# Database Systems, CSCI 4380-01
# Homework # 1
# Due Thursday February 4, 2016 at 11:59:59 PM

**Homework Statement.** This homework aims to teach you how to construct complex queries using relational algebra. Please do the parts in sequence, as they build on each other. Suppose you are given the following database that tracks all UFO related events reported by different users of a database[1].

```
Users(userid, email, name)
Sightings(eventid, userid, status, date, time, state, city, country, description,
          shape, color, size, numpeople, visualangle, fromloc, toloc, casereview)
Abductions(eventid, userid, city, state, country, date, numpeople, description,
          casereview)
CropCircles(eventid, userid, city, state, country, picture, radius, date)
Behaviors(eventid, behavior)
```

*Note example behaviors: hovered, orb-like, changed speed, discontinuous trajectory*

```
Tags(eventid, tag)
```

*Note example tags: eurowheel, bird, near railway, affected electricity*

Note: All date fields are formatted as `mon-day-year`, e.g. `01-31-2016`. You can assume that you can check if a date time `X` comes after another date time `Y` by checking whether `X` $>$ `Y`. Time is formatted similarly, e.g. `08:00:00` and can be compared using $>$ as date.

Each relation stores a UFO based event (sightings, abductions and cropcircles) reported by different users. An event with the same eventid can be of many types, for example both a UFO sighting and an abduction, or a crop circle. The remaining attributes will be explained as they come in specific queries.

Write the following queries using relational algebra:

**Part 1.** The following queries only need a single SELECT ($\sigma$), followed by a PROJECT ($\pi$) and RENAMING ($\rho$) as necessary:

(a) Return the countries in which there was a UFO sighting in 2015 with a shape "equilateral triangle", color "black" involving 2 or more people.

(b) Return the city and state in which a crop circle is seen in "United States" since 2010 (inclusive) or a radius of at least 10.

(c) Return the id of all events involving a behavior of "hovering".

**Part 2.** The following queries combine SELECT ($\sigma$), SET operations ($\cap, \cup, -$), PROJECTION ($\pi$) and RENAMING ($\rho$) as necessary:

[1]Truth is out there! In honor of the X-Files reboot. The database is loosely based on the MUFON database of "actual" UFO sightings. Of course, our queries will be much more advanced than theirs.

(a) Return the id of all users who have reported either a sighting or an abduction.

(b) Return all countries that have had sightings but no abductions and no crop circles.

(c) Return the id all events involving both an abduction and a crop circle in the same state and country.

(d) Return the id of all events involving an abduction and a crop circle occurring in different locations (i.e. not the same state and country). They could be in the same country and different states or even different countries.

**Part 3.** The following queries combine SELECT ($\sigma$) statements with a JOIN ($\bowtie$) (or CARTESIAN PRODUCT), followed by a PROJECT ($\pi$) and RENAMING ($\rho$) as necessary:

(a) Return the name of all users who have reported a sighting in "Lake Okobogee, Iowa, United States" (city, state and country respectively).

(b) Return the name of all users who have reported a crop circle in 2015 with tag "bird".

(c) Return the name of all users who have reported a sighting and an abduction on the same date.

**Part 4.** Freeform, you decide which combination is needed. Any relational algebra operator is fine.

(a) Find sighting events with tags "bird" or "eurowheel", involving a "round" or "oval" white object with "hovered" behavior. Return the state, city, country and description of these events.

(b) Find pair of events (with different ids) involving sightings in the same state, city, country, date, involving at least 2 people, occurring before "17:00" that are reported by at least two different users. Return the pair of event ids.

(c) **Bonus/optional.** Find pairs of events e1, e2 such that e2 has all the behaviors reported for e1 (in the Behaviors relation). Return the event ids for e1 and e2.

Note: If you really want to challenge yourself, try also restricting it to only the sighting events and require e1 and e2 to have exactly the same set of behaviors. This part is not part of the homework, but feel free to show me your solutions for discussion.

**SUBMISSION INSTRUCTIONS.** Submit a PDF or text document for this homework using Homework Submission Server. No other format, especially no WORD documents or hardcopy please. The penalty for late submissions in effect, make yourself familiar with the course policies.

The course website for homework submissions will become available midweek next week. We will announce it on Piazza.