On Definition of Privacy:

Over the course of the day, a number of people had the common temptation to try to modify the definitions of differential privacy. This is healthy, and the basic notion of differential privacy is not always the "right" answer. But it is also important recognize that definitions of privacy are very subtle (as brought out in several of the talks, such as those by Kobbi Nissim and Ashwin Machanavajjhala), and formulating new notions may be best done in collaboration with people who have thought about this problem for a long time. Often once you have a qualitative notion of information should be protected, and what kinds of information need not be protected (eg because it already is considered "public"), it is possible to define a corresponding analogue of differential privacy.

A couple of interesting points that came out during the day:
- Differential privacy was designed on the principle that is *not* OK to violate the privacy of "just a few" individuals, and that is important to protect outliers (they may even be the most important people to protect). One may be willing to give up these principles in some circumstances, but it should be done with care.

- As commented by Heiko Mantel, the methods used to sample from networks (such as the snowball method mentioned by Eric Kolacyzk) do not compose nicely with differential privacy, meaning (edge-level or node-level) differential privacy on the sampled subgraph does not imply differential privacy with respect to the original social network graph.   This stands in contrast with standard tabular data, where traditional iid sampling preserves, and can even improve, the level of differential privacy. The take-away message is that we cannot ignore the sampling process when evaluating privacy.

On Utility:

It is important to keep in mind that this area of research is still at an early stage, with a lot of progress happening just in the past year - such as the first algorithms with node-level differential privacy. Thus I think we should be optimistic about the future, even if the first attempts at implementing differential privacy for network analysis do not achieve everything we might want.

One question for discussion between network analysts and differential privacy researchers is how well the descriptive statistics studied so far in the differential privacy literature (subgraph counts, degree distributions, cut values) correspond to the descriptive statistics of interest in network analysis. It does seem that some of the statistics that measure "cohesion" in network analysis - eg identifying "central nodes" - may be incompatible with differential privacy, and indeed it is a question whether one can formulate any meaningful notion of privacy for such computations.

An interesting point raised by Eric Kolszyk's talk and Jon Ullman's discussion of it is the distinction between the case when the object of study is a specific network and its properties (eg "network mapping") and the case when we are trying to study a "population" from which the network is drawn (eg "model-based network inference"). As pointed out by Jon, the latter seems more likely to be compatible with differential privacy.