Chapitre I: Introduction à l'optimisation mathématique

1. Introduction

L'optimisation fait partie de la vie des êtres vivants. Le processus d'évolution dans la nature révèle qu'il suit l'optimisation. Par exemple, les animaux qui vivent dans des climats plus froids ont des membres plus petits que ceux qui vivent dans des climats plus chauds, afin d'obtenir un rapport surface/volume minimal.

Dans notre vie quotidienne, nous prenons des décisions qui, selon nous, peuvent maximiser ou minimiser notre ensemble d'objectifs, comme prendre un raccourci pour minimiser le temps nécessaire pour atteindre une destination particulière, trouver le meilleure logement possible qui peut satisfaire un maximum de conditions avec des contraintes de coût, ou trouver un article au prix le plus bas dans un magasin. La plupart de ces décisions sont fondées sur nos connaissances du système que nous avions acquises pendant des années, et sans recourir à une théorie mathématique systématique. Cependant, à mesure que le système se complique, qu'il faut prendre de plus en plus de décisions simultanément et qu'il est contraint par divers facteurs, dont certains sont nouveaux pour le système, il est difficile de prendre des décisions optimales sur la base d'une heuristique et de connaissances antérieures. En outre, les enjeux sont souvent élevés et il faut satisfaire de multiples parties prenantes.

La tâche d'optimisation est au cœur de tout problème impliquant une prise de décision, que ce soit en ingénierie, en économie, en agriculture, en médecine, etc. La tâche de prise de décision consiste à choisir parmi diverses alternatives. Ce choix est régi par notre désir de prendre la "meilleure" décision. La mesure de la qualité des alternatives est décrite par une *fonction objectif* ou un *indice de performanc*e. La théorie et les méthodes d'optimisation traitent de la sélection de la meilleure alternative au sens de la fonction objectif donnée. La théorie de l'*optimisation mathématique* offre une meilleure alternative pour la prise de décision dans diverses situations, à condition de pouvoir représenter mathématiquement les décisions et le système.

L'optimisation peut se décrire comme un processus décisionnel en trois étapes, où la première étape consiste à connaître le système, ce qui, en termes mathématiques, peut être considéré comme une *modélisation du processus*. La deuxième étape consiste à trouver une mesure de l'efficacité du système, ou la *fonction objectif*. La troisième étape repose sur la *théorie de l'optimisation* afin de, selon la nature du problème considéré, maximiser ou minimiser cette fonction.

Le domaine de l'optimisation a reçu une énorme attention depuis les deux dernières décennies, principalement en raison de l'évolution rapide de la technologie informatique, notamment le développement et la disponibilité de logiciels conviviaux, de processeurs rapides et parallèles, et de modèles d'optimisation non linaire qui peuvent être facilement ajustés aux donnés du problème considéré, notamment les réseaux de neurones artificiels . Un exemple clair de ce phénomène est la grande accessibilité des outils logiciels d'optimisation tels que ADMB, CUTER, Scilab, CPLEX, TOMLAB, SAS, Optimization Toolbox de MATLAB, etc.

L'objectif de ce chapitre est donner une vue d'ensemble sur l'optimisation et son domaine majeur d'application qui est la recherche opérationnelle. Cette dernière est une discipline moderne qui utilise des modèles mathématiques, statistiques et des algorithmes pour modéliser et résoudre des problèmes complexes, en déterminant la solution optimale et en améliorant la prise de décisions.

2. Origine de l'optimisation mathématique

La théorie de l'optimisation trouve son origine dans le problème *isopérimétrique* auquel la reine Didon a été confrontée en 1000 avant Jésus-Christ. Elle a procuré, pour la fondation de Carthage, la plus grande surface de terre pouvant être entourée de la peau d'un taureau. De cette peau, elle fit une corde, qu'elle disposa en demi-cercle avec les extrémités contre la mer. La solution intuitive de la reine Didon était correcte. Mais il a fallu attendre plusieurs siècles avant qu'une preuve formelle ne soit présentée, et la solution mathématique et systématique de ce problème s'est avérée être un problème très difficile dans le calcul différentiel. Le calcul différentiel traite essentiellement des problèmes où la variable de décision est un vecteur. La figure I.1 montre un rectangle, un carré et un cercle ayant le même périmètre mais des aires différentes. On peut voir que la surface maximale est couverte par le cercle.

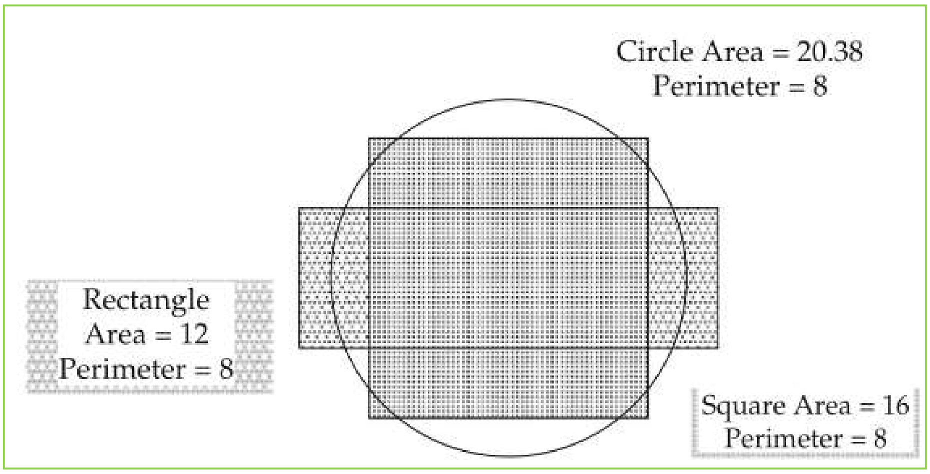

Figure I.1: Objets isopérimétriques.

Le *calcul différentiel*[1], ou la première *théorie systématique de l'optimisation*, est né le 4 juin 1694, lorsque John Bernoulli a posé le problème du *Brachistochrone* (mot grec signifiant "temps le plus court") et a publiquement mis le monde mathématique au défi de le résoudre. Le problème posé était le suivant: "Quelle est la trajectoire de glissement sur laquelle un objet sans frottement glisserait en un minimum de temps (voir figure I.2) ?". De nombreux scientifiques de renommée ont déjà tenté de résoudre ce problème, notamment Galilée, qui a proposé un arc de cercle comme solution, une solution incorrecte, et Leibnitz, qui a présenté des équations différentielles ordinaires sans les résoudre. Puis John Bernoulli a prouvé que le chemin optimal était une *cycloïde* (i.e., courbe décrite par un point appartenant à la circonférence d'un cercle, lorsque celui-ci roule sur une droite tracée dans un plan horizontal). À partir de ce moment, les efforts se sont poursuivis dans le domaine du calcul différentiel, conduisant à l'étude des intégrales multiples, des équations différentielles, de la théorie de la commande (ou contrôle), de la transformation des problèmes, etc. Bien que ces recherches portent principalement sur la théorie et les solutions analytiques, elles constituent la base de l'optimisation numérique développée pendant et après la seconde guerre mondiale. La Seconde guerre mondiale a sensibilisé les scientifiques à l'optimisation numérique et aux solutions aux problèmes de physique et d'ingénierie. En 1947, Dantzig a proposé *l'algorithme du simplexe* pour les problèmes de programmation linéaire. Les conditions nécessaires ont été

[1] *Le calcul différentiel* est une branche des mathématiques qui traite du calcul des dérivées et de leurs applications. Il est essentiel dans de nombreux domaines, notamment en mathématiques, en sciences physiques, en ingénierie et en économie.

présentées par Kuhn et Tucker au début des années 1950, ce qui a constitué un point focal pour la recherche sur la programmation non linéaire. Aujourd'hui, les techniques d'optimisation numérique constituent une partie fondamentale des sciences et de l'ingénierie théoriques et pratiques.

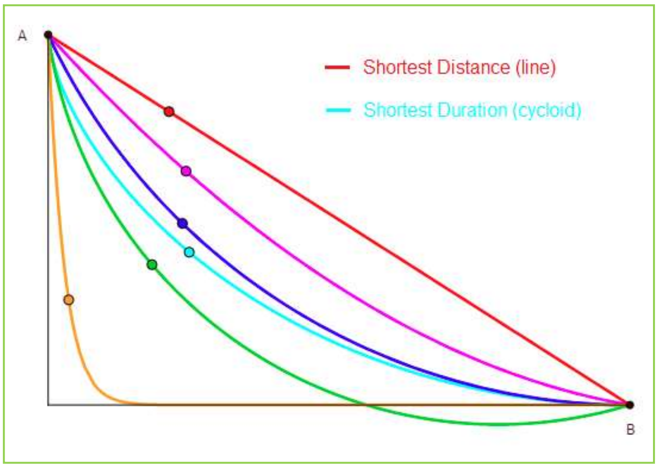

Figure I.2: Illustration du problème Brachistochrone.

3. Définition d'optimisation mathématique

L'*optimisation*, également connue sous le nom de *programmation mathématique*, est un ensemble de principes et de méthodes mathématiques utilisés pour modéliser, analyser et résoudre des problèmes quantitatifs dans de nombreuses disciplines, notamment la physique, la biologie, l'ingénierie, l'économie et le commerce. Le sujet est né de la prise de conscience que des problèmes quantitatifs dans des disciplines manifestement différentes ont des éléments mathématiques importants en commun. En raison de ce point commun, de nombreux problèmes peuvent être formulés et résolus en utilisant l'ensemble unifié d'idées et de méthodes qui constituent le domaine de l'optimisation.

Le terme historique de programmation mathématique, largement synonyme d'optimisation, a été inventé dans les années 1940, avant que la programmation ne soit assimilée à la programmation informatique. La *programmation mathématique* comprend l'étude de la structure mathématique des problèmes d'optimisation, l'invention de méthodes pour résoudre ces problèmes, l'étude des propriétés mathématiques de ces méthodes et la mise en œuvre de ces méthodes sur des ordinateurs. La rapidité des ordinateurs a considérablement augmenté la taille et la complexité des problèmes d'optimisation qui peuvent être résolus. Le développement des techniques d'optimisation a suivi les progrès réalisés non seulement en informatique, mais aussi en recherche opérationnelle, en analyse numérique, en théorie des jeux, en économie mathématique, en théorie de la commande et en combinatoire.

Les problèmes d'optimisation comportent généralement trois éléments fondamentaux. Le premier est une quantité numérique unique, ou *fonction objectif*, qui doit être maximisée ou minimisée. L'objectif peut être le rendement attendu d'un portefeuille d'actions, les coûts de production ou les bénéfices d'une entreprise, le temps d'arrivée d'un véhicule à une destination donnée ou la part de voix d'un candidat politique. Le deuxième élément est un ensemble de variables, qui sont des quantités dont les valeurs peuvent être manipulées afin d'optimiser l'objectif. Il s'agit par exemple des quantités de stock à acheter ou à vendre, des quantités de diverses ressources à allouer à différentes activités de production, de l'itinéraire à suivre par un véhicule dans un réseau de circulation, ou des politiques à défendre par un candidat. Le troisième élément d'un problème d'optimisation est un ensemble de contraintes, qui sont des restrictions sur les valeurs que peuvent prendre les variables. Par exemple, un processus de fabrication ne peut pas nécessiter plus de ressources que celles qui sont disponibles, ni employer moins de ressources que zéro. Dans ce cadre général, les problèmes d'optimisation peuvent avoir différentes propriétés mathématiques. Les problèmes dans lesquels les variables sont des quantités continues (comme dans l'exemple de l'allocation des ressources) nécessitent une approche différente des problèmes dans lesquels les variables sont des quantités discrètes ou combinatoires (comme dans la sélection d'un itinéraire de véhicule parmi un ensemble prédéfini de possibilités).

4. Les types des problèmes d'optimisation

Les problèmes d'optimisation peuvent être divisés en grandes catégories selon le type de variables de décision, de fonction(s) objective(s) et de contraintes.

4.1. Programmation linéaire

Une classe importante d'optimisation est connue sous le nom de *programmation linéaire (PL)*. *Linéaire* signifie que la fonction objectif est les contraintes d'optimisation sont linéaires. Pour cette classe, les problèmes consistent à minimiser (ou maximiser) une fonction objectif linéaire dont les variables sont des nombres réels contraints de satisfaire un système d'égalités et d'inégalités linéaires. La programmation linéaire fait l'objet unique de ce cours.

<u>Exemple de programme linéaire</u>

$$Maximiser\;\; 3x + y \tag{I.1}$$
$sujet$ à:
$$\begin{cases} 2x + 4y \leq 220 \\ 3x + 2y \leq 150 \\ \quad x \geq 0 \\ \quad y \geq 0 \end{cases}$$

4.2. Programmation non linéaire

Une autre classe importante d'optimisation est connue sous le nom de *programmation non linéaire (PNL)*. Dans la programmation non linéaire, les variables sont des nombres réels, et la fonction objectif ou certaines des contraintes sont des fonctions non linéaires (impliquant éventuellement des carrés, des racines carrées, des fonctions trigonométriques ou des produits des variables, etc.).

<u>Exemple de programme non linéaire</u>

$$Minimiser\;\; 3x^3 + \log y - \sin(z) \tag{I.2}$$
$sujet$ à:
$$\begin{cases} 2x + 23\cos(y) \leq 20 \\ e^x + \sin(y) \leq 15 \\ \quad x \geq 0 \\ \quad y > 0 \end{cases}$$

4.3. Programmation en nombres entiers

Un *problème d'optimisation en nombres e*ntiers est un programme mathématique d'optimisation ou de faisabilité dans lequel certaines ou toutes les variables de décision sont limitées à des nombres entiers.

Dans de nombreux cas, ce terme fait référence à la programmation linéaire en nombres entiers (PLNE), dans laquelle la fonction objectif et les contraintes (autres que les contraintes entières) sont linéaires.

Exemple de programme linéaire en nombres entiers

$$\begin{aligned}
&Maximiser\ x + y \qquad \text{(I.3)}\\
&\text{sujet à:}\\
&\begin{cases}
-x + y \ \leq 1\\
5x + 2y \leq 15\\
2x + 5y \leq 15\\
x \geq 0\\
y \geq 0\\
x, y \in \mathbb{Z}
\end{cases}
\end{aligned}$$

On peut distinguer également les problèmes d'optimisation linéaire mixte en nombres entiers (PLMNE): la fonction objectif et les contraintes sont linéaires. Les variables de décision sont scalaires; certaines d'entre elles sont des entiers tandis que d'autres sont des variables continues.

Exemple de programme linéaire mixte en nombres entiers

$$\begin{aligned}
&Minimize\ -x + 2y + 4z \qquad \text{(I.4)}\\
&\text{sujet à:}\\
&\begin{cases}
-x + y \ \leq 1\\
5x + y - z \leq 12\\
3x + \frac{1}{2}z \geq 10\\
x \geq 0\\
y \geq 0\\
x, y \in \mathbb{Z}
\end{cases}
\end{aligned}$$

4.4. Programmation non linéaire mixte en nombres entiers (PNLMNE)

Il s'agit d'un problème d'optimisation non linéaire impliquant des variables de décision entières et continues.

$$\begin{aligned}
&Maximiser\ x^{1/2} + \log_{10} y + \frac{x}{z} \qquad \text{(I.5)}\\
&sujet\ \text{à:}\\
&\begin{cases}
xy + z \leq 20\\
e^{x} + e^{-y} \leq 15\\
x \geq 0\\
y > 0\\
x, z \in \mathbb{Z}
\end{cases}
\end{aligned}$$

4.5. Optimisation discrète

Elle concerne tous les problèmes d'optimisation impliquant des variables de décision discrètes (entières). Cela inclut les PLNE, PLMNE et PNLENE.

