

ДИСЦИПЛІНА «БАЗИ ДАНИХ»

Реляційні бази даних

ЛЕКЦІЯ 9

Функціональні залежності.
Нормальні форми

доц. Мазурова О.О.,
ст. викл. Широкопетлева М.С.

Кафедра Програмної інженерії,
факультет Комп'ютерних наук, ХНУРЕ

Функціональні залежності.

Нормальні форми

Мета лекції:

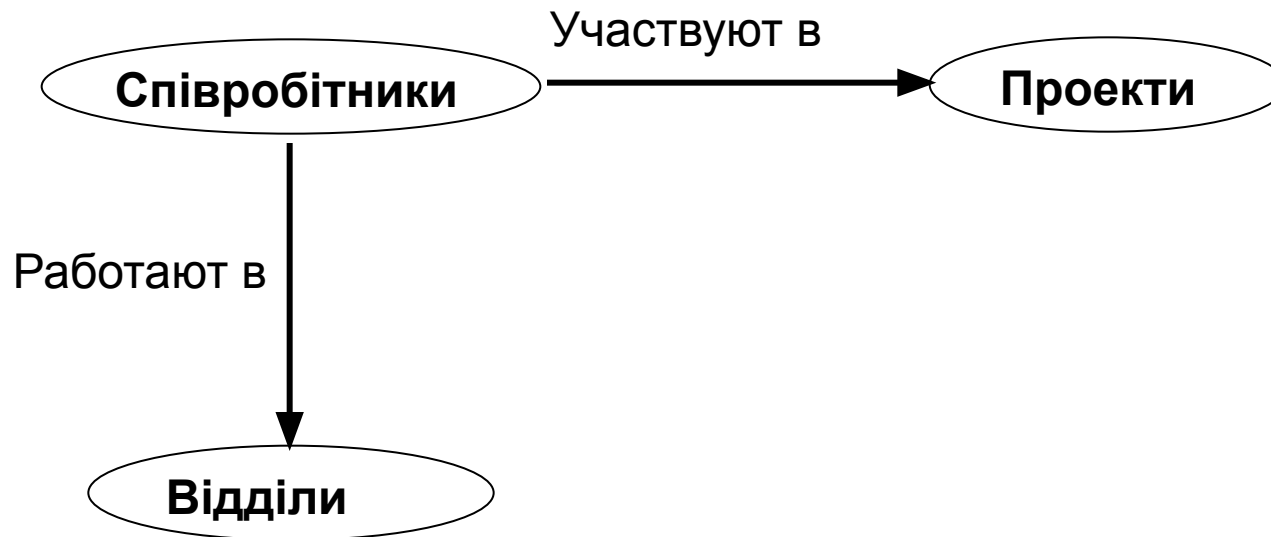
Розглянути види функціональних залежностей, основні нормальні форми та етапи нормалізації відношень.
Навчитися визначати функціональні залежності за моделлю предметної галузі, нормалізувати реляційні відношення.

Зміст:

