**CS522 Advanced Database Systems**
Introduction to DBMS

Chengyu Sun
California State University, Los Angeles

---

**Database Management Systems (DBMS)**

- Query
- Results

**DBMS**
- Data

---

**Terminology**
- Database
- Database Management System (DBMS), Database System
- Relation, table
- Tuple, record, row
- Attribute, field
- Column

---

**DBMS Implementation**

Introducing the **MEGATRON 3000**
Database Management System
- Latest relational database technology!
- Unix compatible

---

**Megatron 3000 Implementation – Relations**
- Relations are stored in text files
  - file name is the same as the relation name
  - one tuple per line
  - fields are separated by #
- e.g., /usr/db/students

<table>
<thead>
<tr>
<th>Smith#123#CS Jones#522#EE</th>
</tr>
</thead>
</table>

---

**Megatron 3000 Implementation – Schemas**
- Schemas are stored in a catalog
- file name is the same as the relation name
  - one schema per line
  - "relation#attribute#type#attribute#type#...
- e.g., /usr/db/catalog

<table>
<thead>
<tr>
<th>students#name#STR#id#INT#dept#STR</th>
</tr>
</thead>
<tbody>
<tr>
<td>depts#name#STR#office#STR</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

---
Megatron 3000 – Selection

SELECT * FROM R WHERE condition; 
◆ Read the catalog to get the attributes of R 
◆ Read file R, and for each line 
  ■ Check condition 
  ■ Output tuple if the condition is true

Megatron 3000 – Join

SELECT * FROM R, S WHERE condition; 
◆ Read the catalog to get the attributes of R and S 
◆ Read file R, and for each line 
  ■ Read file S, and for each 
    ■ create join tuple 
    ■ check condition 
    ■ Output tuple if the condition is true

What’s Wrong with Megatron 3000?

??

A Real DBMS

Class Goals

◆ What this class is about 
  ■ All aspects of DBMS implementation 
  ■ Established algorithms 
◆ What this class is not 
  ■ Distributed systems (??) 
  ■ New research and development (CS594)