4.6. Commande optimale

La théorie de la commande (contrôle) optimale est une branche de l'optimisation mathématique qui se consacre à la détermination d'une commande d'un *système dynamique* sur une période de temps qui minimise ou (maximise) un critère de performance (i.e., une fonction objectif), éventuellement sous des contraintes pouvant porter sur la commande ou sur l'état du système. Elle a de nombreuses applications en science, en ingénierie et en recherche opérationnelle. Les variables de décision dans ce cas sont des vecteurs. Par exemple, le système dynamique peut être un vaisseau spatial avec des commandes correspondant à des propulseurs de fusée, et l'objectif peut être d'atteindre la lune avec une dépense minimale de carburant, ou le système dynamique peut être l'économie d'une nation, avec l'objectif de minimiser le chômage; les commandes dans ce cas peuvent être la politique fiscale et monétaire.

Exemple : Commande optimale d'un véhicule électrique

Supposons que nous concevions la commande optimale pour un véhicule électrique afin de maximiser son autonomie tout en respectant une contrainte de temps. Le but est de trouver le profil de vitesse optimal pour le véhicule pendant un trajet donné.

Variables :

- $x(t)$: La position du véhicule au temps t.
- $v(t)$: La vitesse du véhicule au temps t.
- $u(t)$: L'accélération (commande de puissance) du véhicule au temps t.

Dynamique du système :

- La vitesse du véhicule évolue selon l'équation : $\dot{v}(t)=a(u(t))-b(v(t))$, où $\dot{v}(t)$ est la dérivée de la vitesse par rapport au temps, $a(u(t))$ représente l'accélération produite par la commande de puissance $u(t)$, et $b(v(t))$ représente la décélération due à la résistance au roulement et à l'air.

Objectif :

- Maximiser l'autonomie du véhicule, c'est-à-dire la distance parcourue avant que la batterie ne soit complètement épuisée, tout en respectant un temps de trajet maximum fixé, T.

Contraintes :

- Le véhicule ne peut pas dépasser une vitesse maximale, v_{max}.
- Le temps de trajet ne doit pas dépasser T.

Solution :

- Pour résoudre ce problème de commande optimale, nous pourrions utiliser des techniques d'optimisation numérique ou des méthodes analytiques comme le principe du maximum de Pontryagin[2]. L'objectif serait de trouver le profil d'accélération u(t) optimal qui permet de maximiser l'autonomie du véhicule tout en respectant les contraintes de vitesse et de temps.
- La solution obtenue pourrait être un profil d'accélération qui minimise la consommation d'énergie tout en respectant les contraintes, ce qui permettrait au véhicule électrique de parcourir la plus grande distance possible dans le temps imparti.

Cet exemple illustre comment la commande optimale peut être appliquée dans le domaine des véhicules électriques pour maximiser l'autonomie, mais elle est également utilisée dans de

[2] Le *principe du maximum de Pontryagin* énonce que pour trouver le contrôle optimal, il existe une fonction auxiliaire appelée "fonction de Hamiltonienne" (ou simplement "Hamiltonien") qui doit être maximisée le long de la trajectoire optimale. Le Hamiltonien est généralement défini comme la somme de la fonction de coût et du produit scalaire des contraintes et des multiplicateurs de Lagrange associés aux contraintes du problème.

nombreux autres domaines, tels que la gestion de processus industriels, la robotique, la finance, et bien d'autres, pour résoudre des problèmes d'optimisation complexes.

4.7. Optimisation stochastique

Dans l'*optimisation stochastique* (appelée aussi *optimisation sous incertitudes*) la fonction objectif ou les contraintes dépendent de variables aléatoires, de sorte que l'optimum est trouvé dans un certain sens "attendu" ou probabiliste. Elle implique souvent les types d'optimisation ci-dessus comme sous-catégories.

Exemple : Planification de la production sous incertitude de la demande

Imaginez une entreprise qui fabrique un produit et doit décider de son plan de production pour les prochains mois. Cependant, la demande pour ce produit est incertaine et peut varier d'un mois à l'autre en raison de facteurs tels que les saisons ou les fluctuations économiques. L'entreprise souhaite maximiser ses bénéfices en ajustant sa production chaque mois en fonction de la demande anticipée.

Variables :

- x_t : La quantité de produit à produire au mois t.
- D_t : La demande anticipée pour le mois t.
- P_t : Le prix de vente du produit au mois t.
- C_t: Le coût de production unitaire au mois t.

Objectif :

- Maximiser les bénéfices totaux sur une période de plusieurs mois, en tenant compte de l'incertitude de la demande. L'objectif est donc de maximiser l'espérance des bénéfices.

$$\max \sum_{t=1}^{T} \mathbb{E}[P_t \cdot \min(x_t, D_t) - C_t \cdot x_t]$$

Contraintes :

- La quantité produite ne peut pas être négative : $x_t \geq 0$ pour tout t.
- La quantité produite ne peut pas dépasser la capacité de production maximale : $x_t \leq X_{max}$ pour tout t.
- Les décisions sont prises mois par mois, en fonction de la demande anticipée D_t.

Incertain :

- La demande anticipée D_t est incertaine et suit une distribution de probabilité, par exemple une distribution normale, i.e. $D_t \sim N(\mu, \sigma^2)$, où μ est la moyenne (espérance) de la taille des étudiants et σ^2 est la variance de D_t.

Solution :

- Pour résoudre ce problème d'optimisation stochastique, nous pourrions utiliser des techniques d'optimisation stochastique telles que la programmation stochastique ou la programmation dynamique récursive avec la théorie des files d'attente. Le processus consisterait à déterminer la quantité optimale à produire chaque mois, en tenant compte de l'incertitude de la demande et des coûts de production.

- La solution finale serait un plan de production qui maximise l'espérance des bénéfices sur la période de planification tout en respectant les contraintes de capacité de production.

Ce scénario montre comment l'optimisation stochastique peut être utilisée pour prendre des décisions sous incertitude, ce qui est courant dans de nombreuses situations réelles, telles que la gestion de la chaîne d'approvisionnement, la gestion des investissements financiers et la logistique, où les paramètres du problème peuvent varier de manière aléatoire.

4.8. Optimisation du réseau

L'*optimisation du réseau* impliquant l'optimisation d'une certaine propriété d'un flux à travers un réseau, comme la maximisation de la quantité de matériel qui peut être transportée entre deux emplacements donnés dans le réseau, ou l'augmentation du débit d'un réseau informatique.

Exemple : Optimisation du réseau de transport de marchandises

Supposons que vous travailliez pour une entreprise de logistique qui doit optimiser le transport de marchandises entre plusieurs entrepôts et destinations. Vous avez un réseau routier ou ferroviaire à votre disposition, et votre objectif est de minimiser les coûts totaux de transport tout en respectant les contraintes de capacité et de temps.

Variables :

- x_{ij} : La quantité de marchandises à transporter de l'entrepôt i à la destination j.
- c_{ij} : Le coût de transport unitaire de l'entrepôt i à la destination j.

Objectif :

- Minimiser les coûts totaux de transport, c'est-à-dire la somme des coûts unitaires de transport pondérés par les quantités transportées.

$$\min \sum_i \sum_j c_{ij} \cdot x_{ij}$$

Contraintes :

- Les quantités transportées ne peuvent pas dépasser la capacité de l'entrepôt i, ce qui se traduit par des contraintes de capacité :$\sum_j x_{ij} \leq Capacité_i$.
- Les quantités transportées ne peuvent pas être négatives : $x_{ij} \geq 0$.
- Les délais de livraison doivent être respectés : $x_{ij} \leq Délai_j$.

Solution :

- Pour résoudre ce problème d'optimisation de réseau de transport, nous pouvonz utiliser des méthodes d'optimisation linéaire, telles que le simplexe ou le programme linéaire en nombres entiers (PLNE) si les variables sont discrètes (comme le nombre de produits transportés). Vous pourriez également utiliser des techniques plus avancées, telles que la programmation linéaire en nombres entiers mixtes (PLNEM) ou des algorithmes d'optimisation basés sur des heuristiques et des métaheuristiques pour les problèmes plus complexes.
- La solution obtenue serait un plan de transport optimal qui détermine la quantité de marchandises à transporter de chaque entrepôt à chaque destination de manière à minimiser les coûts totaux tout en respectant les contraintes.

Ce type d'optimisation de réseau est essentiel dans le domaine de la logistique et du transport pour réduire les coûts, améliorer l'efficacité opérationnelle et réduire l'impact environnemental. Des problèmes similaires d'optimisation de réseau peuvent également être trouvés dans d'autres domaines tels que les réseaux de télécommunications, l'optimisation de la distribution d'énergie électrique et la gestion des réseaux informatiques.

4.9. Optimisation combinatoire

L'*optimisation combinatoire* consiste à chercher la solution optimale (la meilleure solution) parmi un ensemble fini mais très large de valeurs possibles, par exemple, déterminer la meilleure affectation de 20 usines à 20 emplacements, à raison d'une usine par emplacement. Il y a de nombreuses façons possibles d'affecter 20 usines de fabrication à 20 emplacements, leur nombre total est égal à 20!

4.10. Programmation dynamique

La *programmation dynamique* est à la fois une méthode d'optimisation mathématique et une méthode de programmation informatique.

C'est une méthode algorithmique pour résoudre des problèmes d'optimisation. Le concept a été introduit au début des années 1950 par Richard Bellman[3]. À l'époque, le terme «programmation» signifie planification et ordonnancement. La programmation dynamique consiste à résoudre un problème en le décomposant en sous-problèmes, puis à résoudre les sous-problèmes, des plus petits aux plus grands en stockant les résultats intermédiaires. Elle a d'emblée connu un grand succès, car de nombreuses fonctions économiques de l'industrie étaient de ce type, comme la conduite et l'optimisation de procédés chimiques, ou la gestion de stocks.

En termes d'optimisation mathématique, la programmation dynamique fait généralement référence à la simplification d'une décision en la décomposant en une séquence d'étapes de décision dans le temps. Pour ce faire, on définit une séquence de fonctions de valeur $V_1, V_2, ..., V_n$ prenant y comme argument représentant l'état du système aux instants i de 1 à n. La définition de $V_n(y)$ est la valeur obtenue dans l'état y au dernier instant n. Les valeurs V_i aux instants précédents $i = n -1, n - 2, ..., 2, 1$ peuvent être trouvées en travaillant à reculons, à l'aide d'une relation récursive appelée équation de Bellman. *Pour* $i = 2, ..., n$, V_{i-1} à tout état y est calculé à partir de V_i en maximisant une fonction simple (généralement la somme) du gain d'une décision au temps $i - 1$ et de la fonction V_i au nouvel état du système si cette décision est prise. Comme V_i a déjà été calculé pour les états nécessaires, l'opération ci-dessus donne V_{i-1} pour ces états. Enfin, V_1 à l'état initial du système est la valeur de la solution optimale. Les valeurs optimales des variables de décision peuvent être retrouvées, une par une, en retraçant les calculs déjà effectués.

Exemple : Problème du sac à dos

Supposons que vous ayez un sac à dos de capacité maximale C (en poids) et que vous soyez dans un magasin avec n articles différents, chaque article ayant un poids w_i et une valeur v_i. L'objectif est de déterminer quels articles vous devez mettre dans le sac à dos pour maximiser la valeur totale des articles tout en respectant la capacité du sac à dos.

Variables :

[3] Richard Bellman, Dynamic Programming, Princeton, Princeton University Press, 1957. — Réimpression 2003, Dover Publication, Mineola, New-York, (ISBN 0-486-42809-5).

- x_i : Une variable binaire (0 ou 1) indiquant si l'article i est pris ou non.

Objectif :

- Maximiser la valeur totale des articles pris, c'est-à-dire la somme des valeurs pondérées par les variables x_i.

$$\max \sum_{i=1}^{n} v_i \cdot x_i$$

Contraintes :

- La somme des poids des articles pris ne peut pas dépasser la capacité du sac à dos, ce qui se traduit par une contrainte de capacité : $\sum_{i=1}^{n} w_i x_i \leq C$.
- Les variables x_i sont binaires : $x_i \in \{0,1\}$.

Solution par programmation dynamique :

- La programmation dynamique peut être utilisée pour résoudre ce problème en construisant un tableau qui représente les sous-problèmes. Chaque case du tableau stocke la valeur maximale que vous pouvez obtenir avec un sous-ensemble spécifique d'articles et une capacité spécifique du sac à dos.
- L'algorithme de résolution par programmation dynamique commence avec un tableau initialisé à zéro et remplit progressivement chaque case du tableau en utilisant les résultats des sous-problèmes précédents. La valeur maximale se trouve dans la case en bas à droite du tableau.

Voici comment nous pouvons procéder :

1. Créez un tableau à deux dimensions pour représenter les articles et les capacités du sac à dos. Les colonnes du tableau représenteront les capacités du sac à dos allant de 0 à la capacité maximale, et les lignes représenteront les articles.
2. Remplissez la première ligne du tableau avec des zéros, car il n'y a aucun article à placer dans le sac à dos.
3. Pour chaque article, remplissez les colonnes du tableau en comparant la capacité du sac à dos avec le poids de l'article. Si le poids de l'article est supérieur à la capacité, copiez la valeur de la cellule précédente dans la même colonne. Sinon, comparez la valeur de la cellule précédente dans la même colonne avec la somme de la valeur de l'article et de la valeur de la cellule dans la colonne correspondant à la capacité restante après avoir placé l'article.
4. Une fois que vous avez rempli tout le tableau, la valeur maximale se trouvera dans la dernière cellule du tableau.
5. Pour trouver les articles qui ont été placés dans le sac à dos, parcourez le tableau de bas en haut et de droite à gauche. Si la valeur de la cellule est différente de celle de la cellule précédente dans la même colonne, cela signifie que l'article correspondant a été placé dans le sac à dos. Ajoutez cet article à votre liste de solution.
6. Répétez les étapes 3 à 5 pour tous les articles jusqu'à ce que vous ayez parcouru tout le tableau.

7. La liste de solution contiendra les articles qui doivent être placés dans le sac à dos pour obtenir la valeur maximale.

   L'implémentation de cette procédure est illustrée dans le l'encadré suivant :

```
#include <iostream>
#include <vector>
#include <algorithm>
using namespace std;
vector<int> v, c;
vector< vector<int> > DP;
//----------------------------------------------------------
//Fonction pour sélectionner les articles à mettre dans le sac
void pick(int i, int w)
{
if (i <= 0 || w <= 0) return;
int k = DP[i][w];
if (k != DP[i - 1][w]) {
cout << i << " "; // select!
pick(i - 1, w - c[i]); // capacity decreases
} else {
// move on to next item; capacity no change
pick(i - 1, w);
}
}
//-----------------------------------------------------------
int main()
{
int S, N; cin >> S >> N;//N is the number of items
// slotted for easy array access; values won't be used.
c.push_back(-1);
v.push_back(-1);
int st, vt;
for (int i = 0; i < N; i++) {
cin >> st >> vt;
c.push_back(st);
v.push_back(vt);
}
DP.resize(N + 1, vector<int>(S + 1, 0));
for (int i = 1; i <= N; i++) { // i is scope of items in consideration
for (int w = 1; w <= S; w++) { // S is max size of bag
if (c[i] > w) {
DP[i][w] = DP[i - 1][w];
} else {
DP[i][w] = max(DP[i - 1][w], v[i] + DP[i - 1][w - c[i]]);
}
}
}
cout << "Selected items: ";
pick(N, S);
}
```

Par exemple, pour un problème de sac à mains de 4 articles avec une capacité maximale de 7 :

- Article 1 : Valeur = 5, Poids = 3
- Article 2 : Valeur = 7, Poids = 5
- Article 3 : Valeur = 1, Poids = 1
- Article 4 : Valeur = 5, Poids = 4

i	w / [v] c	0	1	2	3	4	5	6	7
0	[-] -	0	0	0	End 0	0	0	0	0
*1	[4] 3	0	0	0	4	4	4	4	4
2	[7] 5	0	0	0	4	4	7	7	7
3	[1] 1	0	1	1	4	5	7	8	8
*4	[5] 4	0	1	1	4	5	7	8	9 Start

Dans cet exemple, nous inclurions les articles 4 et 1 pour obtenir une valeur totale de 9 avec un poids total de 7.

