technische universität
dortmund

MAGIC
Major Atmospheric
Gamma Imaging
Cerenkov Telescopes

Database Knowledge and the autoMAGIC Database

autoMAGIC Workshop Padova

Jan Lukas Schubert

28th - 30th August 2024

Contents:

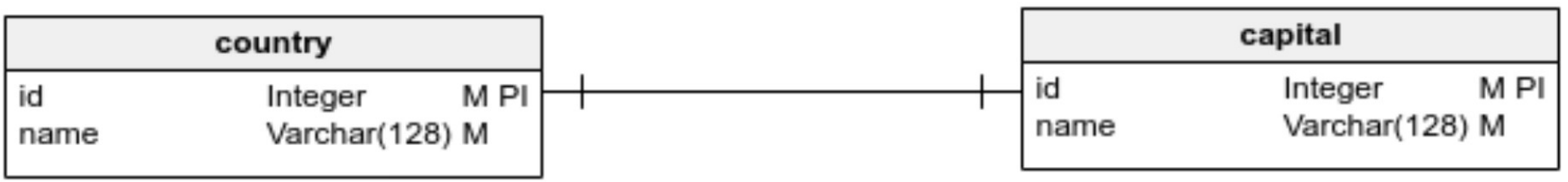
- Database: Basic concept
- Database relationships
- The autoMAGIC database structure:
  - Organization of the Data
  - Organization of the MCs
  - Organization of the Jobs
  - Examples for relationships
- Basic and advanced database commands

Databases – Basic Concept

- Tables storing information
  - Columns: parameters
  - Rows: Entries
- Each row is assigned a unique id
- Tables are linked to each other with relationships

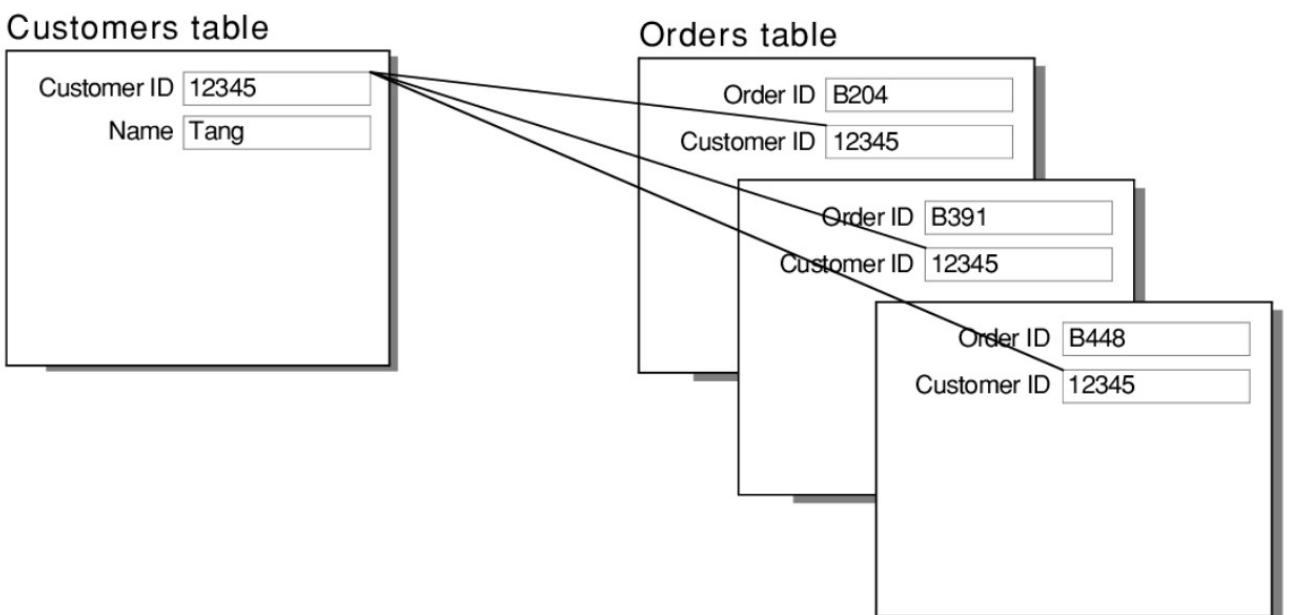
Databases - Relationships

- One-to-One relationship: One object of a DB table can only be linked to one object of another table (uncommon in practice!)