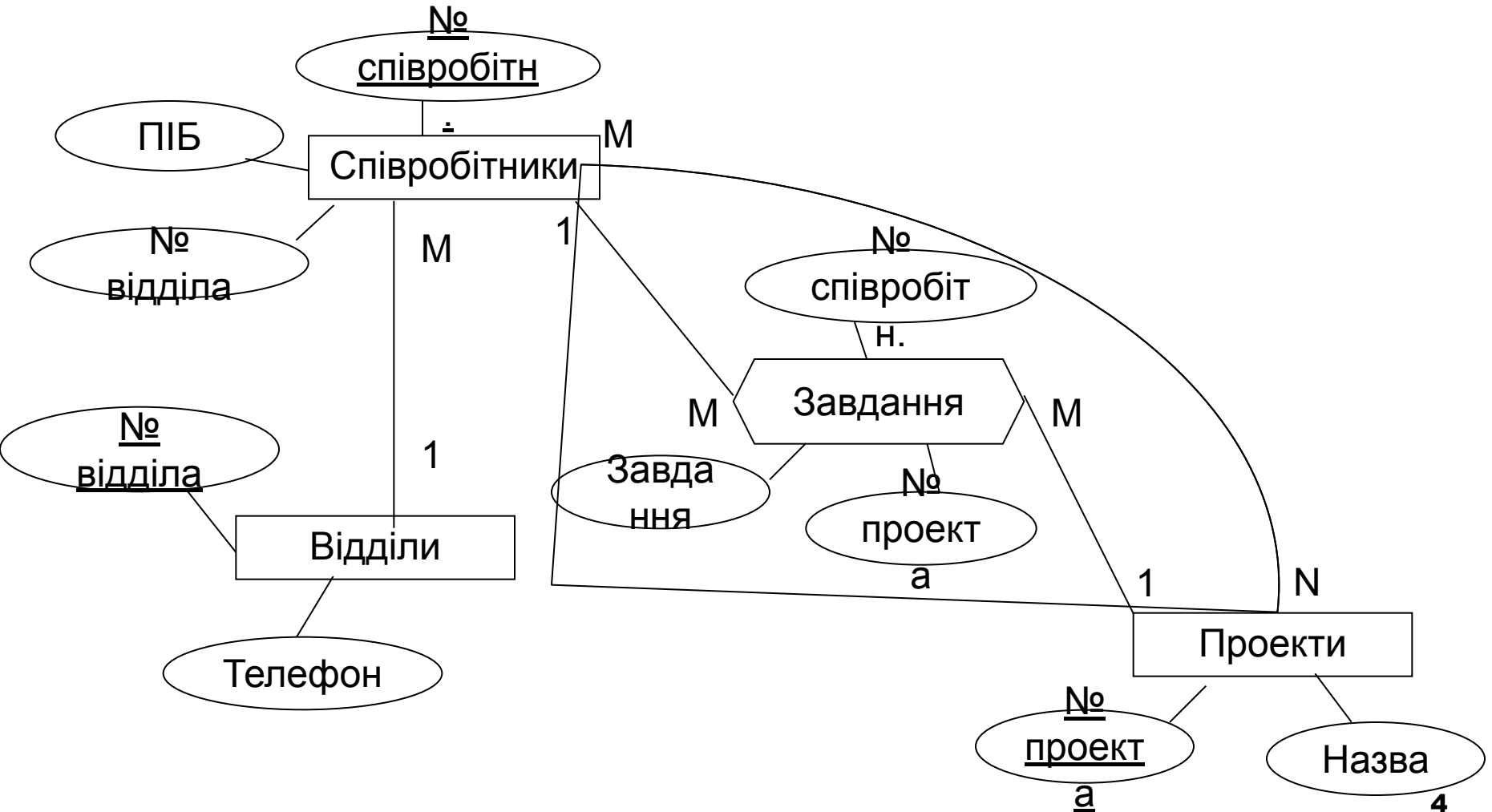
- Аномалії баз даних. Проблеми проектування баз даних
- Функціональні залежності
 - Повні; багатозначні; транзитивні ФЗ
 - Приклади визначення ФЗ
- Раціональна схема та нормалізація.
- Основні нормальні форми
 - Перша нормальна форма (1НФ)
 - Друга нормальна форма (2НФ)
 - Третя нормальна форма (3НФ)
 - Нормальна форма Бойса-Кодда
 - Четверта нормальна форма (4НФ)
 - П'ята нормальна форма (5НФ)

- Алгоритми нормалізації
- ХНУ ІТЕ, факультет КІІ, кафедра ІІІ, тел. 70-21-446, e-mail: mazurova_ok@mail.ru ©

Проектування баз даних (1-й крок – аналіз та концептуальне моделювання)



Проектування баз даних (2-й крок – інфологічне моделювання (ER-діаграма))



Проектування баз даних (3-й крок – логічне моделювання)

Відділи

<i>Н_ВІД</i>	ТЕЛ
1	11-22-33
2	33-22-11

Проекти

<i>Н_ПРО</i>	ПРОЕКТ
1	Космос
2	Климат

Співробітники

<i>Н_СПІВР</i>	ПІБ	<i>Н_ВІД</i>
1	Іванов	1
2	Петров	1
3	Сидоров	2

Завдання

<i>Н_СПІВР</i>	<i>Н_ПРО</i>	<i>Н_ЗАВД</i>
1	1	1
1	2	1
2	1	2
3	1	3
3	2	2

1

M

1

M

M

1

Проектування баз даних (4-й крок – фізичне моделювання)

Сотрудники : таблица

Имя поля	Тип данных	Описание
Номер сотрудника	Счетчик	
Фамилия	Текстовый	
Номер отдела	Числовой	

