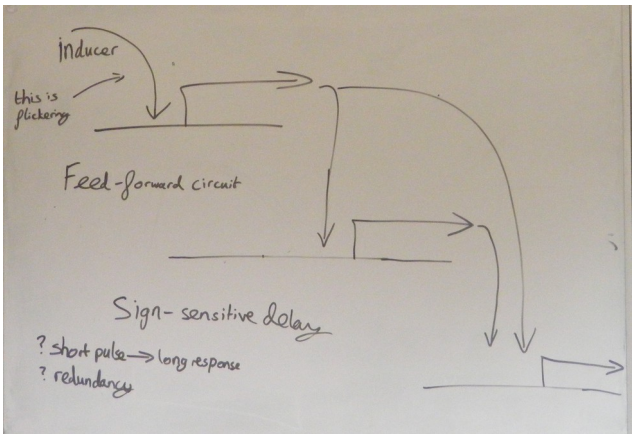
Class notes from DB4 Cellular Mechanisms

Session 3 - 9th October 2015.

We began with you reporting back on your tasks: analysing the three gene networks below.

I also illustrated their action (for two of them, anyway) by using directly comparable electronic analogues, based on 1 transistor modelling 1 gene.

This was a feed-forward network:

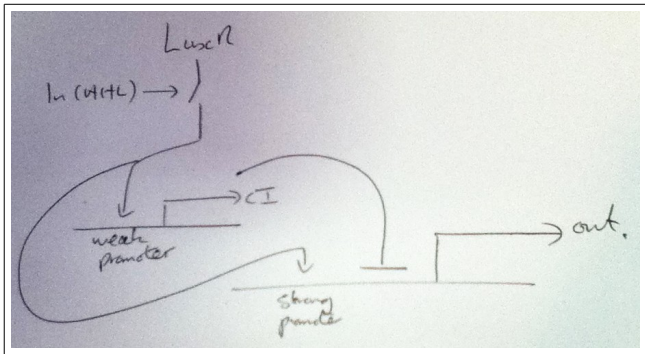


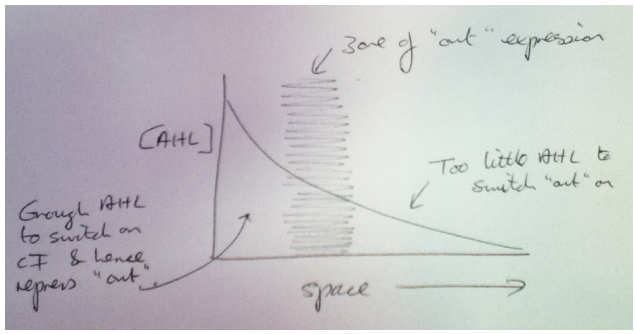
This may well have been the hardest one of the three. In the digital domain, it functions as a sign-sensitive time delay. Starting with everything 'off', when the input switches to 'on', the output also switches to on after one transcription delay. Starting with everything 'on', when the input switches to 'off' the output switches to 'off' only after TWO protein decay delays (because the intermediate gene will be on until the protein of the first has decayed, and the output gene will be on until the protein from the intermediate gene has then decayed). This can be very useful for robustness.

In the analogue domain, it causes hysteresis.

Note that the network analyzed has all-positive connections, and is 'coherent': a mix of positive and negative gives an 'incoherent' feedforward loop, which behaves differently of course. One version is effectively like (3) below. References for this system: PMID: 14607112 , PMID: 23335016 (quite theoretical)

This one was a band-sensitive activator, active for moderate levels of signal but not very high or very low.

