JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

JACQUES PINEL

Une nouvelle méthode de classification et de recherche

Journal de la société statistique de Paris, tome 93 (1952), p. 181-192
<http://www.numdam.org/item?id=JSFS_1952__93__181_0>

© Société de statistique de Paris, 1952, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (http://publications-sfds.math.cnrs.fr/index.php/J-SFdS) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

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

II

UNE NOUVELLE MÉTHODE DE CLASSIFICATION ET DE RECHERCHE

Je voudrais vous exposer rapidement les principes et les résultats d'une nouvelle méthode de classification et de recherche que j'ai appelée sélection automatique. Bien qu'elle soit plus générale, je ne prendrai comme exemples que des problèmes d'empreintes digitales car c'est pour eux qu'elle a été réalisée.

Le classement des individus par les empreintes digitales repose sur l'établissement de la formule digitale suivant des règles variant d'un pays à l'autre, d'un service à l'autre. Mais, en général, deux individus ayant la même formule dans une méthode donnée ont toutes chances d'avoir la même dans une autre méthode.

A Paris, nous utilisons 5 formes fondamentales (fig. 1) affectées chacune d'un chiffre de 1 à 5. Ce sont l'arc (1), la boucle à gauche (2), la boucle à droite (3),

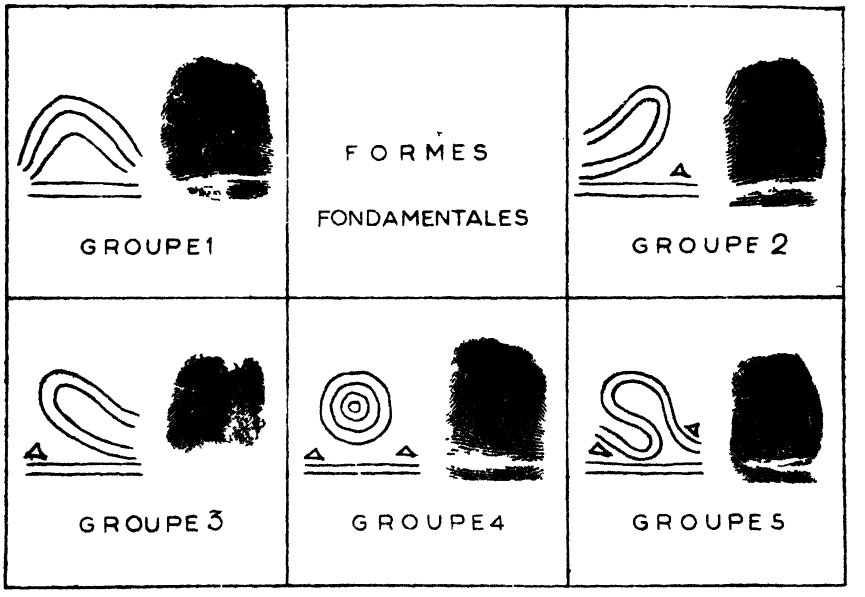

Fig. 1 — Les 5 formes fondamentales d'empreintes digitales.

la volute simple (4) et la volute double (5). La formule digitale est un nombre de 10 chiffres obtenu en écrivant à la suite les indices de chaque doigt pris dans l'ordre où on les obtient en posant les mains à plat sur une table, paumes tournées vers le bas.

Théoriquement, on a près de 10 millions de formules possibles. Pratiquement, certaines sont très chargées, et l'une d'elles atteint 3,6 %. Il s'agit des indi-

vidus ayant des boucles cubitales (ouvertes vers l'auriculaire, ou vers le cubitus) aux dix doigts.

Pour subdiviser ces formules, on utilise le nombre de lacets obtenu en comptant le nombre de crêtes qui coupent une ligne droite tracée sur l'empreinte et joignant deux points appelés « centre de figure » et « delta » (fig. 2). Nous classions d'après le nombre de lacets du pouce droit, puis les groupes importants étaient subdivisés par l'index droit, puis par le médius droit. On n'allait pas plus loin. On obtient ainsi une excellente répartition des fiches dans de petits paquets, mais la recherche est extrêmement pénible, car le nombre de lacets peut varier d'une détermination à l'autre. Bien que l'on ait fait des règles aussi complètes que possible pour définir la position du delta

