

BASIC PLUS

ENCYCLOPEDIE D'INITIATION
A L'INFORMATIQUE ET AUX ORDINATEURS

12

DANS CE NUMERO, VOUS TROUVEREZ

LES CARTES-REponses DES DEUX
CONCOURS DE BASIC PLUS

CONCOURS

THOMSON

100 MICROS TO 7
A GAGNER

BASIC PLUS-GFI.

CONCOURS

VOCATION BULL

DE FANTASTIQUES
VOYAGES
EN CALIFORNIE

Plus de 480 000 F de PRIX

EDITIONS DU HENNIN

BASIC PLUS

ENCYCLOPEDIE D'INITIATION A L'INFORMATIQUE ET AUX ORDINATEURS

Parution hebdomadaire publiée par
les Editions du Hennin
20-22, rue de Clichy 75009 Paris.

Directeur général: Charles-Jean Pradelle.
Directeur délégué: Italo Milani.

Avec la collaboration de **Bull** et
Thomson Micro-Informatique Grand Public.

Direction éditoriale: Nathalie Perrin.

Traduction et adaptation
sous la direction de Martine Allain.

© Editions du Hennin S.A.R.L. 1984.
© Armando Curcio - Editore - Rome.
Titre original: Basic, Enciclopedia dell'Informatica
dei Mini e Personal Computer.
Réalisé par le Département Grands Ouvrages
d'Armando Curcio Editore, sous la direction
technique de Sante Senni et la direction éditoriale
de Gabriella Costarelli.

Dépôt légal: 1^{er} trimestre 1984, ISSN: en cours.
Composition: Typelec, Paris. Imprimé en Belgique
par Van Cortenbergh, Bruxelles.
Photo de couverture: écran réalisé par GFI (Groupe Français
d'Informatique) pour les concours de Basic Plus. Photo: Alain Garrigou.

Directeur de la publication: Maurice Brébart.

Administration

France: Éditions du Hennin, 20-22, rue de Clichy
75009 Paris. Tél. (1) 280.64.65.
Belgique: Femmes d'Aujourd'hui S.A., avenue Frans Van
Kalken, 9 B-1070 Bruxelles, Tél. (02) 523.20.60.

Distribution

France: NMPP. Belgique: AMP S.A. Suisse: NAVILLE S.A.

Vente au numéro

Les fascicules peuvent être obtenus chez les
marchands de journaux ou, à défaut, chez l'Éditeur, au
prix en vigueur au moment de la commande. Pour toute
commande par lettre, joindre le règlement (chèques
bancaire ou postal) majoré de 10% représentant la
participation aux frais de port.

Vente par souscription

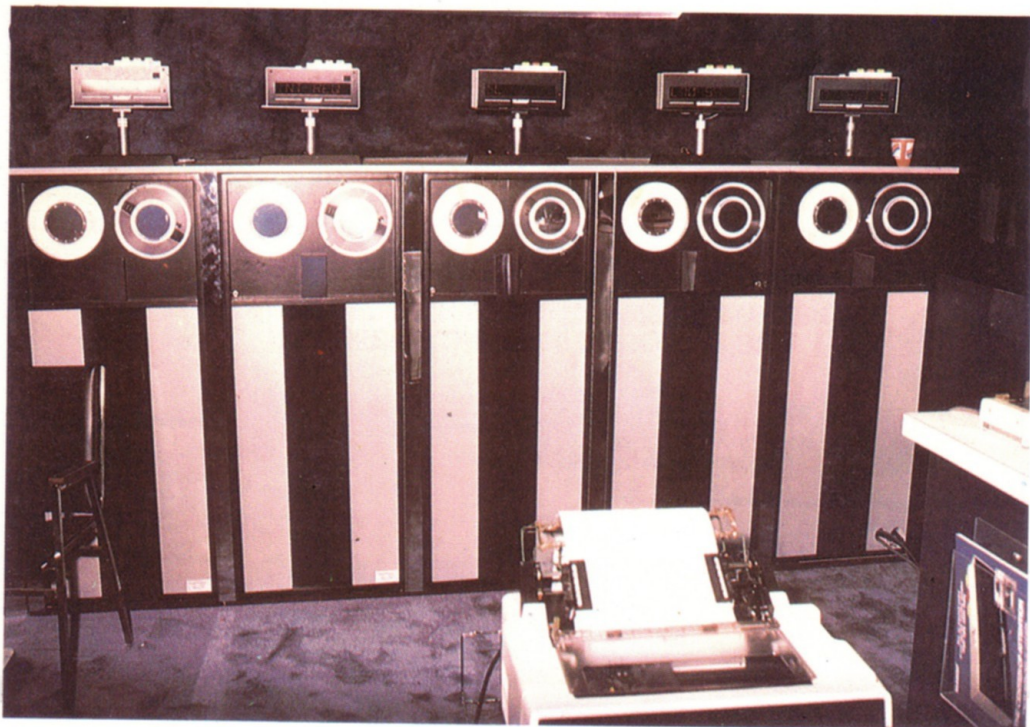
France: 72 fascicules + 6 reliures mobiles pour un prix
global de 1.200 FF dont 108,00 FF de frais de participation
au conditionnement et de port (pour la France
métropolitaine).

Payable par prélèvement automatique en
24 mensualités de 50,00 FF. Envoi de 4 fascicules à la
fois (x 18) sous étui cartonné. La première reliure mobile vous
parviendra avant la parution du douzième fascicule, qui
achève le premier volume. Adressez votre courrier à:
AZ Distribution, 20-22, rue de Clichy, 75009 Paris.
N.B. Avec votre commande, n'envoyez pas d'argent (ni
chèque, ni mandat). Toute souscription concerne l'intégralité
de l'ouvrage et commence au fascicule n° 1.

Belgique: Basic Plus, avenue Frans Van Kalken, 9,
B-1070 Bruxelles.

BASIC PLUS

*Le succès de BASIC PLUS
a fait que les premiers numéros se sont vite trouvés épuisés.
Prévenez vos amis qui n'ont pu se les procurer:
un coffret-cadeau contenant les 12 premiers fascicules avec
leur reliure mobile sera en vente dès la semaine prochaine
dans les maisons de la presse.*



Marka

Vue partielle (dérouleurs de bande) de la mémoire de masse d'une banque de données.

Inversement si, après la donnée 176, nous déclenchons par erreur l'exécution du module 1000, la donnée suivante sera écrite en position 1 (puisque le compteur aura été remis à 1 par le module 1000) et non, comme elle aurait dû l'être, en position 177. En s'écrivant par-dessus le précédent contenu de la position 1, elle l'« écrasera », c'est-à-dire l'effacera. En outre, la machine ne connaîtra plus le nombre des données réellement présentes.

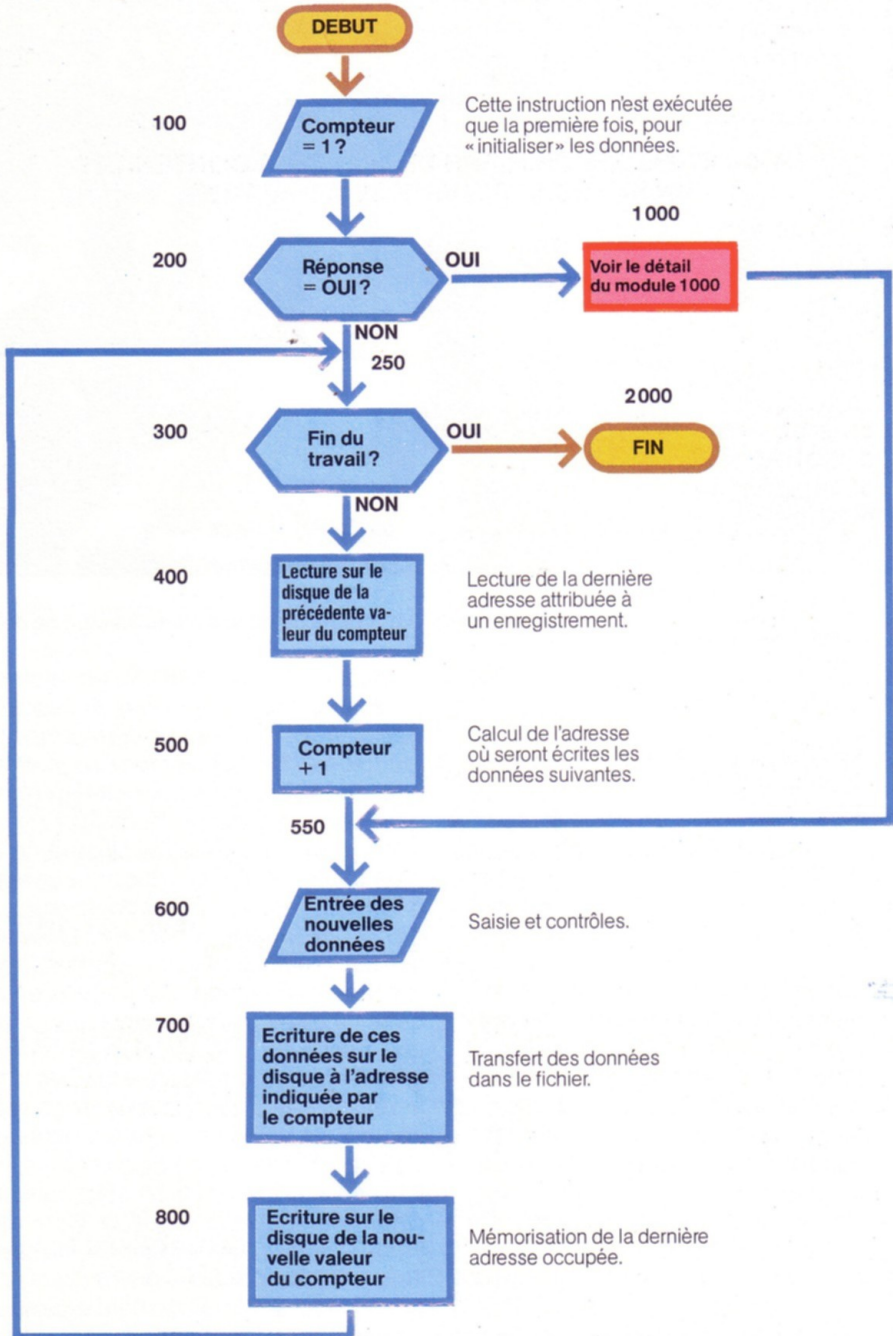
La seule façon de réduire ces risques d'erreur est de subordonner l'exécution du module 1000 à la reconnaissance d'un mot de passe et à une demande de confirmation. On a recours à cette double vérification toutes les fois qu'une distraction ou une mauvaise utilisation risqueraient d'entraîner la perte des données.

Dans notre exemple, la perte des données n'est qu'apparente ou, du moins, limitée, puisqu'il suffit de remettre le compteur à 176 pour les récupérer. Dans d'autres cas, elle est irréversible et lourde de conséquences. Qu'on essaie d'imaginer ce que signifie la destruc-

tion des fichiers contenant les informations économiques d'une grande entreprise. L'organigramme de la page 264 représente le module 1000 amélioré : les contrôles appropriés ont été ajoutés pour protéger les données contre les erreurs humaines.

Venons-en maintenant au second défaut du programme. Imaginons que nous en sommes à l'enregistrement 121 au cours d'une saisie de données. Cette valeur n'a pas été mémorisée sur le disque puisque, comme le montre le schéma de la page 263, ce transfert n'a lieu qu'en fin de travail. C'est alors que se produit une panne de courant. Puis le courant est rétabli. Mais le compteur n'est plus à 121 (il est à zéro, à « blanc » ; ou, pire, il contient un nombre aléatoire) et toutes les données saisies avant la panne ne peuvent plus, malgré leur présence dans le fichier, être considérées comme acquises puisque le système en a perdu l'adresse (la position). Contre ce risque, il n'y a qu'une solution : mémoriser le compteur à chaque introduction de données. L'organigramme définitivement corrigé est représenté page 266.