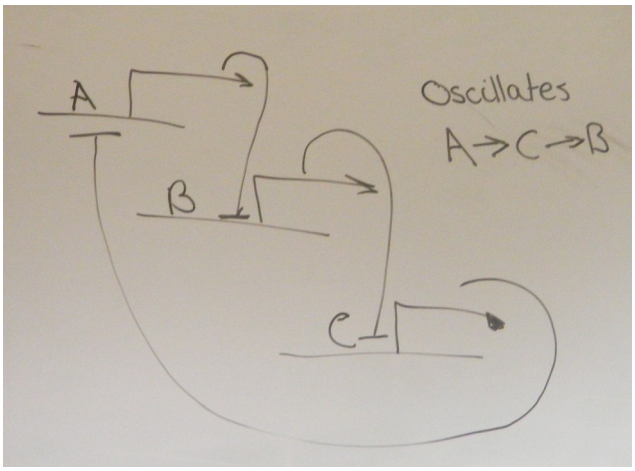




Reference for this system: PMID: 15858574

The embryological relevance of this type of thing is in responding to morphogen gradients, for example turning one gradient into 3 response zones (cf Wolpert's French Flag model). Again, in real life there are additional complications to sharpen gradients.

C was the repressilator (ring oscillator)

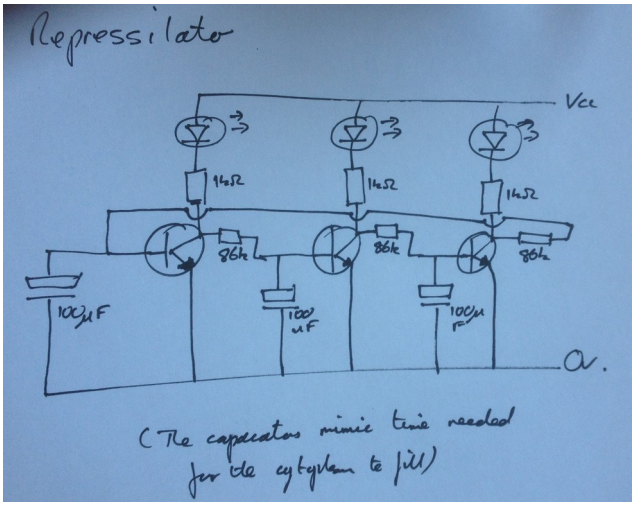


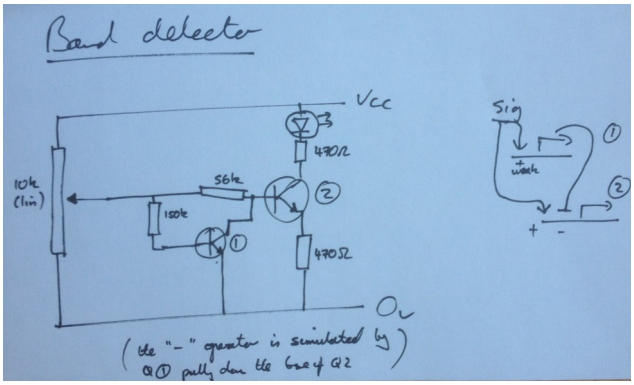
The geometry of this network is pretty straightforward; A represses B, B represses C and C represses A (but each would be on if not repressed). It oscillates. Let's start with A just switching on. After some transcriptional and translational delay, A will hold B off. Once the protein made by B has decayed away, C can switch on. After some transcriptional and translational delay, C will then switch A off. Once the protein made by A has decayed away, B can switch on. After some transcriptional and translational delay, B will then switch C off. Once the protein made by C has decayed away, A can switch on... and we are back where we started. A wave of activation goes backwards (C-B-A-C-B-A...) round the ring.

Reference for this system: PMID: 10659856

You highlighted the somite clock as an example of (more complex) oscillator circuits being used in development. Some of the additional complication in that system is to phase-lock cells together (synchronize them); some is probably for robustness.

By the way – **and of course this is not a remotely examinable part of the course** (I have added it in case any of you has any interest in electronics) – this is how the junk-box circuits I knocked together work:





At this point, our journey to explain the existence of, and transitions between, differentiated cell states has gone from studying control of single genes, through study of very simple positive and negative feedback loops (homeostats and latches), to questioning the existence of 'master regulators', to looking at small gene networks that perform developmentally-relevant tasks (like those above). The final destination is consideration of genome-scale networks.