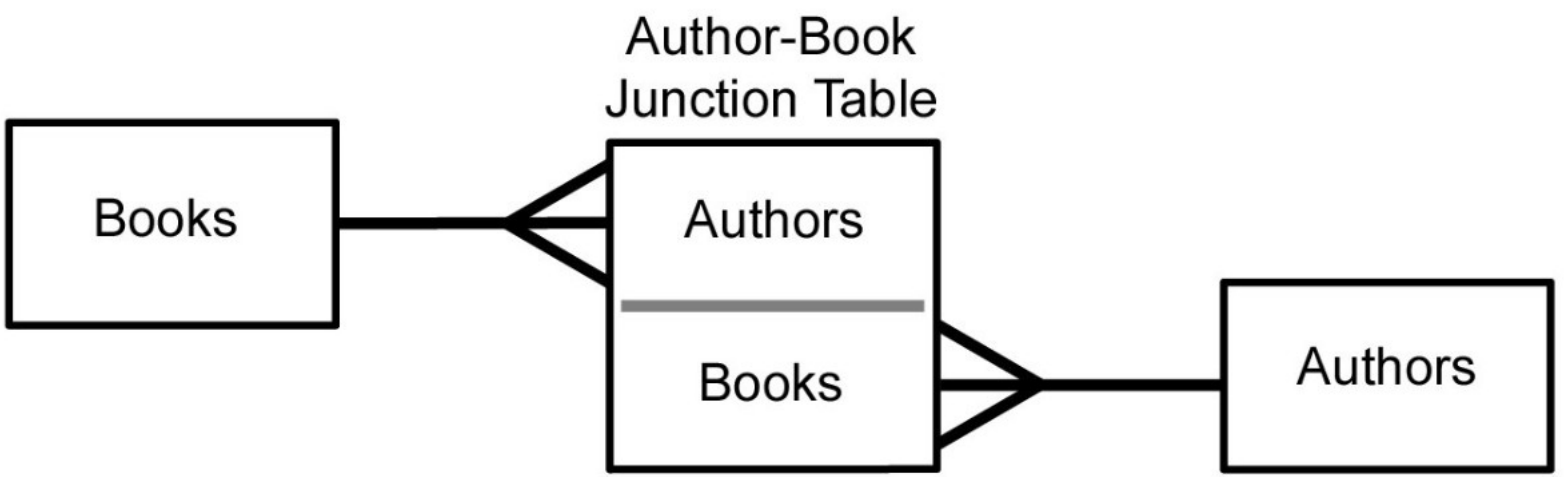
Databases - Relationships

- One-to-Many relationship: One object can be linked to many objects of another table