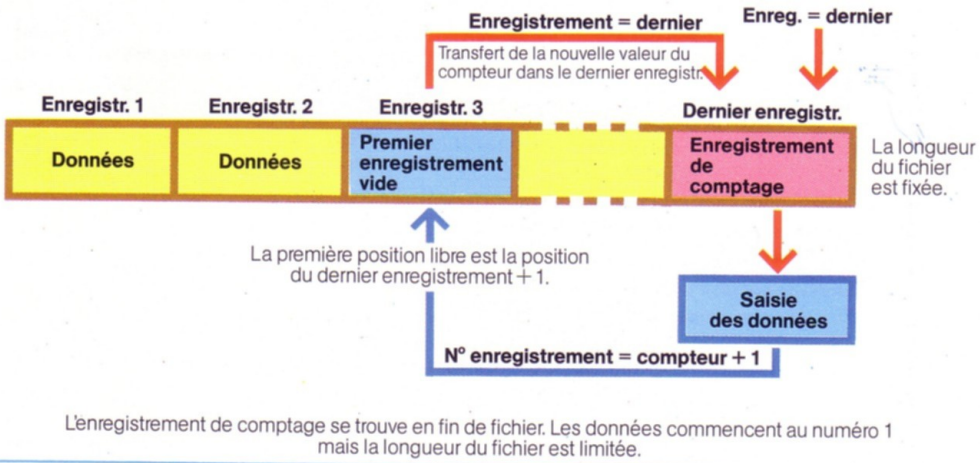
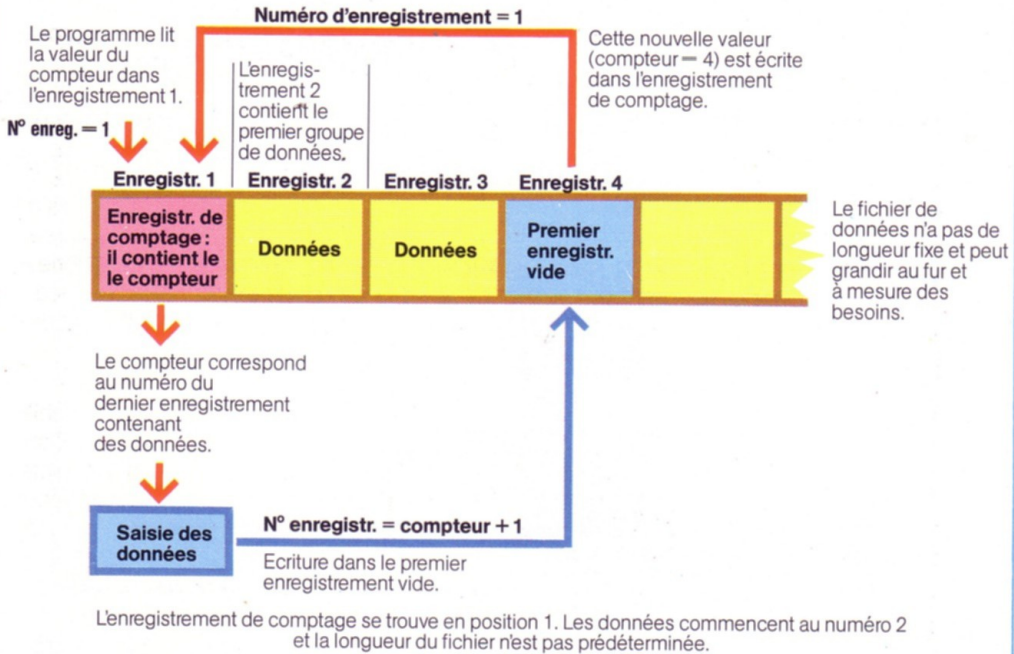
ORGANIGRAMME DE SAISIE DES DONNEES MODIFIE POUR LA MISE A JOUR DU COMPTEUR A CHAQUE ENTREE



C'est alors que se pose la question de savoir où va s'inscrire le compteur. Il est possible d'utiliser un enregistrement donné du fichier, dont toutes les zones seront vides, excepté celle du compteur. Bien qu'appartenant au fichier de données, cet enregistrement n'est qu'une zone de comptage temporaire dans lequel on peut lire ou écrire

le compteur. S'il est placé au début du fichier, le premier enregistrement disponible pour l'écriture des données est alors le numéro 2 et la donnée 1 n'existe pas. Cette solution permet de ne pas spécifier d'avance la longueur du fichier qui pourra s'accroître éventuellement au fur et à mesure des besoins.

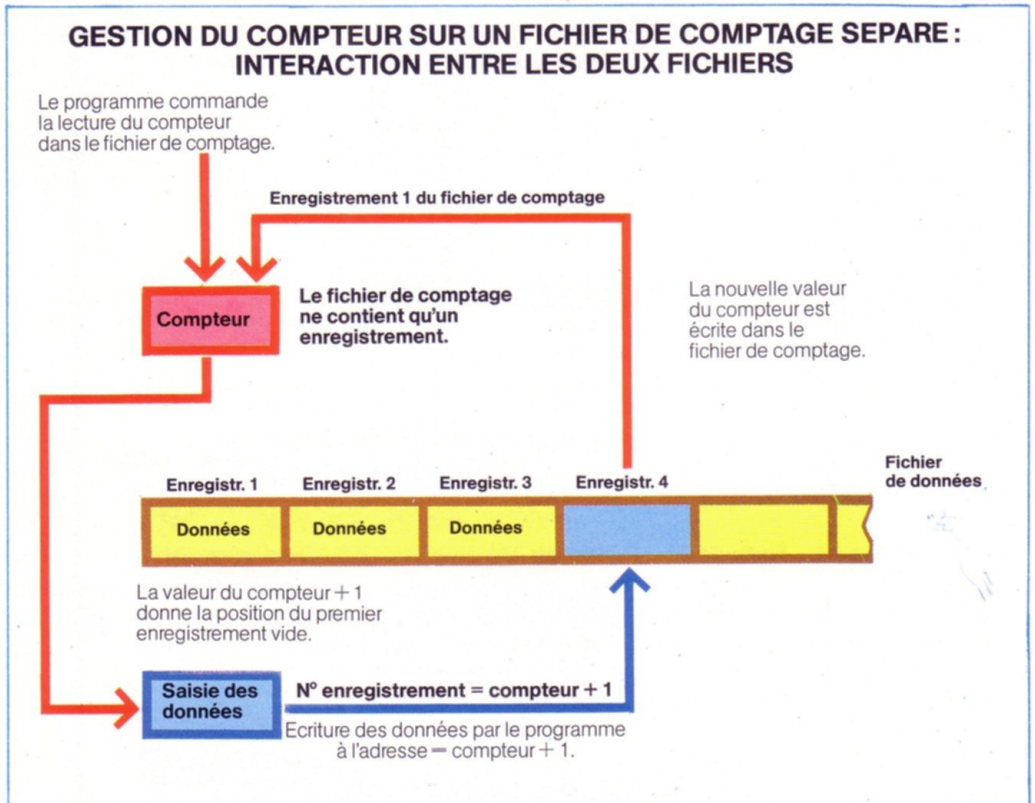
STRUCTURE DES FICHIERS DE DONNEES CONTENANT UN ENREGISTREMENT DE COMPTAGE



En revanche, si l'enregistrement de comptage se situe à la fin du fichier de données, la numérotation commence au numéro 1 puisque le premier enregistrement n'est plus destiné au comptage. Il faut déterminer à l'avance la longueur du fichier, l'adresse de l'enregistrement de comptage étant celle du dernier enregistrement du fichier. Les schémas de la page 267 représentent ces deux possibilités et les voies d'accès aux enregistrements correspondants.

Une seconde solution consiste à créer un nouveau fichier ne contenant qu'un enregistrement: on y mémorise le compteur. Les données sont alors numérotées à partir de 1, et l'on conserve les avantages d'un fichier de longueur indéterminée. Mais on doit alors gérer simultanément deux fichiers: celui des données et celui du comptage.

Le schéma ci-dessous illustre cette structure



et son utilisation. Quelle que soit la méthode choisie, lecture et écriture du compteur sont effectuées par les modules 400 et 800 de

l'organigramme général (page 266). En revanche, le contenu des modules varie en fonction de la méthode adoptée. On peut voir ci-dessous le module de lecture tel qu'il se présente quand on a recours à un fichier de comptage appelé DIR; le module d'écriture (800) se décomposerait de façon analogue.

LECTURE DE LA VALEUR DU COMPTEUR DANS LE FICHIER DIR

Module : 400.

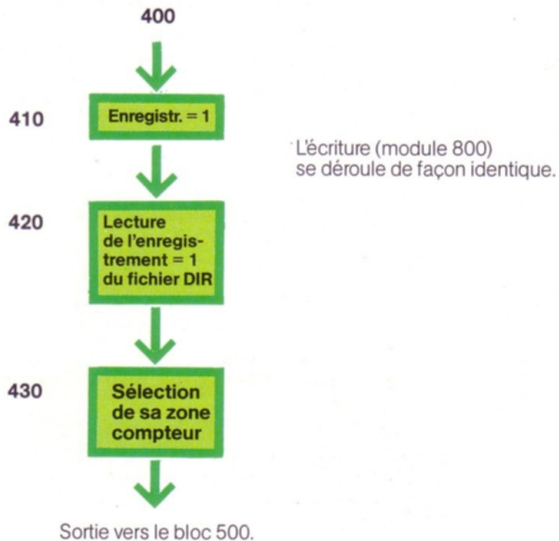
Fonction : lecture de la zone compteur dans le fichier DIR.

Nom du fichier : DIR.

Longueur : 1 enregistrement (60 caractères).

Position de la zone compteur : 1 - 3.

Sortie : valeur numérique du compteur.



Le module de contrôle des données en entrée

Lors de l'introduction (saisie) des données, il faut contrôler l'exactitude formelle des zones Date, Justification et Montant.

