

La marche aléatoire

On peut décrire le cas le plus simple de marche aléatoire est le choix aléatoire entre deux directions, par exemple le haut et le bas. Et on décide à chaque étape si on effectue un pas vers le haut ou vers le bas.

Si on généralise, la marche aléatoire est un déplacement où on choisit sa direction au hasard à chaque « étape », par exemple un homme saoul qui choisirait au hasard la direction de chacun de ses pas.

La marche aléatoire est utilisée en informatique par exemple dans les algorithmes du moteur de recherche Google.

La marche aléatoire est aussi utilisée pour simuler les variations boursières en économie et pour décrire le mouvement de particule dans un fluide (mouvement brownien) en Physique.

Le moteur de recherche Google utilise par exemple la marche aléatoire pour parcourir, identifier et classer les pages du réseau internet.

Réseau bayésien

Un réseau bayésien est un modèle graphique probabiliste qui permet de modéliser des connaissances sur un domaine incertain.

Un réseau bayésien est représenté sous forme d'arbre, chaque branche représentant un événement indépendant et le nœud de deux branches est un événement qui dépend des événements des deux branches.

Un réseau bayésien a pour but de :

- Prévoir
- Diagnostiquer
- Simuler
- Analyser

Par exemple un réseau bayésien peut être utilisé pour résoudre une panne informatique :

On a les informations de la panne : L'écran ne s'allume plus.

Ecran qui ne s'allume pas

Ordinateur allumé

Carte Graphique en panne

Probabilités

Le terme probabilité désigne l'opposé du concept de certitude. C'est la chance qu'un événement se produise.

La probabilité est une science mathématique.

La probabilité d'un événement est un nombre compris entre 0 et 1 et détermine son degré de certitude. Plus ce nombre est grand, plus la chance que l'événement se produise est grande.

Les probabilités dans l'informatique

Les probabilités sont aussi utilisées pour que l'ordinateur se débrouille tout seul (comme dans la correction - relative - des erreurs d'un code ou en informatique décisionnelle).

Code de Hamming

Méthode de gestion des erreurs (notamment lors de l'envoi de messages) basée sur le fait que la probabilité d'avoir une seule erreur dans le code est plus importante qu'en avoir deux.

Codage de données

Pour coder une information en informatique, il faut savoir le nombre "d'états" qu'elle peut avoir (pour le temps, il peut être ensoleillé/pluvieux/orageux/...) afin de savoir quelle place lui allouer (pour calculer la "quantité d'information" (Q) d'une information : $Q = \log_2 (1/P(X))$).

Mais avec cette méthode, chaque "état" a la même probabilité d'exister, ce qui est vrai en théorie mais pas en pratique (l'utilisation du 'e' et du 'z' dans la langue française par exemple).

C'est ce qui a amené à la création d'algorithmes de compression (tels que le code de Huffman, le LZSS, LZ77, etc...) souvent basés sur la notion de probabilité (telle lettre apparaît plus souvent qu'une autre, même chose pour un pixel).

Code de Huffman

Ce code permet de compresser des données (texte, image, son) en codant avec des mots binaires très courts les symboles apparaissant le plus souvent, cela permet en général de diminuer par deux ou trois le nombre de bits nécessaires.

Exemples de secteur d'activité où l'informatique décisionnelle est exercée : Banque, Assurance, Commerce...

Informatique décisionnelle

L'informatique décisionnelle est un ensemble de solutions informatiques permettant l'analyse des données d'une entreprise, afin d'en dégager les informations qualitatives nouvelles qui vont fonder des décisions tactiques ou stratégiques.

Donnée -> Information -> Connaissance -> Sagesse

Donnée : résultat direct d'une mesure.

Information: donnée à laquelle un sens et une interprétation ont été donnés.

Connaissance: Résultat d'une réflexion sur les infos analysées, se base sur les expériences, les idées, valeurs et avis des personnes consultées.

Sagesse: Etat d'esprit de discernement final sur le contenu et de jugement de bon sens

Chaîne décisionnelle: planification -> ETL (extract, transform, load) -> stockage -> ANALYSES -> Restitution

Les probabilités y sont utilisées via les statistiques, le Data Mining.