Fundamentals/ICY: Databases 2013/14

Week 5: Monday

John Barnden
Professor of Artificial Intelligence

School of Computer Science
University of Birmingham, UK

Reminder
(with some re-ordering)

Attribute Determination.

- A collection of one or more attributes ***determines*** another attribute A if only one value for A is possible given the values for the former attributes.

  E.g., the collection DAY-NUMBER, MONTH and YEAR specifying birth-date in a table about people could determine DAY-NAME,

  even though it doesn't determine other attributes such as NATIONALITY: several people could have the same birth-date but be of different nationalities.

- We alternatively say that DAY-NAME is ***functionally dependent on*** DAY-NUMBER, MONTH and YEAR.

Attribute Determination: Special Case

- REMEMBER: ***Rows in a table are uniquely determined (picked out) by the values in some set of columns, i.e. the values of some collection of attributes.***

  That is, given some values for those attributes, there is at most one entity that has those values for those attributes, at any given time.

- Hence, that collection of attributes ***determines*** ALL the other attributes.
- That is, given some values for the determining attributes, there's at most one value for each of the other attributes, at any given time.

New

Attribute Determination, contd.

- The determining collection of attributes is called the ***determinant*** in the determination.

- Write determination as:

  DAY-NUMBER, MONTH, YEAR $\rightarrow$ DAY-NAME

- Determination does ***not*** mean that you have at hand an ***algorithm*** for working out a dependent attribute from the determinant, although you may do.

  E.g. Consider a "father" attribute.

Keys

- ((((A ***key*** for a table is, in general, just a determinant: a collection of one or more attributes that determines at least ***some other*** attribute(s) in that same table.))))
  - **CAUTION:** But often people use "key" as a sloppy abbreviation for "primary key" (see below) or other notions.
- A ***superkey*** for a table is a collection of one or more attributes that determines ***all the other*** attributes in the table, *i.e.* determines a whole row.
  - Trivially, the collection of **all** the attributes is a superkey.
- A ***candidate key*** is a **minimal superkey** (i.e., you can't remove attributes from it and still have a superkey.)

  It does **NOT** necessarily mean a **numerically smallest** superkey.

Superkeys & Candidate Keys: Example

- Suppose a "day" entity type has attributes DAY-NAME, DAY-NUMBER, MONTH, YEAR, IS-UK-HOLIDAY, IS-US-HOLIDAY, …
- Then DAY-NAME, DAY-NUMBER, MONTH, YEAR would be a **superkey** for the day type.
- But it's **not a candidate key** because DAY-NUMBER, MONTH, YEAR is also a superkey.
- This smaller collection **is** a candidate key because no sub-collection of it uniquely identifies a day.

Primary Keys

- A ***primary key*** for a table (entity type) is a **candidate** key that **the DB designer has chosen** as being the main way of uniquely identifying a row (entity). Extra restriction: Its attributes are not allowed to have null values.
  - **It could be that there's only one candidate key in practice anyway, such as a person's ID number.**
- Primary keys are the main way of identifying target entities in entity relationships, *e.g.,* the way to identify someone's employing organization.
- For efficiency (and correctness) reasons, the simpler that primary keys are, the better.
  - **Identity numbers (of people, companies, products, courses, etc.), or combinations of them with one or two other attributes, are the typical primary keys in examples in the textbook and handouts.**

Relationships and Foreign Keys

- Standardly, a relationship is represented by means of **foreign keys.**

- A foreign key in a table **T** is a chosen collection of attributes intended to match the attributes that constitute (usually) the primary key in ***another*** table, **U**, and thereby to refer to entities in **U**.

  Intuitively, the foreign key in **T** is **U**'s "ambassador" [my word] in **T**.

Primary & Foreign Keys

PERS-ID	**NAME**	***PHONE***	***EMPL. ID***	**AGE**
9568876	Chopples	**0121-414-3816**	E22561	37
2544799	Blurp	**01600-719975**	E85704	21
1698674	Rumpel	**07970-852657**	E22561	88

Primary keys are underlined

PHONE	**TYPE**	**STATUS**
0121-414-3816	office	OK
01600-719975	home	FAULT
0121-440-5677	home	OK
07970-852657	mobile	UNPAID

Foreign keys are in italics

EMPL. ID	**EMPL. NAME**	**ADDRESS**	**NUM. EMPLS**	**SECTOR**
E48693	BT	BT House, London, …	1,234,5678	Private TCOM
E85704	Monmouth School for Girls	Hereford Rd, Monmouth, …	245	Private 2E
E22561	University of Birmingham	Edgbaston Park Rd, ….	4023	Public HE

Composite Primary and Foreign Keys

People

PERS-ID	**NAME**	***AREA CODE***	***PHONE BODY***	***EMPL ID***	**AGE**
9568876	Chopples	**0121**	**414-3816**	E22561	37
2544799	Blurp	**01600**	**719975**	E85704	21
1698674	Rumpel	**07970**	**852657**	E22561	88

Phones

AREA CODE	**PHONE BODY**	**TYPE**	**STATUS**
0121	**414-3816**	office	OK
01600	**719975**	home	FAULT
0121	**440-5677**	home	OK
07970	**852657**	mobile	UNPAID

1:1 Connectivity between Tables

People

PERS-ID	**NAME**	***PHONE***	***EMPL ID***	**AGE**
9568876	Chopples	**0121-414-3816**	E22561	37
2544799	Blurp	**01600-719975**	E85704	21
1698674	Rumpel	**07970-852657**	E22561	88
5099235	Biggles		E22561	29

1:1: that is, no more than one phone allowed per person, and vice versa.

Phones

PHONE	**TYPE**	**STATUS**
0121-414-3816	office	OK
01600-719975	home	FAULT
0121-440-5677	home	OK
07970-852657	mobile	UNPAID

Note: the representation is still ***asymmetric*** in that the People table mentions phones but not vice versa – ***symmetry would create extra redundancy***.

NB: Biggles has no phone listed, and 0121-440-5677 has no person recorded. Suggests a possible reason for not combining such tables.

1:M Connectivity between Tables

People

PERS-ID	**NAME**	***PHONE***	***EMPL ID***	**AGE**
9568876	Chopples	**0121-414-3816**	**E22561**	37
2544799	Blurp	**01600-719975**	E85704	21
1698674	Rumpel	**07970-852657**	**E22561**	88
1800748	Dunston	**0121-414-3886**	**E22561**	29

More than one employee allowed per organization, but no more than one employer per person.

NOTE direction of use of the foreign key. Why so??

Organizations

EMPL ID	**EMPL NAME**	**ADDRESS**	**NUM EMPLS**	**SECTOR**
E48693	BT	BT House, London, …	1,234,5678	Private TCOM
E85704	Monmouth School	Hereford Rd, Monmouth, …	245	Private 2E
E22561	University of Birmingham	Edgbaston Park Rd, ….	3023	Public HE