Databases - Relationships

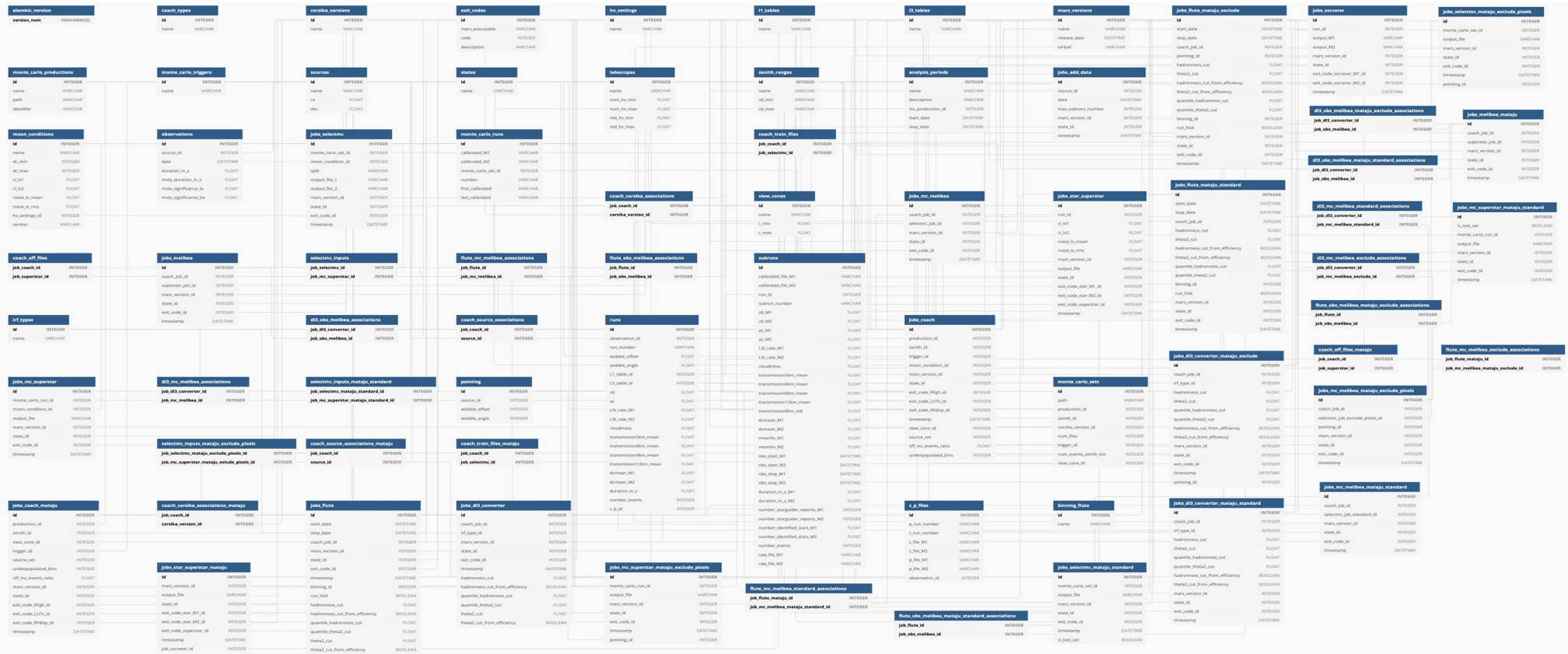
- Many-to-Many relationship: Many objects of one table can be linked to many objects of another table



Databases - Unique Constraints

- DBs tables can use unique constraints to prevent doubling of the entries
- Example: superstar jobs
  - Unique combination of mars_version and run_number

The autoMAGIC database Structure



The autoMAGIC database Structure

- Many tables are just for preventing writing too much redundant information into the DB

- Example: job states
  - Writing the string “created”, “queued”, etc. into the corresponding DB column is possible, but redundant
  - Better: just link the objects in the states table by their id
  - Instead of writing redundant informations, we just store one integer in the column and by the DB structure (relationships!), the information can be retrieved

The autoMAGIC Database Structure

Table Observation

id	source	date	significance
1	Crab Nebula	2020-05-30	15
2	NGC1275	2019-01-31	0.5
3	Crab Nebula	2011-07-04	8
4	NGC1275	2016-08-11	1.2
5	Perseus-MA	2017-06-23	2.9

Table Run

id	run number	zenith	azimuth	transmission	DC	observation id
1	500001	25	118	0.85	1244	1
2	500002	26	117	0.85	1240	1
3	500003	27	116	0.83	1155	1
4	541241	48	50	0.65	508	4
5	555621	7	152	0.92	256	3

Table SubRun

id	subrun number	calibrated file M1	calibrated file M2	run id
1	004	500001.004_M1.root	500001.004_M2.root	1
2	051	500001.057_M1.root	500001.057_M2.root	1
3	006	500001.006_M1.root	500001.006_M2.root	1
4	122	524528.122_M1.root	524528.122_M2.root	7
5	089	597952.098_M1.root	597952.098_M2.root	9

The autoMAGIC Database Structure

Table Analysis Period

id	name	start	stop	production id
1	ST0314	2020-05-14	2020-09-14	1
2	ST0313	2020-02-26	2020-03-13	2
3	ST0307	2017-11-10	2018-06-29	3
4	ST0307	2016-04-29	2017-08-02	3
5	ST0306	2014-11-24	2016-04-28	5