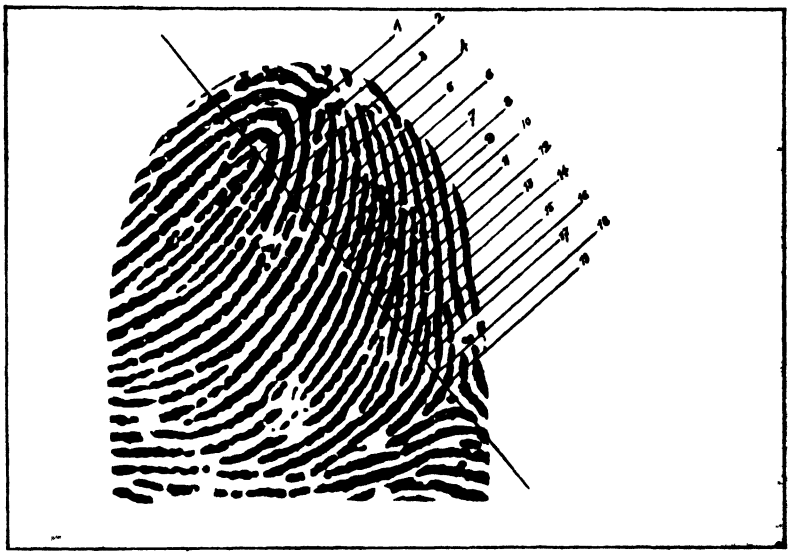

Fig. 2. — Le comptage des lacets.

ou du centre de figure on n'a pas pu tout prévoir. Pour certains doigts on aura des écarts d'une détermination à l'autre. D'autre part, une simple déformation élastique de la peau peut déplacer une ou plusieurs bifurcations par rapport à la ligne de foi et provoquer des écarts. Enfin, il faut tenir compte des empreintes plus ou moins bien relevées, ou abîmées. Un souteneur et un terrassier n'ont pas tout à fait des mains dans le même état. Au total, on estime que, d'une détermination à l'autre du nombre de lacets d'une boucle, on peut avoir un écart de 2 lacets en plus ou en moins. Ceci complique énormément le travail et une recherche dans une grande formule est un travail de longue haleine, bien que cette méthode soit l'une des plus rapides, la sélection automatique mise à part.

Dans toutes les méthodes classiques de classification proposées pour résoudre le problème des grandes formules, on n'utilise qu'un nombre restreint des éléments d'identification et de différenciation.

Le problème est bien plus général que cela. On peut dire qu'il se posera

chaque fois que l'on aura à faire des opérations sur des ensembles d'objets à caractéristiques floues. Voici, bien sommairement les principes de la Sélection Automatique.

PRINCIPE

Quand on pose la question : tel objet possède-t-il telle caractéristique? on a en général deux réponses « oui » ou « non ». Si à ces deux réponses s'ajoute une troisième possibilité « peut-être », et que d'une détermination à l'autre on puisse passer de « oui » à « peut-être » ou de « non » à « peut-être », nous dirons que la caractéristique est floue. Si on peut passer de « oui » à « non », nous dirons que la caractéristique est « instable » et dans le cas contraire qu'elle est « stable ».

Par exemple, la couleur de l'iris de l'œil est une chose stable, mais suivant les cas, on pourra considérer comme possibles plusieurs couleurs différentes, par exemple, si on a choisi comme caractéristiques les couleurs « gris » et « vert », pour certains individus on pourra considérer les deux teintes comme possibles. Par contre la couleur de la cravate peut être nette, mais elle n'est pas stable : on peut en changer.

Nous ne ferons appel qu'aux caractéristiques stables, et nous supposerons épuisés tous les moyens de subdivision par des caractéristiques nettes.

Nous supposerons que, toutes les caractéristiques utilisées étant choisies à l'avance, on sache, pour tous les objets de l'ensemble celles qui sont possibles et celles qui sont impossibles. Supposons maintenant que nous voulions extraire de l'ensemble tous les objets possédant certaines caractéristiques. Ce problème est impossible en toute rigueur, puisqu'il y a des objets pour lesquels on ne sait pas s'ils possèdent ou non l'une ou l'autre des caractéristiques voulues. Nous éliminerons tous ceux qui ne possèdent certainement pas l'une ou l'autre des caractéristiques demandées. Il restera donc tous ceux qui sont possibles, sans que l'on ait pu en oublier un seul par manque d'information.

Il est évident que l'on extrait ainsi d'autres objets que ceux qui correspondent réellement, mais rien ne permet, avec les renseignements dont on dispose, de les éliminer.

RÉALISATION

Il faut disposer d'un système permettant d'enregistrer tous les renseignements dont on dispose, et de les exploiter ensuite commodément. Suivant le cas, on pourra utiliser deux techniques totalement différentes.