Cette opération (détaillée page 270) doit s'effectuer dans le module d'introduction des données (600 sur le schéma de la page 266), afin que seules des données exactes puissent être acheminées vers la suite du programme. Le module (ou « bloc ») 620 concernant le contrôle de la date, fera l'objet d'un développement ultérieur car il revient souvent dans les programmes. Considérons, pour l'instant, qu'il permet de vérifier la compatibilité entre le jour et le mois, et de répondre par une variable d'erreur. Si celle-ci est nulle, la date est exacte; sinon elle est erronée et la

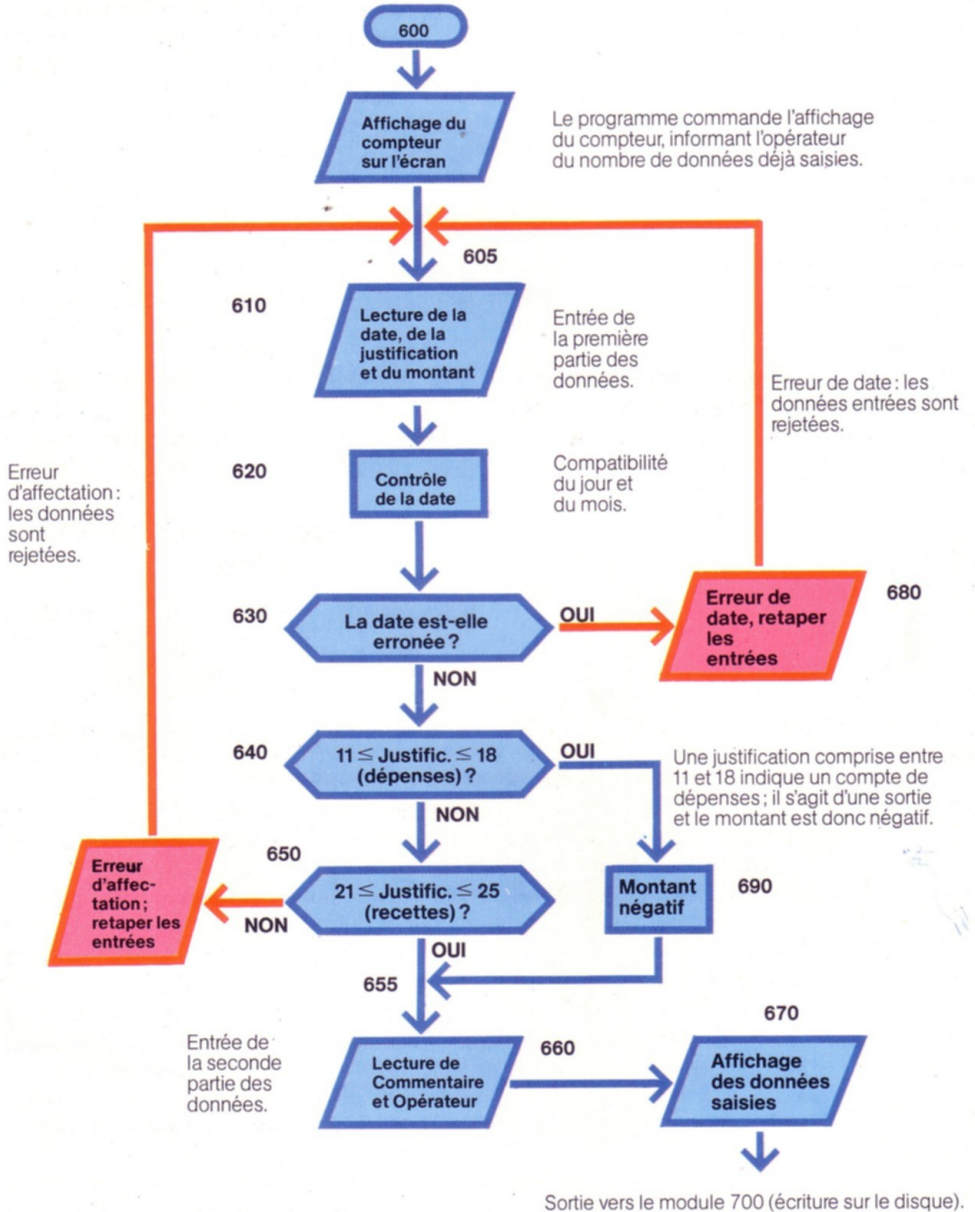
ORGANIGRAMME DETAILLE DU MODULE DE SAISIE ET DE CONTROLE DES DONNEES

Module : 600.

Fonction : entrée au clavier des nouvelles données et contrôle formel.

Entrée : compteur.

Sortie : zones de données prêtes à être transférées sur disque.



valeur de la variable renseigne alors sur le type d'erreur commis. Ce module peut également servir à limiter dans le temps l'utilisation des programmes. Si, par exemple, nous posons comme condition que l'année ne peut être postérieure à 1995, le programme n'acceptera plus les saisies en 1996. Dans le cas où un utilisateur ne dispose que de la version compilée du programme (c'est-à-dire en langage machine), il ne peut modifier lui-même cette date limite. Seule peut le faire la personne qui détient le programme source. Naturellement, après les corrections, on devra recompiler le programme.

La seule ressource de l'utilisateur serait de fournir une fausse date à l'ordinateur, en continuant, par exemple, de taper 1995 au lieu de 1996. Naturellement, cette solution n'est envisageable que pour les programmes à usage personnel et certainement pas sur des programmes d'utilisation officielle (fiscale par exemple).

Cette technique de péremption permet de louer des programmes dont la date d'échéance ne sera remise à jour que si l'utilisateur a payé sa redevance.

Le module 600 permet la saisie des données. Il faut, à présent, compléter la procédure par le développement des calculs.

Ce nouveau traitement doit lire tout le fichier de données et additionner séparément les montants des dépenses et des recettes. Nous aurons ainsi une vue générale des mouvements effectués et le résultat global (la différence entre le total des recettes et le total des dépenses). En plus de ce résultat, il faut prévoir la possibilité d'effectuer des bilans partiels entre deux dates déterminées. Nous pourrions ainsi demander soit une situation globale (prenant en compte les données du début à la fin), soit la situation d'une période déterminée.

L'organigramme de ce traitement est représenté page 272.

Les programmes « auxiliaires » concernant des données particulières (justification, commentaire, opérateur) feront l'objet d'une explication générale ultérieure.

Quant aux modules qui apparaissent page 272, ils seront développés en détail lorsque nous rédigerons le programme après avoir étudié les instructions Basic nécessaires.

En résumé :

- la structure de l'enregistrement de données et les longueurs des zones sont schématisées page 262 ;
- l'organigramme du programme de saisie et de contrôle des données complet se trouve page 266 ;
- l'initialisation est effectuée par le module 1000 dont l'organigramme se trouve page 264 ;
- le compteur de données est mémorisé dans un fichier auxiliaire de comptage (voir schémas pages 268 et 269) ;
- les opérations de saisie et de contrôle des données sont détaillées page 270. En traduisant dans l'ordre les organigrammes des pages 264, 266, 269 et 270 en instructions, nous obtiendrons notre programme.

La recherche des données dans les fichiers

Une fois le fichier constitué (phase de création), il faut pouvoir modifier ses données (phase de mise à jour).

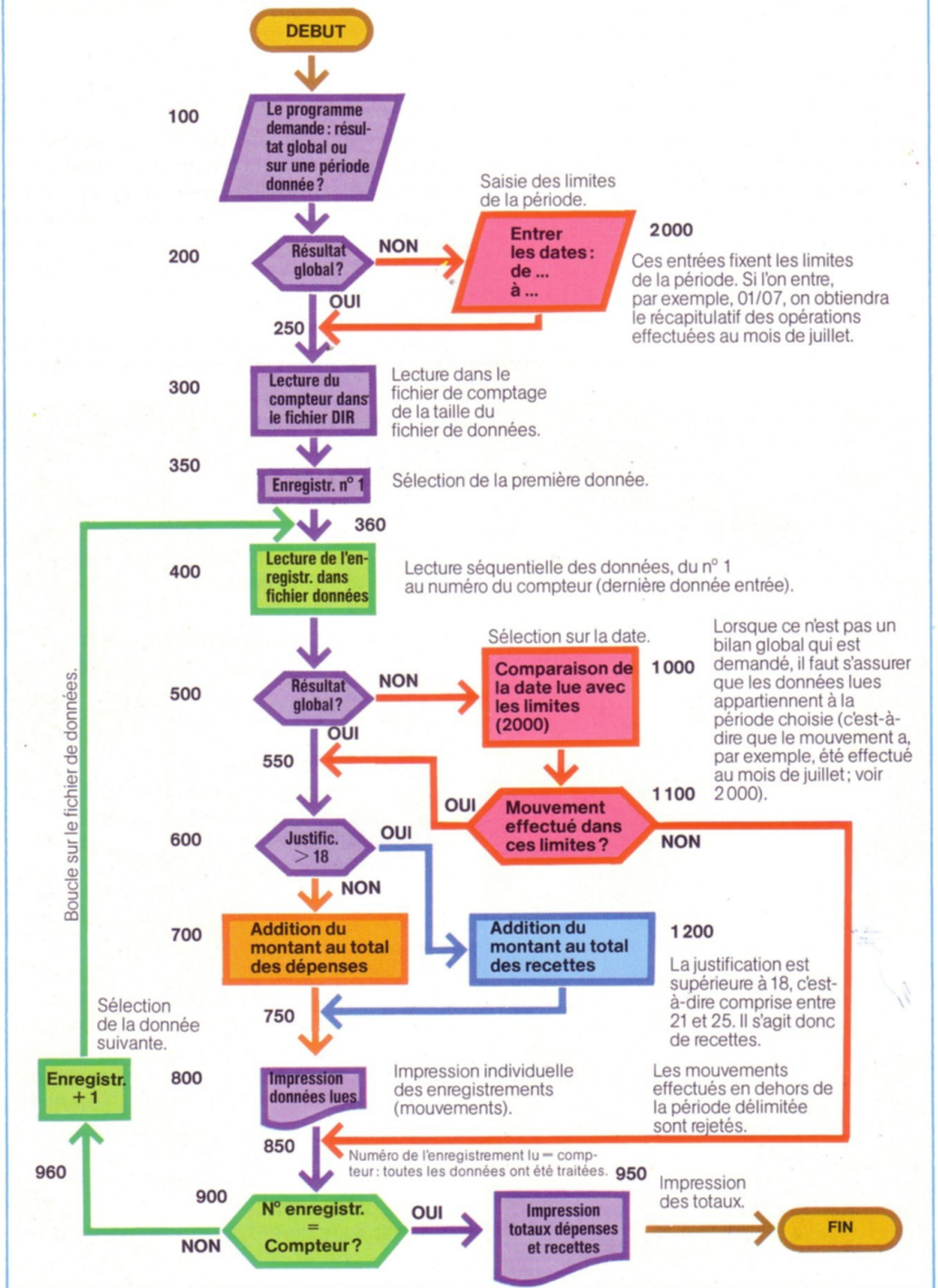
Pour mettre à jour un enregistrement, on doit d'abord le lire, puis, en fonction de son contenu, effectuer les ajouts et les modifications qui s'imposent.

Accéder aux données selon une logique déterminée constitue la phase principale d'utilisation d'un fichier.

Pour la mise à jour, l'utilisateur recherche l'enregistrement possédant la clé voulue et en extrait les données correspondantes pour les modifier. Pour un traitement particulier (listing par exemple), il s'agit d'extraire du fichier une partie des données. On peut ainsi rechercher dans un fichier d'adresses celles qui se situent dans une même ville, ou dans un fichier du personnel tous les employés d'une même catégorie.

Il est donc indispensable de disposer de programmes de recherche qui retrouvent tous les enregistrements dont les données répondent

ORGANIGRAMME DU PROGRAMME DE CALCUL ET D'IMPRESSION DU BUDGET



aux critères voulus.

Les deux méthodes de recherche les plus utilisées sont la recherche séquentielle et la recherche dichotomique.

La recherche séquentielle

La recherche séquentielle d'une donnée implique la lecture successive de tous les enregistrements du fichier et leur comparaison avec la valeur de la clé de recherche. S'il y a égalité, la donnée trouvée est imprimée, sinon la recherche se poursuit jusqu'à la fin des enregistrements.

Le nombre moyen de lectures à effectuer

s'obtient par la formule suivante :

$$\text{nombre moyen de lectures} = \frac{\text{longueur du fichier} + 1}{2}$$

Dans un fichier de 1000 enregistrements, il nous faudra donc en lire 500 en moyenne avant de trouver celui qui nous intéresse. De plus, comme nous l'avons vu, la donnée recherchée peut très bien apparaître dans plusieurs enregistrements; il faudra donc poursuivre la recherche jusqu'à la fin du fichier. Le seul moyen d'accélérer cette opération est d'ordonner préalablement les données.

Deux étapes du travail effectué sur les banques de données. L'information doit être parfaitement planifiée et structurée afin d'utiliser au mieux la capacité de la mémoire.



Marka

La recherche séquentielle s'avère très longue. Aussi ne l'applique-t-on qu'à des petits fichiers, pour lesquels le gain de temps permis par des méthodes plus rapides, mais difficilement mises en œuvre, serait négligeable.

