




La Normalisation



5/7/2021

Z. Kedad

Pourquoi normaliser ?



- ❑ Qu'est-ce qu'un bon schéma relationnel ?
- ❑ Tous les schémas relationnels sont-ils équivalents ?



5/7/2021

Z. Kedad

Pourquoi normaliser ?

❑ Exemple

Propriété	Nom	Prénom	Adresse	Date achat	Numéro	Marque
	Durand	Pierre	5, Rue des..	1/1/99	902FE75	Peugeot
	Dupond	Jean	12, Avenue ..	3/4/89	87JI75	Renault
	Dupond	Jean	12, Avenue	10/12/94	788GF21	Ford

❑ Redondances ⇒ Anomalies

- » de suppression
- » d'insertion
- » de mise à jour

5/7/2021

Z. Kedad

Anomalies de suppression

❑ Exemple

Propriété	Nom	Prénom	Adresse	Date achat	Numéro	Marque
	Durand	Pierre	5, Rue des..	1/1/99	902FE75	Peugeot
	Dupond	Jean	12, Avenue ..	3/4/89	87JI75	Renault
	Dupond	Jean	20, Boulevard	10/12/94	788GF21	Ford

- la suppression du véhicule de la personne « Durand » entraîne la suppression de cette personne

5/7/2021

Z. Kedad

Anomalies d'insertion

❑ Exemple

Propriété	Nom	Prénom	Adresse	Date achat	Numéro	Marque
	Durand	Pierre	5, Rue des..	1/1/99	902FE75	Peugeot
	Dupond	Jean	12, Avenue ..	3/4/89	87JI75	Renault
	Dupond	Jean	20, Boulevard	10/12/94	788GF21	Ford

- l'insertion d'un véhicule ne peut se faire sans connaître le propriétaire
- l'insertion d'une personne ne peut se faire s'il n'a pas de véhicule

5/7/2021

Z. Kedad

Anomalies de modification

❑ Exemple

Propriété	Nom	Prénom	Adresse	Date achat	Numéro	Marque
	Durand	Pierre	5, Rue des..	1/1/99	902FE75	Peugeot
	Dupond	Jean	12, Avenue ..	3/4/89	87JI75	Renault
	Dupond	Jean	12, Avenue	10/12/94	788GF21	Ford

- la mise à jour de l'adresse de la personne « Dupond » doit se faire sur deux tuples

5/7/2021

Z. Kedad

Espace de stockage

- ❑ Le choix des schémas de relation à un impact sur la place occupée par les données

- ❑ Schéma 1
 - Personne (Nom, Prenom, Adresse, Date_Achat, Numero, Marque)

- ❑ Schéma 2
 - Personne (Id, Nom, Prenom, Adresse)
 - Voiture (Numero, Marque, Date Achat, Id)

5/7/2021

Z. Kedad

Autres critères

- ❑ Performances
 - L'accès à des données réparties sur plusieurs tables nécessite des jointures

- ❑ Cohérence
 - La saisie d'une même donnée plusieurs fois peut entraîner des erreurs

5/7/2021

Z. Kedad

Mesures Informelles

- ❑ Simplicité
 - L'interprétation doit être simple
- ❑ Limiter la redondance
 - Anomalies, volume
- ❑ Limiter les valeurs manquantes
- ❑ Eviter les tuples parasites
 - Après une décomposition, les jointures ne doivent pas introduire des tuples parasites

5/7/2021

Z. Kedad

Problème des valeurs manquantes

- ❑ Sémantique ?
 - Valeur existante mais non connue
 - ❖ Age d'une personne
 - Valeur non applicable
 - ❖ Salaire d'un étudiant
 - Indéterminée...
- ❑ En pratique, seule la valeur NULL peut être utilisée

5/7/2021

Z. Kedad

Tuples parasites

❑ Table « Visites »