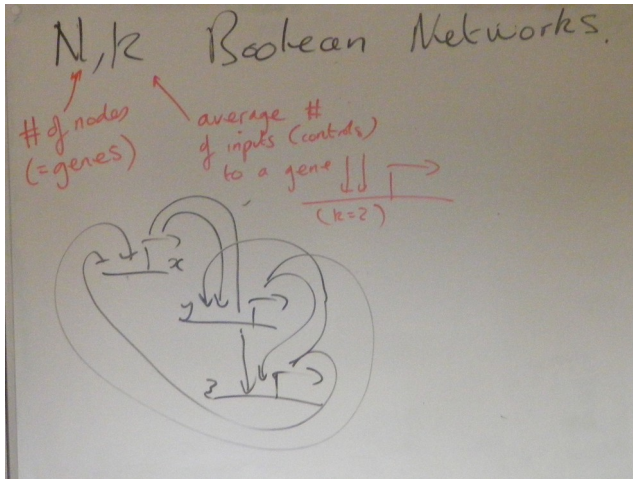
The first class exercise (for which you showed, as a group, a surprising lack of numeracy for BSc Honours students!) just considered what a tiny amount of state space is taken up by differentiated cell states. Perhaps I had better clarify a couple of terms in that sentence: we took a differentiated cell state to be an combination of gene expression states (each gene idealized to being 'off' or being 'on') that is self-sustaining (stable). 'State space' can be considered an N-dimensional space (a graph with N axes) in which the state of the whole network, in terms of all its 1s and 0s, can be specified by a unique point. For example, for 2 genes, a 2-dimensional graph can represent the states 0,0 (the point is at the origin), 0,1 (the point is at level 1 up the y axis but still on the x axis), 1,0 (the point is on the y axis but at level 1 along the x axis) and 1,1 (the point is off both axes, at location 1,1). For 3 genes, a 3-dimensional graph can be used the same way, with 8 distinct possible points. For the 25,000 genes of a human genome (we are rounding numbers all the way through here, for simplicity), we would need a 25,000-axis graph. Of course you cannot literally imagine such a thing, but we did work out that it accommodates $2^{25,000} = 10^{7526}$ different states.

There are roughly 200 differentiated cell states in an adult human (with some room for argument). These occupy an almost unimaginably small part of the 10^{7526}-state statespace. How is stability of just a tiny number of states achieved in such a massive range of possible states?

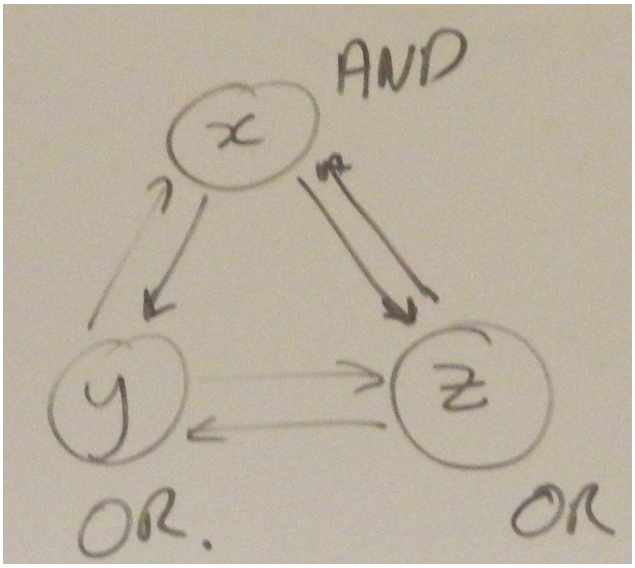
To address this, we examined Stuart Kaufman's N,K Boolean Network models.

