Hidden Markov models

1. Let $G_i^\star$ denote the underlying phase-known genotypes, which we know to follow a Markov chain. (So the G_i at least form a hidden Markov model, where $G_i|G_i^\star$ is deterministic.)

   We wish to show that

   $$\Pr(G_{i+1} = g_{i+1} \mid G_1, \ldots, G_{i-1}, G_i = g_i) = \Pr(G_{i+1} = g_{i+1} \mid G_i = g_i)$$

   In the case that $g_i = AA$ or BB, this is obvious, since then $G_i \equiv G_i^\star$ and $G_{i+1}^\star$ (and thus also G_{i+1}) is conditionally independent of $G_1, \ldots, G_{i-1}$ given $G_i^\star$.

   I've found it difficult to prove this in general. I've convinced myself that it is true, but I've not found an easy way to show it. If you find a proof, please let me know. The key seems to be proving that

   $$\Pr(G_i^\star = g_i^\star \mid G_1, \ldots, G_{i-1}, G_i = g_i) = \Pr(G_i^\star = g_i^\star \mid G_i = g_i)$$

   Of course, it's a simple matter to prove, by complete enumeration, that

   $$\Pr(G_3 = g_3 \mid G_1 = g_1, G_2 = g_2) = \Pr(G_3 = g_3 \mid G_2 = g_2)$$

   I think it's worth doing at least this.

2. I don't really feel like writing down the details for the Viterbi algorithm. If you have any questions about it, please ask me.