La première consistera à utiliser des fiches mécanographiques classiques, une fiche correspondant à chaque objet, et un emplacement de perforation correspondant à chaque caractéristique. On perforera dans chaque carte l'emplacement des caractéristiques possibles. Pour l'extraction, on comparera successivement toutes les cartes à une carte témoin portant des perforations aux endroits des caractéristiques que l'on veut; une machine du type interclasseuse extraira, ou dénombrera, les fiches intéressantes.

Cette méthode sera commode pour réaliser des opérations de statistique, si on doit dénombrer successivement plusieurs sous-ensembles. Des mémoires artificielles à défilement plus rapide pourront même être plus intéressantes.

Si au contraire, on veut extraire un tout petit nombre d'objets dans un grand ensemble, une telle méthode sera peu commode. Pour les problèmes d'identification, l'idéal est de sortir l'objet voulu s'il existe, ou aucun, s'il n'existe pas. La méthode précédente nécessite l'examen *successif* de tous les objets Il sera plus commode d'avoir recours à des méthodes idéologiques permettant un examen *simultané*. Nous avons fait appel à ce procéde, en utilisant les fiches « Sélecto » (fig. 3). Ces fiches comportent 2 000 emplacements de perforations numérotes de 1 à 2 000. Elles sont imprimées et calibrees de telle façon que les perforations correspondant au même numéro soient rigoureusement superposées quand on les empile soigneusement les unes sur les autres. A

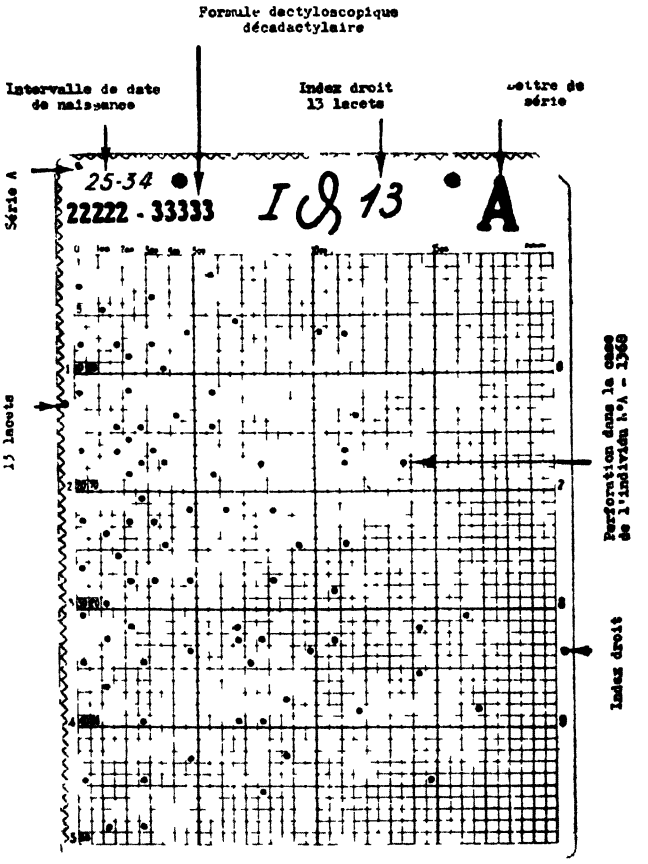

Fig 3 — Reproduction d une fiche " Sélecto
Les perforations marginales servent a identifier la fiche sur les microscopies de sécurité absolues par transparence, et ou le texte disparait

chaque numéro correspond un objet, ou ici un individu pour le cas qui nous intéresse, et réciproquement. Le numéro de classement d'un individu est le même pour toutes les fiches de sélection Chacune d'elles correspond à une caractéristique. Pour établir le fichier, on perfore dans chaque fiche les cases des individus pouvant posséder cette caractéristique, ou plus exactement on ne perfore pas celles de ceux ne la possédant certainement pas. Ainsi l'absence

de perforation dans la case 1250 signifie que tout individu possédant cette caractéristique est différent de celui repertorié sous le n° 1250.

Pour rechercher un individu, on commence par déterminer celles des caractéristiques qu'il possède certainement, puis on empile les fiches « Sélecto » correspondantes. L'alignement des trous sera bouché pour tous ceux pour lesquels il existe une incompatibilité, en sorte que par examen direct à la lumière, on verra le jour à travers les cases de ceux qui peuvent convenir. Il ne reste plus qu'à lire le numéro et à extraire la fiche dactyloscopique correspondante classée d'après ce numéro.

