que l'on a par ailleurs :

$$\nu(4) = 47.$$

et que ni ν_3 ni ν_4 n'est terminal.

Ni l'un ni l'autre n'est par ailleurs éliminable par la comparaison de :

$$h + m - (k + \epsilon - 1) = 2 + 15 - (2 + 3 - 1) = 13 \text{ pour } V_3$$

avec $n - 1 = 5$ (voir 5 a, exemple 4).

Dans ces conditions, le nouvel ensemble de candidats est :

$$Z^2 = \{ V_4, V_2, V_3 \} .$$

L'itération 3 conduit à séparer V_4 et s'achève par :

$$Z^3 = \{ V_5, V_2, V_3, V_6, \}$$

et en poursuivant, il vient de la même façon (voir fig. 6) :

$$Z^4 = \{ V_2, V_7, V_3, V_8, V_6 \}$$

$$Z^5 = \{ V_9, V_7, V_3, V_8, V_6, V_{10} \}$$

$$Z^6 = \{ V_{11}, V_7, V_3, V_8, V_6, V_{12}, V_{10} \}$$

tout cela sans que le calcul de l'évaluation permette d'exhiber un seul ensemble terminal non vide, et sans que la règle de comparaison rapelée plus haut permette d'éliminer un ensemble terminal vide.

En revanche, la séparation de V_{11} conduit à poser (pour l'arbitrage positif) :

$$V_{13} = V(4, \{ (1, 5), (5, 6), (1, 4) \})$$

$$u_{4+\varepsilon} = (2, 6),$$

et il est clair que les trois arêtes dont l'acceptation entre dans la définition de V_{13} , complétées par $(2, 6)$ et l'arête suivante, soit $(2, 3)$, constituent un groupe de 5 arêtes ne formant pas de cycle, et dans lequel chaque sommet n'apparaît que deux fois au plus. C'est donc une chaîne hamiltonienne, elle a une longueur égale à 55, ce qui justifie l'élimination de V_{14} (dont l'évaluation vaut 57). On est ainsi conduit à :

$$Z^7 + \{ V_7, V_3, V_8, V_6, V_{12}, V_{13}, V_{10} \} .$$

La séparation de V_7 conduit à deux nouveaux sous-ensembles terminaux non vides dont l'un (V_{15}) permet d'exhiber la chaîne hamiltonienne :

$$a_0(15) = (3,1), (1, 4), (4, 5), (5, 6), (6,2),$$

dont la longueur est 51. Tous les autres candidats, à l'exception de V_3 peuvent être éliminés, d'où :

$$Z^8 = \{ V_{15}, V_3 \} .$$

Pourvu que l'on donne toujours priorité à un candidat terminal sur un qui ne l'est pas, la sélection suivante porte sur V_{15} , et la procédure s'interrompt : la chaîne $a_0(15)$ est de longueur minimum.

Achevons cet exemple par quelques remarques.

1° Si l'on souhaite obtenir, non pas une, mais toutes les solutions optimales, il suffit de poursuivre dans les mêmes conditions avec $Z^9 = \{ V_3 \}$. Comme le lecteur pourra le vérifier le calcul prouve que la chaîne $a_0(15)$ est la seule solution optimale.

2° Si l'on souhaite obtenir d'autres solutions, par exemple toutes celles qui ne s'écartent pas de l'optimum de plus de deux unités, il suffit :

— de modifier la règle d'élimination afin de n'éliminer un candidat que si son évaluation dépasse de plus de 2 unités la meilleure solution connue au moment de l'élimination : ainsi V_{13} n'autorisera plus l'élimination de V_{14} , V_{13} n'autorisera plus celles de V_8 et V_6 ;

— de poursuivre le calcul — comme il a été dit en 6, a, 4° — aussi longtemps qu'il reste des candidats.

3° Dans cet exemple, la reconnaissance du caractère terminal non vide d'un sous-ensemble s'opère au moment du calcul de l'évaluation, et celle du caractère terminal vide au cours de la phase d'examen ; faisons simplement observer que l'une et l'autre de ces deux reconnaissances peuvent se situer différemment dans le processus.