Nom	Ville	Lieu
Durand	Versailles	Château
Dupond	Versailles	Jardins

❑ Décomposition

- Visiteurs

Nom	Ville
Durand	Versailles
Dupond	Versailles

- Sites

Ville	Lieu
Versailles	Château
Versailles	Jardins

5/7/2021

Z. Kedad

Bilan

❑ Tous les schémas relationnels ne sont pas équivalents

❑ Critères informels

- Donnent une indication de la qualité du schéma
- Difficiles à appliquer sur des schémas de grandes taille

❑ Nécessité de disposer d'un outil précis de mesure de la qualité



Normalisation

5/7/2021

Z. Kedad

Dépendances Fonctionnelles

- ❑ Une dépendance fonctionnelle (DF) sur un schéma de relation R est une expression $X \rightarrow Y$, $X, Y \subseteq R$
- ❑ La DF est vérifiée dans une relation r si et seulement si :
$$\forall t_i, t_j \in r, t_i[X] = t_j[X] \Rightarrow t_i[Y] = t_j[Y]$$

5/7/2021

Z. Kedad

Dépendances Fonctionnelles

- ❑ DF Triviale
 - $X \rightarrow Y$ est dite triviale si $Y \subseteq X$

5/7/2021

Z. Kedad

Dépendances Fonctionnelles

❑ Soit F un ensemble de DF

❑ $F \Rightarrow X \rightarrow Y$

- Pour toute base de données pour laquelle toutes les DF de F sont vérifiées, $X \rightarrow Y$ est également vérifiée

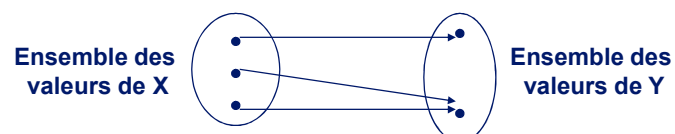
5/7/2021

Z. Kedad

Dépendances fonctionnelles

❑ Exemple

- Véhicule(Numéro, Type, Couleur)
 - ❖ Numéro \rightarrow type
- Personne(Nom, Prénom, Adresse)
 - ❖ Nom, prénom \rightarrow adresse



5/7/2021

Z. Kedad

Exercice

- Personne(NSS, NOM, PRENOM, AGE)
 - ❖ NSS → NOM
 - ❖ NOM → NSS
- Voiture(NUMERO, COULEUR, MARQUE, PUISSANCE, TYPE)
 - ❖ PUISSANCE, TYPE → MARQUE
 - ❖ MARQUE → PUISSANCE
 - ❖ PUISSANCE → TYPE

5/7/2021

Z. Kedad

Exercice (suite)

- Propriété (NUM_VEHICULE, NUM_PROPRIETAIRE, DATE)
 - ❖ NUM_VEHICULE → NUM_PROPRIETAIRE
 - ❖ NUM_PROPRIETAIRE → NUM_VEHICULE
 - ❖ (NUM_VEHICULE , NUM_PROPRIETAIRE) → DATE

5/7/2021

Z. Kedad

Clé

- ❑ Soit R un schéma de relation
- ❑ Un ensemble d'attributs $K \subseteq R$ est clé (ou clé candidate) de R par rapport à l'ensemble de DF F si et seulement si :
 - $F \Rightarrow K \rightarrow R$
 - $\forall X \subset K \ F \not\Rightarrow X \rightarrow R$
- ❑ Un ensemble K d'attributs est une superclé de R s'il contient une clé de R
- ❑ Parmi les clés candidate de R , on choisit une clé primaire

5/7/2021

Z. Kedad

Inférences de DF : Axiomes d'Armstrong

- ❑ Réflexivité :
 - Si $Y \subseteq X$ Alors $X \rightarrow Y$
 - ❖ Nom, Prénom \rightarrow Nom
 - ❖ DF triviale
- ❑ Augmentation :
 - Si $X \rightarrow Y$ Alors $(X, Z) \rightarrow (Y, Z)$
 - ❖ si $Nss \rightarrow Age$ alors $Nss, Nom \rightarrow Age$

5/7/2021

Z. Kedad

Inférences de DF : Axiomes d'Armstrong

□ Transitivité :