In these, N is the number of 'nodes' (genes in our example: the model can be generalized to other things) and K is the <u>average</u> number of inputs that control a node (eg the number of different transcription-factor binding sites in a gene's promoter). The response of each gene (on/off) is determined by a Boolean function of its inputs.

The model begins with the network in any of its possible states. A clock then ticks, and the new state of each gene is calculated according to the states of the inputs to that gene *before the clock ticked*. Then it ticks again, and so on.



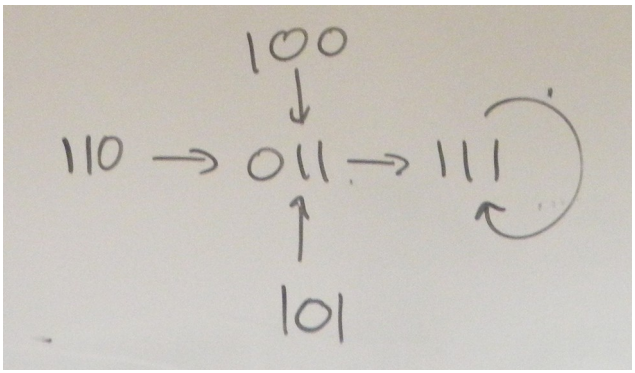
We (you!) explored the functioning of this model with a toy network consisting of just 3 nodes. This is drawn in 'biology-style' above (a mess) and more neatly below.

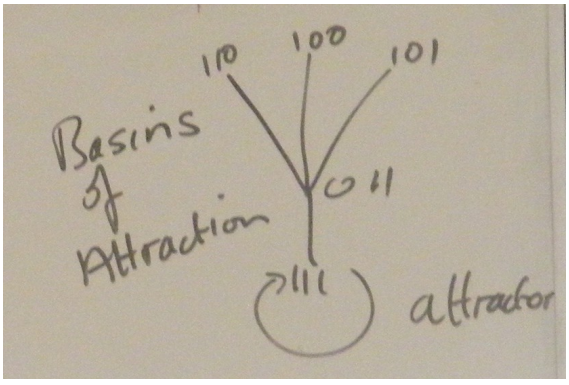


There are 8 (2^3) possible starting states.

You worked through the network from each one, and established that, although there are 8 possible states, only 2 (000 &111) are truly stable and one other oscillates from 010 to 001.

The other 4 possible states all progress to 111. This is called an 'attractor', and the states that lead to it constitute its 'basin of attraction' (see next figures for 2 different styles of drawing this)





So... from our little toy network, we can see that states are not equal. Some are attractors in their own right while others lead to them. This gives an inkling about how small numbers of differentiated states can exist among very many possible initial states.

Kaufman modelled much larger RANDOM N.K Boolean networks by computer, varying K between batches of simulations. He found that when K was large (>5), there were very many attractors – $2^{n/2}$ - and also that most small changes to the network (eg break one connection) cause most attractors to change. On the other hand, when K was small (2-3), the behaviour was very different: there were now only about $2^{\sqrt{n}}$ attractors, working out as 160 for 25,000 genes. What's more, 4/5 of the attractors are robust to (unchanged by) a random break in the network.

None of these models were of actual genomes, of course (they were random), but tis very fact makes the take-home message stronger: the existence of a very few differentiated states in a large network of genes is (if the idealization of genomes as N,K Boolean networks is reasonable) 'automatic' – it emerges from the maths of the networks. Natural selection has no doubt done a great deal to shape what our differentiated states actually are, but the existence of a few stable states in a vast ocean of possibility did not need to evolve – it is am emergent property of such networks.

The take home message from the above is not that we have somehow found a nice neat 'theory of differentiation' and can now go home (we haven't: life is a mess and every case has to be studied the hard way - attempts to predict developmental mechanisms from purely theoretical grounds have almost all failed spectacularly). Rather, it tells us that even random networks of genes show remarkably life-like behaviour; the existence of stable differentiated states may emerge naturally, hierarchical control may not be the norm, and redundancy may be inevitable.