Table Monte Carlo Production

id	name	path	identifier
1	ST0314	/pnfs/path/to/ST0314	M1_AD6.0_MF0.53_M2_AD4.5_MF0.65
2	ST0313	/pnfs/path/to/ST0313	M1_AD6.0_MF0.53_M2_AD4.5_MF0.65
3	ST0307	/pnfs/path/to/ST0307	M1_AD3.5_MF0.68_M2_AD5.0_MF0.71
4	ST0310	/pnfs/path/to/ST0310	M1_AD3.5_MF0.612_M2_AD5.0_MF0.675
5	ST0306	/pnfs/path/to/ST0306	M1_AD5.5_MF0.63_M2_AD5.5_MF0.69

Table Monte Carlo Set

id	trigger	zenith	corsika	view cone	number files	production id
1	standard	low	mmcs699	diffuse 1.5	5000	1
2	standard	medium	mmcs699	diffuse 1.5	5000	1
3	standard	high	mmcs699	diffuse 1.5	5000	1
4	standard	low	mmcs6500	ringwobble	3999	5
5	sumT	medium	mmcs699	ringwobble	4000	3
6	EGAL sumT	high	mmcs699	diffuse 2.5	4999	6

Table Monte Carlo Run

id	number	calibrated file M1	calibrated file M2	number events	MC set id
1	854654	GA_854654_M1.root	GA_854654_M2.root	25417	1
2	875429	GA_875429_M1.root	GA_875429_M2.root	14554	1
3	845701	GA_845701_M1.root	GA_845701_M2.root	42542	1
4	831477	GA_831477_M1.root	GA_831477_M2.root	15497	7
5	842215	GA_842215_M1.root	GA_842215_M2.root	25344	9

The autoMAGIC Database Structure

Table Job Coach

id	MC prod.	moon condition	zenith	view cone	exit code id	state id
1	ST0307	no moon	low	ringwobble	1	5
2	ST0312	no moon	low	ringwobble	1	5
3	ST0312	moderate moon	medium	diffuse1.5	1	5

Table Job Superstar

id	run	exit code id	state id
1	576432	1	5
2	552427	1	5
3	589420	2	6

Table Job Melibea

id	coach job id	superstar job id	exit code id	state id
1	3	1	1	5
2	2	3	1	5
3	1	4		1
4	3	2		4
5	3	5	4	6

Table Exit Code

id	exit code	description
1	0	everything fine
2	6	Starguider issue
3	254	input files not found

Table State

id	name
1	created
2	queued
3	idle
4	running
5	success
6	error
7	stopped

Database Commands

- Start your DB session with `sqlite3 /path/to/your/db.sqlite`
- Sometimes not entering the sqlite shell gives a performance boost:
  - `sqlite3 /path/to/your/db.sqlite 'SELECT * FROM table_name'`
- You can configure your DB to be more convenient with:
  - `.header on`
  - `.mode column`
    - Can also be written into a .sqliterc file as general setup
- You can display all available table names with `.tables`

Basic Database Queries

- Query the whole moon conditions table:
  - `SELECT * FROM moon_conditions`
- Query only certain columns from moon conditions table:
  - `SELECT id FROM moon_conditions`
  - `SELECT id, dc_min FROM moon_conditions`
- Query under certain condition:
  - `SELECT * FROM moon_conditions WHERE dc_min<2200`

Advanced Database Commands

- Joining Tables
  - ```
    SELECT * FROM runs
        JOIN observations ON observation_id=observations.id
        JOIN sources ON source_id=sources.id
    WHERE name='NGC1275';
    ```

- Updating the Database
  - `UPDATE jobs_melibea SET state_id=1 WHERE state_id=4`

- Deleting entries
  - ALWAYS use with the WHERE command!
  - `DELETE FROM jobs_add_data_superstar_pic WHERE id=1`
    ```

Explore the Database

- Have a look into the organization of observations and runs:
  - Table names: observations, runs
- Have a look into the organization of the MCs:
  - Table names: analysis_periods, mc_productions, mc_sets, mc_runs
- Have a look into the job tables (most of them will still be empty!):
  - jobs_add_data_superstar_pic, jobs_star_superstar, jobs_coach, jobs_melibea, jobs_dl3_converter