Calcul en très grande dimension.

Laurent Miclet (ex)-IRISA

Les VSA, pour *Vector Symbolic Architecture* sont des systèmes de calcul fondés sur un calcul vectoriel en très grande dimension d (typiquement $d = 10^4$). Le plus souvent, les composantes sont binaires, par exemple dans $\{-1, +1\}$ et on est alors dans l'espace de représentation $\mathcal{X} = 2^d$. Ces vecteurs sont utilisés à la fois pour représenter des données $\{a, b, \cdots\}$, des variables $\{A, B, \cdots\}$, mais aussi, de manière moins évidente, des relations (variable, valeur) du type $A = b$, des séquences, des arbres, etc.

Ils utilisent la propriété fondamentale que deux vecteurs aléatoires de $\mathcal{X}$ sont en pratique toujours orthogonaux. Le système de calcul consiste à assembler les vecteurs par des opérations internes à $\mathcal{X}$, puis à utiliser la conséquence de cette propriété : si deux vecteurs ne sont pas orthogonaux, ils ont certainement été assemblés par une certaine combinaison de ces opérations. La définition des opérations internes permettra de « décoder » quelle est cette combinaison.

L'exemple classique est le suivant : on décrit un pays, par exemple l'Argentine, par son nom ($Nom = Arg$), sa monnaie ($Mon = Pes$) et le nom de sa capitale ($Cap = BuA$). On tire six vecteurs aléatoires pour caractériser les six éléments
$\{Nom, Mon, Cap, Arg, Pes, BuA\}$ et on en déduit un nouveau vecteur de $\mathcal{X}$ qui résume les connaissances sur ce pays.

La première opération, le *Bind* est utilisée pour calculer trois nouveaux vecteurs $A1 = Bind(Nom, Arg)$, $A2 = Bind(Mon, Pes)$ et $A3 = Bind(Cap, BuA)$. La seconde, le *Bundle*, résume en un dixième vecteur $HolARG = Bundle(A1, A2, A3)$ les trois informations précédentes par une sorte de centre de gravité.

On peut calculer, à partir de trois nouveaux vecteurs aléatoires $\{USA, Dol, Wash\}$ le « résumé » des États-Unis comme :
$HOLUSA + Bundle(Bind(Nom, USA), Bind(Mon, Dol), Bind(Cap, Wash))$

Les propriétés du *Bind* et du *Bundle* permettent alors de répondre à des questions dans $\mathcal{X}$ du type : « que peut-on déduire de la connaissance de $HOLUSA$, $HOLARG$ et Dol ? » et de calculer un vecteur non orthogonal à Pes, ce qui est suffisant pour répondre « Le Peso est la monnaie de l'Argentine ».

Cet exposé présentera les propriétés de l'algèbre construite ces deux opérations (et sur une troisième, permettant de traiter les séquences) et présentera des applications à la reconnaissance d'images et à la robotique.

On parlera pour finir d'une extension de $\mathcal{X} = \{-1, +1\}^d$ à $\mathcal{X} = (\mathbb{Z}/p\mathbb{Z})^d$.

Référence de base : le site `https://www.hd-computing.com/`