Studies of real genomic networks, such as those in yeast, suggest that the number of connections between genes is not a constant 2 or 3, but instead follows a power law. This is (broadly) characteristic of other networks such as the Internet and Air Travel routes.

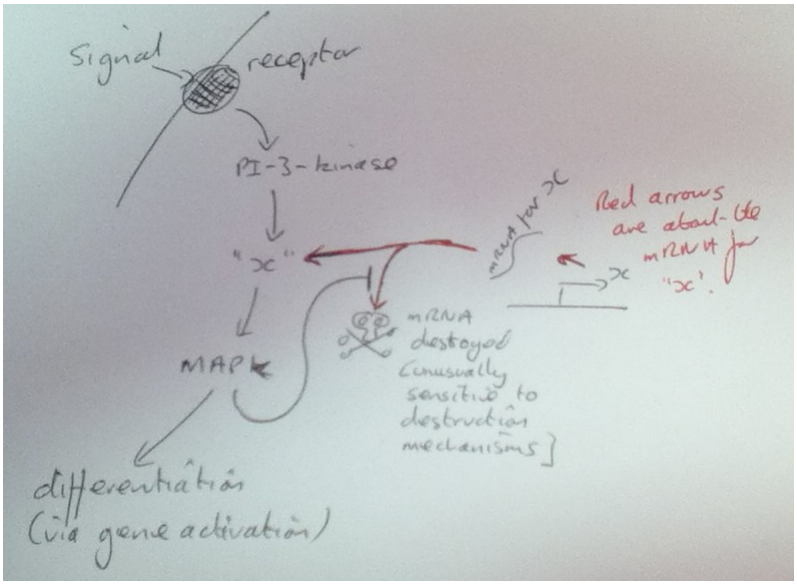
References for N,K networks

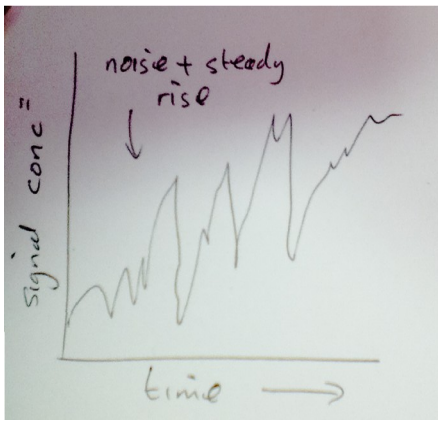
- Attractors in random N,K Boolean networks; Wuensche (1994) - PMID: 9697174
- N,K Boolean networks: Kauffman (1995) *At home in the universe*, Penguin books, chapter 5. This book is a more 'bed-time reading' version of Kauffman (1993) *Origins of Order*, Oxford (a truly excellent book, but harder going). The summary of the science in this chapter is excellent, but be aware that elsewhere in the 1995 book he is pursuing a philosophical/theological agenda alongside the science, which some readers may find distracting.
- Power law networks: Small World (2002), Mark Buchanan (Publ Wiedenfield & Nicholson): again, a pop-science book suitable for 'bed-time reading'.
- Analysis of interactions in the yeast genome Tong et al (2004) Science, Vol 303, Issue 5659, 808-813 , 6 February 2004

Next week, we are going to move out of the cell and into signalling to organize populations of cells.

Homework:

Q3 – here (left panel below) is a cytoplasmic signalling network (I'll tell you what it is from later – I am being mysterious now so you have to think about it and not just look it up). What does it do? Why is this relevant to differentiation in response to an analogue (not 1/0, but continuously-variable) triggering signal like the one on the right panel?





1) Please come up with some exam questions (some short and based on recall, some for long essays based on exploration of a principle, not on simple recall) for the work we have been doing on genes, gene networks and differentiation.

Please e-mail me your questions so I have tie to view them and make any small edits before the next class. Thanks.