Свойства поля

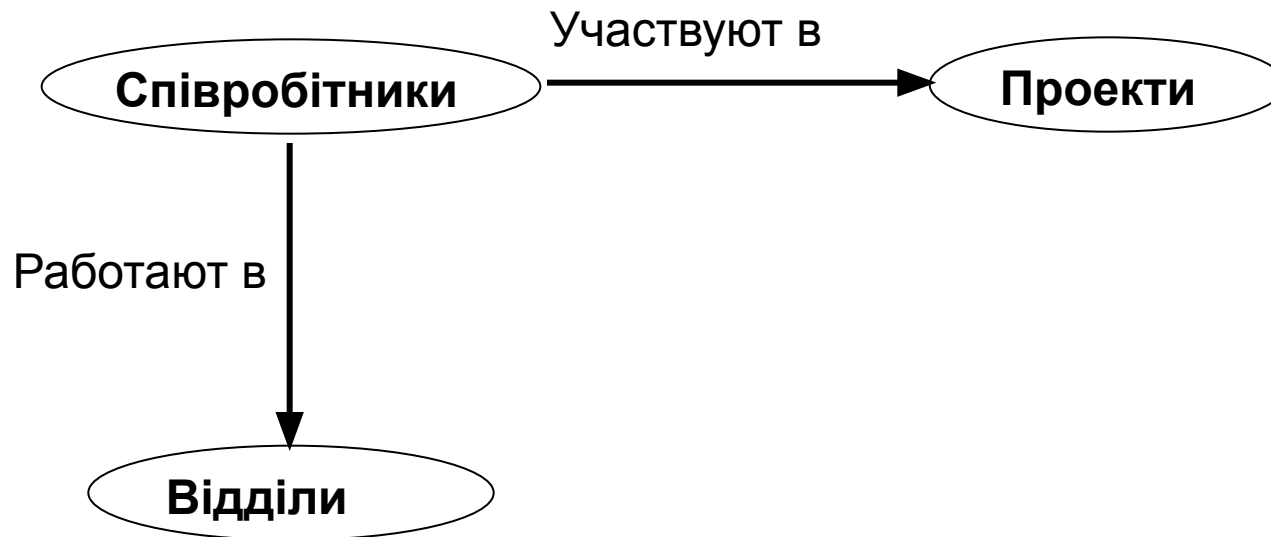
Общие Подстановка

Размер поля	50
Формат поля	
Маска ввода	
Подпись	
Значение по умолчанию	
Условие на значение	
Сообщение об ошибке	
Обязательное поле	Да
Пустые строки	Да
Индексированное поле	Нет
Сжатие Юникод	Да
Режим ИМЕ	Нет контроля
Режим предложений ИМЕ	Нет

Индекс ускоряет поиск и сортировку в данном поле, но замедляет обновление. Чтобы запретить ввод в поле повторяющихся значений, выберите значение "Да (Совпадения не допускаются)". Для справки по индексированным полям нажмите клавишу F1.



Проектування баз даних (1-й крок – аналіз та концептуальне моделювання)



Проектування баз даних (2-й крок – логічне моделювання)

Відділи - Співробітники – Проекти – Завдання = Підприємство

<i>Н_СПР*</i>	ПІБ	Н_ВІД	ТЕЛ	<i>Н_ПРО*</i>	ПРОЕКТ	Н_ЗАВД
1	Іванов	1	11-22-33	1	Космос	1
1	Іванов	1	11-22-33	2	КлимаТ	1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	33-22-11	1	Космос	3
3	Сидоров	2	33-22-11	2	КлимаТ	2

Проблеми проектування баз даних

Проблема логічного проектування БД – яким чином відобразити об'єкти предметної галузі в абстрактні об'єкти моделі даних, щоб це відображення не протирічило семантиці предметної галузі;

Проблема фізичного проектування БД – як забезпечити ефективність виконання запитів до бази даних, як розмістити дані у зовнішній пам'яті.

Аномалії баз даних

Аномалія – протиріччя між моделлю предметної галузі та фізичною моделлю даних, тобто неадекватність моделі даних предметній галузі.

