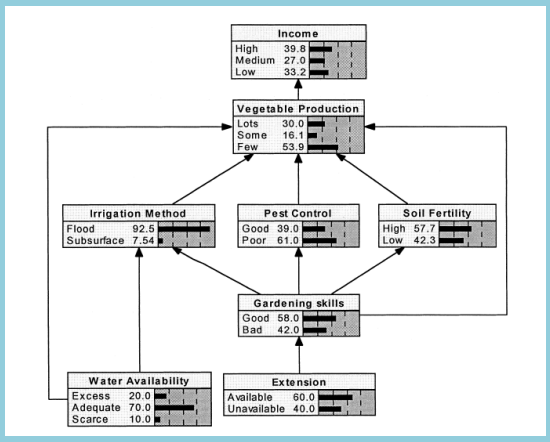
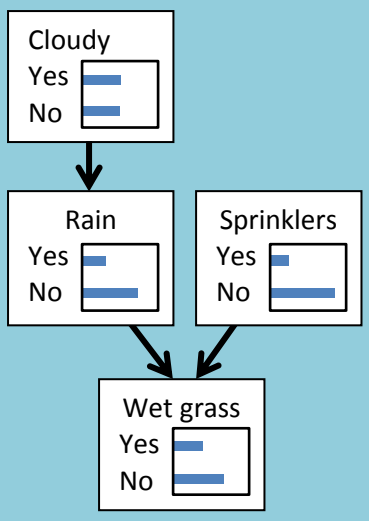
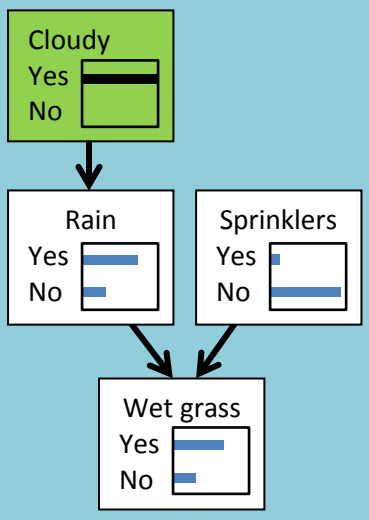
A very short introduction
2: Bayesian networks
e-Bulletin #4



In this, the second bulletin in the 'very short introduction' series, I introduce Bayesian networks. This is the modelling approach I intend for us to try out on the Hebden Water catchment, as a way of better understanding how the different elements of the catchment are linked together. I have not gone into any depth on any of the points, and I am developing a presentation for the first workshop that explains this far better than I can in words. A real example is shown to the top right – a network built to represent a community gardening scheme (from Batchelor and Cain, 1999).

In this bulletin:
- Bayesian networks – introducing key concepts





Bayesian networks in a nutshell

A Bayesian network is a way of graphically laying out different elements of a system, and the cause and effect relationships that exist between them. In the example to the left we see that wet grass can be caused by rain or by sprinklers. Further to this rain can be caused by it being cloudy. It should make sense that whether it is cloudy or not has no effect on the sprinklers (so these are not linked). Each of these *variables* (cloudy, rain, sprinklers and wet grass) has two possible *states* that they can be in, yes or no. The blue bars represent our prior knowledge of the likelihood of each state being true. For example, we could look at weather records and see that it is rainy 40% of the time, and it is cloudy 50% of the time. This would make intuitive sense, as we know from experience that it doesn't rain every time it is cloudy.

In the second example, we look outside and observe that it is cloudy. We now have *evidence* that the variable cloudy is in the state 'yes'. You will notice that a few other changes have happened. Now we know it is cloudy, we have much higher belief that it is raining, and a slightly higher belief that there is wet grass. The belief in rain is not 100% though, as we know that when it is cloudy sometimes it does not rain. Most interestingly, our understanding of the likelihood that the sprinklers are on has decreased. This is because the clouds make rain more likely, and therefore the rain explains away any effect of the sprinklers. In other words, if the grass was wet and we observed clouds we would expect it to be due to rain more often than due to the sprinklers being on.

This is a very simple example given to illustrate some of the key concepts behind Bayesian networks. They can be extended and become increasingly complex. One of the main advantages is that the information used to define the relationships between the variables can come from a wide range of sources; data, expert opinion or personal experience. You can also change an element of the network and see the effects on all other variables immediately, letting you test different possible future scenarios and interventions.

Next steps

I will be e-mailing out dates this week for the workshops, so that we can get started on the project. If you cannot make any of the dates please let me know, and I will do my best to accommodate everyone.

Shaun Maskrey lgxsamas@nottingham.ac.uk 07887 398337