Nous avons appliqué cette méthode à un certain nombre de grandes formules, pour lesquelles on ne pouvait plus utiliser de caractéristiques nettes pour faire des subdivisions. Les éléments utilisés dépendent de la forme fondamentale du doigt.

APPLICATION

Dans le cas des boucles franches, nous ne disposions que du nombre de lacets. Pour chaque doigt, nous avons fait une fiche pour chaque valeur possible du nombre de lacets, soit de 2 à 22 pour tous les doigts, sauf pour les pouces, où nous allons jusque 32. On considère que l'empreinte est bonne si on estime que lors d'une autre détermination du nombre de lacets de ce doigt, on ne peut pas trouver un nombre différent de plus de 2 de la valeur que l'on vient de trouver. En conséquence, quand la détermination est bonne, on perforera pour chaque doigt 5 fiches de sélection correspondant à la valeur trouvée et aux 2 valeurs immédiatement voisines de part et d'autre. Si pour quelque raison que ce soit la détermination n'est pas bonne, on perforera les fiches correspondant aux valeurs que l'on pourrait trouver lors d'une bonne détermination du nombre de lacets de ce même doigt, compte tenu de l'écart qui pourrait alors jouer en sens inverse.

On est sûr ainsi que les valeurs trouvées pour de bonnes empreintes ont été considérées comme possibles lors de déterminations antérieures. Pour la recherche, on empilera les fiches correspondant aux valeurs trouvées pour les seuls bons doigts, en ne tenant pas compte des autres, soit 10 fiches au maximum.

Dans le cas *des arcs tentiformes*, on ne peut pas compter le nombre de lacets. Nous avons eu recours à la morphologie du centre de figure, et nous avons défini 5 formes (fig. 4). Il n'est pas possible de les utiliser directement, car trop de doigts sont « limites » entre deux de ces formes, ce qui rendrait la recherche très complexe. Le cas des volutes est à l'étude actuellement.

Le mode opératoire est le suivant. On établit d'abord le jeu complet de fiches de sélection. On fait, pour une formule, autant de séries qu'il y a de fois 2.000 individus. Pour chaque série, on fait toutes les fiches nécessaires, chaque doigt étant traité séparément. Chaque série est affectée d'une ou plusieurs lettres qui sont reportées sur la fiche de sélection et sur les fiches dactyloscopiques de la série. Les fiches dactyloscopiques sont notées complètement puis entièrement vérifiées. On effectue ensuite la recherche complète dans les séries terminées relatives à la même formule, puis dans la série en cours de constitution, puis quand on est assuré qu'il n'y a pas déjà de fiche pour le même individu, on

effectue les perforations. Pour cela, on empile le paquet de fiches à perforer

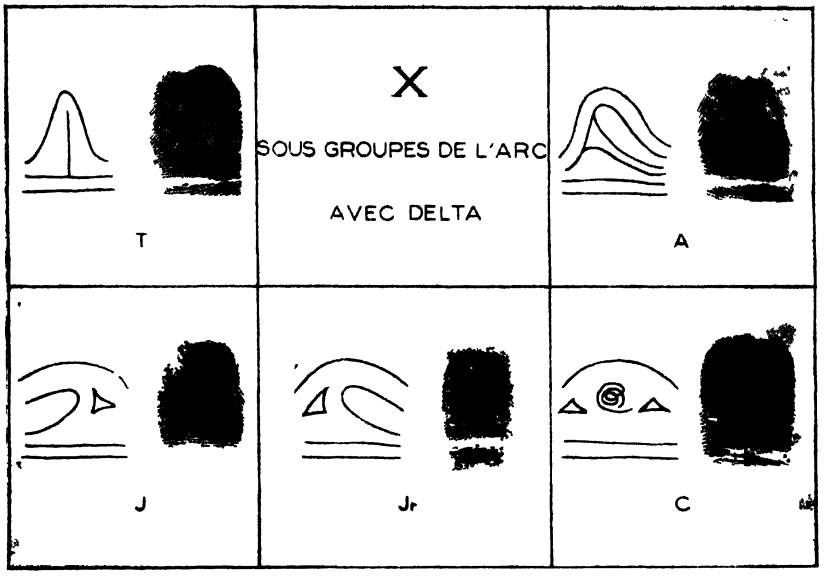

Fig 4 — Les sous groupes de l'arc avec delta

sur le plateau d'une perceuse aménagée spécialement (fig. 5), toutes les perforations étant effectuées en une seule fois ce qui assure un parfait centrage

