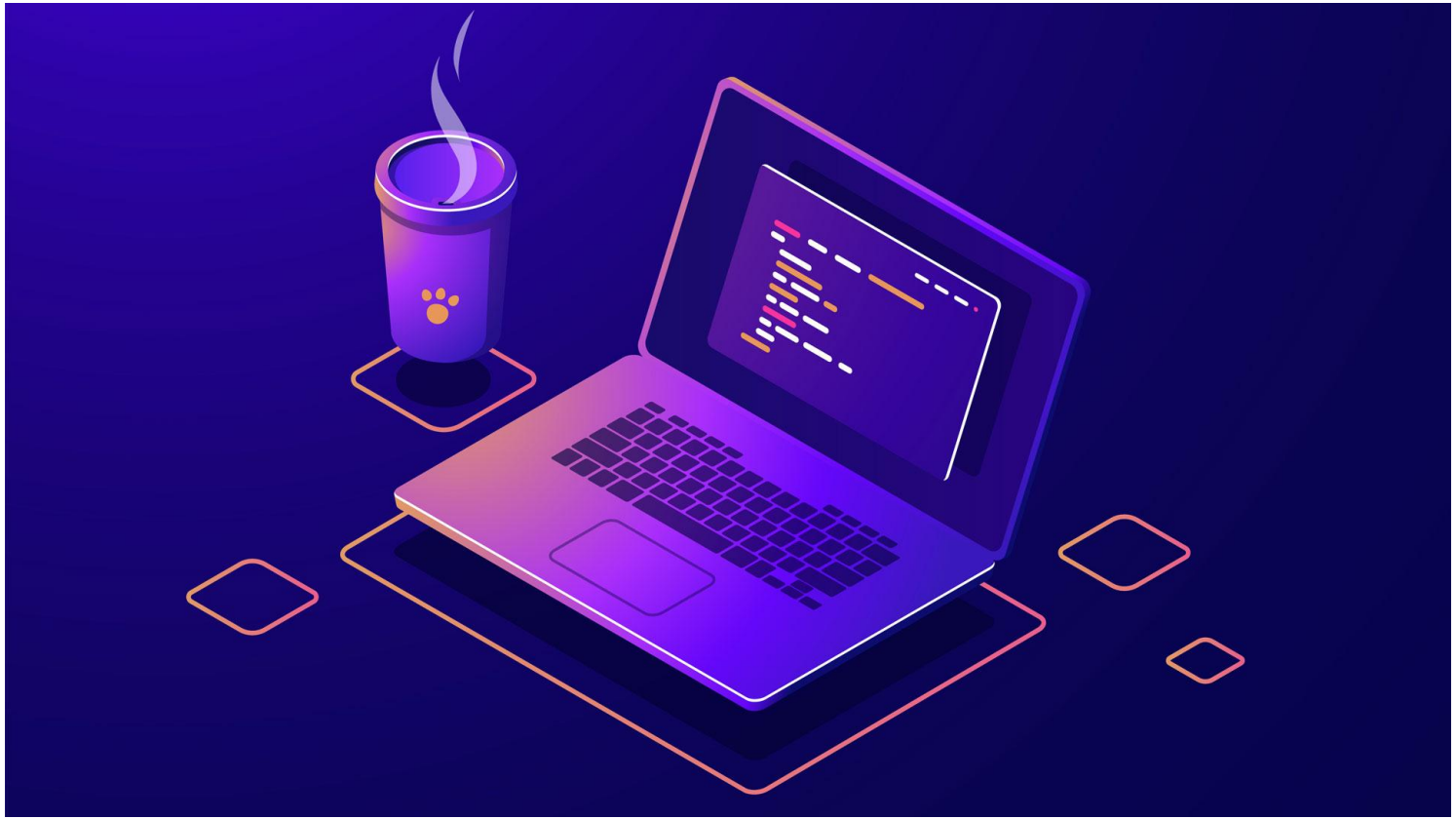
Tasks and Dependencies



Motivation

Consider now that we want to execute the following query:

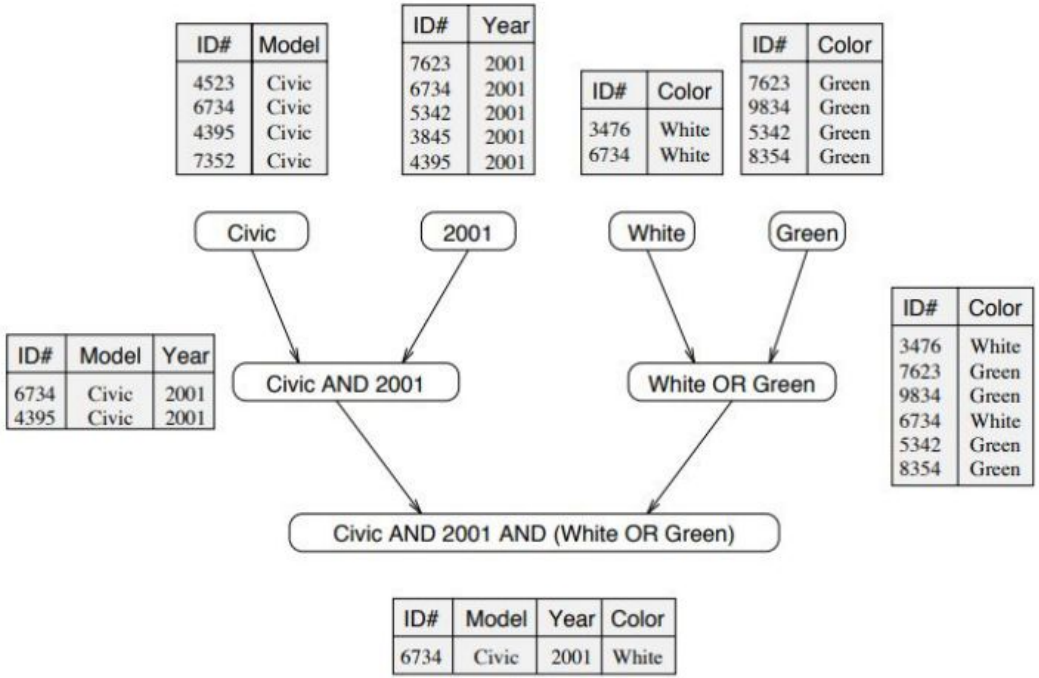
```
MODEL = ‘‘CIVIC’’ AND YEAR = 2001 AND
(COLOR = ‘‘GREEN’’ OR COLOR = ‘‘WHITE)
```