Відділи - Співробітники – Проекти - Завдання

<i>Н_СПР*</i>	ПІБ	Н_ВІД	ТЕЛ	<i>Н_ПРО*</i>	ПРОЕКТ	Н_ЗАВД
1	Іванов	1	11-22-33	1	Космос	1
1	Іванов	1	11-22-33	2	Климат	1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	33-22-11	1	Космос	3
3	Сидоров	2	33 22 11	2	Климат	2

При видаленні співробітника №3 інформація про відділ №2 відсутня

Причини аномалій:

- зберігання в одному відношенні різнорідної інформації;
- надмірність даних.

Види аномалій баз даних

Існують наступні аномалії:

- **Аномалія додавання (INSERT)** – неадекватності, що проявляються під час додавання в БД нового запису;
- **Аномалії оновлення (UPDATE)** - неадекватності, що проявляються під час зміни даних в БД;
- **Аномалії видалення (DELETE)** – неадекватності, що проявляються під час видалення записів з БД.

Процес проектування БД – процес нормалізації похідних схем відношень, таким чином, що кожна наступна нормальна форма має ліпші властивості, нід попередня.

Функціональні залежності (Functional Dependency – FD)

Функціональна залежність . Атрибут В відношення функціонально залежить від атрибута А того ж самого відношення (атрибути можуть бути складовими) в тому і тільки в тому випадку, коли в любу певну мить часу для кожного з різних значень атрибута А обов'язково існує тільки одне з різних значень атрибута В.

Екзаменаційна ведомість

№ зал. кн.*	ПІБ	Дисципліна*	Дата *	Оцінка
111111	Іванов І.І.	Мат.аналіз	10.01.11	Відмінно
111111	Іванов І.І.	Ін. мова	17.01.11	Добре
111113	Петров П.П.	Мат.аналіз	10.01.11	Добре
111113	Петров П.П.	Ін. мова	15.01.11	Відмінно
111115	Сидорова С.С.	Мат.аналіз	12.01.11	Відмінно
111115	Сидорова С.С.	Ін. мова	15.01.11	Задовільно

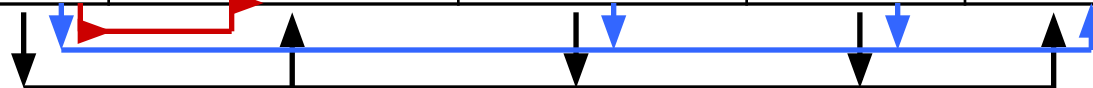


Повна функціональна залежність

Повна функціональна залежність. Атрибут В знаходиться в повній функціональній залежності від складного атрибута А, якщо він функціонально залежить від А та не залежить функціонально від будь-якої підмножини атрибута А.

Екзаменаційна ведомість

№ зал. кн.	ПІБ	Дисципліна	Дата	Оцінка
111111	Іванов І.І.	Мат.аналіз	10.01.11	Відмінно
111112	Іванов І.І.	Ін. мова	15.01.11	Добре
111113	Петров П.П.	Мат.аналіз	10.01.11	Добре
111113	Петров П.П.	Ін. мова	15.01.11	Відмінно
111115	Сидорова С.С.	Мат.аналіз	10.01.11	Відмінно
111115	Сидорова С.С.	Ін. мова	15.01.11	Задовільно

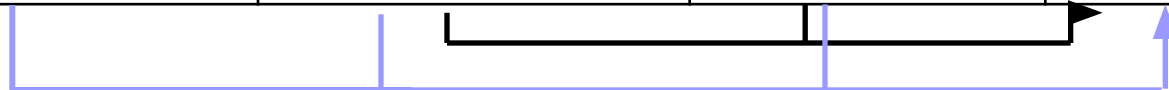


Повна функціональна залежність.

Приклад

Бухгалтерська ведомість

ПІБ	Посада	Наукове звання	Оклад
Іванов І.І.	Ст. викладач	К.т.н.	1500
Павлов П.П.	Доцент	К.т.н.	2000
Сидоров С.С.	Доцент	Д.т.н.	2500
Андрєєв А.А.	Ст. викладач	Без звання	1000
Олексєєв О.В.	Доцент	К.т.н.	2000
Романов Р.Р.	Ст. викладач	К.т.н.	1500
Ніколаєв Н.Н.	Доцент	Д.т.н.	2500
Сергєєв С.С.	Професор	Д.т.н.	3000
Павлов П.П.	Професор	Д.т.н.	3000
Євгєньєв В.В	Ст. викладач	Без звання	1000



Багатозначна залежність

Атрибут А **багатозначно** визначає атрибут В того ж самого відношення, якщо для кожного значення атрибута А існує добре визначена множина відповідних значень В.