- Si $X \rightarrow Y$ et $Y \rightarrow Z$ Alors $X \rightarrow Z$
 - ❖ si $N^{\circ}\text{Véhicule} \rightarrow \text{Type}$ et $\text{Type} \rightarrow \text{Marque}$
alors $N^{\circ}\text{Véhicule} \rightarrow \text{Marque}$

5/7/2021

Z. Kedad

Inférence de DF

□ Soit F un ensemble de DF

- On note $F \vdash X \rightarrow Y$ si $X \rightarrow Y$ peut être dérivée de F en utilisant les axiomes d'Armstrong

5/7/2021

Z. Kedad

Autres règles d'inférences

❑ Pseudo-transitivité :

- Si $X \rightarrow Y$ et $(Y, Z) \rightarrow T$ Alors $(X, Z) \rightarrow T$
 - ❖ si $N^{\circ}\text{Véhicule} \rightarrow \text{NSS}$ et
 - ❖ $\text{NSS}, N^{\circ}\text{Banque} \rightarrow \text{Montant_emprunté}$
 - ❖ Alors $N^{\circ}\text{Véhicule}, N^{\circ}\text{Banque} \rightarrow \text{Montant_emprunté}$

❑ Décomposition :

- Si $X \rightarrow Y$ et $Z \subseteq Y$ Alors $X \rightarrow Z$
 - ❖ si $\text{Nss} \rightarrow (\text{Nom}, \text{Prénom}, \text{Age})$
 - ❖ alors $\text{Nss} \rightarrow \text{Nom}$, $\text{Nss} \rightarrow \text{Prénom}$, $\text{Nss} \rightarrow \text{Age}$

❑ Union :

- Si $X \rightarrow Y$ et $X \rightarrow Z$ Alors $X \rightarrow Y, Z$
 - ❖ si $\text{Nss} \rightarrow \text{Nom}$ et $\text{Nss} \rightarrow \text{Prénom}$ alors
 - ❖ $\text{Nss} \rightarrow (\text{Nom}, \text{Prénom})$

5/7/2021

Z. Kedad

Implication de DF

❑ Soit F un ensemble de DF et $X \rightarrow Y$ une DF, comment vérifier $F \vdash X \rightarrow Y$?

- Montrer que toute relation qui satisfait les DF de F satisfait aussi $X \rightarrow Y$ ou
- Trouver une séquence d'axiomes conduisant à $X \rightarrow Y$ à partir de F ou
- Utiliser la notion de fermeture

5/7/2021

Z. Kedad

Fermeture d'un ensemble d'attributs

- ❑ La fermeture d'un ensemble d'attributs X par rapport à un ensemble de DF F est définie par
- ❑ $X^+ = \{A \in R \mid F \vdash X \rightarrow A\}$
- ❑ $F \vdash X \rightarrow Y$ est équivalent à $Y \subseteq X^+$

5/7/2021

Z. Kedad

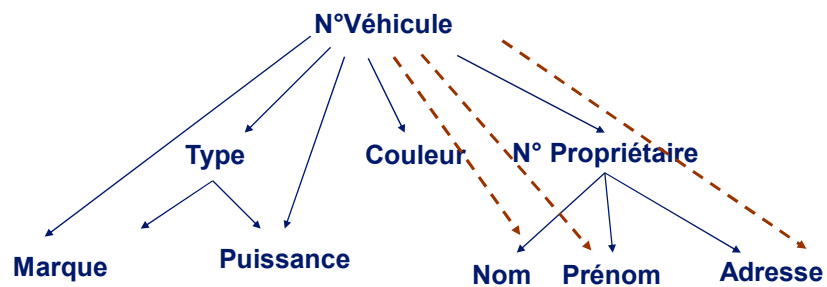
Fermeture d'un ensemble de dépendances

- ❑ Ensemble des dépendances fonctionnelles pouvant se déduire logiquement de cet ensemble (en utilisant les axiomes d'Armstrong et règles d'inférences)
- ❑ La fermeture d'un ensemble de DF F est notée F^+