La recherche dichotomique

Elle ne peut se pratiquer que sur des fichiers ordonnés. On procède à des divisions binaires (en deux parties) successives du fichier. On lit ainsi l'enregistrement situé au milieu, on compare sa clé à la clé de recherche et on rejette alors l'une des parties (la droite si la clé de recherche est inférieure à celle de l'enregistrement lu, la gauche dans le cas contraire, en supposant le fichier classé par ordre croissant). On répète l'opération jusqu'à la découverte de la donnée recherchée. Prenons un exemple (voir schéma page 274).

- Zone clé, 1 caractère alphabétique.
- Prix des articles, 6 caractères numériques*.
- Description de l'article, 20 caractères alphabétiques.
- Quantité en stock, 3 caractères numériques*.

L'utilisateur veut connaître le contenu de l'enregistrement dont la clé est P. La recherche dichotomique s'effectue comme suit.

On divise d'abord (point 1) la longueur du fichier par 2 ; ici, $14 : 2 = 7$.

Ensuite (point 2), la clé de l'enregistrement 7 (N) est comparée à la clé de recherche (P). Or, P est placé après N ($P > N$) ; l'enregistrement de clé P se trouve donc à droite, entre les enregistrements 7 et 14 (fin du fichier). Cette zone est à son tour divisée en deux. $(14 - 7) : 2 = 3$ (on ne prend que la partie entière du quotient). On obtient, en nombre d'enregistrements, l'ampleur du déplacement à effectuer vers la droite, à partir du n° 7.

En se décalant, vers la droite, de trois enregist-

tements à partir du numéro 7, on va lire l'enregistrement $7 + 3 = 10$ (point 3).

Cet enregistrement ayant une clé supérieure à P (point 4), celui qui est recherché se trouve, cette fois, à sa gauche. De même que tout à l'heure notre donnée se trouvait entre les enregistrements 7 et 14, on la localise maintenant entre les enregistrements 7 et 10. Un nouveau découpage en deux de la partie à étudier $(10 - 7) : 2 = 1$ indiquera le déplacement à effectuer vers la gauche à partir de la position actuelle.

L'enregistrement $10 - 1 = 9$ qui est lu contient P dans sa zone clé (point 5). La recherche s'achève et les données de l'enregistrement 9 sont affichées à l'écran.

Avec cette méthode, on n'a lu que trois des quatorze enregistrements du fichier (7, 10 et 9) alors qu'en recherche séquentielle nous aurions dû effectuer neuf lectures (l'enregistrement de clé P et les huit précédents). L'avantage de la recherche dichotomique est évident.

Une réserve s'impose cependant. On ne peut faire une recherche dichotomique que sur un fichier ordonné. Sinon, il faudra d'abord en classer les données. Cette méthode dichotomique impose un programme plus complexe que la méthode séquentielle et ne fait donc pas forcément gagner de temps lorsque le fichier ne comporte qu'un petit nombre d'enregistrements.

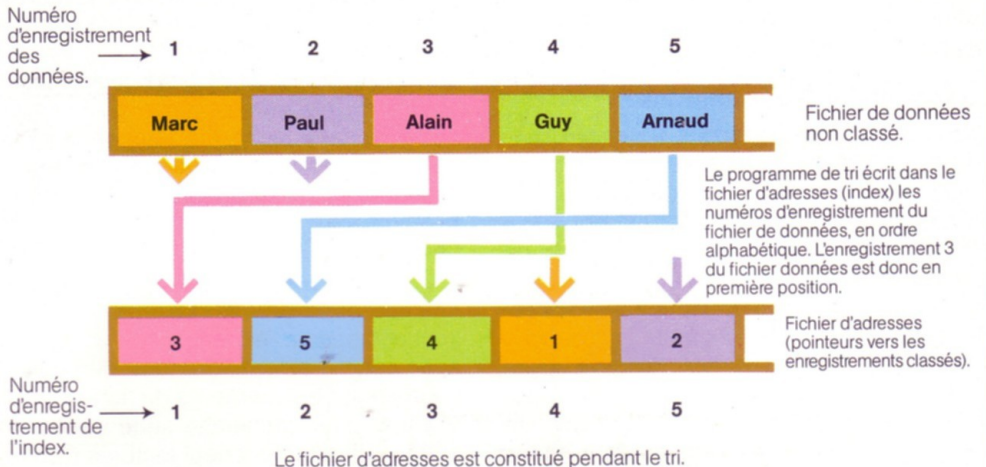
Le classement des données

Le classement des données est l'une des tâches principales de la gestion des fichiers. C'est, en effet, une condition fondamentale pour l'application de certaines techniques de traitement (comme la recherche dichotomique) et pour l'obtention de listings faciles à consulter.

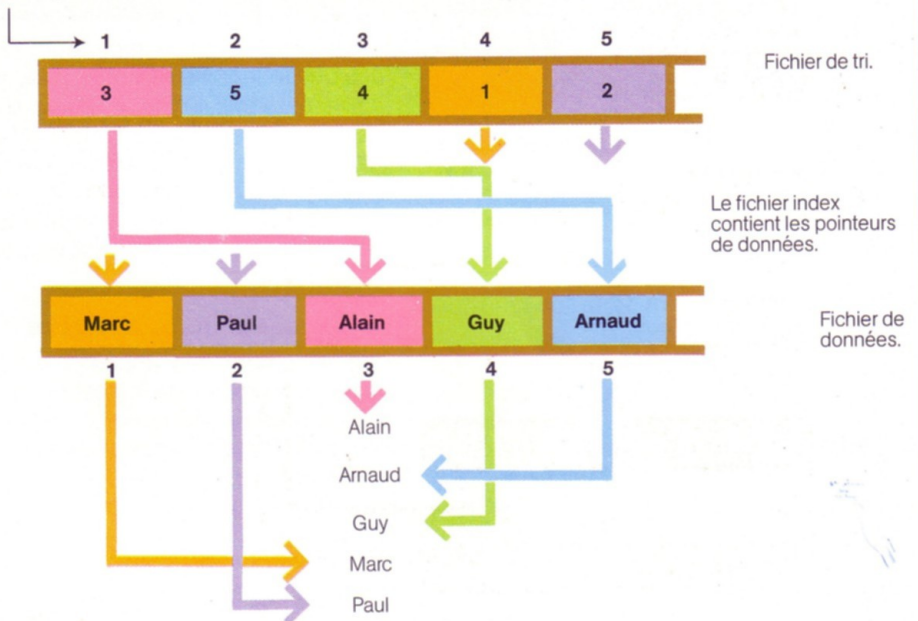
Lorsqu'on ne peut obtenir de listing des données classées, un fichier, même de modestes dimensions, s'avère très difficile à exploiter, car il est pratiquement impossible d'y trouver les éléments que l'on recherche. C'est un peu comme si l'on disposait d'un annuaire téléphonique où les noms ne seraient pas répertoriés par ordre alphabétique mais dans l'ordre chronologique de souscription du contrat d'abonnement ! M. Martin, abonné de longue date, figurerait donc avant M. Leblanc et il serait impossible de retrouver

* Dans les ordinateurs individuels, les nombres sont souvent enregistrés sur les disques comme des caractères. Le nombre 579, par exemple, est enregistré sur trois octets, qui contiennent, dans l'ordre, les chiffres 5, 7 et 9, codés comme des caractères, respectivement par les valeurs hexadécimales 35, 37 et 39 (se reporter au code ASCII). Avec d'autres ordinateurs, on peut, au contraire, enregistrer les nombres tels quels sur les disques, ce qui occupe moins de place. Nous reviendrons plus loin sur cette question.

TRI DES DONNEES A L'AIDE D'UN FICHER INDEX



Sa lecture séquentielle fournira les numéros d'enregistrement des données dans l'ordre de classement adopté.



Pour obtenir une liste classée des données du fichier, il faut lire l'index séquentiellement et utiliser son contenu comme pointeur vers le fichier de données.

un numéro de téléphone donné puisqu'aucune logique de recherche ne serait applicable en temps réel. De même, pour l'utilisation des données d'un fichier; c'est pourquoi toutes les procédures de gestion de fichiers

possèdent des programmes de classement des données, appelés des programmes de tri.

L'importance de ces programmes est telle

TRI DES DONNÉES PAR INSERTION DANS UN FICHER PARALLELE

1 **B L A O N R C** Fichier écrit dans l'ordre des entrées.

2 **B L**

3 **B L**

3-1 **A B L**

4 **A B L O**

5 **A B L O**

5-1 **A B L N O**

6 **A B L N O R**

7 **A B L N O R**

7-1 **A B C L N O R** Fichier de données trié.

Fichier parallèle dans lequel les données sont transférées de façon ordonnée.

qu'ils sont inclus dans presque tous les systèmes d'exploitation, et généralement écrits en Assembleur pour accroître leur rapidité d'exécution.

A ces programmes de tri « utilitaires », c'est-à-dire accessibles à l'utilisateur, ce dernier ne fournit que le nom du fichier, la structure de l'enregistrement (nombre et longueur des zones de données) et la clé en fonction de laquelle il veut ordonner le fichier.

On distingue deux types de programme de tri suivant la structure des résultats qu'ils obtiennent. Certains recopient le fichier, en l'ordonnant, sur une nouvelle partie du disque, créant ainsi un nouveau fichier ordonné, et sous un nom différent du premier. D'autres constituent un fichier d'adresses qui contient (dans l'ordre) les adresses des données du fichier d'origine.

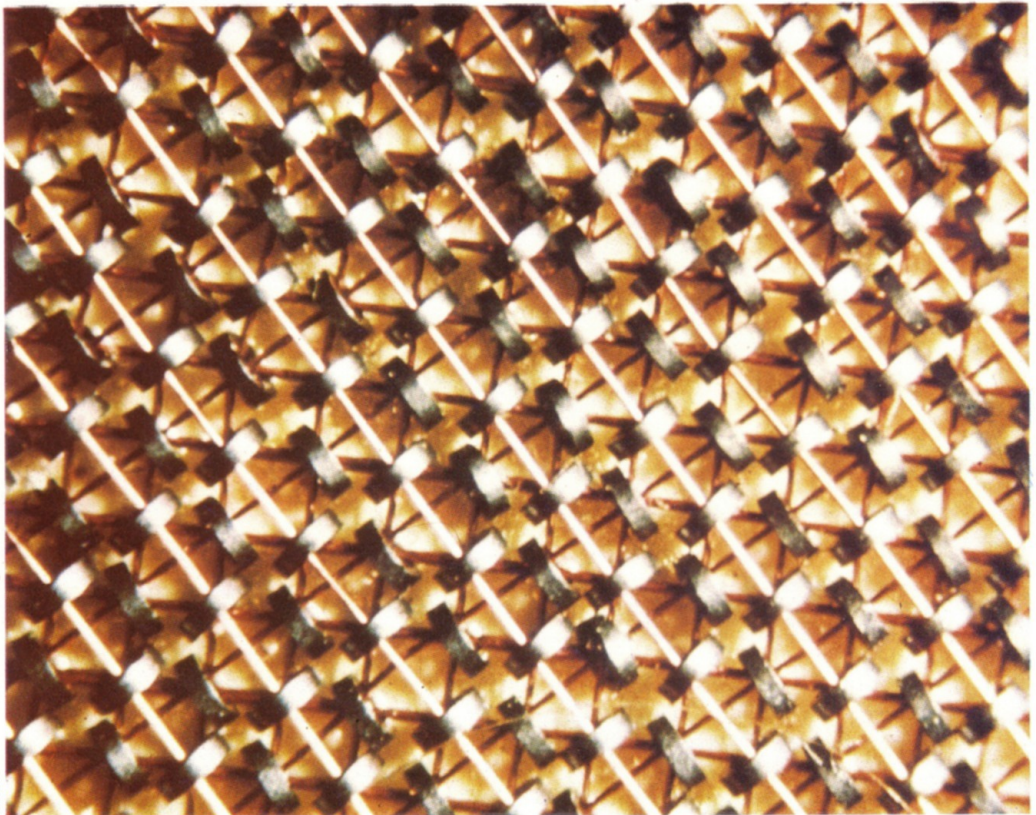
Pour lire les données dans l'ordre, il faudra donc, dans ce dernier cas, lire d'abord le fichier d'adresses, puis l'enregistrement de données

pointé par l'adresse trouvée. Cette méthode est illustrée par le schéma de la page 276.