$A \Rightarrow B$

Дисципліна	Викладач	Підручник
Інформатика	Шипілов П.А.	Форсайт Р. Паскаль для всіх
Інформатика	Шипілов П.А.	Уэйт М. и др. Язык Си
Інформатика	Голованевський Г.Л.	Форсайт Р. Паскаль для всіх
Інформатика	Голованевський Г.Л.	Уэйт М. и др. Язык Си
...

Дисципліна \Rightarrow Викладач

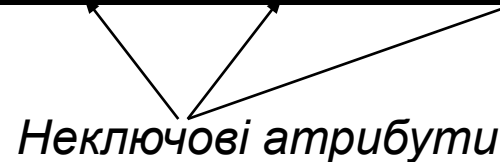
Дисципліна \Rightarrow Підручник

Неключові атрибути

Неключовим атрибутом є любий атрибут відношення, що не входить до складу первинного ключа.

<i>H_студ</i>	ПІБ	H_гр	Шифр_спец
1	Іванов	ПІ-10-1	06.0503
1	Іванов	ПІ-10-1	06.0503
2	Петров	КН-09-1	06.0401
3	Сидоров	ПІ-10-1	06.0503
3	Сидоров	ПІ-10-2	06.0503

Неключові атрибути

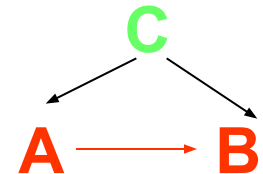


Два або більш атрибута **взаємно незалежні**, якщо ни один з цих атрибутів не є функціонально залежним від інших.

Транзитивна функціональна залежність

Функціонально залежний від А атрибут В **транзитивно залежить** від атрибута А (атрибути можуть бути складовими) в тому та тільки тому випадку, якщо існує такий атрибут С, що є функціональні залежності між атрибутами А та С, а також між В та С.

<i>Н_залік</i>	ПІБ	Н_гр	Шифр_спец
111111	Іванов	ПІ-10-1	06.0503
222222	Іщенко	ПІ-10-1	06.0503
333333	Петров	КН-09-1	06.0401
444444	Сидоров	ПІ-10-1	06.0503
555555	Степанов	ПІ-10-2	06.0503



Приклад визначення функціональних залежностей

Первинний ключ: Н_СПІВР,Н_ПРО

<i>* Н_СПІВР</i>	ПІБ	Н_ВІДД	ТЕЛ	<i>* Н_ПРО</i>	ПРОЕКТ	Н_ЗАВДАН
1	Іванов	1	11-22-33	1	Космос	1
1	Іванов	1	11-22-33	2	Климат	1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	33-22-11	1	Космос	3
3	Сидоров	2	33-22-11	2	Климат	2

*Залежності від
первинного ключа:*

- 1) {Н_СПІВР, Н_ПРО} -> ПІБ
- 2) {Н_СПІВР, Н_ПРО} -> Н_ВІД
- 3) {Н_СПІВР, Н_ПРО} -> ТЕЛ
- 4) {Н_СПІВР, Н_ПРО} -> ПРОЕКТ
- 5) {Н_СПІВР, Н_ПРО} -> Н_ЗАВД

Приклад визначення функціональних залежностей

Первинний ключ: Н_СПІВР,Н_ПРО

<i>Н_СПІВР</i>	ПІБ	<i>Н_ВІДД</i>	ТЕЛ	<i>Н_ПРО</i>	ПРОЕКТ	<i>Н_ЗАВДАН</i>
1	Іванов	1	11-22-33	1	Космос	1
1	Іванов	1	11-22-33	2	Климат	1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	33-22-11	1	Космос	3
3	Сидоров	2	33-22-11	2	Климат	2

Залежності від табельного номера співробітника:

- 6) *Н_СПІВР* -> ПІБ
- 7) *Н_СПІВР* -> *Н_ВІД*
- 8) *Н_СПІВР* -> ТЕЛ

Залежності від номера проекту:

- 9) *Н_ПРО* -> ПРОЕКТ

Залежності від номера відділу:

- 10) *Н_ВІД* -> ТЕЛ

Визначення транзитивних функціональних залежностей

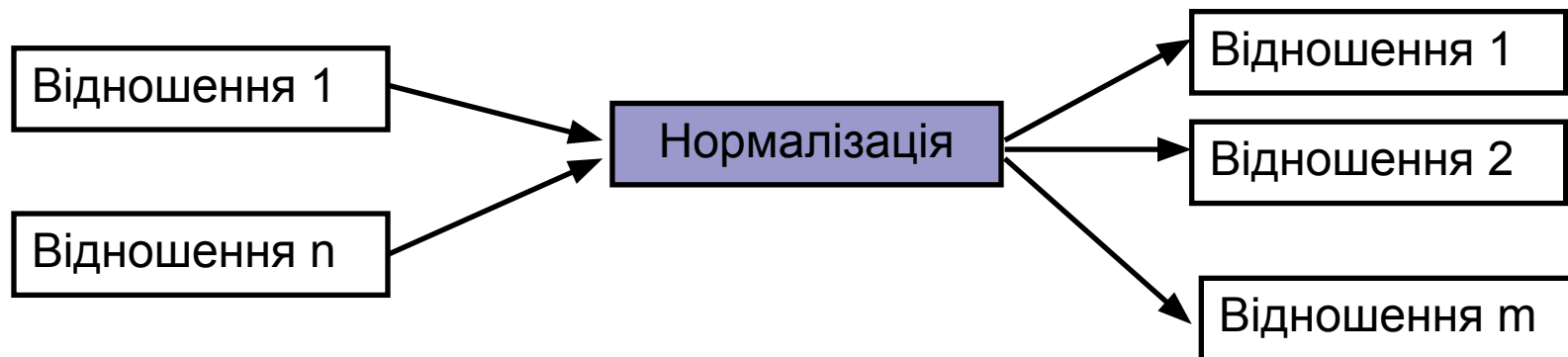
Первинний ключ: Н_СПІВР,Н_ПРО

<i>Н_СПІВР</i>	ПІБ	Н_ВІДД	ТЕЛ	<i>Н_ПРО</i>	ПРОЕКТ	Н_ЗАВДАН
1	Іванов	1	11-22-33	1	Космос	1
1	Іванов	1	11-22-33	2	Климат	1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	33-22-11	1	Космос	3
3	Сидоров	2	33-22-11	2	Климат	2

Нормалізація відношень

Нормалізація – це розділення таблиці на дві або більш, таких, що мають ліпші властивості під час включення, зміни та видалення даних.

Кінцева мета нормалізації – отримання такого проекту бази даних, в якому *кожен факт з'являється лише в одному місті*, тобто виключена надмірність інформації.



Основні нормальні форми

- перша нормальна форма (1NF);
 - друга нормальна форма (2NF);
 - третя нормальна форма (3NF);
 - нормальна форма Бойса-Кодда (BCNF);
 - четверта нормальна форма (4 NF);
 - п'ята нормальна форма, або нормальна форма проєкції-з'єднання)
- Нормальні форми проєкції-з'єднання*
- домено-ключова нормальна форма
 - шоста нормальна форма

Основні властивості нормальних форм:

- кожна наступна нормальна форма в деякому сенсі краще, ніж попередня;
- під час переходу до наступної нормальної форми властивості попередніх нормальних форм зберігаються.

1НФ (Перша Нормальна Форма)

Перша нормальна форма (1НФ) - це звичайне відношення.

Перша нормальна форма - це коли на перетині будь якого стовпця та будь якого запису знаходиться атомарне значення.

Відділи – Співробітники – Проекти – Завдання

Н_ВІД	ТЕЛ	Н СПІВ	ПІБ	Н_ПРО	ПРОЕКТ	Н ЗАВД
1	11-22-33	1	Іванов	1	Космос	1
				2	Климат	1
2	33-22-11	3	Сидоров	1	Космос	3
				2	Климат	2

Н_СПІВР	ПІБ	Н_ВІДД	ТЕЛ	Н_ПРО	ПРОЕКТ	Н_ЗАВД
1	Іванов	1	11-22-33	1	Космос	1
1	Іванов	1	11-22-33	2	Климат	1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	33-22-11	1	Космос	3
3	Сидоров	2	33-22-11	2	Климат	2

2НФ (Друга Нормальна Форма)

Відношення знаходиться у **другій нормальній формі (2NF)** в тому та тільки в тому випадку, коли воно знаходиться у 1NF, та кожен неключовий атрибут повністю функціонально залежить від первинного ключа відношення.