5/7/2021

Z. Kedad

Exemple



5/7/2021

Z. Kedad

Dépendance fonctionnelle élémentaire

❑ C'est une dépendance :

- non triviale,
- telle qu'il n'existe aucun ensemble d'attributs inclus dans sa partie gauche et qui détermine sa partie droite

❑ Exemples

- $(A, B) \rightarrow A$
- $D \rightarrow E, (D, F) \rightarrow E$
- $G \rightarrow H, H \rightarrow I, (G, J) \rightarrow I$

5/7/2021

Z. Kedad

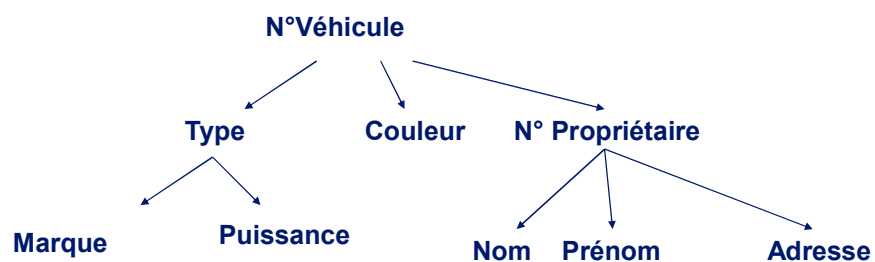
Couverture minimale

- ❑ Un ensemble de dépendances fonctionnelles élémentaires F' constitue une couverture minimale d'un ensemble de dépendances F si :
 - F' et F ont la même fermeture
 - Il n'existe pas d'ensemble F'' inclus dans F' tel que F'' ait la même fermeture que F

5/7/2021

Z. Kedad

Exemple



5/7/2021

Z. Kedad

Exercice

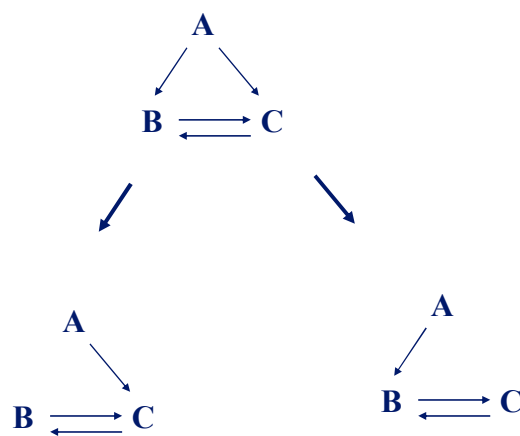
❑ Trouver les couvertures minimales de l'ensemble de dépendances suivant :

❑ $F1 = \{A \rightarrow B, A \rightarrow C, B \rightarrow C, C \rightarrow B\}$

5/7/2021

Z. Kedad

Solution



5/7/2021

Z. Kedad

Exercice

- ❑ Soit la relation $R(A, B, C, D, E)$
- ❑ Soit l'ensemble de dépendances :
 - $A \rightarrow B, C, D$
 - $B \rightarrow D, E$
 - $D \rightarrow E$
 - $C \rightarrow E$
 - $A \rightarrow E$
- ❑ Trouver une couverture minimale de cet ensemble de dépendances fonctionnelles

5/7/2021

Z. Kedad

Egalité et équivalence de deux graphes

- ❑ $G1 = G2$
 - ensembles de dépendances identiques
- ❑ $G1$ est équivalent à $G2$
 - les graphes génèrent la même fermeture



5/7/2021

Z. Kedad

Propriétés d'une décomposition

❑ Décomposition

- Substitution d'un schéma de relations R par plusieurs schémas de relation R_1, R_2, \dots, R_n
- $R = \bigcup_{i=1..n} R_i$

❑ Propriétés d'une décomposition

- Sans perte
 - ❖ La relation initiale doit être recalculée par jointure
- Préservant les DF
 - ❖ Les DF doivent être vérifiées sans jointures