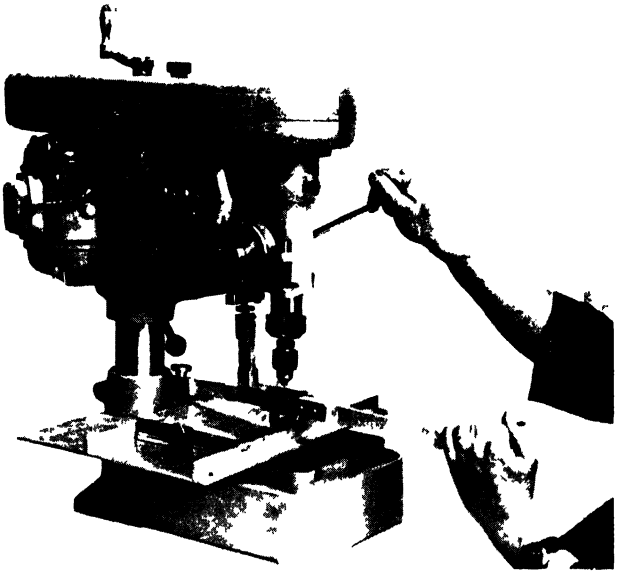

Fig 5. — Perceuse aménagée spécialement pour le perçage des fiches " Sélecto "

des trous en réduisant au minimum les manipulations de fiches. A ce moment on attribue à la fiche dactyloscopique un numéro, à la suite de ceux existant déjà dans la série, et qui correspond à l'emplacement des perforations. Pour

éviter les erreurs, le travail de perforation est effectué par deux opérateurs se contrôlant mutuellement.

COMPARAISON DES MÉTHODES

Nous étudierons d'abord la vitesse de la recherche, nous verrons ultérieurement son intérêt vis-à-vis de la sécurité. Il est difficile de comparer des méthodes sans définir des qualités chiffrables. C'est pourquoi, nous avons défini l'efficacité et la complexité.

Nous définirons l'*efficacité* comme le rapport entre le nombre de fiches d'un groupe et le nombre moyen de comparaisons. Ainsi avec une efficacité de 4, on devra examiner à chaque fois, en moyenne le quart du fichier.

La *complexité* sera le nombre moyen de paquets de fiches que l'on devra examiner pour faire une recherche. Ainsi dans la méthode décrite précédemment, dans une subdivision par un seul doigt, on a en principe 5 sous-groupes à voir pour une recherche. Seulement en pratique dans ce cas, ils sont réunis les uns derrière les autres, ce qui fait que l'on n'a qu'un seul paquet à voir. De même, si on utilise une subdivision de ce type à deux échelons, on aura une complexité de 5 et de 25 dans une subdivision à 3 échelons. On comprend pourquoi on ne pousse pas la subdivision au 4e doigt, d'autant plus qu'une première séparation étant faite par la date de naissance, par groupes de 10 ans, on a en fait 2 ou 3 fois plus de paquets à voir pour tenir compte d'une fraude possible sur l'âge.

Nous avons étudié diverses méthodes dérivant du même principe que celles que nous avons indiquées, mais utilisant d'autres doigts. Nous avons également étudié diverses méthodes en usage ou proposées par divers auteurs. Dans la méthode dérivant de la méthode Galton Henry, on fait pour chaque doigt 2 sous-groupes notés I (interne) ou 0 (externe ou outer). La coupure se situe entre 9 et 10 pour les index, 10 et 11 pour les médius, et 13 et 14 pour les annulaires. On utilise 2, 4 ou 6 doigts, en prenant soit les index seuls, soit les index et médius, soit enfin les index, médius et annulaires.

Nous avons enfin fait les calculs analogues pour une méthode proposée par un Espagnol Giméno Ros, dans laquelle on utilise les sommes des nombres de lacets des trois premiers doigts de chaque main. Les recherches se font avec une marge de sécurité de 3 en plus ou en moins sur cette somme.

La conduite du calcul est la suivante. On détermine le nombre d'individus (N_i) dans chaque sous-groupe en appliquant les règles propres à chaque méthode. Le travail porte sur un nombre important d'individus appartenant à la formule sur laquelle on travaille. Ensuite, on recherche sur le même groupe combien (M_i) de ces individus devraient être recherchés dans chacun des sous-groupes si on devait les rechercher dans cet ensemble, en appliquant les règles de sécurité indispensables. Il ne reste plus qu'à faire les sommes des produits $N_i M_i$. Ce nombre est celui des comparaisons que l'on devrait faire pour rechercher dans cet ensemble tous les individus d'un second groupe de même capacité, en supposant, ce qui est loisible, que les répartitions dans les différents sous-groupes sont les mêmes pour les deux ensembles.