4° A titre d'exercice le lecteur pourra bâtir une P.S.E.S. sur cet exemple en conservant le même principe de séparation et la même fonction d'évaluation ; pour ce qui est de l'ordre transverse il est naturel de définir le premier des deux suivants d'un sommet quelconque comme étant celui qui correspond à l'arbitrage positif ; il est intéressant de comparer l'arborescence ainsi obtenue à celle qui précède.

7. COMMENTAIRES PRATIQUES

La mise en œuvre pratique, à propos d'un problème défini, d'une P.S.E.P., ou d'une procédure heuristique qui en dérive, mérite quelques commentaires.

a) A propos du principe de séparation et de l'application d'évaluation

Une fois définis, l'univers \mathcal{U} et la partie A représentative des décisions décrétées admissibles à propos du problème étudié, le premier pas à franchir, pour bâtir une P.S.E.P., consiste à décider des bases sur lesquelles fonder le principe de séparation et l'application d'évaluation. Les cinq axiomes du paragraphe 5 laissent le plus souvent une assez grande liberté de choix. Les commentaires qui suivent ont précisément pour objet d'éclairer ce genre de choix.

On peut admettre sans perte de généralité notoire que chaque décision $a \in \mathcal{U}$ peut être analysée en décisions (1) élémentaires b_1, \dots, b_n et que l'on est en droit d'écrire :

$$a = (b_1, b_2, \dots, b_n).$$

(1) Nous attirons l'attention du lecteur sur l'ambiguïté que peut faire naître, dans ce qui suit, l'usage de ce mot « décision ». Il peut en effet se référer :

— tantôt à celui des états choisis au sein d'un champ d'états possibles, donc au résultat spécifié du choix : tel élément de A est une décision ;

— tantôt au fait même de devoir choisir au sein d'un champ d'états possibles, donc à la signification globale du choix : on parle de la décision « a ».

Il en sera bien ainsi chaque fois que a se présentera naturellement comme :

— les valeurs prises par des variables y_1, \dots, y_n — comme c'est, par exemple, le cas dans l'exemple 1 ou en programmation mathématique, qu'elle soit linéaire ou non ; la décision élémentaire b_1 s'interprétera alors comme le fait d'attribuer une valeur bien précise à y_1 ;

— les éléments retenus parmi ceux d'un ensemble $E = (e_1, \dots, e_n)$ comme c'est, par exemple, le cas dans l'exemple 2 en typologie ou à propos de telle ou telle classe de sous-ensembles remarquables des sommets d'un graphe ; b_1 signifiera alors que e_1 est retenu ou exclu ; (cf. § 6 b) — un ordre (ou un préordre) défini sur un ensemble $E = (e_1, \dots, e_n)$ — comme c'est le cas dans de nombreux problèmes de classement, d'ordonnement ; b_1 consistera par exemple, à décider de celui des éléments de E qui occupera la première place.

Il est évidemment naturel et très général de vouloir fonder la séparation sur une telle décomposition de « a ». Comme il est facile de s'en rendre compte, il y a en effet bien des manières de considérer une à une de telles décisions élémentaires en vue d'arbitrages progressifs et ramifiés. Nous dirons qu'une décision donne lieu à *arbitrage* si l'on décide de restreindre le champ de ses états possibles ; le résultat d'un arbitrage est donc un sous-ensemble d'un ensemble d'états, donc, éventuellement, un état spécifié. On appellera arbitrage complémentaire, celui qui consiste à restreindre les états possibles au sous-ensemble complémentaire.

Lorsque l'on fonde ainsi la séparation sur une telle décomposition de « a » en décisions élémentaires, les sous-ensembles V de A auxquels on aboutit après une ou plusieurs séparations sont presque toujours caractérisés par une succession de résultats d'arbitrages relatifs à ces décisions élémentaires. (Revoir 6 b, cf. également B. Roy, R. Benayoun, J. Tergny, [16].)