5/7/2021

Z. Kedad

Décomposition sans perte

❑ $R = R1 \cup R2$, F un ensemble de DF sur R

❑ La décomposition est sans perte si

❑ $F \vdash R1 \cap R2 \rightarrow R2$ ou $F \vdash R1 \cap R2 \rightarrow R1$

5/7/2021

Z. Kedad

Décomposition préservant les dépendances

- ❑ Soit F un ensemble de DF sur un schéma de relation R
- ❑ La décomposition $R = \{R_1, R_2, \dots, R_n\}$ préserve les dépendances si l'union des ensembles de DF des schémas R_i est égale à F

5/7/2021

Z. Kedad

Première forme normale

- ❑ Une relation est en 1ère forme normale si chacun de ses attributs prend une valeur atomique
- ❑ Personne(Nom, Adresse, Prénom)
 - Si l'attribut Prénom peut prendre plusieurs valeurs (prénom1, prénom2, prénom3) alors la relation personne n'est pas en 1ère forme normale
 - Si adresse est constituée de N°, Rue, Code postal et Ville, la relation personne n'est pas en 1ère forme normale

5/7/2021

Z. Kedad

Deuxième forme normale

❑ Une relation est en deuxième forme normale si elle est en première forme normale et si chaque attribut non clé dépend de la clé et non d'un sous ensemble de la clé

- $R(\underline{A}, \underline{B}, C, D, E)$
- R est en 2ème forme normale si elle est en 1ère forme normale et s'il n'existe pas de dépendances de la forme :
 - ❖ $A \rightarrow \dots$ ou $B \rightarrow \dots$

5/7/2021

Z. Kedad

Exemple de relation non en 2FN

❑ $\text{Stock}(\underline{\text{Pièce}}, \underline{\text{Entrepôt}}, \text{Quantité}, \text{Adresse})$

Stock	Pièce	Entrepôt	Adresse	Quantité
	DF45	12	5, Rue des..	200
	DF45	45	12, Avenue ..	154
	BTR5	12	5, Rue des	10

❑ Redondance : l'adresse de l'entrepôt est répétée pour chaque pièce qui y est stockée

5/7/2021

Z. Kedad

Exercice

❑ Les relations suivantes sont-elles en 2FN ?

- Voiture(NUMERO, COULEUR, MARQUE, PUISSANCE, TYPE)
- Propriété (NUM_VEHICULE, NUM_PROPRIETAIRE, TYPE, MARQUE, DATE_ACHAT)
 - On suppose qu'un véhicule peut appartenir à plusieurs propriétaires (à des dates différentes) et qu'un propriétaire peut avoir plusieurs véhicules

5/7/2021

Z. Kedad

Troisième forme normale

❑ Une relation est en troisième forme normale si elle est en deuxième forme normale et s'il n'existe pas de dépendance entre attributs non clés (DF transitives)

❑ $R(\underline{A}, \underline{B}, C, D)$

- R est en 3ème forme normale si elle est en 2ème forme normale et s'il n'existe pas de dépendances de la forme :
 - ❖ $C \rightarrow D$ ou $D \rightarrow C$

5/7/2021

Z. Kedad

Exemple de relation non en 3FN

- ❑ Voiture(Numéro, Marque, Puissance, Type)

Voiture	Numéro	Type	Marque	Puissance
	902FE75	Fiesta	Ford	4
	87JI75	Fiesta	Ford	4
	788GF21	Fiesta	Ford	4

- ❑ Redondance : pour chaque occurrence des voitures de même type, on stocke les caractéristiques propres à ce type

5/7/2021

Z. Kedad

Exercice

- ❑ Les relations suivantes sont-elles en 3FN ?
- ❑ Facture(N°FACTURE, DATE, NOM_CLIENT, PRODUIT, QTE, PRIX_UNITAIRE)
- ❑ Propriétaire(N°PROPRIETAIRE, NOM)
- ❑ Conduit(N°CONDUCTEUR, N°VEHICULE, NOM, DATE)

