

## RECONNAISSANCE DES FORMES (H, I7, K9)

La notion générale de reconnaissance des formes comporte deux aspects : l'un est relatif à la définition d'une notion de « **forme** », l'autre à celle de reconnaissance.

(i) Par **forme**, on peut entendre tout « objet » ou « **structure** » définis dans un **domaine de connaissance**, et généralement associés à l'étude d'un **phénomène** qui en relève. Ainsi :

- (a) une forme mathématique peut être une structure algébrique, une structure topologique, une structure mesurable ou un objet géométrique (surface analytique ou variété différentielle) ;
- (b) une forme physique peut être une structure physique : cristal, minéral ;
- (c) une forme biologique peut être une structure biologique : virus, bactérie, champignon, fossile, « carte » ADN.

(ii) La **reconnaissance** est un moyen pour déterminer si des formes données (de même science et de même type) sont (strictement) identiques ou non. En pratique, on se contente parfois de formes seulement équivalentes.

Dans le premier cas, les deux formes considérées seront assimilées (ou identifiées) l'une à l'autre. Dans le second cas, on considère que les formes appartiennent à des « classes » distinctes.

(ii) En **Statistique**, la reconnaissance des formes reçoit plusieurs sens spécialisés (cf aussi **classification, forme**).

(a) **analyse (géométrique) des données** : eg **afc, acp** ou **analyse de PROCUSTE**. Dans cette approche, une « **forme** » est un **facteur** (ou un « axe factoriel »), en général « caché », dont la signification peut correspondre à une interprétation concrète par l'homme de l'art (« intelligence » en psychologie, etc). Le **nuage de points**, qui fait l'objet de ce type d'analyses, est alors « situé » par de telles formes ;

(b) **méthode de classification** (**classification automatique** ou **classification** statistique) : cette méthode consiste tantôt à définir des classes dans une **population statistique** ou dans un **échantillon** donnés, tantôt à affecter une unité statistique supplémentaire à l'une des classes pré-constituées d'une population ou d'un échantillon (cf **discrimination**). Dans les deux cas, les classes considérées peuvent s'interpréter comme des « **formes** », puisque leur « homogénéité » implique une certaine « identité » de forme ;

(c) **méthode d'estimation** : la **théorie de l'estimation** statistique elle-même a pour objet le « repérage » ou l'« identification » d'un **espace probabilisé** au sein d'une famille de tels espaces (ie au sein d'un **modèle statistique**). Dans cette approche, une « **forme** » est soit l'un des espaces en question, soit un **paramètre d'intérêt** relatif à la **loi** de cet espace. La **reconnaissance** consiste à estimer la loi associée à l'espace probabilisé en question, ou seulement à estimer un paramètre de celle-ci. En fait, l'estimation consiste à « reconnaître » le (« vrai ») modèle qui est à l'origine des observations. Les **tests associés** aux problèmes d'estimation permettent alors de confirmer (ou d'infirmer) cette reconnaissance (cf **identification, modèle identifié**).

(iii) Les techniques de reconnaissance des formes interviennent ainsi dans de nombreux domaines, eg :

\* physique : géologie, astrophysique ;

\* biologie : imagerie électronique médicale, catégories de pathologies ;

\* écologie : classification des espèces et biodiversité ;

\* psychologie : **profils de « personnalité »**, identification de comportements extrêmes (à risques) ;

\* sociologie : linguistique (reconnaissance de caractères typographiques ou de racines linguistiques), activités bancaires (reconnaissance de signatures), administration de la police (reconnaissance d'empreintes digitales ou de profil ADN), anthropologie (archéologie, paléontologie).