Відділи – Співробітники – Проекти – Завдання

<i>H_СПІВР*</i>	ПІБ	H_ВІД	ТЕЛ	<i>H_ПРО*</i>	ПРОЕКТ	H_ЗАВДАН
1	Іванов	1	11-22-33	1	Космос	1
1	Іванов	1	11-22-33	2	Климат	1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	33-22-11	1	Космос	3
3	Сидоров	2	33-22-11	2	Климат	2

2НФ (Друга Нормальна Форма)

Відділи – Співробітники – Проекти – Завдання

<i>H_СПІВР*</i>	ПІБ	H_ВІД	ТЕЛ	<i>H_ПРО*</i>	ПРОЕКТ	H_ЗАВДАН
1	Іванов	1	11-22-33	1	Космос	1
1	Іванов	1	11-22-33	2	Климат	1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	33-22-11	1	Космос	3
3	Сидоров	2	33-22-11	2	Климат	2

Функціональні залежності :

Залежності від первинного ключа:

- 1) {*H_СПІВР*, *H_ПРО*} -> ПІБ
- 2) {*H_СПІВР*, *H_ПРО*} -> H_ВІД
- 3) {*H_СПІВР*, *H_ПРО*} -> ТЕЛ
- 4) {*H_СПІВР*, *H_ПРО*} -> ПРОЕКТ
- 5) {*H_СПІВР*, *H_ПРО*} -> H_ЗАВД

Залежності від таб. номера співробітника:

- 6) *H_СПІВР* -> ПІБ
- 7) *H_СПІВР* -> H_ВІД
- 8) *H_СПІВР* -> ТЕЛ

Залежності від номера проекту:

- 9) *H_ПРО* -> ПРОЕКТ

Залежності від номера відділу:

- 10) H_ВІД -> ТЕЛ

Процедура приведення до 2НФ

Замінити $T(K1, K2, F1, F2)$, первинний ключ $(K1, K2)$, ФЗ $K1 \rightarrow F1$
 на $T1(K1, F1)$, первинний ключ $K1$,
 та $T2(K1, K2, F2)$, первинний ключ $(K1, K2)$.

Відділи – Співробітники – Проекти – Завдання

<i>H_СПІВР*</i>	ПІБ	H_ВІД	ТЕЛ	<i>H_ПРО*</i>	ПРОЕКТ	H_ЗАВДАН
1	Іванов	1	11-22-33	1	Космос	1
...
3	Сидоров	2	33-22-11	2	Климат	2

Відділи -Співробітники

<i>H_СПІВР</i>	ПІБ	H_ВІД	ТЕЛ
1	Іванов	1	11-22-33
2	Петров	1	11-22-33
3	Сидоров	2	33-22-11

Завдання

<i>H_СПІВР</i>	<i>H_ПРО</i>	H_ЗАВД
1	1	1
1	2	1
2	1	2
3	1	3
3	2	2

Проекти

<i>H_ПРО</i>	ПРОЕКТ
1	Космос
2	Климат

3НФ (Третя Нормальна Форма)

Відношення знаходиться в **третій нормальній формі (3NF)** в тому та тільки в тому випадку, якщо воно знаходиться у 2NF та між неключовими атрибутами не має транзитивних залежностей.

Процедура приведення до 3НФ:

Замінити $T(K, F1, F2)$, первинний ключ K , ТФЗ $F1 \rightarrow F2$ на $T1(F1, F2)$, первинний ключ $F1$, та $T2(K, F1)$, первинний ключ K .

Відділи-Співробітники

Н_СПІВР	ПІБ	Н_ВІД	ТЕЛ
1	Іванов	1	11-22-33
2	Петров	1	11-22-33
3	Сидоров	2	33-22-11

Залежності від номера відділу:
10) Н_ВІД -> ТЕЛ

Співробітники

Н_СПІВР	ПІБ	Н_ВІД
1	Іванов	1
2	Петров	1
3	Сидоров	2

Відділи

Н_ВІД	ТЕЛ
1	11-22-33
2	33-22-11

Приклад схеми БД в 3НФ

Відділи ●

<i>H_ВІД</i>	ТЕЛ
1	11-22-33
2	33-22-11

Проекти ●

<i>H_ПРО</i>	ПРОЕКТ
1	Космос
2	Климат

Співробітники ●

<i>H_СПІВР</i>	ПІБ	<i>H_ВІД</i>
1	Іванов	1
2	Петров	1
3	Сидоров	2

Завдання ●

<i>H_СПІВР</i>	<i>H_ПРО</i>	<i>H_ЗАВД</i>
1	1	1
1	2	1
2	1	2
3	1	3
3	2	2

1

M

1

M

M

1

Етапи нормалізації бази даних

Крок 1 (Приведення до 1НФ).