4.11. Optimisation multiobjectif

L'optimisation multiobjectif (appelée aussi programmation multi-objective ou optimisation multicritère) est une branche de l'optimisation mathématique traitant spécifiquement des problèmes d'optimisation ayant plusieurs fonctions objectifs. Elle a été appliquée dans de nombreux domaines scientifiques, notamment l'ingénierie, l'économie et la logistique, où des décisions optimales doivent être prises en présence de compromis entre deux ou plusieurs objectifs contradictoires. Minimiser le coût tout en maximisant le confort lors de l'achat d'une voiture, et maximiser les performances tout en minimisant la consommation de carburant et les émissions de polluants d'un véhicule sont des exemples de problèmes d'optimisation multi-objectifs impliquant deux et trois objectifs, respectivement. Dans les problèmes pratiques, il peut y avoir plus de trois objectifs.

5. Formulation du problème

<u>**Exemple I.1**</u>: Un fabricant de produits chimiques confectionne un produit chimique en combinant deux matières premières, M1 et M2. Bien que M_1 puisse être acheté à 5 $ la tonne, M_2 est moins cher et peut être obtenu à 1 $ la tonne. Le fabricant souhaite déterminer la quantité de chaque matière première nécessaire pour réduire au minimum le coût par tonne de produit. Formulez le problème comme un problème d'optimisation.

Solution : Soit x_1 la quantité de M_1 nécessaire pour produire une tonne du produit et x_2 la quantité de M_2 consommée dans le processus. Le problème peut alors être formulé comme suit.

$$Minimiser\ Z = 5x_1 + x_2 \qquad (I.6)$$

$sujet$ à:

$$x_1 \geq 0\ ;\ x_2\ \geq 0$$

À partir de la fonction objectif donnée par l'équation (I.6), il est évident que pour minimiser Z, le coût par tonne de produit fabriqué, le fabricant doit acheter des quantités nulles de M_1 et M_2.

L'*exemple I.1* concerne la fabrication d'un produit chimique à partir des réactifs M_1 et M_2 et nous avons constaté que la solution optimale implique des quantités nulles de M_1 et M_2. Il s'agit d'une solution mathématiquement correcte. Cependant, d'un point de vue réaliste (thermodynamique), si le réactif M_1 n'est pas du tout présent, il n'y a aucune possibilité de fabrication de produit. Qu'avons-nous fait de mal ? Nous avons oublié de fournir des informations sur la quantité minimale de réactifs nécessaire pour former le produit. Comme la formulation du problème n'a pas donné d'informations sur les exigences de la chimie de la réaction, notre solution est mathématiquement correcte mais pratiquement inutile. Cela montre que la formulation correcte du problème est l'étape clé de l'optimisation.

6. Analyse des degrés de liberté

L'analyse des degrés de liberté fournit le nombre de variables de décision que l'on peut modifier pour obtenir la conception optimale, et elle est cruciale dans l'optimisation. Considérons le cas 1 du tableau I.1 ci-dessous. On peut voir qu'il y a deux variables et deux équations, et donc qu'il n'y a pas de problème d'optimisation car les deux ensembles d'équations $D-2$ et $D-3$ peuvent être résolus directement pour obtenir les valeurs des variables de décision x_1 et x_2. L'équation $D-1$ ne joue aucun rôle dans la recherche des valeurs de ces variables. D'autre part, le cas 2 a deux variables et une égalité, et donc un seul degré de liberté. Le cas 3 a deux variables et deux inégalités, ce qui conduit à deux degrés de liberté. Dans les cas 2 et 3, l'optimisation peut être utilisée car le nombre correspondant aux degrés de liberté est supérieur à zéro.

Est-il possible d'utiliser l'optimisation pour un système lorsqu'il y a moins de variables que d'équations ?

Il s'agit d'une situation typique dans le problème d'estimation des paramètres présenté ci-dessous :

$$Y = f(x, \theta) \tag{I.7}$$

Ici, la quantité Y est une fonction des variables x et θ, où θ représente m paramètres inconnus. En général, n expériences sont réalisées pour obtenir diverses valeurs de Y en changeant les valeurs de x. Chaque ensemble de données doit satisfaire l'équation ci-dessus, ce qui donne lieu aux n équations et $n > m$. Dans ce cas, on peut encore utiliser des techniques d'optimisation en combinant les n équations en une fonction objectif contenant le carré des erreurs pour chaque équation.

Le problème des moindres carrés garantit qu'il y a des degrés de liberté disponibles pour résoudre le problème d'optimisation.

Case 1	Case 2	Case 3	Equation
Min $x_1^2 + x_2^2$ $x_1 - x_2 = 4.0$ $x_1 + x_2 = 2.0$	Min $x_1^2 + x_2^2$ $x_1 + x_2 = 2.0$ $-$	Min $x_1^2 + x_2^2$ $x_1 - x_2 \leq 4.0$ $x_1 + x_2 \leq 2.0$	$D-1$ $D-2$ $D-3$
0	1	2	DOF
No	Yes	Yes	Optimization
$x_1 = 3;\ x_2 = -1$	$x_1 = 1;\ x_2 = 1$	$x_1 = 0;\ x_2 = 0$	Solution

Tableau I.1. Analyse des degrés de liberté: DOF (Degrees Of Freedom).

7. Fonction objectif, Contraintes, et Région Réalisable

La figure 1.3 trace le graphique de la fonction objectif Z en fonction d'une variable de décision x. La figure I.3(a) présente un problème de programmation linéaire dont la fonction objectif linéaire ainsi que ses contraintes (lignes AB, BC et CA) sont linéaires. Les contraintes illustrées dans la figure I.3(a) sont des contraintes d'inégalité qui indiquent que la solution doit se trouver au-dessus ou sur la ligne AB, et au-dessous ou sur les lignes BC et CA. ABC représente la région d'opération réalisable dans laquelle la solution doit se situer. Les contraintes lient l'espace objectif, et de ce fait l'objectif linéaire se trouve à la limite de la région réalisable (contrainte). La figure I.3(b) montre un problème sans contrainte, la région réalisable s'étend donc à l'infini. Le minimum de la fonction objectif se situe au point B, où la tangente à la courbe est parallèle à l'axe x, avec une pente nulle (la dérivée de la fonction objectif par rapport à l'axe de décision). C'est l'une des conditions nécessaires d'optimalité pour un problème de programmation non linéaire où la fonction objectif et/ou les contraintes sont non linéaires. Les théories antérieures impliquant le calcul différentiel utilisent cette condition d'optimalité pour atteindre de préférence une solution analytique du problème. Cependant, pour de nombreux problèmes du monde réel, il est difficile d'obtenir une solution analytique et il faut suivre un *schéma itératif*. C'est *l'optimisation numérique*.

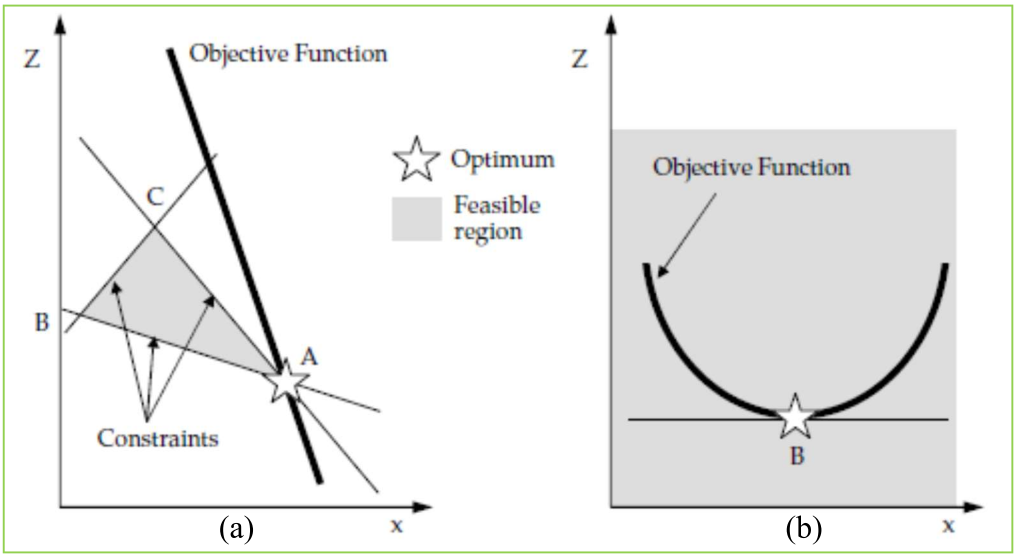

Figure I.3: Problèmes de programmation linéaire et non linéaire.

8. Optimisation numérique

Un problème d'optimisation général peut être formulé comme suit:

$$Optimiser\ Z = z(x) \qquad (I.8)$$
sujet à:
$$h(x) = 0 \qquad (1.9)$$
$$g(x) \leq 0 \qquad (1.10)$$

Le but d'un problème d'optimisation est de déterminer les variables de décision x qui optimisent la fonction objective Z, tout en garantissant que le modèle fonctionne dans les limites établies, imposées par les contraintes d'égalité h (équation (I.9)) et les contraintes d'inégalité g (équation (I.10)).

La figure I.4 illustre de manière schématique la procédure itérative employée dans une technique d'optimisation numérique. Comme on peut le voir sur la figure, l'optimiseur invoque le modèle avec un ensemble de valeurs de variables de décision x. Le modèle simule les phénomènes et calcule la fonction objectif et les contraintes. Ces informations sont utilisées par l'optimiseur pour calculer un nouvel ensemble de variables de décision. Cette séquence itérative est poursuivie jusqu'à ce que les critères d'optimisation relatifs à l'algorithme d'optimisation soient satisfaits.

Il existe un grand nombre de codes logiciels disponibles pour l'optimisation numérique. Il s'agit par exemple de solveurs tels que MINOS, CPLEX, CONOPT et NPSOL. De même, de nombreuses bibliothèques mathématiques, telles que NAG, OSL, IMSL et HARWELL, intègrent différents codes d'optimisation. Les progiciels populaires tels qu'EXCEL, MATLAB et SAS ont également des capacités d'optimisation. Il existe des langages de modélisation algébrique tels que AMPL, LINGO, AIMMS, GAMS et ISIGHT spécifiquement conçus pour résoudre les problèmes d'optimisation et des logiciels tels que Omega et Evolver disposent d'interfaces de tableur. En outre, l'Internet constitue une excellente source d'information. Un groupe de chercheurs du Laboratoire national d'Argonne et de l'Université du Nord-Ouest a lancé un projet appelé NEOS (Network-Enabled Optimization System). Son service web pour l'optimisation numérique est accessible à l'adresse Web suivante: https://neos-server.org/neos/.

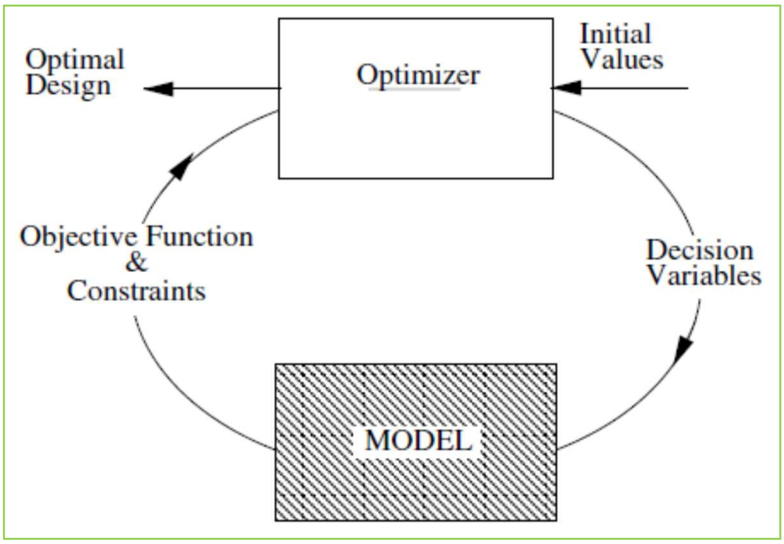

Figure I.4: Illustration schématique du cadre d'optimisation numérique.

9. Model d'optimisation

La modélisation mathématique est la première étape dans le processus d'optimisation**.** Un *modèle* représente les caractéristiques essentielles d'un objet, d'un système ou d'un problème sans les détails non importants. Dans les modèles d'optimisation, les aspects importants sont représentés sous forme mathématique à l'aide de variables, de paramètres et de fonctions. L'analyse et la manipulation du modèle permettent de voir comment le système réel se comporte dans diverses conditions. A partir de là, nous déterminons la meilleure conception du système ou la meilleure action à entreprendre.

Les modèles mathématiques sont moins chers, plus rapides et plus sûrs que la construction et la manipulation de systèmes réels. Supposons que l'on veuille trouver le mélange de déchets de papiers recyclés à utiliser pour produire un type de carton qui minimise le coût. Une entreprise pourrait essayer plusieurs combinaisons différentes, vérifier la qualité et calculer le coût. Comme toutes les combinaisons possibles ne sont pas essayées, la combinaison optimale ne sera probablement pas trouvée. Sinon, à l'aide d'un modèle mathématique, on évalue toutes les combinaisons possibles pour trouver celle qui satisfait aux spécifications du produit au moindre coût. La modélisation mathématique est plus rapide et moins coûteuse que l'approche par essais et erreurs.

Les problèmes de l'emplacement d'installations, d'acheminement et d'ordonnancement des véhicules, d'ordonnancement du personnel, des machines et des tâches, d'assortiment de produits et de gestion des stocks sont formulés sous forme de modèles d'optimisation sous contraintes. Les modèles d'optimisation sous contraintes sont des modèles mathématiques qui trouvent la meilleure solution par rapport à un critère d'évaluation parmi un ensemble de solutions alternatives. Ces solutions sont définies par un ensemble de contraintes mathématiques – des inégalités ou des égalités mathématiques.

9.1. Modèles d'optimisation sous contraintes

Les modèles d'optimisation sous contrainte comportent trois composantes principales : les variables de décision, la fonction objectif et les contraintes.

1. **Les variables de décision** sont des quantités physiques contrôlées par le décideur et représentées par des symboles mathématiques. Par exemple, la variable de décision x_j peut représenter le nombre de tonnes du produit *j* qu'une entreprise produira au cours d'un mois donné. Les variables de décision peuvent prendre n'importe quelle valeur parmi un ensemble de valeurs possibles.
2. **La fonction objectif** définit le critère d'évaluation de la solution. Il s'agit d'une fonction mathématique des variables de décision qui convertit une solution en une évaluation numérique de cette solution. Par exemple, la fonction objectif peut mesurer le bénéfice ou le coût qui se produit en fonction des quantités de divers produits fabriqués. La fonction objectif spécifie également la tendance de l'optimisation, soit la maximisation, soit la minimisation. Une solution optimale pour le modèle est la meilleure solution telle que mesurée par ce critère.
3. **Les contrainte**s sont un ensemble d'égalités ou d'inégalités fonctionnelles qui représentent des restrictions physiques, économiques, technologiques, juridiques, éthiques ou autres sur les valeurs numériques qui peuvent être attribuées aux variables de décision. Par exemple, les contraintes peuvent garantir que l'on n'utilise pas plus de ressources que ce celles qui sont disponibles. Les contraintes peuvent être définitionnelles, définissant le nombre d'employés au début d'une période $t + 1$ comme égal au nombre d'employés au début de la période t, plus ceux ajoutés pendant la période t moins ceux qui quittent l'organisation pendant la période t. Dans les modèles d'optimisation avec contraintes, nous trouvons des valeurs pour les variables de décision qui maximisent ou minimisent la fonction objectif et satisfont toutes les contraintes.

Exemple I.2: Gamme de produits de la société Healthy Pet Food

L'exemple suivant montre comment un problème opérationnel peut être représenté et analysé à l'aide d'un modèle d'optimisation avec contraintes.