Nous avons obtenu au cours de ces calculs quelques résultats intéressants,

Tableau 1

Principaux résultats concernant les subdivisions à un seul échelon de lacet en lacet pour les dix doigts et la méthode de Giméno Ros.

		K	S	N_i max			M_i max		
				i	N_i	%	i	M_i	%
Auriculaire	G . .	2,923	18 156 705	12	712	9,78	12	3 256	44,6
Annulaire	G . . .	2,901	18 326 365	15	766	10,50	14	3 532	48,5
Médius	G . . .	2,626	20 242 281	15	841	11,52	13	3 838	52,6
Index	G . .	2,647	20 080 908	10	808	11,09	11	3 650	50,1
Pouce	G .	3,199	16 618 277	16	687	9,42	16	3 163	43,5
Pouce	D .	3,318	16 022 998	20	672	9,22	19	3 104	42,6
Index	D .	3,044	17 463 707	13	697	9,57	13	3 153	43,2
Médius	D .	2,666	19 940 055	12	829	11,34	13	3 729	51,2
Annulaire	D .	3,238	16 414 232	15	691	9,49	16	3 174	43,6
Auriculaire	D .	3,116	17 059 059	12	609	8,36	12	2 886	39,6
Somme P. I. M	G .	4,584	11 571 247	40	347	4,76	39	2 229	30,5
Somme P. I. M. D	.	4,876	10 922 443	44	315	4,32	41	2 049	28,2

K : efficacité

$S = \Sigma_i (N_i \times M_i)$

Tableau 2

Principaux résultats concernant les subdivisions à deux échelons de lacet en lacet et la méthode de Giméno Ros.

		K	S	N_{ij} max				M_{ij} max			
				i	j	N_{ij}	%	i	j	M_{ij}	%
Pouce	D/G .	7,901	6.727 780	16	19	92	1,26	16	19	1 759	24,2
Index	D/G .	6,518	8 156 433	13	14	114	1,58	12	13	1 991	27,2
Médius	D/G .	5,654	9 400 781	14	13	137	1,88	13	13	2.419	33,2
Annulaire	D/G .	6,858	7 751 388	15	16	120	1,65	15	16	2 048	28,2
Auriculaire	D/G	6,533	8 138 829	12	12	110	1,51	12	12	1.811	24,9
Somme P. I. M.	D/G	15,037	3 533.602	39	42	33	0,45	39	44	921	12,6
				41	44						
Pou. D/I	D .	9,351	5.683.178	18	13	88	1,21	19	13	1.500	20,6
Pou. D/Med	D	8,35	6.388 392								
Pou. D/Ann	D .	10,3									
Pou. D/Aur.	D .	10,2									
Pou. D/Ann.	G .	8,95									
P d/Id/Md . . .		20,474	2.596,332	19	12	12 17	0,23	19	13	13.942	12,9

Tableau 3

Principaux résultats concernant l'efficacité et la complexité de la première classification sous-secondaire utilisée par le F. B. I.

NOMBRE d'échelons	k	S	M	p
2	1,844	28 830 170	13 113	1,80
4	2,84	18 766 576	25 967	3,57
6	4,71	11 291 456	50 453	6,91

p (complexité) $= \Sigma M_i / \Sigma N_i$

$M = \Sigma_i M_i$

qui sont : le nombre maximum de fiches dans un sous-groupe et le nombre maximum de comparaisons à faire dans les cas les plus défavorables (1) (tableau 1).

On voit que de toutes les subdivisions à un seul échelon c'est celle utilisant le pouce droit qui est la meilleure d'assez loin, ou, suivant la méthode de Giméno Ros, la main droite également.

Parmi les subdivisions à deux échelons certaines ont été étudiées par la méthode précédente, et d'autres par une méthode simplifiée substituant une répartition normale à la répartition expérimentale, ce qui est valable pratiquement. On constate, que de celles qui ont été étudiées directement, c'est la subdivision par le pouce et l'index droits qui est la meilleure (c'est celle qui était en usage à l'Identité Judiciaire avant application de la Sélection Automatique) bien que l'on puisse atteindre un meilleur résultat en utilisant l'annulaire ou l'auriculaire de cette même main comme second élément après le pouce. Aucun autre procédé de ce type ne peut donner une meilleure efficacité. La méthode de Giméno Ros a une meilleure efficacité mais nécessite le comptage de 6 doigts (tableau 2).

La subdivision à 3 échelons de ce type a une bonne efficacité, mais, comme nous l'avons vu, sa complexité est déplorable.