Задається одне або декілька відношень, що відображують поняття предметної галузі. На основі моделі предметної галузі (а не по зовнішньому вигляду отриманих відношень!) записуються функціональні залежності, що знайдені. Всі відношення автоматично знаходяться у 1НФ.

Крок 2 (Приведення до 2НФ).

Якщо в деяких відношеннях знайдена залежність атрибутів від частини складного ключа, то проводимо декомпозицію цих відношень на декілька відношень згідно до процедури приведення до 2НФ.

Крок 3 (Приведення до 3НФ).

Якщо в деяких відношеннях знайдена залежність деяких неключових атрибутів від інших неключових атрибутів, то проводимо декомпозицію цих відношень згідно до процедури приведення до 3НФ.

Ефективність нормалізації

Критерій	Відношення слабо нормалізовані (1НФ, 2НФ)	Відношення нормалізовані (3НФ)	Відношення сильно нормалізовані (4НФ, 5НФ)
Адекватність бази даних до предметної галузі	ГІРШЕ (-)	ЛІПШЕ (+)	ЛІПШЕ (+)
Легкість розробки та супроводження бази даних	СКЛАДНІШЕ (-)	ЛЕГШЕ (+)	СКЛАДНІШЕ (-)
Швидкість виконання вставки, оновлення, видалення	ПОВІЛЬНІШЕ (-)	ШВИДЧЕ (+)	ПОВІЛЬНІШЕ (-)
Швидкість виконання вибірки даних	ШВИДЧЕ (+)	ПОВІЛЬНІШЕ (-)	ПОВІЛЬНІШЕ (-)

Література

1. Codd E.F. A relational model of data for large shared data banks. Comm. ACM 1970, V. 13, №. 6, PP. 377-387
2. Codd E.F. Normalized data base structure: a brief tutorial. Proc. ACM-SIGFIDET. 1971, Workshop, San Diego, Calif., Nov. 1971, PP. 1-1.
3. Дейт К.Дж. Введение в системы баз данных. Пер. с англ. 8-е изд. К.: Диалектика, 2006.
4. Джексон Г. Проектирование реляционных баз данных для использования с микро-ЭВМ. Пер. с англ. М.: Мир, 1991
5. Дрибас В.П. Реляционные модели баз данных. Минск: БГУ, БССР, 1982.
6. Коннолли Т., Бегг К. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. М.: Издательский дом "Вильямс", 2003, 1440 с.
7. Мейер М. Теория реляционных баз данных. М.: Мир, 1987
8. Скворцов В.И., Щукин Б.А. Реляционная модель данных. М.: МИФИ, 1983, 92с.
9. Харрингтон Д.Л. Проектирование реляционных баз данных. Просто и доступно. М.: ЛОРИ, 2000
10. Цаленко М.Ш. Моделирование семантики в базах данных М.: Наука, 1989 М.: Наука, 1989.

Висновки

Кожна нормальна форма обмежує певний тип ФЗ і усуває аномалії обробки даних.

Властивості нормальних форм:

- 1НФ - всі атрибути відносини прості;
- 2НФ - відношення знаходиться в 1НФ і не містить часткових ФЗ;
- 3НФ - відношення знаходиться у 2НФ і не містить транзитивних ФЗ від ключа;
- НФБК - відношення знаходиться в 3НФ і не містить ФЗ ключів від неключових атрибутів;
- 4НФ, застосовується при наявності більш ніж однієї багатозначної ФЗ - відношення знаходиться в НФБК або 3НФ і не містить незалежних багатозначних ФЗ;
- 5НФ - відношення знаходиться в 4НФ і не містить ФЗ по з'єднанню.

Запитання та тестові завдання

1. Процесс нормализации это...
2. Первая нормальная форма (1НФ) это...
3. Таблица находится во второй нормальной форме (2НФ), если....
4. Отношение находится в третьей нормальной форме (3НФ) тогда и только тогда, когда...
5. Отношение автоматически находится во 2НФ, если...
6. Отношение автоматически находится в 3НФ, если....