Cette méthode est plus rapide (puisque l'on n'a pas à recopier tout le fichier mais seulement les adresses) et elle permet d'occuper moins de place sur le disque en évitant de recopier des données déjà existantes (redondance). Si tous les mini-ordinateurs sont proposés avec des programmes de tri, certains micro-ordinateurs en sont dépourvus. C'est alors le programmeur qui doit se charger de les écrire lui-même. Ce qui implique une méthode de travail rigoureuse, que nous allons étudier en l'illustrant d'un exemple de tri par insertion (algorithme) avec constitution d'un fichier « parallèle » (recopie). On se reportera au schéma de la page 277.

1 / Après avoir lu le premier enregistrement de données (qui contient B), on le transfère dans le premier enregistrement du fichier de sortie, c'est-à-dire du fichier parallèle qui

Détail d'un ancien dispositif de mémoire centrale à tores de ferrite.



Marika

Test 7



- 1 / Laquelle de ces affirmations s'applique à un fichier séquentiel ?
 - a) on peut accéder aux enregistrements dans n'importe quel ordre ;
 - b) les données sont écrites dans l'ordre croissant de la zone clé ;
 - c) les enregistrements sont écrits et lus l'un à la suite de l'autre.

- 2 / Lesquelles de ces affirmations s'appliquent à un fichier indexé ?
 - a) l'index est une zone spéciale qui contient des informations sur les enregistrements ;
 - b) il permet d'ordonner rapidement les données ;
 - c) un fichier indexé peut posséder plusieurs index.

- 3 / En quoi consiste un tri ?
 - a) à extraire les données dans l'ordre ;
 - b) à chercher une valeur déterminée parmi les données d'un fichier ;
 - c) à recopier un fichier.

- 4 / Enumérez les principales différences entre recherches séquentielle et dichotomique.

- 5 / Un fichier du personnel doit contenir : le prénom, le nom, la qualification, la catégorie, l'ancienneté, le salaire de base, les jours d'absence.
Définissez le format de l'enregistrement en choisissant des longueurs pour les différentes zones.

- 6 / Définissez la structure des enregistrements d'un fichier de gestion des stocks qui doit contenir les informations suivantes : article, quantité en stock, cote d'alerte, date du dernier prélèvement, nom et prénom du fournisseur.

- 7 / Pour la question 6, on peut gagner de la place en codant certaines zones. Essayez de coder la date de prélèvement (jour, mois, année) et le nom du fournisseur (on supposera qu'il y a 100 fournisseurs au maximum).

Les solutions du test se trouvent pages 286 et 287.

- contiendra les données ordonnées.
- 2 / Le deuxième enregistrement (qui contient L) est transféré en deuxième position du fichier de sortie puisque son contenu est supérieur à celui du premier ($L > B$).
 - 3 / L'enregistrement 3 du fichier de données contient A, c'est-à-dire une donnée inférieure à celle qui sont déjà présentes dans le fichier de sortie (B et L). Il faut donc le mettre en position 1.
 - 3-1 / La position 1 étant déjà occupée par la lettre B, il faut la libérer en décalant le contenu du fichier de sortie d'une position vers la droite. Ainsi, la lettre L passe de la position 2 à la position 3, tandis que B vient en position 2.
 - 4 / L'enregistrement 1 est disponible pour A. La lettre O (4^e donnée) est placée en dernière position puisqu'elle est supérieure à toutes les autres lettres du fichier de sortie.
 - 5 / Cette fois, le nouvel enregistrement à recopier (N) doit être placé avant O que l'on décale donc d'une position (étape 5-1).
 - 6 / La lettre R (enregistrement 6 du fichier de données) est placée en dernière position.
 - 7 / Pour insérer la lettre C (enregistrement 7), il faut décaler les quatre derniers enregistrements du fichier de sortie (L, N, O, R) afin de libérer une place entre B et L.
 - 7-1 / Le dernier enregistrement de données (C) est inséré à cette place et le fichier de sortie est complet.

Les banques de données spécialisées

Dans un monde en constante évolution, la **demande d'information** en matière économique et scientifique, tant dans le secteur privé que public, ne cesse de s'accroître.

Or la masse de documentation publiée chaque jour dans le monde entier est telle qu'il est devenu impossible de s'en tenir régulièrement informé. A moins d'en faire un véritable métier.

Les organisateurs de bases et de banques de données rassemblent, sélectionnent et mémorisent une abondante documentation sous une forme condensée. Certaines banques de données extrêmement spécialisées ont choisi de centraliser toute la littérature qui se rapporte à un unique secteur, mais elles sont rares et les banques à caractère pluridisciplinaire sont aujourd'hui beaucoup plus nombreuses. Enfin, mentionnons les banques de données collectant des **informations statistiques, démographiques et commerciales**. Elles sont très nombreuses et leur utilité augmente d'autant plus que l'arrivée des informations se fait à un rythme sans cesse croissant.

Le documentaliste qui travaille dans un petit centre ne peut consulter que quelques-unes des publications susceptibles de l'intéresser. Au contraire, s'il a la possibilité d'interroger une banque de données, il obtient une liste à jour des **références bibliographiques** concernant son sujet et ne laisse ainsi échapper aucune information utile.

Il existe également des banques de données spécialisées dans la **diffusion des brevets**, et qui permettent aux entreprises, même de petite taille, de se tenir au courant de l'apparition, dans leur secteur d'activité, de procédés ou de produits nouveaux et constituent soit un risque de concurrence soit, au contraire, une heureuse opportunité.

Les banques de données favorisent également la transmission rapide des informations dans des domaines aussi variés que la médecine, le commerce international, le droit, entre autres.

Le marché de l'information est particulièrement développé dans les pays dont la structure de production privilégie les industries de

transformation, et qui sont constamment éperonnés par les progrès accomplis à l'extérieur, tant sur le marché des facteurs de production (nouvelles technologies) que sur celui des produits finis. Une telle économie devrait mobiliser toutes ses ressources dans l'innovation technologique et commerciale afin de conserver et (si possible) d'améliorer sa position. Les premières entreprises qui se sont intéressées à la « **documentation en ligne** », sont apparues aux Etats-Unis, pays où des systèmes très performants de traitement des données pouvaient s'appuyer sur le réseau de télécommunications le plus moderne du monde. Lorsque, par la suite, ces services se sont développés et que le nombre des utilisateurs potentiels s'est accru, les sociétés américaines ont commencé à proposer leurs fichiers de l'autre côté de l'océan. Entre-temps, dans tous les pays à haut niveau technologique, de telles entreprises d'information ont été créées en Europe sur le modèle américain, proposant des fichiers comparables, ou revêtant (INSEE) un caractère national et rassemblant des données originales.

Pour permettre à ce patrimoine de se développer sur un plan européen et pour acquérir leur autonomie technologique, les services de Postes et Télécommunications de la CEE ont créé ensemble le premier réseau européen, EURONET. Il est le fruit d'un accord aux termes duquel la libre concurrence est établie entre toutes les banques de données européennes. Les utilisateurs y accèdent, en effet, de façon indifférenciée. Ainsi, l'ouverture des frontières, souvent illusoire dans le domaine des échanges commerciaux, se réalise dans celui des télécommunications.

Opérationnel depuis 1980, EURONET offre le support des télécommunications aux banques de données européennes réunies sous le label DIANE (Direct Information Access Network in Europe). Son caractère d'internationalité est ce qui fait la particularité de ce réseau.

L'utilisateur qui désire interroger un fichier situé dans l'un des pays membres souscrit un contrat d'utilisation d'EURONET et doit disposer d'un terminal connectable au réseau téléphonique. Le raccordement au réseau est ensuite pris en charge automatiquement par les ordinateurs qui constituent les différents

« nœuds » (points d'accès).

Il suffit à l'utilisateur européen de téléphoner au nœud d'une grande ville de son pays pour pouvoir, sans autre frais que celui d'une communication téléphonique interurbaine, se relier aux banques de données d'Allemagne, de France, de Grande-Bretagne, de Hollande, d'Italie, de Belgique ou du Luxembourg. La distance ne retentit pas sur le prix, et l'utilisateur n'a à payer que son appel. De plus, les tarifs d'EURONET sont très souvent inférieurs aux tarifs américains.

Toutefois, il est à peine plus compliqué de consulter les banques de données situées, aux Etats-Unis. Il suffit de s'abonner au réseau de télécommunications assurant le lien entre le pays de départ et les nombreux réseaux américains (networks). A ces frais assez peu élevés s'ajoutent ceux du contrat avec le ou les réseaux américains qui, d'ailleurs, comme la plupart des banques de données reliées à EURONET, ne facturent que l'utilisation effective et n'imposent pas de minimum.

Les fichiers accessibles par ces réseaux ont des contenus, et donc des utilisations, extrêmement variés. Il faut établir une distinction fondamentale entre banques de données et bases de données, en fonction de leur contenu.

Les **bases de données** sont des fichiers organisés que l'on peut interroger à partir d'un terminal et qui contiennent ce qu'on appelle des publications secondaires. Il s'agit de comptes rendus synthétiques de revues, de monographies, de livres, d'exposés de conférences, destinés à fournir à des usagers très différents un moyen de sélection efficace des documents traitant d'un sujet précis.

Elles sont souvent établies par des organismes (associations liées à un secteur de production, institutions universitaires ou encore entreprises privées spécialisées dans le traitement de l'information) qui éditent déjà des résumés de revues ou d'articles. Parfois les fichiers couvrent alors plusieurs disciplines. Pour chaque article, citation, ou monographie recensés, on indique : le titre (en anglais et dans la langue d'origine si elle est différente de l'anglais) ; l'auteur ; la source (revue, article, monographie) et tous les autres éléments qui permettent d'identifier l'original ; une liste de descripteurs ou mots clé, tirés du texte et permettant d'en préciser la portée et de s'assurer qu'il concerne la recherche menée ; et enfin, le résumé, à la fois concis et complet, du document.

Le résultat final d'une recherche parmi ces fichiers fournira l'ensemble des références

L'exploitation des banques de données est particulièrement utile dans le domaine de la recherche scientifique. Ainsi, pour les travaux menés à l'Observatoire de Greenbank, en Virginie (Etats-Unis).



Marka

bibliographiques que le documentaliste aura pu sélectionner en combinant un certain nombre d'éléments qui définissent de façon univoque le sujet choisi.

Ce service s'avère très utile aux personnes éloignées de la source d'information principale, en leur épargnant la difficulté de choisir les publications nécessaires à leur travail. C'est notamment le cas des gens qui ne disposent pas de bibliothèque à proximité de chez eux, ou qui ont besoin, pour une utilisation particulière, d'une revue qu'ils ne parviennent pas à trouver parmi celles qui sont à leur disposition.

Avec les bases de données, ils disposent ainsi d'un outil de sélection rapide des documents en rapport avec leur sujet. Pour préparer l'interrogation d'une base de données, on sélectionne les fichiers spécialisés qui semblent correspondre le mieux au thème de la recherche, et également certains fichiers pluridisciplinaires susceptibles d'être intéressants.

La **stratégie de la consultation** découle de la nature des fichiers interrogés. Si l'on craint de recueillir un trop grand nombre de références par une approche trop générale, on peut restreindre le champ de la recherche en définissant des conditions plus précises (date de mise à jour, type de source, pays d'origine, auteur, etc.). Les instructions de recherche contiennent la suite logiquement ordonnée des mots que le documentaliste s'attend à trouver dans le titre ou dans le texte et qui spécifient les caractéristiques du sujet recherché.