5/7/2021

Z. Kedad

Forme normale de Boyce Codd

- ❑ Une relation est en forme normale de Boyce Codd si elle est en troisième forme normale et s'il n'existe aucune autre dépendance fonctionnelle que celles liant la clé à un attribut non clé
- ❑ $R(\underline{A}, \underline{B}, C)$
 - R est en forme normale de Boyce Codd si elle est en 3NF et s'il n'existe pas de DF de la forme :
 - ❖ $C \rightarrow A$ ou $C \rightarrow B$

5/7/2021

Z. Kedad

Exercice

- ❑ Les relations suivantes sont-elles en forme normale de Boyce Codd ?
- ❑ Employé(N°EMPLOYE, N°ENTREPRISE, NSS, NOM, PRENOM, AGE)
- ❑ Personne(NSS, NOM, PRENOM, AGE)

5/7/2021

Z. Kedad

Degré de normalisation

- ❑ Les redondances sont sources d'anomalies et d'incohérences
- ❑ La présence de redondance est caractérisée par un degré de normalité :
 - plus le degré de normalité est élevé, moins il y a de redondances

5/7/2021

Z. Kedad

Algorithmes de normalisation

- ❑ Algorithme de synthèse
 - 3FN
- ❑ Algorithme de décomposition
 - Forme normale de Boyce Codd, 4FN

5/7/2021

Z. Kedad

Algorithme de synthèse

❑ Notion de relation universelle :

- union de tous les attributs des relations constituant la base de données
 - $R1(A,B,C)$
 - $R2(B,D,E,F)$
 - $U(A,B,C,D,E,F)$

5/7/2021

Z. Kedad

Principe de l'algorithme

❑ A partir d'un graphe de dépendances fonctionnelles :

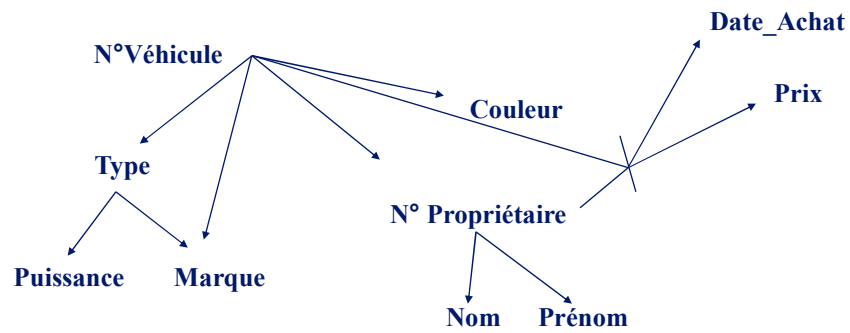
- 1. Rechercher la fermeture
- 2. Rechercher les clés candidates
- 3. Regrouper les clés candidates équivalentes
 - ❖ $K1 \equiv K2$ si $K1 \rightarrow K2$ et $K2 \rightarrow K1$ (cycle)
- 4. Rechercher la couverture minimale
- 5. Synthétiser les relations à partir des groupes de clés