La société Healthy Pet Food fabrique deux types d'aliment pour chiens: Meaties et Yummies. Chaque paquet de Meaties contient 2 kg de céréales et 3 kg de viande ; chaque paquet de Yummies contient 3 Kg de céréales et 1.5 Kg de viande. Healthy pense pouvoir vendre autant de chaque aliment pour chiens qu'elle peut en fabriquer. Les Meaties se vendent 2.80 $ par paquet et les Yummies 2 $ par paquet. La production de Healthy est limitée de plusieurs façons. Premièrement, Healthy ne peut acheter qu'un maximum de 400 000 Kg de céréales par mois à 0.40 $ le kilogramme. Elle ne peut acheter que jusqu'à 300 000 Kg de viande par mois à 1 $ la livre. De plus, une machine spéciale est nécessaire pour fabriquer les Meaties, et cette machine a une capacité de 90 000 paquets par mois. Le coût variable du mélange et de l'emballage de la nourriture pour chiens est de 0.25 $ par emballage pour les Meaties et de 0.20 $ par emballage pour les Yummies. Ces informations sont présentées dans le tableau I.2.

Supposons que vous soyez le directeur de la division des aliments pour chiens de la société Healthy Pet Food. Votre salaire est basé sur le profit de la division, vous essayez donc de maximiser son profit. Comment devez-vous gérer la division pour maximiser son bénéfice et votre salaire ?

Solution:

1) Les variables de décision

Nous commençons par identifier les éléments sur lesquels nous avons un contrôle: *les variables de décision*. Dans ce problème, nous avons un contrôle direct sur deux quantités : le nombre de paquets Meaties à fabriquer chaque mois et le nombre de paquets de Yummies à fabriquer chaque mois. Dans le modèle, ces deux quantités apparaissent de manière répétée, nous les représentons donc de manière simple. Nous désignons ces variables par les symboles M et Y.

M = nombre de paquets de "Meaties" à fabriquer chaque mois

Y = nombre de paquets de Yummies à fabriquer chaque mois

	Meaties	**Yummies**
Prix de vente par paquet	2.80$	2.00$
Matières premières par paquet		
Céréales	2 Kg	3 Kg
Viande	3 Kg	1.5 Kg
Coût variable – mélange et emballage	0.25$ paquet	0.20$ paquet
Ressources		
Capacité de production pour Maties	90 000 paquets par mois	
Céréales disponibles par mois	400 000 Kg	
Viande disponible par mois	300 000 Kg	

Tableau I.2: Les données de Healthy Pet Food.

Notez que les quantités de viande et de céréales utilisées chaque mois ne sont pas de bons choix pour les variables. Tout d'abord, nous ne les contrôlons qu'indirectement par le biais de notre choix de M et Y. Plus important encore, leur utilisation comme variables pourrait conduire à des plans de production ambigus. Déterminer la quantité de céréales et de viande à utiliser dans la production ne nous indique pas comment les utiliser – quelle quantité de chaque aliment pour chiens fabriquer. En revanche, après avoir déterminé les valeurs de M et Y, nous savons ce qu'il faut produire et quelle quantité de viande et de céréales est nécessaire.

2) La fonction objectif

Toute paire de valeurs numériques pour les variables M et Y constitue un plan de production. Par exemple, $M = 10\ 000$ et $Y = 20\ 000$ signifie que nous fabriquons chaque mois 10 000 paquets de "Meaties" et 20 000 paquets de "Yummies". Mais comment savoir s'il s'agit d'un bon plan de production? Nous devons définir un critère d'évaluation – une fonction objectif. La fonction objectif la plus appropriée est de maximiser le profit mensuel. (En fait, il s'agit de la contribution au profit : les coûts fixes sont ignorés car tout plan qui maximise les recettes moins les coûts variables maximise également le profit). Le bénéfice réalisé par Healthy est une fonction directe de la quantité de chaque aliment pour chiens fabriqué et vendu, les variables de décision. Le bénéfice mensuel, désigné par z, s'écrit comme suit:

z = (Bénéfice par paquet de Meaties) × (Nombre de paquets de Meaties fabriqués et vendus par mois) + (Bénéfice par paquet de Yummies) × (Nombre de paquets de Yummies fabriqués et vendus par mois)

Le bénéfice par emballage pour chaque aliment pour chiens est calculé comme suit:

	Meaties	Yummies
Prix de vente	2.80	2.00
Moins		
Viande	1.50	0.75
Céréales	0.40	0.60
Mélange et Emballage	0.25	0.20
Bénéfice par paquet	0.65	0.45

Nous écrivons le bénéfice mensuel comme suit:

$$z = 0.65M + 0.45Y$$

3) Contraintes

Si nous voulons rendre z aussi grand que possible, pourquoi ne pas rendre M et Y égaux à l'infini et réaliser un profit infini ? Nous ne pouvons pas le faire car il existe des limites à la disponibilité des céréales et de la viande et à la capacité de production de Meaties. (En réalité, il existe également une limite à la demande, mais nous l'ignorons ici pour des raisons de simplicité). Nous voulons maximiser z, mais sous réserve de satisfaire les contraintes énoncées. Pour résoudre le problème, nous exprimons ces contraintes sous forme d'égalités ou d'inégalités mathématiques. Commençons par la contrainte de disponibilité des céréales:

- (Le nombre kilogrammes de céréales utilisées dans la production chaque mois) ≤ 400 000 Kg.
- Le côté gauche (c.g.a.) de la contrainte est déterminé par le nombre de paquets de Meaties et de Yummies fabriqués. Plus précisément, le c.g.a est:

  (Quantité en kg de céréales par paquet de Meaties) × (Nombre de paquets de Meaties fabriqués et vendus mensuellement) + (Quantité en kg de céréales par paquet de Yummies) × (Nombre paquets de Yummies fabriqués et vendus mensuellement).
- En substituant la teneur en céréales de chaque produit et les variables de décision dans cette expression, nous écrivons la contrainte comme suit

  $2M + 3Y \leq 400\,000$
- En utilisant un raisonnement similaire, la restriction sur la disponibilité de la viande s'exprime comme suit .

  $3M + 1.5Y \leq 300\,000$
- En plus de ces contraintes, le nombre de paquets de Meaties produits chaque mois ne peut pas dépasser 90 000; c'est-à-dire,
  $M \leq 90\,000$
- Enfin, les niveaux de production négatifs n'ont pas de sens, nous exigeons donc que M ≥ 0 et Y ≥ 0.

En mettant tout cela ensemble, on obtient le modèle d'optimisation sous contrainte suivant:

Maximizer $z = 0.65M + 0.45Y$
Sujet à:
$2M + 3Y \leq 400\,000$
$3M + 1.5Y \leq 300\,000$
$M \leq 90\,000$

$M, Y \geq 0$

Ce type de modèle est appelé modèle de *programmation linéa*ire ou *programme linéaire* car la fonction objectif est linéaire et les fonctions de toutes les contraintes sont linéaires. La solution optimale pour le problème des aliments pour animaux de la sociétée Healthy est $M = 50\ 000$, $Y = 100\ 000$ et $z = 77\ 500\ \$$. Autrement dit, Healthy doit fabriquer 50 000 paquets de Meaties et 100 000 paquets de Yummies chaque mois, ce qui lui permettra de réaliser un bénéfice mensuel de 77 500 $.

9.2. Avantages et inconvénients de l'utilisation de modèles d'optimisation

Le principal avantage des modèles d'optimisation est la possibilité d'évaluer les solutions possibles de manière rapide, sûre et peu coûteuse sans avoir à les construire et à les expérimenter. Les autres avantages sont les suivants:

- ***Structurer le processus de réflexion***

  La construction d'un modèle d'optimisation d'un problème oblige le décideur à réfléchir au problème d'une manière concise et organisée. Le décideur détermine les facteurs qu'il contrôle, c'est-à-dire les variables de décision. Le décideur précise comment la solution sera évaluée (la fonction objectif). Enfin, le décideur décrit l'environnement de la décision (les contraintes). La modélisation permet d'organiser et de clarifier le problème.

- ***Augmenter l'objectivité***

  Les modèles mathématiques sont plus objectifs puisque toutes les hypothèses et tous les critères sont clairement spécifiés. Bien que les modèles reflètent les expériences et les préjugés de ceux qui les construisent, ces préjugés peuvent être identifiés par des observateurs extérieurs. En utilisant un modèle comme point de référence, les personnes concernées peuvent concentrer leurs discussions et leurs désaccords sur ses hypothèses et ses composantes. Une fois que l'on s'est mis d'accord sur le modèle, les parties prenantes (i.e., les personnes concernées) ont tendance à s'en tenir aux résultats.

- ***Rendre les problèmes complexes plus faciles à traiter***

  De nombreux problèmes liés à la gestion d'une organisation sont larges et complexes et traitent d'interrelations subtiles, mais importantes, entre les unités organisationnelles. Par exemple, pour déterminer les quantités optimales de divers produits à expédier d'entrepôts géographiquement dispersés à des clients géographiquement dispersés et les itinéraires à emprunter, l'esprit humain ne peut pas faire les milliards de compromis simultanés nécessaires. Dans ces cas, le décideur utilise souvent de simples règles empiriques, qui peuvent aboutir à des solutions non optimales. Les modèles d'optimisation facilitent la résolution de problèmes complexes à l'échelle d'une organisation.

- ***Rendre les problèmes accessibles à une solution mathématique et informatique***

  En représentant un problème réel sous la forme d'un modèle mathématique, nous utilisons des techniques de solution et d'analyse mathématiques et des ordinateurs d'une manière qui ne serait pas possible autrement.

- ***Facilite l'analyse "what if"***

  Les modèles mathématiques permettent de trouver de manière relativement facile la solution optimale pour un modèle et un scénario spécifiques. Ils facilitent également l'analyse "what if". Avec l'analyse "what if", nous reconnaissons que les prix, les demandes et les disponibilités de produits supposés lors de la construction du modèle sont simplement des estimations et peuvent différer dans le temps peuvent différer dans la pratique. Par conséquent, nous voulons savoir comment la solution optimale évolue lorsque la valeur de ces paramètres varie par rapport aux estimations initiales. Autrement dit, nous voulons connaître la sensibilité de la solution optimale aux hypothèses du modèle. L'analyse "what if" est également appelée ***analyse de sensibilité*** ou ***analyse paramétrique***.

Bien que la modélisation mathématique présente de nombreux avantages, elle comporte également des inconvénients. La formulation ou la construction du modèle est l'étape la plus cruciale de la modélisation mathématique. Comme les problèmes ont tendance à être très complexes, il est possible de mal modéliser le problème réel. Des variables de décision ou des relations importantes peuvent être omises ou le modèle peut être inadapté à la situation. La solution optimale d'un mauvais problème n'a aucune valeur.

Un deuxième inconvénient est de ne pas comprendre le rôle de la modélisation dans le processus décisionnel. La solution optimale pour un modèle n'est pas nécessairement la solution optimale pour le problème réel. Les modèles mathématiques sont des outils qui nous aident à prendre de bonnes décisions. Toutefois, ils ne sont pas le seul facteur à prendre en compte dans la décision finale. Parfois, le modèle n'évalue les solutions que par rapport à des critères quantitatifs. Dans ce cas, des facteurs qualitatifs (par ex. maturité de l'expérience) doivent également être pris en compte lors de la prise de décision finale.

L'essentiel pour évaluer un modèle est de savoir s'il aide ou non le décideur à identifier et à mettre en œuvre de meilleures solutions. Le modèle doit renforcer la confiance du décideur dans sa décision et sa volonté de la mettre en œuvre.

10. Recherche opérationnelle

10.1. Définitions de la recherche opérationnelle (RO)

Il existe plusieurs définitions pour la recherche opérationnelle, on peut en citer quelques-unes:

1. La recherche opérationnelle (RO), appelée aussi "aide à la décision", est une discipline dont le but est d’aider les gestionnaires à prendre des décisions dans des situations complexes grâce à l’utilisation de méthodes scientifiques, en particulier de modèles mathématiques.
2. L'étude systématique de la meilleure façon de résoudre les problèmes dans le monde des affaires et industriel**.**
3. C'est l'application de méthodes scientifiques à la gestion et à l'administration de processus militaires, gouvernementaux, commerciaux et industriels organisés.
4. La recherche opérationnelle est une collection de techniques et méthodes, issues du champ des mathématiques appliquées, destinées à représenter des situations où un ou plusieurs acteurs ont un certain nombre de choix à effectuer, et à guider ces acteurs dans leur décision de façon à ce qu'ils satisfassent au mieux un ou plusieurs critères tout en respectant un ensemble de contraintes prédéfinies.
5. L recherche opérationnelle est l'ensemble des méthodes et techniques rationnelles d'analyse et de synthèse des phénomènes d'organisation utilisables pour élaborer de meilleures décisions, chaque fois que des hommes, des machines, des produits se trouvent en relations actives, on dira qu'on a affaire à un phénomène d'organisation.
6. La recherche opérationnelle peut être définie comme l'utilisation de méthodes quantitatives pour aider les analystes et les décideurs à concevoir, analyser et

améliorer la performance ou le fonctionnement des systèmes. Les systèmes étudiés peuvent être des systèmes financiers, des systèmes scientifiques ou d'ingénierie, ou des systèmes industriels; mais quel que soit le contexte, pratiquement tous ces systèmes se prêtent à un examen dans le cadre systématique de la méthode scientifique.

7. La recherche opérationnelle est une discipline exploitant ce qu'il y a de plus opérationnel dans les mathématiques, l'économie et l'informatique. Elle est en prise directe avec l'industrie et joue un rôle-clé dans le maintien de la compétitivité.

8. La RO est une approche quantitative permettant de produire de meilleures décisions. Elle fournit des outils pour rationaliser, simuler et optimiser l'architecture et le fonctionnement des systèmes industriels et économiques. Elle propose des modèles pour analyser des situations complexes et permet aux décideurs de faire des choix efficaces et robustes.

10.2. Les aspects fondamentaux de la RO

La RO tente de fournir à ceux qui gèrent des systèmes organisés une base objective et quantitative pour la prise de décision; elle est normalement menée par des équipes de scientifiques et d'ingénieurs issus de diverses disciplines. Ainsi, la recherche opérationnelle n'est pas une science en soi mais plutôt l'application de la science à la résolution de problèmes de gestion et d'administration, et elle se concentre sur la performance des systèmes organisés pris dans leur ensemble plutôt que sur leurs parties prises séparément. Généralement concernée par les systèmes dans lesquels le comportement humain joue un rôle important, la RO diffère à cet égard de l'ingénierie des systèmes qui, en utilisant une approche similaire, tend à se concentrer sur les systèmes dans lesquels le comportement humain n'est pas important. À l'origine, la recherche opérationnelle s'intéressait à l'amélioration du fonctionnement des systèmes existants plutôt qu'au développement de nouveaux systèmes; l'inverse était vrai pour l'ingénierie des systèmes. Cette différence s'est toutefois estompée au fur et à mesure de la maturation des deux domaines.

Le sujet de la recherche opérationnelle consiste en des décisions qui contrôlent le fonctionnement des systèmes. Elle s'intéresse donc à:

- La manière dont les décisions de gestion sont et devraient être prises;
- La manière d'acquérir et de traiter les données et les informations nécessaires pour prendre des décisions de manière efficace;
- La manière de contrôler les décisions une fois qu'elles sont mises en œuvre;
- La manière d'organiser le processus de prise de décision et de mise en œuvre des décisions.

Elle fait largement appel à des disciplines plus anciennes telles que la logique, les mathématiques et les statistiques, ainsi qu'à des développements scientifiques plus récents tels que la théorie des communications, la théorie de la décision, la cybernétique, la théorie des organisations, les sciences du comportement et la théorie générale des systèmes.

Au XIXe siècle, la révolution industrielle a impliqué la mécanisation ou le remplacement de l'homme par la machine comme source de travail physique. L'étude et l'amélioration de ce type de travail ont constitué la base du domaine du génie industriel. De nombreuses questions contemporaines concernent l'automatisation ou la mécanisation du travail mental. Les principales technologies concernées sont la mécanisation de la génération de symboles (observation par des machines telles que le radar et le sonar), la mécanisation de la transmission de symboles (communication par téléphone, radio et télévision) et la mécanisation de la manipulation logique des symboles (traitement des données et prise de décision par ordinateur). La recherche opérationnelle applique la méthode scientifique à l'étude du travail mental et fournit les connaissances et la compréhension nécessaires à l'utilisation efficace du personnel et des machines pour le réaliser.