Voici un bref exemple de stratégie de recherche avec ses instructions d'extraction et l'impression de l'une des références bibliographiques trouvées.

Le but de la recherche, menée sur la base de données PAPERCHEM, développée par The Institute of Paper Chemistry et accessible par le réseau américain DARDO, était de trouver le compte rendu d'une conférence intitulée *Energy from Biomass Conference*.

L'abréviation CT (conference title) utilisée dans cet exemple est un opérateur qui désigne un certain type de documents et que PAPERCHEM connaît.

Voici les instructions de consultation :

SELECT CT = ENERGY (1W) CT = BIOMASS,
ce qui signifie : « Sélectionne, parmi les titres

de conférence, le mot ENERGY ; tu dois trouver le mot BIOMASS dans la même zone ». L'opérateur (1W) précise que les deux mots doivent se trouver dans l'ordre indiqué (Biomass doit venir après Energy dans le titre) et qu'ils peuvent être reliés par un troisième mot.

```
?SELECT CT = ENERGY (1W) CT = BIOMASS  
33 83 CT = ENERGY (1W) CT = BIOMASS
```

L'impression de la référence trouvée permet d'obtenir :

1. le titre du compte rendu ;
2. son auteur ;
3. le titre de la conférence (c'est la zone dans laquelle la recherche a été effectuée) ;
4. les pages du document d'origine qui contiennent ce compte rendu ;
5. la date de la conférence ;
6. le nombre total de comptes rendus ;
7. la classification du document ;
8. la langue d'origine ;
9. le résumé ;
10. les descripteurs.

Quant aux **banques de données**, elles présentent une caractéristique fondamentale.

Leur interrogation fournit directement les données finales recherchées, car les fichiers des banques de données ne contiennent pas de publications secondaires (descripteurs d'originaux) mais, essentiellement, des copies de documents originaux consultables à partir d'un terminal.

On accède ainsi par l'écran aux publications des organismes de statistiques nationales ou des centres universitaires qui rassemblent eux-mêmes des données provenant de leurs propres recherches ou d'institutions internationales telles que la FAO ou l'OCDE.

Ces données peuvent servir à des traitements ultérieurs.

Ces fichiers sont destinés à des utilisateurs spécialisés, et associent souvent aux procédures normales d'interrogation (comparables à celles des bases de données bibliographiques), des fonctions parfois très complexes de traitement numérique des données trouvées. On peut ainsi combiner plusieurs séries numériques, insérer, traiter et conserver des données à usage confidentiel, obtenir des représentations graphiques, intercaler des données manquantes, extrapoler à partir de modèles mathématiques fournis par la banque ou mis au point par l'utilisateur.

(D'après A. Beverini. Extrait de BANCA DATI, in CELLULOSA E CARTA n° 4, juillet-août 1983.)

Les différents points exposés mettent en évidence les phases essentielles d'un programme de tri : comparaison des données et déplacements consécutifs au sein du fichier de sortie. Nous reviendrons sur ce sujet lorsque nous aurons étudié les principales instructions du Basic.

Recopier le fichier de données présente deux inconvénients : d'une part, on occupe davantage de place dans la mémoire de masse puisqu'à l'issue du tri on a deux fichiers au lieu d'un et, d'autre part, l'exécution du tri est beaucoup plus lente puisqu'il faut déplacer les données du fichier de sortie pour inscrire presque tous les enregistrements. En utilisant un index (ou fichier d'adresses), non seulement on économise la mémoire de masse, mais surtout on a la possibilité d'exécuter les déplacements en mémoire centrale (l'index ne contient que des pointeurs : il est beaucoup plus petit et l'ensemble d'un fichier index peut donc rester en mémoire centrale), ce qui est beaucoup plus rapide.

Le compactage des données

Pour certaines applications, on est obligé de condenser ou « compacter » les données avant de les inscrire sur les supports de mémoire physique.

Sur les gros systèmes qui traitent d'énormes quantités de données, le fait de gagner quelques caractères par donnée peut entraîner une réduction importante de l'espace occupé. Une meilleure exploitation des mémoires justifie alors une plus grande complexité des programmes (qui doivent compacter les données à l'écriture, puis les étendre à nouveau à la lecture). Sur les micro-ordinateurs, au contraire, ces techniques compliquées sont rarement adaptées, exception faite de quelques cas particuliers que nous étudierons par la suite.

Les principales méthodes de compactage reposent sur l'élimination des zones et des caractères vides et sur la codification de certaines combinaisons de caractères.

Les zones de données vides. Il vaut mieux les éliminer que les remplir de blancs. On obtient ainsi un enregistrement plus court que les autres. En contrepartie, on devra utiliser des programmes plus lourds, capables de traiter des enregistrements de longueur variable.

L'élimination des caractères inutiles. Dans les zones de format ASCII, on peut éliminer les blancs qui se trouvent entre deux caractères ainsi que les zéros dans les zones numériques. Une fois encore, le programme sera plus complexe car il devra être conçu de façon à retrouver combien d'espaces ou de zéros ont été éliminés afin de pouvoir reconstituer les données sous leur forme initiale avant de les présenter en sortie.

Le codage des combinaisons de caractères. Lorsque des combinaisons de caractères reviennent souvent, on peut les représenter par un symbole. Ainsi, dans le fichier du personnel, nous trouvons les noms Rousseau, Lenoir et Leblanc.

En associant, par exemple, « ousseau » à 1 et « blanc » à 2, les données deviennent : R1, Lenoir et Le2.

Cette technique n'est intéressante que si les séquences ainsi codifiées reviennent fréquemment.

Les banques de données des gros systèmes

Les gros ordinateurs et certains mini-ordinateurs possèdent des programmes de gestion de données très complexes.

Grâce à ces programmes l'utilisateur n'a plus à s'occuper de la structure physique des fichiers mais simplement à définir leur structure logique.

Les programmes-système se chargeront ensuite de la traduire en schémas physiques et de gérer fichiers, zones et enregistrements. Le programmeur ne connaît plus l'adresse physique des données puisque l'espace mémoire disponible (disques, tambours) est géré par le système. Les schémas de la page 284 illustrent les différences entre ces deux méthodes.

Dans le premier cas (pas de programme de gestion des données), le programmeur doit définir la structure logique de son fichier (ici, le nom, le prénom et l'adresse), la structure physique des enregistrements et leur position dans le fichier.

Dans le second cas, il n'a que la structure logique à définir. C'est le système qui décide où et comment écrire les données, à l'aide des programmes de gestion.

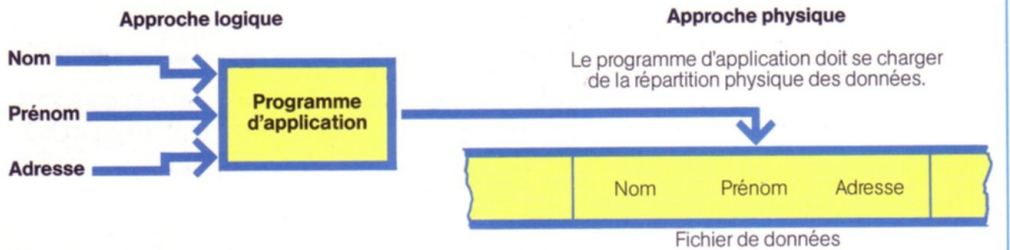
Sur les systèmes les plus évolués, la gestion des données se fait parfois au moyen d'un

langage spécialement étudié, différent des langages d'application.

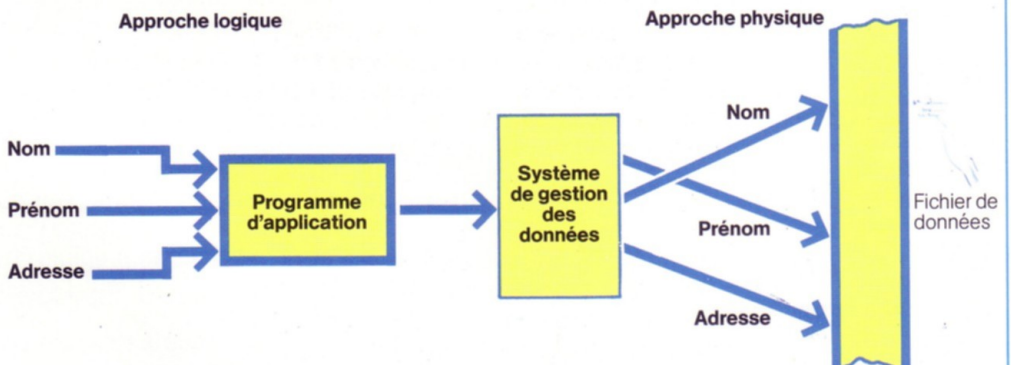
Dans ce cas, le langage de gestion des données est lié à un langage normal de programmation (Basic, Cobol, etc.) et l'utilisateur peut se servir indifféremment des instructions des deux langages.

L'importance de cette question a été à l'origine de la création à Washington, en 1959, du CODASYL (Conference on Data System Language), organisme chargé d'unifier les recherches effectuées dans le monde entier et de publier des normes à observer dans le développement des systèmes de bases de données (structures logiques et relations entre les données).

GESTION DES DONNEES ET PROGRAMMES D'APPLICATION

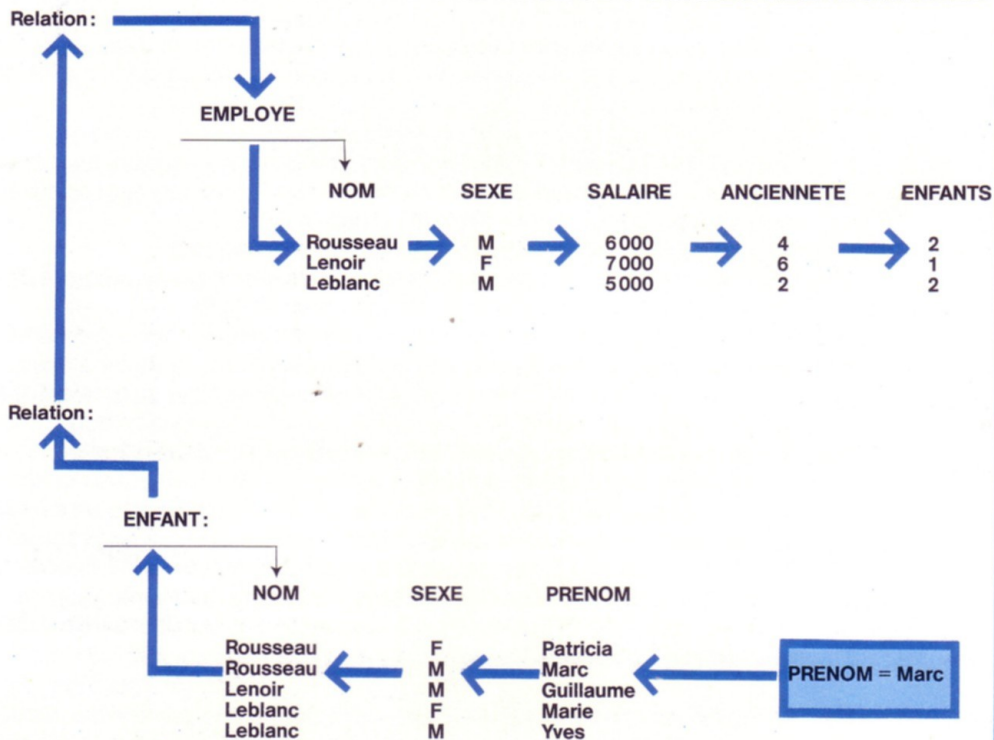


En l'absence de système de gestion des données évolué, les programmes d'application doivent tout prendre en charge, y compris l'adressage des données dans les mémoires de masse.