TABLEAU 4

Valeur des nombres $N_{(\equiv\equiv\equiv)}$ et $M_{(\equiv\equiv\equiv)}$ pour la subdivision sous-secondaire à 6 échelons utilisée par le F. B. I.

G \ D	III	II0	I0I	I00	0II	0I0	00I	000
III	455 1.303	103 710	88 816	27 565	143 1 046	43 672	79 900	39 721
II0	81 618	106 672	23 463	53 586	20 544	45 589	16 549	42 704
I0I	100 862	44 597	89 773	48 660	50 847	30 664	108 1.003	76 1.001
I00	36 565	87 733	38 570	172 930	17 569	52 736	59 786	237 1.356
0II	96 513	29 381	38 434	25 388	147 816	40 607	112 903	60 859
0I0	26 391	37 441	10 351	21 469	35 555	63 650	30 681	118 974
00I	55 590	30 490	73 627	51 649	116 1 003	54 976	392 1.550	376 1.823
000	28 511	74 668	42 592	227 1.042	67 768	184 1.214	216 1.434	1 813 3 192

Dans ce tableau, comme dans les deux suivants, chaque case correspond à un sous-groupe. La formule de la main droite est donnée par la colonne, celle de la main gauche par la ligne, les doigts étant pris dans l'ordre; index, médius et annulaire pour chaque main.

Dans chaque case, le nombre supérieur est N_i, le nombre inférieur est M_i.

(1) Le calcul porte sur un ensemble de 7.291 individus ayant des boucles cubitales franches aux dix doigts.

TABLEAU 5

Valeurs des nombres $N_{(\equiv\equiv\equiv)}$ et $M_{(\equiv\equiv\equiv)}$ pour la subdivision secondaire à 4 échelons utilisée par le F. B. I.

G \ D	II	I0	0I	00
II	745 1.766	191 1.203	251 1 427	176 1.360
I0	267 1.359	347 1.459	149 1.337	480 2.021
0I	188 1.167	94 952	285 1.436	320 1.683
00	187 1.196	393 1.554	421 1.860	2.797 4.187

TABLEAU 6

Valeurs des nombres $N_{(\equiv\equiv\equiv)}$ et $M_{(\equiv\equiv\equiv)}$ pour la subdivision sous-secondaire à 2 échelons utilisée par le F. B. I.

G \ D	I	0
I	1 550 2.750	1.056 2.934
0	862 2.341	3.823 5.088

Enfin les méthodes dérivant de la méthode Galton Henry ont des performances qui ne sont pas bonnes du tout. La complexité est bien trop grande pour l'efficacité obtenue (tableau 3).

Ces méthodes (1) de calcul ne sont plus valables pour déterminer l'efficacité de la sélection automatique. Nous l'avons obtenue par sondage. En reclassant les fiches d'un ensemble, nous faisons les recherches et nous notons combien nous devons faire de comparaisons dans chaque cas. L'efficacité se déduit simplement de ces nombres.

Nous l'avons déterminée dans 3 cas (tableau 7). La première colonne correspond au groupe 22222 33 333 (boucles cubitales franches aux 10 doigts), la seconde au groupe 22 2 X 2 3 X 333 (arcs tentiformes aux index et boucles cubitales aux autres doigts) sans utiliser les index, et enfin la dernière colonne à la même formule en utilisant les 10 doigts. On voit que les valeurs de l'efficacité obtenue sont hors de proportions avec celles que nous avions précédemment.

Le groupe des individus ayant des boucles cubitales franches aux 10 doigts, de tous le plus important comporte actuellement près de 16.000 individus tous différents. Les recherches journalières que l'on y effectue confirment les résul-

(1) L'emploi de cette méthode de calcul pour des méthodes ayant une très grande efficacité conduirait à supposer que l'on a toujours une comparaison au moins à effectuer, alors qu'en fait on n'en a souvent aucune quand l'efficacité est du même ordre de grandeur, voire plus grande que le nombre d'individus du groupe.

tats de ce tableau. Le nombre réel de comparaisons est très variable, mais leur nombre moyen est conforme aux prévisions de ce tableau, c'est-à-dire inférieur à 4.

TABLEAU 7

Calcul de l'efficacité de la Sélection automatique.