5/7/2021

Z. Kedad

Exemple d'application de l'algorithme de synthèse

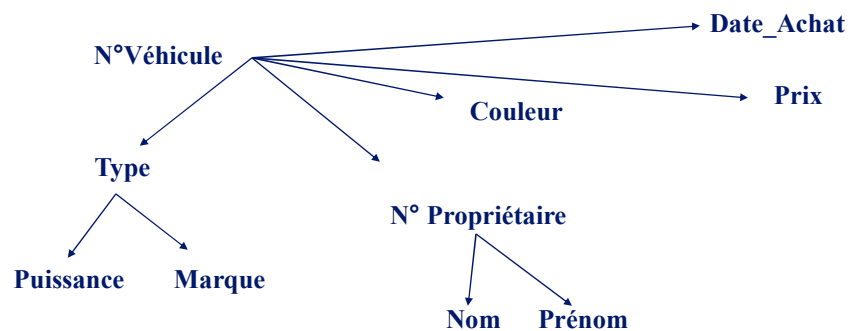
□ Graphe de dépendances fonctionnelles



5/7/2021

Z. Kedad

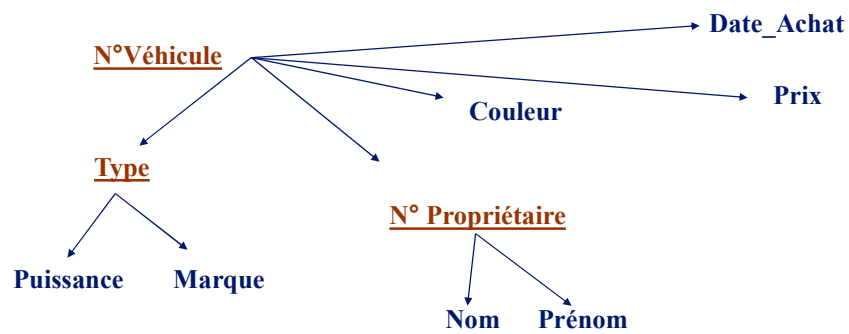
Couverture minimale



5/7/2021

Z. Kedad

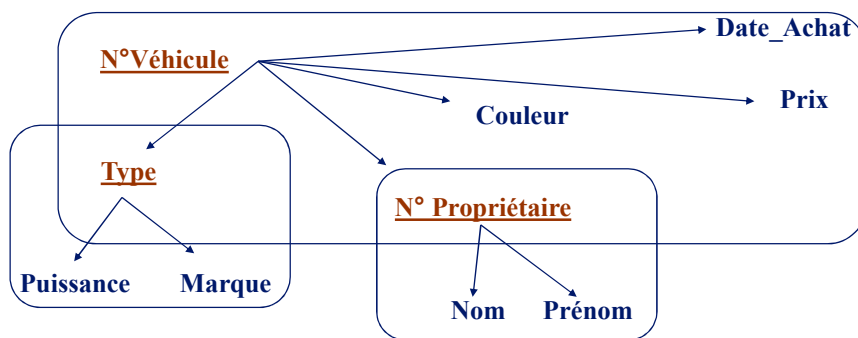
Recherche des clés candidates



5/7/2021

Z. Kedad

Synthèse des relations



- Véhicule(N°Véhicule, Type, Couleur, N°Propri., Date_achat, Prix)
- Marque(Type, Puissance, Marque)
- Personne(N°Propriétaire, Nom, Prénom)

5/7/2021

Z. Kedad

Exercice

❑ Appliquer l'algorithme de synthèse à la relation universelle suivante :

- $R(A, B, C, D)$

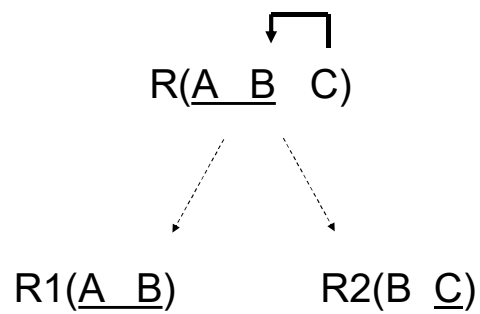
❑ Avec l'ensemble de dépendances :

- $A \rightarrow B$
- $C \rightarrow B$
- $(C, D) \rightarrow A$
- $(C, D) \rightarrow B$

5/7/2021

Z. Kedad

Décomposition



5/7/2021

Z. Kedad

Dépendances Multivaluées

❑ $R(A, B, C, D)$

❑ $A \twoheadrightarrow B$

- à A correspond un ensemble de valeurs pour B, indépendant des valeurs des autres attributs de R

- $A \twoheadrightarrow B$???

A	B	C
1	11	110
1	11	110
1	12	112
3	33	330
3	33	440

5/7/2021

Z. Kedad

Quatrième Forme Normale

- ❑ Une relation R est en quatrième forme normale par rapport à un ensemble de DF et de DMV si pour chaque DM $A \twoheadrightarrow B$ non triviale, A est une superclé de R.

5/7/2021

Z. Kedad