10.3. Historique

Les premières activités officielles de la recherche opérationnelle (RO) ont été lancées en Angleterre pendant la seconde guerre mondiale, lorsqu'une équipe de scientifiques britanniques a entrepris d'évaluer la meilleure utilisation du matériel de guerre en se fondant sur des principes scientifiques plutôt que sur des règles *ad hoc* (*spécifiques*). C'est pour cette raison qu'il est souvent considéré que la recherche opérationnelle, en tant que technique récente d'aide à la décision, date tout au plus de la seconde guerre mondiale. Et, en fait, c'est bien à son application aux opérations militaires qu'elle doit son nom.

En réalité, elle est bien plus ancienne, car, dès le dix-septième siècle, B. Pascal et P. de Fermât, inventeurs de la notion d'espérance mathématique (1654), cherchaient, suivis de peu par Jacques Bernoulli et Waldegrave, à résoudre dès problèmes de décision dans l'incertain. Avant la fin de l'ancien régime, Monge s'était proposé et avait résolu analytiquement un problème économique de nature combinatoire: celui des déblais et remblais (1776). Sous la monarchie de Juillet, Augustin Cournot s'était attaqué à la théorie mathématique des richesses (1838), devenant ainsi le précurseur de 1'économétrie. Plus récemment, Emile Borel introduisait la théorie mathématique des jeux, sous sa forme moderne, à l'Académie des Sciences (1921-25), tandis qu'Erlang fondait celle des files d'attente, qu'il utilisait à la conception des réseaux téléphoniques (1917). Enfin, à la veille de la guerre 1939-45, L. Kantorovitch concevait et appliquait la programmation linéaire à la planification peu après que Kônig se fut intéressé aux graphes (1936).

En un sens, tous les efforts déployés pour appliquer la science à la gestion des systèmes organisés, et à leur compréhension, ont été les prédécesseurs de la RO. Cette discipline distincte a toutefois vu le jour en 1937 en Grande-Bretagne, à l'initiative d'A.P. Rowe, directeur de la station de recherche de Bawdsey, qui a amené des scientifiques britanniques à enseigner aux chefs militaires comment utiliser le radar, alors nouvellement développé, pour localiser les avions ennemis. En 1939, la Royal Air Force a officiellement commencé à s'efforcer d'étendre la portée des équipements radar afin d'augmenter le temps entre le premier avertissement fourni par le radar et l'attaque de l'avion ennemi. Ils ont d'abord analysé les équipements physiques et les réseaux de communication, puis ils ont examiné le comportement du personnel d'exploitation et des cadres concernés. Les résultats de ces études ont révélé des moyens d'améliorer les techniques des opérateurs et ont également mis en évidence des limites insoupçonnées du réseau.

Le physicien anglais Blackett fut, en 1940, appelé à présider la première équipe de chercheurs opérationnels et eut l'immense mérite de trouver l'organisation lui permettant de traiter rapidement et avec succès la difficile question de l'implantation optimale des radars de surveillance britanniques, qui devaient jouer un rôle si déterminant dans la bataille d'Angleterre. On sait qu'il attribua lui-même l'efficacité de son entreprise aux faits que l'équipe qu'il avait rassemblée était très hétérogène (autrement dit les points de vue qui s'y manifestèrent étaient très variés), qu'aucune information (même secrète) ne fut jugée trop noble pour échapper à sa compétence et qu'enfin il réserva la décision à ses commettants.

Un développement de la recherche opérationnelle parallèle à celui de la Grande-Bretagne a eu lieu en Australie, au Canada, en France et, plus important encore pour les développements futurs, aux États-Unis, qui ont bénéficié d'un certain nombre de contacts avec les chercheurs britanniques. Sir Robert Watson-Watt, qui, avec A.P. Rowe, a lancé les deux premières études opérationnelles du radar en 1937 et qui prétend avoir donné son nom à cette discipline, s'est rendu aux États-Unis en 1942 et a insisté pour que la recherche opérationnelle soit introduite dans les départements de la guerre et de la marine. Des rapports sur les travaux britanniques avaient déjà été envoyés de Londres par des observateurs américains, et James B. Conant, alors président du National Defense Research Committee, avait été sensibilisé à la recherche opérationnelle lors d'une visite en Angleterre dans la seconde moitié de 1940. Un autre stimulant fut le mémorandum de Blackett, "Scientists at the Operational Level", de décembre 1941, qui fut largement diffusé dans les départements des services américains.

La première activité organisée de recherche opérationnelle aux États-Unis a débuté en 1942 au Naval Ordnance Laboratory. Ce groupe, qui s'occupait des problèmes de guerre des mines, a ensuite été transféré au département de la marine, d'où il a conçu le blocus minier aérien de la mer intérieure du Japon.

Comme en Grande-Bretagne, le radar a stimulé les développements dans l'armée de l'air américaine. En octobre 1942, tous les commandements de l'armée de l'air ont été invités à inclure des groupes de recherche opérationnelle dans leurs états-majors. À la fin de la Seconde Guerre mondiale, l'armée de l'air comptait 26 groupes de ce type. En 1943, le général George Marshall a suggéré à tous les commandants de théâtre de former des équipes pour étudier les opérations amphibies et terrestres.

À la fin de la Seconde Guerre mondiale, un certain nombre de travailleurs britanniques spécialisés dans la recherche opérationnelle ont rejoint le gouvernement et l'industrie. La nationalisation de plusieurs industries britanniques a été un facteur important. L'un des premiers groupes industriels a été créé au National Coal Board. L'électricité et les transports, deux industries nationalisées, ont commencé à utiliser la recherche opérationnelle peu de temps après. Certaines parties du secteur privé ont commencé à suivre le mouvement, en particulier dans les industries dotées d'associations de recherche coopérative, par exemple la British Iron and Steel Research Association.

Le développement initial de la recherche opérationnelle industrielle a été prudent, et pendant quelques années, la plupart des groupes industriels étaient assez petits. À la fin des années 1950, largement stimulé par les développements aux États-Unis, le développement de la recherche opérationnelle industrielle en Grande-Bretagne s'est considérablement accéléré.

Bien qu'aux États-Unis la recherche militaire ait augmenté à la fin de la guerre et que les groupes aient été élargis, ce n'est qu'au début des années 1950 que l'industrie américaine a commencé à prendre au sérieux la recherche opérationnelle. L'avènement de l'ordinateur a fait prendre conscience d'une foule de problèmes systémiques généraux et de la possibilité de les résoudre, et au cours de la décennie, environ la moitié des grandes entreprises des États-Unis ont commencé à utiliser la recherche opérationnelle. Ailleurs, la technique s'est également répandue dans l'industrie.

Des sociétés ont été organisées, à commencer par l'Operational Research Club of Britain, créé en 1948, qui est devenu en 1954 l'Operational Research Society. La Operations Research Society in America a été créée en 1952. De nombreuses autres sociétés nationales sont apparues ; la première conférence internationale sur la recherche opérationnelle s'est tenue à l'université d'Oxford en 1957. En 1959, une Fédération internationale des sociétés de recherche opérationnelle a été créée.

La recherche opérationnelle a fait sa première apparition en tant que discipline universitaire en 1948, lorsqu'un cours sur les techniques non militaires a été introduit au Massachusetts Institute of Technology de Cambridge. En 1952, un programme d'études menant à une maîtrise et à un doctorat a été établi au Case Institute of Technology (aujourd'hui Case Western Reserve University) à Cleveland. Depuis lors, de nombreux établissements universitaires importants des États-Unis ont mis en place des programmes. Au Royaume-Uni, des cours ont été lancés à l'Université de Birmingham au début des années 1950. La première chaire de recherche opérationnelle a été créée à la toute nouvelle université de Lancaster en 1964. Des développements similaires ont eu lieu dans la plupart des pays dans lesquels une société nationale de recherche opérationnelle existe.

La première revue savante, Operational Research Quarterly, publiée au Royaume-Uni, a été lancée en 1950 ; en 1978, son nom a été changé en Journal of the Operational Research Society. Elle a été suivie en 1952 par le Journal of the Operations Research Society of America, qui a été rebaptisé Operations Research en 1955. La Fédération internationale des sociétés de recherche opérationnelle a lancé les International Abstracts in Operations Research en 1961.

Malgré sa croissance rapide, la recherche opérationnelle est encore une activité scientifique relativement jeune. On peut s'attendre à ce que ses techniques et méthodes, ainsi que les domaines auxquels elles sont appliquées, continuent de se développer rapidement. La majeure partie de son histoire se situe dans le futur.

10.4. Les caractéristiques principales de la RO

Trois caractéristiques essentielles de la recherche opérationnelle sont une orientation-système, le recours à des équipes interdisciplinaires et l'application de la méthode scientifique aux conditions dans lesquelles la recherche est menée.

10.4.1. L'orientation-système

L'approche systémique des problèmes reconnaît que le comportement de toute partie d'un système a un certain effet sur le comportement du système dans son ensemble. Cependant, même si les composants individuels fonctionnent bien, le système dans son ensemble ne fonctionne pas nécessairement aussi bien. Par exemple, si l'on assemble les meilleures pièces de chaque type d'automobile, quelle que soit la marque, on n'obtient pas nécessairement une bonne automobile ou même une automobile qui fonctionne, car les pièces peuvent ne pas fonctionner ensemble. C'est l'interaction entre les pièces, et non les actions d'une seule d'entre elles, qui détermine la performance d'un système.

Ainsi, la recherche opérationnelle tente d'évaluer l'effet des changements dans une partie quelconque d'un système sur la performance du système dans son ensemble et de rechercher les causes d'un problème qui survient dans une partie d'un système dans d'autres parties ou dans les interrelations entre les parties. Dans l'industrie, un problème de production peut être abordé par un changement de politique commerciale. Par exemple, si une usine fabrique quelques produits rentables en grandes quantités et de nombreux articles moins rentables en petites quantités, il faudra peut-être interrompre les longues séries de production efficaces d'articles à fort volume et à forte rentabilité pour produire des séries courtes d'articles à faible volume et à faible rentabilité. Un chercheur en opérations pourrait proposer de réduire les ventes des articles moins rentables et d'augmenter celles des articles rentables en plaçant les vendeurs dans un système d'incitation qui les rémunère particulièrement pour la vente d'articles particuliers.

10.4.2. L'équipe interdisciplinaire

Les disciplines scientifiques et technologiques ont proliféré rapidement au cours des 100 dernières années. Cette prolifération, qui résulte de l'énorme augmentation des connaissances scientifiques, a fourni à la science un système de classement qui permet une classification systématique des connaissances. Ce système de classification est utile pour résoudre de nombreux problèmes en identifiant la discipline appropriée à laquelle on doit faire appel pour trouver une solution. Des difficultés apparaissent lorsque l'on rencontre des problèmes plus complexes, comme ceux qui se posent dans les grands systèmes organisés. Il est alors nécessaire de trouver un moyen de rassembler les divers points de vue disciplinaires. En outre, comme les méthodes diffèrent d'une discipline à l'autre, le recours à des équipes interdisciplinaires permet de disposer d'un arsenal de techniques et d'outils de recherche beaucoup plus vaste que celui dont on disposerait autrement. Ainsi, la recherche opérationnelle peut se caractériser par des combinaisons plutôt inhabituelles de disciplines dans les équipes de recherche et par l'utilisation de procédures de recherche variées.

10.4.3. Méthodologie

Jusqu'au 20e siècle, les expériences en laboratoire étaient la principale et presque la seule méthode de recherche scientifique. Mais les grands systèmes tels qu'ils sont étudiés en recherche opérationnelle ne peuvent pas être amenés dans les laboratoires. En outre, même si les systèmes pouvaient être amenés en laboratoire, ce qui serait appris ne s'appliquerait pas nécessairement à leur comportement dans leur environnement naturel, comme l'ont montré les premières expériences sur les radars. Les expériences sur les systèmes et sous-systèmes menées dans leur environnement naturel ("expériences opérationnelles") sont possibles grâce aux méthodes expérimentales développées par le statisticien britannique R.A. Fisher en 1923-24. Cependant, pour des raisons pratiques ou même éthiques, il est rarement possible d'expérimenter de grands systèmes organisés dans leur ensemble dans leur environnement naturel. Il en résulte un dilemme apparent: pour mieux comprendre les systèmes complexes, l'expérimentation semble nécessaire, mais elle ne peut généralement pas être réalisée. Cette difficulté est résolue par l'utilisation de modèles, des représentations du système étudié. Si le modèle est bon, il est possible de réaliser des expériences (appelées "simulations") sur celui-ci, ou d'utiliser d'autres méthodes pour obtenir des résultats utiles.

10.5. Les phases de la recherche opérationnelle

La démarche de la recherche opérationnelle est illustrée dans la figure I.5.

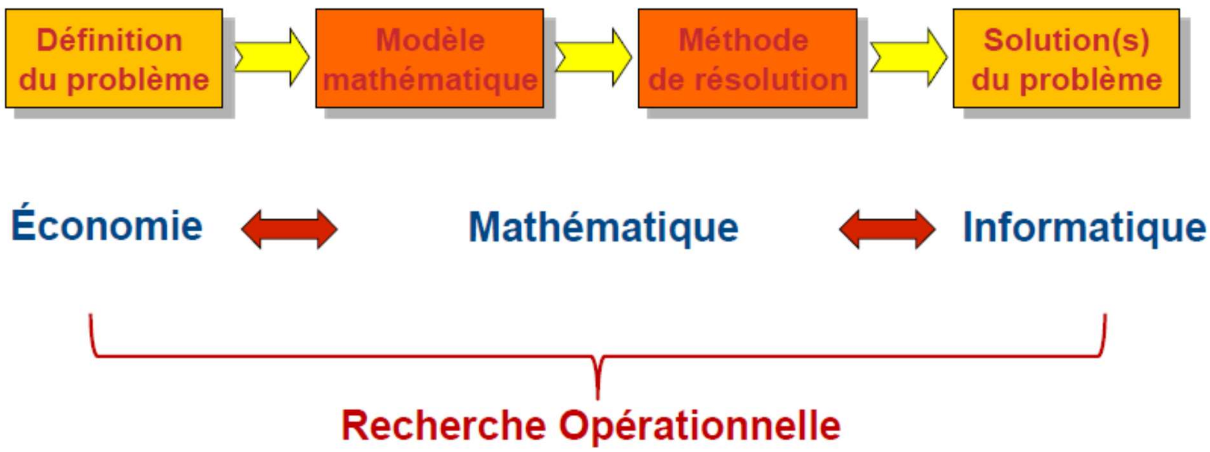

Figure I.5: Les principales étapes de la démarche de la RO.

10.5.1. Formulation (ou définition) du problème

Pour formuler un problème de recherche opérationnelle, il faut concevoir une mesure appropriée de la performance, définir les différentes possibilités d'action (c'est-à-dire les variables contrôlées et leurs contraintes) et identifier les variables non contrôlées pertinentes.

Pour concevoir une mesure de la performance, les objectifs sont identifiés et définis, puis quantifiés. Si les objectifs ne peuvent être quantifiés ou exprimés en termes rigoureux (généralement mathématiques), la plupart des techniques de recherche opérationnelle ne peuvent être appliquées. Par exemple, un chef d'entreprise peut avoir l'objectif acquisitif de lancer un nouveau produit et de le rendre rentable en un an. L'objectif identifié est le profit en un an, qui est défini comme les recettes moins les coûts, et serait probablement quantifié en termes de ventes. Dans le monde réel, les conditions peuvent changer avec le temps. Ainsi, bien qu'un objectif donné soit identifié au début de la période, il est souvent nécessaire de le modifier et de le reformuler.