NOMBRE de comparaisons infructueuses	1	2	3
0	1.721	862	1.010
1	204	126	29
2	41	32	2
3	23	11	0
4	6	0	0
5	4	0	0
6	0	0	0
7	1	0	0
8 et plus	0		
N	2.000	1.041	1.041
S	447	212	33
k	4.460	2.440	16.400

Très souvent, on en a du reste moins à effectuer, mais il arrive souvent parfois que leur nombre devienne relativement important. Il est rare cependant que l'on doive en faire plus d'une vingtaine, et il ne semble pas que ce nombre doive jamais dépasser la 1/200 partie du fichier dans les cas les plus défavorables. Dans le cas des arcs tentiformes, il est vraisemblable que l'ensemble sur lequel a porté la détermination de l'efficacité soit un peu faible, pour le cas où on utilise les 10 doigts. Cependant, le nombre de comparaisons n'est jamais bien grand, quoique ce groupe comporte 8.000 fiches environ. Dans d'autres formules, moins importantes, on a des résultats analogues, bien que nous n'ayons pas fait de dénombrement complet.

Au total, une recherche dans la formule la plus chargée demande largement moins de 10 minutes alors qu'il fallait des heures avec la méthode ancienne.

CAS DES MAUVAISES EMPREINTES

La sélection automatique est particulièrement intéressante. En effet, dans les méthodes classiques, si l'on a une incertitude sur l'un des doigts utilisés pour la subdivision, on doit étudier tous les sous-groupes que l'on peut former avec toutes les valeurs possibles du nombre de lacets de ce doigt, ce qui peut donner des recherches extrêmement longues. Avec la sélection automatique, il n'y a aucune difficulté. Il suffit de se contenter de ne pas utiliser les fiches de sélection des mauvais doigts. Rien n'est plus simple. En fait les statistiques précédentes ont été données en tenant compte des mauvais doigts pour la sélection automatique, tandis que pour faciliter les calculs dans les méthodes classiques, nous avons supposé que les déterminations étaient toutes bonnes, afin de simplifier des calculs déjà fort longs. En fait, dans les méthodes classiques, on fait un sous-groupe intitulé « lacets non comptés » où sont classées ces mauvaises fiches. On devra l'examiner pour toute recherche, même si on

dispose de belles empreintes. On a là un allongement de la recherche qui n'est pas négligeable. Bien entendu, même avec la sélection automatique, il reste des fiches que l'on ne peut pas classer, car elles sont si mauvaises que l'on devrait prendre de telles marges de sécurité que l'on serait obligé de les sortir ensuite très souvent pour des comparaisons. De telles fiches ne sont pas classées et mises de côté. Pratiquement, il y en a extrêmement peu, à peine 10 % de ce que l'on devait mettre de côté avant, bien que nous considérions assez facilement qu'une empreinte nécessite une marge de sécurité anormale. En fait, nous substituons totalement à la notation fixe et nette ancienne de l'empreinte une notation floue tenant compte des variations d'appréciation dues aux divers facteurs qui peuvent intervenir. Cette notation floue correspond certainement plus à la réalité que la notation classique.

SÉCURITÉ

Nous n'avons pas parlé jusque maintenant de cette qualité, la plus importante de toutes, au fond. Théoriquement, la sélection automatique présente une sécurité absolue, puisque pour manquer une recherche, il faut que l'on ait des écarts de notation vraiment anormaux, d'autant plus que l'on peut prendre des sécurités supplémentaires quand une empreinte présente la moindre difficulté. Il en est de même aussi en principe dans toutes les méthodes. Cependant, nous évitons ici bien des erreurs. Le nombre de comparaisons est tellement restreint que, quand on doit en faire une, l'attention est attirée et vraiment on ne risque pas de manquer l'identification, tandis qu'une telle inattention doit se produire fréquemment, par suite de la fatigue, dans les méthodes classiques. Le mode opératoire limite les erreurs matérielles. On ne peut pas comme dans les méthodes classiques négliger par inadvertance de voir un sous-groupe. Si une fiche est sortie et mal reclassée, on s'en apercevra, puisqu'elle manquera à son numéro. Une telle erreur ne se soupçonnerait même pas avec les méthodes classiques. Le temps de recherche étant réduit à rien, on peut donner plus d'attention au contrôle du travail. Nous avons ainsi la possibilité de faire une double notation complète qui ne se faisait pas auparavant, les opérations de recherche et de perforation se font suivant un certain rite qui permet de contrôler le travail, tandis que la recherche par les méthodes classiques consiste en un défilé rapide de fiches devant l'œil du chercheur, travail fastidieux et combien peu sûr. De reste, au cours des opérations de modernisation, nous avons trouvé un nombre considérable de fiches en double ne correspondant pas toujours au même état civil.

Il n'est évidemment pas possible d'espérer avoir une méthode présentant une sécurité absolue mais on peut penser qu'il est humainement difficile de faire mieux, compte tenu de nos connaissances dactyloscopiques actuelles.

Jacques Pinel.