Sur les gros systèmes, des programmes de gestion des données simplifient beaucoup le travail de programmation. Le système prenant totalement en charge leur gestion physique, l'utilisateur n'a plus à définir que la structure logique de ses données.

EXEMPLE DE RECHERCHE DANS UNE BASE DE DONNEES RELATIONNELLE



Le « domaine » PRENOM (PRENOM = Marc), à travers les relations ENFANT et EMPLOYE, permet de sélectionner toutes les données de l'employé Martin.

Les bases de données relationnelles

Elles permettent l'indépendance totale des structures logique et physique des fichiers. Dans une base de données relationnelle, il n'est défini qu'une structure logique et toutes les données sont reliées entre elles par des « relations ». On peut, en partant de n'importe quelle donnée ou « domaine » (qui constitue la clé de la relation), accéder à toutes les autres données associées, au moyen des relations qui existent entre elles.

Le schéma ci-dessus illustre l'exemple d'un fichier du personnel où coexistent deux relations : Employé et Enfant. A partir d'une valeur (Marc) du domaine PRENOM, on obtient toutes les informations sur l'employé.

Cette base de données, qui est la forme la plus évoluée de gestion des données, n'est proposée que sur de gros ordinateurs.

Solutions du test 7

1 / c) les enregistrements ne peuvent être ni lus, ni écrits sélectivement dans un fichier séquentiel. Les données sont enregistrées séquentiellement, c'est-à-dire dans l'ordre de leur introduction.

2 / b et c) l'index est une zone du fichier (ou un fichier à part) où sont enregistrés les clés et les pointeurs vers les enregistrements de données. Il peut, en général, y avoir plusieurs index, correspondant chacun à une clé d'accès différente.

Dans un fichier du personnel qui contient les informations suivantes :

– prénom et nom – qualification, catégorie, service – salaire de base, taux horaire des heures supplémentaires – jours de congé, d'absence, de maladie –

il peut être utile de disposer de plusieurs index : un premier pour les noms (qui constituent la clé principale), d'autres pour les données secondaires, de façon à pouvoir accéder aux enregistrements selon un autre critère que le nom. Ainsi, on peut, grâce à l'index des jours de congé, obtenir la visualisation de tous les employés qui devront bénéficier d'une certaine période de vacances, ou encore se servir de l'index des services pour établir la liste du personnel d'un même service. Sur les systèmes les plus évolués, on peut procéder à une recherche « croisée », qui combine les zones clé à l'aide d'opérateurs logiques. On pourra ainsi, par exemple, sélectionner les employés du service d'entretien qui ont droit à 20 jours de congé et n'appartiennent pas à la catégorie A. D'un point de vue logique, cette instruction correspond à la formule suivante : (service = entretien) ET (congrés = 20) ET (catégorie ≠ A). Le symbole ≠ signifie « différent de » (en Basic, on écrira « catégorie < > A »).

3 / a) c'est ainsi qu'on appelle l'opération de classement. Rappelons que le tri d'un fichier consiste à extraire les données dans l'ordre, selon une zone définie comme clé. Dans l'exemple précédent (2/), l'existence de plusieurs index (nom, congés, service) permet d'obtenir des listes du personnel ordonnées différemment (autant qu'il y a de zones clé). On peut également combiner plusieurs conditions au cours d'un même tri. On pourra ainsi demander une liste du personnel par service, où les noms des employés apparaîtront par ordre alphabétique à l'intérieur de chaque service.

4 / La recherche séquentielle consiste à lire tous les enregistrements d'un fichier jusqu'à la découverte de celui que l'on cherche. La recherche dichotomique ne lit que les enregistrements obtenus par divisions binaires successives, ce qui va beaucoup plus vite. On ne peut l'utiliser que sur des fichiers ordonnés, ce qui nécessite souvent un tri préalable. Lorsqu'un fichier contient peu de données, le tri et la recherche prennent autant de temps que la recherche séquentielle.

5 / Cette solution n'a qu'une valeur indicative puisque le résultat dépend des longueurs des zones utilisées. Une moyenne raisonnable donnerait :

prénom et nom	50 caractères
qualification	2 caractères (code de deux lettres)
catégorie	2 caractères (code)
ancienneté	2 caractères (de 0 à 99 ans)
salaire	5 caractères (jusqu'à 99999 francs)
jours d'absence	2 caractères (jusqu'à 99)
TOTAL	63 caractères

On utiliserait donc ainsi 63 caractères. En adoptant une longueur de 80 caractères, on disposera d'une marge de 17 caractères pour d'éventuels ajouts.

Pour avoir un ordre de grandeur de la capacité de stockage d'une disquette, calculons le nombre d'employés dont on pourrait stocker les données sur des disquettes de 250 Ko et d'1 Mo (1 Ko = 1000 octets; 1 Mo = 1 méga-octet = 1000 Ko).

La disquette de 250 Ko contient 250 000 caractères, ce qui correspond à :

$$\frac{250\,000}{80} = 3\,125 \text{ données de 80 caractères.}$$

La disquette d'1 Mo contient 1 000 000 de caractères, ce qui correspond à :

$$\frac{1\,000\,000}{80} = 12\,500 \text{ données.}$$

Une seule disquette peut donc contenir le fichier du personnel d'une entreprise de 12 000 employés.

Dans notre exemple, des données numériques comme le salaire ont été considérées comme des chaînes de caractères. En effet, les systèmes d'exploitation les plus répandus en Basic représentent en mémoire de masse les valeurs numériques sous la forme de caractères. Certaines machines utilisent cependant le format numérique où les nombres sont traités comme tels. On économise alors de la place car le format numérique binaire en occupe moins que le format caractère (ASCII). Ainsi, en binaire le nombre 60 n'occupe qu'un octet (60 en base 10 = 00111100 en base 2), alors qu'en code ASCII, il lui faudrait deux octets, un pour chaque caractère (6 et 0, représentés respectivement par 0110110 et 0110000).

6 / Cet exercice est semblable au précédent. Voici un exemple de solution :

article	20 caractères (désignation)
quantité	4 caractères (jusqu'à 9999)
cote d'alerte	4 caractères
date	8 caractères (jour = 2 caractères, mois = 2, année = 4; ex. : 21 07 1983)
fournisseur	20 caractères
TOTAL	56 caractères, ce qui donne 70 si l'on prévoit une marge raisonnable.

7 / Il y a deux façons très simples de gagner de la place dans le fichier précédent. On peut déjà coder la date par les deux derniers chiffres de l'année : 83 au lieu de 1983. On peut également désigner les fournisseurs par des numéros : Rousseau = 0, Lenoir = 1, etc. On établit ainsi une codification des 100 fournisseurs à l'aide des nombres compris entre 0 et 99. Ces deux modifications permettent d'économiser 2 caractères pour la date et 18 pour le fournisseur. Les données n'occupent plus que 36 caractères et l'on pourra donc adopter une longueur de 50 caractères au total pour conserver une marge. On pourra ainsi faire tenir 5 000 enregistrements de ce fichier sur les petites disquettes et 20 000 sur les plus grandes.

GLOSSAIRE

Les mots clés de la gestion des fichiers

Accès. Opération de lecture ou d'écriture des données sur les mémoires de masse (disques durs ou souples, etc.).

Adressage indirect. Il permet de repérer une donnée à l'aide d'un fichier index contenant des pointeurs. La recherche se fait en deux temps: le premier pointeur désigne un enregistrement de l'index qui, lui-même, contient l'adresse de la donnée.

Adresse. Valeur numérique indiquant la position d'un enregistrement.

Algorithme. Méthode de calcul utilisée pour la résolution d'un problème donné.

Banque de données. Ensemble de données relatives à un certain nombre de sujets et organisées en vue de leur utilisation par des programmes correspondant à des applications distinctes. Elles présentent une caractéristique fondamentale: les informations que l'on obtient quand on les interroge d'un terminal correspondent aux données finales recherchées. En effet, ces banques ne contiennent pas de publications secondaires mais des copies de documents originaux.

Base de données. Ensemble de fichiers organisé selon un même schéma logique. On distingue bases et banques de données, même si les deux termes sont parfois confondus. Les bases de données désignent, notamment, les fichiers où sont enregistrées les publications secondaires. On peut les interroger à partir d'un terminal. Ce sont des résumés normalisés et concis de revues, livres, monographies ou rapports de conférences, destinés à fournir à des usagers très divers un outil de sélection efficace de documents en rapport avec leur recherche.

Buffer. Voir **Zone tampon**.

Chainage. Les enregistrements sont dits chaînés lorsque chacun d'eux contient un pointeur qui désigne l'enregistrement suivant. Cette technique, utilisée notamment dans les banques de données, permet de former une liste dans laquelle chaque enregistrement est chaîné à son successeur, quelle que soit sa position physique en mémoire. Le pointeur qui sert au chaînage constitue alors une zone supplémentaire de chaque enregistrement.

Clé. Zone d'un enregistrement permettant de le rechercher.

Compactage (ou condensation). Technique qui permet d'écrire une donnée sur un nombre plus réduit de caractères.

DAM (Direct Access Method). Méthode d'accès direct.

Data base. Voir **Base de données**.

Dictionnaire. Liste des noms ou étiquettes de données prévus dans le cadre d'un fichier ou d'un programme.

Fichier. Espace physique et logique destiné à rassem-

bler un ensemble de données homogènes.

Fusion. Ce terme désigne l'opération (et, par extension, le programme correspondant) qui consiste à réunir les données contenues dans des fichiers différents.

ISAM (Indexed Sequential Access Method). Méthode d'accès séquentiel indexé.

Liste. Voir **Chainage**.

Mémoire virtuelle. Des blocs entiers du programme ne pouvant tenir simultanément en mémoire centrale sont transférés vers la mémoire de masse, puis ramenés en mémoire centrale lorsqu'ils sont nécessaires à l'exécution. Ces opérations, effectuées par le système, sont extrêmement rapides et permettent d'augmenter artificiellement la taille de la mémoire centrale.

Méthodes d'accès. Méthodes de lecture et d'écriture des données. Les principales sont l'accès aléatoire (random access), l'accès séquentiel (SAM) et l'accès séquentiel indexé (ISAM).

Pointeur. Valeur numérique indiquant la position d'un enregistrement ou d'une donnée.

Random. Voir **Méthodes d'accès**.

Recherche dichotomique. Méthode de recherche des données fondée sur des divisions successives en deux parties.

Relationnelle (Base de données). Forme particulière de structuration logique des données. Les zones (ou domaines) sont liées logiquement par des relations.

Répertoire. Index des données ou des fichiers existants et de leurs caractéristiques physiques (longueur et emplacement dans la mémoire de masse).

SAM (Sequential Access Method). Méthode d'accès séquentiel.

Secteur. Division physique d'un disque. C'est la plus petite unité adressable par le système d'exploitation, lors des opérations de lecture et d'écriture.

Temps d'accès. Temps nécessaire à l'acquisition d'une donnée. Il comprend le temps de positionnement de la tête de lecture et le temps de transfert de la donnée de la mémoire de masse à l'unité centrale.

Tri. Classement d'un fichier en fonction d'une clé.

VTOC (Volume Table Of Content) ou Directory list. Liste du contenu d'un disque.

Zone. Donnée élémentaire. Un enregistrement de fichier stock comprend, par exemple, une zone de description de l'article, une zone pour le prix, etc.

Zone tampon. Zone de mémoire utilisée pour le stockage momentané des données. Sur les micro-ordinateurs, les données transitent toujours par une zone tampon au cours des échanges avec la mémoire de masse. Ainsi, lorsqu'on écrit sur une disquette, la donnée passe d'abord de la mémoire centrale à une zone tampon, avant d'être transférée sur la disquette.