Une connaissance détaillée du fonctionnement réel du système étudié et de son environnement est essentielle. Cette connaissance est normalement acquise par une analyse du système, un processus en quatre étapes qui consiste à déterminer quels sont les besoins ou les désirs que l'organisation tente de satisfaire, comment ils sont communiqués à l'organisation, comment l'information sur les besoins et les désirs pénètre dans l'organisation, et quelles sont les actions entreprises, comment elles sont contrôlées, et quelles sont les exigences en matière de temps et de ressources de ces actions. Ces informations peuvent généralement être représentées graphiquement dans un organigramme, ce qui permet aux chercheurs d'identifier les variables qui affectent les performances du système.

Une fois que les objectifs, les décideurs, leurs plans d'action et les variables non contrôlées ont été identifiés et définis, il est possible de développer une mesure de la performance et de sélectionner une fonction quantitative de cette mesure à utiliser comme critère pour la meilleure solution.

Le type de critère de décision qui convient à un problème dépend de l'état des connaissances concernant les résultats possibles. La certitude décrit une situation dans laquelle on pense que chaque ligne d'action aboutira à un résultat particulier. Le risque est une situation dans laquelle, pour chaque ligne d'action, d'autres résultats sont possibles, dont les probabilités sont connues ou peuvent être estimées. L'incertitude décrit une situation dans laquelle, pour chaque ligne d'action, des probabilités ne peuvent être attribuées aux résultats possibles.

Dans les situations de risque, qui sont les plus courantes dans la pratique, l'objectif est normalement de maximiser le gain net ou le gain brut attendu (moyenne à long terme) pour des coûts spécifiés, ou de minimiser les coûts pour des bénéfices spécifiés. Une entreprise, par exemple, cherche à maximiser les bénéfices attendus ou à minimiser les coûts attendus. D'autres objectifs, pas nécessairement liés, peuvent être recherchés ; par exemple, un planificateur économique peut souhaiter maintenir le plein emploi sans inflation ; ou différents groupes au sein d'une organisation peuvent être amenés à faire des compromis entre leurs différents objectifs, comme lorsqu'une armée et une marine, par exemple, doivent coopérer en matière de défense.

En abordant des situations incertaines, on peut tenter de maximiser le gain minimum ou de minimiser la perte maximum résultant d'un choix ; c'est l'approche "minimax". On peut aussi peser les résultats possibles en fonction de son optimisme ou de son pessimisme, puis appliquer le principe du minimax. Une troisième approche, le "regret minimax", tente de minimiser l'écart maximal par rapport au résultat qui aurait été choisi si un état de certitude avait existé avant que le choix ne soit fait.

Chaque variable identifiée doit être définie en termes de conditions dans lesquelles, et d'opérations de recherche par lesquelles, les questions concernant sa valeur doivent être répondues ; cela inclut l'identification de l'échelle utilisée pour mesurer la variable.

10.5.2 Construction du modèle

Un modèle est une représentation simplifiée du monde réel et, en tant que tel, ne comprend que les variables pertinentes pour le problème en question. Un modèle de corps en chute libre, par exemple, ne fait pas référence à la couleur, à la texture ou à la forme du corps concerné. En outre, un modèle peut ne pas inclure toutes les variables pertinentes, car un petit pourcentage de celles-ci peut expliquer la majeure partie du phénomène à expliquer. Bon nombre des simplifications utilisées produisent une certaine erreur dans les prédictions dérivées du modèle, mais celles-ci peuvent souvent rester faibles par rapport à l'ampleur de l'amélioration des opérations qui peut en être extraite. La plupart des modèles de recherche opérationnelle sont des modèles symboliques, car les symboles représentent les propriétés du système. Les premiers modèles étaient des représentations physiques telles que des maquettes de bateaux, d'avions, de réservoirs de remorquage et de souffleries. Les modèles physiques sont généralement assez faciles à construire, mais seulement pour des objets ou des systèmes relativement simples, et sont généralement difficiles à modifier.

L'étape suivante au-delà du modèle physique est le graphique, plus facile à construire et à manipuler mais plus abstrait. La représentation graphique de plus de trois variables étant difficile, les modèles symboliques sont apparus. Il n'y a pas de limite au nombre de variables qui peuvent être incluses dans un modèle symbolique, et ces modèles sont plus faciles à construire et à manipuler que les modèles physiques.

Les modèles symboliques sont complètement abstraits. Lorsque les symboles d'un modèle sont définis, le modèle reçoit un contenu ou une signification. Cela a des conséquences importantes. Les modèles symboliques de systèmes dont le contenu est très différent révèlent souvent une structure similaire. Par conséquent, la plupart des systèmes et des problèmes qu'ils posent peuvent être classés de manière fructueuse en fonction d'un nombre relativement restreint de structures. En outre, comme les méthodes d'extraction de solutions à partir de modèles ne dépendent que de leur structure, certaines méthodes peuvent être utilisées pour résoudre une grande variété de problèmes d'un point de vue contextuel. Enfin, un système qui a la même structure qu'un autre, même si leur contenu est différent, peut être utilisé comme modèle de l'autre. Un tel modèle est appelé un analogue. En utilisant de tels modèles, une grande partie de ce que l'on sait du premier système peut être appliquée au second.

Malgré les avantages évidents des modèles symboliques, il existe de nombreux cas où les modèles physiques sont encore utiles, comme pour tester des structures et des mécanismes physiques; il en va de même pour les modèles graphiques. Les modèles physiques et graphiques sont fréquemment utilisés dans les phases préliminaires de la construction de modèles symboliques de systèmes.

Les modèles de recherche opérationnelle représentent la relation de cause à effet entre les variables contrôlées et non contrôlées et la performance du système; ils doivent donc être explicatifs et non simplement descriptifs. Seuls les modèles explicatifs peuvent fournir les moyens nécessaires pour manipuler le système afin de produire les changements de performance souhaités.

L'analyse de la recherche opérationnelle vise à établir des relations de cause à effet. Bien que les expériences portant sur les opérations réelles de tout ou partie d'un système soient souvent utiles, elles ne constituent pas le seul moyen d'analyser les causes et les effets. Il existe quatre modèles de construction de modèles, dont deux seulement impliquent l'expérimentation: l'inspection, l'utilisation d'analogues, l'analyse opérationnelle et les expériences opérationnelles. Ils sont considérés ici par ordre de complexité croissante.

Dans certains cas, le système et son problème sont relativement simples et peuvent être appréhendés soit par une inspection, soit par une discussion avec des personnes familiarisées avec le système. En général, seuls les problèmes de fonctionnement de bas niveau et répétitifs, ceux dans lesquels le comportement humain joue un rôle mineur, peuvent être traités ainsi.

Lorsque le chercheur éprouve des difficultés à représenter symboliquement la structure d'un système, il est parfois possible d'établir une similitude, sinon une identité, avec un autre système dont la structure est mieux connue et plus facile à manipuler. Il peut alors être possible d'utiliser soit le système analogue lui-même, soit un modèle symbolique de celui-ci comme modèle du système problématique. Par exemple, une équation dérivée de la théorie cinétique des gaz a été utilisée comme modèle du mouvement des trains entre deux cours de classement. On a construit des analogues hydrauliques de l'économie et des analogues électroniques du trafic automobile avec lesquels on peut faire des expériences pour déterminer les effets de la manipulation des variables contrôlables. Ainsi, des analogues peuvent être construits aussi bien que trouvés dans des systèmes existants.

Dans certains cas, l'analyse des opérations réelles d'un système peut révéler sa structure causale. Les données sur les opérations sont analysées pour produire une hypothèse explicative, qui est testée par l'analyse des données d'exploitation. Ce test peut conduire à une révision de l'hypothèse. Le cycle se poursuit jusqu'à ce qu'un modèle explicatif satisfaisant soit développé.

Par exemple, une analyse des voitures s'arrêtant dans les stations-service urbaines situées aux intersections de deux rues a révélé que la plupart d'entre elles provenaient de quatre des 16 routes possibles à travers l'intersection (quatre façons d'entrer et quatre façons de sortir). L'examen du pourcentage de voitures de chaque itinéraire qui s'arrêtent pour s'approvisionner suggère que ce pourcentage est lié au temps perdu en s'arrêtant. Des données ont ensuite été recueillies sur le temps perdu par les voitures dans chaque itinéraire. Cela a révélé une relation inverse étroite entre le pourcentage d'arrêts et le temps perdu. Mais cette relation n'était pas linéaire, c'est-à-dire que l'augmentation de l'un n'était pas proportionnelle à l'augmentation de l'autre. Il a ensuite été constaté que le temps perdu perçu était supérieur au temps perdu réel, et que la relation entre le pourcentage de voitures s'arrêtant et le temps perdu perçu était étroite et linéaire. L'hypothèse a été systématiquement testée et vérifiée et un modèle a été construit, qui met en relation le nombre de voitures s'arrêtant aux stations-service avec la quantité de trafic sur chaque route traversant son intersection et avec les caractéristiques de la station qui affectent le temps nécessaire pour obtenir un service.

Dans les situations où il n'est pas possible d'isoler les effets des variables individuelles par l'analyse des données d'exploitation, il peut être nécessaire de recourir à des expériences opérationnelles pour déterminer quelles variables sont pertinentes et comment elles affectent les performances du système.

C'est le cas, par exemple, des tentatives de quantifier les effets de la publicité (quantité, moment et média utilisé) sur les ventes d'un produit de consommation. La publicité faite par le producteur n'est qu'une des nombreuses variables contrôlées et non contrôlées qui affectent les ventes. Par conséquent, dans de nombreux cas, son effet ne peut être isolé et mesuré que par des expériences contrôlées sur le terrain.

Il en va de même pour déterminer comment la taille, la forme, le poids et le prix d'un produit alimentaire affectent ses ventes. Dans ce cas, les expériences en laboratoire sur des échantillons de consommateurs peuvent être utilisées dans les étapes préliminaires, mais les expériences sur le terrain sont finalement nécessaires. Les expériences ne produisent cependant pas de théories explicatives. Elles ne peuvent être utilisées que pour tester les hypothèses explicatives formulées avant de concevoir l'expérience et pour suggérer d'autres hypothèses à tester.

Il est parfois nécessaire de modifier un modèle par ailleurs acceptable parce qu'il n'est pas possible ou pratique de trouver les valeurs numériques des variables qui y figurent. Par exemple, un modèle destiné à guider la sélection de projets de recherche peut contenir des variables telles que "la probabilité de réussite du projet", "le coût prévu du projet" et son "rendement prévu". Mais aucune de ces variables ne peut être calculée avec une quelconque fiabilité.

Les modèles n'aident pas seulement à résoudre les problèmes, ils sont également utiles pour les formuler. En d'autres termes, les modèles peuvent servir de guides pour explorer la structure d'un problème et révéler des pistes d'action possibles qui pourraient autrement passer inaperçues. Dans de nombreux cas, la ligne de conduite révélée par l'application d'un modèle est si manifestement supérieure aux possibilités précédemment envisagées que la justification de son choix est à peine nécessaire.

Dans certains cas, le modèle d'un problème peut être soit trop compliqué, soit trop vaste pour être résolu. Il est souvent possible de diviser le modèle en parties solubles individuellement et de prendre la sortie d'un modèle comme entrée dans un autre. Comme les modèles sont susceptibles d'être interdépendants, plusieurs répétitions de ce processus peuvent être nécessaires.

10.5.3. Détermination de solutions à partir de modèles

Les procédures d'élaboration de solutions à partir de modèles sont soit déductives, soit inductives. Avec la déduction, on passe directement du modèle à une solution sous forme symbolique ou numérique. De telles procédures sont fournies par les mathématiques, par exemple, le calcul. Une procédure analytique explicite pour trouver la solution est appelée un algorithme.

Même si un modèle ne peut être résolu, et beaucoup sont trop complexes pour être résolus, il peut être utilisé pour comparer des solutions alternatives. Il est parfois possible de mener une séquence de comparaisons, chacune étant suggérée par la précédente et chacune étant susceptible de contenir une meilleure alternative que celle contenue dans toute comparaison précédente. Une telle procédure de recherche de solutions est appelée heuristique.

Les procédures inductives consistent à essayer et à comparer différentes valeurs des variables contrôlées. Ces procédures sont dites itératives (répétitives) si elles passent par des solutions successivement améliorées jusqu'à ce qu'une solution optimale soit atteinte ou que la poursuite du calcul ne soit pas justifiée. Une base rationnelle pour mettre fin à un tel processus - connue sous le nom de "règles d'arrêt" - implique la détermination du point auquel l'amélioration attendue de la solution lors de l'essai suivant est inférieure au coût de l'essai.

Des algorithmes aussi connus que la programmation linéaire, non linéaire et dynamique sont des procédures itératives basées sur la théorie mathématique. La simulation et l'optimisation expérimentale sont des procédures itératives basées principalement sur les statistiques.

10.5.4. Tester le modèle et la solution

Un modèle peut être déficient parce qu'il inclut des variables non pertinentes, exclut des variables pertinentes, contient des variables mal évaluées, est mal structuré ou contient des contraintes mal formulées. Les tests pour les déficiences d'un modèle sont de nature statistique ; leur utilisation requiert des connaissances en théorie de l'échantillonnage et de l'estimation, en plans d'expériences et en théorie des tests d'hypothèses (voir aussi statistiques).

La théorie de l'échantillonnage et de l'estimation concerne la sélection d'un échantillon d'éléments dans un grand groupe et l'utilisation des propriétés observées pour caractériser le groupe dans son ensemble. Pour économiser du temps et de l'argent, l'échantillon prélevé est aussi petit que possible. Il existe plusieurs théories du plan d'échantillonnage et de l'estimation, chacune donnant des estimations aux propriétés différentes.

La structure d'un modèle consiste en une fonction reliant la mesure de la performance aux variables contrôlées et non contrôlées ; par exemple, une entreprise peut tenter de montrer la relation fonctionnelle entre les niveaux de profit (la mesure de la performance) et les variables contrôlées (prix, montant dépensé en publicité) et non contrôlées (conditions économiques, concurrence). Afin de tester le modèle, les valeurs de la mesure de la performance calculées à partir du modèle sont comparées aux valeurs réelles dans différentes conditions. S'il existe une différence significative entre ces valeurs, ou si la variabilité de ces différences est importante, le modèle doit être réparé. De tels tests n'utilisent pas les données qui ont été utilisées pour construire le modèle, car cela permettrait de déterminer dans quelle mesure le modèle s'adapte aux données de performance à partir desquelles il a été dérivé, et non dans quelle mesure il prédit la performance.

La solution dérivée d'un modèle est testée pour déterminer si elle donne de meilleures performances qu'une alternative, généralement celle qui est utilisée actuellement. Le test peut être prospectif, par rapport aux performances futures, ou rétrospectives, en comparant les solutions qui auraient été obtenues si le modèle avait été utilisé dans le passé avec ce qui s'est réellement passé. Si aucun test prospectif ou rétrospectif n'est possible, il peut être possible d'évaluer la solution par une "analyse de sensibilité", une mesure de la mesure dans laquelle les estimations utilisées dans la solution devraient être erronées pour que la solution proposée donne des résultats moins satisfaisants que la procédure de décision alternative.

Le coût de la mise en œuvre d'une solution doit être soustrait du gain attendu de son application, ce qui permet d'obtenir une estimation de l'amélioration nette. Si des erreurs ou des inefficacités dans l'application de la solution sont possibles, elles doivent également être prises en compte dans l'estimation de l'amélioration nette.

10.5.5. Mise en œuvre et contrôle de la solution

L'acceptation d'une solution recommandée par le gestionnaire responsable dépend de la mesure dans laquelle il croit que cette solution est meilleure que les autres solutions. Cela dépend à son tour de sa confiance dans les chercheurs impliqués et leurs méthodes. Par conséquent, la participation des gestionnaires au processus de recherche est essentielle au succès.

