Q1: What is the earliest paper on your reading list (in terms of publication date) that advocated use of cell phones as sensing devices? (2 points)

Q2: What is the earliest paper on your reading list (in terms of publication date) that advocated use of cars as mobile sensors? (2 points)

Q3: Your colleague asked you to recommend an infrastructure-free solution for tracking objects to a 5cm accuracy in a warehouse of about 15 by 20 meters. Please choose an appropriate response. (2 points)

a) Sure, I read a paper about a system that can deliver such performance. It is called:

   Name the system (1 point):

   Cite the paper (1 point):

b) This is going to be really hard. According to my understanding of the recent state of the art, infrastructure-free localization will have errors of about

   Specify the error you expect (1 point):

   Cite the paper (1 point):
Q4. A theme that emerged in different types of group behavior monitoring on the reading list is that to detect certain types of group behaviors, it is not a good idea to simply compute correlations among time-series data collected from group members for purposes of clustering similar data. Rather, one needs to carefully consider certain key events in each time series and compute correlations among those key events only. These key events often represent some form of change. The insight was that correlated “changes” in data of individuals (according to some appropriate definition of “change”) are much more indicative of membership of the same “group” (according to some appropriate definition of “group”) than correlations computed over the whole data time-series. Find three papers on the reading list that directly or indirectly made the above observation. For each paper, describe the nature of “group behavior” the paper was trying to uncover and the nature of “key events” or “changes” that made really good features in detecting membership of the same group/behavior. (6 points)

Paper 1 (title):

The group behavior it aims to detect:

Key features to correlate:

Paper 2 (title):

The group behavior it aims to detect:

Key features to correlate:

Paper 3 (title):

The group behavior it aims to detect:

Key features to correlate:

Fake example:

Paper (title): “Detecting Groups of Dining Friends in Communal Dining Halls”

The group behavior it aims to detect: Friends eating together as opposed strangers just being seated at the same table.

Key features to correlate: Time of arrival to table and time of departure from table (as opposed to proximity at table or total amount of time overlap at table).
Q5: Maximum likelihood estimation was found to be a powerful technique in extracting information from social media. Several papers were based on the observation that sources (on social media) have certain “properties” or “affinities” that predispose them to generating content with particular “attributes”. Hence, a graph can be constructed linking the set of sources to the set of content artifacts they generate. A maximum likelihood estimator could then jointly uncover both (i) the property/affinity of each source, and (ii) the attribute of each content artifact. Name two papers from the reading list that adopted this general idea. For each paper, describe (i) what property/affinity of each source was being estimated and (ii) what attribute of each content item was being uncovered. (4 points)

Paper 1 (title):

*The graph:*

*The source property/affinity estimated:*

*The content attribute uncovered:*

Paper 2 (title):

*The graph:*

*The source property/affinity estimated:*

*The content attribute uncovered:*

Fake example:

Paper (title): “Joint estimation of gender and sexual orientation from overnight contacts”

*The graph: A bipartite graph connecting a set of host nodes to the set of their overnight guest nodes as detected from their respective phone (GPS/location) traces over time.*

*The host property/affinity estimated: Sexual orientation*

*The guest attribute uncovered: Gender*

Q6: Invent a paper idea that follows the above recipe. To invent the idea, you need to think of a (bi-partite) graph that is the input to your problem. That graph will connect a set of source nodes, each having some property/affinity to be estimated, with a set of artifacts they generate or activities they do, each having some attribute to be uncovered. Your answer should (i) describe the needed graph, (ii) indicate the source property/affinity being estimated, and (iii) indicate the attribute to be detected for each artifact/activity. (3 points)
Example:

The graph: A bipartite graph connecting multilingual sources to tweets they posted as collected from Twitter.

The source property estimated: Probability of this source posting tweets in English, French, German, and Other (i.e., four different probabilities).

The artifact (tweet) attribute: The language of the tweet, classified as English, French, German, or Other.

Your paper idea:

The graph:

The source property estimated:

The artifact/activity attribute uncovered:

Q7: Which paper on your reading list described an approach for hierarchical clustering of data objects in a way that is independent of the data object type? (2 points)

Q8: Event detection from social media posts requires that these posts be clustered based on attributes such as (i) location, (ii) time, and (iii) content. Different papers differ in the way clustering is done and the order in which these attributes are considered. Name a paper from the reading list that clusters content by location first, then applies additional processing to determine if a location-based content cluster actually constitutes a real event. (2 points)
Q9: You are developing a crowd-sensing app that allows people to comment on how full or empty a restaurant is. You consider two options for the interface. In the first, the app asks the diners to estimate the number of empty tables in the restaurant. In the second, the app asks the diners to estimate the percentage of empty tables in the restaurant. In either case, the collected responses are aggregated and the resulting information shared with other prospective customers who have the app. Which option would you recommend? (2 points)

a) Ask for the number of empty tables.

b) Ask for the percentage of empty tables.

Cite your source of inspiration for the above answer from the reading list (and briefly explain):

Q10: Name three papers on your reading list that follow the “surrogate sensing” writing recipe as defined in class? (3 points)

Q11: From your reading list, mention four key IoT research challenges. (2 points)

1.

2.

3.

4.

Total: 30 points.