Si vous avez entre 15 et 18 ans*, participez au GRAND CONCOURS VOCATION BULL et partez à la découverte de la Californie.

Bull



* Ne peuvent participer au concours Vocation Bull que les personnes nées entre le 25 janvier 1965 et le 25 janvier 1969 inclus.

Répondez aux 20 questions ci-dessous et à la question subsidiaire qui départagera les ex aequo. Reportez vos réponses sur la carte-réponse CONCOURS VOCATION BULL que vous devez poster avant le 14 avril minuit. Bonne chance!

1^o QUESTION

Qu'est-ce qu'un tableur ?

- a un programme de calcul des tables de logarithmes
- b un appareil de lecture automatique de dessins
- c un logiciel de gestion de tableaux.

2^o QUESTION

Qu'est-ce qu'un bit ?

- a une prise d'entrée sur un ordinateur
- b une unité élémentaire de calcul (0 ou 1)
- c un signal sonore émis par le clavier d'un terminal.

3^o QUESTION

La Télématique est-elle :

- a l'association de l'informatique et des Télécommunications
- b l'enseignement des mathématiques assisté par ordinateur
- c une émission de télévision traitant d'informatique.

4^o QUESTION

Un organigramme informatique est :

- a l'assemblage de circuits intégrés
- b une boîte aux lettres électronique
- c le découpage logique d'un programme.

5^o QUESTION

Qu'est-ce que TRANSPAC ?

- a un langage permettant d'interroger l'ordinateur à l'aide du vocabulaire courant
- b un jeu vidéo
- c un réseau national de transmission d'informations entre ordinateurs.

6^o QUESTION

Qu'est-ce qu'une puce ?

- a le plus petit élément d'information
- b le composant électronique monolithique le plus dense
- c une instruction spéciale permettant de sauter d'un programme à l'autre.

7^o QUESTION

Les « ordinateurs de la cinquième génération ». Ce concept a été lancé par :

- a un projet de la Silicon Valley
- b un projet japonais
- c un projet européen.

8^o QUESTION

Le projet « ESPRIT », est-ce ?

- a un projet de recherche européen commun ICL, SIEMENS, BULL
- b une étude portant sur un ensemble de programmes d'intelligence artificielle
- c la tentative de réduire au maximum le matériel au profit du logiciel et des services.

9^o QUESTION

Qu'est-ce qu'une microseconde ?

- a un millionième de seconde (10^{-6} seconde)
- b un milliardième de seconde (10^{-9} seconde)
- c un millième de seconde (10^{-3} seconde).

10^o QUESTION

La capacité actuelle des disques souples se compte en :

- a milliers de caractères
- b millions de caractères
- c milliards de caractères.

11^o QUESTION

Le Logo est :

- a un langage de programmation
- b un programme scientifique
- c une base de calcul informatisée des logarithmes.

12^o QUESTION

Le temps partagé, c'est :

- a l'utilisation d'un ordinateur à distance par l'intermédiaire de moyens de transmission automatique sur ligne téléphonique ou télégraphique
- b l'utilisation simultanée d'un même ordinateur par plusieurs utilisateurs par allocation séquentielle de tranches de temps
- c une procédure d'utilisation d'un ordinateur consistant à utiliser les temps morts au cours du déroulement d'un programme pour exécuter un autre programme.

13^o QUESTION

C.A.O. signifie :

- a Conseil en Automatismes et Organisation
- b Conception Assistée par Ordinateur
- c Création Authentique et Originale de logiciels.

14^o QUESTION

Qu'est-ce qu'un progiciel ?

- a un programme complet et documenté conçu pour être fourni à de nombreux utilisateurs, en vue d'une même application ou d'une même fonction
- b un système permettant l'exécution imbriquée de plusieurs programmes menés de front
- c un logiciel mémorisé dans le processeur frontal.

15^o QUESTION

Qu'est-ce qu'un composant discret ?

- a un composant électronique élémentaire
- b une cassette pouvant programmer un code secret
- c une touche Reset qui permet d'effacer la mémoire.

16^o QUESTION

E.A.O. désigne :

- a une Entrée Automatique Obligatoire
- b Une Extension Aléatoire Optimisée
- c l'Enseignement Assisté par Ordinateur.

Bull aide les jeunes et leur offre dans le cadre de la « **vocation Bull** » 4 voyages aux USA. Ces voyages réservés aux 15-18 ans, se dérouleront début août pendant 10 jours. Dix jours d'un programme prodigieux à travers toute la Californie. Pour les passionnés d'informatique, un grand moment : **Silicon Valley**, la plus grande concentration d'entreprises d'électronique et d'informatique.

Disneyland, monde enchanté où rêve et réalité se confondent : tous les sympathiques personnages de Walt Disney vous y feront un accueil chaleureux. Un pas dans la légende, et ce sera **Hollywood**, le cinéma grandeur nature, des décors fantastiques, des noms prestigieux. Vous vivrez en direct des moments exaltants : la « **vocation Bull** » profite en effet de l'occasion pour vous inviter à assister aux **Jeux Olympiques** de Los Angeles. Un petit détour, et vous plongerez en pleine féerie aquatique avec les célèbres dauphins du parc d'attraction de **Marineland**. Lors de ce périple, la « **vocation Bull** » organisera aussi des rencontres avec de jeunes Américains, tout comme vous passionnés d'informatique.

17^o QUESTION

La mémoire volatile désigne :

- a une mémoire dont le contenu est altéré par une coupure de courant
- b une mémoire dont le contenu n'est pas altéré par une coupure de courant
- c ni l'une ni l'autre.

18^o QUESTION

Le courrier électronique est :

- a une lettre rédigée sur des systèmes de traitement de texte
- b un moyen de faire parvenir le contenu d'un document par télétransmission
- c une technique de dactylographie assistée par ordinateur.

19^o QUESTION

Le traitement en temps réel est :

- a le traitement des données dans lequel plusieurs utilisateurs effectuent simultanément des travaux indépendants
- b le traitement des données dans lequel le traitement est assuré au fur et à mesure des entrées fournies à la machine
- c le traitement des données pendant lequel l'ordinateur produit sans interruption un travail utile.

20^o QUESTION

Qu'est-ce que la carte CP8 ?

- a une carte perforée à 8 trous
- b la carte à microprocesseur de Bull
- c une carte à mémoire magnétique.

QUESTION SUBSIDIAIRE

Rédigez un court slogan publicitaire de vingt mots maximum, grammaticalement correct, sur l'informatique et ses applications en utilisant notamment les lettres du mot **VOCATION**, chaque lettre valant attribution de points – soit :
 $V = 4$ pts – $O = 3$ pts – $C = 2$ pts – $A = 1$ pt – $T = 3$ pts – $I = 2$ pts – $N = 2$ pts.

Attention : n'utilisez pas le mot **VOCATION** lui-même dans le slogan. Les autres lettres sont autorisées mais ne valent pas attribution de points.

Tous les mots employés doivent se trouver dans l'édition 1983 du **PETIT ROBERT**, à l'exclusion des symboles chimiques, des abréviations et des onomatopées.



Avec BASIC PLUS, gagnez votre ordinateur

GRAND CONCOURS THOMSON

MICRO-INFORMATIQUE GRAND PUBLIC

Ouvert à tous les lecteurs de BASIC PLUS

Répondez aux 20 questions ci-dessous et à la question subsidiaire qui départagera les ex aequo, complétez la carte-réponse CONCOURS THOMSON MICRO-INFORMATIQUE GRAND PUBLIC qui se trouve à l'intérieur du fascicule et postez-la avant le 14 avril minuit.

1^{re} QUESTION

L'ordinateur est né :

- a de la machine à écrire
- b de la machine à calculer
- c du magnétophone.

2^e QUESTION

BASIC est :

- a une marque d'ordinateurs
- b un langage de programmation informatique
- c l'abréviation de : Base Analogique des Systèmes d'Informatique Comparée.

3^e QUESTION

Le système Vidéotex de transmission des informations est :

- a italien
- b français
- c anglais.

4^e QUESTION

Le mot «software» signifie :

- a disquette
- b fichier
- c logiciel.

5^e QUESTION

Octet signifie :

- a opération centralisée de télétraitement
- b groupe de huit bits
- c octave du clavier de la console.

6^e QUESTION

La puce est :

- a en sillicone
- b en silicium
- c en sillique.

7^e QUESTION

Le CPU est :

- a le centre de programmation unitaire
- b l'unité centrale
- c le contrôle unifié de programmation.

8^e QUESTION

L'horloge est :

- a le programme indiquant l'heure
- b une montre analogique
- c un dispositif fournissant des signaux périodiques sous forme d'impulsions régulières.

9^e QUESTION

Quel est le poids du TO7 ?

- a 3,400 kg
- b 6,200 kg
- c 8,220 kg.

10^e QUESTION

Qu'est-ce qu'un logiciel d'application ?

- a le moniteur résident
- b un programme éducatif, de jeu, de gestion etc. géré par l'unité centrale
- c une représentation binaire de l'image-écran.

11^e QUESTION

Combien y-a-t-il d'octets dans 64 K ?

- a 64 000
- b 512
- c 65 536.

12^e QUESTION

Le mot Ada évoque :

- a Artificial Device Assignment (Langage d'entrée-sortie virtuelle)
- b Air force Defence Algorithm (Norme informatique du Pentagone)
- c prénom de l'assistante de Babbage, premier informaticien.

13^e QUESTION

Combien de pattes a le micro-processeur 6809 dans le TO7 ?

- a 48
- b 16
- c 40.

14^e QUESTION

Le Cobol est :

- a un jeu
- b un programme de gestion
- c un langage.

15^e QUESTION

Un circuit composite est :

- a un ensemble de circuits reliés entre eux pour obtenir une fonction logique
- b un assemblage complexe donnant des images disparates
- c un réseau composé de circuits de familles différentes.

16^e QUESTION

Une Eprom est :

- a une mémoire morte reprogrammable
- b une marque de circuits intégrés
- c une mémoire électronique définitivement programmée.

17^e QUESTION

Le traitement de texte est :

- a la traduction d'un texte en langage machine
- b le programme qui facilite la dactylographie informatisée
- c l'interprétation d'un texte donné par la machine.

18^e QUESTION

Combien y-a-t-il de touches sur le clavier noir du TO7 ?

- a 36
- b 54
- c 58.

19^e QUESTION

Que signifie le mot DATA ?

- a donnée
- b date
- c adresse.

20^e QUESTION

La cryptographie est :

- a la traduction d'informations en langage machine
- b la recherche de l'erreur dans un programme
- c un système de codage qui rend plus difficile le décodage d'un programme.

QUESTION SUBSIDIAIRE

Rédigez un court slogan publicitaire de vingt mots maximum, grammaticalement correct, sur l'intérêt de l'encyclopédie BASIC PLUS en utilisant notamment les lettres du mot INFORMATIQUE, chaque lettre valant attribution de points - soit :

I = 2 pts - N = 2 pts - F = 2 pts - O = 3 pts - R = 3 pts - M = 2 pts - A = 1 pt - T = 2 pts - L = 2 pts - Q = 4 pts - U = 3 pts - E = 1 pt.

Attention : ne pas utiliser le mot INFORMATIQUE lui-même dans le slogan. Les autres lettres sont autorisées mais ne valent pas attribution de points.

Tous les mots employés doivent se trouver dans l'édition 1983 du PETIT ROBERT, à l'exclusion des symboles chimiques, des abréviations et des onomatopées. Le titre de l'encyclopédie BASIC PLUS compte pour un mot.



100 GAGNANTS

**POUR UN TO 7,
2 LOGICIELS (PICTOR ET BASIC)
ET LE MANUEL D'INITIATION AU TO 7**