Technologies for

Information Systems

prof. F. A. Schreiber, prof. L. Tanca – Feb. 23, 2010

Available Time 1h 45m

Last Name _____

First Name _____

Student ID _____ **Signature** _____

Merge is an online wedding agency that stores detailed information about persons' habits and preferences so as to help people find the perfect spouse. As an internal policy, Merge considers a match possible only if the age difference is at most 10 years.
Patty is a popular dating website with lightweight information regarding millions of customers.
AnemOne is a new company that now owns both Merge and Patty and integrates their services. A specific design requirement for the integrated schema posed by AnemOne's CTO is that there needs to be an explicit relationship for possible matches between persons. In particular, a match is considered possible if age requirements are met and all mandatory habit preferences (wherever available) are satisfied. We report the original relational schemas.

Merge:

person(pid, email, name, birthdate, isMale) // personal data
picture(pid, picture, isProfilePic) // pictures associated with a person
features(fid, description) // habits and features like "smoker", "self confident" etc.
hasFeature(pid, fid) // a person has the given feature
prefersFeature(pid, fid, isMandatory) // a person wants the potential partner to have the
// given feature (mandatorily if the flag is checked)

Patty:

user(email, age, gender, path_to_pic) // personal data
haveMet(email1, email2) // persons 1 and 2 have reported to have actually met
target(email, gender, ageFrom, ageTo) // preferred gender and age range of a person's match

Consider the query Q "Return all middle-aged[1] potential matches for Mr. Bert O' Lazy (bert@lazy.it)".

1. Provide, **for each** input data source, the reverse engineering from the logical to the conceptual schema (ER graph).
2. Design an integrated global conceptual schema (ER graph) for AnemOne capturing **all** the data coming from both Merge and Patty, and provide the corresponding logical schema.
3. Write GAV mappings between AnemOne's schema and the sources either in Datalog or SQL.
4. Consider query Q posed on AnemOne's schema and write it either in Datalog or SQL.
5. Show the rewriting of Q on the data sources either in Datalog or SQL.

Important: *1)* Spell out all your assumptions.
2) Avoid information loss as much as possible when defining the new schema.
3) List clearly all conflicts you detect during schema integration, if any.

[1] Between the ages of 40 and 60 according to Collins Dictionary.