On attend normalement des chercheurs opérationnels qu'ils supervisent la mise en œuvre d'une solution acceptée. Cela leur donne un test ultime de leur travail et l'occasion de faire des ajustements si des déficiences apparaissent dans l'application. L'équipe de recherche opérationnelle prépare des instructions détaillées pour les personnes chargées de mettre en œuvre la solution et les forme à suivre ces instructions. La coopération de ceux qui mettent en œuvre la solution et de ceux qui seront affectés par elle doit être recherchée au cours du processus de recherche, et non après que tout soit terminé. Les plans et les calendriers de mise en œuvre sont testés au préalable et les lacunes sont corrigées. Les performances réelles de la solution sont comparées aux attentes et, en cas de divergence importante, les raisons en sont déterminées et les ajustements appropriés sont effectués.

La solution peut ne pas donner les résultats escomptés pour une ou plusieurs raisons : le modèle peut être mal construit ou mal utilisé ; les données utilisées pour construire le modèle peuvent être incorrectes ; la solution peut être mal exécutée ; le système ou son environnement peut avoir changé de manière inattendue après l'application de la solution. Une action corrective est nécessaire dans chaque cas.

Pour contrôler une solution, il faut décider ce qui constitue une déviation significative des performances par rapport aux attentes, déterminer la fréquence des contrôles, la taille et le type d'échantillon d'observations à effectuer, et les types d'analyses des données résultantes qui doivent être réalisées, et prendre les mesures correctives appropriées. La deuxième étape doit être conçue de manière à minimiser la somme des coûts d'exécution des procédures de contrôle et des erreurs qui pourraient être commises.

Comme la plupart des modèles impliquent une variété d'hypothèses, celles-ci sont vérifiées systématiquement. Ce contrôle nécessite une formulation explicite des hypothèses formulées lors de la construction du modèle.

Des contrôles efficaces permettent non seulement de rendre possible mais aussi de mieux comprendre la dynamique du système concerné. Grâce aux contrôles, le système de résolution de problèmes dont la recherche opérationnelle fait partie apprend de sa propre expérience et s'adapte plus efficacement aux conditions changeantes.

10.6. Les techniques de recherche opérationnelle

Les techniques de la recherche opérationnelle se base essentiellement su deux catégories de modèles: modèles déterministes et modèles stochastiques (probabilistes). Toutefois, certains des ses modèles sont considérés de façons plus accommodée comme hybrides, (i.e., modèles comportant des éléments déterministes et probabilistes).

En fonction des ces types de modèles, les techniques de la RO se déclinent sur plusieurs sujets, comme le montre la figure I.6.

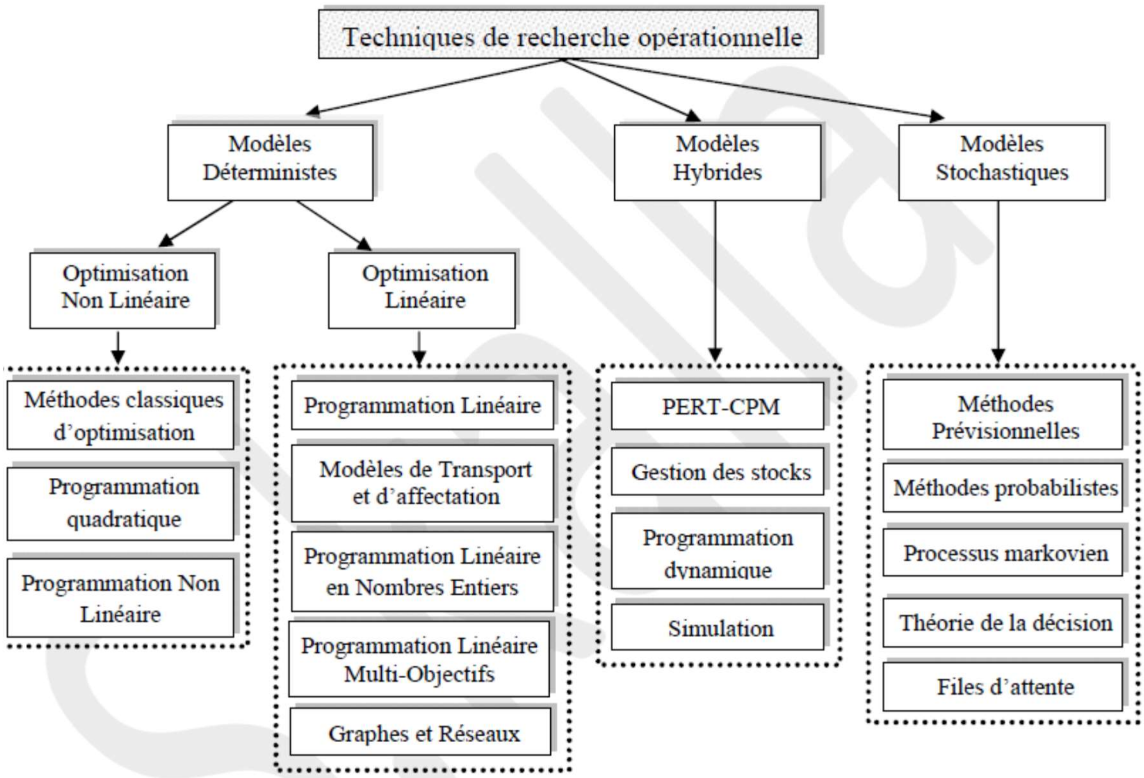

Figure I.6: Les techniques de la RO.

10.7. La RO et l'optimisation

L'une des caractéristiques principales de la RO est la tentative de trouver la meilleure solution possible au problème/modèle et d'identifier le plan d'action qui crée une performance optimale du système réel. En d'autres termes, la RO vise à recommander des décisions optimales. Cette "recherche de l'optimalité" est un moteur important qui, bien qu'il ne soit pas toujours réalisable dans la pratique, est inhérent au paradigme de la RO et distingue l'approche de la RO axée sur les objectifs de la prise de décision sur la base de "règles de base" représentant l'expertise et les (meilleures) pratiques.

La détermination de la meilleure solution possible fait appel aux techniques d'optimisation. Donc, l'optimisation constitue un domaine d'application majeur de la RO.

10.8. Applications de la RO

10.8.1. Allocation des ressources

Les problèmes d'allocation impliquent la répartition des ressources entre des alternatives concurrentes afin de minimiser les coûts totaux ou de maximiser le rendement total. De tels problèmes comportent les éléments suivants: un ensemble de ressources disponibles en quantités données ; un ensemble de tâches à accomplir, chacune consommant une quantité spécifiée de ressources ; et un ensemble de coûts ou de rendements pour chaque tâche et chaque ressource. Le problème consiste à déterminer la quantité de chaque ressource à allouer à chaque tâche.

S'il y a plus de ressources disponibles que nécessaires, la solution doit indiquer quelles ressources ne doivent pas être utilisées, en tenant compte des coûts associés. De même, s'il y a plus de tâches que ce qui peut être fait avec les ressources disponibles, la solution doit indiquer quelles tâches ne doivent pas être faites, en tenant compte des coûts associés.

Si chaque tâche requiert exactement une ressource (par exemple, une personne) et que chaque ressource ne peut être utilisée que pour une seule tâche, le problème résultant est un problème d'affectation. Si les ressources sont divisibles et si les emplois et les ressources sont exprimés en unités à la même échelle, on parle de problème de transport ou de distribution. Si les emplois et les ressources ne sont pas exprimés dans les mêmes unités, il s'agit d'un problème d'affectation général.

Un problème d'affectation peut consister à affecter des travailleurs à des bureaux ou à des emplois, des camions à des itinéraires de livraison, des chauffeurs à des camions ou des classes à des salles. Un problème de transport typique consiste à distribuer des wagons de marchandises vides là où ils sont nécessaires ou à affecter des commandes aux usines pour la production. Le problème général de répartition peut consister à déterminer quelles machines doivent être employées pour fabriquer un produit donné ou quel ensemble de produits doit être fabriqué dans une usine pendant une période donnée.

Dans les problèmes de répartition, les coûts ou les rendements unitaires peuvent être indépendants ou interdépendants ; par exemple, le rendement de l'investissement d'un dollar en efforts de vente peut dépendre du montant dépensé en publicité. Si les répartitions effectuées au cours d'une période influent sur celles des périodes suivantes, le problème est dit dynamique et le temps doit être pris en compte dans sa solution.

10.8.2. Contrôle des stocks

Les stocks comprennent les matières premières, les pièces détachées, les travaux en cours, les produits finis, les matériaux d'emballage et de conditionnement et les fournitures générales. Le contrôle des stocks, vital pour la solidité financière d'une entreprise, consiste en général à décider à quels moments du système de production les stocks doivent être maintenus et quelles doivent être leur forme et leur taille. Étant donné que certains coûts unitaires augmentent avec la taille des stocks - notamment le stockage, l'obsolescence, la détérioration, l'assurance, l'investissement - et que d'autres coûts unitaires diminuent avec la taille des stocks - notamment les coûts d'installation ou de préparation, les retards dus à des pénuries, etc. Une autre partie du problème général des stocks consiste à décider des niveaux (points de commande) auxquels les commandes de réapprovisionnement des stocks doivent être lancées.

Le contrôle des stocks s'intéresse à deux questions : quand réapprovisionner le magasin et dans quelle mesure. Il existe deux principaux systèmes de contrôle. Le système à deux bacs (parfois appelé système min-max) implique l'utilisation de deux bacs, physiquement ou sur papier. Le premier bac est destiné à répondre à la demande actuelle et le second à satisfaire la demande pendant la période de réapprovisionnement. Lorsque le stock du premier emplacement est épuisé, une commande pour une quantité donnée est générée. Le système de cycle de réapprovisionnement, ou système de revue cyclique, consiste à commander à intervalles réguliers fixes. Diverses combinaisons de ces systèmes peuvent être utilisées dans la construction d'une procédure de contrôle des stocks. Un système pur à deux bacs, par exemple, peut être modifié pour exiger une révision cyclique plutôt que continue du stock, les commandes n'étant générées que lorsque le stock tombe en dessous d'un niveau spécifique. De même, un système à cycle de réapprovisionnement pur peut être modifié pour permettre de générer des commandes si le stock tombe en dessous du niveau de réapprovisionnement entre les révisions cycliques. Dans une autre variante encore, la quantité à commander dans le système à cycle de commande dépend du niveau de stock à la période de révision ou de la nécessité de commander d'autres produits ou matériaux au même moment, ou les deux.

Le problème classique des stocks consiste à déterminer la quantité d'une ressource à acquérir, soit en l'achetant, soit en la produisant, et à déterminer si et quand l'acquérir pour minimiser la somme des coûts qui augmentent avec la taille des stocks et ceux qui diminuent avec l'augmentation des stocks. Les coûts du premier type comprennent le coût du capital investi dans les stocks, la manutention, le stockage, l'assurance, les taxes, la dépréciation, la détérioration et l'obsolescence. Les coûts qui diminuent avec l'augmentation des stocks comprennent les coûts de pénurie (découlant des ventes perdues), les coûts de mise en place de la production et le prix d'achat ou les coûts directs de production. Les coûts de mise en place comprennent le coût de la passation d'un ordre d'achat ou du lancement d'une production. Si de grandes quantités sont commandées, les stocks augmentent mais la fréquence des commandes diminue, d'où une baisse des coûts de préparation. En général, plus la quantité commandée est importante, plus le prix d'achat unitaire est bas en raison des remises sur quantité et du coût de production unitaire plus faible résultant de l'efficacité accrue des longues séries de production. Parmi les autres variables pertinentes figurent la demande de la ressource et le délai entre la passation et l'exécution des commandes.

Les problèmes d'inventaire se posent dans une grande variété de contextes ; par exemple, pour déterminer les quantités de marchandises à acheter ou à produire, le nombre de personnes à embaucher ou à former, la taille d'une nouvelle installation de production ou de vente au détail ou le nombre de personnes à fournir, et le capital fluide (d'exploitation) à garder disponible. Les modèles d'inventaire pour les articles uniques sont bien développés et sont normalement résolus par des calculs. Lorsque les quantités commandées pour de nombreux articles sont interdépendantes (comme, par exemple, lorsque l'espace de stockage ou le temps de production sont limités), le problème est plus difficile. Certains des plus gros problèmes peuvent être résolus en les décomposant en problèmes d'inventaire et d'allocation interactifs. Dans les très grands problèmes, la simulation peut être utilisée pour tester diverses règles de décision pertinentes.

10.8.3. Remplacement et maintenance

Les problèmes de remplacement concernent les articles qui se détériorent avec l'utilisation ou le passage du temps et ceux qui tombent en panne après une certaine durée d'utilisation ou de temps. Les articles qui se détériorent sont généralement de grande taille et coûteux (par exemple, les machines-outils, les camions, les navires et les appareils ménagers). Les articles qui ne se détériorent pas ont tendance à être petits et relativement peu coûteux (par exemple, les ampoules électriques, les tubes à vide, les cartouches d'encre). Plus un article en voie de détérioration est utilisé longtemps, plus il nécessite d'entretien pour maintenir son efficacité. En outre, plus un tel article est conservé longtemps, moins sa valeur de revente est élevée et plus il est susceptible d'être rendu obsolète par un nouvel équipement. En revanche, si l'article est remplacé fréquemment, les coûts d'investissement augmentent. Le problème consiste donc à déterminer le moment où il convient de remplacer ces articles et l'ampleur de la maintenance (en particulier préventive) à effectuer de manière à minimiser la somme des coûts d'exploitation, de maintenance et d'investissement.

Dans le cas d'articles qui ne se détériorent pas, le problème consiste à déterminer s'il faut les remplacer en groupe ou remplacer individuellement au fur et à mesure de leur défaillance. Bien que le remplacement en groupe soit un gaspillage, le coût de la main-d'œuvre pour les remplacements est plus élevé lorsqu'ils sont effectués individuellement ; par exemple, les ampoules électriques d'un grand système de métro peuvent être remplacées en groupe pour économiser de la main-d'œuvre. Les problèmes de remplacement qui impliquent de minimiser les coûts des articles, des défaillances et de la main-d'œuvre de remplacement peuvent être résolus par analyse numérique ou par simulation.

Les "éléments" impliqués dans les problèmes de remplacement peuvent être des personnes. Dans ce cas, la maintenance peut être interprétée comme une formation ou une amélioration du salaire, du statut ou des avantages sociaux. L'échec peut être interprété comme un départ, et l'investissement comme des coûts de recrutement, d'embauche et de formation initiale. Il existe de nombreuses autres complexités dans de tels cas ; par exemple, l'effet de la démission ou de la promotion d'une personne sur le comportement des autres. Des aspects de l'environnement aussi contrôlables que le lieu et les horaires de travail peuvent avoir un effet considérable sur la productivité et les taux d'échec. Dans les problèmes de ce type, les apports des sciences comportementales sont particulièrement utiles.

10.8.4. Mise en file d'attente

Une file d'attente est une ligne d'attente, et la mise en file d'attente implique de traiter des objets ou des personnes en séquence. Ainsi, un problème de file d'attente consiste soit à déterminer les installations à fournir, soit à planifier leur utilisation. Le coût de la prestation du service et le temps d'attente des utilisateurs sont minimisés. Parmi les exemples de tels problèmes, on peut citer la détermination du nombre de caisses de sortie à prévoir dans un supermarché, de pistes d'atterrissage dans un aéroport, de places de stationnement dans un centre commercial ou de guichets dans une banque. De nombreux problèmes de maintenance peuvent être traités comme des problèmes de file d'attente ; les éléments nécessitant une réparation sont comme les utilisateurs d'un service. Certains problèmes d'inventaire peuvent également être formulés comme des problèmes de file d'attente dans lesquels les commandes sont comme des utilisateurs et les stocks comme des installations de service.

10.8.5. Séquençage des tâches dans les ateliers

Dans les problèmes de files d'attente, l'ordre dans lequel les utilisateurs en attente de service sont servis est toujours spécifié. La sélection de cet ordre de manière à minimiser une fonction du temps d'exécution de toutes les tâches est un problème d'ordonnancement. La mesure de performance peut prendre en compte le temps total écoulé, le retard total dans le respect des délais ou des dates d'échéance, et le coût des stocks en cours.