Le principe de séparation qui a été proposé au 5 a exemple 4 s'appuie sur un ordre *a priori* des décisions élémentaires. On remarquera que cet ordre n'est pas sans présenter des liens étroits avec le problème. Chaque fois que l'on pourra trouver un tel ordre il sera possible de transposer la façon de faire ci-dessus.

Dans d'autres cas on aura intérêt, plutôt qu'à définir à *priori* un ordre arbitraire et commun à tous les chemins de l'arborescence, à choisir celle des décisions devant intervenir comme support du nouvel arbitrage d'après les résultats et particularités du sous-ensemble que l'on se propose de séparer. C'est ainsi que ce choix peut, comme dans l'application faite par B. Roy, R. Benayoun, J. Tergny, J. de Buchet [16] aux problèmes d'acheminement et d'implantation, dépendre d'une décision particulière liée à l'ensemble que l'on sépare ; R. Descamps et P. Chevignon dans [9] le font dépendre à propos des problèmes d'ordonnement, d'une estimation comparative de l'effet des diverses décisions élémentaires susceptibles d'intervenir.

Pour bâtir une application d'évaluation conforme aux axiomes du 5 b, on peut toujours s'inspirer de l'une, ou l'autre, ou des deux démarches suivantes :

— déduire de Φ une application plus simple dont on sache calculer le minimum sur les ensembles V engendrés par la séparation (c'est ce qui a été fait à propos des problèmes, déjà évoqués, d'implantation et acheminement, en linéarisant la fonction Φ) ;

— plonger le sous-ensemble V dans un ensemble W (plus vaste) sur lequel on saura calculer le minimum de Φ : c'est une voie assez bien adaptée aux problèmes avec contraintes complexes (comme le sont certains problèmes d'ordonnement) car l'extension de V se fait aisément en négligeant certaines de ces contraintes pour ce qu'elles impliquent sur les décisions élémentaires n'intervenant pas directement pour circonscrire V . De même, lorsque des inconnues sont astreintes à ne prendre que certaines valeurs (entières par exemple) d'un intervalle, on peut définir W en abandonnant ce genre de restriction (cf. par exemple Herve [10], ou B. Roy, R. Benayoun, J. Tergny, J. de Buchet [16]).

Les commentaires qui précèdent font en particulier ressortir l'énorme latitude dont on dispose au moment de bâtir une procédure S.E.P. ; cela conduit à se demander si l'on peut avancer quelques principes simples et généraux permettant de juger *a priori* de la plus ou moins grande efficacité numérique d'un principe de séparation joint à une application d'évaluation.

Il nous paraît important que le principe de séparation favorise l'apparition rapide de sous-ensembles (aussi riches que possible) dont l'évaluation dépasse le minimum de Φ sur A ; (il est clair que de tels sous-ensembles, même s'ils ne sont pas éliminés à la phase d'examen, ne seront jamais séparés). Cela suppose en particulier que :

— le principe de séparation — tout en restant simple — dissémine le moins possible parmi les divers sous-ensembles qu'il engendre, les décisions proches de l'optimum ou plus généralement satisfaisantes : l'idéal serait en effet de pouvoir isoler dans un même sous-ensemble, et à chaque séparation, les décisions les plus satisfaisantes du sous-ensemble que l'on sépare ; on remarquera que c'est sur un espoir de cet ordre qu'est fondée la méthode des centres (cf. § 2) ;

— l'application par défaut ν — tout en restant d'un calcul rapide — présente par rapport au minimum de Φ un biais de même ordre sur chaque sous-ensemble ν_1 , afin que l'ordre ainsi induit sur les différents sous-ensembles candidats soit aussi proche que possible de celui associé au minimum exact de Φ sur chacun d'eux ; nous verrons plus loin (cf. § c) que d'autres raisons militent en faveur d'une bonne concordance des deux ordres.