on our car dealer database:

ID#	Model	Year	Color	Dealer	Price
4523	Civic	2002	Blue	MN	$18,000
3476	Corolla	1999	White	IL	$15,000
7623	Camry	2001	Green	NY	$21,000
9834	Prius	2001	Green	CA	$18,000
6734	Civic	2001	White	OR	$17,000
5342	Altima	2001	Green	FL	$19,000
3845	Maxima	2001	Blue	NY	$22,000
8354	Accord	2000	Green	VT	$18,000
4395	Civic	2001	Red	CA	$17,000
7352	Civic	2002	Red	WA	$18,000

Tasks

- A possible query plan could be:



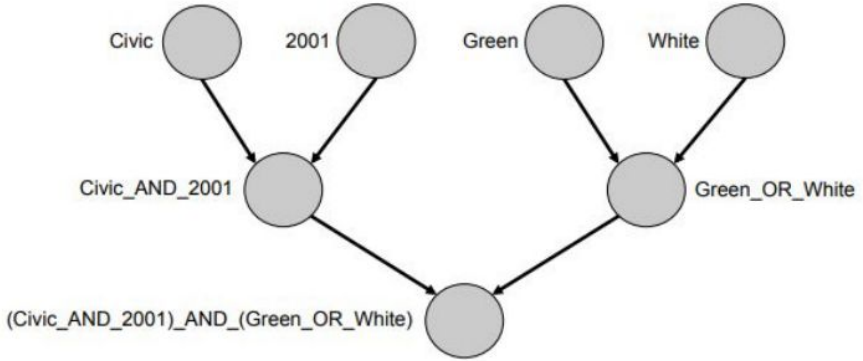
Tasks and Dependencies

- Each of these operations in the query plan could be a task, each computing an intermediate table of entries that satisfy particular conditions

- Are they independent?
  - Some of them are, for example tasks *"Civic"*, *"2001"*, *"Green"* and *"White"*
  - Others are not independent. For example, task *"Civic_AND_2001"* can not start its execution until both tasks *"Civic"* and *"2001"* complete

Tasks and Dependencies

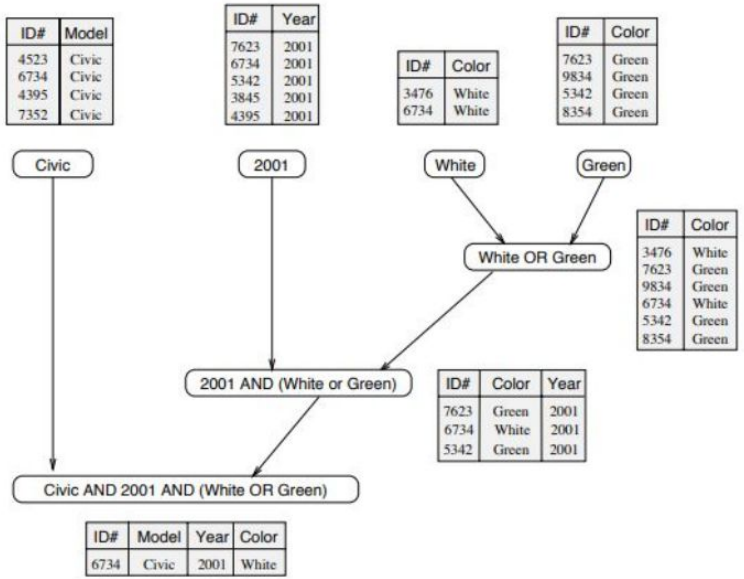
- Dependences impose task execution ordering constraints that need to be fulfilled in order to guarantee correct results



- **Task dependence graph**: graphical representation of the task decomposition

Tasks and Dependencies

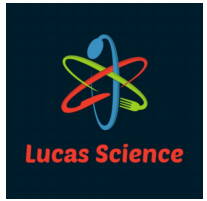
- Other query plans are possible, for example



... with different task dependence graphs and potential to execute tasks in parallel

Instructor Social Media

Youtube: Lucas Science



Instagram: lucaasbazilio

Twitter: lucasebazilio