Le contexte le plus courant pour les problèmes de séquençage est celui d'une installation de production par lots, ou atelier de fabrication, qui traite de nombreux produits différents avec de nombreuses combinaisons de machines. Dans ce contexte, il peut être nécessaire de tenir compte de facteurs tels que le chevauchement des services (c'est-à-dire que si une commande est composée d'un certain nombre d'articles devant passer par plusieurs étapes d'un processus, les premiers articles qui terminent l'étape initiale peuvent commencer la deuxième étape avant que le dernier n'ait terminé la première), le temps de transport entre les installations de service, la correction des pannes de service, les pannes d'installation et les pénuries de matériaux.

Un problème simplifié de séquençage d'atelier, avec deux tâches et quatre machines, est illustré dans la figure I.7. En haut de cette figure se trouve la séquence d'opérations des deux travaux. Le travail A doit d'abord aller sur la machine 1, puis sur la machine 2, puis sur la machine 3, et enfin sur la 4 machine, et l'ordre de traitement sur les quatre machines ne peut pas être modifié. Le temps de traitement de la tâche est d'une heure sur chaque machine, pour un total de quatre heures d'usinage. Dans cet exemple, la tâche ne peut être traitée que sur une seule machine à la fois, comme si la tâche consistait en un seul produit traité par quatre machines-outils.

Le travail B doit suivre une séquence différente. Il commence également sur la machine 1, mais passe ensuite à la machine 4, puis à la machine 2, et enfin à la machine 4. Chaque opération d'usinage de la tâche B nécessite également une heure.

Sous les graphiques indiquant la séquence d'opérations requise, deux programmes alternatifs sont présentés pour les deux travaux. (Dans un diagramme à barres, le temps est indiqué sur la ligne horizontale, et les barres ou blocs représentent le temps pendant lequel chaque opération est programmée sur chacune des quatre machines). Le premier programme suppose que le travail A est exécuté en premier. Une fois la tâche A disposée sur le planning, les opérations de la tâche B sont placées sur le graphique aussi loin que possible vers la gauche, sans violer les contraintes de séquence. Dans ce cas, le graphique montre que les deux tâches (huit heures de travail) peuvent être réalisées en cinq heures. Cela est possible en exécutant les deux tâches en même temps (sur des machines distinctes) pendant les deuxième, troisième et quatrième heures. Le deuxième programme suppose que le travail B est exécuté en premier. Ce programme nécessite un total de six heures, soit une de plus que le programme précédent. Si le temps total écoulé pour la réalisation des deux tâches est un critère important, le premier planning sera supérieur au second.

Bien que ce problème soit facile à résoudre, les solutions aux problèmes réels de séquençage dans les ateliers nécessitent l'utilisation de modèles sophistiqués et la puissance de calcul des ordinateurs. Il n'est pas rare que les ateliers aient 5 000 commandes de clients en cours de traitement à un moment donné, chaque commande nécessitant 50 ou 60 opérations distinctes de traitement ou de machine. Le nombre de combinaisons de séquences réalisables est astronomique dans de tels problèmes, et ils posent de nombreux problèmes de modélisation et de développement de systèmes aux chercheurs en opérations et aux ingénieurs industriels.

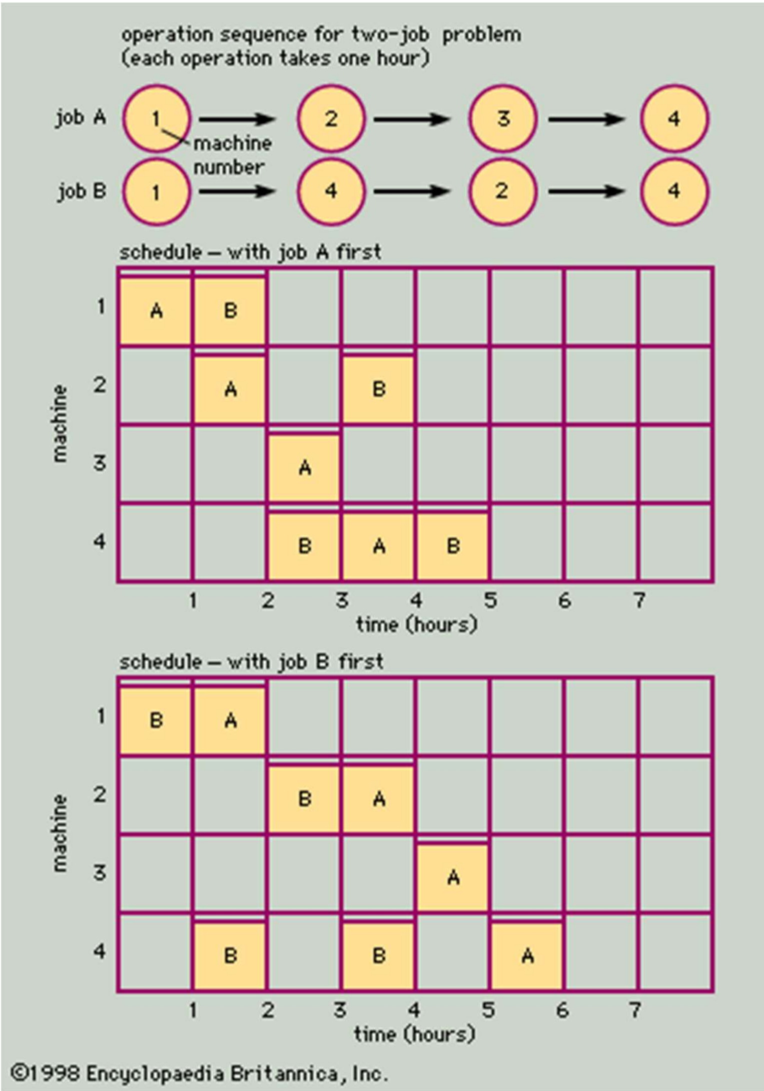

Figure I.7: Exemple de séquençage de tâches dans un atelier.

10.8.6. Routage de réseau

Un réseau peut être défini par un ensemble de points, ou "nœuds", qui sont reliés par des lignes, ou "liens". Une façon d'aller d'un nœud (l'"origine") à un autre (la "destination") est appelée "route" ou "chemin". Les liens, qui peuvent être unidirectionnels ou bidirectionnels, sont généralement caractérisés par le temps, le coût ou la distance nécessaires pour les traverser. Le temps ou le coût de déplacement dans différentes directions sur le même lien peut différer.

Un problème de routage réseau consiste à trouver une route optimale entre deux ou plusieurs nœuds en fonction du temps total, du coût ou de la distance. Diverses contraintes peuvent exister, comme l'interdiction de revenir à un nœud déjà visité ou la stipulation de ne passer qu'une seule fois par chaque nœud.

Les problèmes de routage en réseau se posent couramment dans les systèmes de communication et de transport. Les retards qui se produisent aux nœuds (par exemple, les gares de classification des chemins de fer ou les standards téléphoniques) peuvent être fonction des charges qui leur sont imposées et de leurs capacités. Les pannes peuvent se produire dans les liens ou les nœuds. Le "problème du voyageur de commerce" a été beaucoup étudié. Il consiste à partir d'un nœud désigné à établir un itinéraire qui ne passe qu'une seule fois par chaque nœud (par exemple, une ville) et qui retourne au point d'origine en un minimum de temps, de coût ou de distance. Ce problème se pose lors de la sélection d'un ordre de traitement d'un ensemble de tâches de production lorsque le coût de mise en place de chaque tâche dépend de la tâche qui l'a précédée. Dans ce cas, les tâches peuvent être considérées comme des nœuds, chacun d'entre eux étant connecté à tous les autres, les coûts de préparation étant l'équivalent des distances entre eux. L'ordre qui permet d'obtenir le coût total de préparation le plus faible équivaut donc à une solution au problème du voyageur de commerce. La complexité des calculs est telle que même avec l'utilisation d'ordinateurs, il est très coûteux de traiter plus de 20 nœuds. Il existe cependant des procédures d'approximation moins coûteuses. Les problèmes de routage plus typiques consistent à se rendre d'un endroit à un autre en un minimum de temps, de coût ou de distance. Des procédures graphiques et analytiques sont disponibles pour trouver de telles routes.

10.8.7. Problèmes concurrentiels

Les problèmes concurrentiels traitent du choix dans des situations interactives où le résultat du choix d'un décideur dépend du choix, utile ou nuisible, d'un ou de plusieurs autres. La guerre, le marketing et les appels d'offres en sont des exemples. Les problèmes concurrentiels peuvent être classés comme certains, risqués ou incertains, en fonction de l'état de la connaissance qu'a un décideur des choix de son adversaire. Dans des conditions de certitude, il est facile de maximiser les gains ou de minimiser les pertes. Les problèmes concurrentiels de type risqué nécessitent l'utilisation d'une analyse statistique pour leur résolution ; l'aspect le plus difficile de la résolution de tels problèmes réside généralement dans l'estimation des probabilités des choix des concurrents ; par exemple, dans la soumission d'un contrat pour lequel les concurrents et leurs offres sont inconnus.

La théorie des jeux a été développée pour traiter une grande classe de situations compétitives de type incertitude dans lesquelles chaque participant connaît les choix qui s'offrent à lui et à chaque autre participant. Il existe un "état final" bien défini qui met fin à l'interaction (par exemple, victoire, défaite ou match nul), et les gains associés à chaque état final sont spécifiés à l'avance et sont connus de chaque participant. Dans les situations où toutes les alternatives sont ouvertes à la concurrence, ou si certains de leurs résultats ne sont pas connus à l'avance, les jeux opérationnels peuvent parfois être utilisés. Les militaires construisent depuis longtemps des jeux opérationnels ; leur utilisation par les entreprises est plus récente.

10.8.8. Problèmes de recherche

Les problèmes de recherche consistent à trouver le meilleur moyen d'obtenir les informations nécessaires à une décision. Bien que tout problème contienne un problème de recherche dans un sens, il existe des situations dans lesquelles la recherche elle-même est le processus essentiel; par exemple, dans la vérification des comptes, les procédures d'inspection et de contrôle de la qualité, l'exploration minière, la conception de systèmes d'information et les problèmes militaires impliquant la localisation de menaces telles que les navires, avions, mines et missiles ennemis.

Deux types d'erreurs sont impliqués dans la recherche: celles d'observation et celles d'échantillonnage. Les erreurs d'observation, quant à elles, sont de deux types généraux : la commission, qui consiste à voir quelque chose qui n'est pas là, et l'omission, qui consiste à ne pas voir quelque chose qui est là. En général, plus le risque de commettre l'une de ces erreurs diminue, plus le risque de commettre l'autre augmente. En outre, si des ressources fixes sont disponibles pour la recherche, plus l'échantillon est grand (et donc plus l'erreur d'échantillonnage est faible), moins il y a de ressources disponibles par observation (et donc plus l'erreur d'observation est grande).

Le coût de la recherche se compose du coût de mise en place ou de conception, du coût des observations, du coût d'analyse des données obtenues et du coût de l'erreur. L'objectif est de minimiser ces coûts en manipulant la taille de l'échantillon (quantité d'observations), le plan d'échantillonnage (comment les choses ou les lieux à observer sont sélectionnés), et la manière d'analyser les données (la procédure inférentielle).

Presque toutes les branches de la statistique fournissent des techniques utiles pour résoudre les problèmes de recherche. Dans les problèmes de recherche qui impliquent la localisation d'objets physiques, en particulier ceux qui se déplacent, la physique et certains domaines des mathématiques (par exemple, la géométrie et la trigonométrie) sont également applicables.

Un problème de "recherche inversée" se pose lorsque la procédure de recherche n'est pas maîtrisée mais que l'objet de la recherche l'est. La plupart des détaillants, par exemple, ne peuvent pas contrôler la manière dont les clients recherchent des marchandises dans leurs magasins, mais ils peuvent contrôler l'emplacement des marchandises. Ce type de problème se pose également dans la conception de bibliothèques et de systèmes d'information, ainsi que dans la pose de mines terrestres et marines. Il s'agit là aussi de problèmes de recherche, et les techniques de résolution décrites ci-dessus leur sont applicables.

10.8.9. Fonction de progression de la fabrication

En raison de l'énorme complexité d'une chaîne de production de masse typique et du nombre presque infini de changements qui peuvent être apportés et d'alternatives qui peuvent être poursuivies, un corps de théorie quantitative des systèmes de fabrication de production de masse n'a pas encore été développé. Cependant, le volume des données d'observation disponibles augmente, et des faits qualitatifs émergent qui pourraient éventuellement servir de base à une théorie quantitative. La "fonction de progression de la fabrication" en est un exemple. Cette fonction a été reconnue pour la première fois dans l'industrie aéronautique. Les premiers constructeurs d'avions ont observé qu'à mesure qu'ils produisaient un nombre croissant d'un modèle d'avion donné, leurs coûts de fabrication diminuaient de manière prévisible, d'abord de manière abrupte, puis à un rythme plus lent. Lorsqu'un graphique des coûts réels est tracé sur du papier à double logarithme, le logarithme du coût par unité en fonction du logarithme du nombre total d'unités produites donne des points de données qui forment presque une ligne droite. Au fil des ans, des relations similaires ont été trouvées pour de nombreux produits fabriqués par des techniques de production de masse. La pente de la ligne droite varie d'un produit à l'autre. Toutefois, pour une catégorie donnée de produits et un type donné de technologie de production, la pente semble remarquablement constante.

Les fonctions de progression de la fabrication peuvent être d'une grande valeur pour le fabricant, car elles constituent un outil utile pour estimer les coûts futurs. En outre, le fait que les coûts ne suivent pas une fonction de progression bien établie peut être le signe qu'il faut accorder plus d'attention à l'opération afin de rendre ses performances en matière de coûts conformes aux attentes.

Bien que les fonctions de progression de la fabrication soient parfois appelées "courbes d'apprentissage", elles reflètent bien plus que l'amélioration de la formation des opérateurs de fabrication. L'amélioration des compétences des opérateurs est importante au démarrage de la production, mais la majeure partie de l'amélioration des coûts à long terme provient des améliorations apportées à la conception du produit, aux machines et à la planification technique globale de la séquence de production.

10.10.10. Simulation

Les ordinateurs ont eu un impact considérable sur la gestion des systèmes de production industrielle et sur les domaines de la recherche opérationnelle et du génie industriel. La vitesse et les capacités de traitement des données des ordinateurs permettent aux ingénieurs et aux scientifiques de construire des modèles plus grands et plus réalistes de systèmes organisés et d'obtenir des solutions significatives à ces modèles grâce à l'utilisation de techniques de simulation.

La simulation consiste à calculer les performances d'un système en évaluant un modèle de celui-ci pour des valeurs de variables choisies au hasard. La plupart des simulations en recherche opérationnelle concernent des variables "stochastiques", c'est-à-dire des variables dont les valeurs changent de manière aléatoire dans une certaine distribution de probabilité au fil du temps. L'échantillonnage aléatoire employé dans la simulation nécessite soit une réserve de nombres aléatoires, soit une procédure pour les générer. Il faut également un moyen de convertir ces nombres en distribution de la variable concernée, un moyen d'échantillonner ces valeurs et un moyen d'évaluer la performance qui en résulte.

Une simulation dans laquelle la prise de décision est effectuée par un ou plusieurs décideurs réels est appelée "jeu opérationnel". De telles simulations sont couramment utilisées dans l'étude des interactions entre les décideurs comme dans les situations de compétition. Les jeux militaires sont utilisés depuis longtemps comme dispositif de formation, mais ce n'est que relativement récemment qu'ils ont été utilisés à des fins de recherche. Il reste cependant très difficile de tirer des conclusions des jeux opérationnels pour le monde réel.

L'optimisation expérimentale est un moyen d'expérimenter sur un système afin de trouver la meilleure solution à un problème qui s'y pose. Ces expériences, menées simultanément ou séquentiellement, peuvent être conçues de différentes manières, dont aucune n'est la meilleure dans toutes les situations.