Quoi qu'il en soit, on a toujours énormément de difficultés pour savoir, *a priori*, si tel raffinement du principe de séparation et/ou de

l'application d'évaluation conduisant à des itérations plus longues sera ou non justifié par une diminution de leur nombre.

b) Sous ensembles terminaux. Fonctions δ_{seuil} et δ^*

Les bases de la séparation et de l'évaluation ayant été jetées, il se peut fort bien que l'on dispose encore d'une certaine liberté pour achever de définir la procédure.

L'un des points les plus délicats concerne l'arrêt de la séparation, autrement dit la définition des ensembles terminaux. Très fréquemment on sera en mesure de construire, pour chaque sous-ensemble V_q engendré par la séparation, une décision $a_0(q) \in V_q$ d'autant meilleure que la séparation aura été poussée plus loin. D'après les axiomes d'arrêt (cf. 5 a) et de coïncidence (cf. 5 b), il suffit pour déclarer V_q terminal soit que l'on ait la preuve de l'optimalité de $a_0(q)$ dans ce sous-ensemble, soit que cette décision apparaisse comme satisfaisante relativement à V_q . Il y a évidemment maintes façons de faire ce genre de constatation :

- V_q se réduit (par construction) à un seul élément ;
- $\Phi[a_0(q)]$ coïncide avec $\nu(q)$ comme au paragraphe 5 à propos de l'exemple 4 ;
- $a_0(q)$ possède telle propriété caractéristique permettant d'affirmer son optimalité au sein de V_q ou l'impossibilité de trouver dans V_q un élément significativement plus satisfaisant ;
- $\Phi[a_0(q)]$ (et $\psi_1[a_0(q)]$, ... éventuellement) comparé à $\nu(q)$ (et à $\nu_1(q)$, ..., respectivement) indique que V_q ne contient pas de décision significativement plus satisfaisante : on peut pour cela introduire une (ou plusieurs) application δ .

Il est enfin un dernier choix sur lequel nous ne nous étendrons guère, c'est celui relatif à l'introduction d'une éventuelle application δ^* dans le cadre de la phase de sélection. Celle-ci est à envisager chaque fois que des calculs, faits à propos d'un sous-ensemble V_q (notamment pour construire $a_0(q)$ ou calculer $\nu(q)$), doivent être refaits, ou plus simplement peuvent être utiles, à propos des sous-ensembles engendrés par séparation de V_q . En effet, il est souvent impossible (en raison des capacités limitées de stockage de l'information, cf. 3^o ci-après) de conserver, à propos de chaque candidat, de tels résultats intermédiaires. En revanche, on peut les conserver à propos du dernier calcul dans la mesure où ils doivent servir au suivant. L'application δ^* a précisément pour objet de favoriser ce que l'on peut appeler « la progression le long d'une branche » de l'arborescence aux dépens « des sauts de branche en branche ».

Une telle fonction apparaît sans objet dans les exemples 4 et 5.

c) Interruption prématurée de la procédure, procédures heuristiques dérivées d'une S.E.P.

Lorsque l'on s'engage dans l'exécution des consignes opératoires du 6 a, on ignore le plus souvent le nombre d'itérations nécessaires avant l'arrêt, ainsi que le nombre maximum d'éléments que renfermera Z^p .

Que l'on procède manuellement, ou que l'on utilise un calculateur électronique, on s'expose donc à une interruption prématurée des calculs

soit parce qu'ils sont trop longs, soit parce que la place fait défaut pour stocker l'information à conserver d'une itération à la suivante. En fait, ce sont là deux difficultés étroitement liées et inhérentes à de nombreux calculs ; les commentaires qui suivent indiquent comment les surmonter dans le cadre d'une P.S.E.P.

