

Algorithme 1 : Algorithme ID3.**Fonction** ID3(S_p, S, D)**Entrées** : S_p sous-ensemble des exemples du nœud parent, S sous-ensemble des exemples à considérer, D sous-ensemble des attributs à considérer**Sorties** : Un arbre de décision**début** **si** l'ensemble des exemples S est vide **alors** ↳ **retourner** **si** l'ensemble A des attributs est vide **alors** ↳ **retourner** **si** tous les exemples de S ont une même étiquette y **alors** ↳ **retourner** **sinon** Soit $A \in D$ l'attribut de plus grand gain $G(S, A)$. Construire l'arbre de racine A et de sous-arbre gauche ID3($S, S_{A=NO}, D \setminus \{A\}$) et de sous-arbre droit ID3($S, S_{A=YES}, D \setminus \{A\}$).

Exercice 6. On considère le schéma de base de données suivant, qui décrit un ensemble de fabricants de matériel informatique, les matériels qu'ils vendent, leurs clients et ce qu'achètent leurs clients. Les attributs des clés primaires des six premières relations sont soulignés.

 Production(NomFabricant, Modele) Ordinateur(Modele, Frequence, Ram, Dd, Prix) Portable(Modele, Frequence, Ram, Dd, Ecran, Prix) Imprimante(Modele, Couleur, Type, Prix) Fabricant(Nom, Adresse, NomPatron) Client(Num, Nom, Prenom)

Achat(NumClient, NomFabricant, Modele, Quantite)

Chaque client possède un numéro unique connu de tous les fabricants. La relation **Production** donne pour chaque fabricant l'ensemble des modèles fabriqués par ce fabricant. Deux fabricants différents peuvent proposer le même matériel. La relation **Ordinateur** donne pour chaque modèle d'ordinateur la vitesse du processeur (en Hz), les tailles de la Ram et du disque dur (en Go) et le prix de l'ordinateur (en €). La relation **Portable**, en plus des attributs précédents, comporte la taille de l'écran (en pouces). La relation **Imprimante** indique pour chaque modèle d'imprimante si elle imprime en couleur (vrai/faux), le type d'impression (laser ou jet d'encre) et le prix (en €). La relation **Fabricant** stocke les nom et adresse de chaque fabricant, ainsi que le nom de son patron. La relation **Client** stocke les noms et prénoms de tous les clients de tous les fabricants. Enfin la relation **Achat** regroupe les quadruplets (client c , fabricant f , modele m , quantité q) tels que le client de numéro c a acheté q fois le modèle m au fabricant de nom f . On suppose que l'attribut **Quantite** est toujours strictement positif.

1. Proposer une clé primaire pour la relation **Achat** et indiquer ses conséquences en terme de modélisation.
2. Identifier l'ensemble des clés étrangères éventuelles de chaque table.
3. Donner en SQL des requêtes répondant aux questions suivantes :
 - (a) Quels sont les numéros de modèles des matériels (ordinateur, portable ou imprimante) fabriqués par l'entreprise de nom Durand ?
 - (b) Quels sont les noms et adresses des fabricants produisant des portables dont le disque dur a une capacité d'au moins 500 Go ?
 - (c) Quels sont les noms des fabricants qui produisent au moins 10 modèles différents d'imprimantes ?
 - (d) Quels sont les numéros des clients n'ayant acheté aucune imprimante ?