Homework 1: E/R Modeling (100 points)

Due Date: **Friday, Jan 20 (5:00 PM)**

Submission

All HW assignments should contain both your student ID and your name and must be submitted online, as a PDF file, through their associated HW1 EEE dropbox. See the table below for HW 1 submission opportunities. Note that after Sunday the 22nd, no further submissions will be accepted. That is, we will not accept assignments after that time (since we will be publishing the solutions at that time). Please get all your work in on time! If possible, save your one dropped assignment for the end of the term when you are most likely to want/need it.

Date / Time	Grade Implications
Friday, Jan 20 (5:00 PM)	Full credit will be available
Saturday, Jan 21 (5:00PM)	20 points will be deducted
Sunday, Jan 22 (5:00 PM)	40 points will be deducted

E-R Schema Design [100 pts]

After studying Entrepreneurship at UCI's business school, and inspired by the almost frightening current prominence of Twitter and other social media outlets, four students at UCI have decided to launch a startup company called "TopicalBirds" – and because you are taking CS122a, they have invited you to join the company as their DBA. TopicalBirds.com is going to be a new social media service that allows users ("birds") to share whatever "fowl thoughts" happen to be on their mind by simply "chirping" about them. As we all know, people today are keen on knowing what others are thinking and saying, so birds will be able to listen for the chirps of other birds. The startup service's name stems from the fact that chirps will be topical – i.e., each chirp will include a set of topics that characterize the chirp's subject(s) – and users of the service will also to be able to base their chirp listening on topics. TopicalBirds.com will allow birds to further share chirps that they hear by "parroting" them, and users will be classified as being either just regular birds or "peacocks". (More about that distinction shortly.)

Being business savvy, the founders already have a plausible business model. Much insight can be gleaned by paying attention to what people are saying about things, so companies will be able to sign up on TopicalBirds.com as "bird watchers" (for a fee, of course) in order to gain valuable business feedback and identify newly emerging trends in terms of what's being chirped about. Bird watching will enable a company to identify potential new customers, to target them with topically relevant advertisements, and so on. Again, the inclusion of the notion of topics will

make it easier for bird watchers to assure relevance, and the service's regular vs. peacock user classification will help bird watchers to identify potentially influential users.

Your first major milestone is to setup a database to support TopicalBirds' online service. The first part of your work will be to capture the founders' design requirements, in support of their new service idea and business model, producing an E-R diagram representing a suitable conceptual schema for the TopicalBirds.com database.

The founders have given you the following initial list of requirements:

1. Each bird will have a unique ID ("tag") assigned by TopicalBirds.com. They will also need to specify a current e-mail address and choose a password for the service (which will be stored only in encrypted form). Their initial signup date will be recorded as well.
2. Each bird will be asked to provide their full (first and last) name, current address (with a number, street, city, state, country, and mailcode), gender, date of birth, and optionally their annual gross income.
3. To support the service's user and business models, each bird will be asked to specify one or more topics that they might be interested in (i.e., topics they expect to be chirping about or listening for) and what their interest level is (on a low-to-high scale of 1 to 5).
4. In addition to the above, each bird who is designated as being a peacock will have a specified "variety" attribute that holds one or more values such as politician, musician, actor, professor, or pastor (and so on). A peacock may be "multi-variety", e.g., some particular peacock could be both a musician and an actor.
5. The universe of chirp-worthy topics currently recognized by TopicalBirds.com should be configurable. Examples of such topics might include politics, health, music, movies, travel, weather, or sports.
6. Each topic will have a unique topic id, an associated topic name, and a brief description.
7. Each chirp uttered by a bird will have a user-relative chirp number (1, 2, 3, ...) so that, e.g., chirp number 3 from bird 0012345 is the third chirp uttered by bird 0012345. Each bird may chirp as often as they wish, as there is no maximum chirp rate or limit.
8. Each chirp uttered by a bird will include the chirp's content (text), the date and the time when it was uttered, the location (latitude/longitude) where it was uttered, and one or more topics that its textual content is related to (i.e., about). Each chirp will also have an associated sentiment, ranging from +1.0 (very positive) to -1.0 (very negative), which will be derived from the chirp's content (using machine learning, of course).
9. If a chirp is parroted, we must also be able to identify the chirp that it parrots.
10. Bird watchers are another kind of user, somewhat like birds or peacocks, but they don't have a full name, a gender, a birth date, an income, or a variety. Instead of these, being companies, each bird watcher has a business name, a business sector (e.g., financial, healthcare, government, education, etc.), a founding date, and an annual membership fee. Like all other kinds of users, they have a tag and they specify their topical interests.
11. Any bird can listen for chirps *specifically*, i.e., for chirps coming from a set of specified

other birds (regular birds and/or peacocks).

12. Alternatively, a bird can listen for chirps *topically*, i.e., for any chirps coming from birds or peacocks about a specified topic of interest.
13. Similarly, a bird watcher can listen for chirps topically and/or specifically.
14. The advertisements that each bird watcher may want to show to selected birds are also stored in the database; each bird watcher has its own distinct set of ads.
15. Each ad will have an integer ad id, a textual caption, and an associated picture.

Design an E-R schema to represent the required information and express your design in the form of an E-R diagram. (Please use the E-R notation from either the lectures or the textbook; do not invent your own notation or use UML. You can also find a decent E-R model tutorial on the web at http://www.tutorialspoint.com/dbms/er_model_basic_concepts.htm, but beware of its different ISA notation.) Be sure that your E-R design captures all of the implications of the user and business models, including the aspects below ((a)-(d)), and be sure to make good use of advanced E-R modeling goodies such as weak entities, ISA relationships, composite attributes, set-valued attributes, etc., where appropriate. Please use the book's ISA notation and be sure to use its overlap and covering constraints. *NOTE:* Be sure **NOT** to design your E-R schema by just making a single pass over the bullets above! You will need to carefully analyze all of the requirements in order to identify available attribute and relationship commonalities for ISA purposes; a one-pass design would miss important (inheritance-based) sharing opportunities and lead to a much messier-than-necessary design. Cleanliness matters!

(a) [50pts] All of the relevant entities and their attributes, including their keys. (For Entities, you may **ONLY** use "User", "Bird", "Peacock", "BirdWatcher", "Topic", "Chirp", and "Ad". Do not create more entities or use fewer entities than these.)
(b) [20pts] All of the relevant relationships and any associated attributes.
(c) [15pts] Appropriate cardinality (a.k.a. key) constraints for the relationships.
(d) [15pts] Appropriate participation constraints for the relationships.

Be sure to download the HW#1 PDF schema template from the Attachments area of the CS122a web page and use that as the basis for drawing the E-R schema that you turn in. (Your solution will not be accepted if you do not use our layout for your diagram! Without this standardization it becomes too hard to grade everyone's answers thoroughly.)