Observons tout d'abord que, dès l'instant où l'on introduit à propos de chaque V_q un élément $a_0(q)$ présumé satisfaisant dans cet ensemble, il devient possible d'interrompre le calcul à tout moment en exhibant la ou les solutions les plus satisfaisantes obtenues au moment de l'interruption. Il suffit pour cela de comparer à chaque itération et sur la base du ou des critères, les éléments $a_0(q)$ relatifs aux sous-ensembles engendrés à la phase de séparation, à ceux considérés antérieurement comme les plus satisfaisants. Il est important de souligner que, si $a_0(q)$ est déterminé sans aucune ambiguïté par les seuls caractéristiques de V_q , il ne sera pas nécessaire de conserver ces éléments les plus satisfaisants tout au long des calculs puisqu'ils pourront aisément être retrouvés à tout moment à partir de la seule connaissance des sous-ensembles qui les referment. On aura en revanche intérêt à conserver leur image par Φ (et ψ_1, \dots éventuellement) pour permettre les comparaisons ainsi que l'élimination des candidats dont l'évaluation est supérieure (cf. 2 c).

Lorsque l'on n'a pas cette faculté d'interrompre à tout moment la procédure, (les $a_0(q)$ n'étant définis que pour les seuls V_q terminaux), ou encore lorsqu'on souhaite atteindre rapidement une « bonne solution », on peut, au fur et à mesure de son déroulement, en modifier certaines données afin de hâter l'arrêt.

On peut en premier lieu modifier les conditions d'arrêt par exemple en se contentant d'une approximation plus grossière (application δ).

Si l'on accepte d'abandonner le caractère rigoureux de la procédure, pour la rendre heuristique, on peut en second lieu hâter l'arrêt en introduisant à la phase d'examen, des règles d'élimination qui ne reposent plus sur des certitudes, mais seulement sur des présomptions quant au bien-fondé de l'élimination. On remarquera que du même coup on réduit le nombre de candidats et la masse d'information à stocker. C'est ainsi par exemple que l'on pourra systématiquement éliminer le ou les sous-ensembles engendrés à la précédente séparation pour lesquels l'évaluation est la plus forte, ou encore ceux pour lesquels certaines conditions empiriques ne sont pas remplies.

A moins que l'on ne pousse très loin ce genre d'élimination, (on peut aller jusqu'à ne conserver qu'un seul candidat pour chaque itération : c'est notamment ce qui est fait dans la méthode des centres (cf. fig. 2), il y a toujours le risque de voir grossir Z^p au-delà des possibilités de stockage de l'information. Faisons observer que la règle la plus simple consiste alors à éliminer ceux des candidats qui se trouvent être les derniers dans Z^p (ordonné d'après ν croissant). Cette façon de faire est pleinement justifiée si ν classe les candidats à peu près comme les classe le

minimum de Φ . De façon plus précise, il suffit, pour être en droit de limiter l'ensemble des candidats à K sous-ensembles, que le (ou les) premiers b) selon le minimum de q se trouvent) sans cesse parmi les K premiers selon v . S'il y avait identité des deux ordres tout au long du calcul, on pourrait limiter d'emblée Z^p à son ou ses premiers éléments.

Dans le cadre d'une procédure heuristique, on pourra donc imposer une limite à Z^p (l'excédent étant éliminé comme il vient d'être dit), et éventuellement réduire au fur et à mesure que le calcul s'avance en vue de hâter l'arrêt. Il sera alors toujours intéressant de suivre au fur et à mesure des itérations le minimum de l'évaluation des candidats éliminés parce qu'excédentaires. En effet, si le minimum ainsi trouvé en fin de calcul est supérieur à la valeur de la solution finalement retenue, il est clair que celle-ci aurait été la même sans l'élimination des candidats excédentaires.

En revanche, s'il est inférieur, l'écart entre les deux échelons considérés peut être une information intéressante pour juger de la solution finale.

Pour conclure, nous ne saurions trop insister sur le rôle important que peut jouer la procédure S.E.P. en tant que cadre de référence rigoureux permettant d'imaginer et de juger des procédures heuristiques ; n'est-ce pas en fait cela que l'on demande à la théorie lorsqu'on se préoccupe d'applications ?

