A Medical Decision Problem

You want to build a system to help doctors make decisions, by maximizing expected utility.

- What are the states/acts/outcomes?

States:

- Assume a state is *characterized by n binary random variables*, $X_1, \ldots, X_n$:
  - A state is a tuple $(x_1, \ldots, x_n, x_i \in \{0, 1\})$.
  - The X_is describe symptoms and diseases.
    - $X_i = 0$: you haven't got it
    - $X_i = 1$: you have it
- For any one disease, relatively few symptoms may be relevant.
- But in a complete system, you need to keep track of all of them.

Acts:

- Ordering tests, performing operations, prescribing medication

Outcomes are also characterized by m random variables:

- Does patient die?
- If not, length of recovery time
- Quality of life after recovery
- Side-effects of medications

Some obvious problems:

1. Suppose $n = 100$ (certainly not unreasonable).
   - Then there are 2^{100} states
   - How do you get all the probabilities?
     - You don't have statistics for most combinations!
   - How do you even begin describe a probability distribution on 2^{100} states?
2. To compute expected utility, you have to attach a numerical utility to outcomes.
   - What the utility of dying? Living in pain for 5 years?
     - Different people have different utilities
     - Eliciting these utilities is very difficult
       - People often don't know their own utilities
     - Knowing these utilities is critical for making a decision.

Bayesian Networks

Let's focus on one problem: representing probability.

Key observation [Wright,Pearl]: many of these random variables are independent. Thinking in terms of (in)dependence

- helps structure a problem
- makes it easier to elicit information from experts

By representing the dependencies graphically, get

- a model that's simpler to think about
- (sometimes) requires far fewer numbers to represent the probability