BIBLIOGRAPHIE

- [1] ABADIE (J.), *Nonlinear programming*, North-Holland Publishing Company, 1967, chap. VIII, pp. 208-249.
- [2] AUDIBERT (J. M.), HOLL (J. C.) et PLAS (J. P.), « Un modèle de calcul de programme d'investissement (Capri) », *Metra*, vol. VII, n° 2, 1968.
- [3] BALAS (E.), « A note on the Branch-and-Bound Principle », Carnegie Mellon University. Pittsburg, Pennsylvania. *Operations Research*, vol. 16, n° 2, pp. 442-445, 1968.
- [3 bis] BALINSKY (M. L.), « Integer Programming : Methodes, uses and computation », *Management Sciences*, vol. 12, 1965, pp. 253-313.
- [3 ter] BEALE (E.) et SMALL (R.), Mixed integer programming by a branch and bound technique, IFIP Congress 1965, New York.
- [4] BENAYOUN (R.) et ROY (B.), « Programmes linéaires en variables bivalentes et continues sur un graphe » (Le programme Poligami). *Metra*, vol. VI, n° 4, 1967.
- [5] BERTIER (P.), « Procédures pour élaborer des tournées de distribution », *Metra*, série spéciale, n° 8, 1966.
- [6] BERTIER (P.), « Quelques algorithmes pour les problèmes de tournée », *Metra*, vol. IV, n° 4, décembre 1965.
- [7] BERTIER (P.), NGHIEM (Phong Tuan) et ROY (B.), « Programmes linéaires en nombres entiers et procédures S.E.P. », *Metra*, vol. IV, n° 3, 1966.
- [7 bis] DAKIN (R. J.), « A tree search algorithm for mixed integer programming problems », *Computer Journal*, vol. 8, n° 3, october 1965, pp. 250-255.

- [8] DANTZIG (G. B.), *Linear programming and extensions*, Princeton University Press, 1963.
- [8 bis] DAVIS (R.), KENDRICK (D.) et WEITZMANN (M.), « A Branch and Bound algorithm for zero-one mixed Integer Programming Problem », *Development Economic Report*, n° 69, oct. 1967, Harvard University, Cambridge, Massachusetts,
- [9] DESCAMPS (R.) et CHEVIGNON (P.), *Optimisation algorithm for a class of scheduling problems with disjunctive hereditary constraints*, Vienne, Congrès Internet, 1967.
- [10] HERVE (P.), « Résolution des programmes linéaires à variables mixtes par la procédure S.E.P. », *Metra*, vol. VI, n° 1, 1967.
- [11] HERVE (P.), « Les procédures arborescentes d'optimisation », *R.I.R.O.*, n° 14, V 3, 1968.
- [12] HUARD (P.), « Programmation mathématique convexe », *R.I.R.O.*, 1968, n° 7, pp. 43-59.
- [13] KUNZI (H. P.), *Nonlinear programming*, Blaisdel Pub., 1966.
- [13 bis] LAND (A. H.) et DOIG (A. G.), « An automatic method for solving discrete programming problems », *Econometrica*, vol. 28, 1960, pp. 497-520.
- [13 ter] LITTLE (J. D. C.), MURTY (K.), SWEENEY (D. W.) et KAREL (C.), « The Travelling Salesman Problem », *Operations Research*, 11, pp. 972-989 (1963).
- [14] ROY (B.), *Sur quelques aspects méthodologiques des problèmes d'ordonnement* (Un essai de classification), Internet, Vienne, 1967.
- [15] ROY (B.), *Algèbre moderne et Théorie des Graphes. Application aux Sciences économiques et sociales*, Dunod, 1969, chap. X.
- [16] ROY (B.), BENAYOUN (R.) et TERGNY (J.), « From S.E.P. Procedure to Ophélie Mixte ». In : *Integer and Nonlinear Programming* (vol. II) (J. Abadie Editor), North-Holland Publishing Company and Wiley, 1970.
- [17] ROY (B.), et SUSSMANN (B.), *Problèmes d'ordonnement avec contraintes disjonctives*. SEMA, Direction Scientifique, Rapport de Recherche n° 9, oct. 1964.
- [18] SIMONNARD (M.), *Programmation linéaire*, Paris, Dunod, 1962.
- [19] WILDE (J.), *Méthodes de recherche d'un optimum*. Paris, Dunod, 1966.